MONASH University

Rapid and repeatable local adaptation in the global invader *Ambrosia artemisiifolia*

Lotte Anna van Boheemen

BSc (Biology) Utrecht University, Netherlands
MSc (Ecology and Evolution) Groningen University, Netherlands

A thesis submitted for the degree of Doctor of Philosophy at
Monash University in 2018
School of Biological Sciences

Abstract

Local adaptation is a common feature of many species, yet its genetic and molecular basis is still poorly understood. Research on the genetic basis of adaptation increasingly focuses on the speed and repeatability of genomic and phenotypic adaptation and the effect of genomic architecture on evolutionary response. Unravelling such processes could provide insight into the predictability of adaptive shifts, key in the face of on-going environmental change. In this thesis, I aim to shed light on the genetic basis of local adaptation and investigate the speed and the direction of adaptive responses to environmental heterogeneity. To achieve this, I use samples of the invasive weed *Ambrosia artemisiifolia* (common ragweed) across its widespread native North American and the non-native European and Australian ranges and combine phenotypic, genomic and environmental data. I explore neutral variation in molecular markers at a global scale, conduct phenotype-environment, genotype-environment and genotype-phenotype associations to examine signatures of natural selection experienced during multiple range expansions. My results provide strong evidence for rapid local adaptation on a phenotypic and genomic level, where repeated patterns of adaptation evolved in <100 generations. Even though genetic variation and patterns of linkage disequilibrium were distinctly affected by alternative demographic histories, adaptive divergence was seemingly undeterred during range expansion. This research sheds light on processes determining the evolvability of populations and the predictability of adaptive shifts, as well as the underlying genomic basis of local adaptation. My thesis provides insight into the evolutionary processes that occur during the rapid spread of invasive species and the adaptive potential of species more generally.

Publications during enrolment

van Boheemen, L. A., Lombaert, E., Nurkowski, K. A., Gauffre, B., Rieseberg, L. H., & Hodgins, K. A. (2017). Multiple introductions, admixture and bridgehead invasion characterize the introduction history of *Ambrosia artemisiifolia* in Europe and Australia. *Molecular Ecology*, *26*(20), 5421-5434.

van Boheemen, L. A., Atwater, D. Z. & Hodgins, K. A. (2018). Rapid and repeated local adaptation to climate in an invasive plant. *New Phytologist* (2018).

Thesis including published works declaration

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis. This thesis includes one original papers published in peer reviewed journals, one submitted publications and two unpublished papers intended for submission in peer reviewed journals. The core theme of the thesis is evolutionary ecology. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the student, working within the School of Biological Sciences under the supervision of Dr. Kathryn Hodgins. The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research. In the case of Chapters 2-5, my contribution to the work involved the following:

Thesis Chapter	Publication Title	Status	Nature and % of student contribution	Co-author name(s) Nature and % of Co-author's contribution*	Co-author(s), Monash student
2	Multiple introductions, admixture and bridgehead invasion characterize the introduction history of Ambrosia artemisiifolia in Europe and Australia.	Published	70%. Concept, collecting and analysing data and MS preparation	1) Eric Lombaert 5% concept, MS edits 2) Kristin A. Nurkowski 5% lab work, MS edits 3) Bertrand Gauffre 2% concept, MS edits 4) Loren H. Rieseberg 3% concept, MS edits 5) Kathryn A. Hodgins 15% concept, data collections and MS edits	No
3	Rapid and repeated local adaptation to climate in an invasive plant	Published	80% Concept, collecting and analysing data and MS preparation	1) Daniel Atwater 5% concept, data collection, analyses, MS edits 2) Kathryn A. Hodgins 15% concept, MS edits	No
4	Rapid and repeatable local adaptation following multiple introductions	In preparation	80% Concept, collecting and analysing data and MS preparation	1) Kathryn A. Hodgins 20% concept, MS edits	No
5	EICA fails as an explanation of growth and defence evolution following multiple introductions	In preparation	60% Concept, collecting and analysing data and MS preparation	1) Sarah Bou-Assi 15% concept, data collection, MS edits 2) Akane Uesugi 10%, concept, MS edits 3) Kathryn A. Hodgins 15% concept, MS edits	No

I have not renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Student signature: [redacted] Date: ***25-9-2018***

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the student's and co-authors' contributions to this work. In instances where I am not the responsible author I have consulted with the responsible author to agree on the respective contributions of the authors.

Main Supervisor signature: [redacted] **Date:** ***25-9-2018***

Acknowledgements

Kay, thank you for taking a chance on me by enrolling me as your first student. For the last four and a half years, we have been through a lot and I feel our experiences together helped me grow both professionally and personally. You taught me many skills I never imagined I would master, including bioinformatics, complex genomics and taking care of plants. You were patient with my non-native English and have helped me improve my writing by taking time to teach me paragraph structure and how to summarise articles. You kept pushing me, which led to some amazing results and with one paper published and many more on the way. Thank you for introducing me in the wonderful world of ecological genomics, and giving me a skillset that will allow me to continue improving myself. Kristin, you shared your vivacious character with everyone and managed to tame it in the lab. There, you taught me how to perform techniques meticulously (which I never completely mastered). We have given each other friendship and some grief, I especially miss the first.

Carla Sgrò, Paul Sunnucks, Tim Connallon, Roger Cousens and Rob Richardson-Bryson, thanks for your expert feedback and being there when things got rough. You showed extraordinary commitment to aid me through the PhD, especially this final year. Sarah Bou-Assi, thank you for your friendship, your commitment and your greenhouse and lab help. Akane Uesugi, thank you for chatting with me at the greenhouses (making my long days less boring), lending me your students, and watching over me for the last few months. Eric Lombaert, thanks for teaching me ABC and explaining everything so patiently. Daniel Atwater, your amazing dataset and analyses have made a tremendous contribution to our paper. Loren Rieseberg, thanks for your specialist feedback on my first PhD publication. Ros Gleadow, Martin Burd, Kaspar Delhey and Michael Martin, I really appreciated the time you made to chat with me on plants and various techniques. Special thanks to thesis Outi Savolainen and Rose Andrew for very helpful comments on the thesis.

Amazing help in the field or greenhouse by Magda, Jordyn, Anthony, Stephanie, Anoushka, Jennifer, Stevie, Tracy, Guillaume, Emma, Tess, Rose and friends that helped me at one stage or another. Philip, Simon, Stewart, John A. and Fiona, you shine in your job and make PhD life a little less stressful.

Rarely have I had such an amazing group of friends, which have undoubtedly made the last four+ years amazing. I thank you for being there, sharing stories and complaints, and eating loads of dumplings: Stephen (amazing desk mate and donkus), Stef, Mel, Julie, Em, Louise, Toby, Kelso, Michael, Sonia, Niki, Rowan, Ash(s), Mario, Flo, Tom, Hayley, Cass, Cat, Cath, Cathy, Katie, Candice, Karin, Lee(s), Matthew, Marianne, Helena, Paulina, Gerard, Susie, Marie(s), Fran, Mo, Raquel, Beck, Ollie, Nouman, Will, Jo, Manoj, Andrei, Nataly, Mattia, Marcelo, Vignesh, Prachi, Stephanie, Joe, Rasika and Bass, with postgrad committee colleagues and friends Sean, Amanda, Ash, Marcus, Clem, Qam and Evatt. Nancy and Shane,

thank you for taking me into your home and being amazing friends in general. Sunnucks' group Sasha, Hernán, Anders, Katherine, Ash, Emma, Bertrand, Annie, Steph, Maiko and Lynna, you gave me friendship, scientific inspiration and have adopted me part-time. I received external love from amazing friends Savanna, Kate, Joe, Loris, Linda, Gosia and Pape. Liz, Jane and Warrick, you showed me the amazing underwater life of Melbourne and brought me colour and calm.

Luci, you supported me through the most difficult time in my life, stood by me and showed me love. One day soon I will stop working so much and travel the world with you, I promise. Broertje, je bent gewoonweg top en hebt altijd liefde en knuffels mijn kant opgestuurd. Oma, ik was dan wel ver weg, maar jij bracht Europa altijd iets dichterbij met je kaarten en kadootjes. Rune en Myako, jullie zijn de liefste en mooiste knuffeldieren in de hele wereld. Mam en pap, met jullie voel ik mij de meest geliefde en gesteunde dochter. Dank jullie voor het begrip (soms…) tijdens alle vakanties en bezoekjes dat ik achter de computer zat en een bordje 'daughter at work' had staan. Ik hou zielsveel van jullie. Rust zacht grote beer.

Table of contents

Chapter 1 – General introduction

Local adaptation, the genetically based fitness advantage of home relative to foreign populations, occurs through spatial heterogeneity in selection pressures and is found in many species (Hereford, 2009). The growing interest in revealing the genetic and molecular basis of local adaptation arises as insights can be used to predict species ability to adapt to a changing climate, human induced disturbances, invasive spread and aid conservation efforts (Kinnison & Hairston, 2007; Hoffmann & Sgrò, 2011; Fauvergue *et al.*, 2012; Chown *et al.*, 2014; Harrisson *et al.*, 2014; Hoffmann *et al.*, 2017; Lässig *et al.*, 2017). Although many studies have provided important insight into the genetic basis of local adaptation (e.g. Turner *et al.*, 2008; Fournier-Level *et al.*, 2011; Turchin *et al.*, 2012; Andrew *et al.*, 2013; Colautti & Barrett, 2013; Savolainen *et al.*, 2013; Hämälä & Savolainen, 2018), much remains to be resolved, including the speed at which adaptation occurs (Hoffmann & Sgrò, 2011; Bay *et al.*, 2017), the genomic and phenotypic repeatability of adaptation (Elmer & Meyer, 2011; Losos, 2011; Conte *et al.*, 2012; Lenormand *et al.*, 2016; Bailey *et al.*, 2017) and the genomic architecture that underpins adaptive shifts (e.g. the number, association and distribution of loci responding to selection and their effect size; Kawecki & Ebert, 2004; Hansen, 2006; Stapley *et al.*, 2010; Savolainen *et al.*, 2013; Connallon & Hall, 2018).

Contemporary adaptation

Local adaptation may evolve rapidly due to recent exposure to novel ecological factors, mediated through introductions to new environments, anthropogenic disturbance and climate change (Reznick & Ghalambor, 2001). Rapid local adaptation, when it occurs, is expected in response to strong selective pressures (Hoffmann & Sgrò, 2011), with a species' evolvability depending on factors including gene flow, effective population size and the amount of additive genetic variation for traits under selection (Barrett & Hoekstra, 2011; Bay *et al.*, 2017; Hoffmann *et al.*, 2017). Adaptation can arise through selection on new mutations or selection on standing genetic variation, both resulting in different evolutionary dynamics and distinct genetic signatures (Barrett & Schluter, 2008; Prentis *et al.*, 2008; Pritchard *et al.*, 2010). Standing genetic variation is predicted to contribute more to rapid adaptation than new mutations due to the immediate availability and higher fixation probability of beneficial alleles (Barrett & Schluter, 2008; Prentis *et al.*, 2008; Rockman, 2012; Lee & Coop, 2017; MacPherson & Nuismer, 2017). Indeed, empirical studies are beginning to reveal the importance of standing variation in rapid adaptive change (e.g. Burke *et al.*, 2010; Jones *et al.*, 2012; Marques *et al.*, 2018). Despite these advances however, the genetic basis of adaptation to rapid environmental change is still poorly understood. Such knowledge can provide insight into the evolutionary potential of species to respond to on-going environmental change, as well as the genetic factors

that may contribute to and constrain rapid evolutionary response (Reusch & Wood, 2007; Jump *et al.*, 2009; Hoffmann & Sgrò, 2011).

Repeatable adaption

The reemergence of comparable phenotypic patterns across a species' range suggests that evolution is not solely governed by stochastic processes (Wood *et al.*, 2005; Arendt & Reznick, 2008; Elmer & Meyer, 2011; Lee *et al.*, 2014) and is therefore considered proof of adaptation (Seehausen *et al.*, 2014). Parallel phenotypes can evolve through changes in the same genes or genetic variants (Conte *et al.*, 2012; Martin & Orgogozo, 2013), or through alternative genetic changes leading to similar phenotypes (Elmer & Meyer, 2011; Smith & Rausher, 2011; Losos, 2011). Similar genetic and phenotypic adaptation to comparable selective pressures suggests that the response to selection may be biased or constrained (Weinreich *et al.*, 2006; Gompel & Prud'homme, 2009; Chevin *et al.*, 2010b; Losos, 2011; Connallon & Hall, 2018), for instance through the initial frequency of selected alleles or the underlying genetic architecture. However, the extent of genomic repeatability of adaptive shifts and constraints on evolvability are still major questions in evolutionary biology (Elmer & Meyer, 2011; Losos, 2011; Conte *et al.*, 2012; Lenormand *et al.*, 2016; Bailey *et al.*, 2017; Connallon & Hall, 2018). Studying the genomic basis of repeated phenotypic adaptation could give insight in the predictability of evolution (Stern & Orgogozo, 2008; Stern & Orgogozo, 2009; Lässig *et al.*, 2017), which will be essential in managing responses to climate change, extinctions or rapid invasive spread.

Genomic architecture

The physical organization of, and interaction between, genes in the genome, often referred to as genomic architecture, can have profound consequences for adaptation and speciation (Hansen, 2006; Slatkin, 2008; Nosil & Feder, 2012; Yeaman, 2013; Flaxman *et al.*, 2014; Ortiz-Barrientos *et al.*, 2016). Local adaptation often occurs via allele frequency changes of polygenic loci that underlie quantitative traits (Pritchard & Di Rienzo, 2010; Pritchard *et al.*, 2010). Theory predicts that adaptive changes will depend on the number of loci underlying a functional trait and their mutational effect size, such that few loci with large fitness effects should contribute more to local adaptation than many small effect loci, given intermediate levels of gene flow (Orr, 1998; Yeaman & Whitlock, 2011). Moreover, loci with large phenotypic effect are likely to respond first to selective changes (MacPherson & Nuismer, 2017) and are thus predicted to contribute more to repeatable patterns of adaptation (Lässig *et al.*, 2017; Yeaman *et al.*, 2018). Genetic interactions and reduced recombination among adaptive loci may also enhance evolvability under migration-selection balance (Maynard Smith, 1977; Pylkov *et al.*, 1998; Lenormand & Otto, 2000; Yeaman & Whitlock, 2011). Such associations elevate repeatable genomic patterns, especially when the genomic context is shared due to gene flow

or in recently diverged lineages (Renaut *et al.*, 2014; Holliday *et al.*, 2016; Storz, 2016; Lässig *et al.*, 2017). On the other hand, a locus underlying a trait strongly favoured by selection could also drag neutral or detrimental loci through such associations (Hartfield & Otto, 2011) and may have negative pleiotropic effects on other traits (Otto, 2012). Yet, empirical studies examining the number and effect size of loci underlying functional traits and the contribution or constraints of genomic associations (e.g. linkage and epistasis) on evolvability largely remain untested (Hoban *et al.*, 2016; Ahrens *et al.*, 2018).

Introduced species as a tool to study adaptation

Introduced species provide an exceptional system to study rapid and repeated local adaptation (Bock *et al.*, 2015; Hodgins *et al.*, 2018). Although plasticity and ecological tolerance could contribute to the success of invasive species (e.g. Geng *et al.*, 2007; Zhang *et al.*, 2010), their global spread is often associated with rapid evolutionary events (Huey *et al.*, 2000; Lee, 2002; Maron *et al.*, 2004; Lachmuth *et al.*, 2011; Lawson Handley *et al.*, 2011; Colautti & Barrett, 2013; Chown *et al.*, 2014; Turner *et al.*, 2014; Bock *et al.*, 2015; Colautti & Lau, 2015; Oduor *et al.*, 2016; Szűcs *et al.*, 2017). Invasive species frequently re-establish along similar environmental gradients found in their native ranges, or flourish in new environments (Sax & Brown, 2000; Allendorf & Lundquist, 2003; Atwater *et al.*, 2018). Consequently, biological invasions provide an opportunity to study contemporary repeated and divergent adaptive processes, and thus have the potential to provide insight into the processes underpinning rapid and repeatable adaptation.

Several distinct demographic events characterize the invasion processes and could affect the evolutionary potential of introduced species (Lee, 2002; Facon *et al.*, 2006; Prentis *et al.*, 2008; Rius & Darling, 2014; Estoup *et al.*, 2016). Within small founding populations, genetic drift can reduce genetic variation (Wright, 1931; Dlugosch & Parker, 2008a). Depending on the size of the bottleneck and subsequent gene flow, genetic load can increase or decrease (Frankham, 1995; Glémin, 2003; Facon *et al.*, 2011; Blackburn *et al.*, 2015; Marchini *et al.*, 2016; Wang *et al.*, 2017). Moreover, bottlenecks could lead to the conversion of non-additive to additive genetic variance (Neiman & Linksvayer, 2006). Through a comparable mechanism, gene surfing, the frequency increase of genetic variants at the expansion wave front due to strong genetic drift (Klopfstein *et al.*, 2006; Excoffier & Ray, 2008; Excoffier *et al.*, 2009a) can lead to the accumulation of deleterious mutations (Peischl *et al.*, 2013; Peischl *et al.*, 2015), and this effect is especially strong if mutations are recessive (Peischl & Excoffier, 2015). Additionally, admixture of distinct genotypes is considered beneficial for the adaptive potential of introduced populations, through increased genetic variation (Anderson & Stebbins, 1954), the creation of novel or transgressive phenotypes (Stebbins, 1969; Rieseberg *et al.*, 1999) and heterosis, stimulating growth and vigour, and potentially 'catapulting' populations past the

lag phase of colonisation (Facon *et al.*, 2005; Drake, 2006; Keller *et al.*, 2014; Wagner *et al.*, 2017).

These demographic processes can affect the genomic architecture underlying adaptive traits, with important consequences for the adaptive potential of invasive species. Admixture and drift are predicted to affect variances and covariances among loci throughout the genome (Lande, 1980; Keller & Taylor, 2010), which could interfere with the adaptive response if the associated alleles have opposing selection coefficients or opposing effects (Fisher, 1930; Muller, 1964; Hill & Robertson, 1968; Felsenstein, 1974; McVean & Charlesworth, 2000; Slatkin, 2008). Moreover, rapid invasive changes are predicted to be led by large effect loci (Dlugosch *et al.*, 2015a). Dissection of the various evolutionary processes and their interplay is required to shed light on rapid invasive spread (Keller *et al.*, 2009; Bonhomme *et al.*, 2010; Lachmuth *et al.*, 2011; Agrawal *et al.*, 2015; Cristescu, 2015; Dlugosch *et al.*, 2015a) and species' adaptive potential overall.

Study species

Ambrosia artemisiifolia L. (Asteraceae) is a wind pollinated, outcrossing hermaphrodite that has aggressively spread from its native North America that has invaded every continent except Antarctica (Laaidi *et al.*, 2003; Oswalt & Marshall, 2008; Smith *et al.*, 2013). This self-incompatible annual weed (Brandes & Nitzsche, 2006) is commonly found in disturbed habitats (Bassett & Crompton, 1975; Lommen *et al.*, 2017) and can produce large quantities of seed that can stay dormant for many years (Bassett & Crompton, 1975). This species is an agricultural pest and has significant effects on crop productivity. Competition experiments in corn and soybean have showed yield losses reaching 70% (Weaver, 2001; Brandes & Nitzsche, 2006). In addition, its wind-spread pollen is a leading cause of hay fever worldwide (Laaidi *et al.*, 2003) and costs millions of dollars each year in medical treatment (Taramarcaz *et al.*, 2005). Rising CO_2 levels and ongoing climate change are predicted to stimulate *A. artemisiifolia* growth and pollen production (Chapman *et al.*, 2014), increasing the impact on public health (Emberlin, 1994; Ziska & Caulfield, 2000). These detrimental factors provide considerable incentive to understand the mechanisms impacting the invasive spread of this species.

Within the native North American range, a shift in the population genetic structure using historic and contemporary samples was attributed to human-mediated admixture of western and eastern native genotypes (Martin *et al.*, 2014). Pre-adaptation of genotypes associated with anthropogenic-disturbed environments (Hufbauer *et al.*, 2012) is suggested to have facilitated its global introduction and spread elsewhere. The first records documenting the European *A. artemisiifolia* invasion are in central France around 1850 (Chauvel *et al.*, 2006). These initial founders likely originated from contamination of imported seeds from North America. Later major introductions have been tied to imports during the two World Wars (Chauvel *et al.*,

2006). The multiple introductions from distinct native sources to both east and west Europe resulted in levels of genetic variation equivalent to those found in the native range (Genton *et al.*, 2005a; Chun *et al.*, 2010; Gladieux *et al.*, 2010; Gaudeul *et al.*, 2011; van Boheemen *et al.*, 2017b). In Australia, *A. artemisiifolia* was first observed at the beginning of 1900 century and has become increasingly abundant in southeastern Queensland and New South Wales since around 1950 (Palmer & McFadyen, 2012). A previous *A. artemisiifolia* study testing the genetic structure of global introductions revealed the two sampled Australian populations were strongly differentiated from the rest of the populations around the world. However, the introduction pathway into this range remains unclear.

Latitudinal clines in phenology have been observed within the native North American range and the introduced ranges of Europe (Chun *et al.*, 2011; Leiblein-Wild & Tackenberg, 2014) and China (Li *et al.*, 2014), with earlier reproduction and greater relative investment in reproductive biomass in high-latitudinal compared to low-latitudinal populations (Chun *et al.*, 2011; Hodgins & Rieseberg, 2011; Li *et al.*, 2014). Comparison of multiple introduced ranges, together with an extensive genomic analysis will provide insight into the repeatability of ecological functional trait evolution.

General approach

In this thesis, I aim to provide insight into the following major evolutionary questions: *i)* what is the speed of adaptation? (Hoffmann & Sgrò, 2011; Bay *et al.*, 2017); *ii)* to what extent is genomic and phenotypic adaptation repeatable? (Elmer & Meyer, 2011; Losos, 2011; Conte *et al.*, 2012; Lenormand *et al.*, 2016; Bailey *et al.*, 2017) and *iii)* how does the genomic architecture affect evolvability? (Kawecki & Ebert, 2004; Hansen, 2006; Stapley *et al.*, 2010; Savolainen *et al.*, 2013; Connallon & Hall, 2018). To achieve this, I link several lines of evidence for local adaptation, combining phenotypic, genomic and environmental data, to examine signatures of natural selection and identify evolutionary processes (Sork *et al.*, 2013; Rellstab *et al.*, 2015; Hoban *et al.*, 2016; Ahrens *et al.*, 2018). As such, I explore neutral variation at a global scale (Chapter 2), conduct phenotype-environment correlations (Chapter 3 & 5), and perform genotype-environment and genotype-phenotype associations (Chapter 4). This research is supported by global sampling of the invasive weed *Ambrosia artemisiifolia* (common ragweed) across its widespread native North American and the non-native European and Australian ranges. Invasive weeds are a powerful study system, as these species are experimentally easy to grow in glasshouses, allowing for convenient studies on associations between genomics, phenotypes and (experimental) environmental conditions. By examining the adaptive and neutral genomic diversity in native and introduced populations of an invasive species, I will gain novel insights into the genomic mechanisms associated with adaptation to new environments. Moreover, identification of the role that genomic changes play in the rapid

spread of invasive species will improve our understanding of invasive species and could aid in our capacity to predict and manage future invasion scenarios (Prentis *et al.*, 2008; Estoup & Guillemaud, 2010; Bock *et al.*, 2015; Cristescu, 2015; Dlugosch *et al.*, 2015a).

Chapter outline

Results of this project will be presented in the following chapters:

Chapter 1 – General introduction

Current chapter, outline the general topics discussed in this thesis, mark knowledge gaps and overall aim of the Doctorate.

Chapter 2 – Multiple introductions, admixture and bridgehead invasion characterize the introduction history of Ambrosia artemisiifolia in Europe and Australia.

We explore the demographic history of *A. artemisiifolia* as this has potential consequences for it's adaptive potential and invasion success. Invasive populations have established through multiple introductions from the native range into Europe and subsequent bridgehead invasion into Australia. We propose several evolutionary mechanisms that could promote or impede local adaptation as a result of such admixture or bridgehead introductions.

Chapter 3 – Rapid and repeated local adaptation to climate in an invasive plant

Biological invasions provide opportunities to examine adaptation along similar climatic gradients over contemporary timescales. To gain insight into the speed and repeatability of adaptation, we explore multiple geographical clines in quantitative traits of *A. artemisiifolia*, whilst benchmarking against neutral patterns using SNP data. We identify repeated adaptive trait differentiation along latitudinal clines within the native and introduced ranges. These results suggest local adaptation of *A. artemisiifolia* occurs in a repeatable manner within short evolutionary time-scales (50-150 years). These patterns evolved despite the distinct introduction history of Europe (multiple introductions from North America) compared to Australia (single bottlenecked introduction from Europe) suggesting that genetic bottlenecks did not constrain this species' capacity to adapt to climate variation.

Chapter 4 – Rapid and repeated local genomic adaptation in an invasive plant

Recent range expansions along similar climatic gradients across multiple geographic ranges provide opportunities to assess the rate and repeatability of adaptation. Evidence of parallel geographic clines in quantitative traits has frequently been identified (e.g. Ch2), but few studies have explored if this repeatability extends to the genomic level. Our analysis of 83,559 SNPs in 852 *Ambrosia artemisiifolia* individuals identified substantial repeatability among three

ranges at putative selected loci (ranging from 25-41%). These data provide evidence of substantial constraints and/or biases that limit the diversity of genetic forms contributing to rapid and repeated bouts of adaptation.

Chapter 5 – Repeated tests of enemy escape following multiple introductions along latitudinal clines

Escape from specialist herbivores has been invoked by multiple hypotheses related to trait evolution and invasion success, including the Evolution of Increased Competitive Ability and the Shift Defence Hypothesis. We tested for genetic differentiation in herbivore-defence related traits in *Ambrosia artemisiifolia* to explore the potential impact of escape from specialist herbivores on adaptive trait divergence. Defence related traits show divergence in Europe and Australia from the native trait values. These patterns are not explained by divergence at neutral markers suggesting that adaptive divergence of traits has occurred. Although specialist herbivores are largely absent in Europe and Australia, escape from specialist herbivores is likely not responsible for the shifts as defence related traits as well as traits related to performance are diverging in opposing directions in each introduced range.

Chapter 6 – General conclusion and discussion

Summary of chapter findings, synergy of the concepts and future directions.

Chapter 2 – Multiple introductions, admixture and bridgehead invasion characterize the introduction history of Ambrosia artemisiifolia in Europe and Australia.

Lotte A. van Boheemen[a], Eric Lombaert[bcd], Kristin A. Nurkowski[e], Bertrand Gauffre[afg], Loren H. Rieseberg[e], Kathryn A. Hodgins[a]

[a]School of Biological Sciences, Monash University, Clayton, VIC 3800, Australia

[b]Inra, UMR 1355 ISA, Sophia-Antipolis, France

[c]Université de Nice Sophia Antipolis, UMR ISA, Sophia-Antipolis, France

[d]CNRS, UMR 7254 ISA, Sophia-Antipolis, France

[e]Department of Botany, University of British Columbia, 1316-6270 University Blvd., Vancouver, BC, V6T 1Z4, Canada

[f]UMR 7372, Centre d'Etudes Biologiques de Chizé, CNRS – Université de La Rochelle, Villiers-en-Bois, France

[g]USC1339, Centre d'Etudes Biologiques de Chizé, INRA, Villiers-en-Bois, France

Corresponding author: Lotte A. van Boheemen[a], School of Biological Sciences, Monash University, Victoria 3800, Australia, lotte.van.boheemen@monash.edu

Keywords: admixture, introduction history, bridgehead invasion, *Ambrosia artemisiifolia*, approximate Bayesian computation (ABC), Random Forests

Manuscript published in August 2017, *Molecular Ecology* 26 (20)

Abstract

Admixture between differentiated populations is considered to be a powerful mechanism stimulating the invasive success of some introduced species. It is generally facilitated through multiple introductions; however, the importance of admixture prior to introduction has rarely been considered. We assess the likelihood that the invasive *Ambrosia artemisiifolia* populations of Europe and Australia developed through multiple introductions or were sourced from a historical admixture zone within native North America. To do this, we combine large genomic and sampling datasets analysed with approximate Bayesian computation and Random Forest scenario evaluation to compare single and multiple invasion scenarios with pre- and post-introduction admixture simultaneously. We show the historical admixture zone within native North America originated before global invasion of this weed, and could act as a potential source of introduced populations. We provide evidence supporting the hypothesis that the invasive populations established through multiple introductions from the native range into Europe and subsequent bridgehead invasion into Australia. We discuss the evolutionary mechanisms that could promote invasiveness and evolutionary potential of alien species from bridgehead invasions and admixed source populations.

Introduction

Globalisation has been accompanied by worldwide flourishing of alien species. Non-natives can have disruptive and disastrous ecological consequences by out-competing their indigenous counterparts and affecting invaded ecosystems. During an invasion process, several demographic events can lead to changes in genetic variation, and thereby influence the evolutionary potential and invasiveness of alien species (Lee, 2002; Facon *et al.*, 2006; Prentis *et al.*, 2008; Rius & Darling, 2014; Estoup *et al.*, 2016). Within small founding populations and on the invasion front, genetic drift can lead to reduced genetic diversity, potentially impacting additive genetic variation (Wright, 1931; Dlugosch & Parker, 2008a; Excoffier & Ray, 2008; Peischl *et al.*, 2013; Bock *et al.*, 2015). Such bottlenecks can also result in increased genetic load due to the reduced efficacy of selection (Frankham, 1995; Blackburn *et al.*, 2015). An apparent contradiction between the expected negative effects of introduction on genetic variation and fitness, and the invasion success of some species, has been termed the 'genetic paradox of invasion' (Frankham, 1995; Allendorf & Lundquist, 2003, for review see Estoup *et al.* (2016)).

One potential resolution to the genetic paradox of invasion for some invaders may be admixture: the mixing of genotypes from differentiated genetic backgrounds (Roman & Darling, 2007; Dlugosch & Parker, 2008a; Prentis *et al.*, 2008; Rius & Darling, 2014). Such admixture can arise within the native range prior to colonization (Keller *et al.*, 2014; Martin *et al.*, 2014), or as a more frequently described consequence of multiple introductions (Dlugosch

& Parker, 2008a; Uller & Leimu, 2011). Admixed genotypes in an introduced range could subsequently act as bridgehead—a successful invasion acting as a source for further introductions (Lombaert *et al.*, 2011). Depending on the influx of genetic material, beneficial outcomes of intra-specific hybridization include higher genetic variance within populations, heterosis, as well as novel and/or transgressive phenotypes (Verhoeven *et al.*, 2011; Rius & Darling, 2014; Bock *et al.*, 2015). Disentangling the demographic history of an invader and identifying the putative source populations provides valuable information about the amount and type of genetic variation present in the introductions, which can have significant consequences on the success of invasion (Dlugosch & Parker, 2008a; Dlugosch *et al.*, 2015a; Estoup *et al.*, 2016). Knowledge of these processes are key for studies concerned with understanding how and if alien populations have adapted since their introduction and the role of pre-adaptation to invasion success (Facon *et al.*, 2006; Rius & Darling, 2014; Barrett, 2015; Estoup *et al.*, 2016). The potential role of admixture to invasion success, be it via admixture in the native range, through multiple introductions, or through bridgehead introductions, illustrates the importance of unravelling introduction history. Understanding introduction history requires broad sampling of the native range (Bossdorf *et al.*, 2005; Dlugosch & Parker, 2008a; Cristescu, 2015) along with a comparison with the introduced areas (Dlugosch & Parker, 2008a). However, sampling limitations have often hampered reconstructions of evolutionary changes that accompany invasion (Lombaert *et al.*, 2011; Cristescu, 2015).

The invasive annual weed *Ambrosia artemisiifolia* L. (common ragweed, Asteraceae) is native to North America and introduced to Europe, South America, Australia and Asia where it is considered as an agricultural pest (Oswalt & Marshall, 2008). In Europe, the first known introduction of *A. artemisiifolia* was in France around 1850 and most likely originated from contamination of imported seeds from North America. Later major introductions have been tied to imports during the two World Wars (Chauvel *et al.*, 2006). Given the early expansion of this weed in Europe, introduced European populations could have acted as a bridgehead for subsequent introductions elsewhere. This pathway has however not been tested within this species. In Australia, *A. artemisiifolia* was first observed at the beginning of the 20th century, and has become increasingly abundant in southeastern Queensland and New South Wales since around 1950 (Palmer & McFadyen, 2012). The population structure of this weed within the introduced European range has previously been described and accredited to multiple introductions from various source populations originating from North America (Genton *et al.*, 2005a; Chun *et al.*, 2010; Gaudeul *et al.*, 2011). Recently, Martin *et al.* (2014) suggested that these signatures could be due to pre- rather than post-introduction admixture. These authors observed a shift in genetic structure in a population genetic analysis of historic and contemporary samples of this species from across the native North American range. This pre-introduction shift in the native range was attributed to human-mediated admixture of western

and eastern native genotypes (Martin *et al.*, 2014). As admixture in the native range seems to have been associated with land use change and human mediated dispersal (Martin *et al.*, 2014), this raises the possibility that pre-adaptation to anthropogenic-disturbed environments (Hufbauer *et al.*, 2012) in North America facilitated its introduction and spread elsewhere.

In this study, we aimed to assess the contribution of historical admixture, multiple introductions and bridgehead invasion to the successful introductions of *A. artemisiifolia* into Europe and Australia. We reconstructed the introduction history of *A. artemisiifolia* within Europe and Australia using approximate Bayesian computation (ABC) based on Random Forest (RF) algorithms, using 1022 SNP loci identified in 466 samples collected at 85 locations from across the entire native range of North America, and much of the European and Australian introduced ranges. ABC-RF analyses are superior to other approaches because they have the capacity to differentiate among competing complex introduction scenarios (Pudlo *et al.*, 2016; Fraimout *et al.*, 2017) while using the increased resolution of genomic datasets (e.g. Momigliano *et al.* (2017). This approach might be particularly informative in cases like ragweed where traditional approaches with few markers may be unable to identify source populations especially given complex introduction histories and weak population structure in the native range. We complemented this analysis with an assessment of genetic population structure (1022 SNPs) and diversity (10,100 SNPs) within the native and introduced ranges. We provide evidence that the invasive European population likely established through multiple introductions from two major genetic clusters that we identified, while Australian populations appear to have been sourced from a subsequent bridgehead invasion from the European introduction.

Materials and methods

Study species

Ambrosia artemisiifolia is a monoecious annual commonly found in disturbed habitats (Oswalt & Marshall, 2008). A full-grown *A. artemisiifolia* plant can produce large quantities of seed that can stay dormant for many years (Bassett & Crompton, 1975). This species is an agricultural pest and has significant effects on crop productivity with competition experiments showing yield losses in corn and soybean reaching 70% (Weaver, 2001; Brandes & Nitzsche, 2006). Its wind-spread pollen is a leading cause of hay fever worldwide (Laaidi *et al.*, 2003). Rising CO_2 levels and ongoing climate change are predicted to stimulate *A. artemisiifolia* growth and pollen production, increasing the impact on public health (Emberlin, 1994; Ziska & Caulfield, 2000).

DNA extraction and sequence filtering

We collected leaf samples between 2008 and 2014 from three continents: the native range of North America (230 individuals, 42 sampling locations) and introduced ranges of Europe (195 individuals, 36 sampling locations) and Australia (41 individuals, 7 sampling locations) (Fig. 1, Table S1, supporting information). At each sampling locations, we randomly selected up to 7 plants growing at least 1 m apart to reduce the chance of the plants being close relatives due to local pollen and seed dispersal. For each plant, we placed two green leaves into paper envelopes, which were then stored at room temperature in a sealed plastic bag containing silica gel. We extracted genomic DNA (gDNA) from 20-30 mg dried leaf tissue for 374 samples using the NucleoSpin Plant II kit (Macherey-Nagel, Düren, Germany) and an additional 92 samples using the DNAeasy Plant Mini Kit (Qiagen, Valencia, CA, USA). We assessed DNA quantity (>8.5 ng/µl) using a QuBit broad-sensitivity DNA quantification system (Invitrogen, Carlsbad, CA, USA).

Genotype-by-sequencing and SNP calling

We performed genotype-by-sequencing library construction with single restriction enzyme *Pst-1* using an amended version of the Elshire *et al.*, 2011 protocol (see Appendix S1, supporting information). We explored quality statistics of the raw reads using FastQC (http://hannonlab.cshl.edu/fastx_toolkit). We de-multiplexed reads using STACKS process_radtags (Catchen *et al.*, 2011). After removing adapter sequences, we trimmed reads using Sickle (Joshi & Fass, 2011) with a Q-score of ≥20 and read length of ≥50 base pair. We filtered reads with FASTQ quality filter (http://hannonlab.cshl.edu/fastx_toolkit), allowing for Q-score of 20 or higher for ≥90% of the reads. We aligned filtered reads to a draft reference genome using the Burrows-Wheeler Aligner (Li & Durbin, 2009). The unpublished draft genome was derived from a single diploid individual from the northwest part of the native range (location A19, see supplementary materials for the location information). This same population was included in our current study. The species is diploid, with a gametic chromosome number of 18 (http://www.tropicos.org/Project/IPCN) and a genome size of ~1,135 Mbp (Kubešová *et al.*, 2010). Multiple whole genome shotgun libraries were sequenced at 110X coverage using 100 bp paired end reads and two Dovetail Chicago libraries were sequenced (14.5X coverage) to create the draft genome assembly used in this study (Putnam *et al.*, 2016). All sequencing was conducted on the Illumina HiSeq2000. The genome was assembled using Meraculous (Chapman *et al.*, 2011) and HiRise for scaffolding

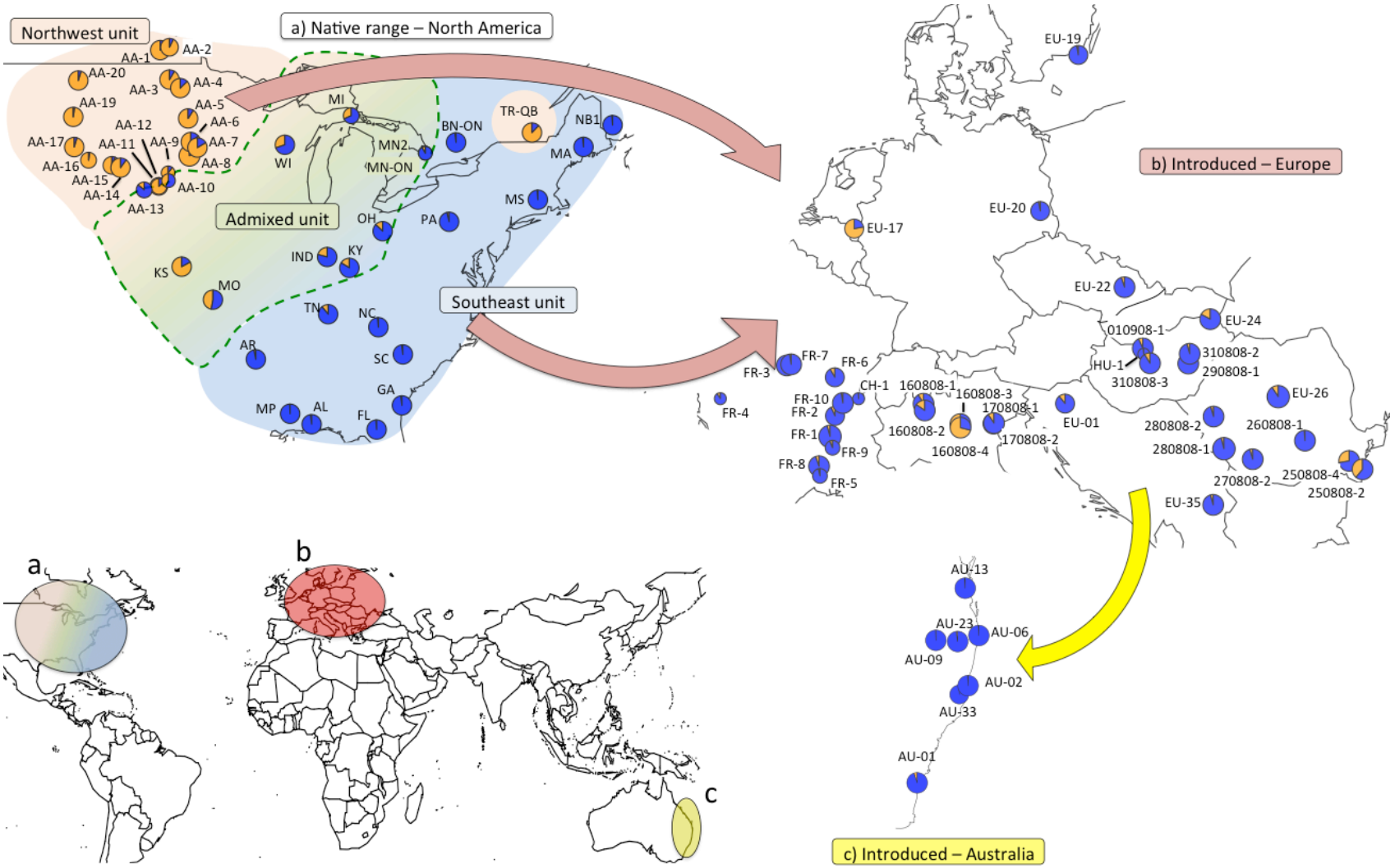


Figure 1. Pie charts of STRUCTURE assignment for K=2 (Figure 5) for *Ambrosia artemisiifolia* sampling locations in North America (a) Europe (b) and Australia (c) with invasion routes as inferred from ABC-RF. Assignment to genetic units for ABC-RF is shown in (a). Pie sizes are proportional to number of samples per sampling location. Geographic distances are not to scale.

(Putnam *et al.*, 2016). The scaffold N50 is 522 Kb in 807 scaffolds, while the N90 is 88 Kb in 3,190 scaffolds. The total scaffold number is 16,702, with a total length of 1,420 Mbp, and with called bases over 893 Mbp (37% Ns). Local realignment around indels was implemented with Picard (http://picard.sourceforge.net) and GATK (McKenna *et al.*, 2010). We called variants with GATK UnifiedGenotyper at the quality threshold of a Q-score ≥50 and used GATK hard filtering of variants; a minimum quality by depth of 2, a maximum Fisher-Strand bias of 60.0, minimum mapping quality rank sum test of -12.5, minimum root mean square mapping quality of 40.0, and a minimum read position rank sum test of -8.0. We subsequently filtered variant calls using VCFtools (Danecek *et al.*, 2011) and custom scripts on a genotype and variant quality of ≥20, depth of 5-240 and a minor allele frequency of 0.05. Finally, we removed 50 loci displaying heterozygosity frequencies of >0.7 to remove potential paralogues. This latter filtering step is lenient in regard to heterozygosity frequency expectations under Hardy-Weinberg equilibrium (with a maximum heterozygous genotype frequency of 0.5 for biallelic loci), and had no considerable effect on F_{IS} values (results not shown). We identified a total of 10,100 polymorphic SNPs with a call rate of 50% or more. We then selected a total of 1022 unlinked biallelic SNPs by shuffling the full SNP table and randomly drawing from each contig as STRUCTURE requires the use of unlinked loci for clustering (Pritchard *et al.*, 2000). To streamline between interdependent methods, we used this unlinked SNP set to select and subset genetic units (STRUCTURE, PCA, pairwise F_{ST} analyses) and to reconstruct invasion history using ABC.

Genetic clustering

We inferred population genetic structure with STRUCTURE v2.3.4, a Bayesian clustering method that allocates individuals into clusters on the basis of their genotypes (Pritchard *et al.*, 2000). We ran STRUCTURE on 1022 unlinked SNPs of 466 individuals from 85 sampling locations. Additionally, we ran analysis on subsets of the data to explore sub-clustering within each continent. We performed the analysis using the admixture model, correlated allele frequencies, no location prior, an uniform alpha individually defined for each population (Wang, 2017) for the number of clusters (K) ranging from 1 to 10, with 20 independent runs per K. Each run comprised of a burn-in of 200,000 followed by 1,000,000 iterations. StructureHarvester v0.6.94 (Earl & vonHoldt, 2011) was used to format the STRUCTURE output. We used log probability and delta K statistic to determine the uppermost clustering level (Evanno *et al.*, 2005). Nevertheless, we present STRUCTURE results for K from 2 to 8 in order to evaluate differentiation at higher levels of K in the supplementary materials. We processed the 20 runs using the Greedy algorithm in CLUMPP v1.1.2 (Jakobsson & Rosenberg, 2007) testing 1000 random input order repeats per K. We found no evidence for multimodality for the most likely K (2) in the global dataset as tested using CLUMPAK (Kopelman *et al.*, 2015).

Finally, we visualized data with DISTRUCT v1.1 (Rosenberg, 2003). To further explore relationships between native and introduced sampling locations, we summarized genetic differentiation using a principle component analysis (PCA) within the adegenet package (Jombart, 2008). We applied this method on population means for the complete dataset to examine genetic variation both within and between continents. Missing values were substituted with mean allele frequencies for the PCA (Jombart, 2008).

ABC and population pooling

We applied approximate Bayesian computation (ABC) (Beaumont *et al.*, 2002) using DIYABC v2.1.0 (Cornuet *et al.*, 2008; Cornuet *et al.*, 2014) to simulate provenance from non-admixed versus admixed source populations, as well as post-introduction admixture during multiple introductions for both the introduced ranges of Europe and Australia and bridgehead introduction from Europe into Australia. Briefly, within ABC posterior probabilities of different invasion scenarios are obtained by comparing the observed dataset to a large number of simulated datasets defined under invasion models given a set of demographic and historical parameters (Cornuet *et al.*, 2008). This approach has been shown to be successful in inferring demographic history in systems that have undergone multiple introductions, bottlenecks and/or admixtures (Estoup & Guillemaud, 2010).

The number of scenarios to be compared within the ABC framework increases drastically with the number of potential source and target population units included in the analysis (Estoup & Guillemaud, 2010; Lombaert *et al.*, 2014). To reduce the complexity of the invasion scenarios, we pooled sampling locations into units showing genetic similarity. For the native range, we defined genetic clusters based on global STRUCTURE assignment scores at K=2, the uppermost clustering level (see Results). As the central area in the native range constituted of admixed populations according to global STRUCTURE results (see Results), we were interested in the possibility that this admixed native source had contributed to the introductions. Consequently, we defined three prior genetic units in the native range: southeast (SE), northwest (NW) and admixed (i.e. pre-introduction admixture, AD). Genetic units SE and NW consisted of sampling locations with average proportion of membership to the first STRUCTURE cluster (Q) below 0.15 and above 0.85 respectively, and the admixed unit consisted of sampling locations with $0.15<Q<0.85$ (averaged over all individuals within a sampling location).

Genetic diversity within native and introduced ranges

All analyses were conducted in *R* v3.2.3 (R Core Team, 2017) unless stated otherwise. We calculated genetic diversity indices (allelic richness (A_R), gene diversity (H_S), observed heterozygosity (H_O) and inbreeding coefficients (F_{IS})) for each sampling location using the

hierfstat package (Goudet, 2005). We calculated means and confidence intervals for these indices over all 10,100 loci, excluding loci with <4 individuals and excluding sampling locations with >90% missing values over all loci (9 loci excluded). We summarized A_R, H_S and H_O by averaging across populations within each geographic range (i.e. North America, Europe and Australia) and within the genetic units defined for ABC analyses. We tested for differences among the geographic ranges in allelic richness and F_{IS} using means for each sampling location by implementing the Kruskal-Wallis test and post-hoc Nemenyi test for pairwise multiple comparisons (PMCMR package, Pohlert, 2014).

Genetic differentiation

To identify population genetic differentiation, we estimated Weir & Cockerham's pairwise F_{ST} between sampling locations and between genetic units defined for ABC over the 1022 unlinked SNP dataset using the diveRsity package (Keenan *et al.*, 2013). To test for genetic dependence (relatedness) of sampling locations caused by isolation by distance (IBD) within each geographic range and clusters defined for ABC analysis including all sampling locations. We tested associations between $F_{ST}/(1-F_{ST})$ and log-transformed geographic distances within sampling locations using a Mantel test with 1000 replicates using the ecodist package (Goslee & Urban, 2007). We assumed the native admixed genetic unit did not falsely result from IBD in the native range, as IBD was not significant within the native range as a whole (Results, Table 1) or within the NW and SE cluster. According to the global STRUCTURE results, we pooled European and Australian samples to a single genetic unit within each range (see Results). Pooling diverged sampling locations as a single unit can alter conclusions drawn from ABC (Lombaert *et al.*, 2014) and introduce the Wahlund effect. Accordingly, we evaluated robustness of the scenario choices by building reduced datasets that excluded divergent sampling locations (see Appendix S2 & Table S2, supporting information).

ABC simulations with RF evaluation

We selected among simulated introduction scenarios using ABC-RF (Pudlo *et al.*, 2016), which is a novel approach based on random forest machine learning algorithms. This tool has been shown to outperform existing ABC model selection techniques in precision, computation time and robustness (Marin *et al.*, 2016; Pudlo *et al.*, 2016). A large set of summary statistics can be used, as RF is robust against common issues encountered in other methods related to the 'curse of dimensionality' (Pudlo *et al.*, 2016), such as collinearity (Estoup & Guillemaud, 2010; Blum *et al.*, 2013) and 'noise' in the data (Marin *et al.*, 2014; Marin *et al.*, 2016). Finally, RF can provide more accurate results using only a low number of simulated datasets per scenario (10^3-10^4, Marin *et al.*, 2016; Pudlo *et al.*, 2016; Fraimout *et al.*, 2017) compared to traditional model evaluation approaches (10^5-10^6, Bertorelle *et al.*, 2010). This last advantage become evident in

the current study, as CPU time required for the large dataset (1022 SNPs and 466 individuals, simulating 10^4 datasets per scenario) amounted to 420 hours for the main datasets only. Together with replicate simulations required for reduced datasets, such analyses would not be feasible using traditional methods (Pudlo *et al.*, 2016; Fraimout *et al.*, 2017; Momigliano *et al.*, 2017).

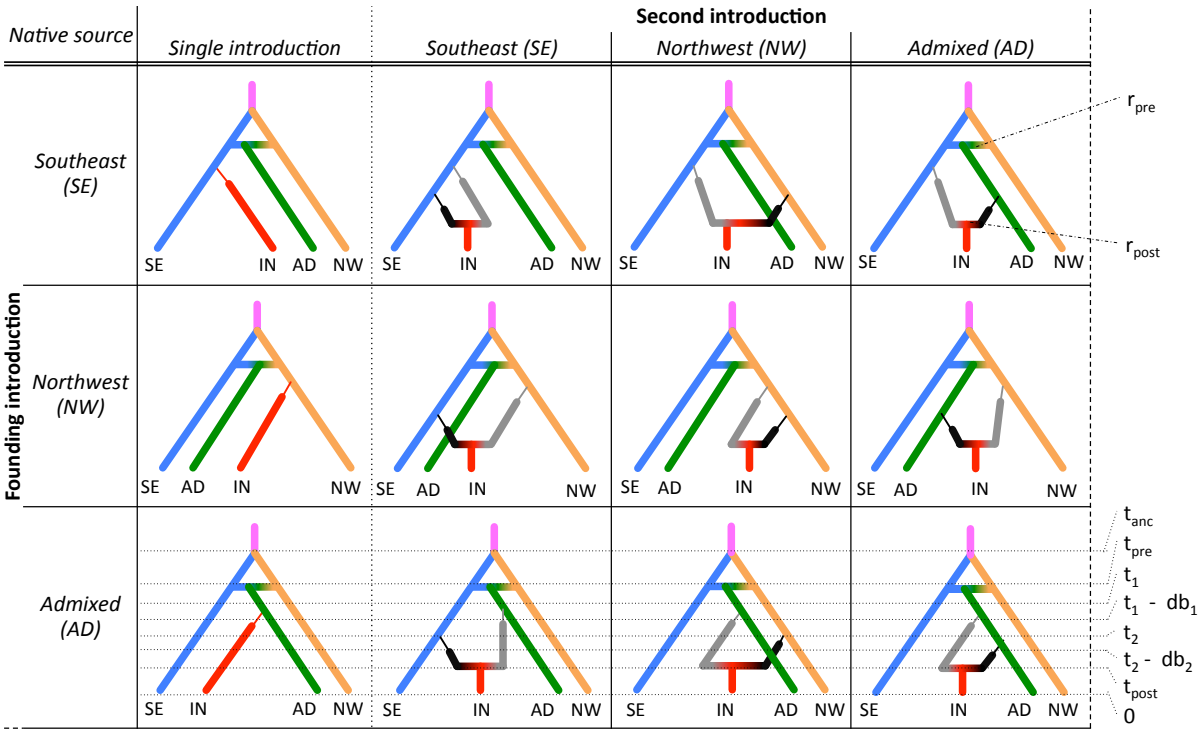


Figure 2. Graphical illustration of *Ambrosia artemisiifolia* introduction scenarios from divergent native North American genetic units (SE, NW and AD) during initial (rows) and secondary (columns) introduction, tested using ABC-RF for Europe and Australia introduced ranges (IN) independently (Table 3). Thin lines indicate bottlenecks of duration db_i with effective population sizes of Nb_i. For parameters descriptions and priors see Table S3, supporting information. Time is not to scale.

Reconstruction of demographic histories

To reconstruct the introduction history of *A. artemisiifolia* from its native North American range into either Europe or Australia, we considered introduction scenarios for each range separately with initial, possibly bottlenecked, introductions from the SE, NE or AD genetic units. In the face of repeated introductions, as is prevalent in *A. artemisiifolia* (Chauvel *et al.*, 2006), we improved the first three single introduction scenarios from different native units by including a secondary introduction from native sources with a second possible bottleneck, leading to a total of 12 scenarios (Fig. 2). As Australian introduction is more recent than European introduction, we included a scenario stipulating a bridgehead invasion (Lombaert *et al.*, 2010)—a successful invasion (here Europe) acting as a source for further introductions (here Australia)(Fig. 3). To reduce the number of scenarios to be compared, we added a bridgehead invasion to the most likely European introduction scenarios according to ABC-RF

and tested this against the most likely Australian introduction scenario from the native range (see Results).

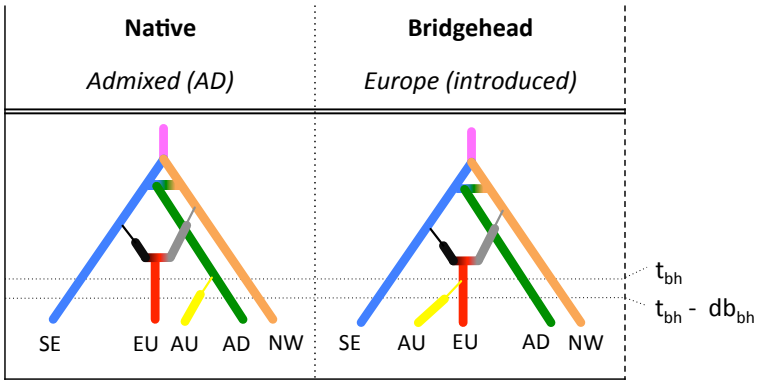


Figure 3. Graphical illustration of introduction scenarios of *A. artemisiifolia* to Australia from the native American admixed zone (AD) and a bridgehead introduction from Europe are based on the European and Australian introduction scenario receiving most votes (Table 3). Non-bridgehead parameters are as in Figure 2.

We set uniform priors on all model parameters (Table S3, supporting information) with the exception of timing of admixture in the native range. As native admixture increased rapidly in more recent years following deforestation and intensification of agriculture (Martin *et al.*, 2014; Martin *et al.*, 2016), prior sampling followed a log-uniform distribution to favour lower values. We specified prior lower bounds for ancestral divergence and upper bounds for the admixture in the native range based on pollen records, which are consistent with southeast and northwest native genetic units prior to 500 years before present (Williams *et al.*, 2004). We set parameter prior upper bounds for the timing of the initial introduction several years before the first known occurrence of *A. artemisiifolia* within the introduced ranges (Europe: 180 years before present (ybp); Australia: 120 ybp), with a lower bound set before the onset of known major secondary introductions (Europe & Australia: 100 ybp)(Chauvel *et al.*, 2006; Palmer & McFadyen, 2012). We set the upper prior bound of secondary introductions equal to that of the initial introduction—180 and 120 ybp for Europe and Australia respectively—with the condition that initial introduction always occurred prior to secondary introduction. This prior would allow a secondary introduction to practically coincide with or be temporally separated from the primary introduction. For Europe, we set the lower prior bound of subsequent introduction events to the end of the last major introduction during the Second World War (60 ybp, Chauvel *et al.*, 2006). For Australia, this prior was set to the last described population increase and spread (50 ybp, Palmer & McFadyen, 2012). Bottleneck priors were set so as to simulate no bottleneck (0 years) to a severe bottleneck (Europe: 0-60 years; Australia: 0-50 years). These priors are bound by the last known population increase as described above. We adjusted lower and upper limits of all other priors by evaluating posterior distributions of preliminary simulated datasets, setting the prior distribution as wide as possible but within biological reason (Bertorelle *et al.*, 2010). We assumed no migration between any of the genetic

units. We conducted analyses on the 1022 unlinked SNP dataset, specifying a MAF criterion of 0.05 for all simulations. We included all possible summary statistics (Table S4, supporting information) provided by DIYABC and ran 10,000 simulations for each scenario.

We evaluated model fit and posterior distributions of ABC simulations using the Random Forest (RF) approach implemented in the abcrf package v1.3 in R (Marin *et al.*, 2016; Pudlo *et al.*, 2016). For each ABC analysis, we grew a classification forest of 1,000 trees based on all simulated datasets. The abcrf package estimates a prior error (measure of classification of votes to the wrong scenario) for each analysis using the out-of-bag-errors (Appendix S2, supporting information). Then, the scenario with the highest classification vote is selected as the most likely scenario for which a posterior probability is calculated. To evaluate the global performance of our ABC scenario choice, we performed posterior model checking in DIYABC (Appendix S2, supporting information). Finally, we inferred posterior distribution values of parameters of the selected scenario under a regression by Random Forest methodology (Marin et al., 2016), with a classification forest of 1,000 trees (see Appendix S2 for additional ABC-RF description and evaluation).

Results

Genetic diversity within sampling locations

We found homozygous excess in a large number of geographic sampling locations within the native and two introduced ranges of *A. artemisiifolia* (Fig. 4, Table S1, Fig. S1, supporting information). Within the native range, homozygous excess was higher in northwestern sampling locations compared to southeastern locations (Latitude Spearman's rho=0.693, $p<0.001$; Longitude Spearman's rho=0.509, p=0.001). Sampling location allelic richness was significantly reduced in the introduced Australian range compared to the native range (Kruskal-Wallis $\chi^2=8.998$, df=2, p=0.011), but no significant difference was present between the European range and the native range (p=0.275), or both introduced ranges (p=0.124, Fig. 4).

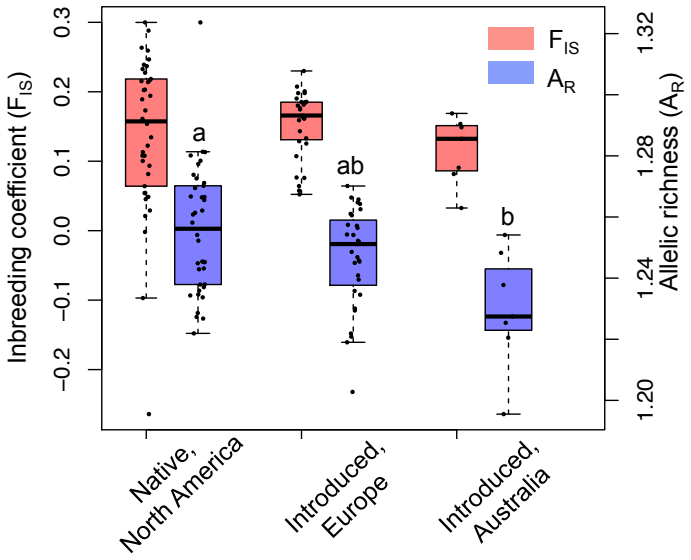


Figure 4. Inbreeding coefficients (F_{IS}, red) and allelic richness (A_R, blue) for 85 *Ambrosia artemisiifolia* sampling locations at the native North American and introduced European and Australian ranges. No significant differences between ranges were identified in F_{IS}. For A_R, letters depict significant differences between ranges (α=0.05).

Population structuring and differentiation

We found low genetic structure within the native and two introduced ranges of *A. artemisiifolia*, with a uppermost clustering level at K=2 in STRUCTURE (Fig. 5, Fig. S2, supporting information) according to the ΔK method (Evanno *et al.*, 2005). This analysis clearly identifies a southeastern and northwestern genetic unit in the native range, with most individuals within these units having assignment of Q>0.85 to their respective cluster. We refer to these genetic units as the southeastern and northwestern genetic units throughout. Individuals from 11 sampling locations geographically located between these two main genetic units showed mixed assignment to either cluster (0.151<Q<0.858). We refer to this unit as the admixed genetic unit (Fig. 1). Some sampling locations (7/36) within the introduced range of Europe showed intermediate cluster assignments, whereas all introduced Australian sampling locations showed assignment similar to southeastern native sampling locations. This assignment of Australian sampling locations and southeastern North American sampling locations to the same cluster was evident to some extent at higher K values (Fig. S2, supporting information). The PCA results reflected the patterns identified in STRUCTURE in that Europe showed most overlap with the admixed and southeast units, whereas Australia showed more extreme differentiation from the native range but some overlap with the southeast unit (Fig. S3, supporting information).

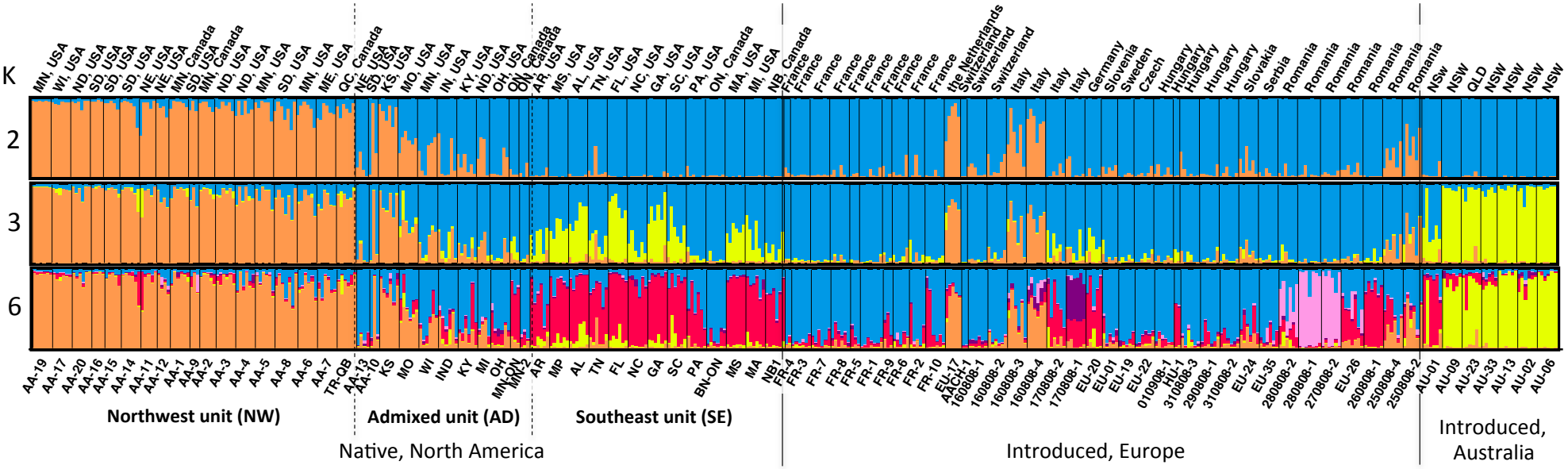


Figure 5. STRUCTURE output for K of 2, 3 and 6 for 20 replicate runs each on 1022 randomly selected unlinked SNPs within 466 *Ambrosia artemisiifolia* samples summarized with CLUMPP and visualized with Distruct. For full overview of STRUCTURE runs see Figure S2, supporting information.

Table 1. Population diversity statistics for *Ambrosia artemisiifolia* within native, introduced European and Australian ranges and genetic units defined for ABC analyses based on full datasets, averaged over sampling locations (± 95% confidence). Abbreviations represent: allelic richness (A_R); gene diversity (H_S); observed heterozygosity (H_O, bold values are significantly (α=0.05) different from H_S); and inbreeding coefficient (F_{IS}, bold values indicate significant (α=0.05) departure from zero).

