Biol Invasions
DOI 10.1007/s10530-013-0514-z

ORIGINAL PAPER

# Genetic evidence for founder effects in the introduced range of houndstongue (*Cynoglossum officinale*)

**Jennifer L. Williams · Lila Fishman**

Received: 12 October 2012 / Accepted: 29 June 2013


**Abstract** Phenotypic differentiation can occur between the native and introduced ranges of a species as a result of novel selective pressures, or by neutral processes and historical events. Our aim was to determine how underlying patterns of genetic diversity and potential population origin might have contributed to phenotypic differentiation between the native and introduced ranges of an herbaceous weed. We combined data from microsatellite markers from 16 native and 16 introduced populations of *Cynoglossum officinale*, a noxious weed of the western US, with previously published phenotypic data from common gardens to investigate genetic diversity in both ranges and relate population structure to phenotypic differentiation. Several lines of evidence suggest loss of genetic diversity during the introduction of *C. officinale*. Despite reduced diversity, introduced plants out-performed natives in a common garden in one environment. We found little evidence that population-level variation in diversity contributed to phenotypic variation (e.g. through inbreeding depression). Our results suggest that establishment, spread, and potentially adaptation of a species to a new range is not prevented by reductions in genetic diversity of the magnitude we observed. Further, we suggest that non-random filtering or biased introduction at the point of emigration may contribute to phenotypic divergence between ranges.

**Keywords** Common garden · *Cynoglossum officinale* · Founder effects · Genetic diversity · Houndstongue · Invasive plant · Native and introduced ranges · Weed

## Introduction

In invasive plants, genotypes from the introduced range are often phenotypically divergent from home range populations. Observed shifts include changes in phenology, vigor, allocation to defense, or the ability to respond favorably to multiple growing conditions (Bossdorf et al. 2005; Novak and Mack 2005; Richards et al. 2006). Phenotypic divergence of introduced populations is often attributed to the action of novel biotic and abiotic selective pressures that may operate along the pathway of introduction, establishment and spread during the invasion process (e.g. Blair and Wolfe 2004; Bossdorf et al. 2005). Alternatively, similar patterns in phenotypic divergence may be

**Electronic supplementary material** The online version of this article (doi:10.1007/s10530-013-0514-z) contains supplementary material, which is available to authorized users.

J. L. Williams (✉) · L. Fishman
Division of Biological Sciences, University of Montana, Missoula, MT 59812, USA
e-mail: jennifer.williams@geog.ubc.ca

J. L. Williams
Department of Geography, University of British Columbia, Vancouver, BC V6T 1Z2, Canada

Published online: 17 July 2013

driven by neutral processes or historical events (Keller and Taylor 2008; Novak and Mack 2005), including genetic drift during bottlenecks (Barrett and Husband 1990; Nei et al. 1975) and non-random filtering of genotypes during the invasion process. For example, non-random filtering may shape phenotypic variation in the invaded range if some genotypes have a higher probability of becoming crop contaminants or horticultural introductions (Keller and Taylor 2008; Kolar and Lodge 2001). Biased introduction, as much as selection among genotypes present in the invaded range, may interact with drift to contribute to invasion-associated shifts in phenotype.

Distinguishing the relative roles of adaptive evolution and chance events in phenotypic divergence during any particular invasion is a challenge (Keller and Taylor 2008). To do so requires integrating analyses of genetic marker variation (e.g. microsatellites, AFLPs) with phenotypic data from across populations in the native and introduced ranges. Studies using genetic markers demonstrate that invasive taxa often have reduced genetic diversity in the introduced range, although this is not true in every case (Bossdorf et al. 2005; Dlugosch and Parker 2008; Novak and Mack 2005; Wares et al. 2005). Those studies that have combined population genetics with phenotypic data from common gardens have generally found that selection in the introduced range has likely altered some traits, but that the mechanisms are specific to each species (e.g. Chun et al. 2009, 2011; Dlugosch and Parker 2008; Lavergne and Molofsky 2007; Maron et al. 2004; Xu et al. 2010). Understanding how variation in genetic diversity has changed between ranges can also offer clues about particular introduction pathways (Gladieux et al. 2011; Hardesty et al. 2012).

