Title

Expansion history and environmental suitability shape effective population size in a plant invasion

Authors

Joseph E Braasch*¹

jbraasch@email.arizona.edu

ORCID 0000-0001-7502-1517

Brittany S Barker¹

Bbarker505@gmail.com

Katrina M Dlugosch¹

kdlugoasch@email.arizona.edu

¹Department of Ecology & Evolutionary Biology, University of Arizona, Tucson, AZ, USA

*Author for correspondence

PO Box 210088, Tucson, AZ, 85712 / tel: 520-336-7623 / fax: 520-621-9190 /

jbraasch@email.arizona.edu

Keywords Ecological Genetics, Invasive Species, Contemporary Evolution, Effective Population Size, RAD-Seq

Running Title Range expansion and effective population size

Article Type Original Article

Main text words 5078

Abstract

The margins of an expanding range are predicted to be challenging environments for adaptation. Marginal populations should often experience low effective population sizes (N_e) where genetic drift is high due to demographic expansion and/or census population size is low due to unfavorable environmental conditions. Nevertheless, invasive species demonstrate increasing evidence of rapid evolution and potential adaptation to novel environments encountered during colonization, calling into question whether significant reductions in N_e are realized during range expansions in nature. Here we report one of the first empirical tests of the joint effects of expansion dynamics and environment on effective population size variation during invasive range expansion. We estimate contemporary values of N_e using rates of linkage disequilibrium among genome-wide markers within introduced populations of the highly invasive plant *Centaurea solstitialis* (yellow starthistle) in North America (California, USA), as well as in native European populations. As predicted, we find that N_e within the invasion is positively correlated with both expansion history (time since founding) and habitat quality (abiotic climate). History and environment had independent additive effects with similar effect sizes, supporting an important role of both factors in this invasion. These results support theoretical expectations for the population genetics of range expansion, though whether these processes can ultimately arrest the spread of an invasive species remains an open question.

Introduction

Adaptation is expected to be a critical component of how species respond to novel environmental conditions, such as those encountered during colonization and range expansion (Mayr 1963; Kirkpatrick & Barton 1997; Griffith & Watson 2006; Colautti & Barrett 2013; Bock et al. 2015; Hamilton et al. 2015). At the same time, it has been suggested that colonizing species might often experience small population sizes that limit the ability of founding populations to respond to natural selection (Elam et al. 2007; Dlugosch et al. 2015b; Welles & Dlugosch 2018). Small population sizes could result from both founder events and maladaptation to novel environments. A failure to adapt under these conditions could slow or limit range expansion and contribute to the formation of range limits (Bridle & Vines 2007; Eckert et al. 2008; Sexton et al. 2009). These effects are currently an active area of theoretical and experimental research (Gilbert et al. 2017; Szűcs et al. 2017a; Szűcs et al. 2017b), but there is little empirical information about the dynamics of population size and its influence on evolution during ongoing range expansions in the wild (Ramakrishnan et al. 2010; Wootton and Pfister 2015).

Population genetic models predict that adaptation can fail during range expansion due to the strong effects of genetic drift during colonization (Lehe et al. 2012; Peischl et al. 2013; Peischl et al. 2015). Range expansions are expected to involve a series of founding events, resulting in reduced effective population size (N_e) and increased sampling effects as the range boundary advances (Le Corre & Kremer 1998; Excoffier 2004; Slatkin & Excoffier 2012). In particular, low N_e at the leading edge can cause random alleles, including deleterious mutations, to ‘surf’ to high frequency regardless of patterns of selection (Hallatschek et al. 2007; Excoffier and Ray 2008; Excoffier et al. 2009; Moreau et al. 2011). This can create an ‘expansion load’ of deleterious alleles at the wave front, but it is also possible that beneficial mutations can surf to

high frequency and aid in local adaptation ([Lehe et al. 2012](#); [Peischl et al. 2013](#); [Peischl et al. 2015](#)). The effects of range expansion on adaptation have been empirically observed with greatest detail in bacterial culture, where manipulative experiments have shown that the strength of genetic drift is key to determining whether allele surfing promotes or hinders adaptation ([Hallatschek & Nelson 2010](#); [Gralka et al. 2016](#)).

Ecological conditions should also shape N_e during range expansion via their impact on population (census) size and demography. If leading edge environments are novel relative to those experienced by source populations, then founding genotypes will not be pre-adapted and are likely to experience lower absolute fitness. When populations are at relatively low abundance and/or fluctuate in size due to unfavorable conditions, N_e will be reduced relative to larger or more stable populations ([Wright 1938](#); [Crow & Morton 1955](#); [Kimura & Crow 1963](#); [Frankham 1996](#)). In a rare empirical example, [Micheletti & Storfer \(2015\)](#) found that streamside salamander (*Ambystoma barbouri*) populations on the periphery of the range were also on the margins of their climatic niche and tended toward lower N_e . Similarly, peripheral populations of North American *Arabidopsis lyrata* possess greater genetic load and appear to exist at their ecological, and perhaps evolutionary limits ([Willi et al. 2018](#)). These studies address a set of long-debated hypotheses proposing that range limits form in part because they include ecologically and/or genetically marginal populations ([Kirkpatrick & Barton 1997](#); [Phillips 2012](#); [Chuang & Peterson 2016](#)), which fail to support local adaptation and further expansion (i.e. the ‘central-marginal’, ‘center-periphery’ and ‘abundant center’ hypotheses: [Sagarin and Gaines 2002](#); [Eckert et al. 2008](#); [Pironon et al. 2015](#)). Importantly, all the above hypotheses share the prediction that colonization will be associated with reduced response to selection for ecological reasons independent of any population genetic effects of range expansion itself. The relative

importance of these two factors (ecology and expansion) for shaping N_e at range margins is unknown, but both have the potential to reduce opportunities for local adaptation.

Although N_e has long been used as a fundamental measure of the scale of genetic drift in populations ([Wright 1931](#); [Robertson 1960](#); [Kimura and Crow 1963](#); [Kimura 1964](#); [Ohta 1992](#); [Charlesworth 2009](#)), there is surprisingly little known about how N_e changes during the process of range expansion. Most empirical population-level estimates come from the field of conservation genetics, where N_e is used to infer the potential for genetic drift to exacerbate the decline of threatened populations ([Lynch et al. 1995](#); [Frankham 1996](#); [Sung et al. 2012](#)). These studies have demonstrated that N_e can be highly variable within species, sensitive to local demography and modes of reproduction, and captured only in part by measures of census size ([Frankham 1995](#); [Palstra and Ruzzante 2008](#)). For example, in recovering Chinook salmon (*Oncorhynchus tshawytscha*) populations, [Shrimpton and Heath \(2003\)](#) found up to a three-fold difference in both N_e and its ratio with census size across spawning sites. While low N_e is generally expected in declining populations, many of the same demographic factors are likely to affect N_e in founding populations ([Colautti et al. 2017](#)).