	Summary statistics			**F_{ST}**	**IBD**	
Range	**A_R**	**H_S**	**H_O**	**Mean**	**r**	**p**
Native, full	1.256 (0.007)	0.262 (0.002)	0.213 (0.004)	0.052	0.017	0.770
Native, NE	1.243 (0.006)	0.252 (0.003)	0.176 (0.002)	0.032	0.137	0.253
Native, AD	1.270 (0.018)	0.273 (0.009)	0.243 (0.015)	0.033	**0.253**	**0.028**
Native, SE	1.264 (0.009)	0.267 (0.001)	0.243 (0.002)	0.066	0.012	0.929
Introduced, Europe	1.247 (0.006)	0.254 (0.002)	0.199 (0.002)	0.059	0.023	0.683
Introduced, Australia	1.230 (0.014)	0.235 (0.002)	0.196 (0.002)	0.116	0.072	0.838

To reveal subpopulation structure within each range, we conducted the structure analysis separately for North America, Europe and Australia (Fig. S2, supporting information). For North America the uppermost clustering level was two, although we found higher clustering emerging in the European subset at K=4 and K=6 (represented by peaks in ΔK), which could be ascribed to a few distinct sampling locations in Romania, Italy, the Netherlands and Germany. A peak in ΔK appeared in the Australian subset at K=7.

We found low genetic differentiation within the native and two introduced ranges (Table 1, Table S5). Genetic differentiation was highest in the introduced Australian range (F_{ST}=0.116), which was also reflected by a wide spread of Australian individuals in the PCA (Fig. S3, supporting information), and a high number of pairwise F_{ST}>0.1 within this range (Table S5, supporting information). We found no evidence of IBD within the native range (r=0.017, p=0.770) as a whole, although IBD was present within the admixed genetic unit (r=0.253, p=0.028). This would be expected if the amount of admixture between the clusters was a function of geographic distance to the SE or NW cluster. No such pattern was apparent within the introduced ranges (Europe: r=0.023, p=0.683; Australia: r=0.072, p=0.838) or within the main native genetic units (northwestern: r=0.137, p=0.253; southeastern: r=0.012, p=0.929; Table 1). Differentiation between North American genetic units was low for all pairwise comparisons (F_{ST}=0.013–0.034)(Table 2). Europe was not highly differentiated from the native units (F_{ST}=0.004–0.021), though Australia showed higher differentiation from the native units (F_{ST}=0.028–0.052)(Table 2).

Table 2. Pairwise F_{ST} for genetic units defined for ABC analyses based on full datasets.

	Native, northwest	Native, admixed	Native, southeast	Introduced, Europe	Introduced, Australia
Native, NE		0.016	0.034	0.021	0.052
Native, AD	0.016		0.013	0.004	0.038
Native, SE	0.034	0.013		0.012	0.028
Introduced, Europe	0.021	0.004	0.012		0.034
Introduced, Australia	0.052	0.038	0.028	0.034	

Invasion history of Ambrosia artemisiifolia using ABC-RF

The European invasion was characterised by multiple introductions (concurrent or temporally separated) from the southeast and northwest native North American range as revealed by ABC-RF model selection. The invasion scenario receiving most votes (23.4%) described an initial introduction sourced from the NW and a secondary introduction from the SE genetic units (posterior probability *P*=0.501, Table 3). Overall, European invasion scenarios including an introduction of the native admixed source (6/12 scenarios) received 19.9% (Table 3) of the votes opposed to 80.1% (Table S6, supporting information) of the votes received by scenarios not including an admixed source, suggesting pre-introduction admixture did not play a large role in shaping the European invasion. The admixed region of the native range contributed to a

large extent to the Australian populations, where the single-introduction scenario received the most ABC-RF votes (44.2%, *P*=0.599, Table 3). With the addition of a bridgehead into Australia however, the bridgehead introduction was found to be most likely (92.6% of votes, *P*=0.998, Table 3). This scenario including a bridgehead introduction summarises the 'best' introduction scenarios from preceding analyses, and was used for posterior parameter inferences. ABC-RF inferences were robust to data subsetting (Table S6, supporting information).

Table 3. Summarized results of ABC-RF model selection for introduction scenarios of *Ambrosia artemisiifolia* from divergent native sources of North America (northwest NW, admixed AD and southeast SE) during single or multiple introductions into Europe (a) and Australia (b) (Fig. 2) and single introduction from native AD and bridgehead Europe (EU) into Australia (Fig. 3). Scenarios in c are based on the European and Australian introduction scenario receiving most votes (a&b). Analyses consisted of a number of competing scenarios, summarized with all summary statistics available in DIYABC, for which a prior error rate was computed. The best scenario was selected based on a number of random forest (RF) votes, where the RF posterior probability gives the degree of confidence in this selected scenario.

Analysis	Number of scenarios	Number of summary statistics	Best scenario	Number of RF votes (/1000)	Prior error	RF posterior probability
a) European introduction from NW, AD and/or SE	12	112	Multiple introductions from SE & NW	234	44.01%	0.501
b) Australian introduction from NW, AD and/or SE	12	112	Single introduction from AD	442	47.30%	0.599
c) Australia single introduction from AD or EU	2	220	Single introduction from EU	926	0.10%	0.998

Most posterior parameter estimates for the scenario receiving most votes were quite wide and close to the set prior boundaries (Table S7, supplementary information)—a common finding in ABC (Csilléry *et al.*, 2010). Nevertheless, some important inferences could be made from a select few posterior parameter distributions. Firstly, the formation of the native admixture zone was predicted to be around 218 years ago (median, Table S7, supplementary information), well before the first reported introduction of *Ambrosia artemisiifolia* into Europe, suggesting this zone could have contributed to invasive populations worldwide. Moreover, the relatively narrow posterior parameter distribution of the Australian introduction from the bridgehead Europe suggests this single introduction was followed by long bottleneck period (d_{bh}) of at least 15 years (2.5% quantile) with a relatively small founding population (Nb_5) of maximum 564 individuals (97.5% quantile). Conversely, bottleneck size (Nb_6 & Nb_7) and duration (db_1 & db_2) for the European introduction from the native North America were approaching set prior boundaries and therefore not conclusive. Finally, this bridgehead scenario

shows narrow posterior distributions of admixture rates, suggesting a median 0.76 contribution of SE to the admixed native source AD (r_{pre}) and 0.74 contribution (1-r_{post}) of a founder population derived from the SE unit to Europe and subsequently Australia (Table S7, supplementary information).

Discussion

We compared the contribution of pre-introduction admixture, multiple introductions and bridgehead invasion of the global weed *A. artemisiifolia* within two of the introduced ranges. ABC simulations indicated admixture following multiple introductions from distinct genetic clusters found in North America played a significant role to the successful European invasion of this weed, subsequently acting as a bridgehead for the Australian introduction. Our study emphasizes the need for careful examination of the population structure and demographic history of invasive species, as introduction pathways–such as bridgehead introduction–may affect evolutionary processes in the introduced ranges. This type of human-assisted spread may become more common, particularly for species like *A. artemisiifolia*, which is frequently associated with human-modified habitats and intensive agriculture.

Genetic diversity and differentiation in native and introduced ranges

The introduction history of common ragweed in Australia has not yet been the subject of much investigation. Although common ragweed has persisted on this continent for over 100 years, it has not become as geographically widespread as its European counterpart. We found lower allelic richness in Australia compared to the native range, while no significant reduction in allelic richness was found in Europe. In addition, Australian sampling locations were more differentiated from each other and from the native and European ranges compared to the other ranges. This result suggests less frequent and/or more bottlenecked Australian introduction(s), as is confirmed by ABC-RF parameter estimation, showing a single Australian introduction followed by a long bottleneck of >15 years. This corresponds to the known colonisation history of *A. artemisiifolia,* as repeated introductions of this invader within Europe commenced during the 19th century (Chauvel *et al.*, 2006), whereas the Australian introduction was more recent (Palmer & McFadyen, 2012). Moreover, Australia introduced strict quarantine laws (c. 1908, Quarantine Act, 1908) potentially limiting the frequency of introduction and size of founder populations.

Our results show that a large number of sampling locations were heterozygote deficient, which is consistent with previous studies of *A. artemisiifolia* genetic diversity (Chun *et al.*, 2010; Gladieux *et al.*, 2010; Chun *et al.*, 2011; Gaudeul *et al.*, 2011; Martin *et al.*, 2014). Northwestern locations within the native range showed higher heterozygote deficiency compared to southeastern locations. A similar longitudinal pattern in *A. artemisiifolia* observed

by Martin *et al.* (2014) has been interpreted by the authors to be a result of higher rates of selfing and/or biparental inbreeding due to local pollen and seed dispersal in the smaller, low density, isolated western populations opposed to outbreeding in larger, high density, inter-connected eastern populations. A shift in mating system could arise in low-density conditions or during colonization (Barrett *et al.*, 2008) and might explain the geographic pattern in F_{IS}. However, *A. artemisiifolia* in Ontario exhibit high levels of outcrossing and a sporophytic self-incompatibility system (Friedman & Barrett, 2008) and no evolutionary shift towards selfing has been found in the introduced range of China (Li *et al.*, 2012). The Wahlund effect resulting from local sub-structuring could be a cause for the observed heterozygote deficiency. However, pollen dispersal within this species is likely high as it is wind pollinated (Martin *et al.*, 2009) and so sub-structuring might be expected to be limited within each sampling location (Genton *et al.*, 2005a; Martin *et al.*, 2014). These contrasting observations remain curious, and geographic variation in selfing and biparental inbreeding has yet to be investigated in this species.

Pre- versus post- introduction admixture during invasion

We found some mixed assignment of European individuals in STRUCTURE and PCA and population allelic richness equal to the native range. This complex European genetic structure has been previously observed and was accredited to multiple introductions from various source populations and post-introduction admixture (Genton *et al.*, 2005a; Chun *et al.*, 2010; Gaudeul *et al.*, 2011), which is a common conclusion for invasive species with admixed genetic backgrounds (Dlugosch & Parker, 2008a; Facon *et al.*, 2008; Uller & Leimu, 2011; Cristescu, 2015). Conversely, a human-mediated intra-specific hybrid zone within native North America has been proposed as a possible introduction source (Martin *et al.*, 2014).

According to ABC-RF posterior parameter analyses, the pre-admixed native zone originated well before the first reported introductions (>180 years ago), suggesting this zone could have been an introduction source for the global *A. artemisiifolia* invasions. However, ABC-RF analysis encompassing diverse introduction scenarios support the establishment of European populations through multiple introductions of the two non-admixed native units, with less likely contribution of native pre-introduction admixture. Although the analysis without a bridgehead scenario suggested Australia originated from the native admixture zone, further investigation indicated the invaded European range instead acted as a bridgehead for the single introduction into Australia. Our results underline the importance of testing combinations of contrasting introduction scenarios simultaneously.

Using simplified versions of the true demographic histories is a widely recognised limitation of ABC inference (e.g. Csilléry *et al.*, 2010) and could possibly contribute to the relative low posterior probabilities (0.501-0.599) found for scenarios depicting introduction

from the native range as only source (i.e., the non-bridgehead scenarios). Similar posterior probabilities have been reported in other studies, with contrasting interpretations by the authors (Pudlo *et al.*, 2016; Fraimout *et al.*, 2017; Momigliano *et al.*, 2017).

In our study, we grouped sampling locations based on relatively homogeneous genetic clusters identified by STRUCTURE (NE mean F_{ST}=0.032; SE mean F_{ST} =0.066) with no evidence of isolation by distance within these two genetic clusters. Although it is possible more complex cluster schemes could support different introduction histories, our results were consistent among stringently filtered subsets where divergent populations within each cluster were excluded. Moreover, the analysis was powerful in distinguishing between single versus multiple introductions and pre- versus post-introduction admixture or bridgehead invasion, but we cannot exclude the existence of additional introductions from other sources. For example, introductions might have been sourced from native locations that were not captured in our broad sampling of the North American range (e.g., south Florida). In addition, the Australian introduction, which occurred more recently than the European introduction, may have had contributions from another as yet unsampled introduction such as Asia, South America or Eastern Europe. Other possibilities include re-introductions from Europe or anywhere else in the world back into the native North America. Given the very short time since the initial invasions, we find it unlikely that a large number of non-native individuals could have had a significant impact on the well-established native populations. Moreover, we are limited in the complexity of the models, and could not explore the myriad of possibilities imaginable through ABC simulations.

The discovery of multiple introductions and admixture or bridgehead invasions to the non-native European and Australian populations does not confirm the causality of these processes in driving successful invasion in *A. artemisiifolia.* Patterns discovered in this study might simply reflect historical trade routes and admixture may not have contributed to the invasiveness of this species. However, our finding of admixture as a predominant feature in Europe through multiple introductions and in Australia through a bridgehead introduction raises the question of fitness benefits of hybridisation to a successful invasion. Moreover, given the global nature of the invasion, the native admixture zone should be considered as potential source for non-sampled invasive introductions in Asia and elsewhere. Although the beneficial consequences of pre- and post- introduction admixture are similar (described below), the role of pre-admixed source populations as a pathway to invasion has been given less attention in the literature than multiple introductions and admixture post-invasion (Keller & Taylor, 2010; Keller *et al.*, 2014; Rius & Darling, 2014).

The immediate fitness benefits of admixture are most widely established to be a consequence of heterosis, stimulating invasiveness through increased growth and vigour, which may in turn ‘catapult’ populations past the lag phase of colonisation (Facon *et al.*, 2005; Drake,

2006; Keller *et al.*, 2014; Wagner *et al.*, 2017). Through this process, admixed populations could outperform single-source invasions, as has been shown in an experimental study on the beetle *Callosobruchis maculatus* (Wagner *et al.*, 2017). Within an experimental study of *A. artemisiifolia*, certain native crosses displayed heterosis whereas other crosses between native populations had no fitness impacts or led to mild outbreeding depression (Hahn & Rieseberg, 2017). By contrast, no heterosis was found in crosses among the European populations. This heterosis experiment aligns with the genomic analysis presented in the current study, as crosses between non-admixed native genetic clusters (here northwest X southeast) were more likely to display heterosis as opposed to crosses to the historically admixed populations or between European populations. In addition, populations from the northwestern cluster in the native range consistently underperformed in traits associated with invasive success compared to introduced European populations in common gardens under a range of environments (Hodgins & Rieseberg, 2011).

Admixture could also facilitate adaptation through increased genetic variation (Anderson & Stebbins, 1954) and creation of novel or transgressive phenotypes (Stebbins, 1969; Rieseberg *et al.*, 1999). Patterns of genetic differentiation of traits, such as latitudinal variation in flowering time have evolved repeatedly in *A. artemisiifolia* following introduction (Chun *et al.*, 2011; Hodgins & Rieseberg, 2011; Li *et al.*, 2014) and are consistent with rapid local adaptation within the introduced ranges. In Europe in particular, no evidence of substantial bottlenecks limiting genetic variation are evident. This raises possibility that the influx of genetic material during introduction fuelled rapid local adaptation, perhaps enhancing the rate of spread of the invader in Europe. Even in Australia, where allelic richness is significantly reduced relative to the native range, the ABC analysis is consistent with a European source population, which was likely admixed.

Future directions

Mechanisms that could promote invasiveness through bridgehead invasions could also play an important role in the success of pre-introduction admixture. We recommend that admixed source populations should be considered as a process potentially impacting invasiveness both for initial and subsequent colonisations. In addition, this study underlines the necessity of more intricate testing of introduction scenarios (Cristescu, 2015), as classical population approaches might prove to be misleading (Lombaert *et al.*, 2014; Falush *et al.*, 2016). For example, the role of the European invasion as a source of the Australian introduction would not have been considered without the ABC analysis. Identifying the specific source populations of introduced species might play a role in predicting its invasive success and could mean management effort should be focused on preventing introductions from admixed or multiple divergent source populations.

Better understanding of the importance of admixture to invasion requires: *i)* an accurate reconstruction of invasion routes (Cristescu, 2015); *ii)* examination of admixture costs and benefits within native and invasive populations; and *iii)* fitness assessments of admixed versus non-admixed genotypes in the invasive range(s) over multiple generations (Cristescu, 2015). As our data elucidate the first of the three requirements, future studies need to further explore the costs and consequences of admixture for invasion using genomic and experimental data in this highly diverse and widespread invader.

Acknowledgements

We would like to thank R. Baucom and three anonymous reviewers for their invaluable comments, J. Stephens and A. Wetherhill for sample collection, G. Boinnard and S. Mouw for assistance in the lab, P. Sunnucks and C. Sgrò for early manuscript suggestions and M. Kourtidou for graphical assistance. We would also like to thanks Mike Martin and Dovetail Genomics for their assistance with the assembly of the unpublished draft genome. Funds were provided by a Monash University Dean's International Postgraduate Research Scholarship to LAB, a Monash University Startup Grant to KAH and a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery grant 327475 to LHR. B. G. was supported by the AgreenSkills+ fellowship programme which has received funding from the EU's Seventh Framework Programme under grant agreement N° FP7-609398 (AgreenSkills+ contract).

Data accessibility

Sequence data is available at the National Center for Biotechnology Information (NCBI) Sequence Read Archive under Bioproject PRJNA374597.

Author contributions

LB and KH conceived the project with input from EL, BG and LR. KH, LB and KN collected samples, KN performed molecular lab work. LB analysed data and drafted manuscript. All authors contributed to and approved the final manuscript.

Supporting information to Chapter 2

Appendix S1. Genotype-by-sequencing protocol adapted from Elshire et al. (2011).

We added 100 ng of high-quality DNA in 11.7 µL water to 2.5 uL Buffer 4 (NEB), 0.5 uL Pst1 HF (NEB), 0.5 uL BSA (NEB), 0.8 uL T4 DNA Ligase (NEB), 5 uL Ligase Buffer (NEB), 4.5 uL barcoded adapters and 4.5 uL common adapter. We digested and ligated samples simultaneously for 3h at 37°C, 3h at 22°C and 20 minutes at 65°C. After incubation, we transferred 30 µL from each well into two 2.0mL tubes, and cleaned up these tubes using the three columns of the QIAquick PCR Purification Kit (Qiagen, Valencia, CA). We eluted the purified product in 150 µL TE buffer. We amplified 8 reactions each with 1 µL of elution and 10.5 µL H2O, 12.5 µL Phusion High-Fidelity PCR Master Mix with HF Buffer, 0.5 µL of each GBS PCR primers A&B. We cycled reactions 3 min at 94°C, followed by 18 repeats of 30 s at 94°C, 30 s at 65°C, 30 s at 68°C with a final step of 10min at 68°C. To each reaction, we added an additional 1 µL Phusion High-Fidelity PCR Master Mix with HF Buffer and 0.5uL of each GBS PCR primers A&B and cycled the reactions for 3 min at 94C, 2 min at 65C, 12 min at 68C and hold at 4C. We cleaned the PCR product using QIAquick PCR Purification Kit (Qiagen, Valencia, CA) and eluted in 30 uL H2O. We performed a size selection by running the cleaned PCR product on a 2% agarose gel and removing the 400 bp to 600 bp fragment. This gel fragment was cleaned up using QIAquick Gel Extraction Kit (Qiagen, Valencia, CA). and eluted in 20 uL H2O. The five GBS libraries were paired-end sequenced by the Biodiversity Sequencing Centre at UBC on the Illumina HiSeq 2000 platform, each on a single lane.

Appendix S2. Additional ABC-RF analyses

Exclusion criteria applied to ABC-RF data subsets

Pooling diverged sampling locations as a single unit can alter conclusions drawn from ABC (Lombaert *et al.*, 2014) and introduce the Wahlund effect. Accordingly, we evaluated robustness of the scenario choices by building reduced datasets, removing more differentiated sampling locations from each genetic unit. Genetic structure was low, particularly in native range (see Results, main text), and few sampling locations were highly differentiated (see Table S5). We thus first applied a stringent filtering and removed sampling locations with pairwise F_{ST}>0.1 to any other sampling location within each unit. This allowed us to test the impact of high levels of substructure on the outcome of the ABC analysis. These datasets included 39 out of 42 North American sampling locations and 28 out of 36 European sampling locations. As many Australian sampling locations were highly divergent (pairwise F_{ST}>0.1), we simulated Australian demographic history using two separate units of sampling locations based on significant sampling location pairwise F_{ST}. For the first group (hereafter "group 1") we included AU-01, AU-02, AU-09 and AU-23; the second group (hereafter "group 2") included AU-06, AU-09 and AU-23. A second filtering step looked at congruent clustering in STRUCTURE and

PCA analyses (39/42 native; 23/36 European; and 2/7 Australian sampling locations. This arbitrary cut-off allowed us to further explore the effect of removing populations that were genetically distinct from the other unit members).

For Europe and North America the datasets consisted of the following:

i) All populations without removing highly divergent populations (no F_{ST} cut-off for either range)(results reported in main text);

ii) Pairwise F_{ST}<0.1 within the same genetic unit;

iii) Pairwise F_{ST}<0.1 within the same unit, as well as a population assignment to a unique structure cluster (Q>0.5) at K=6 (2^{nd} ΔK peak) in the full structure analysis;

Australian and North American populations:

i) All populations without removing highly divergent populations (no F_{ST} cut-off for either range)(results reported in main text);

ii) Pairwise F_{ST}<0.1 within the same unit, group 1 (AU-01, AU-02, AU-09 and AU-23);

iii) Pairwise F_{ST}<0.1 within the same unit, group 2 (AU-06, AU-09 and AU-23);

iv) Pairwise F_{ST}<0.1 within the same unit, group 1, as well as a population assignment to a unique structure cluster (Q>0.5) at K=7 (2^{nd} ΔK peak) in the structure analysis of the Australian populations.

Finally, we ran the ABC-RF simulations for the Australian, European and North American sampling locations, including a 'bridgehead' invasion from Europe to Australia. The datasets were defined as follows:

i) All populations without removing highly divergent populations (no F_{ST} cut-off for any range)(results reported in main text);

ii) Pairwise F_{ST}<0.1 within the same unit, with Australian group 1 (see above);

iii) Pairwise F_{ST}<0.1 within the same unit, with Australian group 2 (see above);

iv) Pairwise F_{ST}<0.1 within the same unit, with Australian group 1, as well as a population assignment to a unique structure cluster as described above (Q>0.5 at K=6 for Europe and at K=7 for Australia).

Excluded sampling locations are shown in Table S3, numbers of datasets correspond to ABC-RF results of European, Australian and bridgehead introduction (Table S6). Population sub-structure does not seem to have a large effect on the conclusions drawn by ABC-RF, as datasets

with different levels of differentiation yielded similar results (Table S6). A low effect of inbreeding was similarly reported in an ABC simulation-based study (Lippens *et al.*, 2017).

ABC-RF as a new tool to investigate demographic history

The random forest machine learning algorithm is a special case of classification and regression trees (CART), where *n* number of trees (the forest) are being grown through classification of the bootstrapped data (here the simulated datasets) using predictors (here the summary statistics, Table S5) (Breiman, 2001; Pudlo *et al.*, 2016). The majority vote of classification of all trees is then used to determine the ABC scenario best suited to the data (Pudlo *et al.*, 2016). Data excluded from each bootstrap sample are used to produce the 'out-of-bag' predictions (OOB), which are unbiased accuracy estimates. The OOB predictions can subsequently be used to determine the importance of predictor variables used for tree classification (Breiman, 2001).To determine most likely invasion scenario and summary statistic variable importance we grew 1000 trees based on all simulated datasets, including linear discriminant analyses (LDA) scores (Estoup *et al.*, 2012) as the latter step has been shown to decrease error rates (Pudlo *et al.*, 2016). Prior analyses showed growing larger forests and/or including a higher number of simulations did not further reduce OOB.

Model checking

To evaluate the global performance of our ABC scenario choice, we performed posterior model checking in DIYABC by simulating 10^6 datasets under these best invasion scenarios, from which we selected a posterior sample of 10^4 values closest to the observed data. We used a logit transformation of parameters and included all summary statistics and simulated 10^4 new datasets with parameter values drawn from the posterior sample (see Cornuet *et al.* (2010) and Fraimout *et al.* (2017) for detailed description of method).

The 'best' European introduction scenario showed 12 of the 112 (10.7%) observed values lay in the either tail of the distribution ($p<0.05$ or >0.95) of simulated values after Benjamini and Hochberg's sequential p-value correction for multiple comparisons. Graphical representation of these results in a PCA show the observed dataset falls well within the parameter space of the simulated datasets for the first four principle components, which together explain 70.4% of the total variation (Figure S4). The 'best' Australian introduction scenario (preliminary analysis, from the pre-admixed native source) showed 6.3% (over 112 summary statistics) of observed values falling into the tails of the distribution of simulated summary statistics, with the observed values falling well within the simulated PCA of the first 4 PC axes, cumulatively explaining 73.1% of total variation (Figure S4).

The final scenario based on these best European and Australian invasion scenarios, including a bridgehead invasion from Europe into Australia, showed 11 (0.05%) out of the 220

observed values fell into the tail of the distribution of simulated summary statistics. Again, the observed values fall well within the simulated data as shown by a graphical PCA representation, where the first 4 PC axes cumulatively explained 62.4% of total variation (Figure S4). Taken together, these model-checking results suggest this final selected scenario is a good representation of the invasion history of *A. artemisiifolia* into Europe and Australia.

Table S1. Overview of sites of *Ambrosia artemisiifolia* within native and introduced European and Australian ranges, including state or province (for native and Australian ranges) and country, location code (as used in text and figures), geographic coordinates (WGS84), year of sampling and sample sizes (n). Local population diversity indices are mean (± 95% confidence) number of individuals with called genotypes (N_G); allelic richness (A_R); gene diversity (H_S); observed heterozygosity (H_O, bold values are significantly different from H_S); and inbreeding coefficient (F_{IS}, bold values indicate significant ($\alpha = 0.05$) departure from zero). Missing values (–) indicate >90% of loci had <4 called individuals within the sampling location had and population statistics could not reliably be calculated.

Sampling location	Location code	Latitude	Longitude	Year	n	A_R	H_O	H_S	F_{IS}
Native range, North America, northwest unit (NW)									
ND, USA	AA-19	46.298	-103.918	2008	6	1.234 (0.016)	**0.184 (0.016)**	**0.241 (0.017)**	**0.173 (0.043)**
SD, USA	AA-17	44.764	-103.853	2008	6	1.228 (0.017)	**0.167 (0.016)**	**0.236 (0.018)**	**0.215 (0.045)**
ND, USA	AA-20	48.147	-103.572	2008	6	1.237 (0.014)	**0.179 (0.014)**	**0.245 (0.015)**	**0.203 (0.040)**
SD, USA	AA-16	44.095	-102.871	2008	4	1.242 (0.034)	**0.173 (0.033)**	**0.254 (0.037)**	**0.238 (0.098)**
SD, USA	AA-15	43.834	-101.271	2008	5	1.227 (0.018)	**0.169 (0.018)**	**0.235 (0.019)**	**0.202 (0.052)**
SD, USA	AA-14	43.715	-100.684	2008	6	1.252 (0.014)	**0.182 (0.013)**	**0.261 (0.015)**	**0.227 (0.037)**
NE, USA	AA-11	42.765	-98.063	2008	6	1.226 (0.017)	**0.173 (0.018)**	**0.234 (0.018)**	**0.194 (0.051)**
NE, USA	AA-12	42.763	-98.054	2008	4	1.265 (0.023)	**0.202 (0.024)**	**0.276 (0.024)**	**0.189 (0.064)**
MN, Canada	AA-1	49.688	-97.965	2008	6	1.221 (0.016)	**0.153 (0.015)**	**0.231 (0.017)**	**0.259 (0.045)**
SD, USA	AA-9	43.451	-97.420	2008	3	–	–	–	–
MN, Canada	AA-2	49.838	-97.329	2008	5	1.266 (0.022)	**0.190 (0.023)**	**0.278 (0.023)**	**0.232 (0.062)**
ND, USA	AA-3	48.195	-97.329	2008	6	1.254 (0.017)	**0.187 (0.017)**	**0.263 (0.018)**	**0.216 (0.046)**
MN, USA	AA-4	47.761	-96.593	2008	6	1.258 (0.015)	**0.185 (0.015)**	**0.267 (0.016)**	**0.236 (0.040)**
MN, USA	AA-5	46.217	-96.050	2008	6	1.245 (0.017)	**0.181 (0.017)**	**0.254 (0.018)**	**0.213 (0.045)**
MN, USA	AA-8	44.325	-95.958	2008	7	1.245 (0.013)	**0.178 (0.012)**	**0.253 (0.014)**	**0.218 (0.034)**
ND, USA	AA-6	45.021	-95.870	2008	6	1.235 (0.017)	**0.155 (0.015)**	**0.247 (0.018)**	**0.288 (0.047)**
MN, USA	AA-7	44.738	-95.412	2008	6	1.234 (0.016)	**0.164 (0.015)**	**0.244 (0.017)**	**0.246 (0.045)**
QC, Canada	TR-QB	45.500	-72.500	2008	6	1.244 (0.017)	**0.158 (0.016)**	**0.255 (0.018)**	**0.299 (0.049)**
Native range, North America, admixed unit (AD)									
NE, USA	AA-13	42.585	-99.061	2008	4	1.260 (0.021)	**0.206 (0.021)**	**0.269 (0.022)**	**0.160 (0.052)**
SD, USA	AA-10	43.057	-97.400	2008	3	–	–	–	–
KS, USA	KS	38.686	-96.493	2013	6	1.237 (0.016)	**0.191 (0.016)**	**0.242 (0.016)**	**0.153 (0.041)**
MO, USA	MO	37.006	-94.350	2013	6	1.270 (0.017)	**0.231 (0.019)**	**0.274 (0.017)**	**0.113 (0.041)**
WI, USA	WI	44.879	-89.424	2013	6	1.278 (0.015)	**0.231 (0.017)**	**0.285 (0.016)**	**0.134 (0.039)**
IN, USA	IND	39.175	-86.527	2008	6	1.323 (0.057)	0.405 (0.081)	0.311 (0.054)	-0.26 (0.095)
KY, USA	KY	38.626	-85.007	2013	6	1.269 (0.016)	**0.227 (0.018)**	**0.274 (0.016)**	**0.122 (0.039)**
MI, USA	MI	46.358	-84.881	2013	4	1.245 (0.026)	0.211 (0.028)	0.250 (0.027)	**0.107 (0.067)**
OH, USA	OH	40.488	-82.727	2013	6	1.271 (0.015)	**0.238 (0.017)**	**0.275 (0.015)**	**0.093 (0.036)**
ON, Canada	MN-ON	44.455	-79.809	2008	3	–	–	–	–

ON, Canada	MN-2	44.447	-79.804	2013	3	–	–	–	–
Native range, North America, southeast unit (SE)									
AR, USA	AR	33.976	-91.413	2013	6	1.272 (0.018)	0.268 (0.021)	0.273 (0.018)	-0.00 (0.039)
MS, USA	MP	31.208	-89.066	2013	6	1.277 (0.017)	**0.242 (0.018)**	**0.281 (0.017)**	**0.100 (0.038)**
AL, USA	AL	30.675	-87.591	2013	6	1.273 (0.018)	0.253 (0.021)	0.276 (0.018)	**0.054 (0.042)**
TN, USA	TN	36.268	-86.461	2013	6	1.261 (0.018)	0.236 (0.020)	0.264 (0.019)	**0.064 (0.040)**
FL, USA	FL	30.406	-83.140	2013	6	1.281 (0.022)	0.263 (0.026)	0.283 (0.022)	0.045 (0.051)
NC, USA	NC	35.607	-83.022	2013	6	1.243 (0.019)	0.265 (0.024)	0.240 (0.019)	-0.09 (0.040)
GA, USA	GA	31.634	-81.407	2013	6	1.280 (0.019)	0.269 (0.024)	0.281 (0.020)	0.021 (0.045)
SC, USA	SC	34.225	-81.343	2013	6	1.266 (0.016)	0.237 (0.018)	0.270 (0.016)	**0.081 (0.040)**
PA, USA	PA	40.966	-78.175	2013	6	1.281 (0.015)	0.261 (0.018)	0.283 (0.015)	**0.048 (0.035)**
ON, Canada	BN-ON	45.000	-77.700	2008	6	1.237 (0.015)	**0.164 (0.014)**	**0.247 (0.016)**	**0.263 (0.041)**
MS, USA	MS	42.088	-72.096	2013	7	1.233 (0.016)	0.221 (0.018)	0.235 (0.016)	0.029 (0.040)
MA, USA	MA	44.771	-68.971	2013	6	1.262 (0.018)	0.226 (0.020)	0.266 (0.019)	**0.108 (0.044)**
NB, Canada	NB1	45.879	-66.978	2013	6	1.266 (0.016)	0.246 (0.019)	0.269 (0.017)	**0.054 (0.038)**
Introduced range, Europe									
France	FR-4	46.180	-0.161	2008	2	–	–	–	–
France	FR-3	47.145	2.819	2008	6	1.244 (0.015)	**0.189 (0.015)**	**0.253 (0.015)**	**0.184 (0.040)**
France	FR-7	47.176	3.015	2008	6	1.256 (0.017)	**0.202 (0.017)**	**0.264 (0.018)**	**0.164 (0.041)**
France	FR-8	44.217	4.264	2008	6	1.264 (0.012)	**0.221 (0.013)**	**0.270 (0.013)**	**0.125 (0.030)**
France	FR-5	43.935	4.306	2008	3	–	–	–	–
France	FR-1	45.080	4.757	2008	7	1.257 (0.013)	**0.199 (0.013)**	**0.264 (0.014)**	**0.174 (0.033)**
France	FR-9	44.753	4.871	2008	3	–	–	–	–
France	FR-6	46.800	4.972	2008	5	1.254 (0.015)	**0.207 (0.016)**	**0.260 (0.015)**	**0.143 (0.039)**
France	FR-2	45.679	4.980	2008	5	1.262 (0.014)	**0.207 (0.015)**	**0.270 (0.015)**	**0.159 (0.037)**
France	FR-10	46.052	5.335	2008	6	1.265 (0.013)	**0.211 (0.014)**	**0.273 (0.014)**	**0.164 (0.035)**
the Netherlands	EU-17	51.120	5.840	2013	6	1.221 (0.022)	0.200 (0.025)	0.225 (0.022)	0.064 (0.069)
Switzerland	AACH-1	46.178	6.037	2008	2	–	–	–	–
Switzerland	160808-1	46.033	8.949	2008	6	1.230 (0.016)	**0.175 (0.015)**	**0.237 (0.017)**	**0.200 (0.042)**
Switzerland	160808-2	45.835	9.005	2008	6	1.246 (0.015)	**0.179 (0.014)**	**0.255 (0.015)**	**0.229 (0.040)**
Italy	160808-3	45.441	10.602	2008	6	1.245 (0.015)	**0.190 (0.015)**	**0.252 (0.015)**	**0.185 (0.040)**

Italy	160808-4	45.337	10.609	2008	6	1.219 (0.015)	**0.174 (0.015)**	**0.224 (0.016)**	**0.161 (0.042)**
Italy	170808-2	45.454	12.084	2008	6	1.251 (0.015)	**0.195 (0.015)**	**0.258 (0.015)**	**0.184 (0.039)**
Italy	170808-1	45.470	12.109	2008	6	1.202 (0.015)	**0.161 (0.015)**	**0.208 (0.016)**	**0.165 (0.045)**
Germany	EU-20	51.633	14.184	2014	5(6)	1.235 (0.016)	0.216 (0.018)	0.238 (0.017)	**0.052 (0.040)**
Slovenia	EU-01	46.036	15.296	2014	5(6)	1.252 (0.030)	0.197 (0.032)	0.261 (0.031)	**0.176 (0.083)**
Sweden	EU-19	56.168	15.890	2014	5	1.239 (0.033)	0.207 (0.036)	0.244 (0.034)	**0.107 (0.085)**
Czech Republic	EU-22	49.418	17.962	2014	6	1.220 (0.023)	0.202 (0.026)	0.223 (0.023)	0.057 (0.062)
Hungary	010908-1	47.642	18.783	2008	6	1.264 (0.015)	**0.203 (0.015)**	**0.272 (0.016)**	**0.184 (0.038)**
Hungary	HU-1	47.471	18.825	2008	2	–	–	–	–
Hungary	310808-3	47.208	19.105	2008	6	1.266 (0.013)	**0.199 (0.014)**	**0.275 (0.014)**	**0.207 (0.036)**
Hungary	290808-1	47.199	20.821	2008	6	1.248 (0.015)	**0.191 (0.015)**	**0.256 (0.015)**	**0.185 (0.038)**
Hungary	310808-2	47.483	20.884	2008	6	1.260 (0.014)	**0.203 (0.014)**	**0.268 (0.014)**	**0.180 (0.036)**
Slovakia	EU-24	48.489	21.806	2014	6	1.245 (0.015)	0.218 (0.017)	0.250 (0.016)	**0.076 (0.037)**
Serbia	EU-35	43.092	21.938	2014	6	1.261 (0.013)	**0.216 (0.014)**	**0.266 (0.014)**	**0.128 (0.032)**
Romania	280808-2	45.659	21.954	2008	6	1.254 (0.014)	**0.193 (0.014)**	**0.262 (0.015)**	**0.198 (0.037)**
Romania	280808-1	44.725	22.421	2008	7	1.234 (0.014)	0.212 (0.016)	0.237 (0.015)	**0.075 (0.036)**
Romania	270808-2	44.418	23.710	2008	6	1.229 (0.014)	**0.179 (0.014)**	**0.236 (0.015)**	**0.181 (0.038)**
Romania	EU-26	46.237	24.854	2014	7	1.251 (0.014)	**0.193 (0.014)**	**0.259 (0.015)**	**0.190 (0.037)**
Romania	260808-1	44.958	26.053	2008	7	1.240 (0.014)	**0.201 (0.015)**	**0.246 (0.015)**	**0.132 (0.037)**
Romania	250808-4	44.369	28.035	2008	6	1.270 (0.013)	**0.204 (0.013)**	**0.278 (0.014)**	**0.198 (0.033)**
Romania	250808-2	44.125	28.634	2008	6	1.257 (0.016)	**0.204 (0.016)**	**0.264 (0.016)**	**0.161 (0.039)**
Introduced range, Australia									
Australia	AU-01	-35.641	150.127	2014	6	1.237 (0.022)	0.190 (0.024)	0.245 (0.023)	**0.168 (0.064)**
Australia	AU-09	-28.869	151.167	2014	6	1.248 (0.017)	**0.200 (0.017)**	**0.255 (0.017)**	**0.153 (0.044)**
Australia	AU-23	-28.926	152.374	2014	6	1.254 (0.017)	**0.210 (0.017)**	**0.260 (0.018)**	**0.132 (0.039)**
Australia	AU-33	-31.442	152.465	2014	5(6)	1.220 (0.016)	0.210 (0.019)	0.221 (0.016)	0.032 (0.048)
Australia	AU-13	-26.391	152.794	2014	6	1.195 (0.018)	**0.158 (0.018)**	**0.201 (0.018)**	**0.148 (0.055)**
Australia	AU-02	-31.025	152.953	2014	6	1.227 (0.015)	0.200 (0.016)	0.231 (0.015)	**0.090 (0.037)**
Australia	AU-06	-28.641	153.551	2014	6	1.225 (0.015)	0.201 (0.016)	0.228 (0.016)	**0.081 (0.036)**

Table S2. Sampling locations excluded from datasets used in ABC-RF analyses testing introduction scenarios of *Ambrosia artemisiifolia* from the native range of North America to the introduced ranges of Europe and Australia. Location codes correspond to those in Table S1.

Analysis	**Description of dataset**	**Removed geographic sampling locations**		
		Native	*Europe*	*Australia*
Native - Europe	Full datasets	–	–	NA
	Remove populations with pairwise $F_{ST} > 0.1$	NC, TN, AA-19	160808-1, 160808-4, 170808-1, 270808-2, 280808-1, EU17, EU19, EU22	NA
	Remove populations with pairwise $F_{ST} > 0.1$ & geographic sampling locations with different STRUCTURE assignment	NC, TN, AA-19	160808-1, 160808-4, 170808-1, 270808-2, 280808-1, EU17, EU19, EU22, EU26, EU20, 260808-1, 280808-2, 270808-1	NA
Native - Australia	Full datasets	–	NA	
	Remove populations with pairwise $F_{ST} > 0.1$ (group 1)	NC, TN, AA-19	NA	AU6, AU13, AU33
	Remove populations with pairwise $F_{ST} > 0.1$ (group 2)	NC, TN, AA-19	NA	AU1, AU2, AU13, AU33
	Remove populations with pairwise $F_{ST} > 0.1$ & geographic sampling locations with different STRUCTURE assignment	NC, TN, AA-19	NA	AU1, AU6, AU9, AU13, AU33
Native, Europe & Australia	Full datasets	–	–	
	Remove populations with pairwise $F_{ST} > 0.1$	NC, TN, AA-19	160808-1, 160808-4, 170808-1, 270808-2, 280808-1, EU17, EU19, EU22	AU6, AU13, AU33
	Remove populations with pairwise $F_{ST} > 0.1$	NC, TN, AA-19	160808-1, 160808-4, 170808-1, 270808-2, 280808-1, EU17, EU19, EU22	AU1, AU2, AU13, AU33
	Remove populations with pairwise $F_{ST} > 0.1$ & geographic sampling locations with different STRUCTURE assignment	NC, TN, AA-19	160808-1, 160808-4, 170808-1, 270808-2, 280808-1, EU17, EU19, EU22, EU26, EU20, 260808-1, 280808-2, 270808-1, 260808-1	AU1, AU6, AU9, AU13, AU33

Table S3. Parameters, parameter priors and prior distribution used to simulate introduction scenarios of *Ambrosia artemisiifolia*. Parameters shown are: common to all ABC-RF analyses (a); specific to introduction from the native North America to Europe (b); Australia (c)(see Fig. 2, main text)); or bridgehead introduction from Europe to Australia (d, see Fig. 3, main text). We defined timing of events (in generations assuming one generation a year) as: post-introduction admixture (t_{post}, multiple introduction scenarios only); initial introduction (t_1); secondary introduction (t_2, multiple introduction scenarios only); pre-introduction admixture in the native range (t_{pre}); bridgehead introduction (t_{bh}, bridgehead scenarios only); and split of the ancestral population into the northwest and southeast native units (t_{anc}). *N* characterizes the stable effective population sizes where *i* is the: southeast native unit (*SE*); northwest native unit (*NW*); admixed native unit (*AD*); introduced unit (*IN*); ancestral population (*anc*); unsampled introduced U_1 (only present in scenarios where second introduction is not bottlenecked); or unsampled introduced U_2 (only present in scenarios where second introduction is bottlenecked). Nb_i represents the effective population of size of population *4* (single introduction scenarios only), U_1 or U_2 of duration db_i at primary (t_1) or secondary (t_2) introduction or of duration db_{bh} at bridgehead introduction (t_{bh}, bridgehead scenarios only). The genetic contribution (admixture rates) *r* of population *i* are defined as: r_{pre} of population *SE* to *AD* at pre-introduction t_{pre}; and r_{post} of *i* to *IN* post-introduction at t_{post}. For the latter, r_{post} is defined for the lowest numerical *i* admixing (e.g. *i=2* when *2 x 3*, *i=3* when $3xU_2$). All parameters were defined and sampled within and between scenarios independently, except for conditions restricting events to follow in a chronological order. Prior sampling followed a uniform (UN) or log-uniform (LU) distribution.

	Parameter	Prior
a)	N_i	UN[5000,20000]
	Nb_i	UN[150,1000]
	t_{pre}	LU[100,500]
	$r_{post,pre}$	UN[0.1,0.9]
	t_{anc}	UN[500, 2000]
	N_{anc}	UN[1000,3000]
b)	t_{post}	UN[10,180]
	db_i	UN[0,60]
	t_2	UN[60,180]
	t_1	UN[100,180]
c)	t_{post}	UN[10,120]
	db_i	UN(0,50]
	t_2	UN[50,120]
	t_1	UN[100,120]
d)	t_{bh}	UN[50,120]
	db_{bh}	UN(0,50]
	t_{post}	UN[10,180]
	t_1	UN[100,180]
	db_i	UN(0,60]

Table S4. Summary statistics used in all simulations, as per DIYABC (Cornuet *et al.*, 2014).

DIYABC abbreviation	Description
Single sample statistics for each sampled population	
HP0	Proportion of loci with zero gene diversity
HM1	Mean gene diversity across polymorphic loci (Nei, 1987)
HV1	Variance of gene diversity across polymorphic loci
HMO	Mean gene diversity across all loci
Two sample statistics for each pairwise sample combination	
FP0	Proportion of loci with zero F_{ST} distance (Weir & Cockerham, 1984)
FM1	Mean across loci of non-zero F_{ST} distances
FV1	Variance across loci of non-zero F_{ST} distances
FMO	Mean across loci of F_{ST} distances
NP0	Proportion of loci with zero Nei's distance (Nei, 1972)
NM1	Mean across loci of non-zero Nei's distances
NV1	Variance across loci of non-zero Nei's distances
NMO	Mean across loci of Nei's distances
Admixture statistics (Maximum likelihoods (Choisy et al., 2004)) for each combination of parental and admixed populations	
AP0	Proportion of loci with zero admixture estimates
AM1	Mean across loci of non-zero admixture estimate
AV1	Variance across loci of non-zero admixture estimated
AMO	Mean across all locus admixture estimates

Table S5. Weir & Cockerham (1984) pairwise FST for geographic sampling sites included in this study within genetic units defined for ABC. Location codes correspond to Table S1. High F_{ST}= green and low F_{ST} = red.

Native, northwest (NW)

	AL	AR	BN-ON	FL	GA	MA	MP	MS	NB1	NC	PA	SC	TN
AL		0.05	0.06	0.04	0.05	0.06	0.03	0.08	0.06	0.09	0.02	0.00	0.07
AR	0.05		0.06	0.06	0.07	0.07	0.03	0.08	0.06	0.12	0.03	0.04	0.06
BN-ON	0.06	0.06		0.08	0.06	0.05	0.04	0.08	0.08	0.11	0.07	0.06	0.08
FL	0.04	0.06	0.08		0.04	0.07	0.03	0.09	0.05	0.12	0.06	0.04	0.12
GA	0.05	0.07	0.06	0.04		0.05	0.02	0.09	0.07	0.14	0.06	0.05	0.11
MA	0.06	0.07	0.05	0.07	0.05		0.05	0.07	0.06	0.11	0.06	0.03	0.10
MP	0.03	0.03	0.04	0.03	0.02	0.05		0.05	0.04	0.10	0.04	0.01	0.06
MS	0.08	0.08	0.08	0.09	0.09	0.07	0.05		0.08	0.12	0.06	0.05	0.10
NB1	0.06	0.06	0.08	0.05	0.07	0.06	0.04	0.08		0.12	0.03	0.04	0.10
NC	0.09	0.12	0.11	0.12	0.14	0.11	0.10	0.12	0.12		0.07	0.10	0.13
PA	0.02	0.03	0.07	0.06	0.06	0.06	0.04	0.06	0.03	0.07		0.04	0.07
SC	0.00	0.04	0.06	0.04	0.05	0.03	0.01	0.05	0.04	0.10	0.04		0.07
TN	0.07	0.06	0.08	0.12	0.11	0.10	0.06	0.10	0.10	0.13	0.07	0.07	

Native, southeast (SE)

	AA-1	AA-11	AA-12	AA-14	AA-15	AA-16	AA-17	AA-19	AA-2	AA-20	AA-3	AA-4	AA-5
AA-1		0.05	0.06	0.05	0.06	0.07	0.05	0.13	0.08	0.06	0.07	0.05	0.05
AA-11	0.05		0.04	0.02	0.03	0.02	0.02	0.08	0.03	0.05	0.02	0.02	0.03
AA-12	0.06	0.04		0.03	0.02	0.02	0.01	0.07	0.02	0.03	0.02	0.03	0.02
AA-14	0.05	0.02	0.03		0.02	0.01	0.01	0.07	0.01	0.03	0.02	0.01	0.02
AA-15	0.06	0.03	0.02	0.02		0.02	0.00	0.07	0.02	0.04	0.02	0.02	0.03
AA-16	0.07	0.02	0.02	0.01	0.02		0.00	0.06	0.03	0.04	0.03	0.02	0.03
AA-17	0.05	0.02	0.01	0.01	0.00	0.00		0.06	0.01	0.02	0.01	0.00	0.02
AA-19	0.13	0.08	0.07	0.07	0.07	0.06	0.06		0.06	0.09	0.06	0.06	0.08
AA-2	0.08	0.03	0.02	0.01	0.02	0.03	0.01	0.06		0.03	0.02	0.01	0.03
AA-20	0.06	0.05	0.03	0.03	0.04	0.04	0.02	0.09	0.03		0.04	0.05	0.04
AA-3	0.07	0.02	0.02	0.02	0.02	0.03	0.01	0.06	0.02	0.04		0.01	0.02
AA-4	0.05	0.02	0.03	0.01	0.02	0.02	0.00	0.06	0.01	0.05	0.01		0.02
AA-5	0.05	0.03	0.02	0.02	0.03	0.03	0.02	0.08	0.03	0.04	0.02	0.02	

AA-6	0.06	0.02	0.04	0.02	0.02	0.02	0.02	0.08	0.04	0.05	0.03	0.02	0.03
AA-7	0.06	0.03	0.02	0.02	0.01	0.03	0.01	0.06	0.01	0.06	0.03	0.01	0.02
AA-8	0.07	0.02	0.02	0.01	0.02	0.02	0.01	0.07	0.02	0.04	0.01	0.02	0.03
AA-9	0.06	0.01	0.01	0.02	0.00	-0.01	0.00	0.07	0.02	0.05	0.01	0.01	0.02
TR-QB	0.05	0.04	0.02	0.02	0.03	0.02	0.02	0.08	0.04	0.05	0.03	0.02	0.03

Native, admixed (AD)

	AA-10	AA-13	IND	KS	KY	MI	MN-2	MN-ON	MO	OH	WI
AA-10		0.02	-0.01	0.04	0.01	0.04	0.01	0.04	0.04	0.04	0.04
AA-13	0.02		0.00	0.06	0.03	0.06	0.01	0.03	0.03	0.06	0.03
IND	-0.01	0.00		0.04	0.01	0.07	0.00	0.02	0.01	0.02	-0.02
KS	0.04	0.06	0.04		0.04	0.06	0.06	0.05	0.03	0.07	0.06
KY	0.01	0.03	0.01	0.04		0.04	0.00	0.03	0.03	0.03	0.01
MI	0.04	0.06	0.07	0.06	0.04		0.05	0.05	0.07	0.06	0.03
MN-2	0.01	0.01	0.00	0.06	0.00	0.05		0.00	0.02	0.02	0.03
MN-ON	0.04	0.03	0.02	0.05	0.03	0.05	0.00		0.03	0.04	0.04
MO	0.04	0.03	0.01	0.03	0.03	0.07	0.02	0.03		0.03	0.05
OH	0.04	0.06	0.02	0.07	0.03	0.06	0.02	0.04	0.03		0.06
WI	0.04	0.03	-0.02	0.06	0.01	0.03	0.03	0.04	0.05	0.06	

Table S5. *Continued*

Introduced, Europe

	010908-1	160808-1	160808-2	160808-3	160808-4	170808-1	170808-2	250808-2	250808-4	260808-1	270808-2	280808-1	280808-2	290808-1	310808-2	310808-3	AACH-1	EU-01	EU-17	EU-19	EU-20	EU-22	EU-24	EU-26
010908-1		0.07	0.03	0.03	0.10	0.13	0.05	0.02	0.01	0.04	0.07	0.07	0.01	0.01	0.00	0.01	0.01	0.02	0.11	0.05	0.05	0.07	0.02	0.04
160808-1	0.07		0.09	0.08	0.17	0.21	0.11	0.07	0.08	0.11	0.13	0.14	0.08	0.08	0.06	0.06	0.07	0.11	0.16	0.09	0.11	0.13	0.07	0.11
160808-2	0.03	0.09		0.05	0.13	0.15	0.08	0.02	0.03	0.06	0.09	0.10	0.03	0.04	0.03	0.02	0.03	0.07	0.12	0.05	0.07	0.09	0.04	0.07
160808-3	0.03	0.08	0.05		0.07	0.14	0.07	0.03	0.03	0.06	0.08	0.08	0.03	0.04	0.03	0.03	0.04	0.05	0.10	0.04	0.06	0.11	0.03	0.05
160808-4	0.10	0.17	0.13	0.07		0.22	0.13	0.10	0.08	0.14	0.14	0.16	0.09	0.11	0.10	0.11	0.13	0.14	0.19	0.16	0.13	0.20	0.11	0.11
170808-1	0.13	0.21	0.15	0.14	0.22		0.15	0.13	0.14	0.15	0.18	0.19	0.13	0.12	0.13	0.13	0.16	0.17	0.23	0.18	0.15	0.22	0.15	0.16
170808-2	0.05	0.11	0.08	0.07	0.13	0.15		0.07	0.05	0.07	0.11	0.12	0.06	0.07	0.06	0.06	0.07	0.06	0.15	0.09	0.08	0.13	0.04	0.07
250808-2	0.02	0.07	0.02	0.03	0.10	0.13	0.07		0.01	0.04	0.07	0.07	0.01	0.02	0.02	0.02	-0.01	0.04	0.09	0.04	0.05	0.06	0.03	0.04
250808-4	0.01	0.08	0.03	0.03	0.08	0.14	0.05	0.01		0.04	0.06	0.06	0.01	0.02	0.02	0.03	0.00	0.03	0.08	0.02	0.04	0.08	0.00	0.04
260808-1	0.04	0.11	0.06	0.06	0.14	0.15	0.07	0.04	0.04		0.09	0.10	0.05	0.04	0.04	0.05	0.04	0.07	0.14	0.10	0.07	0.11	0.03	0.04
270808-2	0.07	0.13	0.09	0.08	0.14	0.18	0.11	0.07	0.06	0.09		0.04	0.03	0.08	0.06	0.06	0.08	0.09	0.17	0.11	0.11	0.14	0.07	0.09
280808-1	0.07	0.14	0.10	0.08	0.16	0.19	0.12	0.07	0.06	0.10	0.04		0.02	0.08	0.08	0.07	0.07	0.11	0.14	0.13	0.11	0.14	0.08	0.09
280808-2	0.01	0.08	0.03	0.03	0.09	0.13	0.06	0.01	0.01	0.05	0.03	0.02		0.03	0.02	0.00	0.01	0.02	0.09	0.05	0.05	0.06	0.02	0.03
290808-1	0.01	0.08	0.04	0.04	0.11	0.12	0.07	0.02	0.02	0.04	0.08	0.08	0.03		0.02	0.02	0.00	0.03	0.11	0.05	0.06	0.08	0.02	0.04
310808-2	0.00	0.06	0.03	0.03	0.10	0.13	0.06	0.02	0.02	0.04	0.06	0.08	0.02	0.02		0.01	-0.01	0.04	0.10	0.03	0.06	0.07	0.01	0.04
310808-3	0.01	0.06	0.02	0.03	0.11	0.13	0.06	0.02	0.03	0.05	0.06	0.07	0.00	0.02	0.01		-0.01	0.02	0.11	0.04	0.05	0.07	0.01	0.05
AACH-1	0.01	0.07	0.03	0.04	0.13	0.16	0.07	-0.01	0.00	0.04	0.08	0.07	0.01	0.00	-0.01	-0.01		0.06	0.12	0.09	0.04	0.09	0.02	0.04
EU-01	0.02	0.11	0.07	0.05	0.14	0.17	0.06	0.04	0.03	0.07	0.09	0.11	0.02	0.03	0.04	0.02	0.06		0.14	0.11	0.09	0.12	0.05	0.06
EU-17	0.11	0.16	0.12	0.10	0.19	0.23	0.15	0.09	0.08	0.14	0.17	0.14	0.09	0.11	0.10	0.11	0.12	0.14		0.16	0.14	0.20	0.11	0.12
EU-19	0.05	0.09	0.05	0.04	0.16	0.18	0.09	0.04	0.02	0.10	0.11	0.13	0.05	0.05	0.03	0.04	0.09	0.11	0.16		0.10	0.17	0.05	0.08
EU-20	0.05	0.11	0.07	0.06	0.13	0.15	0.08	0.05	0.04	0.07	0.11	0.11	0.05	0.06	0.06	0.05	0.04	0.09	0.14	0.10		0.11	0.06	0.07
EU-22	0.07	0.13	0.09	0.11	0.20	0.22	0.13	0.06	0.08	0.11	0.14	0.14	0.06	0.08	0.07	0.07	0.09	0.12	0.20	0.17	0.11		0.11	0.10
EU-24	0.02	0.07	0.04	0.03	0.11	0.15	0.04	0.03	0.00	0.03	0.07	0.08	0.02	0.02	0.01	0.01	0.02	0.05	0.11	0.05	0.06	0.11		0.05
EU-26	0.04	0.11	0.07	0.05	0.11	0.16	0.07	0.04	0.04	0.04	0.09	0.09	0.03	0.04	0.04	0.05	0.04	0.06	0.12	0.08	0.07	0.10	0.05	
EU-35	0.02	0.09	0.06	0.04	0.11	0.13	0.07	0.01	0.02	0.06	0.07	0.06	0.01	0.02	0.01	0.02	0.01	0.06	0.11	0.04	0.06	0.09	0.02	0.05
FR-1	0.01	0.07	0.03	0.04	0.10	0.12	0.05	0.04	0.02	0.04	0.06	0.08	0.03	0.01	0.01	0.01	-0.01	0.03	0.12	0.02	0.04	0.09	0.02	0.04
FR-10	0.02	0.08	0.07	0.05	0.11	0.12	0.06	0.03	0.03	0.05	0.08	0.09	0.03	0.03	0.02	0.05	0.00	0.05	0.10	0.05	0.05	0.09	0.03	0.05

	010908-1	160808-1	160808-2	160808-3	160808-4	170808-1	170808-2	250808-2	250808-4	260808-1	270808-2	280808-1	280808-2	290808-1	310808-2	310808-3	AACH-1	EU-01	EU-17	EU-19	EU-20	EU-22	EU-24	EU-26
FR-2	0.00	0.09	0.03	0.02	0.10	0.12	0.06	0.01	0.00	0.04	0.07	0.07	0.03	0.00	0.03	0.02	0.00	0.05	0.10	0.03	0.05	0.07	0.01	0.03
FR-3	0.03	0.10	0.04	0.04	0.11	0.13	0.06	0.04	0.03	0.07	0.07	0.09	0.03	0.03	0.03	0.03	0.03	0.06	0.14	0.05	0.06	0.11	0.04	0.06
FR-4	0.01	0.08	0.05	0.04	0.12	0.16	0.07	0.01	0.01	0.07	0.08	0.08	0.00	0.02	0.03	0.02	0.03	0.07	0.13	0.05	0.06	0.08	0.04	0.05
FR-5	0.03	0.10	0.05	0.06	0.13	0.15	0.09	0.01	0.04	0.05	0.09	0.09	0.03	0.02	0.02	0.04	0.01	0.07	0.11	0.06	0.06	0.09	0.04	0.05
FR-6	0.01	0.07	0.04	0.03	0.11	0.14	0.07	0.01	0.01	0.06	0.08	0.07	0.01	0.01	0.02	0.02	0.01	0.04	0.11	0.04	0.06	0.08	0.02	0.04
FR-7	0.00	0.06	0.03	0.02	0.09	0.13	0.06	0.01	0.01	0.03	0.07	0.06	0.00	0.01	0.00	0.00	-0.01	0.03	0.10	0.03	0.02	0.08	0.01	0.04
FR-8	0.01	0.07	0.04	0.03	0.09	0.13	0.06	0.04	0.01	0.04	0.06	0.07	0.01	0.02	0.02	0.01	0.02	0.04	0.10	0.04	0.05	0.08	0.02	0.03
FR-9	0.01	0.07	0.03	0.01	0.11	0.12	0.04	0.00	0.00	0.02	0.06	0.07	0.00	0.00	0.01	0.01	-0.01	0.02	0.09	0.04	0.05	0.08	0.01	0.01
HU-1	0.03	0.11	0.04	0.05	0.14	0.16	0.06	0.01	-0.02	0.05	0.08	0.11	0.01	0.05	0.03	0.03	0.04	0.09	0.15	0.08	0.04	0.10	0.04	0.05

Table S5. *Continued*

Introduced, Europe. ***Continued***

	EU-35	FR-1	FR-10	FR-2	FR-3	FR-4	FR-5	FR-6	FR-7	FR-8	FR-9	HU-1
010908-1	0.02	0.01	0.02	0.00	0.03	0.01	0.03	0.01	0.00	0.01	0.01	0.03
160808-1	0.09	0.07	0.08	0.09	0.10	0.08	0.10	0.07	0.06	0.07	0.07	0.11
160808-2	0.06	0.03	0.07	0.03	0.04	0.05	0.05	0.04	0.03	0.04	0.03	0.04
160808-3	0.04	0.04	0.05	0.02	0.04	0.04	0.06	0.03	0.02	0.03	0.01	0.05
160808-4	0.11	0.10	0.11	0.10	0.11	0.12	0.13	0.11	0.09	0.09	0.11	0.14
170808-1	0.13	0.12	0.12	0.12	0.13	0.16	0.15	0.14	0.13	0.13	0.12	0.16
170808-2	0.07	0.05	0.06	0.06	0.06	0.07	0.09	0.07	0.06	0.06	0.04	0.06
250808-2	0.01	0.04	0.03	0.01	0.04	0.01	0.01	0.01	0.01	0.04	0.00	0.01
250808-4	0.02	0.02	0.03	0.00	0.03	0.01	0.04	0.01	0.01	0.01	0.00	-0.02
260808-1	0.06	0.04	0.05	0.04	0.07	0.07	0.05	0.06	0.03	0.04	0.02	0.05
270808-2	0.07	0.06	0.08	0.07	0.07	0.08	0.09	0.08	0.07	0.06	0.06	0.08
280808-1	0.06	0.08	0.09	0.07	0.09	0.08	0.09	0.07	0.06	0.07	0.07	0.11
280808-2	0.01	0.03	0.03	0.03	0.03	0.00	0.03	0.01	0.00	0.01	0.00	0.01
290808-1	0.02	0.01	0.03	0.00	0.03	0.02	0.02	0.01	0.01	0.02	0.00	0.05
310808-2	0.01	0.01	0.02	0.03	0.03	0.03	0.02	0.02	0.00	0.02	0.01	0.03
310808-3	0.02	0.01	0.05	0.02	0.03	0.02	0.04	0.02	0.00	0.01	0.01	0.03
AACH-1	0.01	-0.01	0.00	0.00	0.03	0.03	0.01	0.01	-0.01	0.02	-0.01	0.04
EU-01	0.06	0.03	0.05	0.05	0.06	0.07	0.07	0.04	0.03	0.04	0.02	0.09
EU-17	0.11	0.12	0.10	0.10	0.14	0.13	0.11	0.11	0.10	0.10	0.09	0.15
EU-19	0.04	0.02	0.05	0.03	0.05	0.05	0.06	0.04	0.03	0.04	0.04	0.08
EU-20	0.06	0.04	0.05	0.05	0.06	0.06	0.06	0.06	0.02	0.05	0.05	0.04
EU-22	0.09	0.09	0.09	0.07	0.11	0.08	0.09	0.08	0.08	0.08	0.08	0.10
EU-24	0.02	0.02	0.03	0.01	0.04	0.04	0.04	0.02	0.01	0.02	0.01	0.04
EU-26	0.05	0.04	0.05	0.03	0.06	0.05	0.05	0.04	0.04	0.03	0.01	0.05
EU-35		0.01	0.03	0.01	0.04	0.02	0.03	0.01	0.01	0.02	0.02	0.05
FR-1	0.01		0.03	0.01	0.01	0.01	0.02	0.01	0.01	0.01	-0.01	0.01
FR-10	0.03	0.03		0.01	0.05	0.03	0.04	0.03	0.02	0.03	0.01	0.01
FR-2	0.01	0.01	0.01		0.01	0.01	0.02	0.01	0.00	0.01	-0.01	0.02

	EU-35	FR-1	FR-10	FR-2	FR-3	FR-4	FR-5	FR-6	FR-7	FR-8	FR-9	HU-1
FR-3	0.04	0.01	0.05	0.01		0.04	0.05	0.02	0.02	0.03	0.03	0.03
FR-4	0.02	0.01	0.03	0.01	0.04		0.00	0.02	-0.02	0.03	0.00	0.05
FR-5	0.03	0.02	0.04	0.02	0.05	0.00		0.03	0.03	0.01	0.01	-0.01
FR-6	0.01	0.01	0.03	0.01	0.02	0.02	0.03		0.00	0.01	0.02	0.03
FR-7	0.01	0.01	0.02	0.00	0.02	-0.02	0.03	0.00		0.01	-0.01	0.03
FR-8	0.02	0.01	0.03	0.01	0.03	0.03	0.01	0.01	0.01		-0.03	0.03
FR-9	0.02	-0.01	0.01	-0.01	0.03	0.00	0.01	0.02	-0.01	-0.03		-0.02
HU-1	0.05	0.01	0.01	0.02	0.03	0.05	-0.01	0.03	0.03	0.03	-0.02	

Table S5. *Continued*

Introduced, Australia

	AU-01	AU-02	AU-06	AU-09	AU-13	AU-23	AU-33
AU-01		0.06	0.11	0.07	0.16	0.06	0.10
AU-02	0.06		0.11	0.05	0.16	0.06	0.12
AU-06	0.11	0.11		0.10	0.18	0.10	0.16
AU-09	0.07	0.05	0.10		0.18	0.07	0.13
AU-13	0.16	0.16	0.18	0.18		0.14	0.22
AU-23	0.06	0.06	0.10	0.07	0.14		0.12
AU-33	0.10	0.12	0.16	0.13	0.22	0.12	

Table S6. ABC-RF votes per scenario (/1000 decision trees) for introduction scenarios of *Ambrosia artemisiifolia* from divergent native sources of North America during initial (rows) and secondary (columns) into Europe (a) Australia (b)(Fig. 2) and Australian introduction from native AD and bridgehead EU (c, Fig. 3) within on multiple pseudo-datasets (Table S3). For c) scenarios are based on the European and Australian introduction scenario receiving most votes (a&b). The RF posterior probability corresponds to the scenario with highest number of votes. Event parameters descriptions and priors are in Table S5.

a) European introduction from native genetic units

i) Full dataset (result reported in main text)

			Second introduction		
	Native source	*Single introduction*	*Southeast (SE)*	*Northwest (NW)*	*Admixed (AD)*
Founding introduction	*Southeast (SE)*	86	112	178	28
	Northwest (NW)	98	234	93	54
	Admixed (AD)	8	42	57	10

Posterior probability: 0.501; prior error: 44.01%

*ii) P*airwise F_{ST}<0.1 within the same genetic unit

Founding introduction	*Southeast (SE)*	127	115	161	46
	Northwest (NW)	76	164	94	80
	Admixed (AD)	15	41	68	13

Posterior probability: 0.499; prior error: 44.32%

*iii) P*airwise F_{ST}<0.1 within the same unit & similar STRUCTURE assignment

Founding introduction	*Southeast (SE)*	135	118	150	58
	Northwest (NW)	109	148	101	45
	Admixed (AD)	13	64	54	5

Posterior probability: 0.516; prior error: 44.85%

Table S6. *Continued*

b) Australian introduction from native genetic units

i) Full dataset (result reported in main text)

			Second introduction		
	Native source	*Single introduction*	*Southeast (SE)*	*Northwest (NW)*	*Admixed (AD)*
Founding introduction	*Southeast (SE)*	16	4	30	68
	Northwest (NW)	27	11	9	59
	Admixed (AD)	442	60	58	216

Posterior probability: 0.599; prior error: 47.30%

ii) Pairwise F_{ST}<0.1 within the same unit, AU group 1

Founding introduction	*Southeast (SE)*	9	6	29	87
	Northwest (NW)	28	24	8	62
	Admixed (AD)	354	88	71	234

Posterior probability: 0.500; prior error: 48.44%

iii) Pairwise F_{ST}<0.1 within the same unit, AU group 2

Founding introduction	*Southeast (SE)*	34	8	19	95
	Northwest (NW)	55	24	6	69
	Admixed (AD)	344	171	60	115

Posterior probability: 0.473; prior error: 49.12%

iv) Pairwise F_{ST} <0.1 within the same unit & similar STRUCTURE assignment

Founding introduction	*Southeast (SE)*	47	21	38	95
	Northwest (NW)	46	29	5	44
	Admixed (AD)	343	133	68	131

Posterior probability: 0.490; prior error: 50.78%

Table S6. *Continued*

c) Australian introduction from native AD and bridgehead EU

i) Full dataset (result reported in main text)

Native	**Bridgehead**
Admixed (AD)	*Europe (introduced)*
74	926

Posterior probability: 0.998; prior error: 0.10%

ii) Pairwise F_{ST}<0.1 within the same unit, AU group 1

72	928

Posterior probability: 0.999; prior error: 0.16%

iii) Pairwise F_{ST} <0.1 within the same unit, AU group 2

132	868

Posterior probability: 0.999; prior error: 0.14%

iv) Pairwise F_{ST} <0.1 within the same unit & similar STRUCTURE assignment

Admixed (AD)	*Europe (introduced)*
94	906

Posterior probability: 0.999; prior error: 0.15%

Table S7. Posterior parameter median and quantile distribution (2.5-97.5) for *Ambrosia artemisiifolia* Australian introduction through European bridgehead (Fig. 3 main text), based on 'best' introduction scenarios in preceding analyses (Table 3, main text). Parameters and parameter prior distributions are described in Table S4.