Here, we use microsatellite markers to assess genetic diversity in the native and introduced ranges of houndstongue (*Cynoglossum officinale),* a noxious weed of the Intermountain Western US (Upadhyaya et al. 1988). Using common garden experiments in both ranges, we previously found that plants from native and introduced populations differ in relative performance rank at each site (Williams et al. 2008). Plants from introduced populations responded strongly to favorable growing conditions, achieving higher lifetime fitness in a European garden with high water availability (within the native range), whereas plants from native populations had higher fitness in a Montana garden (within the introduced range) where growing conditions were harsher. This non-intuitive crossing of fitness reaction norms suggests that factors beyond local adaptation may contribute to the overall divergence between native and introduced *C. officinale* populations. In addition, we have established that *C. officinale* has undergone a major life history shift between ranges, with introduced populations containing a high fraction of iteroparous individuals compared to the native range, where semelparity dominates and iteroparity is extremely rare (Williams 2009). We therefore wanted to know how the underlying patterns of genetic diversity and potential population origin might have contributed to phenotypic differentiation, and in particular, how they might have contributed to the observed fitness differences between range in the common gardens. We examined microsatellite markers from 16 native and 16 introduced populations to ask: (1) How is genetic variation distributed across the native and introduced ranges? (2) Is there evidence for loss of genetic diversity occurring with introduction? (3) Do changes in the patterns of genetic diversity relate to changes in phenotype?

## Methods

### Study system

*Cynoglossum officinale* L. (Boraginaceae), commonly known as houndstongue, is native to central Europe, where it occurs in open woodlands, meadows, sand dunes, and marginal habitats adjacent to tilled fields (de Jong et al. 1990). Its range extends from central Asia (where published documentation is sparse) west to the Atlantic Ocean, and latitudinally across Europe, although it does not grow in the extreme south or north of southern Scandinavia or Britain (de Jong et al. 1990). It is considered a rare plant in parts of its native range, including in central and northern Germany (Enßlin et al. 2011). It was introduced to North America in the mid-nineteenth century as a feed contaminant, where it now occurs across the United States, with the exception of the most southerly states, and southern Canada (Upadhyaya et al. 1988). It is a particular problem in the Intermountain West, where it has been placed on the noxious weed list in a number of states, in part due to its toxicity to livestock

(Upadhyaya et al. 1988). In the introduced range, *C. officinale* commonly occurs in riparian areas and grasslands, as well as in grazed and clear-cut areas.

*Cynoglossum officinale* is a self-compatible, facultatively monocarpic perennial (de Jong et al. 1990). Although flowers are self-compatible, seed set is increased by insect pollination, which is carried out mainly by bumble-bees (*Bombus* spp.) (de Jong et al. 1990). Pollination also can result in geitonogamous selfing, or pollen transfer within a plant, and geitonogamy can account for as much as 70 % of seeds (Vrieling et al. 1999). Each flower produces up to four large nutlets that are barbed and dispersed on mammal fur; seeds that are not dispersed typically germinate within 2 m (Boorman and Fuller 1984). Although most native plants die after flowering, a small percentage of individuals may flower a second time in the subsequent year, after which they die (de Jong et al. 1990). In contrast, flowering in subsequent years (iteroparity) is common in the introduced range, occurring, on average, in 20 % of the individuals within a population (Williams 2009).

## Sample collection and preparation

Seeds of *C. officinale* were collected from 16 field sites in each range in 2003–2005 (Fig. 1; Table 1) from 10 to 20 individuals at each site, with individuals separated by at least 1 m. Field sites occurred in broadly representative habitats in central Europe (the center of the native range) and in the region of the introduced range (the Intermountain West) where *C. officinale* is considered noxious, and were chosen haphazardly based on recommendations from local botanists and land managers. In 2007, seeds were placed into cold stratification for 6 weeks to break dormancy and then sown in potting soil in small pots in a greenhouse at the University of Montana (Missoula, MT). A subset of sampled populations (10 from each range) were used in previous common garden experiments that were conducted in Missoula, MT and Bad Lauchstädt, Germany from 2004 to 2005 (Table 1; for more details, see Williams et al. 2008).

We collected approximately 200 mg tissue from the first or second true leaf (2–3 cm) of one plant from each maternal family, and stored tissues in a −80 °C freezer. Due to variable seed germination, the number of maternal families sampled from each population ranged from 5 to 20 (Table 1). We extracted genomic DNA following a CTAB/chloroform extraction protocol that was modified for 96-well format and using a bead beater (Geno/Grinder 2000, Spex Certiprep, Meutchen, NJ) for tissue homogenization (Doyle and Doyle 1990; Fishman and Willis 2005).