Despite the potential obstacle low N_e might pose to adaptation, many species -- including large numbers of invaders -- have been successful at colonization and show evidence of adaptive evolution during range expansion ([Rice and Mack 1991](#); [Dlugosch and Parker 2008](#); [Linnen et al. 2009](#); [Colautti and Barrett 2013](#); [Vandepitte et al. 2014](#); [Colautti and Lau 2015](#); [Li et al. 2015](#)). At the same time, detailed studies of range expansion have found some evidence of serial founding events and associated increases in genetic drift ([Ramakrishnan et al. 2010](#); [Graciá et al. 2013](#); [White et al. 2013](#); [Pierce et al. 2014](#)), and it is notable that few invasions have been convincingly shown to have expanded beyond the fundamental niches of their native range

([Petitpierre et al. 2012](#); [Tingley et al. 2014](#)). Taken together, it appears that adaptive evolution might be achievable in many invading species, but that perhaps expansion load and ecological mismatch may act, either independently or in concert, to prevent further expansion under some conditions. An understanding of how founding dynamics and marginal environments have shaped N_e in individual wave front populations is needed to connect theoretical expectations to these observed patterns of successful range expansion.

Here we estimate contemporary N_e for populations of the obligately outcrossing annual plant *Centaurea solstitialis* (yellow starthistle) across its invasion of California (USA) and its native range in Eurasia. In California, *C. solstitialis* was initially introduced in the mid 19th century into the San Francisco Bay area as a contaminant of alfalfa seed ([Gerlach 1997](#); [DiTomaso et al. 2006](#)). By the mid 20th century, the species was rapidly expanding through California's Central Valley and Sierra Nevada foothill grasslands, and the current leading edge of this invasion lies above 4000 m in elevation along the west side of the Sierra Nevada Mountains ([Pitcairn et al. 2006](#)). During this history of invasion, *C. solstitialis* has crossed climatic environmental gradients that are largely independent in direction from the pathway of colonization (Fig. 1), allowing us to quantify the influence of both climate and expansion history on estimates of N_e across populations.

We used Restriction-site Associated DNA sequencing (RADseq) to estimate contemporary N_e in *C. solstitialis* populations sampled at a single time point. In addition to testing for the joint influence of expansion dynamics and ecological conditions on N_e in this system, we explored solutions for general problems associated with using large genome-wide marker data sets like this to estimate N_e . Linkage disequilibrium N_e (LD- N_e) is a powerful method for inferring contemporary N_e from single time sampled data, and does so by utilizing the

frequency of statistical linkage across loci ([Waples and Do 2008](#); [Gilbert and Whitlock 2015](#)).

This method requires that loci segregate independently of each other, and while RADseq is widely used to produce population genetic datasets in non-model systems ([Narum et al. 2013](#); [Catchen et al. 2017](#)), it is likely to violate this assumption of independence, resulting in biased calculations of N_e .

We used marker resampling and rarefaction approaches to improve inferences of variation in N_e across populations. We tested for effects of expansion history (time since founding) and habitat quality (climatic variables) on rarified N_e estimates, and compared these values to those from populations in the native range. We also explored the ability of genetic diversity measures to predict values of N_e , given that non-equilibrium population dynamics may in the short term decouple contemporary N_e from its expected long term effects on genetic variation (e.g. [Nei et al. 1975](#); [Varvio et al. 1986](#); [Alcala et al. 2013](#); [Epps and Keyghobadi 2015](#)). Together these analyses shed light on the factors shaping the fundamental parameters of evolution during colonization and range expansion.

Materials and Methods

Genomic Data

Genome-wide markers for *C. solstitialis* in this study were sampled from single nucleotide polymorphisms in double-digest RADseq (ddRADseq; ([Peterson et al. 2012](#)), previously published by Barker and colleagues ([Barker et al. 2017](#); Dryad doi:10.5061/dryad.pf550). All sequences were obtained from *C. solstitialis* individuals germinated in the laboratory from wild collected seed. Seeds were sampled in 2008 from maternal plants along a linear transect in each population, with >1m separation between

individuals. Populations included at least 13 individuals each grown from different maternal
 180 plants, from 12 invading populations in California and 7 native populations in Europe (451
 individuals total; Table S1).

Briefly, sequence data published by Barker and colleagues (2017) were generated as
 follows. Genomic DNA was extracted with a modified CTAB protocol ([Webb and Knapp 1990](#))
 and fragmented using *Pst*I and *Mse*I restriction enzymes. Samples were individually barcoded,
 185 cleaned and size selected for fragments between 350 and 650 bp. Size selected fragments were
 amplified through 12 PCR cycles and sequenced on an Illumina HiSeq 2000 or 2500 platform
 (Illumina, Inc., San Diego, CA USA) to generate 100 bp paired-end reads. Reads were de-
 multiplexed with custom scripts and cleaned with the package SNOWHITE 2.0.2 ([Dlugosch et
 al. 2013](#)) to remove primer and adapter contaminants. Barcode and enzyme recognition
 190 sequences were removed from individual reads, and bases with phred quality scores below 20
 were clipped from the 3' end. Reads were trimmed to a uniform length of 76 base pairs. The R2
 (reverse) reads from the data set were removed due to variable quality, and all analyses in this
 study were conducted using R1 (forward) reads only.

We used the denovo_map.pl pipeline in STACKS 1.20 ([Catchen et al. 2011; Hohenlohe
 195 et al. 2011](#)) to identify putative alleles within individuals, allowing a maximum of two nucleotide
 polymorphisms when merging reads (-M parameter in STACKS) a maximum of two alleles per
 locus (-X), and a minimum coverage depth of 5 (-m). A catalog of loci and single nucleotide
 polymorphisms (SNPs) was generated across individuals, allowing two polymorphisms (-n)
 between individuals within a stack. The population.pl module in STACKS was used to calculate
 200 the population level nucleotide diversity (π) ([Nei & Li 1979; Allendorf 1986](#)). We restricted our

analyses to those loci that were sequenced in 80% of individuals within a population and 90% of all populations (-r and -p parameters respectively).

Estimates of N_e

205 We used SNPs identified by STACKs to calculate N_e for each population using a method based on linkage disequilibrium among loci with a correction for missing data ([Waples & Do 2008](#)) implemented in the program NeEstimator v.2.01 ([Do et al. 2014](#)). This method derives estimates of N_e from the frequency of statistical linkage among loci and has been shown to be one of the best predictors of N_e (hereafter LD- N_e) for markers sampled at a single time point 210 ([Gilbert & Whitlock 2015; Wang 2016; Waples 2016](#)). The LD- N_e method is not strongly influenced by the total genetic diversity in the sample ([Charlesworth 2009; Do et al. 2014](#)), making it particularly well suited to analyses of invading populations where low genetic diversity might arise from founder effects unrelated to the number of individuals currently reproducing in the population. We used a minimum allele frequency threshold of 0.05 for 215 including a locus in the analyses, which was the lowest threshold that did not result in excessive loss of loci and infinite estimates of N_e at some study sites.

The ddRADseq data yielded thousands of SNPs across the genome, 622 of which passed our screening requirements (see Results). Some of these loci were located in the same RAD 76bp sequence, and we expect that these and many others do not segregate independently in our data 220 set, either due to physical proximity or the influence of selection on multi-locus allele combinations (*C. solstitialis* has a genome size of 850Mbp, distributed across 8 chromosomes ([Bancheva & Greilhuber 2005; Widmer et al. 2007](#))). To increase the likelihood of basing our

estimates on independent loci, we re-sampled random sets of 20 polymorphic SNPs from unique sequences to obtain distributions of LD- N_e estimates for each population. Each population was
 225 resampled 30,000 times. Sampling distributions were generally lognormal (Supporting Information Fig. S1), and we used medians to identify the peak estimate.