Parameter	prior	median	q2.5	q97.5
N_1	*5000-20,000*	15,693	8,483	19,801
N_2	*5000-20,000*	14,257	7,041	19,691
N_3	*5000-20,000*	13,789	6,107	19,682
N_4	*5000-20,000*	11,220	5,143	19,657
N_5	*5000-20,000*	12,075	5,202	19,667
N_6	*5000-20,000*	12,381	5,175	19,637
N_7	*5000-20,000*	12,605	5,482	19,692
t_{bh}	*50-120*	78	51	115
d_{bh}	*0-50*	34	15	50
Nb_5	*150-1,000*	332	158	564
t_{post}	*10-180*	101	59	151
r_{post}	*0.1-0.9*	0.24	0.11	0.48
t_2	*60-180*	126	81	174
db_2	*0-60*	34	3	59
Nb_6	*150-1,000*	534	167	974
t_1	*100-180*	147	106	179
db_1	*0-60*	36	3	59
Nb_7	*150-1,000*	567	160	962
t_{pre}	*100-500*	218	124	385
r_{pre}	*0.1-0.9*	0.74	0.50	0.89
t_{anc}	*500-2,000*	682	506	1,027
N_{anc}	*1,000-3,000*	1,673	1,019	2,838

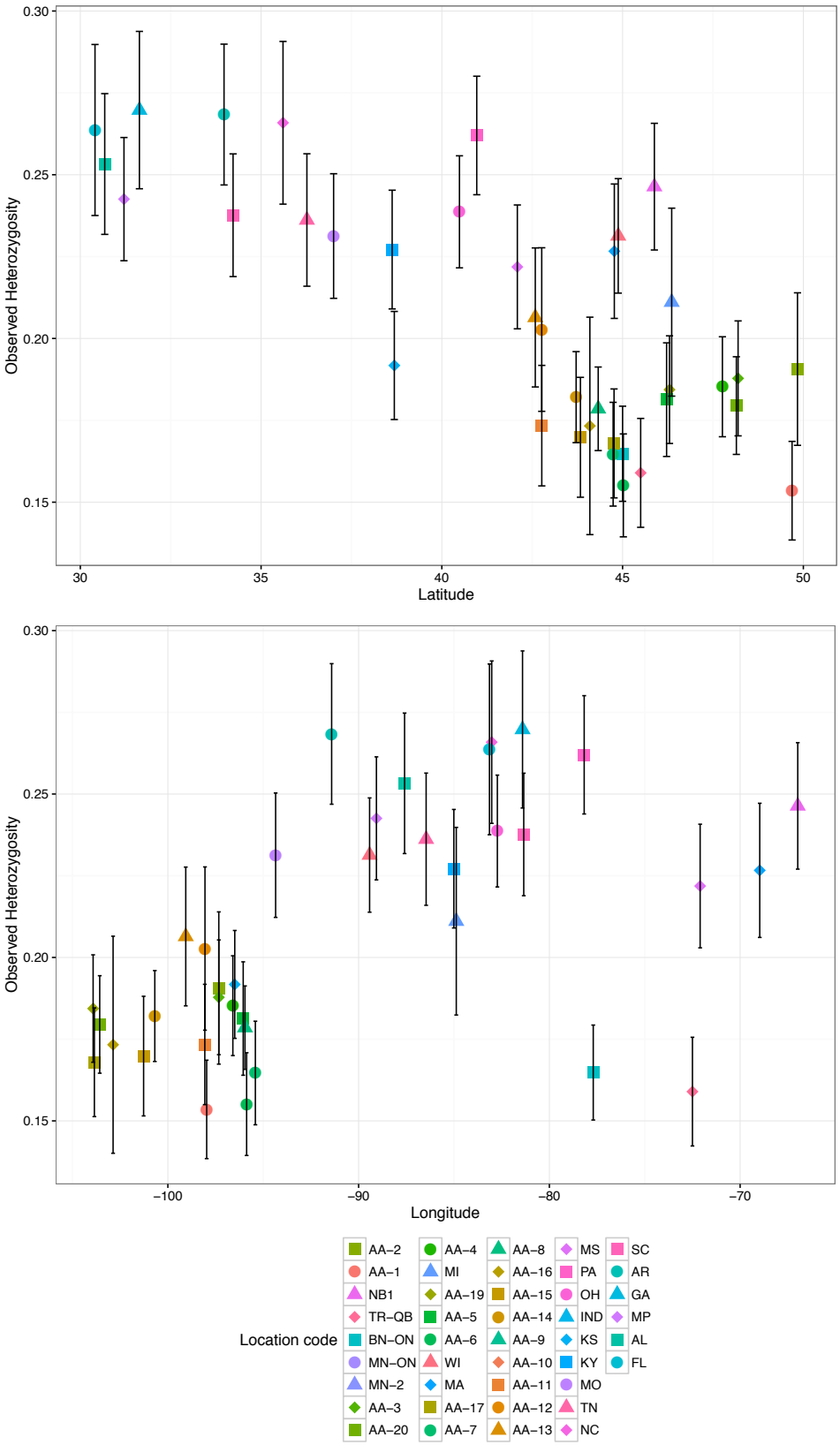


Figure S1a. Observed heterozygosity versus latitude (top figure) and longitude (bottom figure) in the native range for each sampling location (excluding sampling locations with >90% of loci with <4 called genotypes), with 95% confidence intervals, calculated across loci. Points are colored and shaped according to location code (latitude Spearman's rho=0.673, p<0.001; longitude Spearman's rho=0.525, p<0.001).

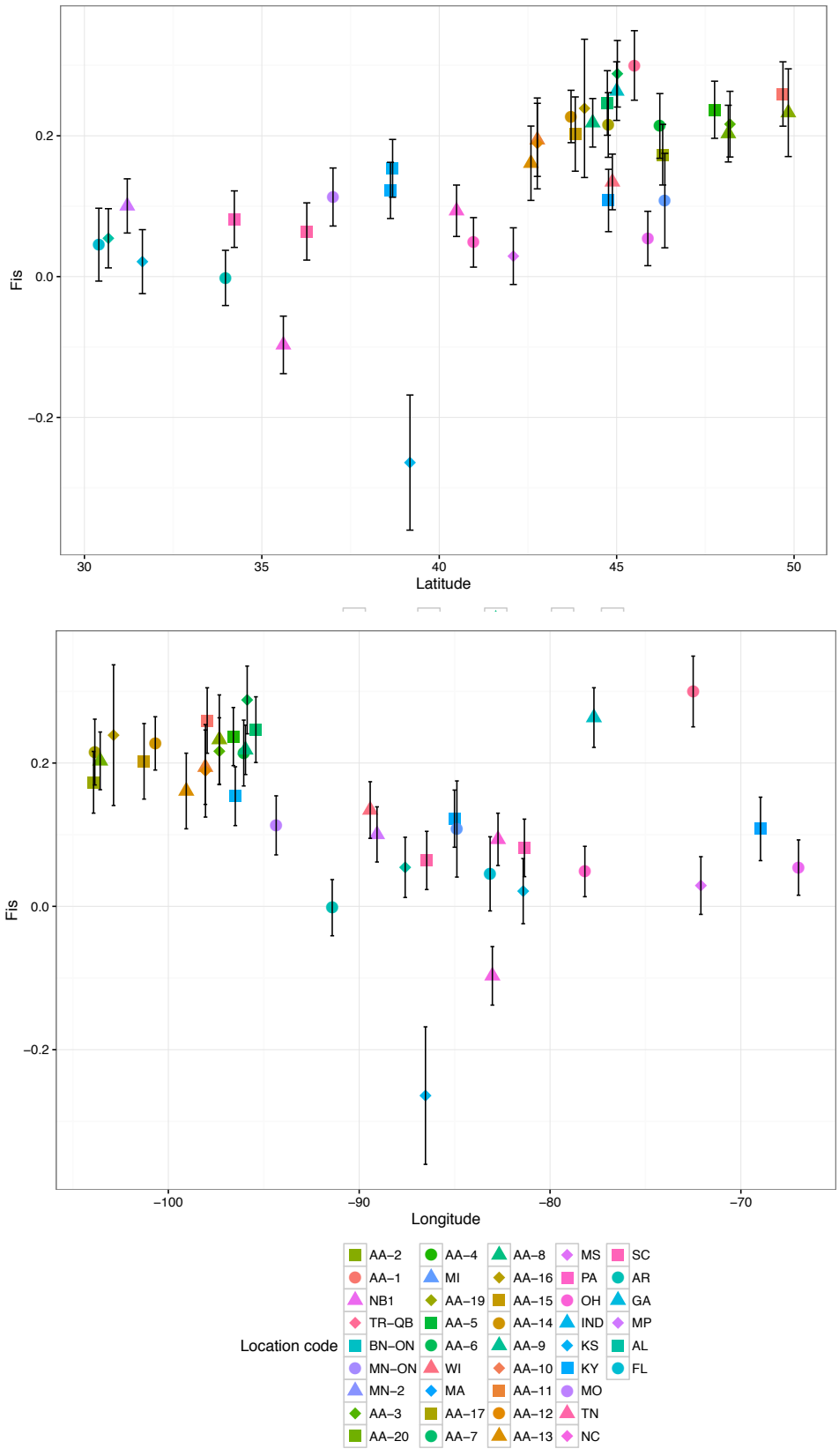


Figure S1b. Inbreeding coefficient (F_{IS}) versus latitude and longitude in the native range for each sampling location (excluding sampling locations for which population statistics could not reliably be calculated), with 95% confidence intervals, calculated across loci. Points are coloured and shaped according to location code (latitude Spearman's rho=0.693, p<0.001; longitude Spearman's rho=0.509, p=0.001).

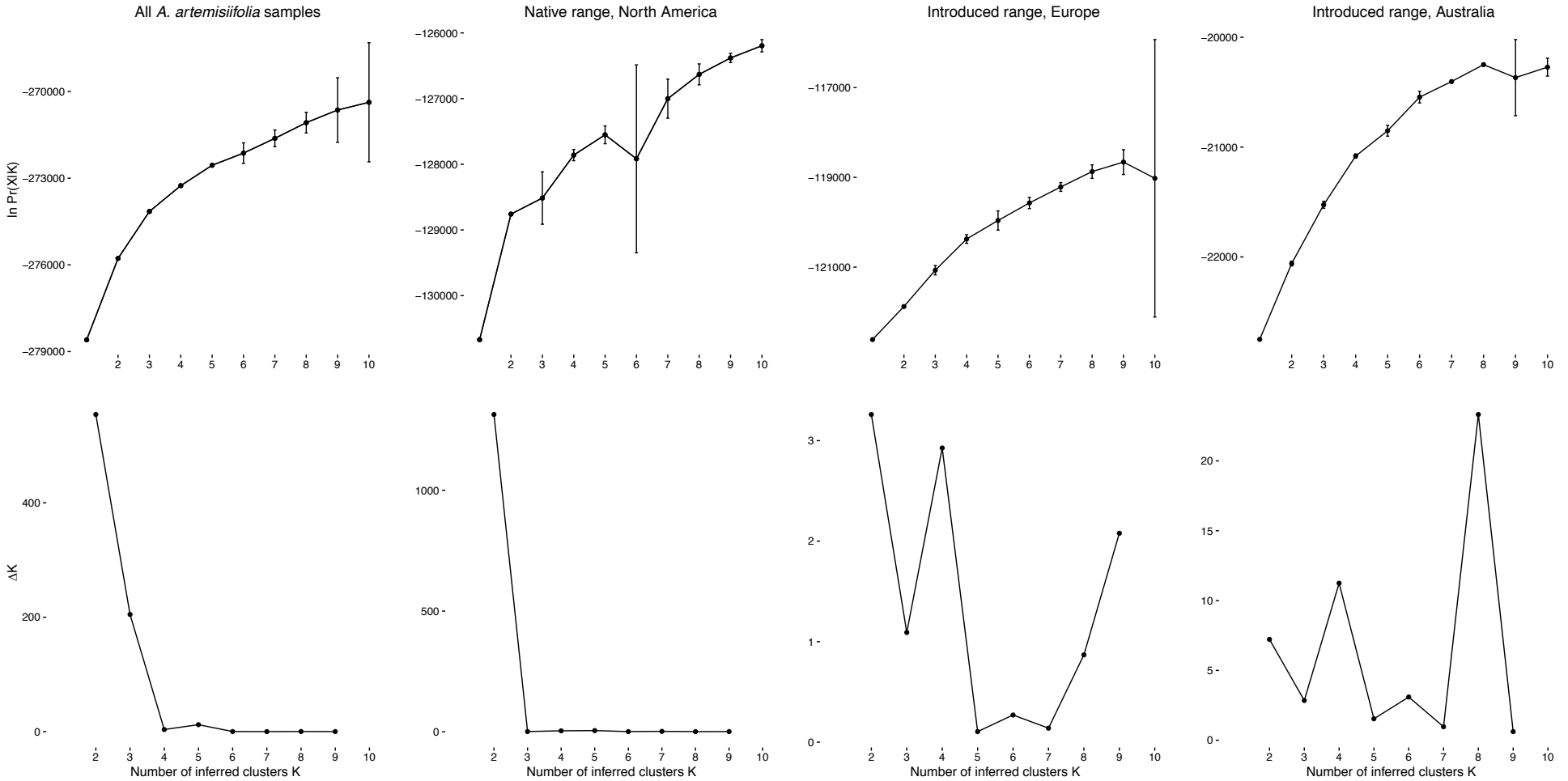


Figure S2a. STRUCTURE analyses of 20 replicate runs on 1022 randomly selected unlinked SNPs within 466 *A. artemisiifolia* samples as well as on subsets of the data (Native range of North America, and introduced ranges of Europe and Australia). with post-hoc evaluation (a, Evanno *et al.*, 2005); top row shows log probability of the data for every inferred number of units K according, bottom row shows the delta K value (i.e. the second order rate of change in this log probability for every K).

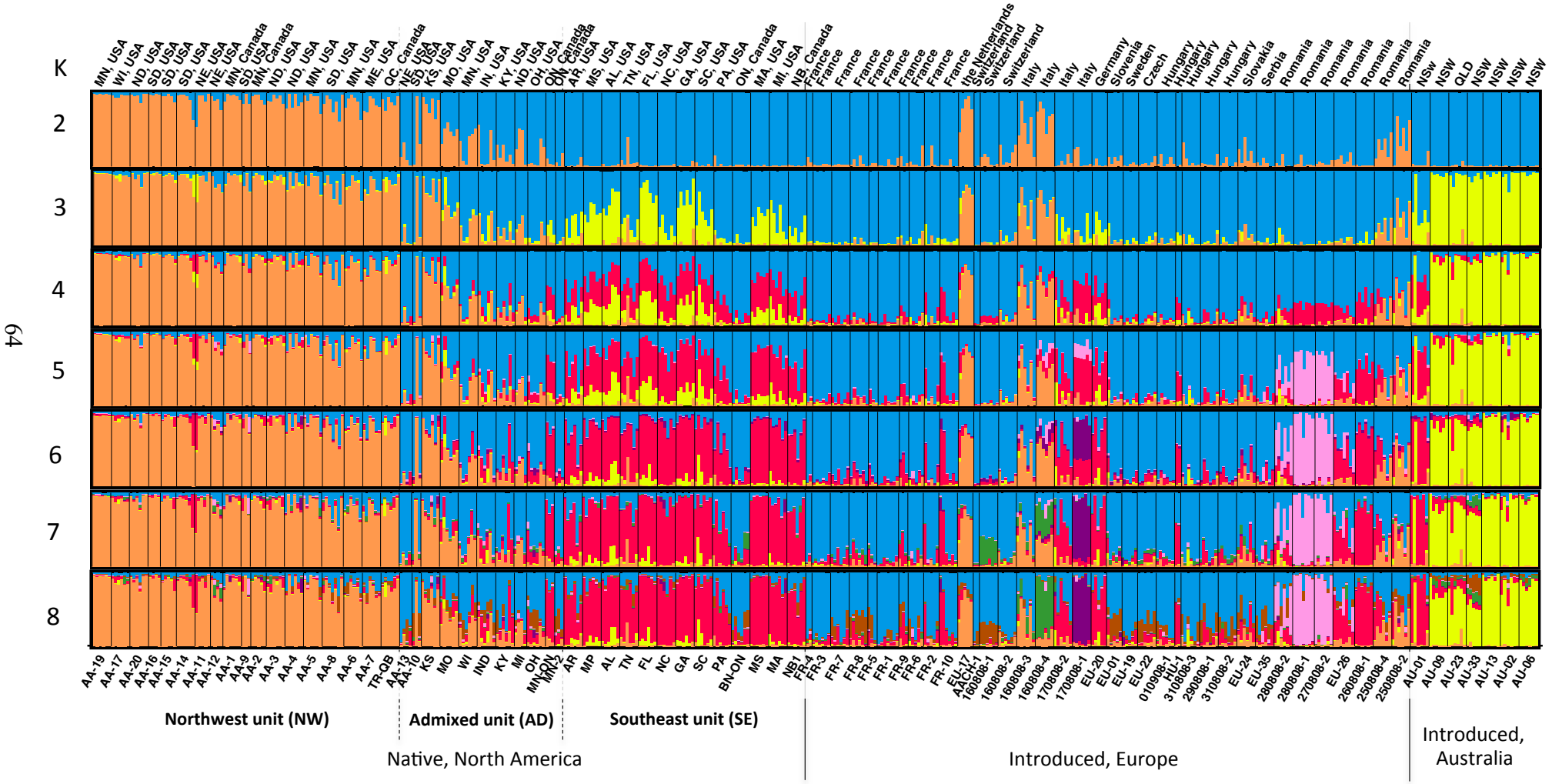


Figure S2b Complete overview of STRUCTURE output for K from 2 to 8 summarized with CLUMPP and visualized with Distruct for all (b, corresponds to Fig. 5 in main text), North American (b), European (c) and Australian (d) populations.

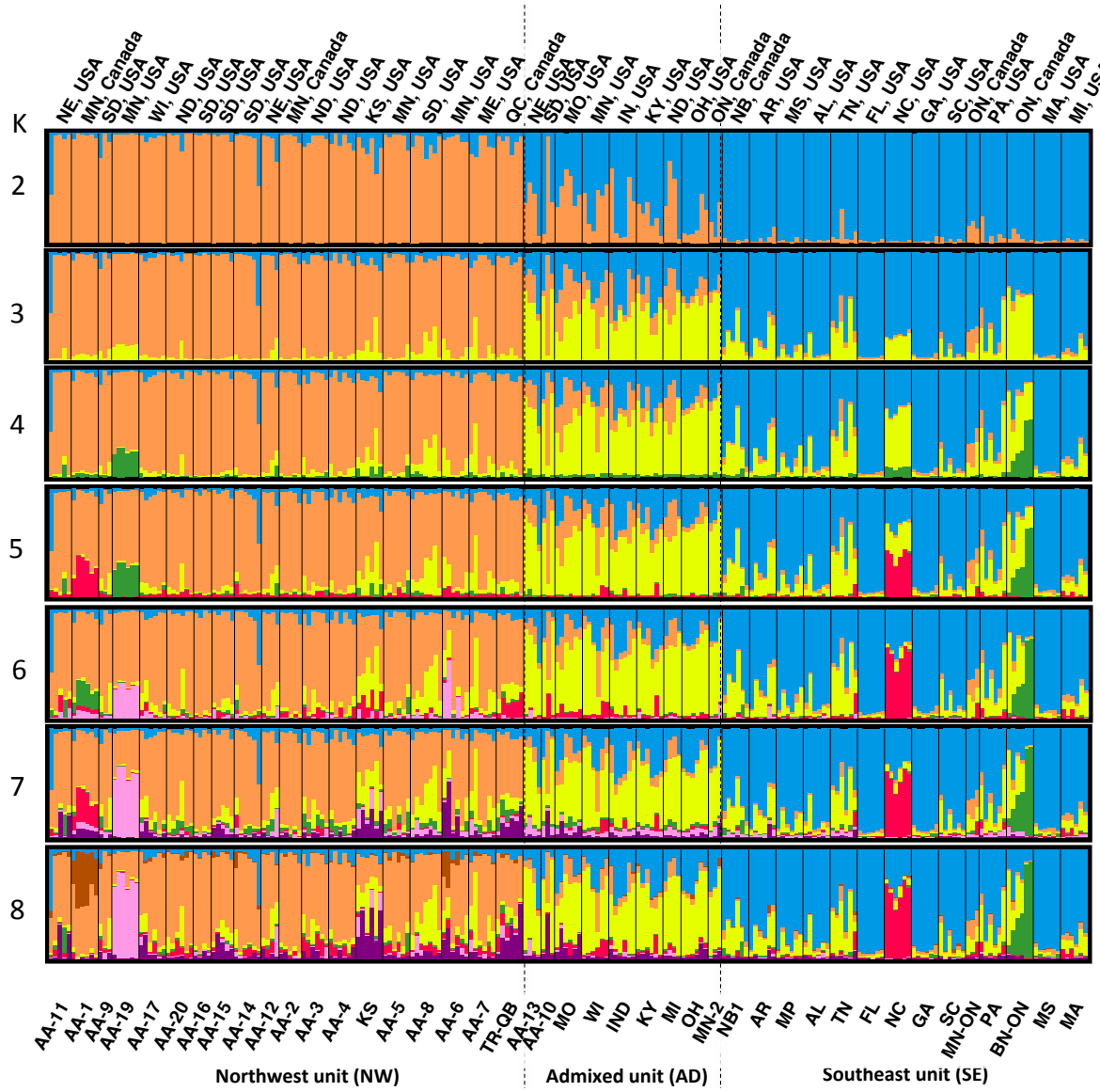


Figure S2c. STRUCTURE output for K from 2 to 8 for North American populations

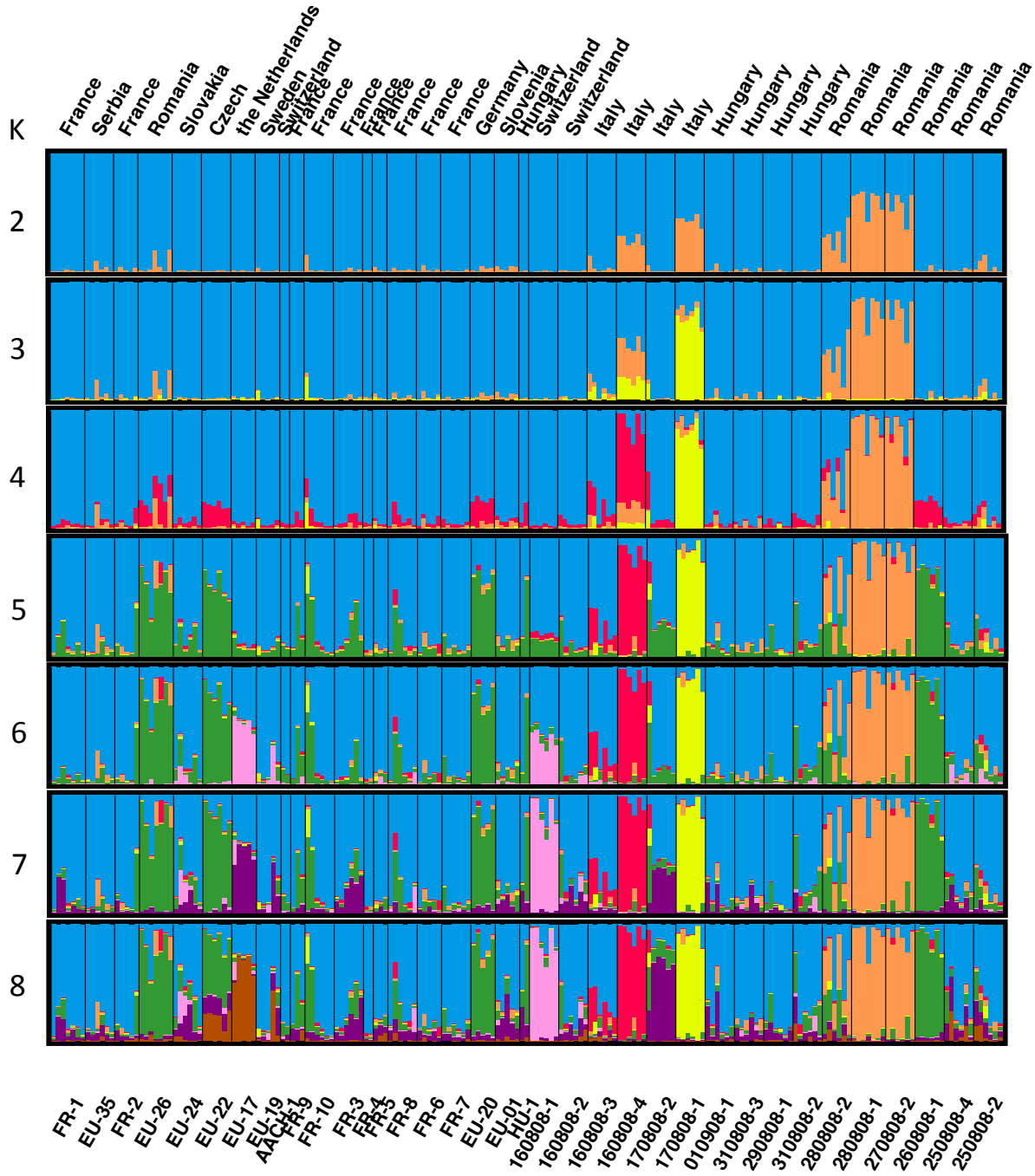


Figure S2d. STRUCTURE output for K from 2 to 8 for European populations

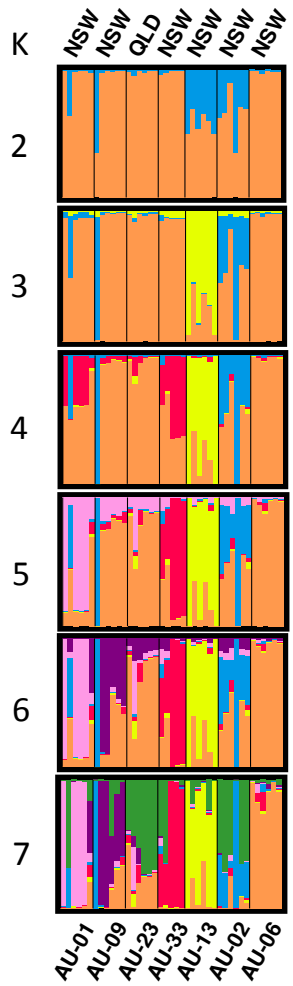


Figure S2e. STRUCTURE output for K from 2 to 8 for Australian populations. Only shown for K2-7, as when K exceeds the number of sampling locations (Australia had 7 locations), one or more clusters will have a mean assignment score no higher than 0.5 within any population. Such clusters should be removed from the analysis (Puechmaille, 2016).

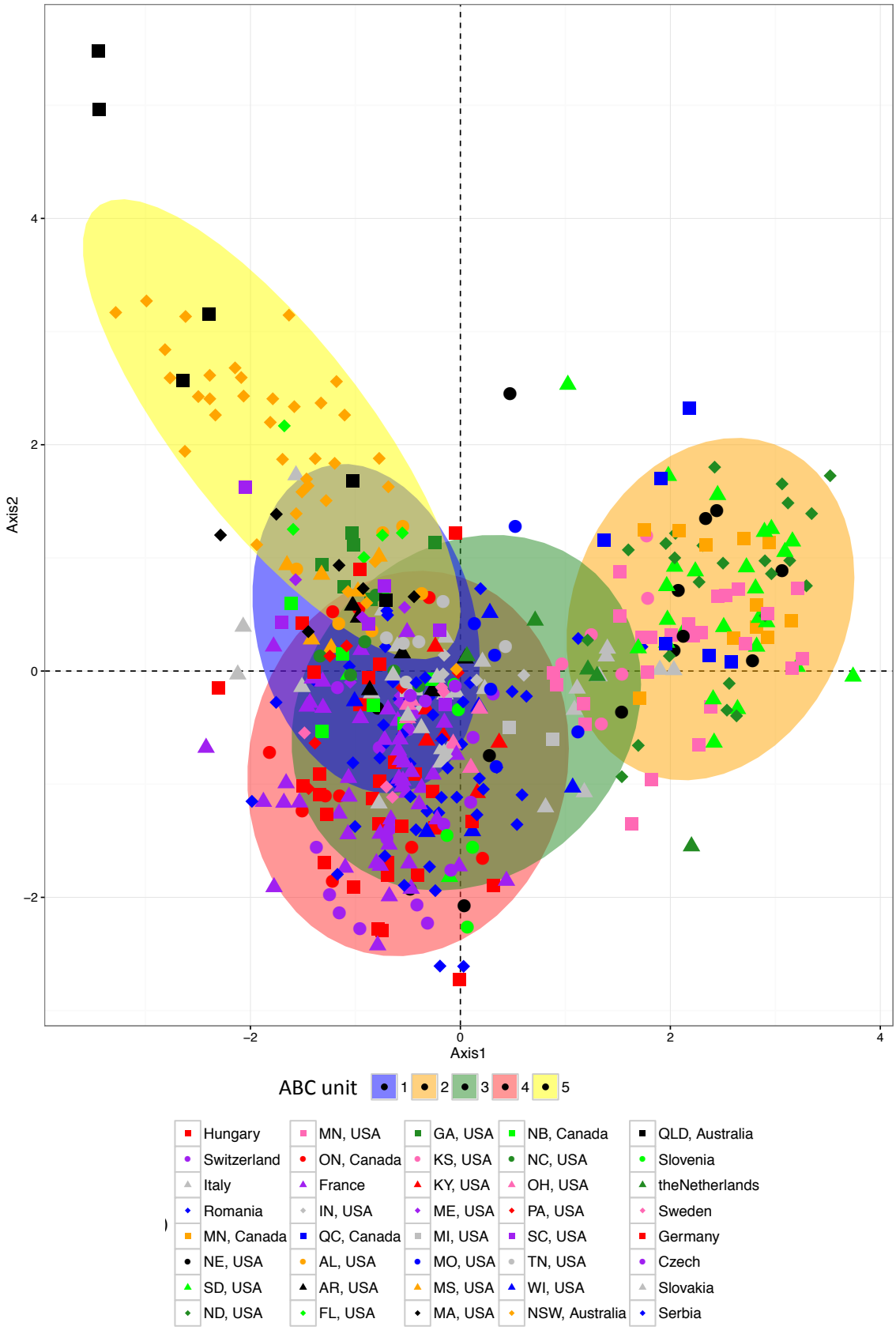


Figure S3. Genetic principle component analysis on allelic frequencies across 1022 randomly selected unlinked SNPs within 466 *Ambrosia artemisiifolia* samples. Points are color-coded and shaped by sampling State or Province/Country (Table S1). Ellipses represent 95% confidence interval of the data for genetic units as defined for ABC-RF. The first two principle components are shown (percentage of variation explained by axis 1=1.75%; axis 2 =1.12%).

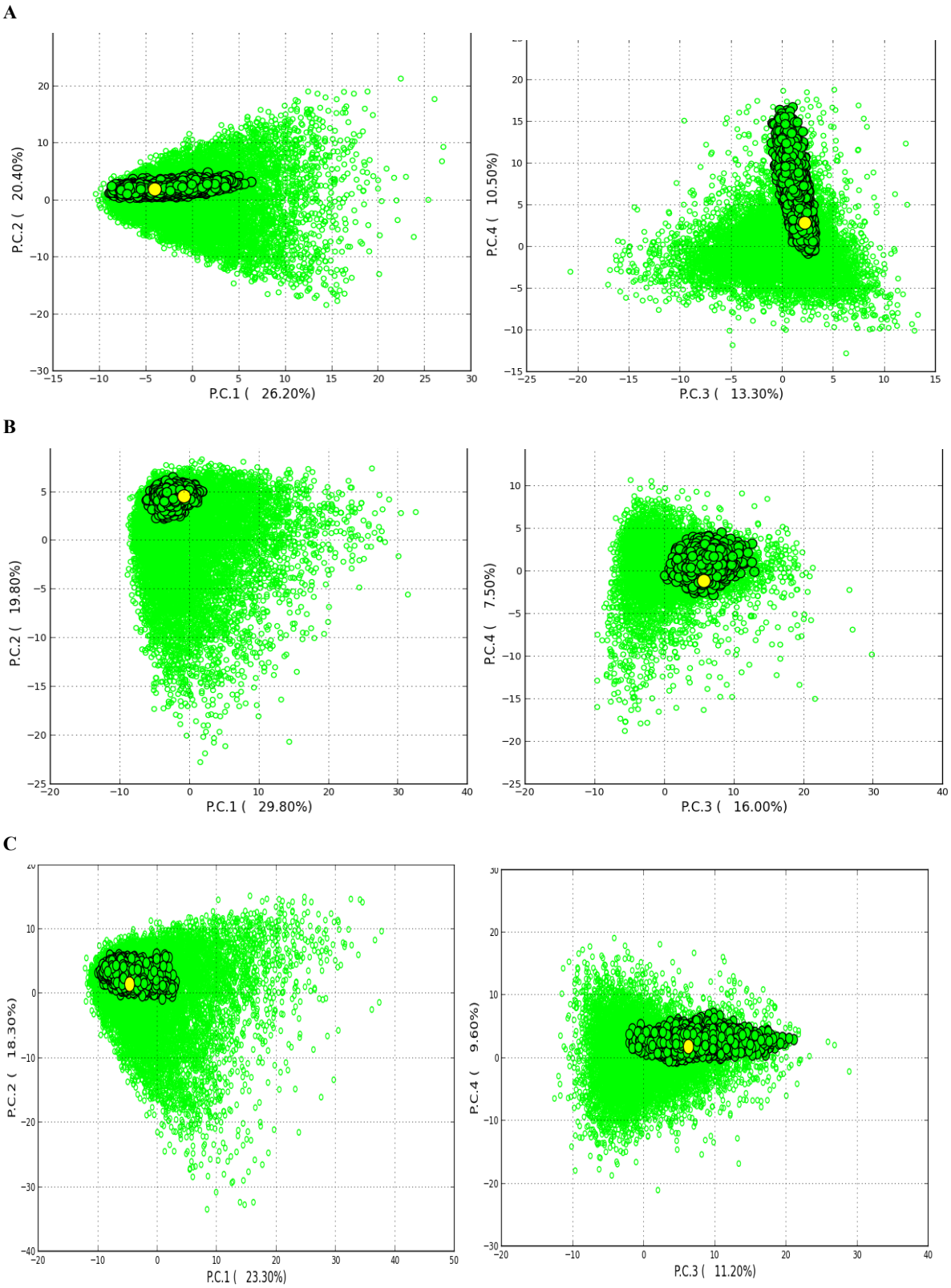


Figure S4. Principle component analyses illustrating model posterior checking of the 'best' scenarios representing the *Ambrosia artemisiifolia* introduction history: a) multiple introductions from native NE & SW units into Europe; b) single introduction from native admixed unit into Australia; c) Australian introduction from bridgehead EU. The plots show the observed dataset (yellow circle) with 10,000 datasets simulated from prior (green open circles) and posterior distributions (green closed circles) in the space of the first four principle components summarizing the 112 (a&b) and 220 (c) summary statistics, representing 70.4% (a), 73.1% (b) and 62.4% (c) of the total variation.

Chapter 3 – Rapid and repeated local adaptation to climate in an invasive plant

Lotte A. van Boheemen[1*], Daniel Z. Atwater[2], Kathryn A. Hodgins[1]

[1]School of Biological Sciences, Monash University, Clayton, VIC 3800, Australia
[2]Department of Biology, Earlham College, Richmond, IN 47374, USA

*Corresponding author, la.vanboheemen@gmail.com

Published in *New Phytologist* on October 26th, 2018

Summary

- Biological invasions provide opportunities to study evolutionary processes occurring over contemporary timescales. To explore the speed and repeatability of adaptation, we examined the adaptation of life-history traits to climate, using latitude as a proxy, in the native North American and introduced European and Australian ranges of the annual plant *Ambrosia artemisiifolia*.
- We explored niche changes following introductions using climate niche dynamic models. In a common garden, we examined trait divergence by growing seeds collected across three ranges with highly distinct demographic histories. Heterozygosity-fitness associations were used to explore effect of invasion history on potential success. We accounted for non-adaptive population differentiation using 11,598 SNPs.
- We revealed a centroid shift to warmer, wetter climates in the introduced ranges. We identified repeated latitudinal divergence in life-history traits, with European and Australian populations positioned at either end of the native clines.
- Our data indicate rapid and repeated adaptation to local climates despite the recent introductions and a bottleneck limiting genetic variation in Australia. Centroid shifts in the introduced ranges suggest adaptation to more productive environments, potentially contributing to trait divergence between the ranges.

Keywords

Invasion, trait evolution, climate niche dynamics, local adaptation, latitudinal clines, climate adaptation, heterozygosity fitness correlations

Introduction

During biological invasions species commonly spread over large and climatically diverse geographic areas. In doing so, they often re-establish within climatic niches found in their native ranges, or flourish in new environments (Sax & Brown, 2000; Allendorf & Lundquist, 2003; Atwater *et al.*, 2018). Although plasticity and broad ecological tolerance may facilitate success of invaders in many cases (e.g. Geng *et al.*, 2007; Zhang *et al.*, 2010), a growing number of empirical examples suggest that rapid adaptation to local conditions can also enable the establishment and spread of these species in the face of new selective dynamics (Huey *et al.*, 2000; Lachmuth *et al.*, 2011; Colautti & Barrett, 2013; Chown *et al.*, 2014; Turner *et al.*, 2014; Oduor *et al.*, 2016). As such, invasions provide an opportunity to study contemporary adaptive processes, which is key in an era of rapid, human-induced, environmental change.

Comparisons of clinal patterns in life history traits within native and introduced ranges provide an important opportunity to observe if rapid and repeated adaptation can occur in response to similar climatic gradients (Moran & Alexander, 2014). However, many factors could impact the trait evolution of native and invasive populations evolving in response to similar climatic gradients resulting in divergent outcomes. Mating system and demographic events, such as bottlenecks, genetic drift and admixture could differentially affect native and invasives' adaptive capacity or influence the route by which adaptive evolution proceeds (Lee, 2002; Facon *et al.*, 2006; Prentis *et al.*, 2008; Rius & Darling, 2014; Estoup *et al.*, 2016). Also, the probability of observing trait clines in the introduced range would be low if the strength of climate mediated selection is weak and/or the invasion is recent. By contrast, the invasion history could create patterns of trait variation (e.g., through climate matching (Maron *et al.*, 2004)) mimicking adaptive population divergence (Colautti *et al.*, 2009; Colautti & Lau, 2015). Shifts in the biotic environment could lead to divergence of multiple traits linked by trade-offs (Blossey & Notzold, 1995). Dissection of these various mechanisms is required to advance our understanding of the role of rapid evolution in invasive spread (Keller *et al.*, 2009; Bonhomme *et al.*, 2010; Lachmuth *et al.*, 2011; Agrawal *et al.*, 2015; Cristescu, 2015; Dlugosch *et al.*, 2015a). More generally, as many single species have invaded several distinct regions of the globe, comparisons of the native range to multiple successful introductions could illuminate if and when traits evolve in parallel.

Across environmental gradients, trade-offs among life history traits can shape adaptive trait divergence in response to local conditions and impact the evolutionary trajectory of trait combinations in invasive populations (Etterson & Shaw, 2001; Griffith & Watson, 2005; Colautti *et al.*, 2010; Hodgins & Rieseberg, 2011; Colautti & Barrett, 2013). For example, reductions in season-length at higher latitudes are frequently reported to select for early flowering at the cost of diminished plant size (Colautti *et al.*, 2010; Li *et al.*, 2014). Coordinated shifts in life-history traits along latitudinal gradients within ranges have been documented in

several invasive plants (Dlugosch & Parker, 2008b; Hodgins & Rieseberg, 2011; Colautti & Barrett, 2013).

Latitudinal patterns in plant size could have important evolutionary consequences for other plant traits. Variation in plant size can influence optimal resource allocation to male and female sex function (Charnov, 1982; De Jong & Klinkhamer, 1989; Klinkhamer *et al.*, 1997). In wind-pollinated plants, height can affect fitness directly through more effective pollen dispersal in taller plants (Burd & Allen, 1988; Klinkhamer *et al.*, 1997), as well as indirectly through increased availability of resources (Lloyd & Bawa, 1984; De Jong & Klinkhamer, 1989; Klinkhamer *et al.*, 1997; Zhang, 2006). Outcrossing wind-pollinated plants are predicted to adaptively change sex allocation to be more male-biased with increase in size (Lloyd, 1984; De Jong & Klinkhamer, 1989; de Jong & Klinkhamer, 1994; Klinkhamer *et al.*, 1997), as local seed dispersal should lead to saturating female gain curves (Lloyd & Bawa, 1984; Sakai & Sakai, 2003), yet more linear male function gain curves are expected (Klinkhamer *et al.*, 1997; Friedman & Barrett, 2009). Latitudinal clines in height could therefore be expected to lead to adaptive shifts in sex allocation. However, studies investigating the evolution of sex allocation patterns over wide geographic ranges are rare (Guo *et al.*, 2010; Barrett & Hough, 2012).

We examine the repeatability and divergence of important life-history traits in the native North American and introduced European and Australian ranges of *Ambrosia artemisiifolia*. We raised seeds collected at 77 locations from broad climatic scales in a common garden and accounted for non-adaptive genetic differentiation using 11,598 genotype-by-sequencing SNPs, as neutral processes could impact trait variation. We investigate four specific questions: *1) Do native and introduced populations occur in similar climates niches?* As climate is likely important in governing trait variation in this species, we examined climatic niche shifts following introduction to assess how traits might be predicted to diverge within and among the ranges. *2) Is there evidence for rapid parallel adaptation to latitude?* Repeatable trait clines for each range along latitudinal gradients, strongly associated with many aspects of climate, would provide strong evidence that rapid adaptation to similar selective environments has occurred. We additionally explore coordinated shifts in traits potentially linked by trade-offs. *3) Is there evidence for trait differentiation between native and introduced ranges?* By examining patterns across multiple introduced ranges, we can explore if novel recipient communities generated trait divergence during introduction, or if adaptation to local climates dominates patterns of trait variation. 4*) Is there a correlation between heterozygosity and fitness related traits?* Significant correlations would reveal if demographic changes such as bottlenecks, admixture and inbreeding have likely impacted the evolution of traits during this species' extensive range expansion. Such correlations are predicted at the individual and population level in regions that have recently expanded their range, including those that have undergone admixture (Peischl & Excoffier, 2015).

Material and methods

Study species

Ambrosia artemisiifolia is a wind pollinated, outcrossing, hermaphrodictic annual, which has aggressively spread from its native North America to many regions worldwide (Laaidi *et al.*, 2003; Smith *et al.*, 2013). The first records documenting the invasion are in central France around 1850 (Chauvel *et al.*, 2006). Multiple introductions from distinct native sources ensued to both east and west Europe, resulting in levels of genetic variation equivalent to those found in the native range (Chun *et al.*, 2010; Gladieux *et al.*, 2010; Gaudeul *et al.*, 2011; van Boheemen *et al.*, 2017b). Genetic analysis suggests the Australian populations originate from a subsequent single introduction event around 80 years ago, derived from the European introduction, although the exact source is unknown (van Boheemen *et al.*, 2017b). Range expansion likely occurred both north- and southward following this south-Queensland introduction (Palmer & McFadyen, 2012; van Boheemen *et al.*, 2017b)

Latitudinal clines in phenology have been observed within the native North American range and the introduced ranges of Europe (Chun *et al.*, 2011; Leiblein-Wild & Tackenberg, 2014) and China (Li *et al.*, 2014), with earlier reproduction and greater relative investment in reproductive biomass in high-latitudinal compared to low-latitudinal populations (Chun *et al.*, 2011; Hodgins & Rieseberg, 2011; Li *et al.*, 2014). The wind-spread pollen is a major cause of hayfever worldwide and its medical treatment costs millions of dollars each year (Taramarcaz *et al.*, 2005), providing considerable incentive to understand the factors impacting pollen production in this species.

Climate niche dynamics

To estimate the climatic niche occupied by *A. artemisiifolia* in its native North American, introduced Eurasian and Australasian ranges, we used ordination-based species distribution models. Models were taken from a larger study of 835 species (Atwater *et al.*, 2018), where the methods are described in detail. Briefly, occurrence data were collected from the Global Biodiversity Information Facility and plotted in 2-dimensional climate space based on rotated component variables (RCA1: temperature and RCA2: precipitation) of the 19 WorldClim variables (Hijmans *et al.*, 2005) representing annual trends, seasonality and means in temperature or precipitation (Table S1, Supplementary Material). We selected these variables as they are commonly used in studies on species distribution and local adaptation, including studies of *A. artemisiifolia* (e.g. Leiblein-Wild & Tackenberg, 2014; Sun *et al.*, 2017). Geographic and climatic sampling bias was removed to produce estimates of occurrence probability in each set of climatic variables, for each of the three ranges (North American, Eurasian, and Australasian). We used Monte Carlo resampling (n = 120) to compare observed

niche dynamics to those expected using randomly resampled occurrence points (Atwater *et al.*, 2018). We used Schoener's *D* (Schoener, 1968) to estimate niche overlap between the native range and each introduced range, and we estimated niche stability, expansion (species occupying climates in their introduced range that are unoccupied in the native range), and unfilling (species occupying climates in the native range that are unoccupied in the introduced range)(Guisan *et al.*, 2014). Finally, we tested whether the location of the niche centroid differed between native and introduced ranges.

Data collection

To investigate local environments, we described climatic differences between 27 populations in the native range of North America, 32 populations in the introduced European range, and 18 populations in the introduced Australian range (Fig. S1, Supplementary Material). We used the 19 WorldClim variables and added a geographic dimension to the data by including altitude, latitude and longitude, as these variables are shown to be important in *A. artemisiifolia* growth and fitness (Chun *et al.*, 2011; Chapman *et al.*, 2014), with the latter affecting season length and photoperiod (Ziska *et al.*, 2011). To explore associations between climatic and geographic variables in the sampled populations, we applied a principal component analysis (PCA). We opted to present latitudinal trait clines only, as *i)* the primary principle component was highly correlated with latitude (Fig. S2); *ii)* precipitation and temperature variables were strongly associated with latitude (Spearman's ρ^2=0.326-0.417 (Table S2); *iii)* exploratory analyses revealed associations between each trait and many climatic and geographic variable were highly similar to trait~latitude trends (results not shown); and *iv*) is shown to affect *A. artemisiifolia* season length and photoperiod (Ziska *et al.*, 2011).

To assess the potential for adaptive differentiation along latitudinal gradients, we measured trait variation in a common garden of raised seeds collected at broad geographical scales across the three ranges. This is a traditional approach to detect genetic differences among populations (e.g. Bossdorf *et al.*, 2005; Colautti *et al.*, 2009; Hodgins & Rieseberg, 2011; Savolainen *et al.*, 2013; de Villemereuil *et al.*, 2016). We collected seeds in 2013-2014 and randomly selected an average of 12 maternal families with 20 seeds per family from each population (for a full description of data collection methods, see Supplementary Methods). Following a 6-week stratification at 4°C (Willemsen, 1975), we placed seeds in a 30°C germination chamber with 12h light/dark cycle. Two weeks after germination, we planted a randomly selected seedling from each maternal line in a random order into 100ml kwikpot trays with Debco Seed Raising Superior Germinating Mix. We top-watered all plants and artificially manipulated daylight following the light cycle at 47.3°N (median latitude over all sampling locations). One month later, we performed a second transplant (hereafter day 0) to 0.7L pots with Debco mix and 1.5ml slow-release fertilizer (Osmocote Pro, eight to nine months). We

examined variation in life-history traits including phenology, growth, and vegetative and reproductive allocation (Table 1).

Table 1. Traits included in this study.

Trait	Description
Max. height	Maximum measured height
Total biomass	Above- and belowground dry biomass
Max. growth rate	Sensu Chuine *et al.* (2001)
Flowering onset	First recorded day of flowering (number of days after second transplant); first day of pollen release (male function) or receptive female function
Dichogamy	First recorded day of pollen release - first recorded day of receptive female function (a positive value is protogyny, a negative value is protandry)
Floral sex allocation (female/male)	Female (seeds) / male (raceme) dry weight (a value >1 is higher biomass allocation to female function)
Weight per seed	Dry weight per seed in milligrams, averaged over 20 seeds (where available)
Total reproductive biomass	Female (seeds) and male (raceme) dry weight
Relative reproductive biomass	Total reproductive biomass / total plant biomass
Specific leaf area	Leaf area of fully expanded fresh leaf/leaf dry weight

To assess neutral genetic differentiation underlying trait variation resulting from non-adaptive evolutionary processes, we extracted DNA from leaf tissue of 861 individuals and performed double-digest genotype-by-sequencing library preparation (see Supplementary Methods). We aligned and filtered raw sequences following van Boheemen *et al.* (2017b). Briefly, SNPs were aligned using BWA-mem (Li & Durbin, 2009) to a draft reference genome for *A. artemisiifolia* (van Boheemen *et al.*, 2017b). We called variants with GATK HaplotypeCaller and filtered SNPs using GATK hard-filtering recommendations (McKenna *et al.*, 2010; Van der Auwera *et al.*, 2013). We identified a total of 11,598 polymorphic biallelic SNPs with 50% SNP call rate. We inferred population genetic structure and calculated individual and population level q-scores for the most likely number of clusters K (=2) with the Bayesian clustering method STRUCTURE v2.3.4 (Pritchard *et al.*, 2000). We used these STRUCTURE q-values as a proxy for genetic and, therefore, trait differentiation resulting from non-adaptive (neutral) mechanisms.

To explore the impact of heterozygosity on fitness-related traits we calculated heterozygosity (H_O) for each population as the proportion of heterozygous loci out of the total number of called genotypes for each individual. Introductions could diminish or increase heterozygosity, which in turn could inhibit or stimulate invasion (e.g. inbreeding due to genetic drift or heterosis following admixture).

Statistical analyses

We conducted all statistical analyses in R v3.4.3 (R Core Team, 2017). We improved normality and reduced heteroscedacity of the data by square root or log-transforming traits where appropriate. We tested all univariate linear mixed models using the lme4 package (Bates *et al.*,

2014). To account for demographic history in patterns of trait divergence, we included STRUCTURE q-scores (individual or population mean as appropriate) as a random effect in all univariate mixed models. We calculated significance of fixed effects using type III Wald F tests with Kenward-Roger's approximation of denominator degrees of freedom and step-wise removed non-significant ($p>0.05$) fixed effects using the lmerTest package (Kuznetsova *et al.*, 2015), starting with the highest order interactions. To reduce false discovery of associations due to the number of traits being tested, we 'fdr' corrected p-values (hereafter q-values) of the combined final models using the p.adjust function, further reducing models when applicable.

To explore patterns of latitudinal trait divergence among ranges, potentially indicative of local adaptation, we tested population mean trait responses to range (native North America, introduced Europe and Australia), latitude, their interaction and latitude^2 (to account for non-linear trait responses) in multi- and univariate models ((M)ANCOVA). In these analyses, latitude values are not randomly distributed among ranges due to the geographic distribution of *A. artemisiifolia*, suggesting a violation of independence (Miller & Chapman, 2001). However, the values of the covariate (latitude) are observational and not manipulated by the independent variables (range) and the (M)ANCOVA assumptions are thus not violated (Keppel, 1991). We increased the power of the multivariate analysis (Scheiner, 2001) by removing highly correlated traits (Spearman's $\rho^2>0.6$, Table S3, Supplementary Material) and calculated the approximate F-statistics and Wilks' λ (multivariate F-value) to measure the strength of the associations.

To explore differences in latitudinal trait clines between ranges as revealed in the ANCOVAs, we tested for significant two-way interactions between range and latitude which is indicative of non-parallel trait~latitude slopes among the ranges. To further dissect the extent of trait divergence and its dependence on latitude, we compared model estimates of traits in the introduced ranges to the native estimates at the minimum and maximum observed latitude in each range, where applicable (EU_{min}, NA_{max}, NA_{min} and AU_{max}) (Fig. S4, Supplementary Material). We tested overall pairwise range differences in trait values for significant range effects, when higher order interactions involving range were not significant. We explored the highest order significant interactions, using χ^2 tests with Holm p-value correction using the phia package (De Rosario-Martinez, 2015.).

We verified the presence of the well-described plant height-flowering time trade-off and examined associated patterns in other reproductive traits (dichogamy and sex allocation) within ranges. We tested linear relations between individual trait values of plants in mixed models, with plant height, range and their interaction as fixed effects. In addition to individual STRUCTURE q-values, we added population as random effect. A significant interaction between range and height indicated a differential slope between the focal reproductive trait and plant height between ranges.

To investigate geographical patterns in genetic diversity of populations, we tested the effect of range, latitude, their interaction and latitude^2 on population mean H_O. To explore the effect of heterozygosity on chosen traits, we included population mean H_O, latitude, their interactions with range and latitude^2 as fixed effects. To reduce false positive association due to multiple testing, we only tested the response of growth (plant height and biomass) and fitness (total reproductive biomass and average seed weight) related traits. As we identified signatures of repeated local adaptation in floral sex allocation (Results), we also tested the response of this trait.

Results

Climate niche dynamics

Niche overlap (*D*) was significantly lower than expected between the North American native and Eurasian invasive range ($P<0.001$) although the native and Australasian range did not have significantly lower overlap than expected ($P=0.425$). However, niche stability was low between the native and both invasive ranges ($P<0.001$). Climatic niche unfilling and expansion were not significantly different than the null model, except that especially low expansion was found in the Eurasian population ($P=0.017$), meaning that while the niche shifted, the species did not tend to occupy completely novel climates in its Eurasian range. In both introduced ranges the niche centroid shifted significantly towards hotter, wetter climates ($P<0.001$; Fig. 1).

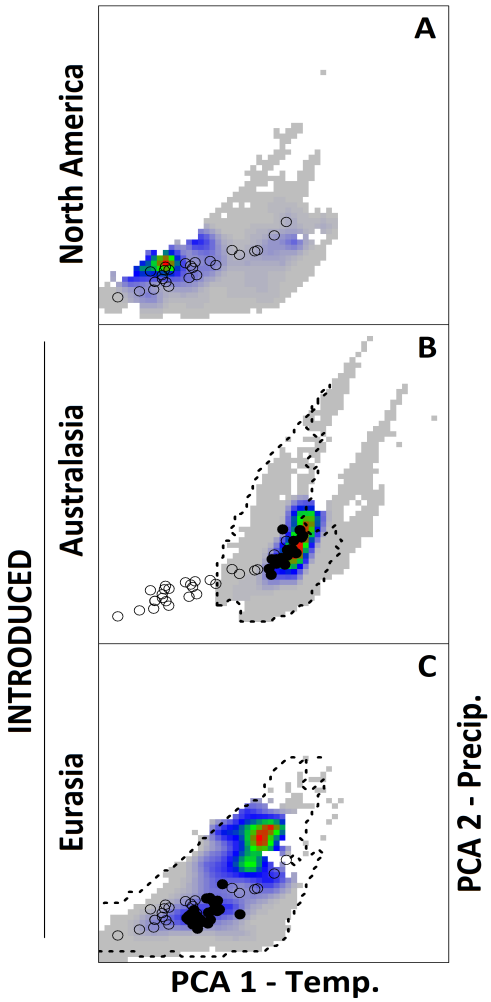


Figure 1. Climatic niche models of the native North American population (A) and introduced Australasian (B) and Eurasian (C) populations. Each panel shows the climate space occupied in the respective range, with a PCA variable corresponding to temperature on the x-axis and a PCA variable corresponding to precipitation on the y-axis. Colour indicates occurrence probability in a given climate (red: high occurrence, green: medium occurrence, blue: low occurrence, grey: no occurrence). Open circles plot climates of the North American source localities. Closed circles plot the climates of the respective introduced range. On panels B and C, the dashed line encloses the climates shared by both the native North American and respective introduced range.

We note that this climate analysis compares the entire Eurasian and Australasian ranges with the native North American range. For our common garden analysis, we focused only on specific latitudinal transects in Europe and Australia to examine how traits have evolved along latitudinal clines during invasion, and did not include any Asian populations. We include this more general analysis of climate niche to assess how climate shifts might contribute to trait divergence among the ranges generally, although our sampling for the common garden was more limited.

Climate of sampled populations

Climatic variables across all populations in the native North American, introduced European and introduced Australian ranges could be effectively summarized by the first two principle components (PC) in the PCA (Fig. S2, Table S2, Supplementary Material), which together explained 70.06% of between-population variation. Here, PC1 was strongly associated with latitude, temperature and seasonality, whereas PC2 was mostly precipitation related (Fig S2, Table S2). The climate experienced by Australian populations was distinct from the North American and European ranges (Fig. S2, Table S2), with higher annual, winter and summer temperatures, and with lower seasonal variation. Moreover, given the sub-tropical location of Australian populations, the sampled populations experienced higher precipitation during the growing season (Fig. S2).

Repeatability in trait clines among ranges

Traits across all North American, European and Australian ranges were well summarized by the PCA, where the first two PCs explained 83.1% of all variation (Fig. S3, Supplementary Material). The main PCs were associated with each trait to a similar extent, so no trait syndromes were apparent. Traits measured in Australian plants were generally distinct from the other populations although there was some overlap in multivariate space among the ranges (Fig. S3, Table S3). Multivariate trait analyses revealed a significant two-way interaction between latitude and range ($F_{14,120}$=1.796, p=0.047, Wilks' λ=0.684) (Table S4) suggesting latitudinal trait clines exist, but do not have the same relationship within ranges for all traits. Further dissection of these patterns in univariate analyses revealed maximum growth rate, flowering onset, dichogamy, average seed weight, total reproductive biomass and specific leaf area (SLA) displayed similar latitudinal clines among ranges, indicated by a significant latitude effect but an absence of a higher-order interaction (Fig. 2, Table 2).