## Microsatellite analysis

We analyzed six microsatellite loci that were previously developed for *C. officinale* and 5′-labeled with the fluorescent dyes noted in parentheses: C2-19 (NED), C3-41 (FAM), C2-43 (HEX), C2-62 (HEX), C2-72 (FAM), and C3-79 (FAM) (Korbecka and Wolff 2004) combined into one multiplex set. Polymerase chain reactions (PCR) were performed in a total volume of 10 μL with 2 μL of 1:20 diluted DNA solution, and 0.2 μL of 10 μM C2-19, C3-41 and C2-72 and 0.1 μL of 10 μM C2-43, C2-62, and C3-79 labeled forward primers and unlabelled reverse primers in the following reaction mix: 2.5 μL $H_20$, 2.0 μL 5 × GoTAQ Flexi buffer, 0.8 μL 25 mM $MgCl_2$, 0.8 μL 2.5 mM dNTPs, and 0.1 μL GoTAQ Flexi. The markers were amplified using a touchdown PCR program: 3 min at 94 °C, 10 touchdown amplification cycles (30 s at 94 °C, 30 s at 58 °C, 45 s at 72 °C) with the annealing temperature decremented by 1 °C each cycle, 30 amplification cycles (30 s at 94 °C, 30 s at 48 °C, 45 s at 72 °C), and a 10-min final extension (72 °C). PCR products were run on an ABI 3130xl automated capillary sequencer (Applied Biosystems, Foster City, CA) and marker genotypes were assigned automatically using GeneMapper software (Applied Biosystems) and then manually verified.

## Statistical analyses

Measures of genetic diversity, including observed ($H_o$) and expected heterozygosity ($H_e$) and allele number ($N_a$), were calculated using GenAlEx 6.1 (Peakall and Smouse 2006), and inbreeding coefficients ($F_{IS}$) were calculated using Fstat 2.9.3 (Goudet 2001). To account for different sample sizes among populations in the calculations of allelic richness, we calculated the rarefied allelic richness for each population using the hierfstat package in R (Goudet 2005). Range means for each diversity statistic were compared using two-sided *t* tests on mean values for each population in R 3.0.0 (R Core Development Team (2013).

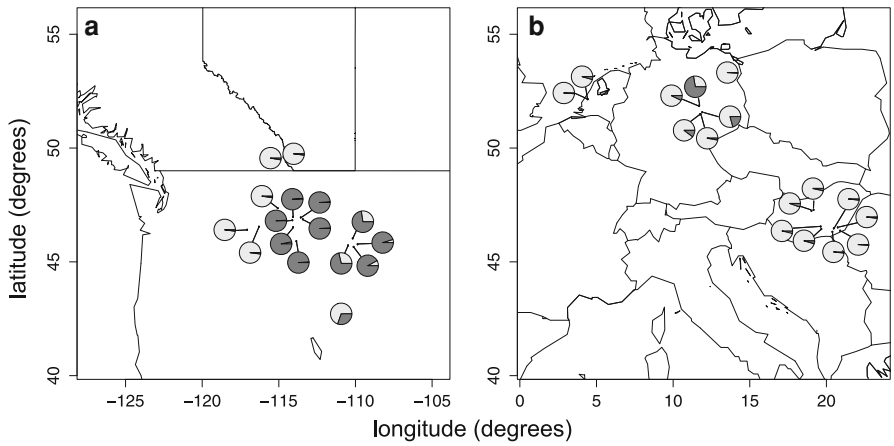


**Fig. 1** Locations of sampled *Cynoglossum officinale* populations in the introduced range (**a** North America) and native range (**b** Europe). Pies represent proportional assignment of populations to one of two groups based on the results of a Bayesian clustering analysis. The native range extends from northern Spain to southern Britain and Sweden and east through central Russia; *C. officinale* is present across southern Canada and the United States, excluding the most southern states, but is of particular concern as an invasive plant in the Intermountain West, where populations were sampled for this study