We observed a strong, positive effect of the number of individuals sampled in each population on median LD- N_e ($F_{1,17}=9.36$, $P=0.007$). Unequal sampling has been shown to decrease the accuracy of LD- N_e estimates ([England et al. 2006; Authors](#)), and NeEstimator implements a
 230 corrective algorithm to address this problem ([Do et al. 2014](#)). To account for persistent sampling effects we produced rarefaction curves of median N_e estimated by subsampling different numbers of individuals (10 to the maximum number available per population) after marker resampling. As above, each marker resampling consisted of 30,000 N_e estimates with 20 loci. Median estimates did not asymptote at our maximum sampling effort and increased linearly (see Results). We fit a
 235 linear mixed model with random intercept and slopes implemented in the Lme4 package in R ([Bates et al. 2014](#)) to obtain population specific functions which describe the relationships between the number of individuals sampled in each population and median LD- N_e values. The estimated slope and intercepts for each population were extracted from the model and used to calculate rarefied N_e for each population at a standard value of 10 individuals. We explored the
 240 relationship between rarefied N_e and measures of genome-wide marker variation using nucleotide diversity (π) at variable sites, as calculated in STACKS. We used linear regression to predict nucleotide diversity from N_e within geographic region (invading California populations or native European populations).

245 *Effects of Expansion History and Environment on N_e*

Braasch et al. – Range expansion and effective population size

We tested for an effect of population age since founding on the rarefied N_e of invading populations. We estimated the date of colonization for each population by searching the Jepson Online Herbarium (<http://ucjeps.berkeley.edu/>) for records of *C. solstitialis* in California since its first record in 1869. For each sampling location, we used the earliest date on record for the county, or for an adjacent county when the sampling location was closer to older collection records there. These dates were subtracted from the year of our seed collections (2008) to produce values of population age used in subsequent analyses. Using herbaria records to assign population ages in this manner may not represent the true population age because of the time between population founding and the first records of the population. Nevertheless, *C. solstitialis* has a relatively well-documented invasion history in California (858 specimens on record, 577 records with GPS data, 61 records prior to 1930 in the Jepson Herbarium), and our population age estimates are in line with historical reconstructions of a general pattern of expansion out from initial establishment in the San Francisco Bay area first to the Central Valley and then to the North, East, and South ([Gerlach 1997](#); [DiTomaso et al. 2006](#); [Pitcairn et al. 2006](#)).

We also tested for the influence of the climatic environment on rarefied N_e in both invading and native populations. To quantify the climatic gradients that might be most relevant to *C. solstitialis* ecology, we used the first two principal component (PC) axes of climatic variation across *C. solstitialis* collection sites in North America and Europe, as previously identified by Dlugosch and colleagues (2015a; Fig. 1). Climatic variables at 18.5 x 18.5 km resolution ([Kriticos et al. 2012](#)) were extracted for each site using ArcGIS. Larger values along the first PC climate axis generally indicate sites with higher temperatures and lower seasonality in total radiation. Larger PC2 values indicate lower annual precipitation and greater seasonality in temperature. Greater seasonal variation in temperature has been shown to be related to

ecologically important traits (plant size and drought tolerance) in *C. solstitialis* in both the native
 270 and invaded ranges ([Dlugosch et al. 2015a](#)).

To quantify the contributions of both population age and climatic environment to
 variation in rarified N_e for the invaded range, we used a general linear model with N_e as the
 dependent variable and population age, climate PC1, climate PC2, and their interactions as
 explanatory variables. We constructed a separate model of rarified N_e in native range populations
 275 using only PC1 and PC2 as variables, since no information about population age is available for
 the native range. We used model decomposition and F-scores to identify the best fit model. To
 explore the relative effect of each variable and their interactions on N_e , effect sizes were
 calculated as partial eta-squared values, which partition the total variance in a dependent variable
 among all independent variables (analogous to R^2 in multiple regression), using the best fit linear
 280 model with the function ‘etasq’ in the R package ‘heplots’ version 1.3-1 ([Fox et al. 2008](#)). We
 also tested for an overall difference in the rarified N_e of invading and native populations using a
 Wilcoxon signed-rank test and a Monte Carlo exact test implemented with the coin package in R
 ([Hothorn et al. 2006](#)).

285 Results

We obtained a dataset of 622 SNPs from our filtering pipeline which were used to
 estimate N_e . For each population, subsampling produced a wide range of estimates of LD- N_e ,
 contingent upon which 20 loci were sampled (Supporting Information Fig. S1). These
 distributions spanned at least 4 orders of magnitude for every population. Distributions peaked
 290 strongly around median estimates (Supporting Information Fig. S1). Median estimates of LD- N_e
 prior to rarefaction varied from 19.5 to 38.5 across the California invasion (Supporting

Information Table S1). In general, estimates were higher in central and northern California and decreased to the East and South (Fig. 1). In native populations, median N_e estimates ranged from 16.2 to 42.7, with the three populations with lowest LD- N_e located on the western side of the range in Spain (Fig. 1).

We found a strong association between median LD- N_e and the number of individuals sampled from a population ($r^2_{\text{adj}}=0.32$, $F_{1,17}=9.36$, $P=0.007$). Rarefaction sampling produced positive relationships between LD- N_e and the number of individuals resampled within each population (Fig. 2). Slopes ranged from ~0.11 to 1.19. Importantly, rarefaction removed the significant effect of sampling effort on N_e values (rarefied N_e vs. total sample size; $r^2_{\text{adj}}=0.12$, $F_{1,17}=3.47$, $P=0.08$).

Both climate and population age predicted rarefied N_e in invading populations. The best fit linear model ($r^2_{\text{adj}}=0.46$, $F_{(4,8)}=4.09$, $P=0.0493$) included significant, additive effects of population age and PC2 (Fig. 3; Table 1). Population age and PC2 were both positively correlated with rarefied N_e values, indicating that N_e is largest in older populations and habitats with more temperature seasonality and lower precipitation. PC2 had a greater influence on rarefied N_e values than age, based its larger effect size (Table 1), although this difference was small. In contrast, rarefied N_e of native range populations was not predicted by either climatic PC variable (Full model: $r^2_{\text{adj}}=0.2314$, $F_{2,4}=1.90$, $P=0.23$) (Interactions: PC1: $P=0.13$, PC2: $P=0.20$).

Invading populations included a narrower range of rarefied N_e values, nested within the distribution observed for native populations (Supporting Information Fig S2), and there was no significant difference between rarefied N_e values in the native and invading ranges (Wilcoxon signed rank test: $W = 43$, $P = 0.97$; Monte-Carlo one-way exact test: $P = 0.93$). Nucleotide diversity (π) also did not differ overall between the native and invaded ranges (Model: $r^2_{\text{adj}}=0$.-

0.04, $F_{(3,15)}=0.78$, $P=0.55$; region term: $t_{(2,15)}=-1.5$, $P=0.15$). There was no significant relationship between π and N_e in invaded range populations ($r^2_{\text{adj}}=-0.037$, $F_{(1,10)}=0.59$, $P=0.46$) and there was a positive, marginally significant relationship between nucleotide diversity and N_e in native range populations ($r^2_{\text{adj}}=0.43$, $F_{(1,5)}=5.45$, $P=0.067$), despite a smaller sample size in this range (Supporting Information Fig S3).