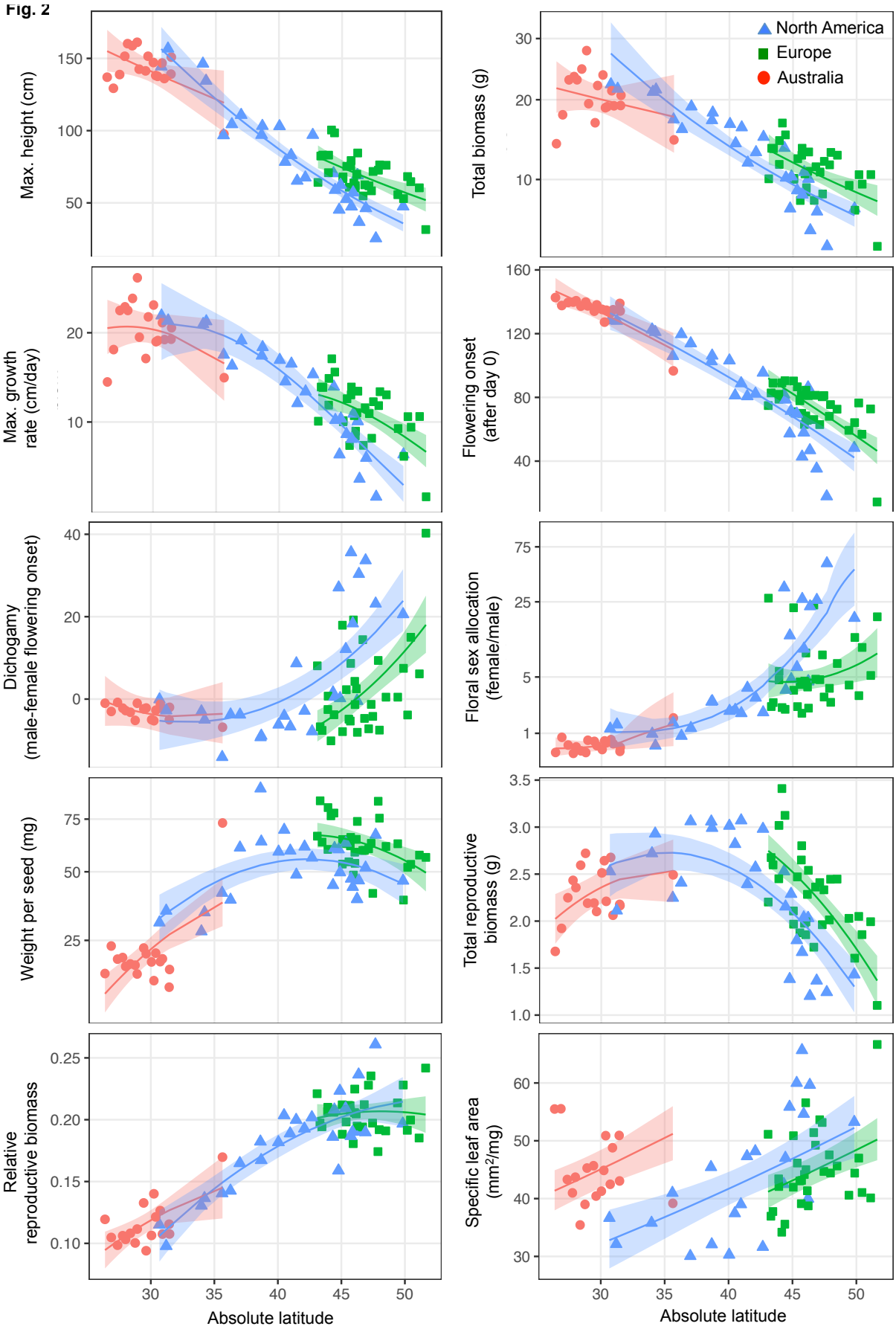


Figure 2. Traits responses (population means) to absolute latitude in the native North American (blue triangles), introduced European (green squares) and Australian (red circles) ranges, with model predictions and 95% shaded confidence intervals from step-wise reduced models (Table 2).

We identified range differences in latitudinal trait clines for maximum height, total biomass, floral sex allocation and relative reproductive biomass, as indicated by significant range:latitude interactions. However, all of these slopes were significantly different from zero and were in the same direction as the native North American patterns (Fig. 2, Table S5a). These results are consistent with local adaptation along a latitudinal gradient in all three ranges.

Table 2. *Ambrosia artemisiifolia* population mean trait responses to range, latitude, their interaction and latitude^2, with population q-values as random effects. We reported Wald type III F-values, with Kenward-Roger degrees of freedom as subscript and symbols specifying significance (fdr corrected q) of effect.

	Range	Latitude	Latitude^2	Range:Latitude
Max. height	$3.886_{2,70.57}$*	$55.963_{1,71.9}$***	$0.581_{1,70.97}$(ns)	$4.068_{2,70.35}$*
Total biomass	$3.344_{2,69.25}$#	$3.227_{1,70.98}$#	$5.214_{1,70.97}$*	$3.667_{2,69.14}$*
Max. growth rate	$6.405_{2,71.98}$**	$21.658_{1,70.76}$***	$2.025_{1,73}$(ns)	$1.101_{2,69.14}$(ns)
Flowering onset	$6.862_{2,71.978}$**	$158.301_{1,70.763}$***	$1.102_{1,73}$(ns)	$3.43_{2,69.14}$#
Dichogamy	$10.521_{1,73}$**	$8.163_{2,71.18}$**	$6.805_{1,72.99}$*	$3.157_{2,69.14}$#
Floral sex allocation (female/male)	$6.354_{2,69.82}$**	$6.048_{1,66.76}$*	$7.949_{1,66.27}$*	$7.094_{2,69.77}$**
Weight per seed	$4.156_{2,66.13}$*	$25.758_{1,72.41}$***	$23.704_{1,72.88}$***	$0.18_{2,67.29}$(ns)
Total reproductive biomass	$27.478_{1,73}$***	$6.856_{2,71.18}$**	$20.509_{1,72.99}$***	$3.427_{2,69.14}$#
Relative reproductive biomass	$6.436_{2,70.57}$**	$22.839_{1,71.9}$***	$2.997_{1,70.97}$(ns)	$6.422_{2,70.35}$**
Specific leaf area	$7.626_{2,67.39}$**	$17.167_{1,66.45}$***	$1.807_{1,63.65}$(ns)	$3.003_{2,65.14}$#

ns q>0.1; # q<0.1; * q<0.05, ** q<0.01; *** q<0.001

At higher latitudes, plants were shorter, weighed less, reached lower maximum growth rates and flowered earlier. Flowering onset extremes were 14-133 days after transplant (population means for EU20 and AU13, Fig. 2). In all ranges, dichogamy (the temporal separation of pollen dispersal and emergence of receptive stigma within an individual plant) was prevalent. Protogyny (emergence of stigma prior to pollen release) predominated at higher latitudinal populations, with receptive stigmas being visible up to 40 days before any pollen was released within the same plant (EU20)(Fig. 2). Conversely, protandry prevailed at latitudes below 40° from the equator, with pollen release occurring up to 14 days before stigmas became receptive (NC)(Fig. 2). Floral sex allocation followed a similar trend across ranges, with a slight male function bias at lower latitudes, shifting towards an extreme female function bias at high-latitude populations (Fig. 2). The biggest seeds were found at mid-latitudinal populations (38.6°N in KY), with seeds decreasing in size towards the high and low latitudes (Fig. 2, Table 2). Total reproductive biomass also showed a similar curved relationship, with the combined weight of racemes (male floral sex function) and seeds (female floral sex function) being up to three times as high at mid-latitudinal populations compared to high-latitudinal plants. In contrast, the relative reproductive biomass increased with latitude. Within each range, plants from lower latitudes had lower SLA (Fig. 2).

Trait divergence between ranges

While latitudinal trait clines were repeatable for many traits as described above, we identified shifts in trait values at comparable latitudes as revealed by significant range effects (Table 2). Maximum growth rates were highest in Europe and lowest in Australia (Fig. 2, Table 3a).

European plants also flowered later than North American and Australian plants at similar latitudes. The temporal separation between pollen release and the appearance of receptive stigma (dichogamy) was greater in the native North America compared to Europe (Fig. 2, Table 3a). European seeds were heavier, and plants had higher total reproductive biomass than those measured in the other ranges. At any given absolute latitude, Australian leaves had higher SLA compared to the native range, with lowest SLA in European populations (Fig. 1, Table 3a).

Table 3. Range differences of *A. artemisiifolia* population mean traits at comparable latitudes for significant (q<0.05) range effects (a, Table 2) and trait differences between ranges at minimum (min) and maximum (max) latitudes (Figure S4, Supplementary Material) for significant range:latitude interactions by comparing trait values (b, Table 2). We reported Wald type III F-values test, Kenward-Roger degrees of freedom as subscript and symbols specifying significance of effect.

a)	**North America - Europe**	**North America - Australia**	**Europe - Australia**
Max. growth rate	12.328_{1}**	4.264_{1}*	1.278_{1}**
Flowering onset	16.465_{1}***	0.123_{1}(ns)	13.269_{1}***
Dichogamy	18.8_{1}***	0.077_{1}(ns)	4.87_{1}#
Weight per seed	7.44_{1}*	2.059_{1}(ns)	6.002_{1}*
Total reproductive biomass	15.228_{1}***	0.811_{1}(ns)	6.516_{1}*
Specific leaf area	2.734_{1}#	14.936_{1}***	14.736_{1}***

b)	**North America (NA) - Europe (EU)**		**North America (NA) - Australia (AU)**	
Trait	**EU_{min}**	**NA_{max}**	**NA_{min}**	**AU_{max}**
Max. height	8.679_{1}**	25.236_{1}***	4.818_{1}#	0.039_{1}(ns)
Total biomass	0.339_{1}(ns)	15.572_{1}***	0.564_{1}(ns)	1.307_{1}(ns)
Floral sex allocation (female/male)	0.012_{1}(ns)	26.417_{1}***	1.586_{1}(ns)	0.096_{1}(ns)
Relative reproductive biomass	4.289_{1}(ns)	7.557_{1}*	0.065_{1}(ns)	0.017_{1}(ns)

ns q>0.1; # q<0.1; * q<0.05, ** q<0.01; *** q<0.001

Dissection of range differences in latitudinal trait clines (maximum plant height, total biomass, floral sex allocation and relative reproductive biomass) revealed most significant interactions between range and latitude were prompted by clinal differences between the introduced European and native North American populations (Table S5, Supplementary Material). For these traits, the discrepancy between North American and European trait values increased with increasing latitudes, such that at high-latitude populations, European plants were taller, heavier and less female-biased in floral sex allocation (Table 3b). Moreover, Australian plants found closest to the equator were significantly shorter than native North American expectations (Table 3b).

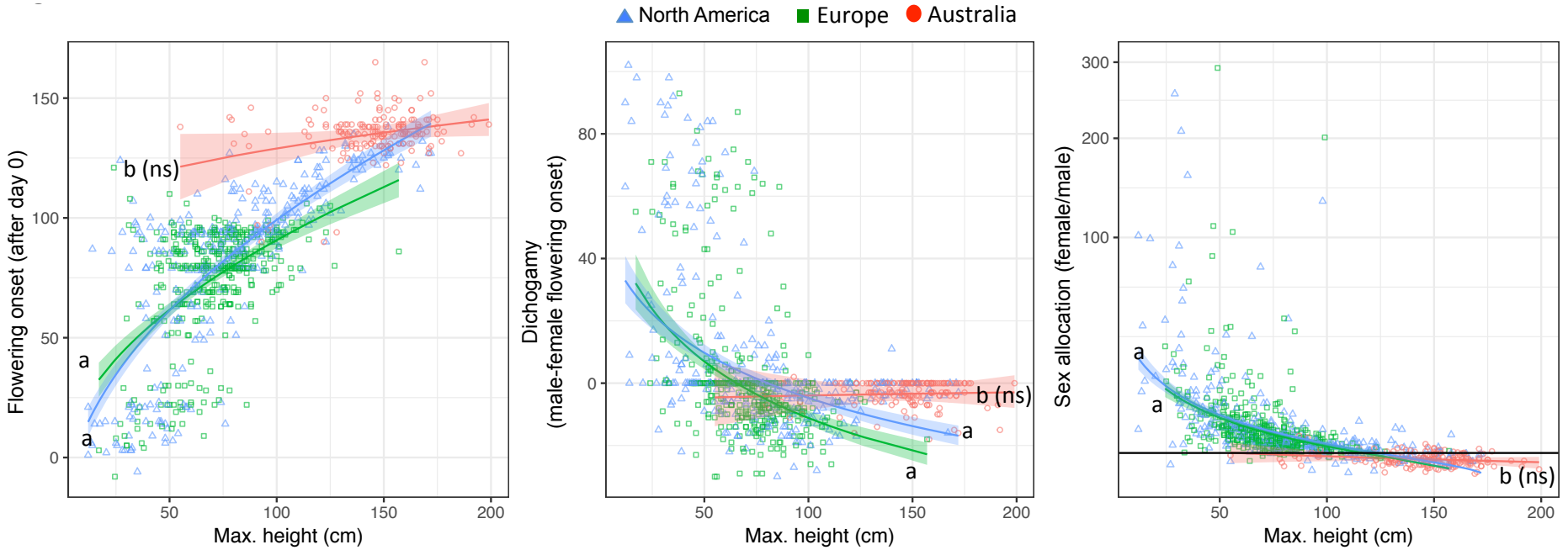


Figure 3. Flowering time, dichogamy and floral sex allocation responses to maximum plant height (individual values) in the native North American (NA, blue triangles), introduced European (EU, green squares) and Australian (AU, red circles) ranges, with model predictions and 95% shaded confidence intervals from step-wise reduced models (Table 4). Differences in slopes are indicated by letters and are significantly different from zero unless otherwise indicated (ns)(Table S5, supporting information).

Table 4. Flowering time and sex function allocation responses to maximum plant height (Height), range and their interaction, with individual q-values and population as random effects. We reported Wald type III F-values test, Kenward-Roger degrees of freedom as subscript and symbols specifying significance of effect.

Trait	Range	Height	Range:Height
Flowering onset	$20.658_{2,606.14}$***	$105.082_{1,766.54}$***	$14.767_{2,671.02}$***
Dichogamy	$6.503_{2,510.44}$**	$51.235_{1,670.96}$***	$10.621_{2,545.18}$***
Floral sex allocation (female/male)	$17.451_{2,508.47}$***	$138.084_{1,659.31}$***	$16.963_{2,508.47}$***

ns q>0.1; # q<0.1; * q<0.05, ** q<0.01; *** q<0.001

Trade-offs between life-history traits

We tested for the presence of a trade-off between plant height and flowering time and investigated associated patterns in dichogamy and floral sex allocation and height. As expected, taller plants started flowering later in both the native and the introduced European range, though this pattern was not significant in introduced Australian populations (Fig. 3, Table 4, Table S6, Supplementary Material). Correspondingly, we observed protogyny and a large female-biased sex allocation in short plants versus protandry with a slight male bias in tall plants. These dichogamy associations with height were not significant in Australia (Fig. 3, Table 4 & S6). However, it is possible that the maximum pot size used in the greenhouse may have limited the growth of some large Australian plants.

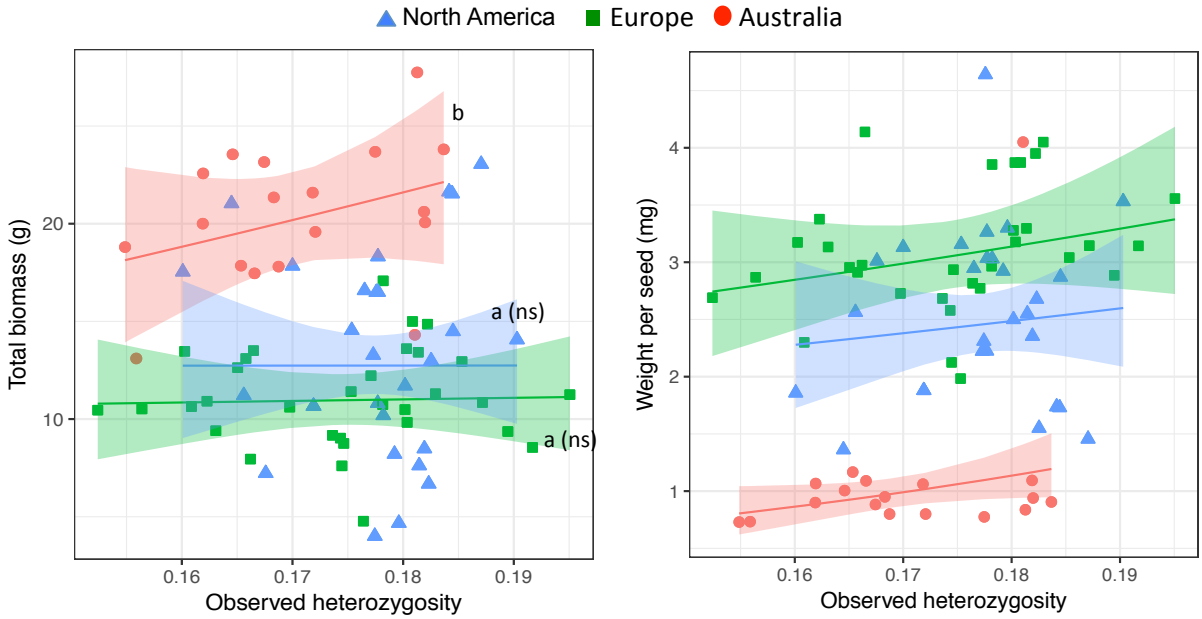


Figure 4. Total biomass and weight per seed response to heterozygosity (population means) in the native North American (blue triangles), introduced European (green squares) and Australian (red circles) ranges, with model predictions and 95% shaded confidence intervals from step-wise reduced models (Table 5). Slopes of predicted lines are significantly different from zero, unless otherwise indicated as (ns).

Associations between heterozygosity and life-history traits

To identify geographic patterns in observed heterozygosity (H_O), we tested the effect of range, latitude and their interaction on H_O. We found no latitudinal patterns in H_O varying within ranges (range:latitude, χ^2_1=3.811, p=0.149) or among all ranges (latitude, χ^2_1=0.000, p=0.986). We did identify variation between ranges (range, (χ^2_1=6.446, p=0.040), resulting from significantly lower H_O in Australia compared to native North America (χ^2_1=6.446; p=0.033). When accounting for latitude and population genetic structure, we found a significant interaction effect between mean population H_O and range on total biomass (Fig. 4, Table 5). Pairwise range comparisons in post-hoc tests revealed a higher H_O that was associated with heavier Australian plants (Table S7, Supplementary Material). Moreover, we found that mean

population H_O was positively correlated with seed size in all ranges (Fig. 4, Table 5). We found no effect of mean observed heterozygosity on plant height, phenology, dichogamy, total or relative reproductive investment and floral sex allocation (Table 5).

Table 5. Trait responses (population mean) to range, latitude, latitude^2, heterozygosity (H_O) and interactions, with population q-values as random effects. We reported Wald type III F-values test values, Kenward-Roger degrees of freedom as subscript and symbols specifying significance of effect.

	H_O	Range	Latitude	Latitude^2	H_O:Range	Range:Latitude
Max. height	$0.007_{1,70.75}$(ns)	$3.435_{2,69.52}$#	$51.007_{1,70.95}$***	$0.198_{1,69.96}$(ns)	$0.095_{2,59.46}$(ns)	$3.699_{2,69.23}$#
Total biomass	**$8.455_{1,53.099}$***	$4.942_{2,65.157}$*	$108.362_{1,68.316}$***	$0.699_{1,69.812}$(ns)	**$3.947_{2,64.456}$***	$2.055_{2,66.152}$(ns)
Total reproductive biomass	$0.263_{1,71.8}$(ns)	$7.123_{2,70.12}$**	$19.116_{1,71.91}$***	$24.912_{1,71.96}$***	$0.274_{2,59.46}$(ns)	$1.721_{2,68.2}$(ns)
Weight per seed	**$5.858_{1,38.45}$***	$3.271_{2,47.3}$#	$23.516_{1,71.93}$***	$20.776_{1,71.84}$***	$0.069_{2,25.95}$(ns)	$0.332_{2,55.71}$(ns)
Floral sex allocation (female/male)	$0.12_{1,69.63}$(ns)	$6.619_{2,68.56}$**	$6.229_{1,69.87}$*	$7.55_{1,69.84}$*	$0.096_{2,59.46}$(ns)	$7.287_{2,68.47}$**

ns q>0.1; # q<0.1; * q<0.05, ** q<0.01; *** q<0.001

Discussion

We show genetically based differentiation along multiple latitudinal clines in all examined traits including plant size, growth, reproductive investment, phenology, dichogamy, SLA, and sex allocation. Remarkably, the clinal patterns apparent in the native range evolved repeatedly within both introduced ranges over the course of only 100-150 years and despite limited neutral genetic variation in the introduced Australian range. These patterns are consistent with rapid adaptation, as we accounted for neutral genetic differentiation and plasticity. Moreover, low effects of maternal environment are expected (Hodgins & Rieseberg, 2011) and the introduction history of this species (van Boheemen *et al.*, 2017b) reveals climate-matching (Maron *et al.*, 2004) is unlikely. The adaptive trait divergence at similar latitudes, together with a centroid shift to warmer and wetter climates in the introduced ranges, could suggest invasive populations have adapted to more productive environments following introduction. The observed rapid evolution has implications for the evolutionary potential of this species and further range expansion following climate change. Furthermore, the divergence of reproductive traits such as flowering time, sex allocation and seed size during recent range expansion should impact the production of allergenic pollen as well as the abundance and dispersal of seed that could impact spread.

Climate niche shifts

Higher resource levels, such as increased water availability, are a known contributor to invasion in many plant species (Blumenthal, 2006; Dlugosch *et al.*, 2015b). Increased resource availability may occur through a shift in the fundamental or realized niche during invasion. The latter can result from reductions in competition, perhaps reflecting the presence of a vacant niche in the introduced range (e.g. Dlugosch *et al.*, 2015b). Climate niche dynamics analysis reveals higher *A. artemisiifolia* abundance in warmer and wetter climates in the introduced

ranges compared to the native range. It is possible that this centroid shift reflects an historic effect where colonization of warmer and wetter environments occurred earlier, or perhaps by genotypes pre-adapted to these climates (*but see* van Boheemen *et al.*, 2017b). Alternatively, the shift might reflect changes in biotic interactions leading to greater abundance of this species in high resource environments or differences in the availability of these climates in the introduced regions. Evolutionary processes that allow introduced species to colonize warmer and wetter environments than those occupied by native plants could also cause centroid shifts. This evolutionary interpretation is supported as Australian populations follow trait trajectories parallel to, but extending beyond, those of the native range.

Repeated latitudinal clines

Our common garden experiments using samples collected across multiple similar latitudinal gradients, reveal that local adaptation can happen quickly and predictably, with repeated evolution of native clines in both of the introduced ranges. Latitudinal clines in phenology and size are a common feature of many geographically widespread plant species (e.g. Colautti *et al.*, 2010; Li *et al.*, 2014), with *A. artemisiifolia* flowering shown to be driven by the association between season length and latitude (Ziska *et al.*, 2011). Short season lengths at high latitudes can select for earlier flowering (Bradshaw & Holzapfel, 2008; Colautti & Barrett, 2013), while the evolution of delayed flowering at low latitudes reflects the trade-off between size and the timing of reproductive maturity, where fitness is maximized by flowering later at a large size (Colautti *et al.*, 2010; Colautti & Barrett, 2013). This correlation between plant size and flowering time has been reported for *A. artemisiifolia* (Hodgins & Rieseberg, 2011; Leiblein-Wild & Tackenberg, 2014; Scalone *et al.*, 2016) and our results are consistent with rapid genetic differentiation in plant size, growth rates and phenology in response to latitude-associated selection pressures such as season length.

We exposed repeated patterns of genetic differentiation in sex allocation strategy over similar latitudinal clines, consistent with rapid adaptation to local climate. Plants sourced from higher latitudes displayed female-biased sex-allocation and protogyny, with more balanced floral sex allocation and a decrease in the temporal separation of male and female function towards the equator. Previous studies on *A. artemisiifolia* showed plasticity for sex-allocation and dichogamy in relation to plant size (Paquin & Aarssen, 2004; Friedman & Barrett, 2009; Friedman & Barrett, 2011) and ample genetic variation for evolution to act on (Friedman & Barrett, 2011). The patterns observed this study are consistent with theory on sex allocation in outcrossing wind-pollinated hermaphrodites with local seed dispersal. Spatial heterogeneity has been observed in animal pollinated plants, where small, resource limited individuals often allocate more resources to male function (Korpelainen, 1998; Guo *et al.*, 2010). However, these studies are on wild populations and cannot separate environmental and genetic effects in

allocation patterns along resource gradients. Our current findings from common garden experiments are therefore novel in identifying genetic differentiation among populations in sex allocation over spatial gradients.

Trait divergence among the ranges

Many hypotheses aim to explain the success of invasive species, including the evolution of increased competitive ability (EICA) through escape from native herbivores (Blossey & Notzold, 1995). Moreover, if trade-offs between performance and abiotic stress tolerance occur, greater resources could facilitate the evolution of more competitive phenotypes in introduced ranges (Grime, 1977; Bossdorf *et al.*, 2005; He *et al.*, 2010; Dlugosch *et al.*, 2015b). Our reported trait shifts in European population compared to natives at equivalent latitudes indeed suggest an increase in competitive ability, commonly measured as elevated plant growth and reproductive effort (Felker-Quinn *et al.*, 2013). As no evidence for EICA is found in Europe (Hodgins & Rieseberg, 2011), these observations might reflect the warmer and wetter European climate (Fig. 1). Notably, although Europe was identified as the introduction source for Australian populations (van Boheemen *et al.*, 2017b), traits measured within each range were highly dissimilar. Most of the sampled Australian populations extended beyond absolute latitudes of the other populations and were situated in warmer, less seasonal climates (Fig. S2). These factors might explain Australian trait variation beyond values observed in source populations.

Heterozygosity and invasion

Genetic drift within small founding populations and on the invasion front can lead to reduced genetic diversity, potentially impacting additive genetic variation (Dlugosch & Parker, 2008a; Excoffier & Ray, 2008; Peischl *et al.*, 2013; Bock *et al.*, 2015). In *A. artemisiifolia*, Australian populations were bottlenecked and potentially subjected to high genetic drift, whereas multiple introductions into Europe from distinct native sources has implicated admixture as a driver of invasion success (van Boheemen *et al.*, 2017b). We found the biomass of Australian, but not European, plants was indeed associated with heterozygosity, providing only partial support for the fitness benefits of heterozygosity during invasion (Peischl & Excoffier, 2015). Admixture and heterosis are unlikely to be main drivers of invasiveness in Europe, as we found no relationship between population level heterozygosity and any trait other than seed size. Indeed, in most experimentally admixed European and native *A. artemisiifolia* crosses heterosis was absent (Hahn & Rieseberg, 2017). These observations suggest demographic processes had very limited consequences (negative or positive) on the fitness of these introduced populations. However, local adaptation of life history traits such as plant size across broad environmental gradients may mask heterozygosity-fitness correlations. Future studies could address this

question by examining the association in single populations (e.g. Conte *et al.*, 2017).

In plants, reduced seed size is one trait that could aid dispersal and might therefore be expected to evolve during range expansion (Bartle *et al.*, 2013; Huang *et al.*, 2015). Spatial sorting for dispersal traits at the expansion front has been well documented in other invasions, such as the cane toads (Estoup *et al.*, 2004; Phillips *et al.*, 2006). In Europe, spatial sorting for increased dispersal, and therefore smaller seeds, could have occurred at the range edge during expansion northwards. However, this mechanism would only explain the seed size decline in low-latitudinal populations in Australia, where range expansion likely occurred both north- and southward (Palmer & McFadyen, 2012; van Boheemen *et al.*, 2017b). The association between seed size and mean population heterozygosity we identified in all three ranges could be expected when small seeds aid dispersal, as founder effects should also reduce heterozygosity during colonization. Though additional factors likely shape seed size evolution, our findings suggest seed size divergence could represent an important difference in life-history strategies between ranges. Moreover, we observed patterns indicating a relationship between genomic and ecological dynamics potentially linked to range expansion and colonization.

Conclusion

Invasive species often exhibit rapid adaptation despite facing novel selective pressures (Lachmuth *et al.*, 2011; Colautti & Barrett, 2013; Chown *et al.*, 2014; Turner *et al.*, 2014). Moreover, the success of invasives is considered paradoxical as strong demographic changes are predicted to enhance inbreeding and reduce genetic variation and, consequently, evolutionary potential (Allendorf & Lundquist, 2003). We investigated these apparent contradictions in a comprehensive study. We compared the native range with multiple introduced ranges with highly distinct demographic histories, characterized similarities and shifts in climatic niches, tested adaptation in a large number of life-history traits and explored heterozygosity-fitness associations while accounting for non-adaptive population differentiation. We found strong evidence for parallel adaptation in all three ranges. This study therefore emphasizes that although introduction dynamics can affect genetic diversity (Dlugosch & Parker, 2008a) the adaptive potential of those traits might not be constrained to a similar extent (Dlugosch *et al.*, 2015a).

Acknowledgements

We thank J. Stephens and A. Wetherhill for sample collection, S. Bou-Assi, M. Kourtidou, J. Taylor, G. Boinnard, T. Freeman and E. Barnett for greenhouse assistance, S. Bou-Assi and K. Nurkowski for genomic analyses and T. Connallon for manuscript suggestions. A Monash University Dean's International Postgraduate Research Scholarship was provided to LAB, a Monash University Start-up Grant to KAH.

Author contributions

LAB and KH designed the project, with data collection and analyses carried out by LAB and refined by KH. DZA developed and carried out the niche distribution modelling. All authors discussed the results and contributed to the MS writing.

Data accessibility

Sequence data are available at the National Center for Biotechnology Information (NCBI) Sequence Read Archive under Bioproject PRJNA449949.

Supporting information to Chapter 3

Fig. S1. Sampling locations of *Ambrosia artemisiifolia* in the native North America (blue) and introduced European (green) and Australian (red) ranges. Size of dots/numbers indicates number of maternal lines from each sampling site.

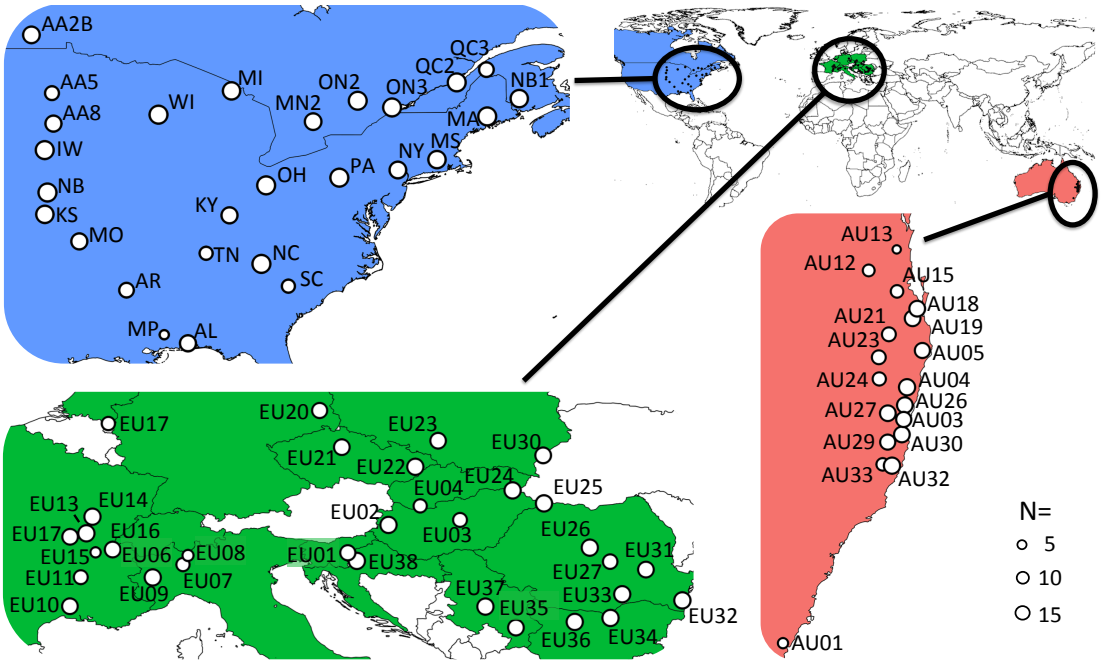


Fig. S2. Principle component analysis for all climatic and geographic variables within the native North American (blue triangles) and invasive European (green squares) and Australian (red circles) *Ambrosia artemisiifolia* ranges, with in brackets the percentage explained by the primary principle components (PC1 and PC2). Abbreviations and PC loadings are in Table S1.

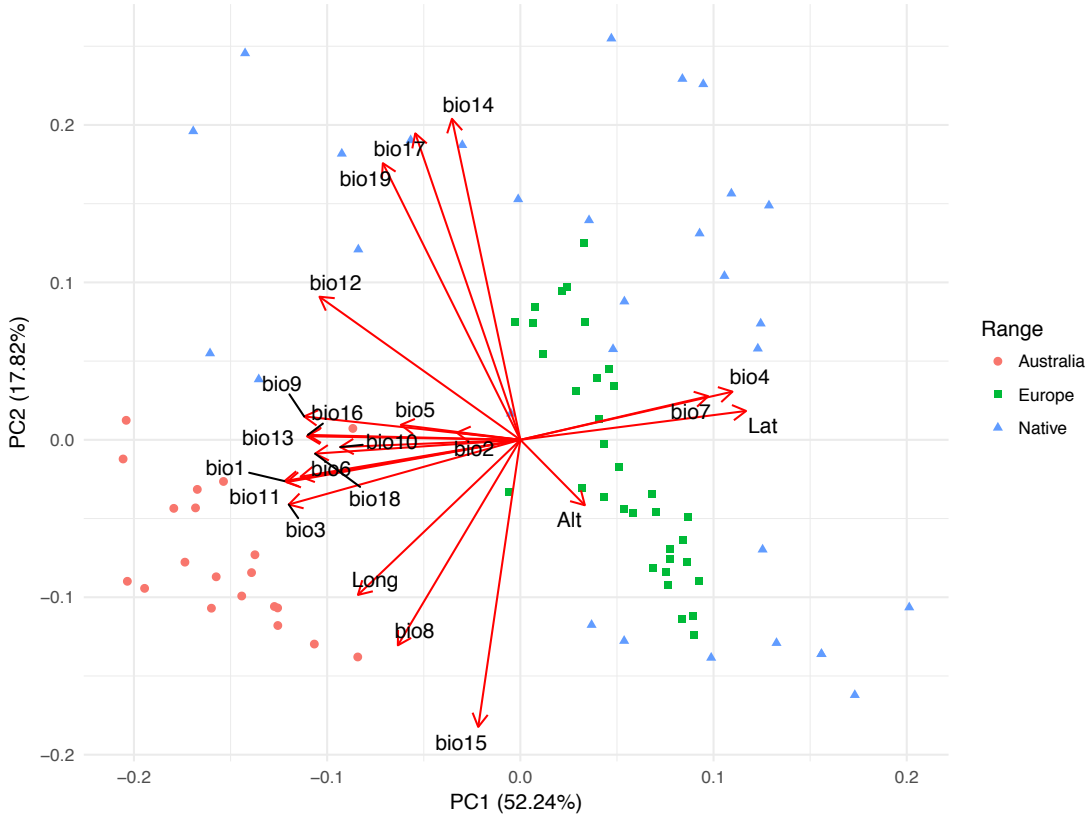


Fig. S3. Principle component analysis for all traits within the native North American (blue triangles) and invasive European (green squares) and Australian (red circles) *Ambrosia artemisiifolia* ranges, with in brackets the percentage explained by the primary principle components (PC1 and PC2) with loadings in Table S2.

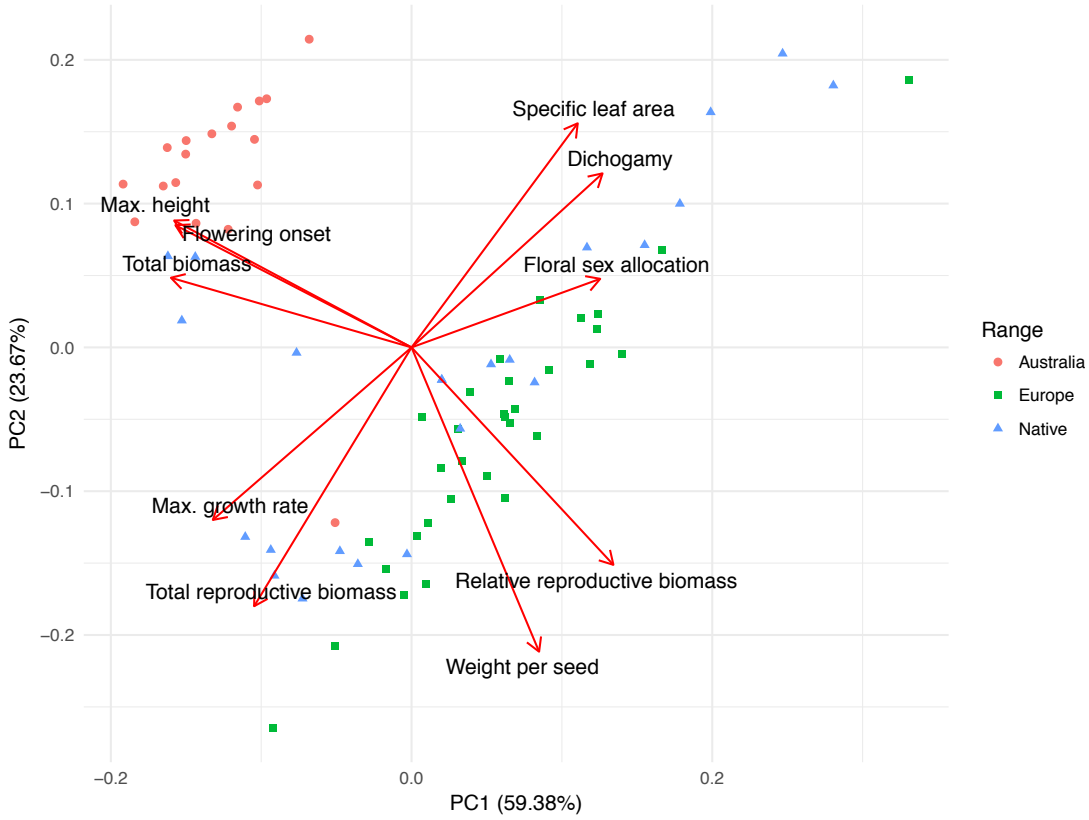


Fig. S4. Graphical representation of performed trait comparisons in the introduced European (EU) and Australian (AU) ranges to the native North American (NA) range at minimum (min) and maximum (max) latitudes observed within each range (when overlapping latitudes were existing).

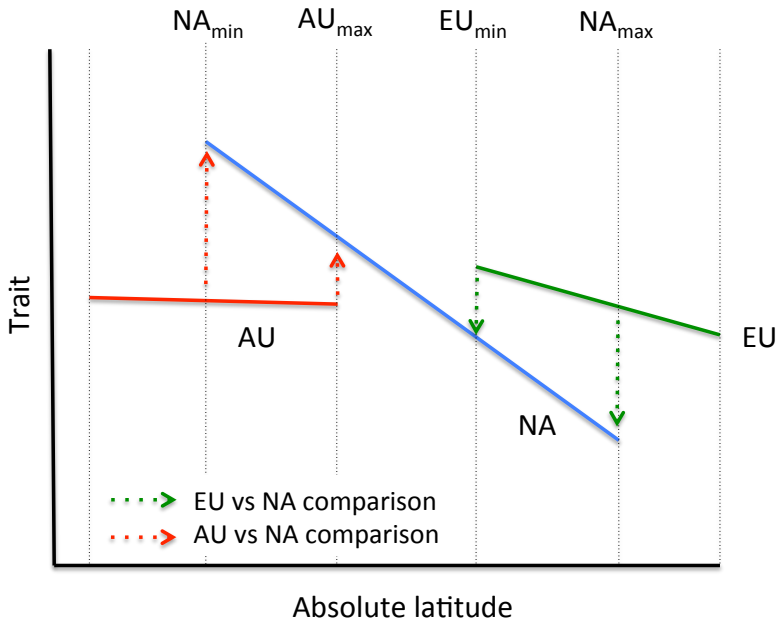


Table S1. List of climatic and geographic variables included with abbreviations.

Abbreviation	Description
bio1	Annual Mean Temperature
bio2	Mean Diurnal Range (Mean of monthly (max temp - min temp))
bio3	Isothermality (BIO2/BIO7) (* 100)
bio4	Temperature Seasonality (standard deviation *100)
bio5	Max Temperature of Warmest Month
bio6	Min Temperature of Coldest Month
bio7	Temperature Annual Range (BIO5-BIO6)
bio8	Mean Temperature of Wettest Quarter
bio9	Mean Temperature of Driest Quarter
bio10	Mean Temperature of Warmest Quarter
bio11	Mean Temperature of Coldest Quarter
bio12	Annual Precipitation
bio13	Precipitation of Wettest Month
bio14	Precipitation of Driest Month
bio15	Precipitation Seasonality (Coefficient of Variation)
bio16	Precipitation of Wettest Quarter
bio17	Precipitation of Driest Quarter
bio18	Precipitation of Warmest Quarter
bio19	Precipitation of Coldest Quarter
Alt	Altitude
Lat	Absolute latitude
Long	Longitude

Table S2. Correlations between climatic and geographic variables (Spearman's ρ^2) and first two principle component loadings (PC) of a principle component analysis (Fig. S2) in native North American, introduced European and introduced Australian *A. artemisiifolia* populations, with abbreviations corresponding to Table S1. Colour intensity increases with the strength of the association. Average correlations (Spearman's ρ^2) between each variable and temperature (bio1-11), precipitation (bio12-19) and all climatic (bio1-19) measures are provided at the bottom of the table.

	bio1	bio2	bio3	bio4	bio5	bio6	bio7	bio8	bio9	bio10	bio11	bio12	bio13	bio14	bio15	bio16	bio17	bio18	bio19	Alt	Lat	Long
bio1		0.061	0.769	0.625	0.396	0.815	0.516	0.287	0.806	0.671	0.898	0.360	0.383	0.021	0.036	0.362	0.058	0.262	0.128	0.103	0.615	0.292
bio2	0.061		0.081	0.034	0.496	0.004	0.100	0.081	0.001	0.362	0.003	0.176	0.317	0.007	0.067	0.307	0.019	0.289	0.024	0.041	0.425	0.084
bio3	0.769	0.081		0.714	0.196	0.772	0.578	0.165	0.712	0.367	0.830	0.422	0.398	0.062	0.015	0.388	0.114	0.318	0.193	0.036	0.535	0.309
bio4	0.625	0.034	0.714		0.008	0.919	0.960	0.135	0.707	0.113	0.854	0.252	0.206	0.023	0.013	0.202	0.050	0.170	0.119	0.084	0.215	0.544
bio5	0.396	0.496	0.196	0.008		0.092	0.001	0.204	0.169	0.869	0.166	0.141	0.239	0.000	0.056	0.215	0.008	0.135	0.021	0.019	0.542	0.005
bio6	0.815	0.004	0.772	0.919	0.092		0.853	0.199	0.846	0.290	0.968	0.295	0.248	0.028	0.013	0.240	0.062	0.177	0.139	0.131	0.346	0.467
bio7	0.516	0.100	0.578	0.960	0.001	0.853		0.132	0.607	0.056	0.760	0.172	0.133	0.007	0.016	0.135	0.022	0.120	0.070	0.089	0.134	0.622
bio8	0.287	0.081	0.165	0.135	0.204	0.199	0.132		0.109	0.341	0.201	0.034	0.142	0.143	0.372	0.147	0.093	0.204	0.031	0.048	0.288	0.204
bio9	0.806	0.001	0.712	0.707	0.169	0.846	0.607	0.109		0.368	0.882	0.291	0.223	0.063	0.000	0.203	0.106	0.117	0.202	0.139	0.357	0.290
bio10	0.671	0.362	0.367	0.113	0.869	0.290	0.056	0.341	0.368		0.396	0.258	0.373	0.002	0.076	0.347	0.020	0.250	0.055	0.061	0.690	0.025
bio11	0.898	0.003	0.830	0.854	0.166	0.968	0.760	0.201	0.882	0.396		0.355	0.315	0.038	0.014	0.305	0.080	0.220	0.161	0.112	0.436	0.401
bio12	0.360	0.176	0.422	0.252	0.141	0.295	0.172	0.034	0.291	0.258	0.355		0.823	0.421	0.005	0.836	0.543	0.735	0.640	0.026	0.511	0.038
bio13	0.383	0.317	0.398	0.206	0.239	0.248	0.133	0.142	0.223	0.373	0.315	0.823		0.111	0.091	0.982	0.200	0.884	0.294	0.002	0.622	0.053
bio14	0.021	0.007	0.062	0.023	0.000	0.028	0.007	0.143	0.063	0.002	0.038	0.421	0.111		0.555	0.118	0.954	0.084	0.833	0.041	0.043	0.017
bio15	0.036	0.067	0.015	0.013	0.056	0.013	0.016	0.372	0.000	0.076	0.014	0.005	0.091	0.555		0.086	0.458	0.116	0.328	0.054	0.057	0.074
bio16	0.362	0.307	0.388	0.202	0.215	0.240	0.135	0.147	0.203	0.347	0.305	0.836	0.982	0.118	0.086		0.206	0.920	0.296	0.002	0.608	0.053
bio17	0.058	0.019	0.114	0.050	0.008	0.062	0.022	0.093	0.106	0.020	0.080	0.543	0.200	0.954	0.458	0.206		0.151	0.918	0.052	0.095	0.007
bio18	0.262	0.289	0.318	0.170	0.135	0.177	0.120	0.204	0.117	0.250	0.220	0.735	0.884	0.084	0.116	0.920	0.151		0.213	0.000	0.512	0.058
bio19	0.128	0.024	0.193	0.119	0.021	0.139	0.070	0.031	0.202	0.055	0.161	0.640	0.294	0.833	0.328	0.296	0.918	0.213		0.148	0.160	0.002
Alt	0.103	0.041	0.036	0.084	0.019	0.131	0.089	0.048	0.139	0.061	0.112	0.026	0.002	0.041	0.054	0.002	0.052	0.000	0.148		0.012	0.089
Lat	0.615	0.425	0.535	0.215	0.542	0.346	0.134	0.288	0.357	0.690	0.436	0.511	0.622	0.043	0.057	0.608	0.095	0.512	0.160	0.012		0.110
Long	0.292	0.084	0.309	0.544	0.005	0.467	0.622	0.204	0.290	0.025	0.401	0.038	0.053	0.017	0.074	0.053	0.007	0.058	0.002	0.089	0.110	
ρ2 with bio1-11	0.584	0.122	0.518	0.507	0.260	0.576	0.456	0.185	0.521	0.383	0.596	0.251	0.270	0.036	0.062	0.259	0.057	0.206	0.104	0.078	0.417	0.295
ρ2 with bio12-19	0.201	0.151	0.239	0.129	0.102	0.150	0.084	0.146	0.151	0.172	0.186	0.572	0.484	0.439	0.234	0.492	0.490	0.443	0.503	0.041	0.326	0.038
ρ2 with bio1-19	0.414	0.135	0.394	0.339	0.189	0.387	0.291	0.168	0.356	0.290	0.414	0.375	0.353	0.193	0.129	0.350	0.226	0.298	0.259	0.063	0.378	0.187

Table S3. Correlations between traits included in this study (Spearman's ρ^2) and associated first two principle components loadings (PC) from PCA (Fig. S3) in native North American, introduced European and introduced Australian *A. artemisiifolia* populations. Traits in bold were removed from MANCOVA due to high correlation ($\rho^2 > 0.6$). Colour intensity increases with the strength of the association.

	Max. height	**Total biomass**	Max. growth rate	Flowering onset	Dichogamy	**Floral sex allocation (female/male)**	Weight per seed	Total reproductive biomass	Relative reproductive biomass	Specific leaf area	**PC1**	**PC2**
Max. height		**0.932**	0.496	**0.842**	0.311	**0.834**	0.203	0.352	0.593	0.274	0.122	0.073
Total biomass	**0.932**		0.453	**0.787**	0.312	**0.796**	0.179	0.437	0.589	0.247	0.124	0.040
Max. growth rate	0.496	0.453		0.319	0.309	0.348	0.004	0.401	0.187	0.503	0.102	0.099
Flowering onset	**0.842**	**0.787**	0.319		0.336	**0.806**	0.277	0.213	0.564	0.141	0.121	0.070
Dichogamy	0.311	0.312	0.309	0.336		0.332	0.006	0.426	0.061	0.237	0.098	0.100
Floral sex allocation (female/male)	**0.834**	**0.796**	0.348	**0.806**	0.332		0.227	0.232	0.540	0.152	0.097	0.039
Weight per seed	0.203	0.179	0.004	0.277	0.006	0.227		0.038	0.394	0.008	0.066	0.175
Total reproductive biomass	0.352	0.437	0.401	0.213	0.426	0.232	0.038		0.028	0.445	0.081	0.149
Relative reproductive biomass	0.593	0.589	0.187	0.564	0.061	0.540	0.394	0.028		0.040	0.104	0.125
Specific leaf area	0.274	0.247	0.503	0.141	0.237	0.152	0.008	0.445	0.040		0.085	0.129

Table S4. Multivariate growth, phenology and reproductive population mean *A. artemisiifolia* traits response to range, latitude, their interactions and latitude^2 for non-correlated (Spearman's ρ^2 <0.6) (a). We reported approximate F-statistic with degrees of freedom as subscript and symbols specifying significance of effect, in addition to Wilk's λ (multivariate F-value).

Range	**Latitude**	**Latitude^2**	**Range: Latitude**
$29.230_{14,120}$*** λ=0.051	$28.377_{7,60}$*** λ=0.232	$8.009_{7,60}$** λ=0.517	$1.796_{14,120}$* λ=0.529

ns p>0.1; # p<0.1; * p<0.05, ** p<0.01; *** p<0.001

Table S5. Trait~latitude associations (adjusted slope of transformed population means +/- SE) within each range (a), slopes significantly different from zero in bold. Pairwise range differences for significant range:latitude interactions (q<0.05, Table 2 main text), with χ^2 test values, degrees of freedom as subscript and symbols specifying Holm adjusted significance of effect (b).

a)	**North America**	**Europe**	**Australia**
Max. height	**-0.804(0.344)**	**-1.116(0.406)**	**-0.663(0.271)**
Total biomass	**0.314(0.186)**	**0.429(0.22)**	**0.287(0.147)**
Floral sex allocation (female/male)	**1.679(0.17)**	**2.407(0.492)**	**2.565(1.04)**
Relative reproductive biomass	**0.006(0.001)**	0.000(0.001)	**0.006(0.002)**

b)	**North America-Europe**	**North America-Australia**	**Europe-Australia**
Max. height	4.758_1#	4.787_1#	0.289_1(ns)
Total biomass	7.010_1*	0.201_1(ns)	2.515_1(ns)
Floral sex allocation (female/male)	14.873_1***	1.64_1(ns)	7.502_1*
Relative reproductive biomass	13.093_1***	0.003_1(ns)	4.912_1#

ns q>0.1; # q<0.1; * q<0.05, ** q<0.01; *** q<0.001

Table S6. Flowering time and sex trait responses to maximum plant height (adjusted slope of transformed individual values +/- SE) within each range (a) and pairwise range differences with χ^2 test values with degrees of freedom as subscript and symbols specifying Holm adjusted significance of effect (b). Slopes significantly different from zero are in bold.

a)

	North America	**Europe**	**Australia**
Flowering onset	**10.682(0.756)**	**8.301(0.974)**	1.356(1.537)
Dichogamy	**-0.078(0.009)**	**-0.107(0.012)**	0.002(0.020)
Floral sex allocation (female/male)	**-0.327(0.0213)**	**-0.329(0.030)**	-0.039(0.046)

b)

	North America – Europe	**North America – Australia**	**Europe - Australia**
Flowering onset	3.73_1#	29.647_1***	14.567_1***
Dichogamy	3.639_1#	13.131_1***	21.415_1***
Floral sex allocation (female/male)	0.030_1(ns)	31.962_1***	27.931_1***

Table S7. Range differences of population mean heterozygosity in response to total biomass, with χ^2 test values, degrees of freedom as subscript and symbols specifying Holm adjusted significance of effect.

North America – Europe	**North America – Australia**	**Europe - Australia**
0.027_1(ns)	5.703_1*	7.330_1*

Methods S1. Detailed methods for common garden set-up and data collection

Common garden set-up

In 2013-2014, we sampled seeds from 30 maternal plants along transects at each sampling location following the protocol of Hodgins & Rieseberg (2011). We collected plants separated by at least 1-2 meters to reduce the possibility of collecting closely related plants due to limited seed and pollen dispersal. From this collection, we randomly selected 4-17 maternal families per sampling location (mean = 11.73) and 20 seeds per family (although for a small number of families fewer seeds were available). We weighed all seeds per family to estimate the average seed weight. We followed stratification procedures of Willemsen (1975) and stratified at 4°C in sand moistened with 1% plant preservative mixture (ppm) for 6 weeks. Stratification induces germination by imitating winter conditions (Willemsen, 1975). Following stratification, we sowed seeds on damp filter paper in Petri dishes with 1% ppm. We placed dishes in a 30°C germination chamber with 12h light/dark cycle and watered twice daily with 1% ppm to maintain damp filter paper.

After 14-15 days from the start of germination (22nd and 23rd of April 2015), we randomly selected two seedlings from each mother and planted these into 30-well kwikpot trays containing 100ml Debco Seed Raising Superior Germinating Mix in a random order. We also transplanted two additional seedlings per maternal line, 1-2 days after this first transplant in case focal seedlings died prior to establishment in the greenhouse. Finally, we transplanted late germinating seedlings (sprouting >14 days after we placed seeds in the germination chamber) 5 days after first transplant. We top-watered all plants by hand twice daily and artificially manipulated daylight following the light cycle at the median latitude over all samples (47.3°N) and adjusted timers fortnightly to accommodate the change in day length. Glasshouse temperature was regulated between 20–30°C. We performed a second transplant of a randomly selected seedling per maternal family on the 21st and 22nd of May, 2015, 1 month after the first transplant, to 100x135mm pots containing 0.7L Debco Seed Raising Superior Germinating Mix and 1.5ml slow-release fertilizer (Osmocote Pro, eight to nine months, 16% N, 4.8% P, 8.3% K, 5.2% S, 1.2% Mg + trace elements). We watered plants twice daily until full senescence (most seeds had ripened). We randomized tray locations with 9 plants each weekly (first 9 weeks) or biweekly, to minimize damage to the larger plants.

Trait measurements

To explore associations between traits and climatic variables, we examined variation in life-history traits including phenology, growth, and vegetative and reproductive allocation (Table 1, main text). We measured start of flowering by monitoring pollen production and emerging female flowering structures through inspection of each plant every 2 days. We calculated dichogamy (the temporal separation of pollen dispersal and emergence of receptive stigma

within an individual plant) as the difference in onset of male to female function flowering. We recorded growth by measuring plant height every 2-4 weeks and once again when plants started showing signs of senescence (first occurrence of end of pollen production or ripe seeds). We calculated maximum growth achieved following Chuine *et al.* (2001), omitting data points after maximum growth was achieved. Once the majority of seeds had ripened, we deconstructed plants for biomass measurements. We separated roots from dirt and washed them under pressure. We placed all components in paper bags and dried these in ovens at 45°C for at least 36h. Before dry weight biomass measures, we dried materials for an additional minimum of 24h to ensure the dry weight was constant at the time of measuring and it was not variable due to humidity in the air or incomplete drying. We separated seeds from other plant matter by pushing dry material through a 0.5mm sieve. We weighed vegetative biomass and roots to the closest 0.1 gram and male flowers and seeds to the closest milligram. Relative reproductive biomass was obtained by dividing total reproductive biomass by total biomass. We calculated relative sex allocation as the dry weight of all seeds over the dry weight of all racemes. We calculated average seed weight over 20 randomly selected seeds (when available) and used this to estimate total seed number per individual. From each plant we scanned one young, fully expanded leaf on day 85 to 88 and calculated leaf area (LA) using ImageJ and the R package LeafArea (Katabuchi, 2015). We dried leaves at 45°C for >7 days and an addition 12h prior to weighing and weighed to the closest milligram. We calculated specific leaf area (SLA) by dividing leaf area by dry leaf weight (mm^2/mg).

Genetic data collection and analyses

Four weeks after second transplant, we harvested a 5-7cm young leaf for DNA extraction and sealed these inside 1.5mL Eppendorf tubes, which were flash-frozen in liquid nitrogen. We stored samples on dry ice before placing them in an -80°C freezer. We extracted DNA from 29-99mg (mean=72.5mg) of leaf tissue of 861 individuals (84 populations) using the Glass Fiber Plate DNA Extraction Protocol (CCDB, Guelph, ON, Canada) and assessed DNA quantity (>8.5 ng/µl) using a QuBit broad-sensitivity DNA quantification system (Invitrogen, Carlsbad, CA, USA). We performed double-digest genotype-by-sequencing library preparation We added 200 ng of high-quality DNA in 7.2 µL water to 2.0 uL CutSmart Buffer 10x, 0.4 uL Pst-1 HF (NEB), 0.4 uL Msp1. We digested samples for 8h at 37°C, 20 minutes at 65°C. To each reaction, we added 2.0 uL 10x CutSmart Buffer, 4.0 uL 10mM ATP, 0.5uL T4 DNA Ligase, 8 uL H2O, 1uL 10mM common adaptor and 5uL 0.6ng/uL barcoded adaptor. We ligated samples for 3h at 22°C and 20 minutes at 65°C. We mixed all samples with 6144 uL Sera-Mag beads (Thermo Fisher). After 15-minute incubation at room temperature, we allotted samples to 7 1.5mL tubes and placed these in Dyna-Mag 2 (Thermo Fisher) magnet for 4 minutes. We removed clear liquid and washed 3 times using 80% EtOH and once with 100% EtOH. We eluted in 150uL

10mM Tris pH 8.0. We amplified 8 reactions each with 3uL of elution and 7.5uL H2O, 12.5 µL KAPA 2x MasterMix, 1uL of 12.5mM each PCR primers f&r. We cycled reactions at 98°C for 1 minute, followed by 20s at 62°C and 30s at 72°C. Following 16 cycles we additionally kept samples at 72°C for 5 minutes. After amplification, we cleaned up 30 µL from each well using the Bioline PCR and Gel kit (Bioline). We eluted the purified product in 30uL buffer. We performed a size selection by running the cleaned PCR product on a 2% agarose gel and removing the 400-600bp fragment. This gel fragment was cleaned up using the Bioline PCR and Gel Kit (Biolin1) and eluted in 20 uL H2O. The eleven ddGBS libraries were paired-end sequenced by the Biodiversity Sequencing Centre at UBC on the Illumina HiSeq 2000 platform, two libraries per lane.

We aligned and filtered raw sequences following van Boheemen *et al.* (2017b). Briefly, SNPs were aligned using BWA-mem (Li & Durbin, 2009) to a draft reference genome for *A. artemisiifolia* (van Boheemen *et al.*, 2017b). We called variants with GATK UnifiedGenotyper (McKenna *et al.*, 2010) and filtered SNPs as follows: quality threshold of a Q-score ≥50; a minimum quality by depth of 2, a maximum Fisher-Strand bias of 60.0, minimum mapping quality rank sum test of -12.5, minimum root mean square mapping quality of 40.0, and a minimum read position rank sum test of -8.0; genotype and variant quality of ≥20, depth of 5-240 and a minor allele frequency of 0.05. We identified a total of 11,598 polymorphic biallelic SNPs with 50% SNP call rate. We calculated individual heterozygosity (H_O) as the proportion of heterozygous loci out of the total number of called genotypes for each individual across 836 control treatment plants and excluded individuals with >80% non-called genotypes (25 out of 861 individuals).

We inferred population genetic structure with STRUCTURE v2.3.4, a Bayesian clustering method that allocates individuals into clusters on the basis of their genotypes (Pritchard *et al.*, 2000). We used this program to calculate population q-scores for the most likely K, used in subsequent analyses to correct for population structure. From the 11,598 SNPs identified, we selected 1024 unlinked SNPs by shuffling the full SNP table and randomly drawing one SNP from each contig. We ran STRUCTURE on these SNPs using the admixture model, correlated allele frequencies, no location prior for the number of clusters (K) ranging from 1 to 10, with 20 independent runs per K. We sampled from a uniform prior for alpha, whilst allowing for alpha to vary between clusters, accounting for unequal sample sizes (Wang, 2017) Each run comprised of a burn-in of 200,000 followed by 1,000,000 iterations. We used log probability and delta K statistic to determine the uppermost clustering level (Evanno *et al.*, 2005). We used CLUMPAK (Kopelman *et al.*, 2015) to test for multimodality (not present) for the most likely K (=2) and visualize the data.

Notes S1. Australian trait divergence

We identified shallower slopes for most latitude-trait associations (e.g. maximum height, total reproductive biomass) in the more recently invaded Australia range (~80 years ago (van Boheemen *et al.*, 2017b)). This may suggest that local optima have not yet been reached in Australia reflecting the recency of the range expansion in the area. Alternatively, the number and latitudinal range of populations was lower in this region, potentially limiting our capacity to detect patterns. Non-linear associations with latitude might also indicate the presence of a local optimum at lower latitudes, which is particularly plausible for traits where a plateauing curve was observed in the native North America at low latitudes (e.g. total biomass, dichogamy).

Chapter 4 – Rapid repeatable phenotypic and genomic adaptation following multiple introductions

Lotte A. van Boheemen[1*], Kathryn A. Hodgins[1]

[1]School of Biological Sciences, Monash University, Clayton, VIC 3800, Australia

*Corresponding author, la.vanboheemen@gmail.com

In preparation for *Molecular Ecology*

Abstract

Uncovering the genomic basis of repeated adaption can provide important insights into the constraints and biases that limit the diversity of genetic forms. Demographic processes such as admixture or bottlenecks affect genetic variation underlying traits experiencing selection, although the impact of these processes on the genetic basis of adaptation remains largely unexamined empirically. We here test repeatability in phenotypes and genotypes along parallel climatic clines within the native North American and introduced European and Australian *Ambrosia artemisiifolia* ranges. To do this, we combined multiple lines of evidence from phenotype-environment associations, F_{ST}-like outlier tests, genotype-environment associations and genotype-phenotype associations. We used 853 individuals grown in common garden from 84 sampling locations, targeting 19 phenotypes, >83k SNPs and 22 environmental variables. We found that 25-41% of loci with adaptive signatures were repeated among ranges, despite alternative demographic histories shaping genetic variation and genetic associations. Our results suggest major adaptive changes can occur on short timescales, with seemingly minimum constraints due to demographic changes linked to introduction. These patterns reveal some predictability of evolutionary change during range expansion, key in a world facing ongoing climate change, and rapid invasive spread.

Introduction

Reoccurrence of similar phenotypic patterns suggests stochastic processes do not solely govern evolution and that population and species divergence is to some extent predictable (Wood *et al.*, 2005; Arendt & Reznick, 2008; Elmer & Meyer, 2011; Lee *et al.*, 2014). But if repeated patterns at the phenotypic level occur, does this also extend to the genetic level? The repeatability of adaptation at multiple levels of biological organization (phenotype, molecular pathway, gene and genetic variant) and the drivers of this repeatability are major questions in evolutionary biology (Smith & Rausher, 2011; Conte *et al.*, 2012; Martin & Orgogozo, 2013). Studying the reappearance of the same adaptive solutions on a genomic and phenotypic level (Rausher & Delph, 2015) to similar selective factors (e.g. parallel environmental clines) could give insight in the predictability of evolution (Stern & Orgogozo, 2008; Stern & Orgogozo, 2009; Lässig *et al.*, 2017), the factors limiting evolutionary outcomes (Christin *et al.*, 2010; Connallon & Hall, 2018; Yeaman *et al.*, 2018), and the mechanisms by which populations adapt.

Rapid evolutionary events are often observed during the spread of invasive species and can contribute to their success (Lee, 2002; Maron *et al.*, 2004; Lawson Handley *et al.*, 2011; Bock *et al.*, 2015; Colautti & Lau, 2015). For instance, the reemergence of native phenotypic clines in ecologically important traits, such as flowering time, has been observed in many alien populations (e.g. Maron *et al.*, 2004; Montague *et al.*, 2008; Colautti & Barrett, 2013; van

Boheemen *et al.*, 2018a). These parallel patterns can be accompanied by the repeated divergence of ecologically important genes (Bock *et al.*, 2015; Colautti & Lau, 2015), but documented evidence of repeatability at the genetic level following introduction to novel ranges has been limited (Hamilton *et al.*, 2015; Gould & Stinchcombe, 2017; Marques *et al.*, 2018). Such information can be harnessed to identify the genetic basis (Dlugosch *et al.*, 2015a; van Kleunen *et al.*, 2018) and predictability of adaptation during range expansion.

In closely related taxa, reoccurrence of the same genetic and trait patterns in response to similar selection pressures is considered a likely response due to shared standing variation (Bollback & Huelsenbeck, 2009; Conte *et al.*, 2012; Bailey *et al.*, 2015; Holliday *et al.*, 2016; Bailey *et al.*, 2017). Standing genetic variation is predicted to contribute more to rapid adaptation than new mutations due to the immediate availability and higher fixation probability of beneficial alleles (Hermisson & Pennings, 2005; Barrett & Schluter, 2008; Prentis *et al.*, 2008; Rockman, 2012; Lee & Coop, 2017; MacPherson & Nuismer, 2017). However, demographic changes charactering introduction events can have important impacts on this standing genetic variation, thus constraining the basis for selection to act on (Lee, 2002; Facon *et al.*, 2006; Prentis *et al.*, 2008; Rius & Darling, 2014; Estoup *et al.*, 2016).