Discussion

Here we report some of the first empirical evidence for the joint effects of both range expansion and climatic environment on contemporary N_e in natural populations. We produced rarified estimates of LD- N_e across 12 populations in the invaded range of *C. solstitialis* and found a significant positive relationship between population age and N_e , a finding in line with theoretical expectations for the population genetics of expanding populations ([Hallatschek et al. 2007](#); [Excoffier & Ray 2008](#); [Excoffier et al. 2009](#); [Moreau et al. 2011](#); [Lehe et al. 2012](#); [Peischl et al. 2013](#); [Peischl et al. 2015](#)). We also find evidence that spatial variation in ecological conditions, specifically climate in our study, has a significant impact on N_e which is independent of population age. These effects of range expansion and climate were similar in magnitude, suggesting that both of these factors are likely to be important for shaping evolutionary outcomes in invading populations.

We emphasize that our rarified LD- N_e values do not reflect a ‘true’ N_e value for the populations in our study. Rather, rarified estimates here represent *relative* values of N_e , and are thus useful for comparisons among populations. We expect asymptotic LD- N_e values for these populations to be larger, because we observed no asymptote with rarefaction for any of the populations in our study. Despite this, our estimates are similar to values reported in other plant *Braasch et al. – Range expansion and effective population size*

and animal species using the same approach (e.g. [Shrimpton & Heath 2003](#); [Coyer et al. 2008](#); [Wang et al. 2013](#); [Álvarez et al. 2015](#)). The LD- N_e estimation method itself also has a tendency
340 to underestimate known values of N_e in simulations ([Gilbert & Whitlock 2015](#))

We found that subsampling of loci was important for using a genome-wide marker dataset to estimate LD- N_e . Simulation studies of LD- N_e typically use on the order of 20 loci ([England et al. 2006](#); [Waples 2006](#) ; [Waples & Do 2008](#); [Gilbert & Whitlock 2015](#)). Our resampling of 20 loci revealed that N_e estimates in *C. solstitialis* vary by at least 4 orders of
345 magnitude when different sets of loci are used. This variation is expected given that particular sets of loci will capture different effects of physical linkage, history of selection, and chance sampling effects ([Daly et al. 2001](#); [Remington et al. 2001](#); [Flint-Garcia et al. 2003](#)). Resampling allowed us to leverage many combinations of loci across the genome to identify a well defined peak in the distribution of estimates. A resampling approach is likely to be generally useful for
350 RAD-seq and other popular genome wide marker methods, particularly where a complete reference genome is not available to control for the physical arrangement of loci.

After accounting for these methodological issues, we found a significant effect of population age on differences in N_e across populations. Rarefied N_e estimates were lower in younger populations, which fits with expectations that a subset of individuals will contribute to
355 range expansion ([Excoffier & Ray 2008](#)) and that genetic drift will be larger at the leading edge ([Lehe et al. 2012](#); [Peischl et al. 2013](#); [Peischl et al. 2015](#)). Estimates of contemporary N_e from our invading populations are within the distribution that we observe from native populations, suggesting that this species did not experience a large initial genetic bottleneck during introduction to the New World, nor exceptionally low N_e during range expansion (relative to
360 values observed in native populations). This lack of evidence for a strong genetic bottleneck is in

line with models of historical demography by Barker and colleagues (2017), who inferred little reduction in population size and maintenance of genetic diversity during the colonization of the Americas by *C. solstitialis*. In general, introduced species often lack strong genetic bottlenecks (Dlugosch & Parker 2008; Uller & Leimu 2011; Dlugosch et al. 2015b), and our results here demonstrate that species which avoid genetic bottlenecks during introduction may still experience declines in N_e at the leading edge of range expansion. Importantly, invading populations of *C. solstitialis* are an order of magnitude higher in density than native populations (Uygur et al. 2004; Andonian et al. 2011), indicating that the fraction of the census population that is contributing to evolutionary dynamics in the invasion is much lower than in the native range.

We also observed an independent positive relationship between climatic PC2 and N_e of invading *C. solstitialis* populations, consistent with an impact of habitat suitability on N_e . High PC2 values reflect greater variation in seasonal temperatures and lower total annual precipitation, which typify areas of especially high *C. solstitialis* density in California (Dlugosch et al. 2015a). Previous studies in this system have proposed that *C. solstitialis* success stems from a lack of effective competitors in more drought prone habitats (Dlugosch et al. 2015a), due in part to the extensive conversion of these habitats to rangeland (Menke 1989, (Stromberg & Griffin 1996)). These patterns of *C. solstitialis* density can be counterintuitive, as studies within the California invasion have found that water availability (both naturally occurring and experimentally manipulated) is strongly and positively correlated with *C. solstitialis* density and fecundity (Enloe et al. 2004; Morghan & Rice 2006; Hulvey & Zavaleta 2012; Eskelinen & Harrison 2014), suggesting that fitness should be highest in wetter areas. Our results are most consistent with the landscape pattern of abundant *C. solstitialis* in drier areas, and might therefore reflect

differences in human land use and the availability of native competitors across the invaded
 385 range. An underlying relationship between N_e and land use in the invasion could also explain
 why we did not find a relationship with climate in the native range, where other features might
 shape habitat suitability for this species.

These differences in rarified N_e among invading populations were not predicted by
 sequence diversity (π). Nonequilibrium populations such as the invasions here are unlikely to
 390 have had sufficient time to reach equilibrium diversity at a given N_e , and will also have been
 changing in size over time ([Nei et al. 1975](#); [Alcala et al. 2013](#); [Epps & Keyghobadi 2015](#)).
 Notably, we did find a marginally significant positive relationship between π and N_e in native
 range populations, which will have had more time to stabilize in population size and reach
 mutation-drift equilibrium. Moreover, rare alleles contribute important equilibrium genetic
 395 variation ([Luikart et al. 1998](#)) and native *C. solstitialis* populations have been previously shown
 to harbor more rare alleles than invading populations in North America ([Barker et al. 2017](#)).
 There is also a tendency for RAD-seq to underestimate π in more diverse genomes ([Arnold et al.](#)
[2013](#); [Cariou et al. 2016](#)), although given the loss of rare alleles from invading populations, we
 might expect this to affect native populations more strongly than invading populations.

400 Our results support the prediction that both range expansion and habitat quality can
 increase the genetic drift experienced by leading edge populations. There is particular interest in
 whether these effects can hinder adaptation, slow further colonization, and establish static range
 boundaries ([Bosshard et al. 2017](#); [Lehe et al. 2012](#); [Peischl et al. 2013](#); [Peischl et al. 2015](#);
[Marculis et al. 2017](#); [Birzu et al. 2018](#)). Recent studies have demonstrated a link between
 405 differences in historical values of N_e and differences in efficacy of selection across species (e.g.
[Slotte et al. 2010](#); [Jensen & Bachtrog 2011](#); [Strasburg et al. 2011](#)), and both theoretical and

experimental studies of bacteria have shown that the process of range expansion can reduce contemporary N_e and impose limits to adaptation and further colonization at the expansion front ([Hallatschek & Nelson 2010](#); [Lehe et al. 2012](#); [Gralka et al. 2016](#); [Peischl et al. 2013](#); [Peischl et al. 2015](#)). Natural populations of *Arabidopsis lyrata* have been shown to demonstrate some of these effects, with greater genetic load in range edge populations associated with a lack of adaptation along an environmental cline ([Willi et al. 2018](#)). Limits to range expansion are expected to be sensitive to the specifics of evolutionary parameters in natural populations, including the magnitudes of N_e and selection, the amount and scale of gene flow across the expansion, and the genetic architecture of adaptive variation ([Hallatschek & Nelson 2010](#); [Lehe et al. 2012](#); [Peischl et al. 2013](#); [Peischl et al. 2015](#); [Gralka et al. 2016](#)).

The expansion ecology of *C. solstitialis* does not support the existence of maladapted edge populations. Populations of *C. solstitialis* closer to the range edge can achieve greater density than more interior populations in some cases ([Swope et al. 2017](#)), which runs counter to any expectations of high genetic load. Additionally, evolution of increased growth and earlier flowering appears to be enhancing the invasiveness of *C. solstitialis* ([Dlugosch et al. 2015a](#)), suggesting that reduced N_e at the range edge has not created a barrier to adaptation and further expansion. On the other hand, in areas where we would expect adaptation to be particularly critical for colonization, such as stressful serpentine soil habitats in California, *C. solstitialis* is notably rare or absent ([Dukes 2002](#); [Gelbard & Harrison 2003](#)), though this might simply reflect the absence of relevant adaptive genetic variation. The availability of adaptive variation and the degree to which this is a limiting factor in species invasions is an active area of debate ([Ellstrand & Schierenbeck 2000](#); [Rius & Darling 2014](#); [Bock et al. 2015](#)), and should be particularly relevant to the colonization of habitats requiring significant niche evolution. The results reported

here emphasize that an understanding of the evolutionary mechanisms that generate boundaries to range expansion in natural populations will require evaluating evidence not only for availability of adaptive variation ([Dlugosch et al. 2015a](#)), but also for an effective response to selection.

References

- Alcala, N., Streit, D., Goudet, J., & Vuilleumier, S. (2013). Peak and persistent excess of genetic diversity following an abrupt migration increase. *Genetics*, **193**(3), 953–971.
- Allendorf, F. W. (1986). Genetic drift and the loss of alleles versus heterozygosity. *Zoo Biology*, **5**(2), 181–190.
- Álvarez, D., Lourenço, A., Oro, D., & Velo-Antón, G. (2015). Assessment of census (N) and effective population size (Ne) reveals consistency of Ne single-sample estimators and a high Ne/N ratio in an urban and isolated population of fire salamanders. *Conservation Genetics Resources*, **7**(3), 705–712.
- Andonian, K., Hierro, J. L., Khetsuriani, L., Becerra, P., Janoyan, G., Villarreal, D., ... Callaway, R. M. (2011). Range-expanding populations of a globally introduced weed experience negative plant-soil feedbacks. *PloS One*, **6**(5), e20117.
- Arnold, B., Corbett-Detig, R. B., Hartl, D., & Bomblies, K. (2013). RAD seq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. *Molecular Ecology*, **22**(11), 3179–3190.
- Atwater, D. Z., Ervine, C., & Barney, J. N. (2018). Climatic niche shifts are common in introduced plants. *Nature Ecology & Evolution*, **2**(1), 34–43.
- Bancheva, S., & Greilhuber, J. (2006). Genome size in Bulgarian *Centaurea s.l.* (Asteraceae). *Plant Systematics and Evolution*, **257**(1), 95–117.
- Barker, B. S., Andonian, K., Swope, S. M., Luster, D. G., & Dlugosch, K. M. (2017). Population genomic analyses reveal a history of range expansion and trait evolution across the native and invaded range of yellow starthistle (*Centaurea solstitialis*). *Molecular Ecology*, **26**(4), 1131–1147.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, **67**(1), 1–48. <http://doi:10.18637/jss.v067.i01>
- Birzu, G., Hallatschek, O., & Korolev, K. S. (2018).. *Proceedings of the National Academy of Sciences of the United States of America*, **115**(16), E3645–E3654.
- Bock, D. G., Caseys, C., Cousens, R. D., Hahn, M. A., Heredia, S. M., Hübner, S., ... Rieseberg, L. H. (2015). What we still don't know about invasion genetics. *Molecular Ecology*, **24**(9), 2277–2297.
- Bosshard, L., Dupanloup, I., Tenaillon, O., Bruggmann, R., Ackermann, M., Peischl, S., & Excoffier, L. (2017). Accumulation of Deleterious Mutations During Bacterial Range Expansions. *Genetics*, **207**(2), 669–684.
- Cariou, M., Duret, L., & Charlat, S. (2016). How and how much does RAD-seq bias genetic diversity estimates? *BMC Evolutionary Biology*, **16**(1), 240.
- Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: building and genotyping Loci de novo from short-read sequences. *G3*, **1**(3), 171–182.
- Catchen, J. M., Hohenlohe, P. A., Bernatchez, L., Funk, W. C., Andrews, K. R., & Allendorf, F. W. (2017). Unbroken: RADseq remains a powerful tool for understanding the genetics of adaptation in natural populations. *Molecular Ecology Resources*, **17**(3), 362–365.

- Charlesworth, B. (2009). Fundamental concepts in genetics: effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, **10**(3), 195–205.
- Chuang, A., & Peterson, C. R. (2016). Expanding population edges: theories, traits, and trade-offs. *Global Change Biology*, **22**(2), 494–512.
- Colautti, R. I., Alexander, J. M., Dlugosch, K. M., Keller, S. R., & Sultan, S. E. (2017). Invasions and extinctions through the looking glass of evolutionary ecology. *Philosophical Transactions of the Royal Society of London. Series B Biological Sciences*, **372**(1712). <http://doi:10.1098/rstb.2016.0031>
- Colautti, R. I., & Barrett, S. C. H. (2011). Population divergence along lines of genetic variance and covariance in the invasive plant *Lythrum salicaria* in eastern North America. *Evolution*, **65**(9), 2514–2529.
- Colautti, R. I., & Lau, J. A. (2015). Contemporary evolution during invasion: evidence for differentiation, natural selection, and local adaptation. *Molecular Ecology*, **24**(9), 1999–2017.
- Coyer, J. A., Hoarau, G., Sjøtun, K., & Olsen, J. L. (2008). Being abundant is not enough: a decrease in effective population size over eight generations in a Norwegian population of the seaweed, *Fucus serratus*. *Biology Letters*, **4**(6), 755–757.
- Crow, J. F., & Morton, N. E. (1955). Measurement of Gene Frequency Drift in Small Populations. *Evolution*, **9**(2), 202–214.
- Daly, M. J., Rioux, J. D., Schaffner, S. F., Hudson, T. J., & Lander, E. S. (2001). High-resolution haplotype structure in the human genome. *Nature Genetics*, **29**, 229.
- DiTomaso, J. M., Kyser, G. B., & Pitcairn, M. J. (2006). Yellow starthistle management guide. California Invasive Plant Council Berkeley, CA.
- Dlugosch, K. M., Cang, F. A., Barker, B. S., Andonian, K., Swope, S. M., & Rieseberg, L. H. (2015a). Evolution of invasiveness through increased resource use in a vacant niche. *Nature Plants*, **1**. <http://doi:10.1038/nplants.2015.66>
- Dlugosch, K. M., Anderson, S. R., Braasch, J., Cang, F. A., & Gillette, H. D. (2015b). The devil is in the details: genetic variation in introduced populations and its contributions to invasion. *Molecular Ecology*, **24**(9), 2095–2111.
- Dlugosch, K. M., Lai, Z., Bonin, A., Hierro, J., & Rieseberg, L. H. (2013). Allele identification for transcriptome-based population genomics in the invasive plant *Centaurea solstitialis*. *G3*, **3**(2), 359–367.
- Dlugosch, K. M., & Parker, I. M. (2008a). Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, **17**(1), 431–449.
- Dlugosch, K. M., & Parker, I. M. (2008b). Invading populations of an ornamental shrub show rapid life history evolution despite genetic bottlenecks. *Ecology Letters*, **11**(7), 701–709.
- Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., & Ovenden, J. R. (2014). NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Molecular Ecology Resources*, **14**(1), 209–214.
- Dukes, J. S. (2002). Comparison of the effect of elevated CO₂ on an invasive species (*Centaurea solstitialis*) in monoculture and community settings. *Plant Ecology*, **160**(2), 225–234.
- Eckert, C. G., Samis, K. E., & Loughheed, S. C. (2008). Genetic variation across species' geographical ranges: the central--marginal hypothesis and beyond. *Molecular Ecology*, **17**(5), 1170–1188.
- Elam, D. R., Ridley, C. E., Goodell, K., & Ellstrand, N. C. (2007). Population size and relatedness affect fitness of a self-incompatible invasive plant. *Proceedings of the National Academy of Sciences of the United States of America*, **104**(2), 549–552.
- Ellstrand, N. C., & Schierenbeck, K. A. (2000). Hybridization as a stimulus for the evolution of invasiveness in plants? *Proceedings of the National Academy of Sciences of the United States of America*, **97**(13), 7043–7050.
- England, P. R., Cornuet, J.-M., Berthier, P., Tallmon, D. A., & Luikart, G. (2006). Estimating effective population size from linkage disequilibrium: severe bias in small samples. *Conservation Genetics*, **7**(2), 303.