Bottlenecks in small founding populations can increase genetic load (Frankham, 1995; Blackburn *et al.*, 2015) and restrict opportunities for adaptive evolution by reducing genetic diversity (Wright, 1931; Dlugosch & Parker, 2008a; Bock *et al.*, 2015). In contrast, such bottlenecks could result in the conversion of non-additive (e.g. epistatic) to additive genetic variance (Neiman & Linksvayer, 2006) or facilitate the purging of genetic load by exposing deleterious recessive mutations (Glémin, 2003; Facon *et al.*, 2011; Marchini *et al.*, 2016). Admixture, the mixing of genotypes from differentiated genetic backgrounds (Dlugosch & Parker, 2008a; Uller & Leimu, 2011) could increase adaptive potential of introduced populations through higher genetic variance within populations, heterosis and novel or transgressive phenotypes (Roman & Darling, 2007; Prentis *et al.*, 2008; Verhoeven *et al.*, 2011; Rius & Darling, 2014; Bock *et al.*, 2015). Comparison of distinct introductions with different demographic histories will allow us to dissect these opposing effects of historical contingency on evolvability.

Suppression of recombination among loci contributing to local adaptation is predicted during local adaptation with gene flow (Lenormand & Otto, 2000; Yeaman & Whitlock, 2011; Ortiz-Barrientos *et al.*, 2016). Such linkage can elevate the probability of parallel genomic patterns underlying recurrent bouts of adaptation, especially in recently diverged lineages sharing genomic context (Renaut *et al.*, 2014; Holliday *et al.*, 2016; Storz, 2016; Lässig *et al.*, 2017; Yeaman *et al.*, 2018). However, it is difficult to ascertain if multiple loci or a single adaptive locus is the cause of signatures of selection in a genomic region due to hitchhiking (Yeaman, 2013; Rougemont *et al.*, 2017). In addition to reduced recombination, selection can

also create patterns of LD among unlinked loci if alternative alleles are favored across multiple independent loci (Ortiz-Barrientos *et al.*, 2016). Admixture and strong drift are expected to cause strong patterns of LD among loci (Lande, 1980; Hudson, 1985; Keller & Taylor, 2010), potentially interfering with one another if associated linked alleles have opposing selection coefficients (Fisher, 1930; Muller, 1964; Hill & Robertson, 1968; Felsenstein, 1974; McVean & Charlesworth, 2000; Slatkin, 2008). Moreover, such demographic processes can reduce the capacity to identify specific loci involved in local adaptation in recently invaded regions. Empirical investigations of these demographic changes on tests of selection (Hodgins *et al.*, 2018), patterns of LD (Ahrens *et al.*, 2018) or repeatability (Marques *et al.*, 2018) are however lacking.

Ambrosia artemisiifolia, an invasive annual plant, provides an excellent system to study the repeatability of rapid adaptation. This native North American weed has successfully established globally (Oswalt & Marshall, 2008), including contemporary introductions to Europe (~160 years ago Chauvel *et al.*, 2006) and Australia (~80 years ago; Palmer & McFadyen, 2012; van Boheemen *et al.*, 2017b). Post-introduction admixture of these distinct native units contributed to the European invasion, followed by a subsequent single, bottlenecked introduction into Australia (van Boheemen *et al.*, 2017b). Repeated latitudinal trait clines have been found in the native range and the two introduced ranges included in the current study. The repeated trait clines in the introduced ranges are unlikely to be due to climate matching (Maron *et al.*, 2004) and have emerged despite the short time since introduction and distinct invasion histories (van Boheemen *et al.*, 2018a).

To test the repeatability of adaptation at multiple hierarchical levels of organization, we here identify phenotypic and genomic variation along parallel climatic clines within the native North American and introduced European and Australian ranges of *Ambrosia artemisiifolia.* We raised 853 individuals in a common garden, collected from 84 sampling locations across broad geographic scales. We used >83k SNPs, 19 phenotypes and 22 environmental variables to explore repeatable signatures of selection from the environment to phenotype to genotype. We estimate the strength and direction of phenotype-environment associations (PEA), 1genomic divergence (F_{ST}-like X^TX), associations between allele frequencies and environment (GEA) and associations between genotypes and phenotypes (GPA). By combining these methods, we can more effectively identify adaptive signatures (Sork *et al.*, 2013; Hoban *et al.*, 2016) and test the genetic basis of parallel evolution (Conte *et al.*, 2015). This study provides insight into the evolutionary mechanisms underlying invasion and contemporary evolution. Moreover, understanding the adaptive potential of invaders is important in making predictive models and adopting strategies to better manage alien species (Estoup & Guillemaud, 2010; Bock *et al.*, 2015; Cristescu, 2015; Dlugosch *et al.*, 2015a).

Methods

Study species

Ambrosia artemisiifolia is a diploid monoecious annual, native to North America, introduced globally and is commonly found in disturbed habitats (Oswalt & Marshall, 2008). The first recorded introduction of *A. artemisiifolia* was in Europe around 1850 and later introductions have been tied to imports during the two World Wars (Chauvel *et al.*, 2006). In Australia, *A. artemisiifolia* was first introduced at the end of the 19th century, and has become increasingly abundant in southeastern Queensland and New South Wales since the mid-1900s (Palmer & McFadyen, 2012). Genomic reconstruction of the introduction history in these ranges indicates high levels of admixture in Europe sourced from distinct native sources, with Australian populations in turn sourced from the European invaded range whilst experiencing genetic bottlenecks (van Boheemen *et al.*, 2017b). Multiple studies of *A. artemisiifolia* populations show genetic differentiation of important life history phenotypes in the multiple introduced ranges, which parallel patterns in the native range (Chun *et al.*, 2011; Hodgins & Rieseberg, 2011; Leiblein-Wild & Tackenberg, 2014; Li *et al.*, 2014; van Boheemen *et al.*, 2018a). Such repeated latitudinal clines have also been identified for the European and Australian populations included in the current study, with evident phenotype divergence (van Boheemen *et al.*, 2018a).

Data collection

We collected seed samples from North American (300 individuals from 30 locations), European (369 individuals from 34 locations) and Australian (183 individuals from 20 locations) ranges between 2013-2014. We measured 19 phenotypes (Supporting Information, Table S1) of common garden raised seedlings according to van Boheemen *et al.* (2018a) and van Boheemen *et al.* (2018b). For environment associations, we included 19 bioclimatic and three geographic variables (latitude, longitude and altitude) (Table S2) for our sampling sites (van Boheemen *et al.*, 2018a) and analyzed genetic data as described by these authors. In brief, we aligned raw sequences using BWA-mem (Li & Durbin, 2009) to a draft reference genome for *A. artemisiifolia* (van Boheemen *et al.*, 2017b) and called variants with GATK HaplotypeCaller and filtered SNPs using GATK hard-filtering recommendations (McKenna *et al.*, 2010; Van der Auwera *et al.*, 2013), with a depth of 5-240 and a minor allele frequency of 0.05. As we were interested in detecting signatures in as much of the genome as possible in the genome scan, we included low call rate SNPs (10%) even if they might have low statistical power for the Bayenv2 analysis (e.g. Yeaman *et al.*, 2016; Lotterhos *et al.*, 2018), resulting in 83,559 SNPs. Statistical power was later increased by combining multiple tests (see below). For the linkage disequilibrium (LD) analyses and genotype-phenotype associations (GPA), we removed all SNPs with good genotypes in <50% of individuals, as the latter analysis proved sensitive to missing data, resulting in 11,497 SNPs. We identified 2854 scaffolds with SNPs,

representing 17% of the reference genome scaffolds. Median length of scaffolds with SNPs was 291,470 bases (ranging from 1000 to 3,127,000).

Methods to detect signatures of selection

All data analyses were performed in R v.3.4.3 (R Core Team, 2017) unless otherwise stated. To explore putative locally adapted phenotypes and associated selective environments, we tested for the occurrence and correlation strength of phenotype-environment associations (PEA) within ranges. We calculated Spearman's ρ between phenotypes and environmental variables using population phenotype means, resulting in a total of 418 PEA combinations. We focused on those PEA showing the strongest putative signatures of selection in each range (hereafter strongest PEA), identified as the top 5% Spearman's ρ per range. To explore the relationship amongst phenotypes or environmental variables, we performed principle component analyses (PCAs).

To detect putatively adaptive SNPs, we applied Bayenv2 (Günther & Coop, 2013), a powerful test to identify outliers in SNP data (Lotterhos & Whitlock, 2014). This model allows for non-independence in gene frequencies between samples, arising through population structure, by calculating a covariance matrix of allele frequencies. The X^TX functionality of the program calculates standardized allele frequencies, used to locate SNPs that strongly deviate from neutral population genetic structure. The X^TX score is a direct F_{ST} analog, but accounts for variance-covariance between populations (Günther & Coop, 2013), outperforming other outlier tests by producing lower false-positive rates (Lotterhos & Whitlock, 2014).

We conducted genotype environmental association tests (GEA) within Bayenv2, identifying SNPs showing a strong correlation with the standardized environmental data as indicated by Bayes Factors (BF). GEA are suggested to be more powerful in detecting signatures of selection than F_{ST} outlier tests (De Mita *et al.*, 2013), although the false-positive rates associated with each test are shown to greatly depend on demographic history (Lotterhos & Whitlock, 2015). Here, we standardized environmental factors by subtracting the mean and dividing by the standard deviation (Günther & Coop, 2013) per range and within the global dataset. We performed X^TX and GEA analyses on 83,624 SNPs (10% SNP call rate) on samples within each range, correcting for population structure by averaging across 3 independent covariance matrices using 2877 unlinked SNPs (from different scaffolds). We ran simulations for $5x10^5$ iterations and averaged 3 independent runs per dataset for each analysis.

Within each method, we targeted putative adaptive phenotypes or SNPs by exploring divergence of each test-value compared to their general distribution. Due to the absence of a theoretical expected null-distribution for X^TX values (Günther & Coop, 2013), we defined an empirical cut-off for this analysis by identifying the top 1% ranked SNPs within each range as outliers. To align methods, we similarly defined the top 1% outlier SNPs in GEA by ranking

SNPs across environmental variables based on Bayes Factors (GEA). By ranking across multiple environmental variables, a higher number of SNP outliers in GEA were identified compared to X^TX.

Genotype-phenotype associations

To determine the putative function of loci showing signatures of selection, we performed genotype-phenotype associations (GPA) using the mixed linear model association analysis in GCTA (Yang *et al.*, 2011), standardizing phenotypes by subtracting the mean and dividing by the standard deviation. We increased power of this analysis by reducing missing genomic data to a ≥50% SNP call rate. We accounted for further population structure using a genetic relatedness matrix (GRM) as a random effect and 10 principle components calculated from this GRM as fixed effects. The GRM was calculated from SNPs pruned for LD with a step of 3 SNPs, window of 10 SNPs and an r^2 cut-off of 0.1 using indep-pairwise in plink v1.07 (Purcell *et al.*, 2007). Sample sizes per range were too small and we further increased the power of this analyses by pooling all samples. However, these steps might not be sufficiently effective in reducing false positives when population structure is strong (Yang *et al.*, 2014). Therefore, we further increased the accuracy of our approach by excluding Australian samples from these tests due to their extreme population genetic structure compared to the other ranges (van Boheemen *et al.*, 2017b) and calculated associations across pooled North American and European samples. We defined the top 1% SNPs by ranking across phenotypes based on p-values. Fewer SNPs in the GPA analysis result from a more stringent SNP call rate for this method (for details, see Supporting information).

Sliding window analysis

To account for physical linkage between SNPs and reduce the false positive detection of outliers, we used a window approach to identify genomic regions showing signatures of selection. We determined linkage disequilibrium by computing pairwise correlations of SNPs with a call rate of ≥50% based on shared occurrence within individuals. Overall breakdown of linkage disequilibrium between all SNPs was determined by evaluating the loss of allelic correlations over genomic distance between SNPs (Results). Accordingly, window size was set to 1000 bases. To account for variation in the number of SNPs between datasets and methods, the first window within each scaffold started at the first SNP within the global ≥10% call rate dataset. For each window within each range and the global dataset, we calculated if the number of outlier SNPs was higher than could be expected from a binomial distribution, where the expected frequency of outlier SNPs (a) per number of SNPs (n) equals $\Sigma_i \frac{a_i}{n_i}$ (Yeaman *et al.*, 2016). Windows falling outside the 99.9th binomial quantile and containing >1 outlier SNP were identified as outlier windows.

Repeatability of signatures of selection between ranges

We tested for repeatable patterns among geographic ranges in PEA, X^TX divergence and GEA using all test-values (outliers and non-outliers). We calculated Pearson's R of test values between range pairs for phenotype-environment pairs (PEA), SNPs or windows (X^TX and GEA), averaging test values per window. To determine whether repeatability depended on the strength of the association (PEA and GEA) or divergence (X^TX), we grouped test values in 5% quantile bins. We counted the number of times phenotype-environment pairs (PEA), SNPs or windows (X^TX and GEA) were allocated to the same bin among range pairs. We compared shared bin allocation to a binomial expectation (*E*) of random distribution among bins, where *P*=bin size^2 and *E*=*P**number of shared phenotype-environment pairs, SNPs or windows (following Renaut *et al.*, 2014). We highlighted the top 5% shared PEA environments in GEA and top 5% shared PEA phenotypes in GPA (see Results), although we performed GEA and GPA on all environments and phenotypes (Results and Supporting Information).

We further explored repeatable genomic signatures of adaptation between ranges with the null-W method (Yeaman *et al.*, 2016). This method allows for identification of genomic regions with increased divergence against the genetic background. Such a test is favorable when LD increased divergence of SNPs in tight linkage with causal SNPs but did not raise their test values (X^TX or BF) enough for a window to be classified as an outlier according to the binomial test. We here targeted windows identified as outlier windows in other ranges and randomly sampled 10,000 'background' SNPs from non-outlier windows per focal range. We calculated null distributions by comparing the test values of the background SNPs to SNP test-values in non-outlier windows and computed W using a Wilcoxon signed-rank test. We standardized W to Z to allow for comparisons: $Z = \frac{2W - n_1 n_2}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{3}}}$. Here, n_1 is 10,000 (number of background SNPs) and n_2 is the number of SNPs in each window. For each window we calculated empirical p-values by counting the number of times their respective z-score exceeded z-scores in the null distribution, dividing by the number of windows in the null-distribution.

To evaluate the repeatability of the specific regions under selection, we tested if the number of windows identified as outliers within range pairs was higher than could be expected. As the number of outliers is not equal within each range, we calculated expectations under the hypergeometric distribution (e.g. Holliday *et al.*, 2016; Yeaman *et al.*, 2018) using the phyper and qhyper function in the stats package (R Core Team, 2017). Here, the probability (*P*) of a given number of outlier windows with the same identity among range-pairs (*x*) equals $P(x) = \frac{\binom{m}{x}\binom{n}{k-x}}{\binom{N}{k}}$, where *m* is the number of outlier windows in range 1, *k* is the number of outlier windows in range 2 and *n* is (the number of total windows shared between ranges-*m*). We

calculated the 99.9th hypergeometric quantiles and computed p-values as 1-*P*, followed by FDR-correction for multiple pairwise range comparisons (Benjamini & Hochberg, 1995). First, we compared shared outlier windows identified in either X^TX or GEA test, followed by a comparison of outlier windows identified in at both the X^TX and GEA tests.

Patterns of linkage disequilibrium

To explore the linkage disequilibrium (LD) between putative adaptive and non-adaptive loci, we tested LD on a local and genome-wide level. For all analyses, we log- and square root transformed variables where appropriate. Although LD is usually depicted as r^2 (Remington *et al.*, 2001), we here had to apply square-root transformations to the absolute Pearson's r to increase normality of the data and satisfy assumptions of the statistical models applied. We tested LD as correlation between SNPs (described above). For these analyses, outlier windows were detected using the binomial method, and only included if identified in both X^TX and GEA outlier analyses. Non-outlier windows were defined as windows not identified in either outlier test. To reduce physical linkage from local and global LD analyses, only one window of each type per scaffold was selected at random. We tested models using a type 3 ANOVA. Since we were interested in within-range effects and pairwise range differences between outlier windows or between non-outlier windows, we dissected the highest order significant interactions in post-hoc tests by only comparing these meaningful interactions using F-tests with Holm p-value correction within the phia package (De Rosario-Martinez, 2015).

We first explored patterns of local LD decay between SNPs within windows and investigated range differences in these trends. We tested LD as a response to distance between SNPs (maximum of 1000 bp due to window size), range, type of window (outlier or not) and all their interactions. To explore range difference in genome-wide LD, we tested whether LD among outlier loci was higher compared to LD among non-outlier loci. To reduce the effect of local LD, we selected a single SNP at random within each window (where we previously selected one window per scaffold). We improved the accuracy of our outlier categorization by selecting SNPs identified as top 1% outlier in both X^TX and GEA tests. Within non-outlier windows, we selected a SNP at random, not identified as top 1% outlier SNP in either X^TX or GEA. We tested LD as response to range, window type and their interaction.

Results

Phenotypic repeatable signatures of selection

We identified a total of 20 PEA with strong associations (top 5% based on ranked Spearman's ρ) in at least two ranges, the top 5% shared PEA environments and phenotypes (Table 1, Fig. S1, Supporting Information). Notable strong repeatable correlations were observed between latitude and temperature measures with size, growth, phenology and sex function allocation

(Table 1, Fig. S1), in line with previous reported latitude-trait clines in *A. artemisiifolia* using a subset of the current data (van Boheemen *et al.*, 2018a). Longitude was associated with various trait categories in Europe and Australia, but not in North America (Table 1). In contrast, winter temperature was strongly correlated with sex function allocation in North America and Australia. In addition to these patterns, we identified a single PEA with a similar strong, but opposite, association within Europe and Australia, where total seed weight was positively associated with mean summer temperature in Europe, but mass decreased with this temperature measure in Australia. This pattern is closely associated with the curved relationship between latitude and average seed weight described previously (van Boheemen *et al.*, 2018a). In addition to these repeatable shifts of top 5% PEA, some weaker phenotype-environment associations also displayed higher than expected repeatability among ranges (Fig. S1&2). However, when considering all PEA together, repeatability among range pairs was generally low or absent (Pearson's R=0.048-0.246, Table S3).

The top 5% shared PEA environments and phenotypes were a well-distributed subset of all the phenotypes and environmental variables tested (Fig. S3 & S4, Supporting Information). Australian phenotypes were quite distinct from the other ranges (Fig. S3), whereas sampling locations for each range were clustered reflecting differences in the PC1 axis which was correlated mainly with growth and reproductive traits (Fig. S3, Table S1). Additionally, sampling locations within each range were clustered along variation in PC1, mainly reflecting differences in latitude and temperature (Fig. S4, Table S2).

Table 1. Repeatability of the strongest (top 5% ranked Spearman's ρ) PEA amongst *Ambrosia artemisiifolia* ranges (N=native North America; E=introduced Europe; A=introduced Australian). Lat=absolute latitude; Lon=longitude; Tmean=annual mean temperature; Tdiur=temperature diurnal range; Tmax=maximum temperature in the warmest month; TQwarm=mean temperature of warmest quarter; TQcold=mean temperature of coldest quarter (Supporting information Table S1&S2). These top 5% shared PEA's are highlighted in further GEA and GPA analyses.

	Lat	**Lon**	**Tmean**	**Tdiur**	**Tmax**	**TQwarm**	**TQcold**
Maximum height	N,E	E,A				N,E	
Total biomass	N,E						
No. branches	N,E						
Day half growth	N,E		N,E		N,E	N,E	
Flowering onset	N,E		N,E,A		N,E,A	N,E,A	
Male function biomass	N,E,A	E,A					
Dichogamy				N,E			
Sex function allocation		E,A	N,A			N,A	N,A

Genomic signatures of selection

We identified a large number of outlier windows (windows with a higher number of outliers SNPs than binomial expectation), with 0.9-1.2% for the X^TX, 5.4-7.6% for GEA and 2.6% for GPA of all windows being identified as outliers (Table S4). We observed a lower number of

SNPs and windows in Australia compared to the other ranges, which aligns with reduced genetic variation within this range (van Boheemen *et al.*, 2017b) as well as lower sample sizes. Further differences in numbers of outlier SNPs and windows are due to outlier classification among environmental variables (GEA) or phenotypes (GPA) or due to a more stringent SNP-call rate (GPA).

Genomic repeatable signatures of selection

We investigated patterns of repeatability on a SNP-by-SNP and window-by-window basis, starting with genomic differentiation compared to the neutral background. We found high repeatability of SNP differentiation, where X^TX values were strongly correlated (Pearson's R=0.434-0.501) in pairwise range comparisons (Table S3). Repeatability at the level of 1000 bp windows was slightly higher (Pearson's R=0.474-0.585; Table S3). For all pairwise range comparisons, repeatability was higher in North America and Europe compared to these ranges and Australia. To identify whether repeatability depended on the level of differentiation, we placed SNPs and windows in 5% quantile bins based on ranked (mean) X^TX values. Strong- and weakly-differentiated SNPs and windows were disproportionately often placed in the same bin among range pairs, with up to 4.7 and 6.1 times more bin sharing of the top 5% differentiated SNPs and windows than could be expected from binomial distribution (Fig. 2). We discarded patterns of balancing selection (low differentiated SNPs and windows) from further analyses as these are known to be sensitive to high false positive rates (Excoffier *et al.*, 2009b; Lotterhos & Whitlock, 2014).

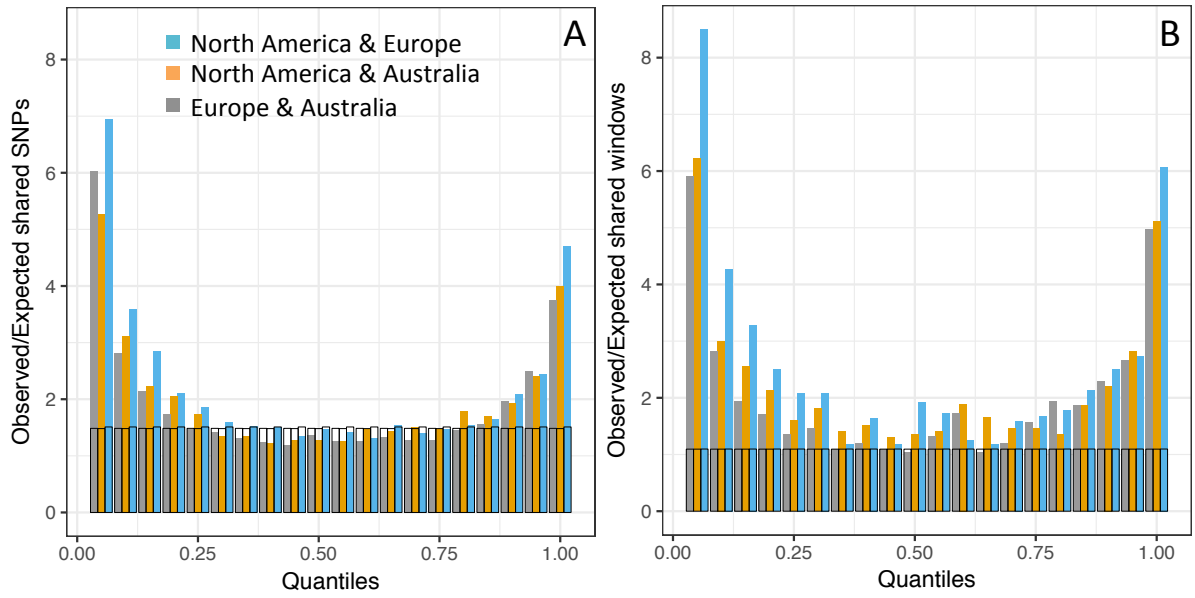


Figure 2. Observed/expected number of SNPs (A) and 1000 bp windows (B) shared between range pairs in the same 5% quantile bin based on X^TX values with 95% confidence intervals of the binomial expectation under random distribution of test values illustrated with black lines.

We explored the number of genomic regions associated with top 5% shared PEA environmental variables using GEA. Compared to the whole pool of environmental variables, the number of outlier windows identified for the top 5% shared PEA environmental variables was relatively low among all ranges (Fig. 3). Instead, we repeatedly found various precipitation measures to have the largest number of outlier windows within each range, with some of these shared among ranges. Furthermore, for each top 5% shared PEA environmental variable, we identified a relative lower number of outlier windows (compared to the outlier windows identified for other PEA variables) in North America compared to Europe or Australia. , we did not reveal higher sharing of outlier binomial windows among range pairs than could be expected by chance alone (Fig. 4A, Fig S5&6). A lack of repeatability can result from dilution of the selective signal in one range due to differences in power between ranges (see Discussion). To overcome this issue, we explored shifts of test values within windows compared to the genomic background, targeting those windows identified as outliers in one of the other ranges using the null-W method. When we included such windows in our pairwise range comparisons, we did find a higher number of shared windows compared to random chance for most top 5% shared PEA environmental variables (Fig. 4B) and other PEA environmental variables (Fig. S6, Supporting Information).

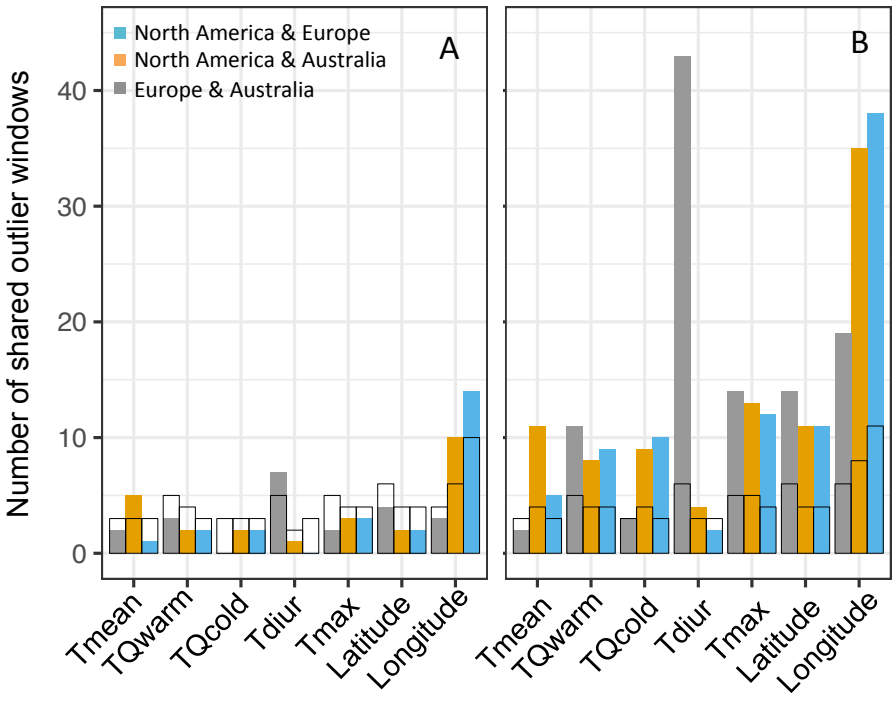


Figure 4. Number of 1000 bp outlier windows as identified with binomial (A) and null-W tests (B) shared among *A. artemisiifolia* range pairs with 99.9th hypergeometric quantiles illustrated with black lines.

Genomic and phenotypic repeatable signatures of selection

To reduce false-positive rates that could be associated with individual outlier tests, we explored range sharing of outlier windows identified in the outlier analyses (X^TX and GEA). For all

pairwise range comparisons, the sharing of outlier windows was significantly higher (q≤0.001) than could be expected under a random hypergeometric distribution. This was true when focusing on top 5% shared PEA environmental variables (Fig. 5), or when investigating overlap for all environmental variables (Fig. S4, Supporting Information). In comparisons including top 5% shared PEA environmental variables, 240 windows showed multiple signatures of selection (X^TX and GEA), 50 (24.5%) were shared amongst ranges, 9 (4.4%) of which were found in all ranges. Six of these windows were also identified as binomial GPA outlier windows, in phenotype associations with growth (outlier in North America only), sex allocation (Europe only) and dichogamy (North America and Australia) traits (Fig. 5). When testing all environmental variables, sharing of outlier windows (X^TX and GEA) amongst ranges reached 40.7% (Fig. S4). Ten of these outlier windows showed significant binomial GPA, with additional associations to relative reproductive allocation, seed size and trichome density. These results support genomic and environmental repeatability amongst ranges when combining multiple lines of evidence supporting local adaptation.

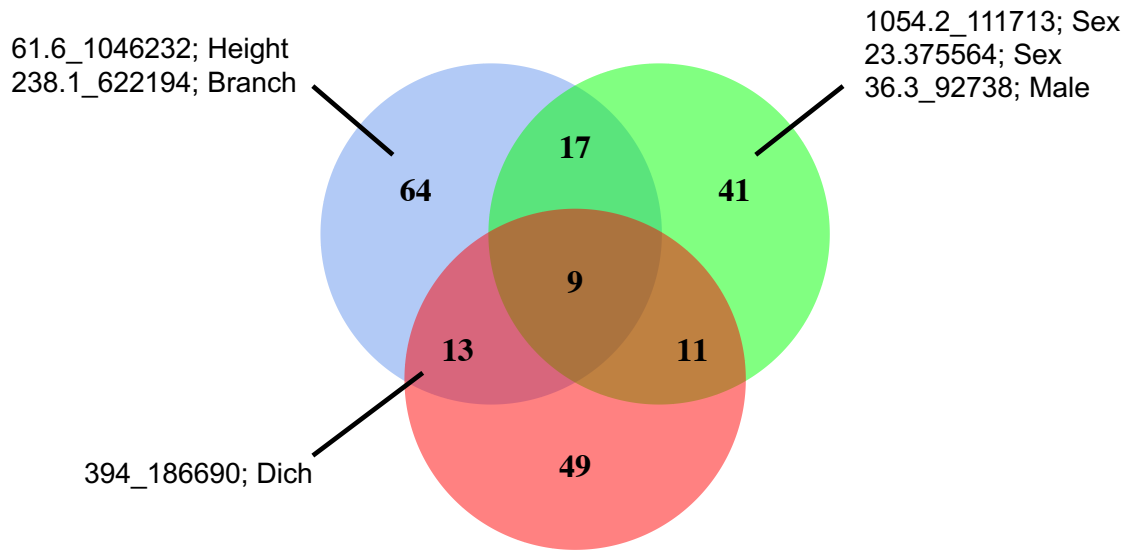


Figure 5. Number of *Ambrosia artemisiifolia* outlier windows (binomial and null-W tests, identified in both X^TX and GEA) within each or multiple ranges. In each diagram: blue/top left=native North America; green/top right=introduced Europe; red/bottom=introduced Australia. An outlier window had to be identified in both X^TX (F_{ST}-like outlier test) and GEA (genotype-environment associations) where windows identified as outliers for multiple environments were counted once. Binomial outlier window ID and associated phenotypes are provided when identified in GPA (phenotype abbreviations correspond to Table S1, Supporting Information; Height=Max. plant height; Branch=Number of branches; Sex=Sex function allocation; Male=Weight of male sex function; Dich=Dichogamy). We only considered top 5% shared PEA phenotypes and environmental variables. The overlap in outlier windows in each pairwise range comparison is significantly higher (q≤0.001) than could be expected under a random hypergeometric distribution.

Patterns of LD

We investigated range differences in local patterns in LD decay between SNPs and tested whether such patterns differed between outlier and non-outlier windows. We found a significant

three-way interaction of distance (bp) between SNPs within a window, range and window type (F_1=7.830, p<0.001). Dissection of specific interaction contrasts revealed that LD decayed over shorter distances in non-outlier compared to outlier windows in the native North America. Such differences were however not apparent in European or Australian individuals (Fig. 6, Table 3). In contrast, correlation between SNPs in outlier windows was higher in the native compared to the Australian samples (Fig. 6, Table 3).

Table 3. Local LD decay in *Ambrosia artemisiifolia*, within outlier and non-outlier windows among pairwise range comparisons (a) or within ranges (b) (N=North America, E=Europe, A=Australia), with adjusted slopes (standard error) $*10^{-5}$)(c). For a &b, we reported F-test values, degrees of freedom in subscript and Holm-corrected p-values (q). *a.*

Pairwise range comparisons

Non-outlier			Outlier		
N-E	N-A	E-A	N-E	N-A	E-A
F_1=0.000, q=1.000	F_1=0.329, q=1.000	F_1=0.313, q=1.000	F_1=5.446, q=0.137	F_1=16.587, q<0.001	F_1=4.060, q=0.220

b.

Non-outlier versus outlier within ranges

N	E	A
F_1=12.927, q=0.003	F_1=0.051, q=1.000	F_1=4.801, q=0.171

c.

10^-5	**Non-outlier**	**Outlier**
North America	-14.758 (0.002)	-10.2 (0.006)
Europe	-14.765 (0.002)	-14.427 (0.009)
Australia	-15.034 (0.002)	-19.616 (0.015)

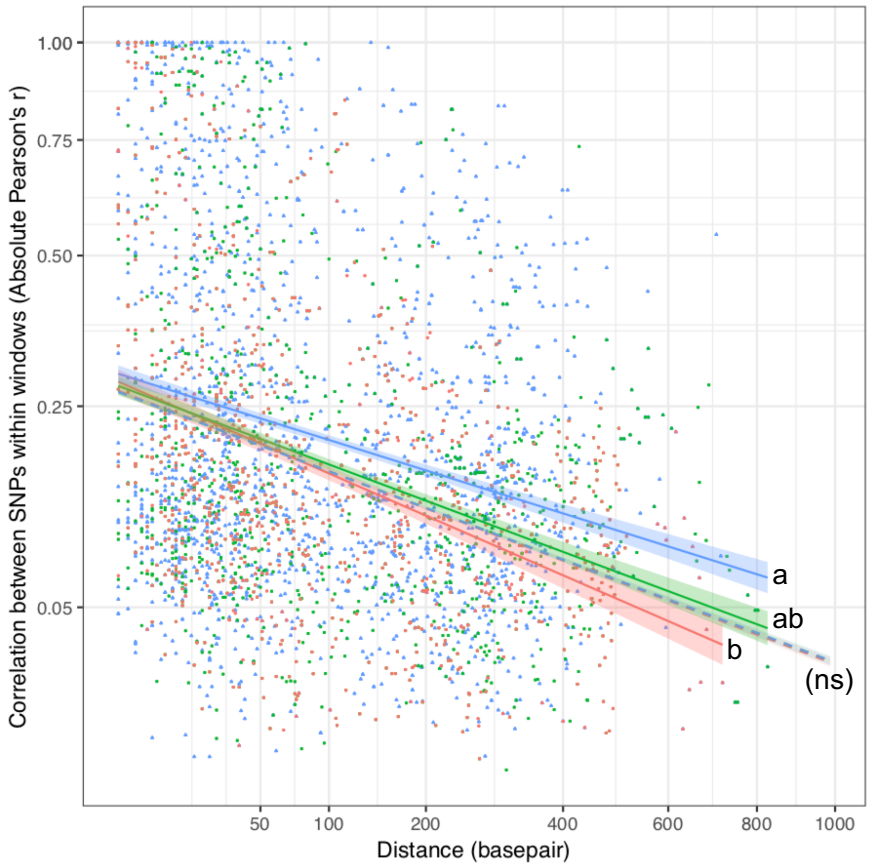


Figure 6. Decay of linkage disequilibrium (measured as pairwise correlation between SNPs in Pearson's r) over increasing genomic distance in base pairs within outlier (solid lines) and non-outlier (dashed line) 1000 bp windows within each *Ambrosia artemisiifolia* range (native North America=blue, introduced Europe=green, introduced Australia=red). Outliers were defined with the binomial method and had to be detected in both X^TX and GEA outlier tests. Letters depict significance of range difference for outlier windows, whereas LD decay was not significantly different among ranges for non-outlier windows (ns) (Table 3). The relationship between LD and distance is non-linear and we transformed axes to highlight differences. The non-transformed figure is shown in the Supporting Information (Fig. S5).

We explored genome-wide (non-local) LD between outlier and non-outlier SNPs (selected from outlier and non-outlier windows respectively) among ranges. We found a significant interaction between window type (outlier or non-outlier) and range (F_2=83.109, p<0.001). Dissection of this interaction in post-hoc tests revealed LD was higher between genome-wide outlier SNPs compared to non-outlier SNPs within each range (Fig 7, Table 4). LD between most pairwise range comparisons of outlier or non-outlier windows was significantly different, suggesting variation in patterns of genome-wide LD among ranges (Fig. 7, Table 4).

a. **Pairwise range comparisons**

Non-outlier			**Outlier**		
N-E	N-A	E-A	N-E	N-A	E-A
F_1=4499.195	F_1=42473.932	F_1=71405.75	F_1=270.244	F_1=0.703, q=0.402	F_1=40.394

b. **Non-outlier versus outlier within ranges**

N	E	A

F_1=2120.989	F_1=170.069	F_1=27.179

c.

	Non-outlier	**Outlier**
North America	0.061 (0.000)	0.102 (0.000)
Europe	0.055 (0.000)	0.072 (0.000)
Australia	0.079 (0.000)	0.106 (0.001)

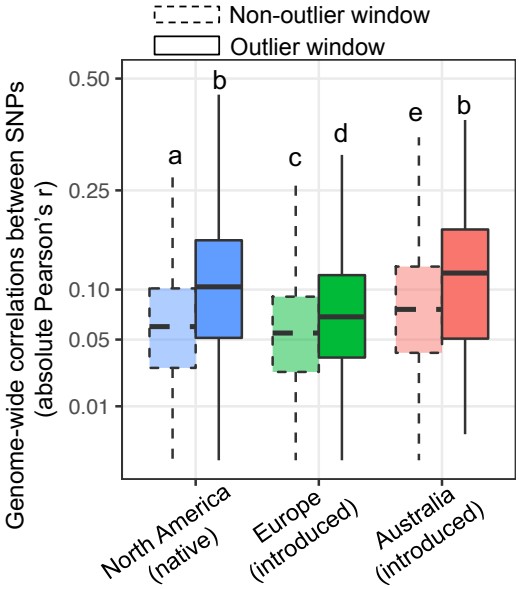


Figure 7. Linkage disequilibrium between SNPs (measured as pairwise correlation in Pearson's r) within outlier and non-outlier windows within each *Ambrosia artemisiifolia* range. Letters indicate significance (Table 4).

Discussion

We provide evidence to support rapid, repeatable genomic and phenotypic adaptation within the *A. artemisiifolia* native North American and introduced European and Australian ranges, despite alternative demographic histories. Among ranges, we identified high repeatability among phenotype-environment associations. We revealed significant sharing among ranges of genomic regions with elevated differentiation, with repeatability increasing with the level of differentiation. Furthermore, loci showing the strongest putative signatures of selection were shared more often between ranges than could be expected by chance alone. When combining the multiple lines of evidence from these phenotype-environment associations, F_{ST}-like outlier tests, genotype-environment associations and genotype-phenotype associations, 25-41% loci showed repeated signatures of selection among ranges.

Repeated patterns of phenotype-environment associations among ranges

We found evidence supporting rapid and repeated phenotypic adaptation over multiple continents to similar environments. Specifically, we identified parallel patterns in phenotype-environment associations among the three ranges, particularly between latitude and temperature measures with size, growth, phenology and sex function allocation (Table 1). The demographic histories of the populations included in the current study indicate climate matching (Maron *et al.*, 2004) is unlikely, as the geographic structuring of neutral markers did not follow the same patterns with climate in each range (van Boheemen *et al.*, 2017b). Moreover, Australian phenotypes are highly divergent from their proposed European source and likely represent a single, bottlenecked introduction (van Boheemen *et al.*, 2017b; van Boheemen *et al.*, 2018a), suggesting that introduction history and pre-adaptation of populations to climate variation across the Australian landscape is an implausible mechanism underlying the observed repeated trait evolution. Phenotypic clines can be produced by drift during range expansion (Colautti & Lau, 2015) but it seems unlikely that this process would create parallel clines repeatedly, and in the same direction in all three ranges and across multiple quantitative traits. Therefore, our phenotypic common garden observations support rapid and repeated parallel patterns of adaptation in response to climate variability within each range.

Genomic repeatability among ranges

We identified notable genomic repeatability using multiple lines of evidence (X^TX and GEA). Repeatability of putative selective signatures at the genetic level has rarely been studied in introduced species, and results vary from little (Gould & Stinchcombe, 2017), to substantial evidence for repeatability (Marques *et al.*, 2018). Our results add to this previous research by comparing multiple introductions, showing high genomic repeatability despite distinct demographic histories. Repeatability increased with the strength of differentiation for the X^TX

analysis. Disproportionate repeatability of high-differentiated SNPs and genes has been previously described in several studies (e.g. Renaut *et al.*, 2014; Trucchi *et al.*, 2017). These patterns suggest biases and/or constraints in the genetic basis of adaptation. Shared standing variation is one bias that can contribute to the increased probability of gene reuse during local adaptation (Conte *et al.*, 2012), as are differences in mutation rate among genes (Conte *et al.*, 2012; Storz, 2016; Yeaman *et al.*, 2018) or shared heterogeneity in recombination rates (Renaut *et al.*, 2014; Conte *et al.*, 2015). As native range expansion (Martin *et al.*, 2014) and introduction to Europe and Australia were recent (van Boheemen *et al.*, 2017b), standing variation is likely to strongly contribute to rapid adaptation due to the waiting time of new mutations (Barrett & Schluter, 2008). In addition to biases that can contribute to repeatability at the genetic level, gene reuse could occur because of low redundancy in the mapping of genotype to phenotype (only a few ways to make the same phenotype), or because of low redundancy in the mapping of genotype to fitness (only a subset of the genotypes yielding the same phenotype are optimal)(Yeaman *et al.*, 2018). Low redundancy in the mapping of genotype to fitness can arise due to indirect effects such as epistatic interactions or pleiotropic effects on other selected traits (Chevin *et al.*, 2010b; Weinreich *et al.*, 2005). Such redundancy can also be influenced by aspects of the genetic architecture such as the number of alleles, recombination rates (Renaut *et al.*, 2014), and effect sizes and is dependent on the interaction between drift, selection and migration (Yeaman *et al.*, 2018). Closely related lineages, such as recent introductions, are more likely to share a similar genomic context, increasing the probability of repeatable patterns emerging (Holliday *et al.*, 2016).

We identified a higher repeatability when comparing genomic windows versus single nucleotide divergence (Table 2, Figure 2). This reduction in repeatability towards lower levels of organization has been observed previously when comparing SNP to allozyme markers (Ravinet *et al.*, 2016), genomic regions or genes (Tenaillon *et al.*, 2012; Renaut *et al.*, 2014), biochemical pathways (Roda *et al.*, 2013) and phenotypes to ecological function (McGee & Wainwright, 2013). Hierarchical decline in parallelism is proposed to result from redundancy in the mapping of one level to the next, where multiple elements could affect the same higher-level function (McGee & Wainwright, 2013; Lenormand *et al.*, 2016). Furthermore, as outlier loci are often likely linked to functional changes under selection instead being the direct cause of adaptive differences, high levels of repeatability at the SNP level may not be unexpected, even if the same functional changes are under selection in the different ranges. Indeed, quantitative traits are often polygenic and classic quantitative genetic theory predicts high levels of redundancy in the mapping of genotype to phenotype (Pritchard & Di Rienzo, 2010; Pritchard *et al.*, 2010). Future work dissecting redundancy at different levels (e.g. Yeaman *et al.*, 2018) will provide insight into the processes that limit the genetic basis of adaptation (Lässig *et al.*, 2017) and the mechanisms by which species evolve.

Integration of genomic and phenotypic repeatability

Overlap in adaptive signals among ranges identified on the genomic and phenotypic level was much higher than could be expected by stochastic processes, such as drift, when we combined the multiple lines of evidence to identify adaptation candidates (Fig. 5). This excessive similarity among ranges was found for every combination of outlier tests and is therefore implausible to have resulted from biases unique to individual tests. These results indicate that the Australian genetic bottleneck and short time since introduction (~80 generations; van Boheemen *et al.*, 2017b) did not appear to strongly impede the repeatability of adaptation. Indeed, rapid adaptation following bottlenecks are not uncommon and has been found in *Hypericum canariense* (Dlugosch & Parker, 2008b). In addition, similar extremely rapid and repeated adaptation was recently shown in threespine sticklebacks, where <20 generations of adaptive divergence paralleled 75% of genomic changes known to have accrued over 12k years, despite the presence of genetic bottlenecks (Marques *et al.*, 2018).

Biases in outlier detection and repeatability

The number of putative adaptive loci associated with each environmental variable (candidate variable) was relatively low compared to those putative adaptive loci identified for non-candidate environmental variables. Moreover, we identified fewer candidate environment-associated outliers in North America compared to Europe or Australia. One explanation for this pattern could be differences in the genomic architecture of traits diverging in response to candidate and non-candidate variables, alongside differences in architecture between ranges. For instance, fewer loci contributing to adaptive divergence responding to the candidate environmental variables could lead to observed pattern. Conversely, these results could indicate our tests were underpowered, particularly in situations where the population structure covaries with adaptive differentiation in the diverging trait and/or the underlying loci are of small effect. Furthermore, the signal of a selective sweep should dissipate over time, whereby the region of the genome impacted by the sweep declines with the time since the sweep (Maynard Smith & Haigh, 1974; Przeworski, 2002; Kim & Nielsen, 2004). This might be the case for the native range where each population has likely been experiencing relatively stable environmental conditions for a longer period of time versus those that have been recently founded in the introduced range. Alternatively, some regions of the genome impacted by climate selection may not have been sampled in the GBS analysis limiting our capacity to detect selection for some traits. Moreover, the limited LD observed in the current study likely results in a reduces selective signal in the genome, thus diminishing the number of putative adaptive SNPs to be discovered. Whole-genome sequencing, larger sample sizes, together with polygenic and

multivariate tests of selection could help to differentiate among these diverse biological and technical causes to differences in the number of putative adaptive loci among traits and ranges.

Further mechanisms could lead to biases of adaptive signatures and genetic repeatability (Renaut *et al.*, 2014; Conte *et al.*, 2015). For instance, pleiotropy could cause overestimates when alternate traits influenced by a single gene are adaptive in different ranges. Hitchhiking can result in biases through the repeated detection of false-positive, neutral loci linked to loci adaptive in multiple ranges. This would likely result in a decrease of genomic repeatability at lower levels of organization (Conte *et al.*, 2015). Similar problems could arise if multiple mutations contribute to adaptation within a single genomic region, as would be expected with the clustering of adaptive loci (Yeaman, 2013; Rougemont *et al.*, 2017). These biases are difficult to disentangle without functional analysis of regions with signatures of selection. Furthermore they are challenging to control for statistically given the fragmentation of the genome assembly and the fact that some of our scaffolds are linked with one another (Hoban *et al.*, 2016).

Overestimates of repeatability are also predicted from underpowered tests of selection as small effect loci, which are more difficult to detect, are thought to be less frequently involved in adaptive gene reuse (Yeaman *et al.*, 2018). A polygenic basis (Chevin & Hospital, 2008; Pritchard *et al.*, 2010) and soft sweeps (Hermisson & Pennings, 2005; Messer & Petrov, 2013) complicates detection of selective signatures in outlier analyses, as allele frequency changes in response to even strong selection at the trait level are often small on individual underlying loci (Hancock *et al.*, 2010; Turchin *et al.*, 2012). As a result, changes in covariances among loci could result in divergence of highly polygenic traits, even with limited genetic differentiation at individual loci (Kremer & Le Corre, 2012; Le Corre & Kremer, 2012). Repeated signatures of selection would be challenging to identify if small effect loci are the main contributors to adaptive divergence (Gould & Stinchcombe, 2017; Trucchi *et al.*, 2017) both because they are more difficult to identify technically, and they are less likely to be involved in repeatability from the theoretical perspective. However, it is possible repeatable phenotypic adaptation in our study involved low-frequency changes. Indeed, elevated patterns of sharing were observed when exploring minor shifts in allele frequencies using the more sensitive null-W method, perhaps in part because of the recency of selection in the introduced ranges, or the complex genetic architecture of the traits likely under selection resulting is more subtle allele frequency changes.

Differences in false-positive rates among the ranges could result from the dramatically distinct demographic histories (Hodgins *et al.*, 2018). For instance, more false-positives could be expected in the introduced ranges compared to the native due to recent allele surfing following recent range expansion (Klopfstein *et al.*, 2006). In addition, the multiple introductions experienced in Europe (van Boheemen *et al.*, 2017b) could conserve false positive

loci identified in North America, resulting in their elevated repeatability. F_{ST} outlier type tests (e.g. X^TX) would be particularly prone to this problem, as repeated false positives from the GEA analysis would require matching of the climate between the source and the introduced ranges, while F_{ST} outlier analyses would not. GEA has been found to outperform F_{ST}-like tests under certain demographic scenarios (Bierne *et al.*, 2013). Such a bias would however not explain the repeatable patterns with Australia, where populations have probably become established following a single introduction (van Boheemen *et al.*, 2017b). Additionally, F_{ST}-like outlier tests are sensitive to low heterozygosity, which is common in regions of low recombination (Tiffin & Ross-Ibarra, 2014). Consequently, F_{ST} outliers can arise through processes other than local adaptation, including selective sweeps or background selection (Ortiz-Barrientos *et al.*, 2016). As populations within the multiple ranges likely share genomic context, and thus patterns of recombination, it is plausible X^TX false-positives are shared among ranges.

Lower repeatability of patterns in GEA compared to X^TX could arise through a relatively high false negative rate of the former as this test underperforms in isolation-by-distance (IBD) scenarios (Lotterhos & Whitlock, 2015). IBD is however negligible or absent in the tested ranges (van Boheemen *et al.*, 2017b) and thus not likely to lead to observed test differences. Alternatively, key selective factors might not have been included in GEA, however the tested phenotypes did show strong associations with the included environmental variables. Regardless of these discrepancies between tests, the high sharing of putative adaptive loci among ranges when combining tests provides convincing evidence for repeated adaptation at the genetic level.

Demographic history, selection and linkage disequilibrium

We identified higher local linkage disequilibrium (LD) in the native range for loci showing putative signatures of selection versus putative non-adaptive loci. Such high local LD excess could be indicative of recent selective sweeps (Przeworski, 2002). As introduced populations face novel selection regimes, local LD as a result of recent sweeps might be expected to be higher in these populations compared to the native range. However, admixture, reduced population sizes and increased genetic drift, all observed in the introduced ranges (van Boheemen *et al.*, 2017b), are expected to elevate associations among alleles at linked loci (Slatkin, 2008), potentially masking LD differences between putative adaptive and non-adaptive loci. Such linkage could interfere with individual loci's response to selection if the associated alleles have opposing selection coefficients (Fisher, 1930; Muller, 1964; Hill & Robertson, 1968; Felsenstein, 1974; McVean & Charlesworth, 2000; Slatkin, 2008). Our results, however, suggest that, although distinct demographic histories have left their signature on local LD over short windows, these processes did not appear to impede rapid adaptive

change both in terms of rapidly generating phenotypic clines and allele frequency changes of putative adaptive loci.

Interestingly, in contrast to these patterns in local LD, we revealed differences in genome-wide LD among SNPs in outlier versus non-outlier windows within each range. Patterns of LD can be created by selection or physical linkage. To reduce the possibility of physical linkage among these SNPs, we selected a single SNP per window, with a single window per scaffold. However, we cannot exclude the possibility these SNPs are still physically linked, as we currently do not have a genetic map for this species. Nevertheless, this elevated genome-wide statistical association among putative adaptive loci is indicative of selection driving parallel responses in multiple putatively independent regions of the genome.

Conclusion

This study is one of an emerging cohort of investigations focused on the repeatability, and thus predictability (Lässig *et al.*, 2017), of climate adaptation on the phenotype and genotype. By gathering multiple lines of evidence on these hierarchical levels of organization, we have strong indications that major evolutionary changes can occur on short timescales in response to climate variability. Importantly, by comparing adaptation among multiple introductions with distinct demographic histories, we make a substantial contribution to the increasing body of evidence showing that introduction dynamics can affect neutral genetic diversity (Dlugosch & Parker, 2008a), while adaptive potential of those traits might not be constrained to a similar extent (Dlugosch *et al.*, 2015a). We support our results by an exploration of the genomic architecture and show elevated LD among putative adaptive loci.

Studies, such as ours, showing repeatability of changes occurring over contemporary timescales, imply a promising future for our ability to predict aspects of a species' genetic and phenotypic response to a changing climate. Future studies could examine features of the genomic architecture that might play a key role in evolvability. For instance, theoretical studies predict that loci with large phenotypic effects will respond first to selective changes (MacPherson & Nuismer, 2017), will contribute more to repeatable adaptation (Lässig *et al.*, 2017) and are considered important in invasive spread (Dlugosch *et al.*, 2015a). Prospective research could benefit from a native-introduced study system with variable invasion history, as presented in the current study, to explore the genomic basis of adaptation and the impact of historical contingency on the genetic route by which adaptation proceeds.

Acknowledgements

We would like to thank J. Stephens and A. Wetherhill for sample collection, M. Kourtidou and J. Taylor for greenhouse assistance and S. Bou-Assi and K. Nurkowski for genomic analyses. A Monash University Dean's International Postgraduate Research Scholarship was provided to

LAB, a Monash University Startup Grant to KAH.

Author contributions

Both authors developed the project, with data collection and analyses carried out by LAB, refined by KH. The authors discussed the results, contributed to the MS writing and gave final approval for publication.

Data accessibility

Sequence data are available at the National Center for Biotechnology Information (NCBI) Sequence Read Archive under Bioproject PRJNA449949.

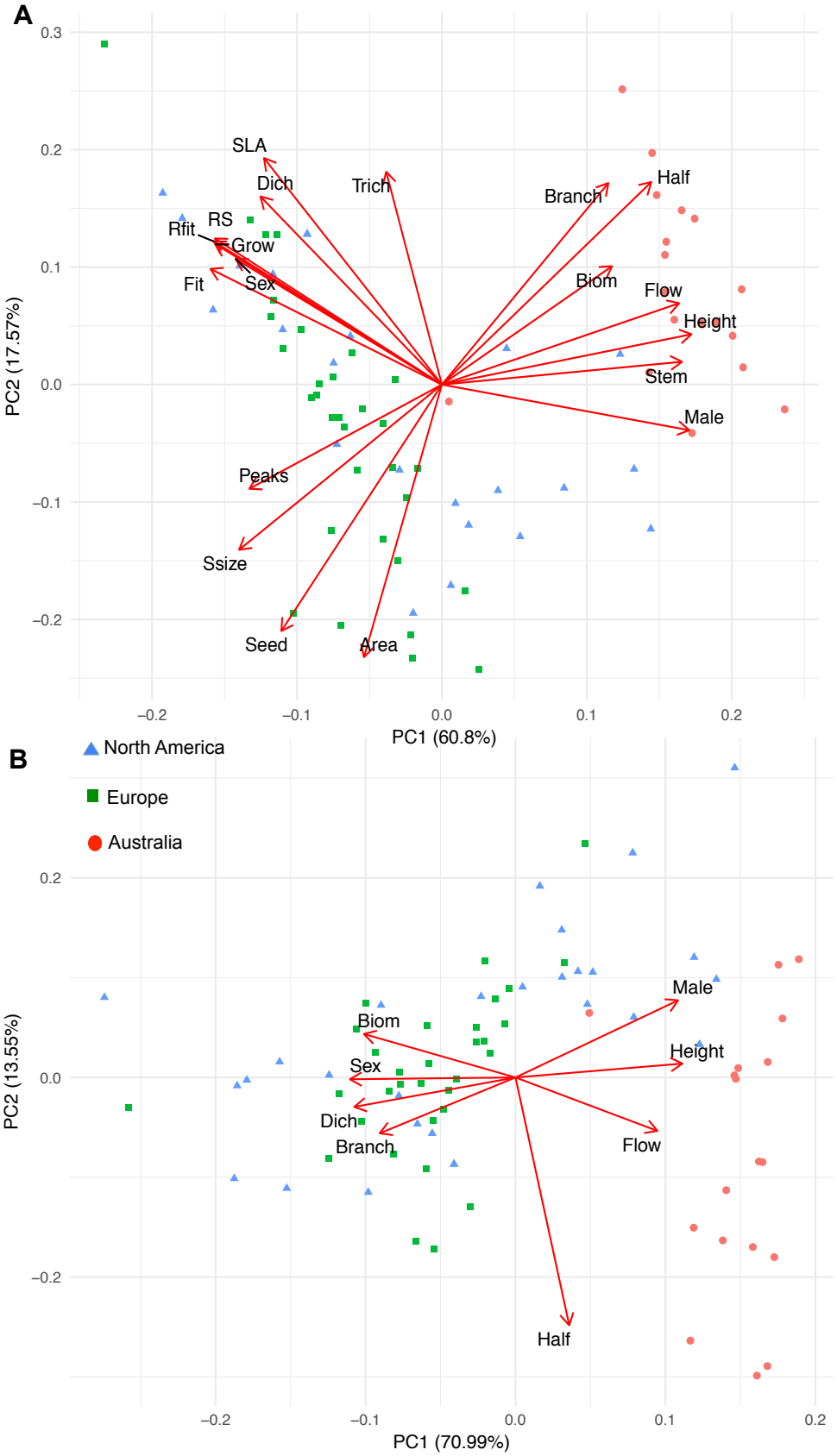


Figure S1. Principle component analysis for all phenotypic variables (A) and shared PEA variables (B) within the native North American (blue triangles) and invasive European (green squares) and Australian (red circles) *Ambrosia artemisiifolia* ranges, with in brackets the percentage explained by the primary principle components (PC1 and PC2) and abbreviations corresponding to Table S2.

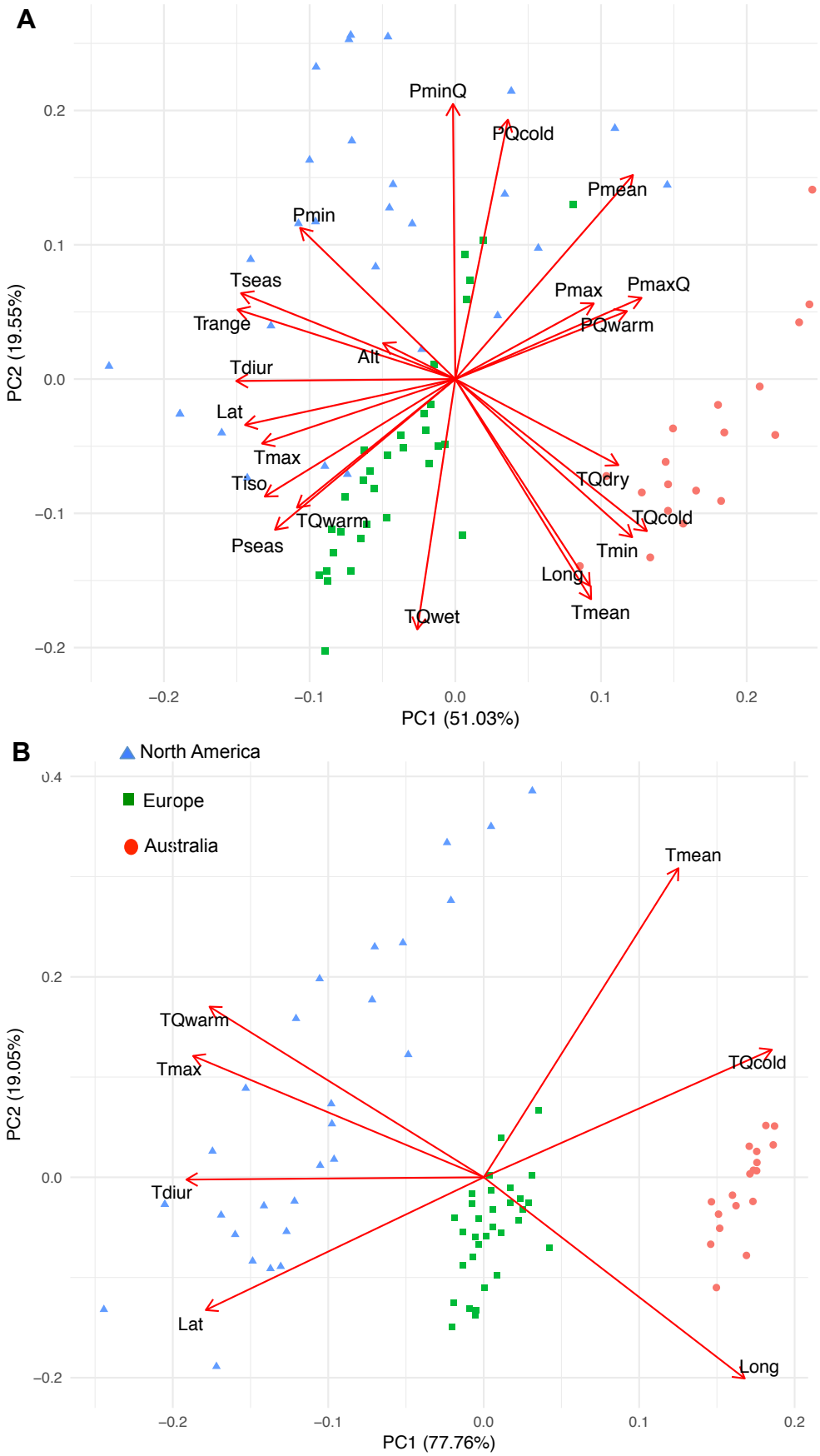


Figure S2. Principle component analysis for all environmental variables (A) and shared PEA variables (B) within the native North American (blue triangles) and invasive European (green squares) and Australian (red circles) *Ambrosia artemisiifolia* ranges, with in brackets the percentage explained by the primary principle components (PC1 and PC2) and abbreviations corresponding to Table S2.

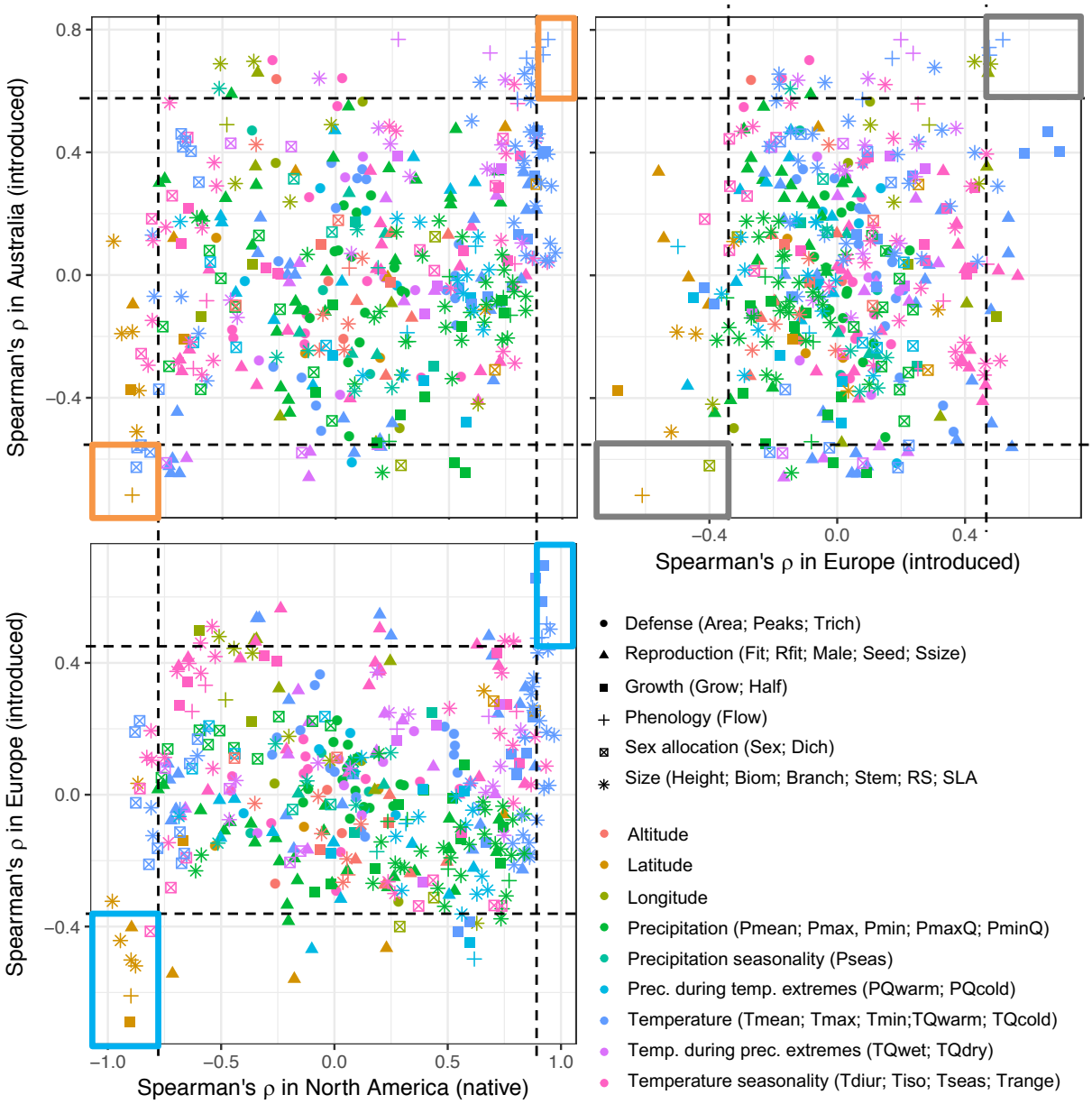


Figure S3. Signatures of repeatable phenotypic local adaptation among *Ambrosia artemisiifolia* ranges. Phenotype-environment associations (PEA) were measured in Spearman's ρ within each range, where the top 5% cut-offs (based on ranked ρ within each range) are depicted by dashed lines. Top 5% PEA shared amongst range pairs (solid squares) are indicative of repeatable adaptation. Categorical classification is for graphical purposes only (see Tables S1 & S2).