- Enloe, S. F., DiTomaso, J. M., Orloff, S. B., & Drake, D. J. (2004). Soil water dynamics differ among
rangeland plant communities dominated by yellow starthistle (*Centaurea solstitialis*), annual grasses,
or perennial grasses. *Weed Science*, **52**(6), 929–935.
- 525 Epps, C. W., & Keyghobadi, N. (2015). Landscape genetics in a changing world: disentangling historical
and contemporary influences and inferring change. *Molecular Ecology*, **24**(24), 6021–6040.
- Eskelinen, A., & Harrison, S. (2014). Exotic plant invasions under enhanced rainfall are constrained by
soil nutrients and competition. *Ecology*, **95**(3), 682–692.
- Excoffier, L. (2004). Patterns of DNA sequence diversity and genetic structure after a range expansion:
530 lessons from the infinite-island model. *Molecular Ecology*, **13**(4), 853–864.
- Excoffier, L., & Ray, N. (2008). Surfing during population expansions promotes genetic revolutions and
structuration. *Trends in Ecology & Evolution*, **23**(7), 347–351.
- Flint-Garcia, S. A., Thornsberry, J. M., & Buckler, E. S., 4th. (2003). Structure of linkage disequilibrium
in plants. *Annual Review of Plant Biology*, **54**, 357–374.
- 535 Fox, J., Friendly, M., & Monette, G. (2008). Visualizing hypothesis tests in multivariate linear models:
the heplots package for R. *Computational Statistics*, **24**(2), 233–246.
- Frankham, R. (1995). Effective population size/adult population size ratios in wildlife: a review.
Genetical Research, **89**(5-6), 491–503.
- Frankham, R. (1996). Relationship of Genetic Variation to Population Size in Wildlife. *Conservation*
540 *Biology*, **10**(6), 1500–1508.
- Gelbard, J. L., & Harrison, S. (2003). Roadless habitats as refuges for native grasslands: interactions with
soil, aspect, and grazing. *Ecological Applications*, **13**(2), 404–415.
- Gerlach, J. D., Jr. (1997). The introduction, dynamics of geographic range expansion, and ecosystem
effects of yellow starthistle (*Centaurea solstitialis*). In *Proceedings of the California Weed Science*
545 *Society* (Vol. 49, pp. 136–141).
- Gilbert, K. J., Sharp, N. P., Angert, A. L., Conte, G. L., Draghi, J. A., Guillaume, F., ... Whitlock, M. C.
(2017). Local Adaptation Interacts with Expansion Load during Range Expansion: Maladaptation
Reduces Expansion Load. *The American Naturalist*, **189**(4), 368–380.
- Gilbert, K. J., & Whitlock, M. C. (2015). Evaluating methods for estimating local effective population
550 size with and without migration. *Evolution*, **69**(8), 2154–2166.
- Graciá, E., Botella, F., Anadón, J. D., Edelaar, P., Harris, D. J., & Giménez, A. (2013). Surfing in
tortoises? Empirical signs of genetic structuring owing to range expansion. *Biology Letters*, **9**(3),
20121091.
- Gralka, M., Stiewe, F., Farrell, F., Möbius, W., Waclaw, B., & Hallatschek, O. (2016). Allele surfing
555 promotes microbial adaptation from standing variation. *Ecology Letters*, **19**(8), 889–898.
- Griffith, T. M., & Watson, M. A. (2006). Is evolution necessary for range expansion? Manipulating
reproductive timing of a weedy annual transplanted beyond its range. *The American Naturalist*,
167(2), 153–164.
- Hallatschek, O., Hersen, P., Ramanathan, S., & Nelson, D. R. (2007). Genetic drift at expanding frontiers
560 promotes gene segregation. *Proceedings of the National Academy of Sciences of the United States of*
America, **104**(50), 19926–19930.
- Hallatschek, O., & Nelson, D. R. (2010). Life at the front of an expanding population. *Evolution*, **64**(1),
193–206.
- Hamilton, J. A., Okada, M., Korves, T., & Schmitt, J. (2015). The role of climate adaptation in
565 colonization success in *Arabidopsis thaliana*. *Molecular Ecology*, **24**(9), 2253–2263.
- Hohenlohe, P. A., Amish, S. J., Catchen, J. M., Allendorf, F. W., & Luikart, G. (2011). Next-generation
RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and
westslope cutthroat trout. *Molecular Ecology Resources*, **11**(s1), 117–122.
- Hothorn, T., Hornik, K., van de Wiel, M. A., & Zeileis, A. (2006). A Lego System for Conditional
570 Inference. *The American Statistician*, **60**(3), 257–263.
- Hulvey, K. B., & Zavaleta, E. S. (2012). Abundance declines of a native forb have nonlinear impacts on
grassland invasion resistance. *Ecology*, **93**(2), 378–388.

- Jensen, J. D., & Bachtrog, D. (2011). Characterizing the influence of effective population size on the rate of adaptation: Gillespie's Darwin domain. *Genome Biology and Evolution*, **3**, 687–701.
- 575 Kimura, M. (1964). Diffusion Models in Population Genetics. *Journal of Applied Probability*, **1**(2), 177–232.
- Kimura, M., & Crow, J. F. (1963). The Measurement of Effective Population Number. *Evolution*, **17**(3), 279–288.
- Kirkpatrick, M., & Barton, N. H. (1997). Evolution of a species' range. *The American Naturalist*, **150**(1), 1–23.
- 580 Kriticos, D. J., Webber, B. L., Leriche, A., Ota, N., Macadam, I., Bathols, J., & Scott, J. K. (2012). CliMond: global high-resolution historical and future scenario climate surfaces for bioclimatic modelling. *Methods in Ecology and Evolution*, **3**(1), 53–64.
- Langley, C. H., Tobari, Y. N., & Kojima, K. I. (1974). Linkage disequilibrium in natural populations of *Drosophila melanogaster*. *Genetics*, **78**(3), 921–936.
- 585 Le Corre, V., & Kremer, A. (1998). Cumulative effects of founding events during colonisation on genetic diversity and differentiation in an island and stepping-stone model. *Journal of Evolutionary Biology*, **11**(4), 495–512.
- Lehe, R., Hallatschek, O., & Peliti, L. (2012). The rate of beneficial mutations surfing on the wave of a range expansion. *PLoS Computational Biology*, **8**(3), e1002447.
- 590 Linnen, C. R., Kingsley, E. P., Jensen, J. D., & Hoekstra, H. E. (2009). On the origin and spread of an adaptive allele in deer mice. *Science*, **325**(5944), 1095–1098.
- Li, X.-M., She, D.-Y., Zhang, D.-Y., & Liao, W.-J. (2015). Life history trait differentiation and local adaptation in invasive populations of *Ambrosia artemisiifolia* in China. *Oecologia*, **177**(3), 669–677.
- 595 Luikart, G., Allendorf, F. W., Cornuet, J. M., & Sherwin, W. B. (1998). Distortion of allele frequency distributions provides a test for recent population bottlenecks. *The Journal of Heredity*, **89**(3), 238–247.
- Lynch, M., Conery, J., & Burger, R. (1995). Mutation Accumulation and the Extinction of Small Populations. *The American Naturalist*, **146**(4), 489–518.
- 600 Marculis, N. G., Lui, R., & Lewis, M. A. (2017). Neutral Genetic Patterns for Expanding Populations with Nonoverlapping Generations. *Bulletin of Mathematical Biology*, **79**(4), 828–852.
- Mayr, E. (1963). Animal species and evolution (Vol. 797). Belknap Press of Harvard University Press Cambridge, Massachusetts.
- Menke, J. W. (1989). Management Controls on Productivity. In L. F. Huenneke & H. A. Mooney (Eds.), Grassland structure and function: California annual grassland (pp. 173–199). Dordrecht: Springer Netherlands.
- 605 Micheletti, S. J., & Storfer, A. (2015). A test of the central–marginal hypothesis using population genetics and ecological niche modelling in an endemic salamander (*Ambystoma barbouri*). *Molecular Ecology*, **24**(5), 967–979.
- 610 Moreau, C., Bhéer, C., Vézina, H., Jomphe, M., Labuda, D., & Excoffier, L. (2011). Deep human genealogies reveal a selective advantage to be on an expanding wave front. *Science*, **334**(6059), 1148–1150.
- Narum, S. R., Buerkle, C. A., Davey, J. W., Miller, M. R., & Hohenlohe, P. A. (2013). Genotyping-by-sequencing in ecological and conservation genomics. *Molecular Ecology*, **22**(11), 2841–2847.
- 615 Nei, M., Maruyama, T., & Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution*, **29**(1), 1–10.
- Nei, M., & Li, W. H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the United States of America*, **76**(10), 5269–5273.
- 620 Nordborg, M., Borevitz, J. O., Bergelson, J., Berry, C. C., Chory, J., Hagenblad, J., ... Weigel, D. (2002). The extent of linkage disequilibrium in *Arabidopsis thaliana*. *Nature Genetics*, **30**(2), 190–193.
- Ohta, T. (1992). The Nearly Neutral Theory of Molecular Evolution. *Annual Review of Ecology and Systematics*, **23**, 263–286.

- Palstra, F. P., & Ruzzante, D. E. (2008). Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? *Molecular Ecology*, **17**(15), 3428–3447.
- Peischl, S., Dupanloup, I., Kirkpatrick, M., & Excoffier, L. (2013). On the accumulation of deleterious mutations during range expansions. *Molecular Ecology*, **22**(24), 5972–5982.
- Peischl, S., Kirkpatrick, M., & Excoffier, L. (2015). Expansion load and the evolutionary dynamics of a species range. *The American Naturalist*, **185**(4), E81–93.
- Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PloS One*, **7**(5), e37135.
- Petitpierre, B., Kueffer, C., Broennimann, O., Randin, C., Daehler, C., & Guisan, A. (2012). Climatic niche shifts are rare among terrestrial plant invaders. *Science*, **335**(6074), 1344–1348.
- Phillips, B. L. (2012). Range shift promotes the formation of stable range edges. *Journal of Biogeography*, **39**(1), 153–161.
- Pierce, A. A., Zalucki, M. P., Bangura, M., Udawatta, M., Kronforst, M. R., Altizer, S., ... de Roode, J. C. (2014). Serial founder effects and genetic differentiation during worldwide range expansion of monarch butterflies. *Proceedings of the Royal Society of London B: Biological Sciences*, **281**(1797). <http://doi:10.1098/rspb.2014.2230>
- Pironon, S., Villellas, J., Morris, W. F., Doak, D. F., & García, M. B. (2015). Do geographic, climatic or historical ranges differentiate the performance of central versus peripheral populations? *Global Ecology and Biogeography*, **24**(6), 611–620.
- Pitcairn, M., Schoenig, S., Yacoub, R., & Gendron, J. (2006). Yellow starthistle continues its spread in California. *California Agriculture*, **60**(2), 83–90.
- Ramakrishnan, A. P., Musial, T., & Cruzan, M. B. (2010). Shifting dispersal modes at an expanding species' range margin. *Molecular Ecology*, **19**(6), 1134–1146.
- Reever Morghan, K. J. R., & Rice, K. J. (2006). Variation in resource availability changes the impact of invasive thistles on native bunchgrasses. *Ecological Applications*, **16**(2), 528–539.
- Remington, D. L., Thornsberry, J. M., Matsuoka, Y., Wilson, L. M., Whitt, S. R., Doebley, J., ... Buckler, E. S., 4th. (2001). Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proceedings of the National Academy of Sciences of the United States of America*, **98**(20), 11479–11484.
- Rice, K. J., & Mack, R. N. (1991). Ecological genetics of *Bromus tectorum*. *Oecologia*, **88**(1), 91–101.
- Rius, M., & Darling, J. A. (2014). How important is intraspecific genetic admixture to the success of colonising populations? *Trends in Ecology & Evolution*, **29**(4), 233–242.
- Robertson, A. (1960). A Theory of Limits in Artificial Selection. *Proceedings of the Royal Society of London B: Biological Sciences*, **153**(951), 234–249.
- Sagarin, R. D., & Gaines, S. D. (2002). The “abundant centre” distribution: to what extent is it a biogeographical rule? *Ecology Letters*, **5**(1), 137–147.
- Sexton, J. P., McIntyre, P. J., Angert, A. L., & Rice, K. J. (2009). Evolution and Ecology of Species Range Limits. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 415–436.
- Shrimpton, J. M., & Heath, D. D. (2003). Census vs. effective population size in chinook salmon: large- and small-scale environmental perturbation effects. *Molecular Ecology*, **12**(10), 2571–2583.
- Slatkin, M., & Excoffier, L. (2012). Serial founder effects during range expansion: a spatial analog of genetic drift. *Genetics*, **191**(1), 171–181.
- Slotte, T., Foxe, J. P., Hazzouri, K. M., & Wright, S. I. (2010). Genome-wide evidence for efficient positive and purifying selection in *Capsella grandiflora*, a plant species with a large effective population size. *Molecular Biology and Evolution*, **27**(8), 1813–1821.
- Strasburg, J. L., Kane, N. C., Raduski, A. R., Bonin, A., Micheltore, R., & Rieseberg, L. H. (2011). Effective population size is positively correlated with levels of adaptive divergence among annual sunflowers. *Molecular Biology and Evolution*, **28**(5), 1569–1580.