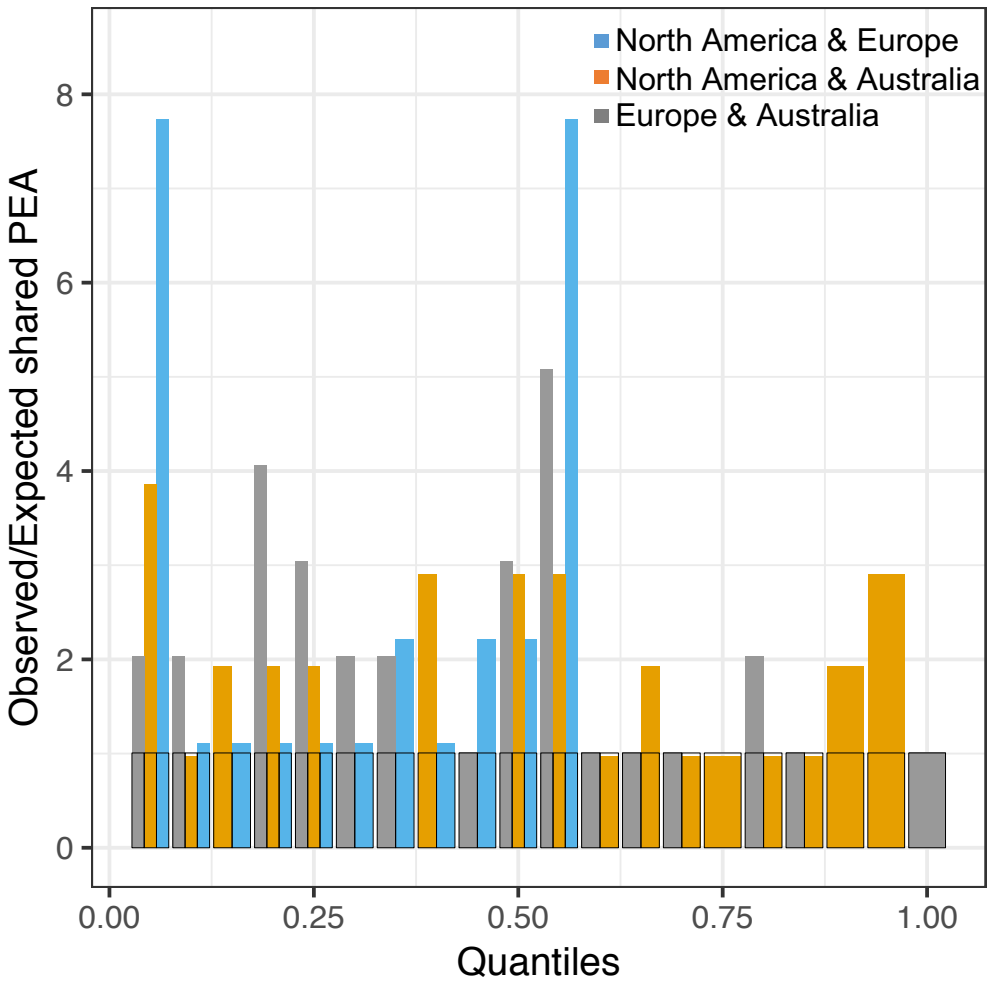


Figure S4. Observed/expected number of phenotype-environment pairs shared between range pairs in the same 5% quantile bin based on the strength (Spearman's rho) of these phenotype-environment associations (PEA) with 95% confidence intervals of the binomial expectation under random distribution of test values illustrated with black lines.

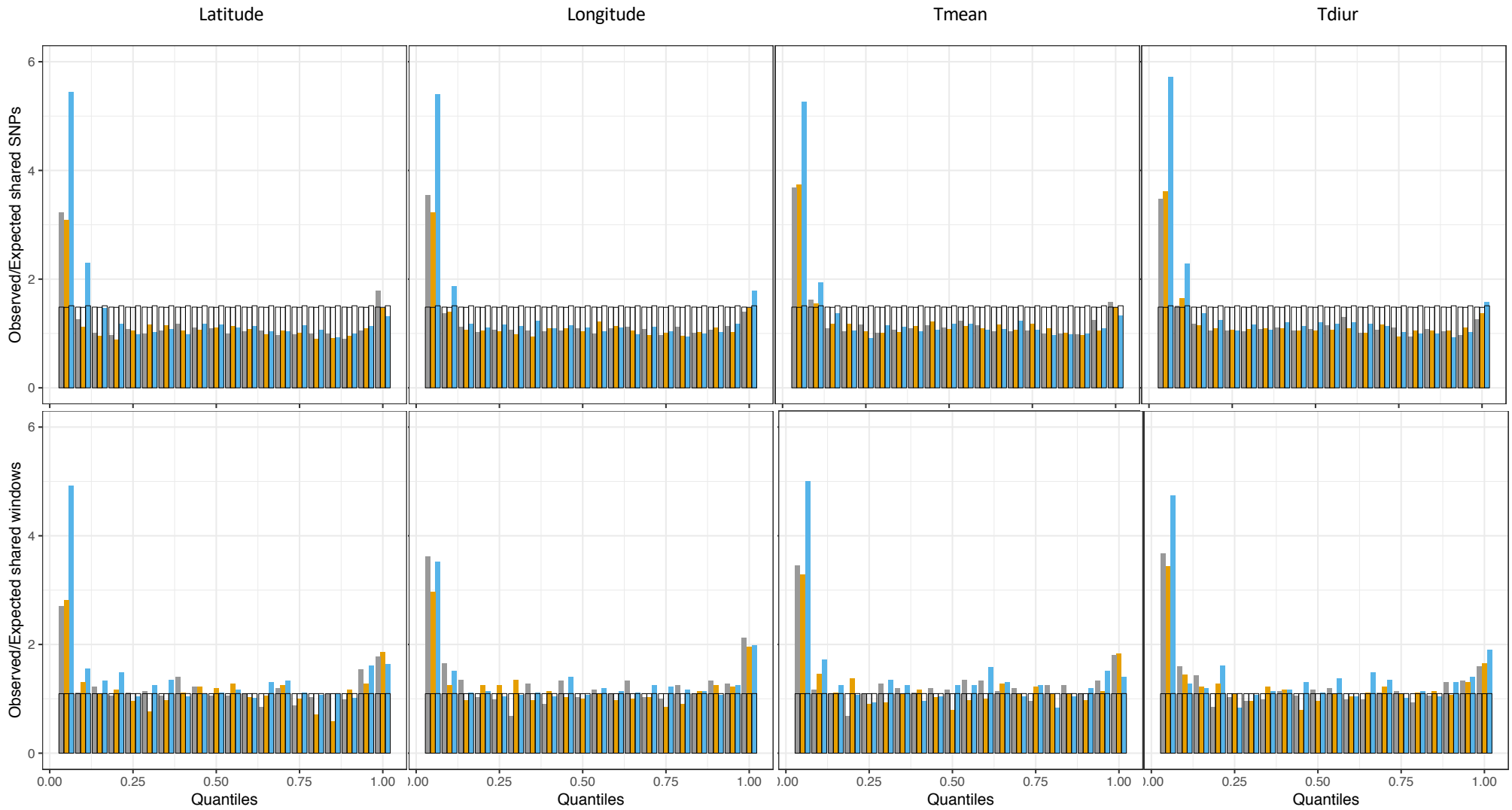
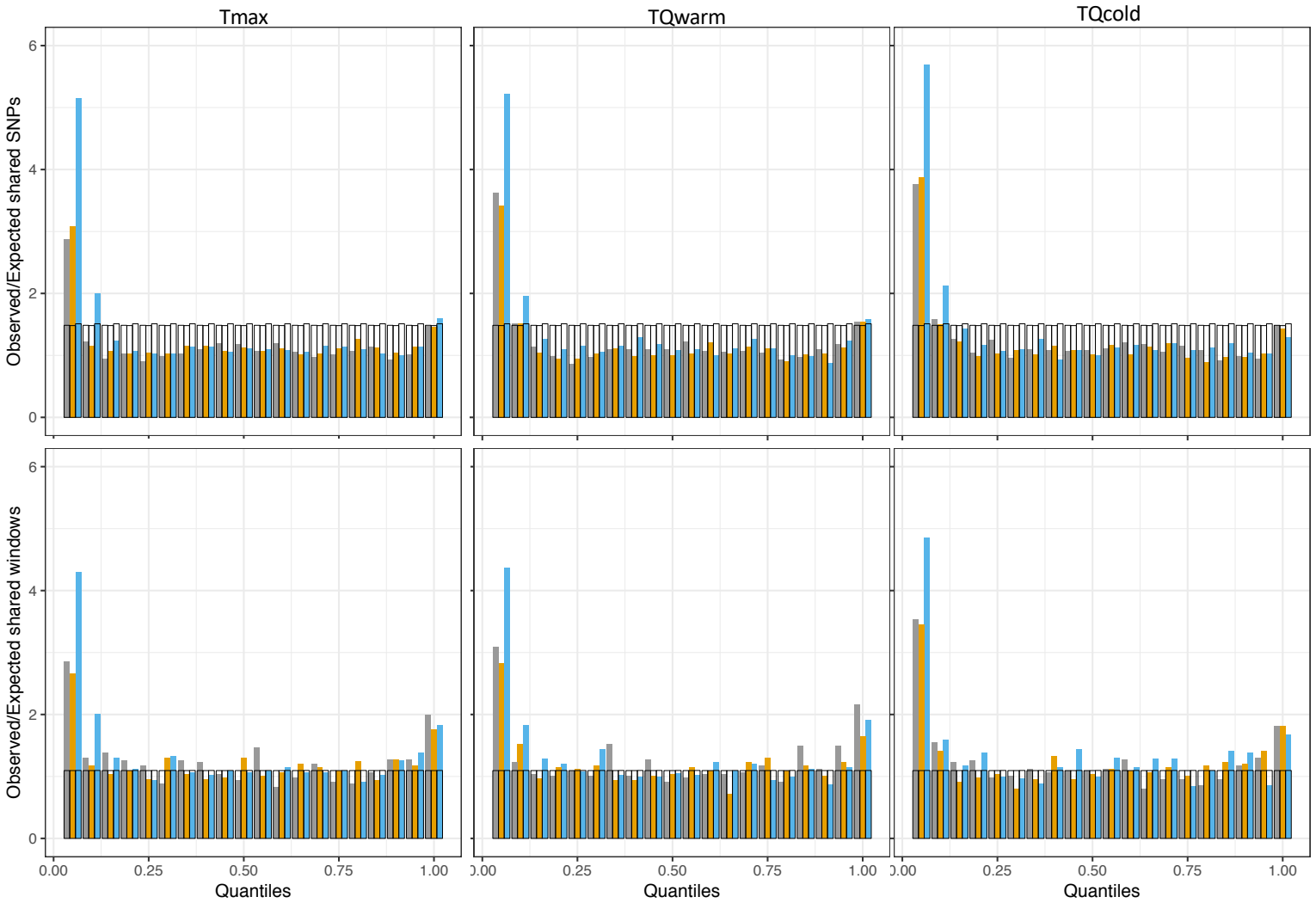


Figure S5. Observed/expected number of SNPs (top panels) and 1000 bp windows (bottom panels) shared between range pairs in the same 5% quantile bin based on BF values with 95% confidence intervals of the binomial expectation under random distribution of test values illustrated with black lines. Presented figures include top 5% PEA environmental variables only.

Figure S5. Continued



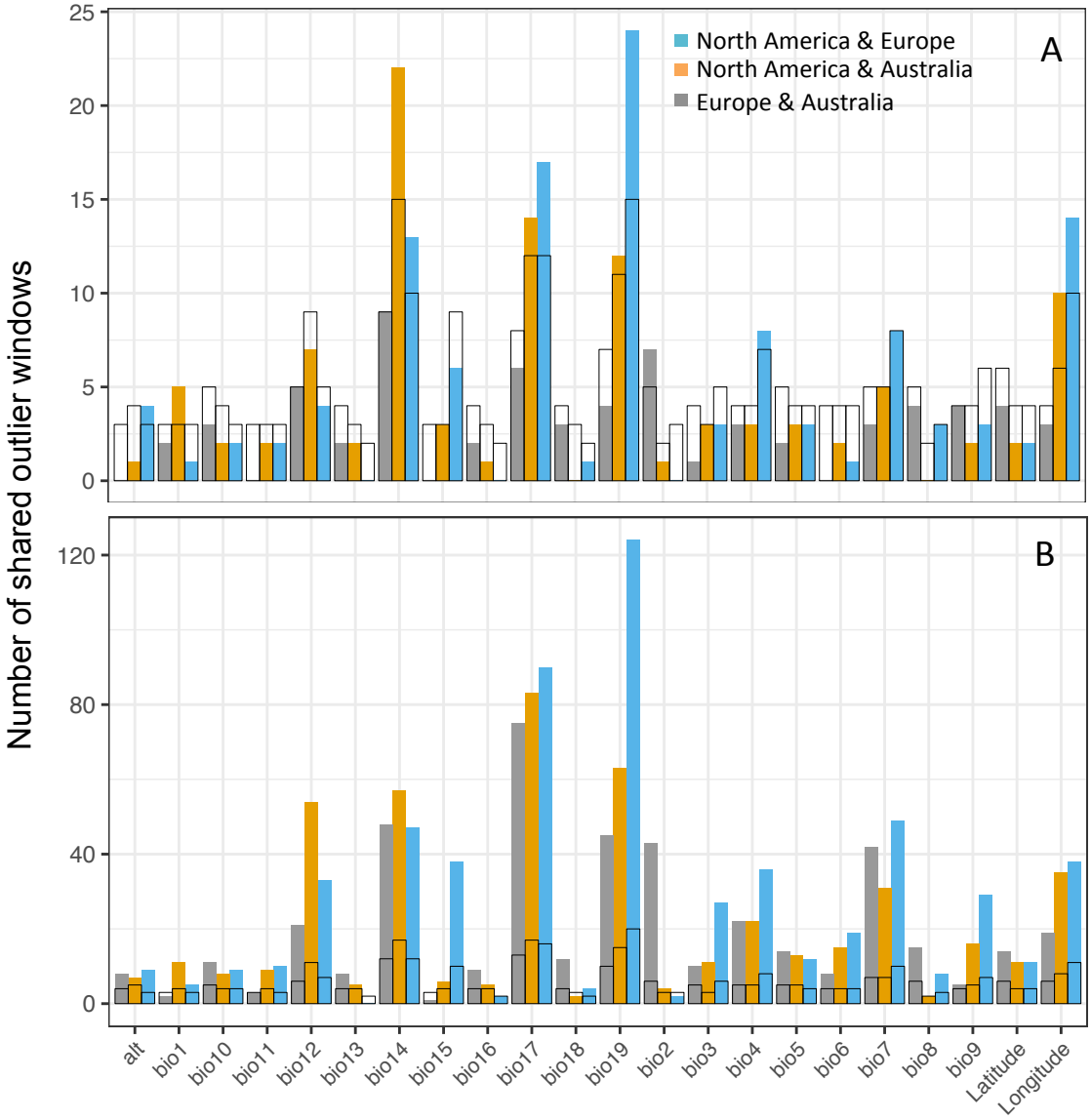


Figure S6. Number of 1000 bp outlier windows as identified with binomial (A) and null-W tests (B) shared among range pairs with 99.9th hypergeometric quantiles illustrated with black lines.

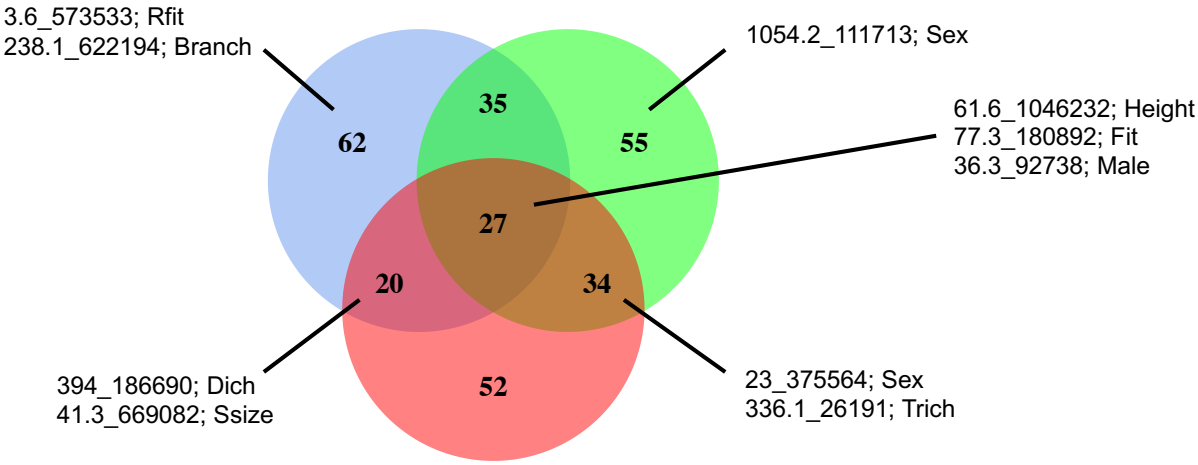


Figure S7. Number of *Ambrosia artemisiifolia* outlier windows (binomial and null-W tests, identified in both X^TX and GEA) within each or multiple ranges. In each diagram: blue/top left=native North America; green/top right=introduced Europe; red/bottom=introduced Australia. An outlier window had to be identified in both X^TX (F_{ST}-like outlier test) and GEA (genotype-environment associations) where windows identified as outliers for multiple environments were counted once. Binomial outlier window ID and associated phenotypes are provided when identified in GPA (phenotype abbreviations correspond to Table S1, Supporting Information; Rfit=Relative reproductive biomass; Height=Max. plant height; Branch=Number of branches; Sex=Sex function allocation; Fit=Total reproductive biomass; Male=Weight of male sex function; Dich=Dichogamy; Ssize=Average seed weight; Trich=Trichome density). The overlap in outlier windows in each pairwise range comparison is significantly higher ($q \leq 0.001$) than could be expected under a random hypergeometric distribution.

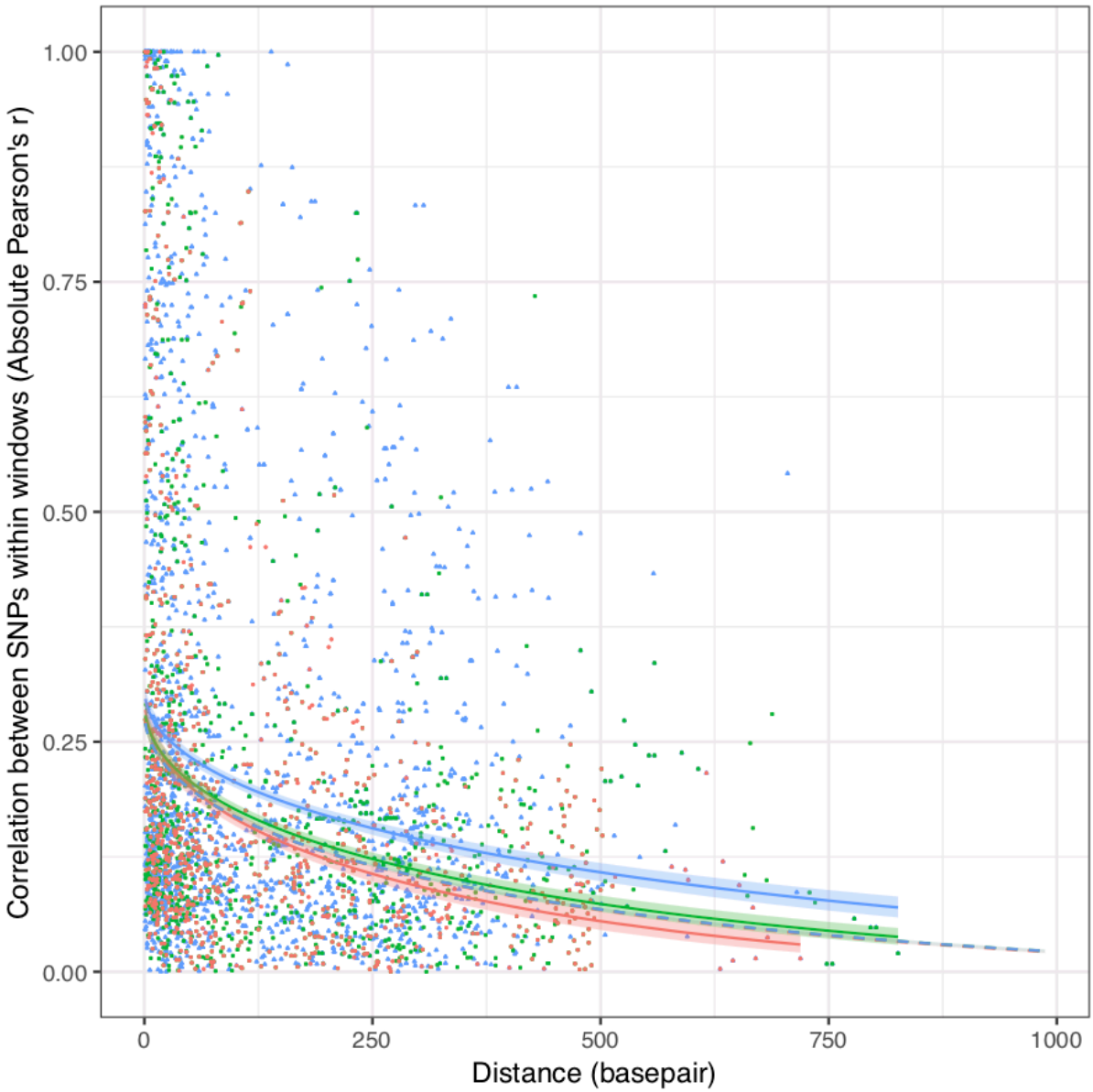


Figure S8. Decay of linkage disequilibrium (absolute correlation between SNPs) over increasing genomic distance in base pairs within (solid lines) 1000 bp outlier windows and within non-outlier windows (dashed line) of the same size within each *Ambrosia artemisiifolia* range (native North America=blue, introduced Europe=green, introduced Australia=red), where outliers were defined with the binomial method and had to be detected in both X^TX and GEA outlier tests.

Table S1. List of traits included in this study with abbreviation used in figures and associated first two principal components loadings (PC) from PCA for all (Fig. S1A) or shared PEA phenotypes (Fig. S1B), with colour scales corresponding to loadings.

			PCA for all phenotypes		PCA for shared PEA phenotypes	
Abbreviation	**Full variable name**	**Description**	**PC1**	**PC2**	**PC1**	**PC2**
Biom	Total biomass	Above- and belowground dry biomass	0.046	0.042	0.133	0.083
Branch	Number of branches	Number of branches from main stem; only branches ending in leaves are counted	0.045	0.072	0.119	0.107
Height	Maximum plant height	Maximum measured height	0.068	0.018	0.147	0.027
RS	Root/shoot ratio	Root/shoot dry biomass	0.062	0.052	-	-
SLA	Specific leaf area	Leaf area of fully expanded fresh leaf/leaf dry weight	0.048	0.081	-	-
Stem	Stem width	Basal stem width	0.065	0.008	-	-
Dich	Dichogamy	First recorded day of pollen release - first recorded day of receptive female function (a positive value is protogyny, a negative value is protandry)	0.049	0.067	0.141	0.056
Sex	Sex function allocation	Female (seeds) / male (raceme) dry weight (a value >1 is higher biomass allocation to female function)	0.056	0.045	0.145	0.003
Flow	Flowering onset	First recorded day of flowering (number of days after second transplant); first day of pollen release (male function) or receptive female function	0.064	0.029	0.125	0.102
Grow	Maximum growth rate	Sensu Chuine *et al.*, 2001	0.061	0.050	-	-
Half	Time of half growth	Sensu Chuine *et al.*, 2001	0.057	0.072	0.047	0.474
Fit	Total reproductive biomass	Female (seeds) and male (raceme) dry weight	0.063	0.041	-	-
Male	Male function reproductive biomass	Dry weight of male function flowering structure (racemes)	0.067	0.016	0.143	0.148
Rfit	Relative reproductive biomass	Total reproductive biomass / total plant biomass	0.062	0.051	-	-
Seed	Total seed weight	Dry weight of female function flowering structure (seeds)	0.044	0.088	-	-
Ssize	Average seed weight	Dry weight per seed in milligrams, averaged over 20 seeds (where available)	0.055	0.059	-	-
Area	Total phenolic peak area	Peak area of >350mAU phenolic peaks identified with high-performance liquid chromatography ()	0.052	0.040	-	-
Peaks	Number of phenolic peaks	Number of >350mAU phenolic peaks identified with high-performance liquid chromatography ()	0.020	0.095	-	-
Trich	Number of trichomes	Trichome density at the mid-point of each plant ()	0.012	0.070	-	-

Table S2. List of bioclimatic and geographic variables included with abbreviations used in text and according to the Worldclim database and associated first two principal components loadings (PC) from PCA for all (Fig. S2A) or shared PEA environmental variables (Fig. S2B) with colour scales corresponding to loadings.

			PCA for all environmental variables (Fig. S2A)		PCA for shared PEA variables (Fig. S2B)	
Abbreviation	**Worldclim code**	**Description**	**PC1**	**PC2**	**PC1**	**PC2**
Tmean	BIO1	Annual Mean Temperature	0.040	0.076	0.103	0.290
Tdiur	BIO2	Mean Diurnal Range (Mean of monthly (max temp - min temp))	0.065	0.001	0.158	0.002
Tiso	BIO3	Isothermality (BIO2/BIO7) (* 100)	0.056	0.041	-	-
Tseas	BIO4	Temperature Seasonality (standard deviation *100)	0.063	0.030	-	-
Tmax	BIO5	Max Temperature of Warmest Month	0.057	0.022	0.154	0.114
Tmin	BIO6	Min Temperature of Coldest Month	0.052	0.055	-	-
Trange	BIO7	Temperature Annual Range (BIO5-BIO6)	0.064	0.024	-	-
TQwet	BIO8	Mean Temperature of Wettest Quarter	0.011	0.087	-	-
TQdry	BIO9	Mean Temperature of Driest Quarter	0.048	0.030	-	-
TQwarm	BIO10	Mean Temperature of Warmest Quarter	0.047	0.045	0.146	0.160
TQcold	BIO11	Mean Temperature of Coldest Quarter	0.057	0.053	0.153	0.120
Pmean	BIO12	Annual Precipitation	0.053	0.071	-	-
Pmax	BIO13	Precipitation of Wettest Month	0.041	0.026	-	-
Pmin	BIO14	Precipitation of Driest Month	0.046	0.052	-	-
Pseas	BIO15	Precipitation Seasonality (Coefficient of Variation)	0.053	0.052	-	-
PmaxQ	BIO16	Precipitation of Wettest Quarter	0.055	0.028	-	-
PminQ	BIO17	Precipitation of Driest Quarter	0.001	0.095	-	-
PQwarm	BIO18	Precipitation of Warmest Quarter	0.051	0.024	-	-
PQcold	BIO19	Precipitation of Coldest Quarter	0.016	0.090	-	-
Alt	-	Altitude	0.021	0.012	-	-
Lat	-	Absolute latitude	0.062	0.016	0.148	0.125
Long	-	Longitude	0.040	0.072	0.139	0.189

Table S3. Patterns of repeatable signatures of selection measured as the strength of phenotype-environment associations (PEA), genomic divergence (X^TX) or genotype-environment associations (GEA) of phenotype-environment pairs (PEA) SNPs and 1000 bp genomic windows (X^TX and GEA) within *A. artemisiifolia* range pairs (N: native North America; E: introduced Europe; A: introduced Australia) with Pearson's correlation coefficients calculated among test

N-E	**N-A**	**E-A**

values. P-values were <0.001 unless indicated in parentheses. Variable abbreviations correspond to Table S2.				
Phenotype-environment pair				
PEA		0.048, p=0.328	0.246	0.059, p=0.154
SNP				
XTX		0.501	0.455	0.434
GEA	Latitude	0.025	0.000, p=0.964	0.019
	Longitude	0.044	0.025	0.022
	Tmean	0.022	0.014	0.018
	Tdiur	0.033	0.016	0.021
	Tmax	0.028	0.014	0.022
	TQwarm	0.026	0.015	0.022
	TQcold	0.031	0.015	0.012
1000 bp window				
XTX		0.585	0.474	0.489
GEA	Latitude	0.059	0.001, p=0.919	0.018, p=0.031
	Longitude	0.055	0.043	0.05
	Tmean	0.045	0.012	0.038
	Tdiur	0.061	0.044	0.032
	Tmax	0.056	0.017	0.029
	TQwarm	0.059	0.013	0.039
	TQcold	0.065	0.021	0.037

Table S4. Total number of SNPs, top 1% ranked 'outlier' SNPs based on test-values, windows and number of 'outlier' windows with more outlier SNPs compared to binomial expectation identified using tests of SNP divergence (X^TX), environment-allele associations (GEA) and genome-wide genotype-phenotype associations (GPA) within the global, native North American (NA) and introduced European (EU) and Australian (AU) datasets. SNPs and windows identified as outliers across multiple environments or phenotypes were counted once.

		Total number				**Number of outliers**			
		Global	**NA**	**EU**	**AU**	**Global**	**NA**	**EU**	**AU**
SNPs	**X^TX**	83,559	82,995	82,097	77,410	835	831	830	783
	GEA					7,562	5,568	8,629	6,655
	GPA	11,496	11,496		NA	3,295	2,188		NA
Windows	**X^TX**	15,549	15,516	15,378	15,090	190	151	152	135
	GEA					978	843	1,161	952
	GPA	2,797	2,797		NA	94	73		NA

Chapter 5 – EICA fails as an explanation of growth and defence evolution following multiple introductions

Lotte A. van Boheemen[1*], Sarah Bou-Assi[1], Akane Uesugi[1], Kathryn A. Hodgins[1]

[1]School of Biological Sciences, Monash University, Clayton, VIC 3800, Australia

*Corresponding author, la.vanboheemen@gmail.com

In review at *Journal of Ecology*

Abstract

1. Rapid adaptation is aiding invasive populations in their competitive success. The evolution of increased competitive ability (EICA) hypothesis posits this enhanced performance results from escape from native enemies, yet its support is equivocal.
2. We here test EICA by investigating adaptive divergence of various constitutive and inducible defence-related traits within the native North America and introduced European and Australian ranges, whilst controlling for divergence due to latitudinal trait clines, individual resource budgets and population differentiation, using >11,000 SNPs.
3. We do not identify a general reduction in defence in concert with an increase in growth among the multiple introduced ranges as predicted by EICA. Yet, rapid, repeated clinal adaptation in defence-related traits was apparent despite distinct demographic histories. Divergence in energy budgets among ranges may explain some, but not all, defence-related trait divergence.
4. *Synthesis*: The rapid spread of invasive species is affected by a multitude of factors, likely including adaptation to climate and escape from natural enemies. Unravelling the mechanisms underlying invasives' success enhances understanding of eco-evolutionary theory and is essential to inform management strategies in the face of ongoing climate change.

Keywords

EICA, latitudinal adaptation, growth-defence trade-offs, constitutive defence, inducible defence, invasive species, phenolic compounds, resource budgets

Introduction

Biological invasions are occurring at an accelerating pace due to the globalisation of anthropogenic activity (Ricciardi, 2007). Individuals colonizing new ranges likely face environments different from those previously experienced (Sax & Brown, 2000; Allendorf & Lundquist, 2003; Chown *et al.*, 2014). Nonetheless, alien populations often display enhanced performance compared to their native counterparts (Blossey & Notzold, 1995; Thébaud & Simberloff, 2001; Parker *et al.*, 2013), and this can be facilitated by rapid adaptation (Chown *et al.*, 2014; Colautti & Lau, 2015; Dlugosch *et al.*, 2015a). The evolution of increased competitive ability (EICA) hypothesis posits that invasives' trait divergence results from release from natural enemies, allowing the allocation of defence resources to growth and reproduction (Blossey & Notzold, 1995). However, empirical support for the predicted evolutionary decrease in invasive plants' herbivory defence in concert with increased competitive ability is equivocal and various between species (Bossdorf *et al.*, 2005; Felker-Quinn *et al.*, 2013). Therefore, EICA as yet falls short as a general hypothesis to explain invasion success.

The complex interplay between the evolutionary mechanisms shaping phenotypic divergence could confound inferences predicted by EICA. Distinct demographic processes, including founder effects, genetic drift and admixture, often characterize introduction and alone can lead to divergence between native and introduced populations (Lee, 2002; Facon *et al.*, 2006; Prentis *et al.*, 2008; Rius & Darling, 2014; Estoup *et al.*, 2016). Dissection of the various evolutionary processes that can contribute to trait divergence is required to advance our understanding of rapid spread in invasive species. In addition, the repeatability of evolutionary patterns associated with introductions is unclear, as the majority of studies on EICA focus on a single invaded range (e.g. Blossey & Notzold, 1995; Joshi & Vrieling, 2005; Hodgins & Rieseberg, 2011; Uesugi & Kessler, 2016, but see Colomer-Ventura *et al.*, 2015). As EICA predicts the evolution of competitive ability following enemy escape is repeated among introductions, tests across multiple introduced ranges are essential to reveal the general applicability of this hypothesis.

Biotic and abiotic clines governing plant resistance within ranges (Endara & Coley, 2011; Moles *et al.*, 2011a) can also obscure the adaptive underpinnings of trait divergence among ranges experienced during invasion. For instance, herbivore pressure in the native range is expected to increase towards lower latitudes and potentially drive clines in plant defence in some species (Moles *et al.*, 2011a). This clinal pattern may be absent in the introduced range due to overall lack of herbivory, resulting in non-parallel defence gradients between ranges (e.g. Cronin *et al.*, 2015; Allen *et al.*, 2017). Moreover, high-resource environments support plant species with faster growth that are more vulnerable to herbivores (Coley *et al.*, 1985; Zandt, 2007; Endara & Coley, 2011), resulting in latitudinal clines in defence traits. Latitudinal clines in resource availability could subsequently lead to the evolution of high growth and reduced chemical defences at lower latitudes (Woods *et al.*, 2012; Moreira *et al.*, 2014), although this interspecific pattern may have limited application to intraspecific variation (Hahn & Maron, 2016, but see Woods et al., 2012). However, taken together these patterns suggest that the evolutionary consequences of herbivore escape could change along latitudinal gradients (Blumenthal, 2006). Geographical clines therefore need to be considered in tests of adaptive divergence between ranges (Colautti *et al.*, 2009).

Herbivore defence may also evolve indirectly if defensive traits are genetically correlated with other traits that are under strong selection. For example, the growth-defence trade-off hypothesis (Coley *et al.*, 1985), a fundamental component of EICA, suggests a negative genetic correlation between plant growth and defence-related traits, because allocation of limited resource to one trait necessarily reduces resource available for the other. Nevertheless, empirical demonstrations of such trade-offs are rare (Uesugi *et al.*, 2017), and may often be obscured by genetic variation in resource acquisition (Van Noordwijk & de Jong, 1986). High variation in resource acquisition among genotypes from across the species range, such as

elevated resource returns of larger individuals (Van Noordwijk & de Jong, 1986) or individuals with high specific leaf area (SLA, Poorter & Remkes, 1990), is expected to mask the underlying trade-offs and result in a positive correlation between growth and defence (Agrawal, 2011; Züst & Agrawal, 2017). To test for the evolution of defence traits driven by allocation trade-offs, we thus need to assess how potential traits governing resource acquisition could impede detection of allocation trade-offs between growth and defence.

The frequency and level of attack could lead to various defence responses (Orrock *et al.*, 2015; Bixenmann *et al.*, 2016), expected to trade off due to their costs and redundancy (Koricheva *et al.*, 2004; Agrawal *et al.*, 2010). Predictable and strong attack should favour constitutive defence, whereas low, infrequent herbivory would favour no, or an inducible response (Agrawal & Karban, 1999; Ito & Sakai, 2009). These responses have been shown to vary over latitudinal clines within ranges (Moreira *et al.*, 2014 Rasmann & Agrawal, 2011). However, the studies exploring evolutionary shifts between native and introduced ranges showed mixed results (e.g. Cipollini *et al.*, 2005; Eigenbrode *et al.*, 2008). Various responses are predicted to result from a decrease in the intensity and frequency of herbivory following introduction (Maron & Vilà, 2001; Agrawal & Kotanen, 2003), including an increase in plasticity (Cipollini *et al.*, 2005; Lande, 2015) or high variability in inducible response among populations (Eigenbrode *et al.*, 2008). Testing such shifts in invasive species would provide insight into the evolution of induced/constitutive trait defence more generally.

Ambrosia artemisiifolia is a highly suitable system to study adaptive divergence in defence-related traits during invasion. This native North American weed has successfully established globally (Oswalt & Marshall, 2008), including recent introductions to Europe (~160 years ago Chauvel *et al.*, 2006) and Australia (~80 years ago; Palmer & McFadyen, 2012; van Boheemen *et al.*, 2017b). Repeated clinal associations were found in *A. artemisiifolia* populations included in the current study, with declines in size and increase in SLA at higher latitudes (van Boheemen *et al.*, 2018a), though differences occurred among ranges. At comparable latitudes, European plants were bigger and had lower SLA than natives, while Australian plants had higher SLA leaves (van Boheemen *et al.*, 2018a).

We test for adaptive trait divergence in 1) physical defence (trichome density), 2) chemical defence (phenolic compounds concentration and richness), and 3) inducibility of chemical defence among the native North American and introduced European and Australian ranges. Trichomes are found on the leaves and the stems of plants and deter herbivores (Kessler & Baldwin, 2002; Dalin *et al.*, 2008; Tian *et al.*, 2012). Phenolics are secondary metabolites that are often thought to confer resistance against herbivores (Bhattacharya *et al.*, 2010; War *et al.*, 2012; War *et al.*, 2018). These compounds are also known to be inducible in response to herbivore damage, as well as simulated herbivory treatments including wounding and methyl jasmonate (MeJA) applications (e.g. Lee *et al.*, 1997; Constabel & Ryan, 1998; Keinänen *et al.*,

2001; Heredia & Cisneros-Zevallos, 2009). We accounted for non-adaptive genetic differences among populations potentially influencing trait variation using >11,000 double-digest genotype-by-sequencing SNPs. Moreover, we controlled for defence-related trait variation along latitudinal clines.

We predict reduced constitutive defence within the introduced ranges together with elevated inducible response due to lower certainty of attack (Cipollini *et al.*, 2005) and a more plastic (inducible) response in recent colonisations (Lande, 2015). We expect non-parallel defence gradients between native and introduced ranges due to divergence of clines in herbivory (Moles *et al.*, 2011b) and/or variable resource gradients (Blumenthal, 2006; Hahn & Maron, 2016). Finally, we explored the association between defence-related trait divergence and divergence in growth and SLA among ranges as EICA predicts greater growth follows from reduced defence, whereas greater defence could be facilitated by enhanced resource acquisition. By considering the complex interplay of the evolutionary mechanisms shaping defence divergence among multiple ranges, we effectively test adaptive evolutionary changes in herbivore defence.

Methods

Study species

Ambrosia artemisiifolia is a highly invasive, monoecious, self-incompatible annual plant (Brandes & Nitzsche, 2006), most commonly found in disturbed habitats (Bassett & Crompton, 1975; Lommen *et al.*, 2017) and is expected to expand its range with ongoing climate change (Chapman *et al.*, 2014). It is the leading cause of hay fever worldwide (Taramarcaz *et al.*, 2005) and has a significant impact on crop yields (Kazinczi *et al.*, 2008). Within Europe, admixture following multiple introductions from distinct native sources has contributed to the success of introduced populations, and genetic variation equals levels observed in North America (Chun *et al.*, 2010; Gladieux *et al.*, 2010; Gaudeul *et al.*, 2011; van Boheemen *et al.*, 2017b). A subsequent single bottlenecked introduction from Europe has been determined as the origin of the Australian invasion, although the exact European source is unknown (van Boheemen *et al.*, 2017b).

Within the native range, around 450 herbivores have been associated with *Ambrosia* species, of which about 30% are specific to the *Ambrosia* genus (Gerber *et al.*, 2011). The North American n ative specialist *Ophraella communa* is shown to exert high levels of damage (Throop, 2005). Up to 50 polyphagous insect species have been associated with *A. artemisiifolia* in Europe, yet most cause little damage (Gerber *et al.*, 2011; Essl *et al.*, 2015). *Ophraella communa* has been sighted in Southern Switzerland and Northern Italy since 2013 (Müller-Schärer *et al.*, 2014), where it greatly affects *A. artemisiifolia* seedling survival and growth (Cardarelli *et al.*, 2018). In Australia, generalists *Zygogramma bicolorata* (leaf-feeding)

and *Epiblema strenuana* (stem-boring) are widespread and seemingly exert some control (Palmer & McFadyen, 2012).

Experimental set-up

To explore the divergence of constitutive quantitative defence traits between native and introduced ranges ("constitutive-defence experiment"), while accounting for divergence along latitudinal clines, we collected *Ambrosia artemisiifolia* seeds in 2013-2014 from broad geographical scales within the native North America and introduced Europe and Australia. We raised seedlings in a common garden (for detailed methods, see Supporting Information). Briefly, we stratified seeds for 6 weeks at 4°C (Willemsen, 1975). After a 2-week germination at 30°C with 12h light/dark cycle, we randomly transplanted into 100ml kwikpot trays with Debco mix, followed by a second transplant to 0.7L pots containing Debco and 1.5ml slow-release fertilizer (Osmocote Pro, eight to nine months) one month later. We top-watered all plants and artificially manipulated daylight following the light cycle at the median latitude for all populations (47.3°N). To explore constitutive defence, we selected a seedling from four maternal lines, originating from 28 North American, 32 European and 20 Australian locations (Supporting Information, Table S1).

A separate greenhouse experiment was conducted to test whether the inducibility of defence response varied among plant origins (hereafter, "induction experiment"). We used a subset of populations used in the constitutive experiment (10 North American, 17 European and 12 Australian locations, Table S1). For each population, we selected four maternal lines, and grew two seedlings per line as above. One seedling per maternal line was allocated to either the control or simulated herbivory treatment. We simulated herbivory by vertically cutting off half of the newest fully formed leaf (wounding) and subsequently spraying the whole plant with 1mM methyl jasmonate (MeJA) (Campos-Vargas & Saltveit, 2002; Heredia & Cisneros-Zevallos, 2009). Control plants were not wounded and were sprayed with distilled water.

Trait measurements

For the constitutive experiment, we recorded trichome density at the mid-point of each plant under a dissecting microscope (Olympus, SZ-PT) using a 1 cm x 0.3 cm stem area at the mid-point of each plant, nine weeks after the second transplant. Three weeks later, we scanned one young, fully expanded leaf from each plant and calculated leaf area using ImageJ and the R package LeafArea (Katabuchi, 2015). We dried leaves at 45 °C for seven days and an addition 12 hours prior to weighing and weighed to the closest milligram. We calculated specific leaf area (SLA) by dividing leaf area by dry leaf weight (mm^2/mg). We deconstructed plants for biomass measurements once the majority of seeds had ripened. We placed aboveground components in paper bags and dried these in ovens at 45 °C for at least 36 hours. Before dry

weight biomass measures, we dried materials for an additional minimum of 24 hours to ensure the dry weight was constant at the time of measuring and it was not variable due to humidity in the air or incomplete drying. We weighed this shoot biomass to the closest 0.1 gram.

Leaf samples for phenolic analyses were collected four weeks after the second transplant by clipping approximately 200 mg of the newest fully expanded leaf, which was flash frozen in liquid nitrogen and stored in a -80 ˚C. In the induction experiment, we collected leaf samples 24 hours after the final treatment. Samples were extracted in 1 ml of 80% methanol (% by volume in water) using a Qiagen TissueLyser II for 30 seconds at 30 rps twice and centrifuged for 30 minutes at 2700 rpm. Phenolic samples from the constitutive-defence experiment were analysed using HPLC Agilent 1200 series (Agilent Technologies Australia, Mulgrave, VIC, Australia) equipped with C18 reverse-phase column (Waters, 5.0 µm, 250 mm x 4.6 mm; Alltech Australia, Baulkham Hills NSW, Australia). The elution system consisting of solvents (A) 0.25% H3PO4 in water (pH 2.2) and (B) acetonitrile was: 0–6 min, 0–12% of B; 6–10 min, 12-18% of B, and 10-30 min, 18-58% of B, with a flow rate of 1 mL/min and injection volume of 15 µL (Keinänen *et al.*, 2001). Samples from the induction experiment were analysed with Agilent Infinity 1260 equipped with C18 reverse-phase column (Poroshell 120 EC-C18, 2.7 µm, 150 mm x 3.0 mm; Agilent Technologies Australia, Mulgrave, VIC, Australia). The elution method was modified from above and was: 0–2 min, 0–12% of B; 2–3.3 min, 12–18% of B, and 3.3–10 min, 18-58% of B, with a flow rate of 0.5 mL/min and injection volume of 5 µL. In both experiments, phenolic compound peaks were identified to their compound classes using UV spectra and relative abundance was quantified at 320 nm. To estimate phenolic compound richness, we counted the number of detectable peaks. The relative concentration of eight major phenolic peaks was estimated as area under each peak divided by sample fresh weight. Results could not be directly compared as the two experiments were performed in different greenhouses and samples from each experiment were run using different HPLC machines.

Statistical analyses

To test if constitutive defence differed among ranges (the constitutive experiment), we examined individual phenolic compound composition in a multivariate analysis of covariance (MANCOVA) and the concentration of individual phenolic compounds, phenolic compound richness, total phenolic concentration, and trichome density in univariate mixed models. To account for latitudinal variation within and among ranges, each multi- and univariate model included range, latitude, their interaction and a latitude^2 effect as fixed factors. To control for neutral population structure, possibly shaping trait variation between populations, univariate models included q-values as a random effect, as obtained from STRUCTURE analysis performed on genetic data. Within multi- and univariate analyses, we improved normality of

the data by square-root- or log-transforming traits where appropriate. Within the MANCOVA, we included the concentration of eight major phenolic compounds (Spearman's ρ among peaks < 0.75) and calculated Wilks' λ (multivariate F-value) to measure the strength of the associations. To measure the variance explained by the fixed effects or the full model within the univariate models, we calculated the marginal and conditional coefficients of determination using the MuMIn package (Bartón, 2018). We computed type III Wald F-values with Kenward-Roger degrees of freedom and step-wise removed non-significant effects, starting with the highest order interaction. For univariate models, we plotted the partial residuals of each response variable by ranges, thus accounting for latitudinal clines and neutral population genetic structure and reported these adjusted means and standard errors for each range, calculated using the *phia* packaged (De Rosario-Martinez, 2015).

To explore the variation in inducibility among ranges (the induction experiment), we repeated the steps for the constitutive experiment, now including treatment and its interactions with range and latitude as fixed effects. For the MANCOVA, we included five peaks (excluding three with Spearman's ρ > 0.75) to increase power of this test (Scheiner, 2001). We retained treatment in these models, as this was the variable of interest. Here, a significant treatment effect would signify an inducible response, whereas a treatment x range interaction would imply this response differs between ranges. A treatment x latitude interaction would indicate different inducibility at different latitudes. To test if variation in induction differed between ranges (Eigenbrode *et al.*, 2008), we compared the coefficient of variation (c_v) using the modified signed-likelihood ratio test for equality with 10^4 simulations in the cvequality package (Krishnamoorthy & Lee, 2014; Marwick & Krishnamoorth, 2018).

To examine associations between defence-related traits and plant growth and to assess if divergence in individual resource budgets could have resulted in range differences in defence-related trait investment, we tested responses of phenolic richness, phenolic concentrations or trichome density to shoot biomass or SLA. Each model included a defence-related trait as response, with shoot biomass or SLA, range and their interaction as predictors. We used individual trait values and included individual STRUCTURE q-values and sampling location as random factors. We explored significant range x defence interactions using a Holm p-value correction in the *phia* package (De Rosario-Martinez, 2015). In these models, a negative association between defence-related traits and shoot biomass would suggest a trade-off, while a positive one might indicate differences in resource acquisition. Range differences at similar values of shoot biomass or SLA would indicate defence-related trait divergence independent of genotypic differences in individual resource budgets.

To explore if constitutive and inducible defence trade off, we first calculated the induced level of total phenolics for each maternal line as the difference between damage and control treatments of the two half-sibs. This estimate of induction is thought to reduce correlations with

control treatment estimates and thus the collinear associations (e.g. due to genotypic biases) will not mask the trade-off associations (Morris *et al.*, 2006). We included population of origin and individual q-values as random factor in these models. A significant negative association between induced and constitutive levels of phenolic concentration and richness would indicate the presence of a trade-off. All statistical analyses were conducted in R v3.4.3 (R Core Team, 2017).

Results

Constitutive defense trait divergence between ranges

We found significant range divergence in constitutive phenolic composition (Table 1a), resulting from differences between the introduced Europe and the native North America ($F_{8,66}$=3.280, p=0.010, Wilks' λ=0.716) (Table 1b).Total phenolic concentration was similar between the native and European populations, but 28% lower in Australia (Table 1b&c, Fig. 1). Phenolic peak richness differed among ranges (Table 1): it was highest in the introduced European range (adjusted mean of 43 peaks) followed by the native North American (40 peaks) and introduced Australian ranges (33 peaks). Trichome density showed no differences between ranges (Table 1a, Fig. 1). The composition of individual phenolic compounds and peak richness depended on latitude, though no such effect was found for the total phenolic concentration or trichome density (Table 1a, Fig. 2). We did not observe range x latitude interactions for any of the defence-related traits (Table 1a), suggesting latitudinal clines, when present, did not differ between ranges.

Table 1. *Ambrosia artemisiifolia* defence-related trait responses (population means) to range, latitude, their interaction and latitude^2 in the constitutive experiment in multivariate (individual phenolic compounds) and univariate analyses (a), with dissection of significant range effects (p<0.05) in post-hoc tests (b) and model adjusted means (standard error) for each range (c). We reported Wald type III F (a) or χ^2 test values (b), Kenward-Roger degrees of freedom (subscript), significance (symbols). In the multivariate analysis, Wilk's λ measure the strength of the association, in univariate analyses, marginal (R^2m) and conditional (R^2c) coefficients measure the variance explained by fixed effects or full models. Models were step-wise reduced starting with the highest order non-significant interaction and univariate analyses included neutral population genetic structure as a random effect. *ns: p>0.1; #: p<0.1, *: p<0.05; **: p<0.01; ***: p< 0.001*

a.	Range	Latitude	Latitude^2	Range:Latitude	R^2m	R^2c
Individual phenolic compounds composition	$4.520_{16,132}$***, λ =0.417	$6.928_{8,66}$***, λ=0.544	$2.814_{8,66}$*, λ =0.746	$0.849_{16,128}$(ns), λ =0.817		
Phenolic concentration	$8.601_{1,71.918}$***	$0.934_{1,73.244}$(ns)	$0.036_{1,72.973}$(ns)	$0.127_{1,70.054}$(ns)	0.189	0.290
Phenolic richness	$7.615_{2,58.48}$**	$7.79_{1,71.66}$**	$0.046_{1,53.51}$(ns)	$2.027_{2,69.64}$(ns)	0.177	0.688
Trichome density	$0.663_{2,71.183}$(ns)	$1.825_{1,66.941}$(ns)	$3.148_{1,74.991}$#	$0.121_{2,69.141}$(ns)	0.055	0.055

a.	Range	Latitude	Latitude^2	Range:Latitude	R^2m	R^2c
Individual phenolic compounds composition	$4.520_{16,132}$***, λ =0.417	$6.928_{8,66}$***, λ=0.544	$2.814_{8,66}$*, λ =0.746	$0.849_{16,128}$(ns), λ =0.817		
Phenolic richness	$7.615_{2,58.48}$**	$7.79_{1,71.66}$**	$0.046_{1,53.51}$(ns)	$2.027_{2,69.64}$(ns)	0.177	0.688
Phenolic concentration	$8.601_{1,71.918}$***	$0.934_{1,73.244}$(ns)	$0.036_{1,72.973}$(ns)	$0.127_{1,70.054}$(ns)	0.189	0.290

Trichome density	$0.663_{2,71.183}$(ns)	$1.825_{1,66.941}$(ns)	$3.148_{1,74.991}$#	$0.121_{2,69.141}$(ns)	0.055	0.055

b.

	North America - Europe	North America - Australia	Europe - Australia
Individual phenolic compounds composition	$3.280_{8,66}$**, λ =0.716	$1.580_{8,66}$(ns), λ=0.840	$1.994_{8,66}$(ns), λ =0.805
Phenolic concentration	1.397_{1}(ns)	9.311_{1}**	17.321_{1}***
Phenolic richness	4.783_{1}*	12.725_{1}***	15.843_{1}***

c.

	North America	Europe	Australia
Phenolic concentration	14.524(0)	15.824(0)	10.928(0)
Phenolic richness	40.72(0.869)	43.503(1.09)	32.699(1.935)
Trichome density	99.893(4.607)	97.973(4.226)	113.437(6.524)

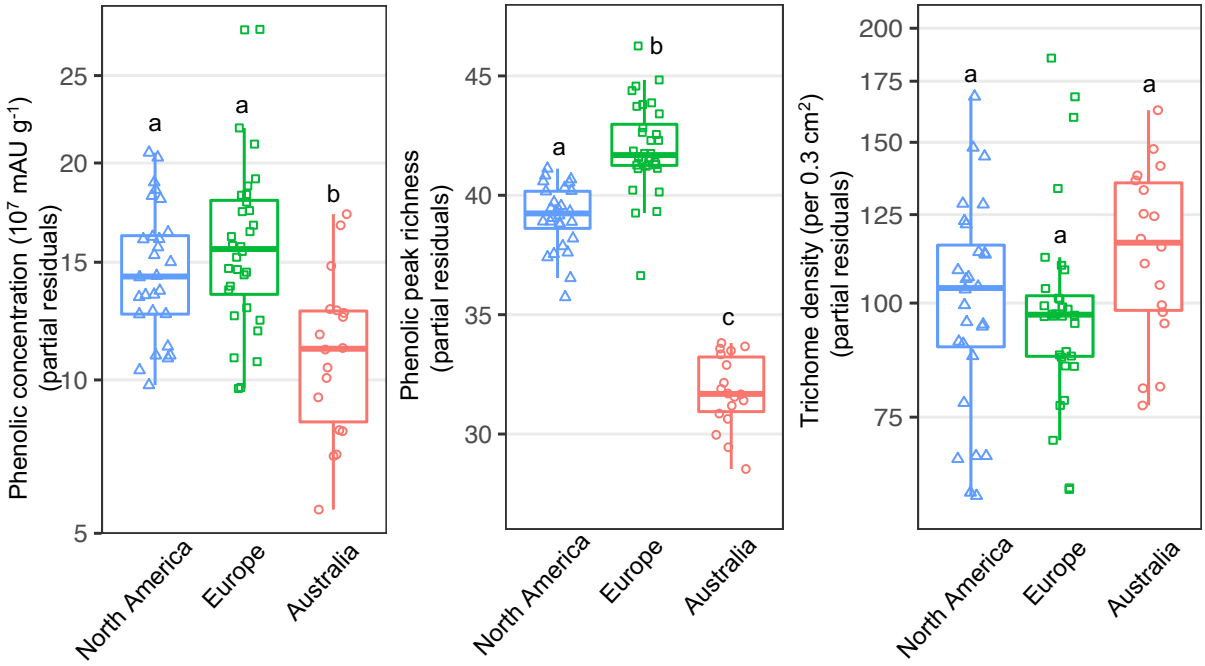


Fig. 1. Partial residual defence trait responses (phenolic concentration, peak richness and trichome density) of *A. artemisiifolia* populations to range, accounting for latitudinal clines and neutral population structure. Different letters indicate significance for pairwise range comparisons (Table 1).

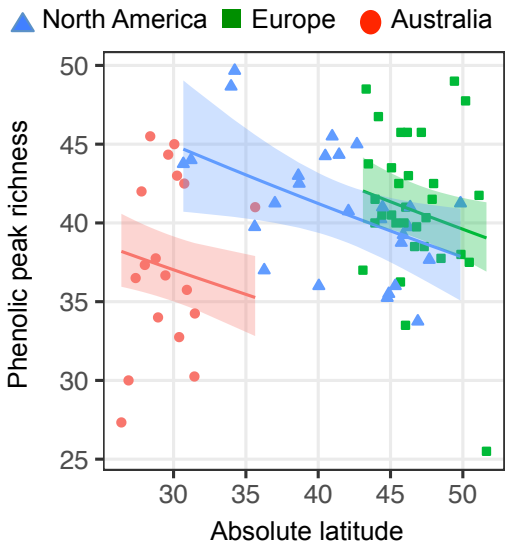


Fig. 2. Population mean response of phenolic peak richness to range (native North America, blue triangles; introduced Europe, green squares; introduced Australia, red circles) and latitude in *Ambrosia artemisiifolia,* with predicted latitudinal clines (+/- 95% confidence interval) corrected for neutral population structure.

Table 2. *Ambrosia artemisiifolia* defence-related trait responses (population means) to range, latitude, treatment, their interactions and latitude^2 (a) and model adjusted means (standard error) for each range (b) in the inducible experiment in multivariate (individual phenolic compounds) and univariate analyses. Range, latitude, their interaction or latitude^2 were included as covariates and significant results were not explored further. We reported Wald type III F, Kenward-Roger degrees of freedom (subscript), significance (symbols) (a). In the multivariate analysis, Wilk's λ measure the strength of the association, in univariate analyses, marginal (R^2m) and conditional (R^2c) coefficients measure the variance explained by fixed effects or full models (a). Models were step-wise reduced starting with the highest order non-significant interaction and univariate analyses included neutral population genetic structure as a random effect. *ns: p>0.1; *: p<0.05; **: p<0.01; ***: p< 0.001*

a.

	Range	Latitude	Latitude^2	Range:Latitude	Treatment	Treatment:Range	Treatment:Latitude	Treatment:Range: Latitude	R^2m	R^2c
Individual phenolic compounds concentration	$7.591_{10,118}$***, λ=0.370	$10.637_{5,59}$***, λ=0.526	$4.818_{5,59}$**, λ=0.710	$2.311_{10,118}$*, λ=0.699	$12.014_{5,59}$***, λ=0.496	$0.326_{10,112}$(ns), λ=0.944	$0.977_{5,58}$(ns), λ=0.922	$1.357_{10,108}$(ns), λ=0.789		
Phenolic concentration	$4.970_{2,30.905}$*	$5.932_{1,31.505}$*	$1.577_{2,28.556}$(ns)	$1.745_{1,31.505}$(ns)	$4.241_{1,35.628}$*	$0.005_{2,32.285}$(ns)	$1.077_{1,35.428}$(ns)	$1.417_{2,31.192}$(ns)	0.201	0.457
Phenolic richness	$3.764_{2,29.93}$*	$4.700_{1,31.08}$*	$1.340_{2,28.24}$(ns)	$6.030_{1,30.95}$*	$0.850_{1,35.33}$(ns)	$0.091_{2,32.1}$(ns)	$0.825_{1,34.89}$(ns)	$1.923_{2,30.78}$ (ns)	0.323	0.723

b.

Range	Native		Europe		Australia	
Treatment	Control	Wounding + MeJa	Control	Wounding + MeJa	Control	Wounding + MeJa
Phenolic concentration	52.803 (10.406)	44.618 (10.407)	69.724 (10.404)	61.539 (10.405)	19.949 (10.402)	11.764 (10.403)
Phenolic richness	24.06 (1.934)	23.261 (1.943)	30.616 (2.18)	29.816 (2.189)	18.569 (3.359)	17.77 (3.354)

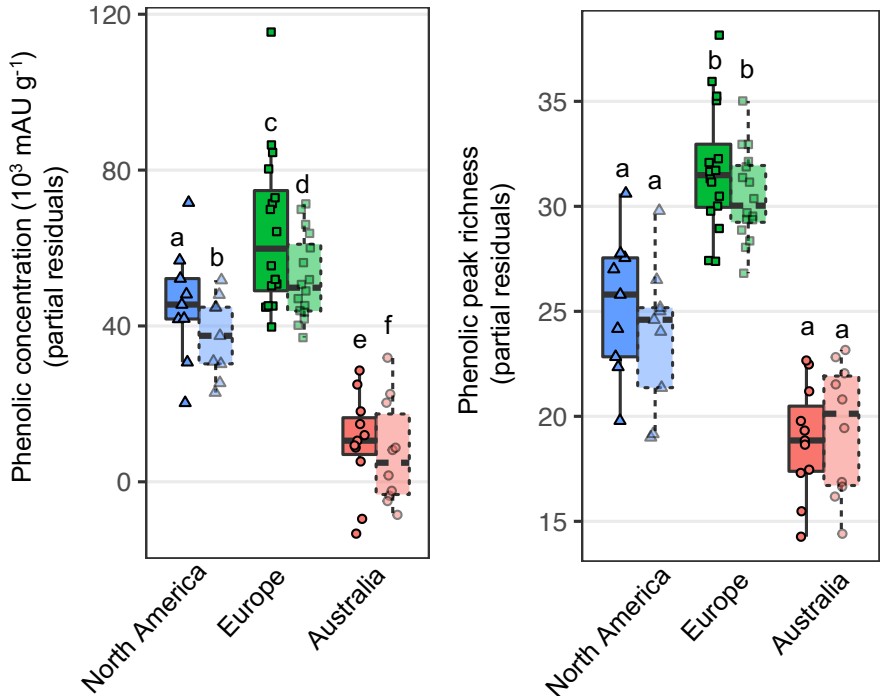


Fig. 3. Partial residual defence trait responses (phenolic concentration and peak richness) of *A. artemisiifolia* populations to control (solid symbols) and herbivore simulating treatment (wounding + MeJA, dashed transparent symbols), with covariates of range, accounting for latitudinal clines and neutral population structure. Letters indicate significance of effect (Table 2).

Inducible defence trait divergence between ranges

We found a significant treatment effect on individual phenolic compound composition in the induction experiment ($F_{5,59}$=12.014, p<0.001, Wilks' λ=0.496) (Table 2). The total phenolic concentration was slightly suppressed in the herbivory-simulating treatment (Table 2, Fig. 3), but the phenolic peak richness did not show a response to experimental treatment (Table 2, Fig. 3). We identified no treatment x range x latitude interactions (Table 2, Fig. 3), suggesting there is no range difference in inducibility clines. Also, the absence of treatment x latitude interactions (Table 2, Fig. 3), suggests an overall lack of latitudinal clines in inducibility. Moreover, no treatment x range interactions (Table 2, Fig. 3) suggests the inducible response did not differ between ranges. We did not find range differences in the variation of inducible phenolic peak richness (c_v=1.401, p=0.496) or concentration (c_v=2.297, p=0.317).

Associations between defence, biomass and specific leaf area (SLA)

Within each range, phenolic concentration and richness was positively correlated with shoot biomass, whereas we found a negative association between trichome density and shoot biomass (Table 3, Fig. 4). We found high-SLA leaves had lower phenolic concentration and peak richness, yet higher trichome density (Table 3, Fig. 4). No interactions were significant between range and predictor variables (shoot biomass or SLA), suggesting these associations among

traits were consistent between ranges (Table 3, Fig. 4). These results emphasize the close relationship between plant growth, physiology and defence.

When controlling for shoot biomass or SLA, total phenolic concentration in European plants was higher compared to North American individuals of comparable weight (Table 3, Fig. 4), whereas no difference existed in latitude models (Table 1, Fig. 1). Conversely, phenolic peak richness was no longer significantly different between North America and Europe (Table 3, Fig. 4) compared to range comparisons accounting for latitude (Table 1, Fig. 1). Australian plants exhibited lower phenolic concentration and peak richness compared to native or European plants of comparable weight or SLA. Yet, at the same plant weight, Australian plants had higher trichome densities than in the other ranges (Table 3, Fig. 4). These patterns match previous analyses including latitude (Table 1, Fig. 1).