- Stromberg, M. R., & Griffin, J. R. (1996). Long-term patterns in coastal California grasslands in relation to cultivation, gophers, and grazing. *Ecological Applications*, **6**(4), 1189–1211.
- Swope, S. M., Satterthwaite, W. H., & Parker, I. M. (2017). Spatiotemporal variation in the strength of density dependence: implications for biocontrol of *Centaurea solstitialis*. *Biological Invasions*, **19**(9), 2675–2691.
- Sung, W., Ackerman, M. S., Miller, S. F., Doak, T. G., & Lynch, M. (2012). Drift-barrier hypothesis and mutation-rate evolution. *Proceedings of the National Academy of Sciences of the United States of America*, **109**(45), 18488–18492.
- Szűcs, M., Melbourne, B. A., Tuff, T., Weiss-Lehman, C., & Hufbauer, R. A. (2017a). Genetic and demographic founder effects have long-term fitness consequences for colonising populations. *Ecology Letters*, **20**(4), 436–444.
- Szűcs, M., Vahsen, M. L., Melbourne, B. A., Hoover, C., Weiss-Lehman, C., & Hufbauer, R. A. (2017b). Rapid adaptive evolution in novel environments acts as an architect of population range expansion. *Proceedings of the National Academy of Sciences of the United States of America*, **114**(51), 13501–13506.
- Tingley, R., Vallinoto, M., Sequeira, F., & Kearney, M. R. (2014). Realized niche shift during a global biological invasion. *Proceedings of the National Academy of Sciences of the United States of America*, **111**(28), 10233–10238.
- Turner, T. F., Wares, J. P., & Gold, J. R. (2002). Genetic effective size is three orders of magnitude smaller than adult census size in an abundant, Estuarine-dependent marine fish (*Sciaenops ocellatus*). *Genetics*, **162**(3), 1329–1339.
- Uller, T., & Leimu, R. (2011). Founder events predict changes in genetic diversity during human-mediated range expansions. *Global Change Biology*, **17**(11), 3478–3485.
- Uygur, S., Smith, L., Uygur, F. N., Cristofaro, M., & Balciunas, J. (2004). Population densities of yellow starthistle (*Centaurea solstitialis*) in Turkey. *Weed Science*, **52**(5), 746–753.
- Varvio, S. L., Chakraborty, R., & Nei, M. (1986). Genetic variation in subdivided populations and conservation genetics. *Heredity*, **57**, 189–198.
- Vandepitte, K., de Meyer, T., Helsen, K., van Acker, K., Roldán-Ruiz, I., Mergeay, J., & Honnay, O. (2014). Rapid genetic adaptation precedes the spread of an exotic plant species. *Molecular Ecology*, **23**(9), 2157–2164.
- Wang, J. (2016). A comparison of single-sample estimators of effective population sizes from genetic marker data. *Molecular Ecology*, **25**(19), 4692–4711.
- Wang, X.-Q., Huang, Y., & Long, C.-L. (2013). Assessing the genetic consequences of flower-harvesting in *Rhododendron decorum* Franchet (Ericaceae) using microsatellite markers. *Biochemical Systematics and Ecology*, **50**, 296–303.
- Waples, R. S. (2006). A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics*, **7**(2), 167–184.
- Waples, R. S. (2016). Making sense of genetic estimates of effective population size. *Molecular Ecology*, **25**(19), 4689–4691.
- Waples, R. S., & Do, C. (2008). Ldne: a program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources*, **8**(4), 753–756.
- Webb, D. M., & Knapp, S. J. (1990). DNA extraction from a previously recalcitrant plant genus. *Plant Molecular Biology Reporter*, **8**(3), 180–185.
- Welles, S. R., & Dlugosch, K. M. (2018). Population Genomics of Colonization and Invasion (Vol. 17, p. 24). Cham: Springer International Publishing.
- White, T. A., Perkins, S. E., Heckel, G., & Searle, J. B. (2013). Adaptive evolution during an ongoing range expansion: the invasive bank vole (*Myodes glareolus*) in Ireland. *Molecular Ecology*, **22**(11), 2971–2985.
- Widmer, T. L., Guermache, F., Dolgovskaia, M. Y., & Reznik, S. Y. (2007). Enhanced Growth and Seed Properties in Introduced Vs. Native Populations of Yellow Starthistle (*Centaurea Solstitialis*). *Weed Science*, **55**(5), 465–473.

- 725 Willi, Y., Fracassetti, M., Zoller, S., & Van Buskirk, J. (2018). Accumulation of mutational load at the
edges of a species range. *Molecular Biology and Evolution*. doi:10.1093/molbev/msy003
Wootton, J. T., & Pfister, C. A. (2015). Processes affecting extinction risk in the laboratory and in nature.
Proceedings of the National Academy of Sciences of the United States of America, **112**(44), E5903.
Wright, S. (1931). Evolution in Mendelian Populations. *Genetics*, **16**(2), 97–159.
730 Wright, S. (1938). Size of population and breeding structure in relation to evolution. *Science*, **87**(2263),
430–431.

Acknowledgements We thank # reviewers for helpful comments on previous versions of this
735 manuscript, and MS Barker for computational assistance. This study was supported by a USDA
ELI Fellowship #2017-67011-26034 to JEB, the National Institute of General Medical Sciences
of the NIH under Award #K12GM000708 through the University of Arizona Center for Insect
Science to BSB, and USDA grant #2015-67013-23000 to KMD. The content is solely the
responsibility of the authors and does not necessarily represent the official views of the National
740 Institutes of Health. All authors declare no conflict of interests.

Data Accessibility All data files to be made accessible online (e.g. via Dryad)

Author Contributions JB and KMD conceived the study; JB and BSB performed the analyses;
745 JB, BSB, and KMD wrote the paper.

750

755

760

Tables and Figures

Table 1. Individual effects for the best fit linear model explaining rarefied effective population size (N_e) in invading populations of *C. solstitialis*, as a function of climatic principal component variables (PC1, PC2) and population age (Age).

Effect	Coefficient	Standard Error	t-value	p-value	Effect Size
PC1	1.45	0.483	-0.66	0.528	0.052
PC2	-0.39	0.438	2.50	0.011	0.579
Age	0.03	0.012	3.31	0.037	0.439

Fig 1. The distribution of rarefied N_e , climatic principal component (PC) gradients, and population age across *C. solstitialis* populations in Eurasia and California. In all panels, circles indicate sampled populations with a diameter proportional to N_e . PC1 is positively correlated with annual temperature and temperature of the driest quarter and negatively correlated with seasonal differences in total radiation in the native (a) and invaded (c) ranges. PC2 is positively correlated with seasonal differences in temperature and negatively correlated with annual precipitation and seasonal differences in precipitation in the native (b) and invaded (d) ranges. In the California invasion, population age (e) reflects a history of expansion beginning in the San Francisco Bay area and expanding first to the North and then to the South and East of the state.

Fig 2. Rarified relationships for the number of subsampled individuals and median values of LD- N_e . Rarefaction was performed by linear mixed model with random slopes and intercepts. N- Native populations; I- Invading populations.

Fig 3. Rarefied N_e values are predicted by climatic variables and population age in invading *C. solstitialis* populations. The second principal component (PC2) of climatic variability, for which larger values represent lower annual precipitation and greater seasonality in temperature, is significantly correlated with rarefied N_e (A; $P=0.011$) as is estimated population age (B; $P=0.037$). Lines show linear model fits and shading indicates the 95% confidence interval. Points represent partial residuals after taking into account other variables in the linear model.

SUPPORTING INFORMATION

825 Supporting information is included as a single PDF document