Table 3. Constitutive defence trait response of *Ambrosia artemisiifolia* individuals to shoot biomass, specific leaf area and their interaction with range (a), with dissection of significant range effects ($p<0.05$) in post-hoc tests (b) and model adjusted means (standard error) for each range (c). We reported corresponding figure, Wald type III F (a) or χ^2 test values (b), Kenward-Roger degrees of freedom (subscript) and significance (symbols). Marginal (R^2m) and conditional (R^2c) coefficients measure the variance explained by fixed effects or full models (a). Models were step-wise reduced starting with the highest order non-significant interaction and included population origin and neutral population genetic structure as random effects.

*ns: p>0.1; #: p<0.1, *: p<0.05; **: p<0.01; ***: p< 0.001*

a.

Predictor	Response	Figure 4	Range	Predictor	Range:Predictor	R^2m	R^2c
Shoot biomass	Phenolic concentration	A	$19.321_{2,79.435}$***	$31.098_{1,181.299}$***	$1.441_{2,184.26}$(ns)	0.160	0.189
	Phenolic richness	B	$18.96_{2,79.29}$***	$53.389_{1,180.51}$***	$1.565_{2,187.82}$(ns)	0.199	0.224
	Trichome density	C	$10.242_{2,79.4}$***	$18.49_{1,174.06}$***	$0.525_{2,180.84}$(ns)	0.093	0.103
Specific leaf area	Phenolic concentration	D	$6.162_{2,71.167}$**	$38.464_{1,208.912}$***	$0.98_{2,202.53}$(ns)	0.202	0.269
	Phenolic richness	E	$5.349_{2,69.16}$**	$42.692_{1,217.97}$***	$0.32_{2,202.6}$(ns)	0.215	0.331
	Trichome density	F	$1.828_{2,71.7}$(ns)	$10.994_{1,204.66}$**	$1.196_{2,206.12}$(ns)	0.049	0.121

b.

Predictor	Response	Figure 4	North America - Europe	North America - Australia	Europe - Australia
Shoot biomass	Phenolic concentration	A	5.846_1*	22.155_1***	40.087_1***
	Phenolic richness	B	2.546_1(ns)	26.667_1***	37.964_1***
	Trichome density	C	2.255_1(ns)	13.049_1***	21.047_1***
Specific leaf area	Phenolic concentration	D	2.152_1(ns)	5.032_1*	12.485_1**
	Phenolic richness	E	0.459_1(ns)	6.655_1*	10.643_1**
	Trichome density	F	-	-	-

c.

		Figure 4	North America	Europe	Australia
Shoot biomass	Phenolic concentration	A	13.881(0)	16.502(0.001)	8.246(0)
	Phenolic richness	B	40.985(0.764)	42.671(0.751)	33.863(1.133)
	Trichome density	C	97.429(20.172)	89.073(18.13)	126.099(39.006)
Specific leaf area	Phenolic concentration	D	13.581(0.001)	15.313(0)	10.367(0.001)
	Phenolic richness	E	40.345(0.967)	41.182(0.817)	36.106(1.325)
	Trichome density	F	99.412(26.303)	96.815(21.298)	113.845(42.536)

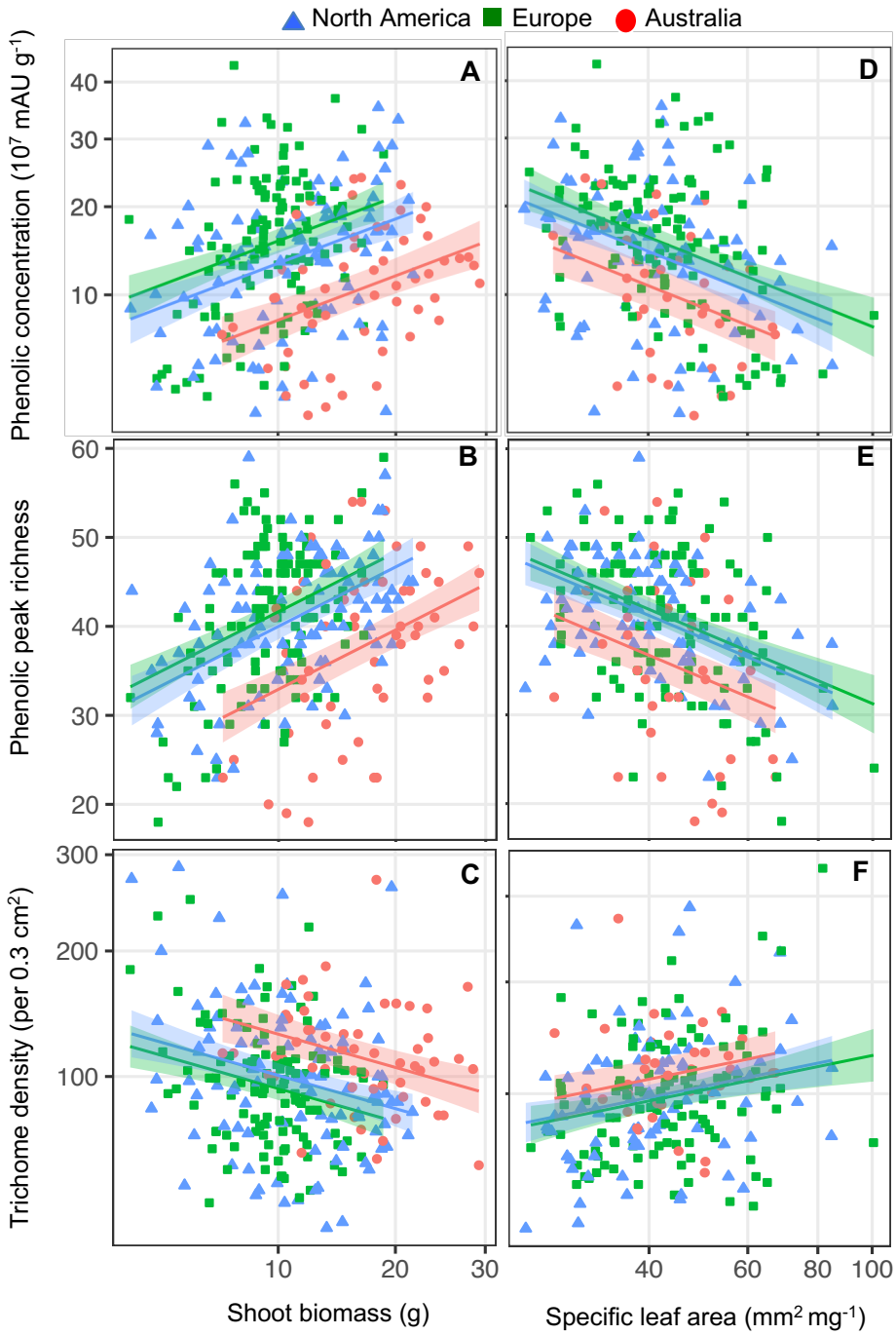


Fig 4. Defence trait responses (phenolic concentration, peak richness and trichome density) of *A. artemisiifolia* individuals to range (native North America (blue triangles); Europe (green squares); Australia (red circles)), shoot biomass (left panels) or specific leaf area (right panels) with model predictions (+/- 95% confidence interval, Table 4).

Constitutive-inducible trade-offs

Induced levels of phenolic concentration and richness, the response variables, were negatively associated with the predictors, the constitutive levels (concentration: $F_{1,141.88}$=76.286, p<0.001; richness: $F_{1,123.81}$=78.126, p<0.001; Fig. 5). We found no range differences in either induced response trait (concentration: $F_{2,31.07}$=0.265, p=0.769; richness: $F_{2,31.518}$=1.719, p=0.196; Fig. 5), nor did we identify interactions between range and the predictor constitutive phenolic concentration ($F_{2,141.11}$=0.866, p=0.423). Range x constitutive phenolic peak richness

($F_{2,129.82}$=3.045, p=0.051) was marginally significant. These results suggest that constitutive and inducible defence trade off, although there is no difference between ranges.

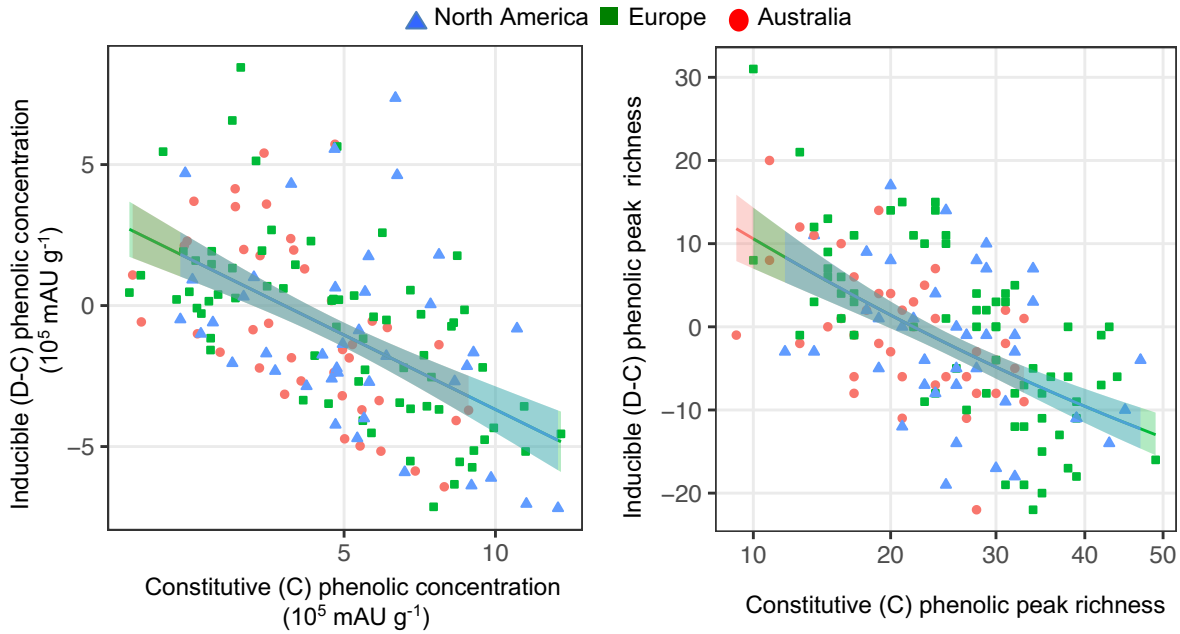


Fig. 5. Inducible (D: wounding + MeJA; C: control) versus constitutive (control) defence trait responses (phenolic concentration and peak richness) of *A. artemisiifolia* populations among ranges (native North America: blue triangles; Europe: green squares; Australia: red circles) with model predictions (+/- 95% confidence interval).

Discussion

In this study, we show the Evolution of Increased Competitive Ability (EICA) hypothesis fails as a general explanation of repeated intraspecific defence-related trait divergence across multiple introduced ranges. Though we observed reduced phenolic concentration and richness in introduced Australia compared to the native plants while controlling for genetic structure, levels were similar or slightly higher in the introduced Europe compared to native populations at comparable latitudes and energy budgets. In addition, trichome density did not differ among ranges. In line with predictions, a trade-off between the constitutive and inducible phenolics was observed together with similar phenolic inducibility among ranges. Inconclusive support for EICA has been shown in inter- (Felker-Quinn *et al.*, 2013) and intraspecific comparisons (Colomer-Ventura *et al.*, 2015). To our knowledge however, this is the first study testing EICA across multiple introductions while exploring the predicted confounding of latitudinal clines, population substructure or genotypic differences in individual resource budgets. Therefore, the apparent absence of the predicted repeated selection against high defence investment following introduction is unlikely to be entirely masked by these factors. We examine these processes in detail and suggest alternative mechanisms driving defence-trait divergence within and among native and introduced ranges.

Divergence in constitutive defence-related traits

We found conflicting patterns of defence-related trait divergence between the native and two introduced ranges. Biotic and abiotic latitudinal clines in focal traits (Colautti *et al.*, 2009), genotypic differences in resource acquisition (Van Noordwijk & de Jong, 1986; Agrawal, 2011; Züst & Agrawal, 2017) and historical contingency (Lee, 2002; Facon *et al.*, 2006; Prentis *et al.*, 2008; Rius & Darling, 2014; Estoup *et al.*, 2016) can obscure trade-offs predicted under EICA. Accordingly, in addition to latitudinal clines in phenolic compound composition and peak richness (Fig. 2), we show trichome density, phenolic concentration and peak richness were strongly associated with plant biomass and specific leaf area (SLA)(Fig. 4). Contrary to EICA predictions, phenolic peak concentration was significantly higher in Europe compared to native North America at comparable shoot biomass, although this difference disappeared when controlling for latitude or SLA. Similarly, phenolic richness was significantly higher in Europe than North America at equivalent latitudes, but this likely reflects the larger size and lower SLA of European plants at similar latitudes (van Boheemen *et al.*, 2018a). However, lower phenolic concentration and peak richness in Australia was still present at similar latitude, biomass or SLA compared to North America. Invasion history in unlikely a major factor in this observed defence-related trait divergence as we accounted for population genetic structure in our analysis.

An adaptive reduction of constitutive defence traits following introduction to Europe and Australia was predicted due to a general release from natural enemies. However, levels of chemical defence-related traits (phenolic concentration and richness) were not consistently lower in introduced ranges compared to native populations. Such unexpected findings could have resulted from variation in contemporary herbivory among introduced ranges. Of particular relevance to the EICA hypothesis are specialist herbivores, as herbivory by specialists, but not necessarily generalists, is hypothesized to consistently decline during invasion (Muller-Scharer *et al.*, 2004; Joshi & Vrieling, 2005; Felker-Quinn *et al.*, 2013). Indeed, introduced Japanese *A. artemisiifolia* populations re-exposed to specialist leaf beetle *Ophraella communa* for >10 years were more resistant than herbivore-free populations (Fukano & Yahara, 2012). However, rapid adaptation to *O. communa* is unlikely to have led to the observed elevated European phenolic concentration and richness, as the seeds used in our experiment were collected in 2014 and this beetle is constrained to southern Europe since introduction in 2013 (Sun *et al.*, 2017).

Alternatively, differences in generalist load between introduced ranges could have resulted in variation in quantitative digestibility-reducing chemicals (e.g. phenolics), which defend against both generalist and specialists (Muller-Scharer *et al.*, 2004). Surveys describe a high diversity of generalist species in Europe (Gerber *et al.*, 2011; Essl *et al.*, 2015) not identified in Australia (Palmer & McFadyen, 2012) suggesting herbivory in this species is higher in Europe than Australia. However, Genton *et al.* (2005b) previously found that

compared to native Ontario, the most common forms of damage (chewing and perforation) together with the generalist herbivore load was reduced in introduced France populations consistent with enemy escape in Europe compared to native North America. Contradicting EICA expectations, but consistent with our findings for Europe, the French plants showed no evolutionary loss of defence (Genton *et al.*, 2005b). Therefore, although reductions in both specialists and generalist herbivores have been documented in both introduced ranges, we did not find parallel changes in defence-related traits as predicted by EICA, suggesting such predictions are perhaps too simplistic. Nevertheless, a more detailed survey of herbivory, resistance and the mechanisms of resistance across all three ranges is warranted, particularly given the contrasting patterns of divergence in phenolics identified among the two introduced ranges.

A key assumption of EICA is a resource allocation trade-off between defence and growth. However, even when these traits have evolved in the EICA predicted direction, negative genetic correlations have yet to be detected (Franks *et al.*, 2008; Schrieber *et al.*, 2017; Hodgins *et al.*, 2018). Furthermore, a direct trade-off might not be evident as resource reallocation from other traits, drawing from the same resource pool, could allow for the elevated investment in defence related traits and growth simultaneously (Züst & Agrawal, 2017; Hodgins *et al.*, 2018). For instance, an analysis of climate niche shifts in *A. artemisiifolia* has revealed that Eurasian and Australasian ranges on average experience warmer, wetter climates compared to the North American range (van Boheemen *et al.*, 2018a). Therefore, reduced investment in abiotic stress tolerance could have allowed for resource reallocation to defence and growth simultaneously. These recently acknowledged complex dynamics underlying competitive ability call for more integrative tests of invasive spread.

The rapid and repeated latitudinal divergence in phenolic compound composition and richness populations suggests direct or indirect selection of latitude-associated factors. Corresponding to our findings, typical reported patterns include high growth and low defence at more productive high-resource (Coley *et al.*, 1985; Zandt, 2007; Endara & Coley, 2011), low-latitude (Woods *et al.*, 2012; Moreira *et al.*, 2014; Hahn & Maron, 2016) environments (Blumenthal, 2006). Native clines in herbivore load could result in such observations, though the predicted herbivore absence following introduction should lead to non-parallel defence clines among native and introduced ranges (Cronin *et al.*, 2015; Allen *et al.*, 2017). However, in our data, latitudinal clines in defence-related traits (phenolic compound composition and peak richness) were parallel, which could suggest consistent patterns of selection with latitude in all three ranges. The absence of the predicted patterns could result from parallel clines in herbivore loads in each range or the presence of alternative evolutionary forces driving latitudinal trait divergence in the multiple ranges. Clinal variation in herbivory is not as

common as previously thought (Moles *et al.*, 2011a) and geographic information on *A. artemisiifolia* herbivore pressure is needed.

Alternatively, latitudinal clines could arise through direct selection on the alternative functions of phenolic compounds, or indirect selection through genetic covariance with traits under climate mediated selection. Climate was previously shown to be a more important driver of trait divergence compared to enemy release (Colautti *et al.*, 2009; Colautti & Barrett, 2013; Colomer-Ventura *et al.*, 2015). Along these lines, climatic differences between the ranges not captured by latitude could contribute to patterns of divergence in Australian defence-related traits. For instance, trichomes protect plants from UV (Bassman, 2004; Hauser, 2014) and selection for this alternate function in high-UV Australia (WHO, 1998) could potentially explain the higher density of trichomes in this range when controlling for plant size. Herbivore exclusion experiments at various latitudes and environments would be required to disentangle how resource availability, herbivory and other climatic factors might interact during invasion and impact the evolution of growth and defence traits.

Constitutive versus induced range divergence

We observed a negative association between constitutive and inducible defence-related traits suggesting a trade-off (Koricheva *et al.*, 2004; Agrawal *et al.*, 2010). A decrease in the level and predictability of attack in the introduced range is expected to cause a reduction in constitutive defence and the maintenance or increase in inducible defence (Cipollini *et al.*, 2005; Orians & Ward, 2010; Lande, 2015). In agreement with this prediction constitutive phenolic levels were reduced in Australia, while inducible response did not differ among ranges. Such maintenance of mean inducibility could result from insufficient herbivore pressure, where a selection-drift imbalance could increase inducible variability (Eigenbrode *et al.*, 2008). Although analysis of neutral markers suggests genetic drift has been particularly strong in Australia (van Boheemen *et al.*, 2017b), we did not reveal any increase in inducible variation. The growing body of literature testing constitutive versus inducible defence in native and introduced ranges frequently report inconsistent results varying from reductions, to maintenance, to increases in either defence (Cipollini *et al.*, 2005; Eigenbrode *et al.*, 2008; Beaton *et al.*, 2011; Carrillo *et al.*, 2012; Cipollini & Lieurance, 2012; Wang *et al.*, 2012; Wang *et al.*, 2013; Fortuna *et al.*, 2014; Gu *et al.*, 2014; Agrawal *et al.*, 2015; Macel *et al.*, 2017) and calls for more detailed research on the costs-benefit trade-offs of the various responses.

Remarkably, we found evidence of a suppression of phenolics in response to herbivore simulation for some populations, especially those with high constitutive levels, in contrast to some previous studies (Lee *et al.*, 1997; Constabel & Ryan, 1998; Keinänen *et al.*, 2001; Heredia & Cisneros-Zevallos, 2009). Conversely, cardenolide suppression was found in various *Asclepias* species at high constitutive levels (Rasmann *et al.*, 2009), though the mechanistic

cause was not discussed (Agrawal *et al.*, 2010). We propose that the retraction of phenolics from damaged leaves could indicate a cost-reducing response when the inducible phenolic compounds have alternative functions (e.g. allelopathic interactions and plant structure; Bhattacharya *et al.*, 2010; Li *et al.*, 2010), or function only in particular aspects of defence response, not induced by the treatment. Nevertheless, gaining insight into such cost-benefit associations might prove difficult due to, for instance, issues identifying and addressing all factors influencing the investment of defence-related traits (Neilson *et al.*, 2013).

Conclusion

We show that escape from natural enemies is not likely the single force driving the evolution of increased competitive ability in this invasive, as enhanced growth in European populations was not in lieu of defence-related trait reduction. Trait evolution in Australian populations, derived from European founders, occurred rapidly (~80 generations), seemingly unconstrained by strong genetic bottleneck identified in this range (van Boheemen *et al.*, 2017b), as measured traits in these two invaded ranges are primarily on opposing ends of the phenotypic spectrum of values. Evidence is growing that adaptation to climate might explain the alarming spread and success of non-natives to a greater extent than release from natural enemies (Colautti *et al.*, 2009; Colautti & Barrett, 2013; Colomer-Ventura *et al.*, 2015). This study emphasizes that intraspecific multi-introduction tests of adaptive trait divergence of invasive species are essential to understand contemporary evolutionary process during range expansion.

Acknowledgements

We would like to thank J. Stephens and A. Wetherhill for sample collection, M. Kourtidou and J. Taylor for greenhouse assistance, K. Nurkowski for genomic analyses and R. Andrew for comments on the MS. A Monash University Dean's International Postgraduate Research Scholarship was provided to LAB, a Monash University Startup Grant to KAH.

Author contributions

All authors developed the project, with data collection and analyses carried out by LAB and SB, refined by AU and KH. All authors discussed the results, contributed to the MS writing and gave final approval for publication.

Data accessibility

Sequence data are available at the National Center for Biotechnology Information (NCBI) Sequence Read Archive under Bioproject PRJNA449949.

Supporting information to Chapter 5

Detailed methods for common garden set-up and data collection

Common garden set-up

In 2013-2014, we sampled seeds from 30 maternal plants along transects at each sampling location following the protocol of Hodgins & Rieseberg (2011). We collected plants separated by at least 1-2 meters to reduce the possibility of collecting closely related plants due to limited seed and pollen dispersal. We followed stratification procedures of Willemsen (1975) and stratified at 4°C in sand moistened with 1% plant preservative mixture (ppm) for 6 weeks. Stratification induces germination by imitating winter conditions (Willemsen, 1975). Following stratification, we sowed seeds on damp filter paper in Petri dishes with 1% ppm. We placed dishes in a 30°C germination chamber with 12h light/dark cycle and watered twice daily with 1% ppm to maintain damp filter paper.

After 14-15 days from the start of germination (22nd and 23rd of April, 2015), we randomly selected two seedlings from each mother and planted these into 30-well kwikpot trays containing 100ml Debco Seed Raising Superior Germinating Mix in a random order. We also transplanted two additional seedlings per maternal line, 1-2 days after this first transplant in case focal seedlings died prior to establishment in the greenhouse. Finally, we transplanted late germinating seedlings (sprouting >14 days after we placed seeds in the germination chamber) 5 days after first transplant. We top-watered all plants by hand twice daily and artificially manipulated daylight following the light cycle at the median latitude over all samples (47.3°N) and adjusted timers fortnightly to accommodate the change in day length. Glasshouse temperature was regulated between 20–30 °C. We performed a second transplant of up to three randomly selected seedlings per maternal family (Table S1) on the 22rd to 24th of May, 2015, 1 month after the first transplant, to 100x135mm pots containing 0.7L Debco Seed Raising Superior Germinating Mix and 1.5ml slow-release fertilizer (Osmocote Pro, eight to nine months, 16% N, 4.8% P, 8.3% K, 5.2% S, 1.2% Mg + trace elements). To explore constitutive defence, we selected one seedling from four maternal lines originating from 28 North American, 32 European and 20 Australian locations (Table S1). A separate greenhouse experiment was conducted to test whether the inducibility of defence response varied among plant origins (hereafter, "induction experiment"). We used a subset of populations used in the constitutive experiment (10 North American, 17 European and 12 Australian locations, Table S1). For each population, we selected four maternal lines, and grew two seedlings per line as above. One seedling per mom was allocated to either the control or simulated herbivory treatment. We simulated herbivory by vertically cutting off half of the newest fully formed leaf (wounding) and subsequently spraying the whole plant with 1mM methyl jasmonate (MeJA)

(Campos-Vargas & Saltveit, 2002; Heredia & Cisneros-Zevallos, 2009). Control plants were not wounded and were sprayed with distilled water. We watered plants twice daily and randomized tray locations weekly.

Genetic data collection and analyses

Four weeks after second transplant, we harvested a 5-7 cm young leaf for DNA extraction and sealed these inside 1.5mL Eppendorf tubes, which were flash-frozen in liquid nitrogen. We stored samples on dry ice before placing them in an -80 °C freezer. We extracted DNA from 29-99mg (mean=72.5mg) of leaf tissue of 861 individuals (84 populations) from the control treatment using the Glass Fiber Plate DNA Extraction Protocol (CCDB, Guelph, ON, Canada) and assessed DNA quantity (>8.5 ng/µl) using a QuBit broad-sensitivity DNA quantification system (Invitrogen, Carlsbad, CA, USA). We performed double-digest genotype-by-sequencing library preparation We added 200 ng of high-quality DNA in 7.2 µL water to 2.0 uL CutSmart Buffer 10x, 0.4 uL Pst-1 HF (NEB), 0.4 uL Msp1. We digested samples for 8h at 37°C, 20 minutes at 65°C. To each reaction, we added 2.0 uL 10x CutSmart Buffer, 4.0 uL 10mM ATP, 0.5uL T4 DNA Ligase, 8 uL H2O, 1uL 10mM common adaptor and 5uL 0.6ng/uL barcoded adaptor. We ligated samples for 3h at 22°C and 20 minutes at 65°C. We mixed all samples with 6144 uL Sera-Mag beads (Thermo Fisher). After 15-minute incubation at room temperature, we allotted samples to 7 1.5mL tubes and placed these in Dyna-Mag 2 (Thermo Fisher) magnet for 4 minutes. We removed clear liquid and washed 3 times using 80% EtOH and once with 100% EtOH. We eluted in 150 uL 10 mM Tris pH 8.0. We amplified 8 reactions each with 3uL of elution and 7.5uL H2O, 12.5 µL KAPA 2x MasterMix, 1uL of 12.5mM each PCR primers f&r. We cycled reactions at 98°C for 1 minute, followed by 20s at 62°C and 30s at 72°C. Following 16 cycles we additionally kept samples at 72°C for 5 minutes. After amplification, we cleaned up 30 µL from each well using the Bioline PCR and Gel kit (Bioline). We eluted the purified product in 30uL buffer. We performed a size selection by running the cleaned PCR product on a 2% agarose gel and removing the 400-600bp fragment. This gel fragment was cleaned up using the Bioline PCR and Gel Kit (Biolin1) and eluted in 20 uL H2O. The eleven ddGBS libraries were paired-end sequenced by the Biodiversity Sequencing Centre at UBC on the Illumina HiSeq 2000 platform, two libraries per lane.

We aligned and filtered raw sequences following van Boheemen *et al.* (2017a). Briefly, SNPs were aligned using BWA-mem (Li & Durbin, 2009) to a draft reference genome for *A. artemisiifolia* (van Boheemen *et al.*, 2017a). We called variants with GATK UnifiedGenotyper (McKenna *et al.*, 2010) and filtered SNPs as follows: quality threshold of a Q-score ≥50; a minimum quality by depth of 2, a maximum Fisher-Strand bias of 60.0, minimum mapping quality rank sum test of -12.5, minimum root mean square mapping quality of 40.0, and a minimum read position rank sum test of -8.0; genotype and variant quality of ≥20, depth of 5-

240 and a minor allele frequency of 0.05. We identified a total of 11,598 polymorphic biallelic SNPs with 50% SNP call rate. We calculated individual heterozygosity (H_O) as the proportion of heterozygous loci out of the total number of called genotypes for each individual across 836 control treatment plants and excluded individuals with >80% non-called genotypes (25 out of 861 individuals).

We inferred population genetic structure with STRUCTURE v2.3.4, a Bayesian clustering method that allocates individuals into clusters on the basis of their genotypes (Pritchard *et al.*, 2000). We used this program to calculate population q-scores for the most likely K, used in subsequent analyses to correct for population structure. From the 11,598 SNPs identified, we selected 1024 unlinked SNPs by shuffling the full SNP table and randomly drawing one SNP from each contig. We ran STRUCTURE on these SNPs for each range, using the admixture model, correlated allele frequencies, no location prior for the number of clusters (K) ranging from 1 to 10, with 20 independent runs per K. We sampled from a uniform prior for alpha, whilst allowing for alpha to vary between clusters, accounting for unequal sample sizes (Wang, 2017) Each run comprised of a burn-in of 200,000 followed by 1,000,000 iterations. We used log probability and delta K statistic to determine the uppermost clustering level (Evanno *et al.*, 2005). We used CLUMPAK (Kopelman *et al.*, 2015) to test for multimodality (non present) for the most likely K (=2).

Table S1. Sampling locations within the native North American and introduced European and Australian *Ambrosia artemisiifolia* ranges, including state or province (North America and Australian) and country (North America & Europe), population ID, geographic coordinates (WGS84) and if sampling location was included in the constitutive (one seedling per maternal line) and/or inducible (two seedlings per maternal line) experiment.

Range	Population ID	State/Country	Latitude	Longitude	Constitutive	Inducible
North America, native	AL	AL, USA	30.6754	-87.5908	yes	
	MP	MS, USA	31.2078	-89.0664	yes	yes
	GA	GA, USA	31.6343	-81.4069	yes	yes
	AR	AR, USA	33.9755	-91.4134	yes	yes
	SC	SC, USA	34.2252	-81.3431	yes	yes
	NC	NC, USA	35.6070	-83.0217	yes	
	TN	TN, USA	36.2678	-86.4606	yes	yes
	MO	MO, USA	37.0064	-94.3501	yes	yes
	KY	KY, USA	38.6259	-85.0074	yes	
	KS	KS, USA	38.6857	-96.4927	yes	yes
	NB	NE, USA	40.0439	-96.3314	yes	yes
	OH	OH, USA	40.4876	-82.7270	yes	yes
	PA	PA, USA	40.9659	-78.1748	yes	
	NY	NY, USA	41.4415	-74.5291	yes	
	MS	MA, USA	42.0883	-72.0962	yes	
	IW	IW, USA	42.6780	-96.5024	yes	yes
	AA8	MN, USA	44.3254	-95.9583	yes	yes
	MN2	ON, Canada	44.4472	-79.8039	yes	yes
	MA	ME, USA	44.7712	-68.9709	yes	
	WI	WI, USA	44.8793	-89.4238	yes	yes
	ON3	ON, Canada	45.3297	-74.8933	yes	yes
	ON2	ON, Canada	45.7445	-77.0294	yes	
	NB1	NB, Canada	45.8786	-66.9785	yes	yes
	AA5	MN, USA	46.2171	-96.0502	yes	yes
	MI	MI, USA	46.3581	-84.8807	yes	yes
	QC2	QC, Canada	46.8867	-70.8684	yes	
	QC3	QC, Canada	47.6788	-69.0220	yes	
	AA2B	MN, Canada	49.8378	-97.3293	yes	yes
Europe, introduced	EU35	Serbia	43.0918	21.9380	yes	
	EU36	Bulgaria	43.3152	24.2598	yes	yes
	EU34	Bulgaria	43.4704	25.6707	yes	yes
	EU37	Serbia	43.9179	20.7331	yes	yes
	EU10	France	43.9324	4.3205	yes	
	EU32	Romania	44.1627	28.5098	yes	
	EU33	Romania	44.4046	26.1348	yes	yes
	EU09	Italy	45.0654	7.5923	yes	
	EU11	France	45.0744	4.7509	yes	yes
	EU31	Romania	45.3740	27.0720	yes	yes
	EU07	Italy	45.47089	8.93683		yes
	EU06	Italy	45.5707	8.7855	yes	yes
	EU27	Romania	45.6870	25.6581	yes	yes
	EU38	Croatia	45.7153	15.6538	yes	
	EU08	Switzerland	45.9309	8.9838	yes	yes
	EU01	Slovenia	46.0361	15.2961	yes	
	EU15	France	46.0522	5.3351	yes	
	EU16	Switzerland	46.1622	6.0094	yes	yes
	EU26	Romania	46.2370	24.8540	yes	
	EU12	France	46.6643	4.3278	yes	
	EU13	France	46.8003	4.9724	yes	yes
	EU02	Hungary	47.1309	16.9031	yes	
	EU03	Hungary	47.3278	19.7307	yes	yes
	EU14	France	47.4556	5.2121	yes	
	EU04	Slovakia	47.8799	18.1545	yes	yes
	EU25	Romania	47.9773	23.0443	yes	
	EU24	Slovakia	48.4892	21.8062	yes	
	EU22	Czech	49.4180	17.9615	yes	yes
	EU30	Poland	49.8685	23.0118	yes	
	EU21	Czech	50.1900	15.0604	yes	yes
	EU23	Poland	50.4430	18.8634	yes	
	EU17	the Netherlands	51.1200	5.8403	yes	yes
	EU20	Germany	51.6328	14.1844	yes	

Range	Population ID	State/Country	Latitude	Longitude	Constitutive	Inducible
Australia, introduced	AU01	NSW	-35.6411	150.1274	yes	
	AU32	NSW	-31.4721	152.6836	yes	
	AU33	NSW	-31.4419	152.4651	yes	yes
	AU29	NSW	-30.9195	152.5825	yes	
	AU30	NSW	-30.7406	152.9150	yes	yes
	AU03	NSW	-30.3881	152.9582	yes	yes
	AU27	NSW	-30.2413	152.5855	yes	
	AU26	NSW	-30.0514	152.9849	yes	yes
	AU04	NSW	-29.6310	153.0368	yes	
	AU24	NSW	-29.4358	152.3848	yes	yes
	AU23	NSW	-28.9263	152.3740	yes	
	AU09	NSW	-28.8688	151.1670	yes	
	AU05	NSW	-28.7668	153.3966	yes	
	AU21	NSW	-28.3870	152.6099	yes	yes
	AU19	QLD	-28.0122	153.1677	yes	
	AU18	QLD	-27.7850	153.2751	yes	
	AU15	QLD	-27.3792	152.8016	yes	yes
	AU12	QLD	-26.8858	152.1369	yes	
	AU13	QLD	-26.3913	152.7937	yes	
	AU11	QLD	-25.3655	152.9156	yes	yes

Chapter 6 – General Discussion

This thesis provides strong evidence for the rapid occurrence of local adaptation on a phenotypic and genomic level in *Ambrosia artemisiifolia.* Comparison of multiple introductions using several lines of evidence for adaptation revealed that adaptive divergence is seemingly unaffected by alternative demographic histories. This is one of the first studies investigating genetic associations and the consequence of this genomic context for the invasion process. Taken together, the chapters of this thesis provide new insight into the underlying genomic basis of local adaptation. In addition, the results strongly suggest that predictable adaptive shifts can occur in <100 generations, thus providing a clue to the evolutionary mechanisms by which invasive species rapidly spread, which can inform management strategies. This thesis exemplifies how testing the genomic basis of repeated adaptation in alien populations can provide important insight into the adaptive potential of species more generally.

Introduction pathways and consequences for adaptation

Distinct demographic processes shape the genetic background of introduced species, affecting their adaptive potential (Lee, 2002; Facon *et al.*, 2006; Prentis *et al.*, 2008; Rius & Darling, 2014; Estoup *et al.*, 2016). In Chapter 2, I used novel statistical tools to disentangle the complex historical spread of *Ambrosia artemisiifolia* from the native range of North America into Europe and Australia. In addition to considering established native populations as sources for these invasions, the reconstruction considered a recent admixed native zone and secondary spread originating from an invaded range as potential introduction sources. I identified the European genotypes established through multiple introductions from distinct native sources (Fig. 1). These highly admixed European populations subsequently acted as a source for the Australian spread. However, the latter was most likely a single, strongly bottlenecked event, resulting in a reduction in the genetic variation in current Australian populations (Fig. 1).

These distinct introduction histories could have affected the evolutionary trajectories of the introduced populations in the European and Australian ranges in distinctive ways. Therefore, the current study system provides an opportunity to examine the effects of distinct non-adaptive evolutionary events on adaptive potential empirically in natural populations. Moreover, this study highlights the necessity to consider admixed populations as a potential source for both initial and subsequent colonizations. For species like *A. artemisiifolia*, which is frequently associated with human-modified habitats and intensive agriculture, this type of human-assisted spread may become increasingly more common. As such distinct introduction pathways may affect evolutionary processes in the introduced ranges, these results advocate the intricate testing of introduction scenarios (Cristescu, 2015) and could mean management effort should be focused on preventing introductions from admixed or multiple divergent source populations.

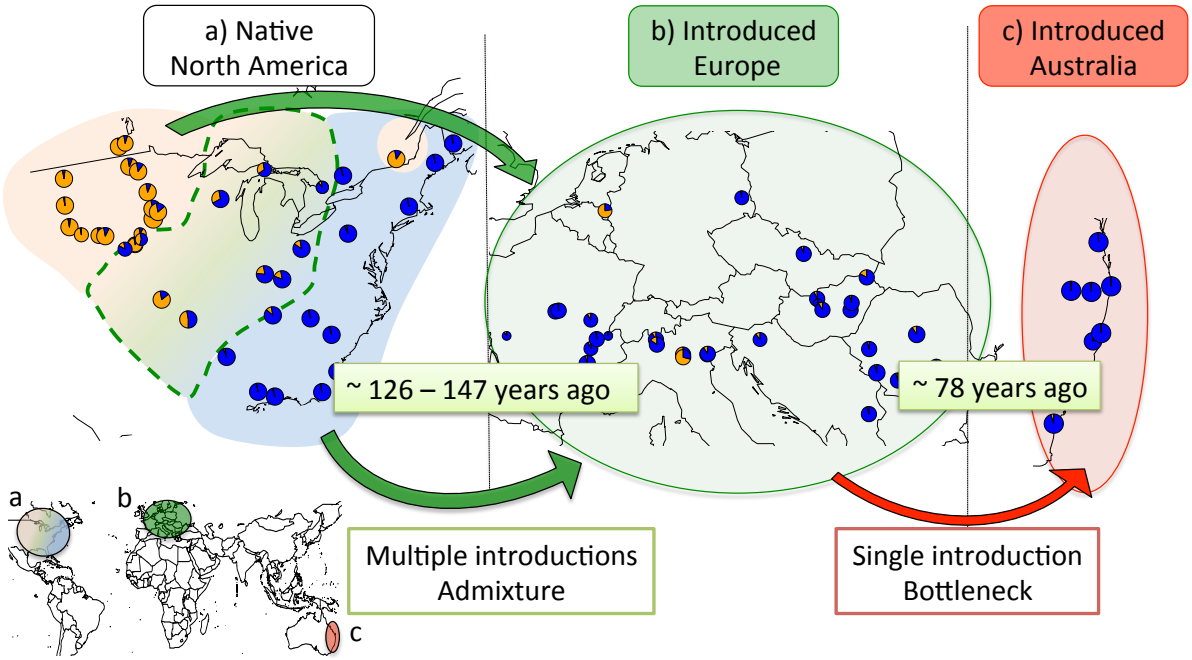


Figure 1. Reconstruction of *Ambrosia artemisiifolia* introduction history from the native North America (a) to Europe (b) and Australia (c) using approximate Bayesian computation with random forest evaluation on 1,022 unlinked genotype-by-sequencing SNPs. Pie charts represent individual cluster assignment at sampling locations by STRUCTURE where K=2. Parameter estimates are medians of posterior distributions.

Rapid repeated phenotypic and genomic adaptation

As the success of introduced species is often associated with rapid evolutionary events (Huey *et al.*, 2000; Lee, 2002; Maron *et al.*, 2004; Lachmuth *et al.*, 2011; Lawson Handley *et al.*, 2011; Colautti & Barrett, 2013; Chown *et al.*, 2014; Turner *et al.*, 2014; Bock *et al.*, 2015; Colautti & Lau, 2015; Oduor *et al.*, 2016; Szűcs *et al.*, 2017; Hodgins *et al.*, 2018) they provide an excellent system to investigate the rate and repeatability of adaptation. In Chapter 3, I explored the extent of repeatability of phenotypes along latitudinal clines within the native North American, introduced European and Australian *A. artemisiifolia* ranges. Latitudinal trait clines in Europe and Australia for all ecologically important traits were parallel to the native range (e.g. maximum plant height, Fig. 2). As neutral genetic differentiation was accounted for (Chapter 2), maternal effects are suspected to be minimal (Hodgins & Rieseberg, 2011) and the reconstructed introduction history (Chapter 2) implies climate matching (Maron *et al.*, 2004) is unlikely, these results indicate rapid adaptation to selective factors closely associated with latitude. These findings are further supported in Chapter 3, with high overlap among the native and invaded ranges in the relative strength and direction of associations among temperature measures and growth, flowering and sex allocation traits.

These results are remarkable as the demographic reconstruction revealed introductions were quite recent (~80-150 generations) and genetic variation in Australia is limited by a genetic bottleneck (Chapter 2). In addition, the Australian populations likely originated from

European introductions (Chapter 2), yet the traits examined in either range were highly dissimilar (Chapter 3). Such rapid evolution has important implications for the evolutionary potential of this species and further range expansion following climate change. For instance, the observed contemporary divergence of reproductive traits such as flowering time, sex allocation and seed size should impact the production of allergenic pollen as well as the abundance and dispersal of seed that could impact spread. Finally, this chapter emphasizes that although introduction can affect genetic diversity of putatively neutral markers the adaptive potential of ecologically important quantitative traits might not be constrained to a similar extent (Dlugosch & Parker, 2008a; Bock *et al.*, 2015; Dlugosch *et al.*, 2015a; Hodgins *et al.*, 2018). Therefore, these results could suggest that confining import of this invasive species into new countries might limit establishment, but would not effectively constrain its adaptive potential once established. Further research is warranted on the relationship between the rate of adaptation and the loss of molecular variation in neutral markers across multiple invasive species, together with investigations on the implications of strong selective sweeps, the effect of mating system, reproductive mode and ploidy on adaptive potential.

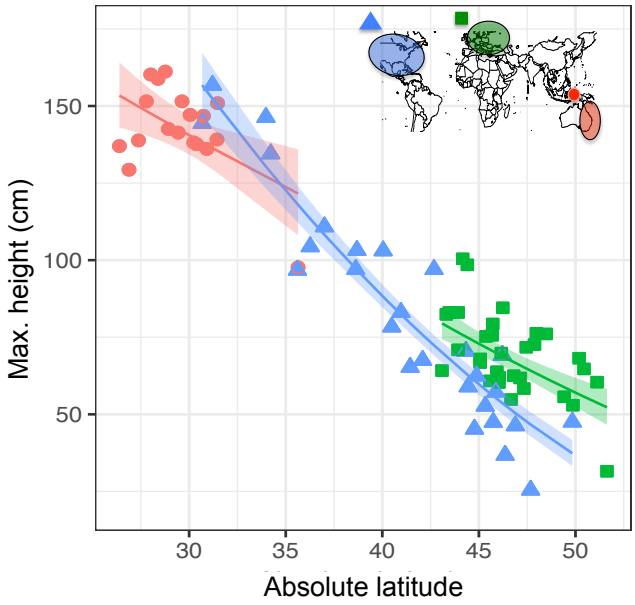


Figure 2. High repeatable phenotypic clines (here maximum plant height) identified in *Ambrosia artemisiifolia* native North American (blue triangles), introduced European (green squares) and introduced Australian (red circles) ranges (A, Chapter 3).

Rapid repeated genomic adaptation

Uncovering the extent of the genetic repeatability underlying repeated phenotypic patterns can provide important insights into biases and constraints on adaptive potential (Weinreich *et al.*, 2006; Gompel & Prud'homme, 2009; Chevin *et al.*, 2010b; Losos, 2011; Connallon & Hall, 2018). In Chapter 4, I delve deeper into the phenotypic results from Chapter 3 and investigate the genomic basis of adaptive differentiation. To identify adaptation at the genomic level, I

combined multiple lines of evidence (Sork *et al.*, 2013; Rellstab *et al.*, 2015) from phenotype-environment associations, F_{ST}-like outlier tests, genotype-environment associations and genotype-phenotype associations. Adaptive signatures were repeated among ranges in a consistent manner, such that 25-41% of putative adaptive loci were shared among ranges (Fig. 3). Moreover, even though the alternative demographic histories left distinct signatures on the genetic associations among loci, these processes seemingly did not impede rapid adaptive changes.

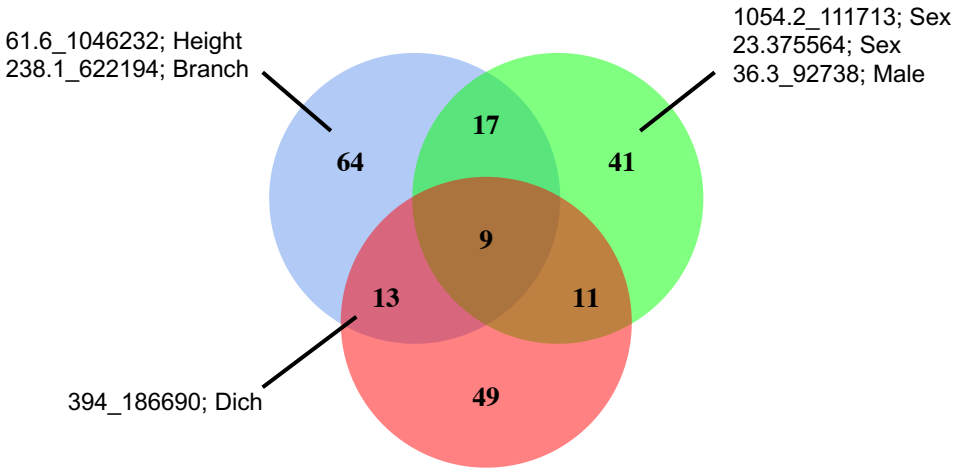


Figure 3. High genomic repeatability among native North American (blue) and introduced European (green) and Australian (red) ranges, as shown by comparing the overlap in outlier regions identified in both F_{ST}-like and genotype-environment associations. Outlier window ID and associated phenotypes are provided when identified in genotype-phenotype associations GPA (phenotype abbreviations correspond to Table S1, Supporting Information; Height=Max. plant height; Branch=Number of branches; Sex=Sex function allocation; Male=Weight of male sex function; Dich=Dichogamy). When only considering traits and environments with the highest phenotype-environment associations (this figure), about 1% of all genomic regions were identified as outlier within each range, yet 25% of these outliers were identified in two or three ranges. When considering all variables, overlap among ranges reached 41% (Chapter 4).

This is one of the first studies showing that adaptive divergence at the genomic and phenotypic level can occur rapidly, with minimum constraints due to demographic processes. A similar exceptional discovery was made recently in sticklebacks, where, despite genetic bottlenecks, 75% of the genomic adaptation of a 12,000 year evolved population was observed in less than 20 generations (Marques *et al.*, 2018). This current study adds to these results by revealing no differences in adaptive potential between highly admixed versus bottlenecks populations, despite distinct signatures on the genetic associations among loci. Therefore, results from this chapter suggest adaptation can be to some extent predictable.

Evolutionary theories to explain the success of invasive species

Chapter 3 revealed trait shifts in European populations compared to natives at equivalent latitudes, while Australian trait values were often outside the phenotypic spectrum observed in the native range (e.g. maximum plant height, Fig. 2). These observations suggest an increase in

competitive ability of the Australian populations following invasion, commonly measured as elevated plant growth and reproductive effort (Felker-Quinn *et al.*, 2013). Indeed, resource allocation trade-offs between life-history traits, such as stress tolerance and growth rate or reproductive output, feature prominently in evolutionary theories developed to explain the success of invasive species (e.g. Hodgins & Rieseberg, 2011; Kumschick *et al.*, 2013; Turner *et al.*, 2014; Colautti & Lau, 2015). In Chapter 5, I explore the well-studied evolution of increased competitive ability hypothesis (EICA), which postulates that release from specialist herbivores within the introduced range favours genotypes allocating resources to growth and reproduction in lieu of defence (Blossey & Notzold, 1995). I found no consistent evidence for the evolution of reduced defence in concert with increase growth following introduction into multiple novel ranges. Instead, adaptation of defence-related trait to climate-associated factors (latitude) was more prominent. Moreover, *A. artemisiifolia* populations in the introduced Eurasian and Australian ranges have experienced a niche centroid shift towards warmer and wetter climates (Chapter 3).

These results suggest that EICA is unlikely to be the sole mechanism determining increased competitive ability. Instead, enhanced growth in introduced populations might reflect to more favourable conditions. Indeed, a recent survey of 815 invasive species show climate niche shifts following introduction are common and alien habitats are generally warmer and wetter (Atwater *et al.*, 2018). Therefore, it might be essential to shift our focus in invasive biology away from enemy escape and towards contemporary climate adaptation, or perhaps to a more complex understanding that considers the combined roles of the biotic and abiotic environment during invasive species trait evolution. The work reported in this thesis highlights the insight gained by comparing multiple introductions to test classical theories of invasion.

Future directions

This thesis breaks ground in revealing remarkable consistency in genomic and phenotypic signatures of adaptive divergence during invasions. However, much more needs to be done to provide insight into the genomic basis of adaptation, the evolutionary biases and constraints that limit adaptation, and how the invasion process affects these. Specifically, future studies should focus on better understanding:

1) The genomic architecture, including the location and size of the genomic regions responding to selection, and the mutational effect size of loci can determine the speed and direction of adaptive shifts and invasive spread (e.g. Dlugosch *et al.*, 2015a), yet theoretical and empirical studies on this topic are lacking. Small effect loci are thought to contribute most to heritable variation in quantitative traits (Pritchard & Di Rienzo, 2010; Pritchard *et al.*, 2010; Rockman, 2012), and adaptation along environmental gradients during range expansion is a function of the genetic variance present for functional quantitative traits (Connallon & Sgrò,

2018; Polechová, 2018). However, adaptation by small effect loci is less likely under selection-drift-migration balance (Yeaman *et al.*, 2018) and small populations with low genetic variance or strong genetic drift can still adapt via few large effect loci (Polechová & Barton, 2015; Gilbert *et al.*, 2017). Thus, loci of large effect are expected be important in adaptive divergence when drift is strong, as might be predicted during the colonization process (Dlugosch *et al.*, 2015a). Consequently, this process should constrain the diversity of genetic forms contributing to adaptation so that the same large effect loci are more likely to contribute repeatedly to adaptation in multiple ranges. However, empirical studies are highly biased towards finding large effect loci and future investigations of small allele frequency shifts, are required to obtain a better understanding of the interaction between mutational effect size, evolvability and parallel evolution. Together with improving statistical power through high sample sizes and high-resolution genomic data, tests could include multivariate (e.g. Lasky *et al.*, 2012; Hecht *et al.*, 2015; Brauer *et al.*, 2016; Sork *et al.*, 2016; Harrisson *et al.*, 2017) or machine learning (e.g. Fitzpatrick & Keller, 2015; Schweizer *et al.*, 2016; Bay *et al.*, 2018) approaches to identify covarying groups of loci responding to environmental variation or covarying allele frequency shifts associated with diverging quantitative traits (Berg & Coop, 2014). Such methods are however still underused (Hoban *et al.*, 2016).

2) Unraveling phenotypic responses to environmental change and the underlying mechanisms is crucial to better predict the future persistence of species (Lässig *et al.*, 2017) and will require identification of trait divergence (this thesis) together with evidence and estimates of selection using reciprocal transplant information and longitudinal field studies. Moreover, assessments of the fitness impact of putative adaptive loci identified in this thesis will provide insight into the genetic basis of adaptation and its role in range expansion.

3) Explicit consideration of all the above factors could alter the outcomes of predictive models of future species distribution in response to climate change. A suite of experimental approaches is currently used to predict phenotypic responses to environmental change, including testing of reaction norms, reciprocal transplants and selection experiments (Hoffmann & Sgrò, 2011; Merilä & Hendry, 2014; Bay *et al.*, 2017). Still, it remains difficult to predict trait responses in complex natural populations from the outcomes of these experiments (Chevin *et al.*, 2010a; Wolkovich *et al.*, 2012). As an alternative, a species' response to ecological change may be predicted with species distribution models using habitat suitability (Guisan & Zimmermann, 2000; Kearney & Porter, 2009; Ehrlén & Morris, 2015; Zurell *et al.*, 2016). However this approach is limited as it assumes constant phenotypic responses to stable environments over time, and assumes that adaptation does not occur (Chevin *et al.*, 2010a; Hoffmann & Sgrò, 2011). Incorporating information about genomic architecture, heritability, strength and variation in selection, plasticity, geographic distribution of alleles and population dynamics into such approaches will allow for high-resolution predictive models.

Enhanced species distribution models will require the combination of a wide range of approaches, including common garden data, demographic studies, plasticity estimates from reaction norms or individual variability in nature, experimental estimates of selection gradients and response dynamics based on the underlying genetic architecture (Merilä & Hendry, 2014; Urban *et al.*, 2016; Bay *et al.*, 2017)

Bibliography

Agrawal, AA (2011). Current trends in the evolutionary ecology of plant defence. *Functional Ecology*, **25**, 420-432.

Agrawal, AA, Conner, JK and Rasmann, S (2010). Tradeoffs and negative correlations in evolutionary ecology. Evolution since darwin: The first 150 years. M. A. Bell, D. J. Futuyma, W. F. Eanes and J. S. Levinton. Sunderland, Massachusetts, USA, Sinauer Associates**:** 243-268.

Agrawal, AA, Hastings, AP, Bradburd, GS, Woods, EC, Züst, T, Harvey, JA and Bukovinszky, T (2015). Evolution of plant growth and defense in a continental introduction. *The American Naturalist*, **186**, E1-E15.

Agrawal, AA and Karban, R (1999). Why induced defenses may be favored over constitutive strategies in plants. The ecology and evolution of inducible defenses. R. Tollrian and C. D. Harvell. Princeton, New Jersey, Princeton University Press**:** 45-61.

Agrawal, AA and Kotanen, PM (2003). Herbivores and the success of exotic plants: A phylogenetically controlled experiment. *Ecology Letters*, **6**, 712-715.

Ahrens, CW, Rymer, PD, Stow, A, Bragg, J, Dillon, S, Umbers, KD and Dudaniec, RY (2018). The search for loci under selection: Trends, biases and progress. *Molecular Ecology*, **27**, 1342-1356.

Allen, WJ, Meyerson, LA, Cummings, D, Anderson, J, Bhattarai, GP and Cronin, JT (2017). Biogeography of a plant invasion: Drivers of latitudinal variation in enemy release. *Global Ecology and Biogeography*, **26**, 435-446.

Allendorf, FW and Lundquist, LL (2003). Introduction: Population biology, evolution, and control of invasive species. *Conservation Biology*, **17**, 24-30.

Anderson, E and Stebbins, G (1954). Hybridization as an evolutionary stimulus. *Evolution*, **8**, 378-388.

Andrew, RL, Bernatchez, L, Bonin, A, Buerkle, CA, Carstens, BC, Emerson, BC, Garant, D, Giraud, T, Kane, NC, Rogers, SM, Slate, J, Smith, H, Sork, VL, Stone, GN, Vines, TH, Waits, L, Widmer, A and Rieseberg, LH (2013). A road map for molecular ecology. *Molecular Ecology*, **22**, 2605-2626.

Arendt, J and Reznick, D (2008). Convergence and parallelism reconsidered: What have we learned about the genetics of adaptation? *Trends in Ecology & Evolution*, **23**, 26-32.

Atwater, DZ, Ervine, C and Barney, JN (2018). Climatic niche shifts are common in introduced plants. *Nature Ecology & Evolution*, **2**, 34-43.

Bailey, SF, Blanquart, F, Bataillon, T and Kassen, R (2017). What drives parallel evolution? *BioEssays*, **39**, 1-9.

Bailey, SF, Rodrigue, N and Kassen, R (2015). The effect of selection environment on the probability of parallel evolution. *Molecular Biology and Evolution*, **32**, 1436-1448.

Barrett, RD and Hoekstra, HE (2011). Molecular spandrels: Tests of adaptation at the genetic level. *Nature Reviews Genetics*, **12**, 767-780.

Barrett, RD and Schluter, D (2008). Adaptation from standing genetic variation. *Trends in Ecology & Evolution*, **23**, 38-44.

Barrett, SC (2015). Foundations of invasion genetics: The baker and stebbins legacy. *Molecular Ecology*, **24**, 1927-1941.

Barrett, SC, Colautti, RI and Eckert, CG (2008). Plant reproductive systems and evolution during biological invasion. *Molecular Ecology*, **17**, 373-383.

Barrett, SC and Hough, J (2012). Sexual dimorphism in flowering plants. *Journal of Experimental Botany*, **64**, 67-82.

Bartle, K, Moles, AT and Bonser, SP (2013). No evidence for rapid evolution of seed dispersal ability in range edge populations of the invasive species *senecio madagascariensis*. *Austral Ecology*, **38**, 915-920.

Bartón, K (2018). Multi-model inference. CRAN, The R Foundation for Statistical Computing**:** R package.

Bassett, IJ and Crompton, CW (1975). The biology of canadian weeds. *Canadian Journal of Plant Science*, **55**, 463-476.

Bassman, JH (2004). Ecosystem consequences of enhanced solar ultraviolet radiation: Secondary plant metabolites as mediators of multiple trophic interactions in terrestrial plant communities. *Photochemistry and Photobiology*, **79**, 382-398.

Bates, D, Maechler, M, Bolker, B and Walker, S (2014). Lme4: Linear mixed-effects models using eigen and s4. *R package version*, **1**, 1-23.

Bay, RA, Harrigan, RJ, Le Underwood, V, Gibbs, HL, Smith, TB and Ruegg, K (2018). Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science*, **359**, 83-86.

Bay, RA, Rose, N, Barrett, R, Bernatchez, L, Ghalambor, CK, Lasky, JR, Brem, RB, Palumbi, SR and Ralph, P (2017). Predicting responses to contemporary environmental change using evolutionary response architectures. *The American Naturalist*, **189**, 463-473.

Beaton, LL, Van Zandt, PA, Esselman, EJ and Knight, TM (2011). Comparison of the herbivore defense and competitive ability of ancestral and modern genotypes of an invasive plant, *lespedeza cuneata*. *Oikos*, **120**, 1413-1419.

Beaumont, MA, Zhang, W and Balding, DJ (2002). Approximate bayesian computation in population genetics. *Genetics*, **162**, 2025-2035.

Benjamini, Y and Hochberg, Y (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289-300.

Berg, JJ and Coop, G (2014). A population genetic signal of polygenic adaptation. *PLoS Genetics*, **10**, e1004412.

Bertorelle, G, Benazzo, A and Mona, S (2010). Abc as a flexible framework to estimate demography over space and time: Some cons, many pros. *Molecular Ecology*, **19**, 2609-2625.

Bhattacharya, A, Sood, P and Citovsky, V (2010). The roles of plant phenolics in defence and communication during *agrobacterium* and *rhizobium* infection. *Molecular Plant Pathology*, **11**, 705-719.

Bierne, N, Roze, D and Welch, JJ (2013). Pervasive selection or is it...? Why are f_{st} outliers sometimes so frequent? *Molecular Ecology*, **22**, 2061-2064.

Bixenmann, RJ, Coley, PD, Weinhold, A and Kursar, TA (2016). High herbivore pressure favors constitutive over induced defense. *Ecology and Evolution*, **6**, 6037-6049.

Blackburn, TM, Lockwood, JL and Cassey, P (2015). The influence of numbers on invasion success. *Molecular Ecology*, **24**, 1942-1953.

Blossey, B and Notzold, R (1995). Evolution of increased competitive ability in invasive nonindigenous plants: A hypothesis. *Journal of Ecology*, **83**, 887-889.

Blum, MG, Nunes, MA, Prangle, D and Sisson, SA (2013). A comparative review of dimension reduction methods in approximate bayesian computation. *Statistical Science*, **28**, 189-208.

Blumenthal, DM (2006). Interactions between resource availability and enemy release in plant invasion. *Ecology Letters*, **9**, 887-895.

Bock, DG, Caseys, C, Cousens, RD, Hahn, MA, Heredia, SM, Hübner, S, Turner, KG, Whitney, KD and Rieseberg, L (2015). What we still don't know about invasion genetics. *Molecular Ecology*, **24**, 2277-2297.

Bollback, JP and Huelsenbeck, JP (2009). Parallel genetic evolution within and between bacteriophage species of varying degrees of divergence. *Genetics*, **181**, 225-234.

Bonhomme, M, Chevalet, C, Servin, B, Boitard, S, Abdallah, JM, Blott, S and San Cristobal, M (2010). Detecting selection in population trees: The lewontin and krakauer test extended. *Genetics*, **186**, 241-262.

Bossdorf, O, Auge, H, Lafuma, L, Rogers, WE, Siemann, E and Prati, D (2005). Phenotypic and genetic differentiation between native and introduced plant populations. *Oecologia*, **144**, 1-11.

Bradshaw, WE and Holzapfel, CM (2008). Genetic response to rapid climate change: It's seasonal timing that matters. *Molecular Ecology*, **17**, 157-166.

Brandes, D and Nitzsche, J (2006). Biology, introduction, dispersal and distribution of common ragweed (*ambrosia artemisiifolia* l.) with special regard to germany. *Nachrichtenblatt des Deutschen Planzenschutzdienstes*, **58**, 286-291.

Brauer, CJ, Hammer, MP and Beheregaray, LB (2016). Riverscape genomics of a threatened fish across a hydroclimatically heterogeneous river basin. *Molecular Ecology*, **25**, 5093-5113.

Breiman, L (2001). Random forests. *Machine learning*, **45**, 5-32.

Burd, M and Allen, T (1988). Sexual allocation strategy in wind-pollinated plants. *Evolution*, **42**, 403-407.

Burke, MK, Dunham, JP, Shahrestani, P, Thornton, KR, Rose, MR and Long, AD (2010). Genome-wide analysis of a long-term evolution experiment with *drosophila*. *Nature*, **467**, 587-590.

Campos-Vargas, R and Saltveit, ME (2002). Involvement of putative chemical wound signals in the induction of phenolic metabolism in wounded lettuce. *Physiologia Plantarum*, **114**, 73-84.

Cardarelli, E, Musacchio, A, Montagnani, C, Bogliani, G, Citterio, S and Gentili, R (2018). *Ambrosia artemisiifolia* control in agricultural areas: Effect of grassland seeding and herbivory by the exotic leaf beetle *ophraella communa*. *NeoBiota*, **38**, 1.

Carrillo, J, Wang, Y, Ding, J, Klootwyk, K and Siemann, E (2012). Decreased indirect defense in the invasive tree, *triadica sebifera*. *Plant Ecology*, **213**, 945-954.

Catchen, JM, Amores, A, Hohenlohe, P, Cresko, W and Postlethwait, JH (2011). Stacks: Building and genotyping loci de novo from short-read sequences. *G3 (Bethesda)*, **1**, 171-182.

Chapman, DS, Haynes, T, Beal, S, Essl, F and Bullock, JM (2014). Phenology predicts the native and invasive range limits of common ragweed. *Global Change Biology*, **20**, 192-202.

Chapman, JA, Ho, I, Sunkara, S, Luo, S, Schroth, GP and Rokhsar, DS (2011). Meraculous: De novo genome assembly with short paired-end reads. *PloS One*, **6**, e23501.

Charnov, EL (1982). The theory of sex allocation. *Monographs in Population Biology*, **18**, 1-355.

Chauvel, B, Dessaint, F, Cardinal-Legrand, C and Bretagnolle, F (2006). The historical spread of *ambrosia artemisiifolia* l. In france from herbarium records. *Journal of Biogeography*, **33**, 665-673.

Chevin, L-M and Hospital, F (2008). Selective sweep at a quantitative trait locus in the presence of background genetic variation. *Genetics*, **180**, 1645.

Chevin, L-M, Lande, R and Mace, GM (2010a). Adaptation, plasticity, and extinction in a changing environment: Towards a predictive theory. *PLoS Biology*, **8**, e1000357.

Chevin, LM, Martin, G and Lenormand, T (2010b). Fisher's model and the genomics of adaptation: Restricted pleiotropy, heterogenous mutation, and parallel evolution. *Evolution*, **64**, 3213-3231.

Choisy, M, Franck, P and Cornuet, JM (2004). Estimating admixture proportions with microsatellites: Comparison of methods based on simulated data. *Molecular Ecology*, **13**, 955-968.

Chown, SL, Hodgins, KA, Griffin, PC, Oakeshott, JG, Byrne, M and Hoffmann, AA (2014). Biological invasions, climate change and genomics. *Evolutionary Applications*, **8**, 23-46.

Christin, P-A, Weinreich, DM and Besnard, G (2010). Causes and evolutionary significance of genetic convergence. *Trends in Genetics*, **26**, 400-405.

Chuine, I, Aitken, SN and Ying, CC (2001). Temperature thresholds of shoot elongation in provenances of *pinus contorta*. *Canadian Journal of Forest Research*, **31**, 1444-1455.

Chun, YJ, Fumanal, B, Laitung, B and Bretagnolle, F (2010). Gene flow and population admixture as the primary post-invasion processes in common ragweed (*ambrosia artemisiifolia*) populations in france. *New Phytologist*, **185**, 1100-1107.

Chun, YJ, V, LEC and Bretagnolle, F (2011). Adaptive divergence for a fitness-related trait among invasive *ambrosia artemisiifolia* populations in france. *Molecular Ecology*, **20**, 1378-1388.

Cipollini, D and Lieurance, DM (2012). Expression and costs of induced defense traits in *alliaria petiolata*, a widespread invasive plant. *Basic and Applied Ecology*, **13**, 432-440.

Cipollini, D, Mbagwu, J, Barto, K, Hillstrom, C and Enright, S (2005). Expression of constitutive and inducible chemical defenses in native and invasive populations of *alliaria petiolata*. *Journal of Chemical Ecology*, **31**, 1255-1267.

Colautti, RI and Barrett, SC (2013). Rapid adaptation to climate facilitates range expansion of an invasive plant. *Science*, **342**, 364-366.

Colautti, RI, Eckert, CG and Barrett, SC (2010). Evolutionary constraints on adaptive evolution during range expansion in an invasive plant. *Proceedings of the Royal Society B: Biological Sciences*, **277**, 1799-1806.

Colautti, RI and Lau, JA (2015). Contemporary evolution during invasion: Evidence for differentiation, natural selection, and local adaptation. *Molecular Ecology*, **24**, 1999-2017.

Colautti, RI, Maron, JL and Barrett, SCH (2009). Common garden comparisons of native and introduced plant populations: Latitudinal clines can obscure evolutionary inferences. *Evolutionary Applications*, **2**, 187-199.

Coley, PD, Bryant, JP and Chapin, FS (1985). Resource availability and plant antiherbivore defense. *Science*, **230**, 895-899.

Colomer-Ventura, F, Martínez-Vilalta, J, Zuccarini, P, Escolà, A, Armengot, L and Castells, E (2015). Contemporary evolution of an invasive plant is associated with climate but not with herbivory. *Functional Ecology*, **29**, 1475-1485.

Connallon, T and Hall, MD (2018). Genetic constraints on adaptation: A theoretical primer for the genomics era. *Annals of the New York Academy of Sciences*.

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

Constabel, CP and Ryan, CA (1998). A survey of wound-and methyl jasmonate-induced leaf polyphenol oxidase in crop plants. *Phytochemistry*, **47**, 507-511.

Conte, GL, Arnegard, ME, Best, J, Chan, YF, Jones, FC, Kingsley, DM, Schluter, D and Peichel, CL (2015). Extent of qtl reuse during repeated phenotypic divergence of sympatric threespine stickleback. *Genetics*, **201**, 1189-1200.

Conte, GL, Arnegard, ME, Peichel, CL and Schluter, D (2012). The probability of genetic parallelism and convergence in natural populations. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 5039-5047.

Conte, GL, Hodgins, KA, Yeaman, S, Degner, JC, Aitken, SN, Rieseberg, LH and Whitlock, MC (2017). Bioinformatically predicted deleterious mutations reveal complementation in the interior spruce hybrid complex. *BMC genomics*, **18**, 970-981.

Cornuet, J-M, Pudlo, P, Veyssier, J, Dehne-Garcia, A, Gautier, M, Leblois, R, Marin, J-M and Estoup, A (2014). Diyabc v2. 0: A software to make approximate bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*, **30**, 1187-1189.

Cornuet, J-M, Ravigné, V and Estoup, A (2010). Inference on population history and model checking using DNA sequence and microsatellite data with the software diyabc (v1. 0). *BMC Bioinformatics*, **11**, 401.

Cornuet, J-M, Santos, F, Beaumont, MA, Robert, CP, Marin, J-M, Balding, DJ, Guillemaud, T and Estoup, A (2008). Inferring population history with diy abc: A user-friendly approach to approximate bayesian computation. *Bioinformatics*, **24**, 2713-2719.

Cristescu, ME (2015). Genetic reconstructions of invasion history. *Molecular Ecology*, **24**, 2212-2225.

Cronin, JT, Bhattarai, GP, Allen, WJ and Meyerson, LA (2015). Biogeography of a plant invasion: Plant–herbivore interactions. *Ecology*, **96**, 1115-1127.

Csilléry, K, Blum, MG, Gaggiotti, OE and François, O (2010). Approximate bayesian computation (abc) in practice. *Trends in Ecology & Evolution*, **25**, 410-418.

Dalin, P, Ågren, J, Björkman, C, Huttunen, P and Kärkkäinen, K (2008). Leaf trichome formation and plant resistance to herbivory. Induced plant resistance to herbivory. A. Schaller, Springer**:** 89-105.

Danecek, P, Auton, A, Abecasis, G, Albers, CA, Banks, E, DePristo, MA, Handsaker, RE, Lunter, G, Marth, GT and Sherry, ST (2011). The variant call format and vcftools. *Bioinformatics*, **27**, 2156-2158.

De Jong, T and Klinkhamer, P (1989). Size-dependency of sex-allocation in hermaphroditic, monocarpic plants. *Functional Ecology*, **3**, 201-206.

de Jong, TJ and Klinkhamer, PG (1994). Plant size and reproductive success through female and male function. *Journal of Ecology*, **82**, 399-402.

De Mita, S, Thuillet, AC, Gay, L, Ahmadi, N, Manel, S, Ronfort, J and Vigouroux, Y (2013). Detecting selection along environmental gradients: Analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecololgy*, **22**, 1383-1399.

De Rosario-Martinez, H (2015). Phia: Post-hoc interaction analysis (v0.2-1). CRAN, The R Foundation for Statistical Computing**:** R package.

de Villemereuil, P, Gaggiotti, OE, Mouterde, M and Till-Bottraud, I (2016). Common garden experiments in the genomic era: New perspectives and opportunities. *Heredity*, **116**, 249-254.

Dlugosch, KM, Anderson, SR, Braasch, J, Cang, FA and Gillette, HD (2015a). The devil is in the details: Genetic variation in introduced populations and its contributions to invasion. *Molecular Ecology*, **24**, 2095-2111.

Dlugosch, KM, Cang, FA, Barker, BS, Andonian, K, Swope, SM and Rieseberg, LH (2015b). Evolution of invasiveness through increased resource use in a vacant niche. *Nature Plants*, **1**, 1-5.

Dlugosch, KM and Parker, IM (2008a). Founding events in species invasions: Genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, **17**, 431-449.

Dlugosch, KM and Parker, IM (2008b). Invading populations of an ornamental shrub show rapid life history evolution despite genetic bottlenecks. *Ecology Letters*, **11**, 701-709.

Drake, JM (2006). Heterosis, the catapult effect and establishment success of a colonizing bird. *Biology Letters*, **2**, 304-307.

Earl, DA and vonHoldt, BM (2011). Structure harvester: A website and program for visualizing structure output and implementing the evanno method. *Conservation Genetics Resources*, **4**, 359-361.

Ehrlén, J and Morris, WF (2015). Predicting changes in the distribution and abundance of species under environmental change. *Ecology Letters*, **18**, 303-314.

Eigenbrode, SD, Andreas, JE, Cripps, MG, Ding, H, Biggam, RC and Schwarzländer, M (2008). Induced chemical defenses in invasive plants: A case study with *cynoglossum officinale* l. *Biological Invasions*, **10**, 1373-1379.

Elmer, KR and Meyer, A (2011). Adaptation in the age of ecological genomics: Insights from parallelism and convergence. *Trends in Ecology and Evolution*, **26**, 298-306.

Elshire, RJ, Glaubitz, JC, Sun, Q, Poland, JA, Kawamoto, K, Buckler, ES and Mitchell, SE (2011). A robust, simple genotyping-by-sequencing (gbs) approach for high diversity species. *PLoS One*, **6**, e19379.

Emberlin, J (1994). The effects of patterns in climate and pollen abundance on allergy. *Allergy*, **49**, 45-20.

Endara, MJ and Coley, PD (2011). The resource availability hypothesis revisited: A meta-analysis. *Functional Ecology*, **25**, 389-398.

Essl, F, Biró, K, Brandes, D, Broennimann, O, Bullock, JM, Chapman, DS, Chauvel, B, Dullinger, S, Fumanal, B and Guisan, A (2015). Biological flora of the british isles: *Ambrosia artemisiifolia*. *Journal of Ecology*, **103**, 1069-1098.

Estoup, A, Beaumont, M, Sennedot, F, Moritz, C and Cornuet, J-M (2004). Genetic analysis of complex demographic scenarios: Spatially expanding populations of the cane toad, *bufo marinus*. *Evolution*, **58**, 2021-2036.

Estoup, A and Guillemaud, T (2010). Reconstructing routes of invasion using genetic data: Why, how and so what? *Molecular Ecology*, **19**, 4113-4130.

Estoup, A, Lombaert, E, Marin, JM, Guillemaud, T, Pudlo, P, Robert, CP and Cornuet, J (2012). Estimation of demo-genetic model probabilities with approximate bayesian computation using linear discriminant analysis on summary statistics. *Molecular Ecology Resources*, **12**, 846-855.

Estoup, A, Ravigné, V, Hufbauer, R, Vitalis, R, Gautier, M and Facon, B (2016). Is there a genetic paradox of biological invasion? *Annual Review of Ecology, Evolution, and Systematics*, **47**, 51-72.

Etterson, JR and Shaw, RG (2001). Constraint to adaptive evolution in response to global warming. *Science*, **294**, 151-154.

Evanno, G, Regnaut, S and Goudet, J (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular Ecology*, **14**, 2611-2620.

Excoffier, L, Foll, M and Petit, RJ (2009a). Genetic consequences of range expansions. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 481-501.

Excoffier, L, Hofer, T and Foll, M (2009b). Detecting loci under selection in a hierarchically structured population. *Heredity*, **103**, 285-298.

Excoffier, L and Ray, N (2008). Surfing during population expansions promotes genetic revolutions and structuration. *Trends in Ecology & Evolution*, **23**, 347-351.

Facon, B, Genton, BJ, Shykoff, J, Jarne, P, Estoup, A and David, P (2006). A general eco-evolutionary framework for understanding bioinvasions. *Trends in Ecology & Evolution*, **21**, 130-135.

Facon, B, Hufbauer, RA, Tayeh, A, Loiseau, A, Lombaert, E, Vitalis, R, Guillemaud, T, Lundgren, JG and Estoup, A (2011). Inbreeding depression is purged in the invasive insect *harmonia axyridis*. *Current Biology*, **21**, 424-427.

Facon, B, Jarne, P, Pointier, J and David, P (2005). Hybridization and invasiveness in the freshwater snail *melanoides tuberculata*: Hybrid vigour is more important than increase in genetic variance. *Journal of Evolutionary Biology*, **18**, 524-535.

Facon, B, Pointier, J-P, Jarne, P, Sarda, V and David, P (2008). High genetic variance in life-history strategies within invasive populations by way of multiple introductions. *Current Biology*, **18**, 363-367.

Falush, D, van Dorp, L and Lawson, D (2016). A tutorial on how (not) to over-interpret structure/admixture bar plots. *bioRxiv*, 066431, doi: dx.doi.org/10.1101/066431.

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

Felker-Quinn, E, Schweitzer, JA and Bailey, JK (2013). Meta-analysis reveals evolution in invasive plant species but little support for evolution of increased competitive ability (eica). *Ecology and Evolution*, **3**, 739-751.

Felsenstein, J (1974). The evolutionary advantage of recombination. *Genetics*, **78**, 737-756.

Fisher, RA (1930). The genetical theory of natural selection, Oxford University Press.

Fitzpatrick, MC and Keller, SR (2015). Ecological genomics meets community-level modelling of biodiversity: Mapping the genomic landscape of current and future environmental adaptation. *Ecology Letters*, **18**, 1-16.

Flaxman, SM, Wacholder, AC, Feder, JL and Nosil, P (2014). Theoretical models of the influence of genomic architecture on the dynamics of speciation. *Molecular Ecology*, **23**, 4074-4088.

Fortuna, TM, Eckert, S, Harvey, JA, Vet, LE, Müller, C and Gols, R (2014). Variation in plant defences among populations of a range-expanding plant: Consequences for trophic interactions. *New Phytologist*, **204**, 989-999.

Fournier-Level, A, Korte, A, Cooper, MD, Nordborg, M, Schmitt, J and Wilczek, AM (2011). A map of local adaptation in *arabidopsis thaliana*. *Science*, **334**, 86-89.

Fraimout, A, Debat, V, Fellous, S, Hufbauer, RA, Foucaud, J, Pudlo, P, Marin, J-M, Price, DK, Cattel, J and Chen, X (2017). Deciphering the routes of invasion of *drosophila suzukii* by means of abc random forest. *Molecular Biology and Evolution*, **34**, 980-996.

Frankham, R (1995). Conservation genetics. *Annual review of genetics*, **29**, 305-327.

Franks, SJ, Pratt, PD, Dray, FA and Simms, EL (2008). Selection on herbivory resistance and growth rate in an invasive plant. *The American Naturalist*, **171**, 678-691.

Friedman, J and Barrett, SC (2008). High outcrossing in the annual colonizing species *ambrosia artemisiifolia* (asteraceae). *Annals of Botany*, **101**, 1303-1309.

Friedman, J and Barrett, SC (2009). Wind of change: New insights on the ecology and evolution of pollination and mating in wind-pollinated plants. *Annals of Botany*, **103**, 1515-1527.

Friedman, J and Barrett, SC (2011). Genetic and environmental control of temporal and size-dependent sex allocation in a wind-pollinated plant. *Evolution*, **65**, 2061-2074.

Fukano, Y and Yahara, T (2012). Changes in defense of an alien plant *ambrosia artemisiifolia* before and after the invasion of a native specialist enemy *ophraella communa*. *PloS One*, **7**, e49114.

Gaudeul, M, Giraud, T, Kiss, L and Shykoff, JA (2011). Nuclear and chloroplast microsatellites show multiple introductions in the worldwide invasion history of common ragweed, *ambrosia artemisiifolia*. *PLoS One*, **6**, e17658.

Geng, Y-P, Pan, X-Y, Xu, C-Y, Zhang, W-J, Li, B, Chen, J-K, Lu, B-R and Song, Z-P (2007). Phenotypic plasticity rather than locally adapted ecotypes allows the invasive alligator weed to colonize a wide range of habitats. *Biological Invasions*, **9**, 245-256.

Genton, B, Shykoff, J and Giraud, T (2005a). High genetic diversity in french invasive populations of common ragweed, *ambrosia artemisiifolia*, as a result of multiple sources of introduction. *Molecular Ecology*, **14**, 4275-4285.

Genton, BJ, Kotanen, PM, Cheptou, PO, Adolphe, C and Shykoff, JA (2005b). Enemy release but no evolutionary loss of defence in a plant invasion: An inter-continental reciprocal transplant experiment. *Oecologia*, **146**, 404-414.

Gerber, E, Schaffner, U, Gassmann, A, Hinz, H, Seier, M and MÜLLER-SCHÄRER, H (2011). Prospects for biological control of *ambrosia artemisiifolia* in europe: Learning from the past. *Weed Research*, **51**, 559-573.

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

Gladieux, P, Giraud, T, Kiss, L, Genton, BJ, Jonot, O and Shykoff, JA (2010). Distinct invasion sources of common ragweed (*ambrosia artemisiifolia*) in eastern and western europe. *Biological Invasions*, **13**, 933-944.

Glémin, S (2003). How are deleterious mutations purged? Drift versus nonrandom mating. *Evolution*, **57**, 2678-2687.

Gompel, N and Prud'homme, B (2009). The causes of repeated genetic evolution. *Developmental Biology*, **332**, 36-47.

Goslee, SC and Urban, DL (2007). The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, **22**, 1-19.

Goudet, J (2005). Hierfstat, a package for r to compute and test hierarchical f-statistics. *Molecular Ecology Notes*, **5**, 184-186.

Gould, BA and Stinchcombe, JR (2017). Population genomic scans suggest novel genes underlie convergent flowering time evolution in the introduced range of *arabidopsis thaliana*. *Molecular Ecology*, **26**, 92-106.

Griffith, TM and Watson, MA (2005). Is evolution necessary for range expansion? Manipulating reproductive timing of a weedy annual transplanted beyond its range. *The American Naturalist*, **167**, 153-164.

Grime, JP (1977). Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. *The American Naturalist*, **111**, 1169-1194.

Gu, X, Siemann, E, Zhu, L, Gao, S, Wang, Y and Ding, J (2014). Invasive plant population and herbivore identity affect latex induction. *Ecological Entomology*, **39**, 1-9.

Guisan, A, Petitpierre, B, Broennimann, O, Daehler, C and Kueffer, C (2014). Unifying niche shift studies: Insights from biological invasions. *Trends in Ecology & Evolution*, **29**, 260-269.

Guisan, A and Zimmermann, NE (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147-186.

Günther, T and Coop, G (2013). Robust identification of local adaptation from allele frequencies. *Genetics*, **195**, 205-220.

Guo, H, Mazer, SJ and Du, G (2010). Geographic variation in primary sex allocation per flower within and among 12 species of *pedicularis* (orobanchaceae): Proportional male investment increases with elevation. *American Journal of Botany*, **97**, 1334-1341.

Hahn, MA and Rieseberg, LH (2017). Genetic admixture and heterosis may enhance the invasiveness of common ragweed. *Evolutionary Applications*, **10**, 241-250.

Hahn, PG and Maron, JL (2016). A framework for predicting intraspecific variation in plant defense. *Trends in Ecology & Evolution*, **31**, 646-656.

Hämälä, T and Savolainen, O (2018). Local adaptation under gene flow: Recombination, conditional neutrality and genetic trade-offs shape genomic patterns in *arabidopsis lyrata*. *bioRxiv*.

Hamilton, JA, Okada, M, Korves, T and Schmitt, J (2015). The role of climate adaptation in colonization success in *arabidopsis thaliana*. *Molecular Ecology*, **24**, 2253-2263.

Hancock, AM, Alkorta-Aranburu, G, Witonsky, DB and Di Rienzo, A (2010). Adaptations to new environments in humans: The role of subtle allele frequency shifts. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **365**, 2459-2468.

Hansen, TF (2006). The evolution of genetic architecture. *Annual Review of Ecology, Evolution, and Systematics*, 123-157.

Harrisson, KA, Amish, SJ, Pavlova, A, Narum, SR, Telonis-Scott, M, Rourke, ML, Lyon, J, Tonkin, Z, Gilligan, DM and Ingram, BA (2017). Signatures of polygenic adaptation associated with climate across the range of a threatened fish species with high genetic connectivity. *Molecular Ecology*, **26**, 6253-6269.

Harrisson, KA, Pavlova, A, Telonis-Scott, M and Sunnucks, P (2014). Using genomics to characterize evolutionary potential for conservation of wild populations. *Evolutionary Applications*, **7**, 1008-1025.

Hartfield, M and Otto, SP (2011). Recombination and hitchhiking of deleterious alleles. *Evolution: International Journal of Organic Evolution*, **65**, 2421-2434.

Hauser, M-T (2014). Molecular basis of natural variation and environmental control of trichome patterning. *Frontiers in Plant Science*, **5**, 320.

He, W-M, Thelen, GC, Ridenour, WM and Callaway, RM (2010). Is there a risk to living large? Large size correlates with reduced growth when stressed for knapweed populations. *Biological Invasions*, **12**, 3591-3598.

Hecht, BC, Matala, AP, Hess, JE and Narum, SR (2015). Environmental adaptation in chinook salmon (*oncorhynchus tshawytscha*) throughout their north american range. *Molecular Ecology*, **24**, 5573-5595.

Heredia, JB and Cisneros-Zevallos, L (2009). The effect of exogenous ethylene and methyl jasmonate on pal activity, phenolic profiles and antioxidant capacity of carrots (*daucus carota*) under different wounding intensities. *Postharvest Biology and Technology*, **51**, 242-249.

Hereford, J (2009). A quantitative survey of local adaptation and fitness trade-offs. *American Naturalist*, **173**, 579-588.

Hermisson, J and Pennings, PS (2005). Soft sweeps: Molecular population genetics of adaptation from standing genetic variation. *Genetics*, **169**, 2335-2352.

Hijmans, RJ, Cameron, SE, Parra, JL, Jones, PG and Jarvis, A (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965-1978.

Hill, W and Robertson, A (1968). Linkage disequilibrium in finite populations. *Theoretical and Applied Genetics*, **38**, 226-231.

Hoban, S, Kelley, JL, Lotterhos, KE, Antolin, MF, Bradburd, G, Lowry, DB, Poss, ML, Reed, LK, Storfer, A and Whitlock, MC (2016). Finding the genomic basis of local

adaptation: Pitfalls, practical solutions, and future directions. *The American Naturalist*, **188**, 379-397.

Hodgins, KA, Bock, DG and Rieseberg, L (2018). Trait evolution in invasive species. *Annual Plant Reviews online*, **1**, 1-37.

Hodgins, KA and Rieseberg, L (2011). Genetic differentiation in life-history traits of introduced and native common ragweed (*ambrosia artemisiifolia*) populations. *Journal of Evolutionary Biology*, **24**, 2731-2749.

Hoffmann, AA and Sgrò, CM (2011). Climate change and evolutionary adaptation. *Nature*, **470**, 479-485.

Hoffmann, AA, Sgrò, CM and Kristensen, TN (2017). Revisiting adaptive potential, population size, and conservation. *Trends in Ecology & Evolution*.

Holliday, JA, Zhou, L, Bawa, R, Zhang, M and Oubida, RW (2016). Evidence for extensive parallelism but divergent genomic architecture of adaptation along altitudinal and latitudinal gradients in *populus trichocarpa*. *New Phytologist*, **209**, 1240-1251.

Huang, F, Peng, S, Chen, B, Liao, H, Huang, Q, Lin, Z and Liu, G (2015). Rapid evolution of dispersal-related traits during range expansion of an invasive vine *mikania micrantha*. *Oikos*, **124**, 1023-1030.

Hudson, RR (1985). The sampling distribution of linkage disequilibrium under an infinite allele model without selection. *Genetics*, **109**, 611-631.

Huey, RB, Gilchrist, GW, Carlson, ML, Berrigan, D and Serra, Ls (2000). Rapid evolution of a geographic cline in size in an introduced fly. *Science*, **287**, 308-309.

Hufbauer, RA, Facon, B, Ravigne, V, Turgeon, J, Foucaud, J, Lee, CE, Rey, O and Estoup, A (2012). Anthropogenically induced adaptation to invade (aiai): Contemporary adaptation to human-altered habitats within the native range can promote invasions. *Evolutionary Applications*, **5**, 89-101.

Ito, K and Sakai, S (2009). Optimal defense strategy against herbivory in plants: Conditions selecting for induced defense, constitutive defense, and no-defense. *Journal of Theoretical Biology*, **260**, 453-459.

Jakobsson, M and Rosenberg, NA (2007). Clumpp: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801-1806.

Jombart, T (2008). Adegenet: A r package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403-1405.

Jones, FC, Grabherr, MG, Chan, YF, Russell, P, Mauceli, E, Johnson, J, Swofford, R, Pirun, M, Zody, MC, White, S, Birney, E, Searle, S, Schmutz, J, Grimwood, J, Dickson, MC, Myers, RM, Miller, CT, Summers, BR, Knecht, AK, Brady, SD, Zhang, H, Pollen, AA, Howes, T, Amemiya, C, Broad Institute Genome Sequencing, P, Whole Genome

Assembly, T, Baldwin, J, Bloom, T, Jaffe, DB, Nicol, R, Wilkinson, J, Lander, ES, Di Palma, F, Lindblad-Toh, K and Kingsley, DM (2012). The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, **484**, 55-61.

Joshi, J and Vrieling, K (2005). The enemy release and eica hypothesis revisited: Incorporating the fundamental difference between specialist and generalist herbivores. *Ecology Letters*, **8**, 704-714.

Joshi, NA and Fass, JN (2011). Sickle: A sliding-window, adaptive, quality-based trimming tool for fastq files https://github.com/najoshi/sickle. **Version 1.33**.

Jump, AS, Marchant, R and Penuelas, J (2009). Environmental change and the option value of genetic diversity. *Trends in Plant Sciences*, **14**, 51-58.

Katabuchi, M (2015). Leafarea: An r package for rapid digital image analysis of leaf area. *Ecological Research*, **30**, 1073-1077.

Kawecki, TJ and Ebert, D (2004). Conceptual issues in local adaptation. *Ecology Letters*, **7**, 1225-1241.

Kazinczi, G, Béres, I, Novák, R, Bíró, K and Pathy, Z (2008). Common ragweed (*ambrosia artemisiifolia*): A review with special regards to the results in hungary. I. Taxonomy, origin and distribution, morphology, life cycle and reproduction strategy. *Herbologia*, **9**, 55-91.

Kearney, M and Porter, W (2009). Mechanistic niche modelling: Combining physiological and spatial data to predict species' ranges. *Ecology letters*, **12**, 334-350.

Keenan, K, McGinnity, P, Cross, TF, Crozier, WW and Prodöhl, PA (2013). Diversity: An r package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, **4**, 782-788.

Keinänen, M, Oldham, NJ and Baldwin, IT (2001). Rapid hplc screening of jasmonate-induced increases in tobacco alkaloids, phenolics, and diterpene glycosides in *nicotiana attenuata*. *Journal of agricultural and food chemistry*, **49**, 3553-3558.

Keller, S, Fields, P, Berardi, A and Taylor, D (2014). Recent admixture generates heterozygosity–fitness correlations during the range expansion of an invading species. *Journal of evolutionary biology*, **27**, 616-627.

Keller, SR, Sowell, DR, Neiman, M, Wolfe, LM and Taylor, DR (2009). Adaptation and colonization history affect the evolution of clines in two introduced species. *New Phytologist*, **183**, 678-690.

Keller, SR and Taylor, DR (2010). Genomic admixture increases fitness during a biological invasion. *Journal of Evolutionary Biology*, **23**, 1720-1731.

Keppel, G (1991). Design and analysis: A researcher's handbook. New Jersey, United States, Prentice-Hall, Inc.

Kessler, A and Baldwin, IT (2002). Plant responses to insect herbivory: The emerging molecular analysis. *Annual Review of Plant Biology*, **53**, 299-328.

Kim, Y and Nielsen, R (2004). Linkage disequilibrium as a signature of selective sweeps. *Genetics*, **167**, 1513-1524.

Kinnison, MT and Hairston, NG (2007). Eco-evolutionary conservation biology: Contemporary evolution and the dynamics of persistence. *Functional Ecology*, **21**, 444-454.

Klinkhamer, PG, De Jong, TJ and Metz, H (1997). Sex and size in cosexual plants. *Trends in Ecology & Evolution*, **12**, 260-265.

Klopfstein, S, Currat, M and Excoffier, L (2006). The fate of mutations surfing on the wave of a range expansion. *Molecular Biology and Evolution*, **23**, 482-490.

Kopelman, NM, Mayzel, J, Jakobsson, M, Rosenberg, NA and Mayrose, I (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across k. *Molecular Ecology Resources*, **15**, 1179-1191.

Koricheva, J, Nykänen, H and Gianoli, E (2004). Meta-analysis of trade-offs among plant antiherbivore defenses: Are plants jacks-of-all-trades, masters of all? *The American Naturalist*, **163**, E64-E75.

Korpelainen, H (1998). Labile sex expression in plants. *Biological Reviews*, **73**, 157-180.

Kremer, A and Le Corre, V (2012). Decoupling of differentiation between traits and their underlying genes in response to divergent selection. *Heredity*, **108**, 375-385.

Krishnamoorthy, K and Lee, M (2014). Improved tests for the equality of normal coefficients of variation. *Computational Statistics*, **29**, 215-232.

Kubešová, M, Moravcova, L, Suda, J, Jarošík, V and Pyšek, P (2010). Naturalized plants have smaller genomes than their non-invading relatives: A flow cytometric analysis of the czech alien flora. *Preslia*, **82**, 81-96.

Kumschick, S, Hufbauer, RA, Alba, C and Blumenthal, DM (2013). Evolution of fast-growing and more resistant phenotypes in introduced common mullein (*verbascum thapsus*). *Journal of Ecology*, **101**, 378-387.

Kuznetsova, A, Brockhoff, PB and Christensen, RHB (2015). Package 'lmertest' (v3.0-1). R package version. CRAN, The R Foundation for Statistical Computing.

Laaidi, M, Laaidi, K, Besancenot, J-P and Thibaunod, M (2003). Ragweed in france: An invasive plant and its allergenic pollen. *Annals of Allergy, Asthma & Immunology*, **91**, 195-201.

Lachmuth, S, Durka, W and Schurr, FM (2011). Differentiation of reproductive and competitive ability in the invaded range of *senecio inaequidens*: The role of genetic allee effects, adaptive and nonadaptive evolution. *New Phytologist*, **192**, 529-541.

Lande, R (1980). Genetic variation and phenotypic evolution during allopatric speciation. *The American Naturalist*, **116**, 463-479.

Lande, R (2015). Evolution of phenotypic plasticity in colonizing species. *Molecular Ecology*, **24**, 2038-2045.

Lasky, JR, Des Marais, DL, McKAY, JK, Richards, JH, Juenger, TE and Keitt, TH (2012). Characterizing genomic variation of *arabidopsis thaliana*: The roles of geography and climate. *Molecular Ecology*, **21**, 5512-5529.

Lässig, M, Mustonen, V and Walczak, AM (2017). Predicting evolution. *Nature Ecology & Evolution*, **1**, 0077.

Lawson Handley, L-J, Estoup, A, Evans, DM, Thomas, CE, Lombaert, E, Facon, B, Aebi, A and Roy, HE (2011). Ecological genetics of invasive species. *BioControl*, **56**, 409-428.

Le Corre, V and Kremer, A (2012). The genetic differentiation at quantitative trait loci under local adaptation. *Molecular Ecology*, **21**, 1548-1566.

Lee, CE (2002). Evolutionary genetics of invasive species. *Trends in Ecology and Evolution*, **17**, 386-391.

Lee, J, Vogt, T, Schmidt, J, Parthier, B and Löbler, M (1997). Methyljasmonate-induced accumulation of coumaroyl conjugates in barley leaf segments. *Phytochemistry*, **44**, 589-592.

Lee, KM and Coop, G (2017). Distinguishing among modes of convergent adaptation using population genomic data. *Genetics*, genetics. 300417.302017.

Lee, YW, Gould, BA and Stinchcombe, JR (2014). Identifying the genes underlying quantitative traits: A rationale for the qtn programme. *AoB Plants*, **6**.

Leiblein-Wild, MC and Tackenberg, O (2014). Phenotypic variation of 38 european *ambrosia artemisiifolia* populations measured in a common garden experiment. *Biological Invasions*, **16**, 2003-2015.

Lenormand, T, Chevin, L-M and Bataillon, T (2016). Parallel evolution: What does it (not) tell us and why is it (still) interesting. *Chance in evolution*, 196.

Lenormand, T and Otto, SP (2000). The evolution of recombination in a heterogeneous environment. *Genetics*, **156**, 423-438.

Li, H and Durbin, R (2009). Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*, **25**, 1754-1760.

Li, X-M, Liao, W-J, Wolfe, LM and Zhang, D-Y (2012). No evolutionary shift in the mating system of north american *ambrosia artemisiifolia* (asteraceae) following its introduction to china. *PloS One*, **7**, e31935.

Li, X-M, She, D-Y, Zhang, D-Y and Liao, W-J (2014). Life history trait differentiation and local adaptation in invasive populations of *ambrosia artemisiifolia* in china. *Oecologia*, **177**, 669-677.

Li, Z-H, Wang, Q, Ruan, X, Pan, C-D and Jiang, D-A (2010). Phenolics and plant allelopathy. *Molecules*, **15**, 8933-8952.

Lippens, C, Estoup, A, Hima, M, Loiseau, A, Tatard, C, Dalecky, A, Bâ, K, Kane, M, Diallo, M, Sow, A, Niang, Y, Piry, S, Berthier, K, Leblois, R, Duplantier, J-M and Brouat, C (2017). Genetic structure and invasion history of the house mouse (*mus musculus domesticus*) in senegal, west africa: A legacy of colonial and contemporary times. *Heredity*.

Lloyd, D (1984). Gender allocations in outcrossing cosexual plants. Perspectives on plant population ecology. D. Dirzo and J. Sarukhan. Sunderland, Massachusetts, Sinauer Associated**:** 277-300.

Lloyd, DG and Bawa, KS (1984). Modification of the gender of seed plants in varying conditions. *Evolutionary Biology*, **17**, 255-338.

Lombaert, E, Guillemaud, T, Cornuet, J-M, Malausa, T, Facon, B and Estoup, A (2010). Bridgehead effect in the worldwide invasion of the biocontrol harlequin ladybird. *PloS One*, **5**, e9743.

Lombaert, E, Guillemaud, T, Lundgren, J, Koch, R, Facon, B, Grez, A, Loomans, A, Malausa, T, Nedved, O and Rhule, E (2014). Complementarity of statistical treatments to reconstruct worldwide routes of invasion: The case of the asian ladybird *harmonia axyridis*. *Molecular Ecology*, **23**, 5979-5997.

Lombaert, E, Guillemaud, T, Thomas, C, Lawson Handley, L, Li, J, Wang, S, Pang, H, Goryacheva, I, Zakharov, I and Jousselin, E (2011). Inferring the origin of populations introduced from a genetically structured native range by approximate bayesian computation: Case study of the invasive ladybird *harmonia axyridis*. *Molecular Ecology*, **20**, 4654-4670.

Lommen, STE, Hallmann, CA, Jongejans, E, Chauvel, B, Leitsch-Vitalos, M, Aleksanyan, A, Tóth, P, Preda, C, Šćepanović, M, Onen, H, Tokarska-Guzik, B, Anastasiu, P, Dorner, Z, Fenesi, A, Karrer, G, Nagy, K, Pinke, G, Tiborcz, V, Zagyvai, G, Zalai, M, Kazinczi, G, Leskovšek, R, Stešević, D, Fried, G, Kalatozishvili, L, Lemke, A and Müller-Schärer, H (2017). Explaining variability in the production of seed and allergenic pollen by invasive *ambrosia artemisiifolia* across europe. *Biological Invasions*, **20**, 1475-1491.

Losos, JB (2011). Convergence, adaptation, and constraint. *Evolution*, **65**, 1827-1840.

Lotterhos, K, Yeaman, S, Degner, J, Aitken, S and Hodgins, K (2018). Modularity of genes involved in local adaptation to climate despite physical linkage. *bioRxiv*, 202481.

Lotterhos, KE and Whitlock, MC (2014). Evaluation of demographic history and neutral parameterization on the performance of f_{st} outlier tests. *Molecular Ecology*, **23**, 2178-2192.

Lotterhos, KE and Whitlock, MC (2015). The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular ecology*, **24**, 1031-1046.

Macel, M, Dostálek, T, Esch, S, Bucharová, A, van Dam, NM, Tielbörger, K, Verhoeven, KJ and Münzbergová, Z (2017). Evolutionary responses to climate change in a range expanding plant. *Oecologia*, **184**, 543-554.

MacPherson, A and Nuismer, S (2017). The probability of parallel genetic evolution from standing genetic variation. *Journal of Evolutionary Biology*, **30**, 326-337.

Marchini, GL, Sherlock, NC, Ramakrishnan, AP, Rosenthal, DM and Cruzan, MB (2016). Rapid purging of genetic load in a metapopulation and consequences for range expansion in an invasive plant. *Biological Invasions*, **18**, 183-196.

Marin, J-M, Raynal, L, Pudlo, P, Ribatet, M and Robert, CP (2016). Abc random forests for bayesian parameter inference. *arXiv preprint arXiv:1605.05537*.

Marin, JM, Pillai, NS, Robert, CP and Rousseau, J (2014). Relevant statistics for bayesian model choice. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **76**, 833-859.

Maron, JL and Vilà, M (2001). When do herbivores affect plant invasion? Evidence for the natural enemies and biotic resistance hypotheses. *Oikos*, **95**, 361-373.

Maron, JL, Vilà, M, Bommarco, R, Elmendorf, S and Beardsley, P (2004). Rapid evolution of an invasive plant. *Ecological Monographs*, **74**, 261-280.

Marques, DA, Jones, FC, Di Palma, F, Kingsley, DM and Reimchen, TE (2018). Experimental evidence for rapid genomic adaptation to a new niche in an adaptive radiation. *Nature Ecology & Evolution*, **2**, 1130.

Martin, A and Orgogozo, V (2013). The loci of repeated evolution: A catalog of genetic hotspots of phenotypic variation. *Evolution*, **67**, 1235-1250.

Martin, MD, Chamecki, M, Brush, GS, Meneveau, C and Parlange, MB (2009). Pollen clumping and wind dispersal in an invasive angiosperm. *American Journal of Botany*, **96**, 1703-1711.

Martin, MD, Olsen, MT, Samaniego, JA, Zimmer, EA and Gilbert, MTP (2016). The population genomic basis of geographic differentiation in north american common ragweed (*ambrosia artemisiifolia* l.). *Ecology and Evolution*, **6**, 3760-3771.

Martin, MD, Zimmer, EA, Olsen, MT, Foote, AD, Gilbert, MT and Brush, GS (2014). Herbarium specimens reveal a historical shift in phylogeographic structure of common ragweed during native range disturbance. *Molecular Ecology*, **23**, 1701-1716.

Marwick, B and Krishnamoorth, K (2018). Cvequality: Tests for the equality of coefficients of variation from multiple groups**:** R package.

Maynard Smith, J (1977). Why the genome does not congeal. *Nature*, **268**, 693.

Maynard Smith, J and Haigh, J (1974). The hitch-hiking effect of a favourable gene. *Genetics Research*, **23**, 23-35.

McGee, MD and Wainwright, PC (2013). Convergent evolution as a generator of phenotypic diversity in threespine stickleback. *Evolution*, **67**, 1204-1208.

McKenna, A, Hanna, M, Banks, E, Sivachenko, A, Cibulskis, K, Kernytsky, A, Garimella, K, Altshuler, D, Gabriel, S, Daly, M and DePristo, MA (2010). The genome analysis toolkit: A mapreduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, **20**, 1297-1303.

McVean, GA and Charlesworth, B (2000). The effects of hill-robertson interference between weakly selected mutations on patterns of molecular evolution and variation. *Genetics*, **155**, 929-944.

Merilä, J and Hendry, AP (2014). Climate change, adaptation, and phenotypic plasticity: The problem and the evidence. *Evolutionary Applications*, **7**, 1-14.

Messer, PW and Petrov, DA (2013). Population genomics of rapid adaptation by soft selective sweeps. *Trends in Ecology & Evolution*, **28**, 659-669.

Miller, GA and Chapman, JP (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, **110**, 40-48.

Moles, AT, Bonser, SP, Poore, AG, Wallis, IR and Foley, WJ (2011a). Assessing the evidence for latitudinal gradients in plant defence and herbivory. *Functional Ecology*, **25**, 380-388.

Moles, AT, Wallis, IR, Foley, WJ, Warton, DI, Stegen, JC, Bisigato, AJ, Cella-Pizarro, L, Clark, CJ, Cohen, PS and Cornwell, WK (2011b). Putting plant resistance traits on the map: A test of the idea that plants are better defended at lower latitudes. *New Phytologist*, **191**, 777-788.

Momigliano, P, Jokinen, H, Fraimout, A, Florin, A-B, Norkko, A and Merilä, J (2017). Extraordinarily rapid speciation in a marine fish. *Proceedings of the National Academy of Sciences*, **114**, 6074-6079.

Montague, JL, Barrett, SC and Eckert, CG (2008). Re-establishment of clinal variation in flowering time among introduced populations of purple loosestrife (*lythrum salicaria*, lythraceae). *Journal of Evolutionary Biology*, **21**, 234-245.

Moran, EV and Alexander, JM (2014). Evolutionary responses to global change: Lessons from invasive species. *Ecology Letters*, **17**, 637-649.

Moreira, X, Mooney, KA, Rasmann, S, Petry, WK, Carrillo-Gavilán, A, Zas, R and Sampedro, L (2014). Trade-offs between constitutive and induced defences drive

geographical and climatic clines in pine chemical defences. *Ecology Letters*, **17**, 537-546.

Morris, WF, Traw, MB and Bergelson, J (2006). On testing for a tradeoff between constitutive and induced resistance. *Oikos*, **112**, 102-110.

Muller, HJ (1964). The relation of recombination to mutational advance. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, **1**, 2-9.

Muller-Scharer, H, Schaffner, U and Steinger, T (2004). Evolution in invasive plants: Implications for biological control. *Trends in Ecology & Evolution*, **19**, 417-422.

Müller-Schärer, H, Lommen, ST, Rossinelli, M, Bonini, M, Boriani, M, Bosio, G and Schaffner, U (2014). *Ophraella communa*, the ragweed leaf beetle, has successfully landed in europe: Fortunate coincidence or threat? *Weed research*, **54**, 109-119.

Nei, M (1972). Genetic distance between populations. *The American Naturalist*, **106**, 283-292.

Nei, M (1987). Molecular evolutionary genetics, Columbia University Press.

Neilson, EH, Goodger, JQ, Woodrow, IE and Møller, BL (2013). Plant chemical defense: At what cost? *Trends in plant science*, **18**, 250-258.

Neiman, M and Linksvayer, T (2006). The conversion of variance and the evolutionary potential of restricted recombination. *Heredity*, **96**, 111.

Nosil, P and Feder, JL (2012). Genomic divergence during speciation: Causes and consequences, The Royal Society**:** 332-342.

Oduor, AM, Leimu, R and Kleunen, M (2016). Invasive plant species are locally adapted just as frequently and at least as strongly as native plant species. *Journal of Ecology*, **104**, 957-968.

Orians, CM and Ward, D (2010). Evolution of plant defenses in nonindigenous environments. *Annual Review of Entomology*, **55**, 439-459.

Orr, HA (1998). The population genetics of adaptation: The distribution of factors fixed during adaptive evolution. *Evolution*, **52**, 935-949.

Orrock, JL, Sih, A, Ferrari, MC, Karban, R, Preisser, EL, Sheriff, MJ and Thaler, JS (2015). Error management in plant allocation to herbivore defense. *Trends in Ecology & Evolution*, **30**, 441-445.

Ortiz-Barrientos, D, Engelstädter, J and Rieseberg, LH (2016). Recombination rate evolution and the origin of species. *Trends in Ecology & Evolution*, **31**, 226-236.

Oswalt, ML and Marshall, GD (2008). Ragweed as an example of worldwide allergen expansion. *Allergy, Asthma and Clinical Immunology*, **4**, 130-135.

Otto, SP (2012). Theoretical perspectives on rapid evolutionary change. Rapidly evoloving genes and genetic systems. R. S. Singh, J. Xu and R. J. Kulathinal, Oxford University Press.

Palmer, B and McFadyen, RE (2012). *Ambrosia artemisiifolia* l. - annual ragweed. Biological control of weeds in australia. M. H. Julien, R. E. McFadyen and J. Cullen. Collingwood, CSIRO**:** 52-59.

Paquin, V and Aarssen, LW (2004). Allometric gender allocation in *ambrosia artemisiifolia* (asteraceae) has adaptive plasticity. *American Journal of Botany*, **91**, 430-438.

Parker, JD, Torchin, ME, Hufbauer, RA, Lemoine, NP, Alba, C, Blumenthal, DM, Bossdorf, O, Byers, JE, Dunn, AM and Heckman, RW (2013). Do invasive species perform better in their new ranges? *Ecology*, **94**, 985-994.

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

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

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

Phillips, BL, Brown, GP, Webb, JK and Shine, R (2006). Invasion and the evolution of speed in toads. *Nature*, **439**, 803-803.

Pohlert, T (2014). The pairwise multiple comparison of mean ranks package (pmcmr).

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

Polechová, J and Barton, NH (2015). Limits to adaptation along environmental gradients. *Proceedings of the National Academy of Sciences*, 201421515.

Poorter, H and Remkes, C (1990). Leaf area ratio and net assimilation rate of 24 wild species differing in relative growth rate. *Oecologia*, **83**, 553-559.

Prentis, PJ, Wilson, JR, Dormontt, EE, Richardson, DM and Lowe, AJ (2008). Adaptive evolution in invasive species. *Trends in Plant Sciences*, **13**, 288-294.

Pritchard, JK and Di Rienzo, A (2010). Adaptation - not by sweeps alone. *Nature Review Genetics*, **11**, 665-667.

Pritchard, JK, Pickrell, JK and Coop, G (2010). The genetics of human adaptation: Hard sweeps, soft sweeps, and polygenic adaptation. *Current Biology*, **20**, R208-R215.

Pritchard, JK, Stephens, M and Donnelly, P (2000). Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-959.

Przeworski, M (2002). The signature of positive selection at randomly chosen loci. *Genetics*, **160**, 1179-1189.

Pudlo, P, Marin, J-M, Estoup, A, Cornuet, J-M, Gautier, M and Robert, CP (2016). Reliable abc model choice via random forests. *Bioinformatics*, **32**, 859-866.

Puechmaille, SJ (2016). The program structure does not reliably recover the correct population structure when sampling is uneven: Subsampling and new estimators alleviate the problem. *Molecular Ecology Resources*, **16**, 608-627.

Purcell, S, Neale, B, Todd-Brown, K, Thomas, L, Ferreira, MA, Bender, D, Maller, J, Sklar, P, De Bakker, PI and Daly, MJ (2007). Plink: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, **81**, 559-575.

Putnam, NH, O'Connell, BL, Stites, JC, Rice, BJ, Blanchette, M, Calef, R, Troll, CJ, Fields, A, Hartley, PD and Sugnet, CW (2016). Chromosome-scale shotgun assembly using an in vitro method for long-range linkage. *Genome Research*, **26**, 342-350.

Pylkov, KV, Zhivotovsky, LA and Feldman, MW (1998). Migration versus mutation in the evolution of recombination under multilocus selection. *Genetics Research*, **71**, 247-256.

Quarantine Act (1908). Act No. 3 of 1908. H. A. a. W. Resources. **C1908A00003**.

R Core Team (2017). R: A language and environment for statistical computing (v3.4.3 "kite-eating tree"). Vienna, Austria, The R Foundation for Statistical Computing.

Rasmann, S and Agrawal, AA (2011). Latitudinal patterns in plant defense: Evolution of cardenolides, their toxicity and induction following herbivory. *Ecology Letters*, **14**, 476-483.

Rasmann, S, Agrawal, AA, Cook, SC and Erwin, AC (2009). Cardenolides, induced responses, and interactions between above-and belowground herbivores of milkweed (*asclepias spp.*). *Ecology*, **90**, 2393-2404.

Rausher, MD and Delph, LF (2015). Commentary: When does understanding phenotypic evolution require identification of the underlying genes? *Evolution*, **69**, 1655-1664.

Ravinet, M, Westram, A, Johannesson, K, Butlin, R, André, C and Panova, M (2016). Shared and nonshared genomic divergence in parallel ecotypes of *littorina saxatilis* at a local scale. *Molecular Ecology*, **25**, 287-305.

Rellstab, C, Gugerli, F, Eckert, AJ, Hancock, AM and Holderegger, R (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, **24**, 4348-4370.

Remington, DL, Thornsberry, JM, Matsuoka, Y, Wilson, LM, Whitt, SR, Doebley, J, Kresovich, S, Goodman, MM and Buckler, ES (2001). Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proceedings of the National Academy of Sciences*, **98**, 11479-11484.

Renaut, S, Owens, GL and Rieseberg, LH (2014). Shared selective pressure and local genomic landscape lead to repeatable patterns of genomic divergence in sunflowers. *Molecular Ecology*, **23**, 311-324.

Reusch, TB and Wood, TE (2007). Molecular ecology of global change. *Molecular Ecology*, **16**, 3973-3992.

Reznick, DN and Ghalambor, CK (2001). The population ecology of contemporary adaptations: What empirical studies reveal about the conditions that promote adaptive evolution. *Genetica*, **112-113**, 183-198.

Ricciardi, A (2007). Are modern biological invasions an unprecedented form of global change? *Conservation Biology*, **21**, 329-336.

Rieseberg, LH, Archer, MA and Wayne, RK (1999). Transgressive segregation, adaptation and speciation. *Heredity*, **83**, 363-372.

Rius, M and Darling, JA (2014). How important is intraspecific genetic admixture to the success of colonising populations? *Trends in Ecology & Evolution*, **29**, 233-242.

Rockman, MV (2012). The qtn program and the alleles that matter for evolution: All that's gold does not glitter. *Evolution*, **66**, 1-17.

Roda, F, Liu, H, Wilkinson, MJ, Walter, GM, James, ME, Bernal, DM, Melo, MC, Lowe, A, Rieseberg, LH and Prentis, P (2013). Convergence and divergence during the adaptation to similar environments by an australian groundsel. *Evolution*, **67**, 2515-2529.

Roman, J and Darling, JA (2007). Paradox lost: Genetic diversity and the success of aquatic invasions. *Trends in Ecology & Evolution*, **22**, 454-464.

Rosenberg, NA (2003). Distruct: A program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137-138.

Rougemont, Q, Gagnaire, PA, Perrier, C, Genthon, C, Besnard, AL, Launey, S and Evanno, G (2017). Inferring the demographic history underlying parallel genomic divergence among pairs of parasitic and nonparasitic lamprey ecotypes. *Molecular Ecology*, **26**, 142-162.

Sakai, A and Sakai, S (2003). Size-dependent ess sex allocation in wind-pollinated cosexual plants: Fecundity vs. Stature effects. *Journal of Theoretical Biology*, **222**, 283-295.

Savolainen, O, Lascoux, M and Merila, J (2013). Ecological genomics of local adaptation. *Nature Review Genetics*, **14**, 807-820.

Sax, DF and Brown, JH (2000). The paradox of invasion. *Global Ecology and Biogeography*, **9**, 363-371.

Scalone, R, Lemke, A, Štefanić, E, Kolseth, A-K, Rašić, S and Andersson, L (2016). Phenological variation in *ambrosia artemisiifolia* l. Facilitates near future establishment at northern latitudes. *PloS One*, **11**, e0166510.

Scheiner, SM (2001). Multiple response variables and multi-species interactions. Design and analysis of ecological experiments. S. M. Scheiner and J. Gurevitch. New York, Chapman & Hall**:** 99-133.

Schoener, TW (1968). The anolis lizards of bimini: Resource partitioning in a complex fauna. *Ecology*, **49**, 704-726.

Schrieber, K, Wolf, S, Wypior, C, Höhlig, D, Hensen, I and Lachmuth, S (2017). Adaptive and non-adaptive evolution of trait means and genetic trait correlations for herbivory resistance and performance in an invasive plant. *Oikos*, **126**, 572-582.

Schweizer, RM, Vonholdt, BM, Harrigan, R, Knowles, JC, Musiani, M, Coltman, D, Novembre, J and Wayne, RK (2016). Genetic subdivision and candidate genes under selection in north american grey wolves. *Molecular Ecology*, **25**, 380-402.

Seehausen, O, Butlin, RK, Keller, I, Wagner, CE, Boughman, JW, Hohenlohe, PA, Peichel, CL, Saetre, G-P, Bank, C and Brännström, Å (2014). Genomics and the origin of species. *Nature Reviews Genetics*, **15**, 176.

Slatkin, M (2008). Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics*, **9**, 477.

Smith, M, Cecchi, L, Skjøth, C, Karrer, G and Šikoparija, B (2013). Common ragweed: A threat to environmental health in europe. *Environment International*, **61**, 115-126.

Smith, SD and Rausher, MD (2011). Gene loss and parallel evolution contribute to species difference in flower color. *Molecular Biology and Evolution*, **28**, 2799-2810.

Sork, VL, Aitken, SN, Dyer, RJ, Eckert, AJ, Legendre, P and Neale, DB (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.

Sork, VL, Squire, K, Gugger, PF, Steele, SE, Levy, ED and Eckert, AJ (2016). Landscape genomic analysis of candidate genes for climate adaptation in a california endemic oak, *quercus lobata*. *American Journal of Botany*, **103**, 33-46.

Stapley, J, Reger, J, Feulner, PG, Smadja, C, Galindo, J, Ekblom, R, Bennison, C, Ball, AD, Beckerman, AP and Slate, J (2010). Adaptation genomics: The next generation. *Trends in Ecology and Evolution*, **25**, 705-712.

Stebbins, GL (1969). The significance of hybridization for plant taxonomy and evolution. *Taxon*, **18**, 26-35.

Stern, DL and Orgogozo, V (2008). The loci of evolution: How predictable is genetic evolution? *Evolution*, **62**, 2155-2177.

Stern, DL and Orgogozo, V (2009). Is genetic evolution predictable? *Science*, **323**, 746-751.

Storz, JF (2016). Causes of molecular convergence and parallelism in protein evolution. *Nature Reviews Genetics*, **17**, 239.

Sun, Y, Brönnimann, O, Roderick, GK, Poltavsky, A, Lommen, ST and Müller-Schärer, H (2017). Climatic suitability ranking of biological control candidates: A biogeographic approach for ragweed management in europe. *Ecosphere*, **8**, e01731.

Szűcs, M, Vahsen, M, Melbourne, B, Hoover, C, Weiss-Lehman, C and Hufbauer, R (2017). Rapid adaptive evolution in novel environments acts as an architect of population range expansion. *Proceedings of the National Academy of Sciences*, **114**, 13501-13506.

Taramarcaz, P, Lambelet, C, Clot, B, Keimer, C and Hauser, C (2005). Ragweed (*ambrosia*) progression and its health risks: Will switzerland resist this invasion? *Swiss Medical Weekly*, **135**, 538-548.

Tenaillon, O, Rodríguez-Verdugo, A, Gaut, RL, McDonald, P, Bennett, AF, Long, AD and Gaut, BS (2012). The molecular diversity of adaptive convergence. *Science*, **335**, 457-461.

Thébaud and Simberloff (2001). Are plants really larger in their introduced ranges? *The American Naturalist*, **157**, 231-236.

Throop, HL (2005). Nitrogen deposition and herbivory affect biomass production and allocation in an annual plant. *Oikos*, **111**, 91-100.

Tian, D, Tooker, J, Peiffer, M, Chung, SH and Felton, GW (2012). Role of trichomes in defense against herbivores: Comparison of herbivore response to woolly and hairless trichome mutants in tomato (*solanum lycopersicum*). *Planta*, **236**, 1053-1066.

Tiffin, P and Ross-Ibarra, J (2014). Advances and limits of using population genetics to understand local adaptation. *Trends in Ecology & Evolution*, **29**, 673-680.

Trucchi, E, Frajman, B, Haverkamp, TH, Schönswetter, P and Paun, O (2017). Genomic analyses suggest parallel ecological divergence in *heliosperma pusillum* (caryophyllaceae). *New Phytologist*, **216**, 267-278.

Turchin, MC, Chiang, CW, Palmer, CD, Sankararaman, S, Reich, D, Genetic Investigation of, ATC and Hirschhorn, JN (2012). Evidence of widespread selection on standing variation in europe at height-associated snps. *Nature Genetics*, **44**, 1015-1019.

Turner, KG, Hufbauer, RA and Rieseberg, LH (2014). Rapid evolution of an invasive weed. *New Phytologist*, **202**, 309-321.

Turner, TL, Levine, MT, Eckert, ML and Begun, DJ (2008). Genomic analysis of adaptive differentiation in *drosophila melanogaster*. *Genetics*, **179**, 455-473.

Uesugi, A, Connallon, T, Kessler, A and Monro, K (2017). Relaxation of herbivore-mediated selection drives the evolution of genetic covariances between plant competitive and defense traits. *Evolution*, **71**, 1700-1709.

Uesugi, A and Kessler, A (2016). Herbivore release drives parallel patterns of evolutionary divergence in invasive plant phenotypes. *Journal of Ecology*, **104**, 876-886.

Uller, T and Leimu, R (2011). Founder events predict changes in genetic diversity during human-mediated range expansions. *Global Change Biology*, **17**, 3478-3485.

Urban, MC, Bocedi, G, Hendry, AP, Mihoub, J-B, Pe'er, G, Singer, A, Bridle, J, Crozier, L, De Meester, L and Godsoe, W (2016). Improving the forecast for biodiversity under climate change. *Science*, **353**, aad8466.

van Boheemen, LA, Atwater, DZ and Hodgins, KA (2018a). Rapid and repeated local adaptation to climate in an invasive plant. *bioRxiv*, doi: https://doi.org/10.1101/420752.

van Boheemen, LA, Bou-Assi, S, Uesugi, A and Hodgins, KA (2018b). Eica fails as an explanation of growth and defence evolution following multiple introductions. *bioRxiv*, doi: https://doi.org/10.1101/435271.

van Boheemen, LA, Lombaert, E, Nurkowski, K, Gauffre, B, Rieseberg, LH and Hodgins, KA (2017a). Multiple introductions, pre-introduction admixture and bridgehead invasion characterize the

introduction history of *ambrosia artemisiifolia* in europe and australia. *Molecular Ecology*, **doi: 10.1111/mec.14293**.

van Boheemen, LA, Lombaert, E, Nurkowski, KA, Gauffre, B, Rieseberg, LH and Hodgins, KA (2017b). Multiple introductions, admixture and bridgehead invasion characterize the introduction history of *ambrosia artemisiifolia* in europe and australia. *Molecular Ecology*, **26**, 5421-5434.

Van der Auwera, GA, Carneiro, MO, Hartl, C, Poplin, R, del Angel, G, Levy-Moonshine, A, Jordan, T, Shakir, K, Roazen, D and Thibault, J (2013). From fastq data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Current Protocols in Bioinformatics*, **43**, 11.10. 11-11.10. 33.

van Kleunen, M, Bossdorf, O and Dawson, W (2018). The ecology and evolution of alien plants. *Annual Review of Ecology, Evolution, and Systematics*.

Van Noordwijk, AJ and de Jong, G (1986). Acquisition and allocation of resources: Their influence on variation in life history tactics. *The American Naturalist*, **128**, 137-142.

Verhoeven, KJ, Macel, M, Wolfe, LM and Biere, A (2011). Population admixture, biological invasions and the balance between local adaptation and inbreeding depression. *Proceedings of the Royal Society of London B: Biological Sciences*, **278**, 2-8.

Wagner, NK, Ochocki, BM, Crawford, KM, Compagnoni, A and Miller, TE (2017). Genetic mixture of multiple source populations accelerates invasive range expansion. *Journal of Animal Ecology*, **86**, 21-34.

Wang, J (2017). The computer program structure for assigning individuals to populations: Easy to use but easier to misuse. *Molecular Ecology Resources*.

Wang, L, Beissinger, TM, Lorant, A, Ross-Ibarra, C, Ross-Ibarra, J and Hufford, MB (2017). The interplay of demography and selection during maize domestication and expansion. *Genome Biology*, **18**, 215.

Wang, Y, Carrillo, J, Siemann, E, Wheeler, GS, Zhu, L, Gu, X and Ding, J (2013). Specificity of extrafloral nectar induction by herbivores differs among native and invasive populations of tallow tree. *Annals of Botany*, **112**, 751-756.

Wang, Y, Siemann, E, Wheeler, GS, Zhu, L, Gu, X and Ding, J (2012). Genetic variation in anti-herbivore chemical defences in an invasive plant. *Journal of Ecology*, **100**, 894-904.

War, AR, Kumar Taggar, G, Hussain, B, Sachdeva Taggar, M, Nair, RM and Sharma, HC (2018). Plant defense against herbivory and insect adaptations. *AoB Plants*.

War, AR, Paulraj, MG, Ahmad, T, Buhroo, AA, Hussain, B, Ignacimuthu, S and Sharma, HC (2012). Mechanisms of plant defense against insect herbivores. *Plant signaling & behavior*, **7**, 1306-1320.

Weaver, SE (2001). Impact of lamb's-quarters, common ragweed and green foxtail on yield of corn and soybean in ontario. *Canadian Journal of Plant Science*, **8**, 821-828.

Weinreich, DM, Delaney, NF, DePristo, MA and Hartl, DL (2006). Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science*, **312**, 111-114.

Weinreich, DM, Watson, RA and Chao, L (2005). Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. *Evolution*, **59**, 1165-1174.

Weir, BS and Cockerham, CC (1984). Estimating f-statistics for the analysis of population structure. *Evolution*, **38**, 1358-1370.

WHO, WHO (1998). Global solar uv index. Global Solar UV Index.

Willemsen, RW (1975). Dormancy and germination of common ragweed seeds in the field. *American Journal of Botany*, **62**, 639-643.

Williams, JW, Shuman, BN, Webb, T, Bartlein, PJ and Leduc, PL (2004). Late-quaternary vegetation dynamics in north america: Scaling from taxa to biomes. *Ecological Monographs*, **74**, 309-334.

Wolkovich, EM, Cook, BI, Allen, JM, Crimmins, TM, Betancourt, JL, Travers, SE, Pau, S, Regetz, J, Davies, TJ and Kraft, NJ (2012). Warming experiments underpredict plant phenological responses to climate change. *Nature*, **485**, 494.

Wood, TE, Burke, JM and Rieseberg, LH (2005). Parallel genotypic adaptation: When evolution repeats itself. Genetics of adaptation, Springer**:** 157-170.

Woods, EC, Hastings, AP, Turley, NE, Heard, SB and Agrawal, AA (2012). Adaptive geographical clines in the growth and defense of a native plant. *Ecological Monographs*, **82**, 149-168.

Wright, S (1931). Evolution in mendelian populations. *Genetics*, **16**, 97-159.

Yang, J, Lee, SH, Goddard, ME and Visscher, PM (2011). Gcta: A tool for genome-wide complex trait analysis. *The American Journal of Human Genetics*, **88**, 76-82.

Yang, J, Zaitlen, NA, Goddard, ME, Visscher, PM and Price, AL (2014). Advantages and pitfalls in the application of mixed-model association methods. *Nature genetics*, **46**, 100.

Yeaman, S (2013). Genomic rearrangements and the evolution of clusters of locally adaptive loci. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 1743-1751.

Yeaman, S, Gerstein, AC, Hodgins, KA and Whitlock, MC (2018). Quantifying how constraints limit the diversity of viable routes to adaptation. *bioRxiv*, 279661.

Yeaman, S, Hodgins, KA, Lotterhos, KE, Suren, H, Nadeau, S, Degner, JC, Nurkowski, KA, Smets, P, Wang, T and Gray, LK (2016). Convergent local adaptation to climate in distantly related conifers. *Science*, **353**, 1431-1433.

Yeaman, S and Whitlock, MC (2011). The genetic architecture of adaptation under migration-selection balance. *Evolution*, **65**, 1897-1911.

Zandt, PAV (2007). Plant defense, growth, and habitat: A comparative assessment of constitutive and induced resistance. *Ecology*, **88**, 1984-1993.

Zhang, D-Y (2006). Evolutionarily stable reproductive investment and sex allocation in plants. Ecology and evolution of flowers. L. D. Harder and R. D. Barrett. Oxford, Oxford University Press**:** 41-60.

Zhang, YY, Zhang, DY and Barrett, SC (2010). Genetic uniformity characterizes the invasive spread of water hyacinth (*eichhornia crassipes*), a clonal aquatic plant. *Molecular Ecology*, **19**, 1774-1786.

Ziska, L, Knowlton, K, Rogers, C, Dalan, D, Tierney, N, Elder, MA, Filley, W, Shropshire, J, Ford, LB and Hedberg, C (2011). Recent warming by latitude associated with increased length of ragweed pollen season in central north america. *Proceedings of the National Academy of Sciences*, **108**, 4248-4251.

Ziska, LH and Caulfield, FA (2000). Rising co_2 and pollen production of common ragweed (*ambrosia artemisiifolia*), a known allergy-inducing species: Implications for public health. *Australian Journal of Plant Physiology*, **27**, 893-898.

Zurell, D, Thuiller, W, Pagel, J, Cabral, JS, Münkemüller, T, Gravel, D, Dullinger, S, Normand, S, Schiffers, KH and Moore, KA (2016). Benchmarking novel approaches for modelling species range dynamics. *Global Change Biology*, **22**, 2651-2664.

Züst, T and Agrawal, AA (2017). Trade-offs between plant growth and defense against insect herbivory: An emerging mechanistic synthesis. *Annual Review of Plant Biology*, **68**, 513-534.