To examine the relationship between isolation and distance, pairwise $F_{ST}$ values were regressed over the geographic distance (km) between pairs of populations within each range; significance was tested using a Mantel test with 1,000 randomizations performed in GenAlEx 6.1 (Peakall and Smouse 2006). We further looked for spatial structuring of genetic variability within and between ranges and populations with an analysis of molecular variance using ARLEQUIN version 3.1 (Excoffier et al. 2005). Within ranges, we grouped populations into regions based on geographic proximity. The native range contained four groups: Netherlands (2 populations), northern Germany (1 population), central Germany (5 populations), and Hungary (8 populations). The introduced range contained 5 groups: Wyoming (1 population), central Montana (4 populations), western Montana (7 populations), southern Canada (2 populations), and Washington/Idaho (2 populations). We further examined the pattern of the relationship of populations using a Bayesian clustering approach in STRUCTURE (Pritchard et al. 2000). The STRUCTURE analysis used an admixture ancestry model with correlated allele frequencies. We estimated the posterior probability (L(K)) for values of K (number of clusters) ranging from 1 to 32 (n = 20 runs each), using a burn-in period and run length of the Markov Chain Monte Carlo (MCMC) of 10,000 and 50,000 iterations, respectively. Parameters (e.g., alpha) generally converged after <2,500 iterations, so these run times were adequate for estimation of K. To assign individual multi-locus genotypes to clusters and estimate population mean membership, we examined L(K) directly and calculated ΔK using the method of Evanno et al. (2005). We also used the program InStruct (Gao et al. 2007), which estimates K (using similar algorithms and identical settings as the STRUCTURE runs) while relaxing the assumption of Hardy–Weinberg equilibrium (HWE) within clusters. Because *C. officinale* can have high rates of geitonogamous selfing, deviations from HWE conditions within populations may be common, violating the assumptions of STRUCTURE and potentially leading to spurious evidence of admixture (Gao et al. 2007).

We used linear models to investigate the link between genotypic diversity and the phenotypic data collected in common garden experiments (Williams et al. 2008). Specifically, models examined whether range, expected heterozygosity ($H_e$), and the interaction were significant predictors of two phenotypic traits: plant volume and seed set (both log-transformed) in common gardens in Germany and Montana. We expected that these two fitness traits might be influenced by the population history of inbreeding. When interaction terms were significant, separate

**Table 1** Location of populations sampled for genetic analyses of *Cynoglossum officinale* in the native range (Europe) and introduced range (North America), number of samples (N), and population averages for: number of alleles ($N_a$), observed heterozygosity ($H_o$), expected heterozygosity ($H_e$), inbreeding coefficient ($F_{IS}$), number of polymorphic loci (maximum = 6), number of polymorphic loci that differ significantly from Hardy–Weinberg equilibrium ($P < 0.05$)

| Continent | State/Country | Collection site | Latitude | Longitude | N | $N_a$ | $H_o$ | $H_e$ | $F_{IS}$ | Polymorphic loci | #Loci different from H–W | Phenotype |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Europe | Hungary | Batida | 46.320 | 20.331 | 14 | 1.67 | 0.024 | 0.102 | 0.782 | 1 | 1 | |
| Europe | Hungary | Mezohegyes | 46.369 | 20.833 | 5 | 1.83 | 0.083 | 0.178 | 0.620 | 2 | 0 | |
| Europe | Hungary | Kiskun | 46.440 | 19.610 | 8 | 2.67 | 0.164 | 0.329 | 0.553 | 5 | 3 | |
| Europe | Hungary | Hodmezovarzehly | 46.470 | 20.417 | 11 | 2.17 | 0.324 | 0.355 | 0.134 | 4 | 1 | cg |
| Europe | Hungary | Korduskut | 46.513 | 20.678 | 17 | 2.83 | 0.100 | 0.347 | 0.728 | 6 | 3 | cg |
| Europe | Hungary | Tazlar | 46.552 | 19.563 | 5 | 2.00 | 0.167 | 0.310 | 0.545 | 5 | 2 | |
| Europe | Hungary | Tokol | 47.244 | 18.964 | 6 | 2.00 | 0.083 | 0.328 | 0.792 | 6 | 1 | |
| Europe | Hungary | Moby Dick (Budapest) | 47.283 | 19.101 | 10 | 3.33 | 0.098 | 0.541 | 0.840 | 6 | 6 | |
| Europe | Germany | Aseleben | 51.480 | 11.691 | 19 | 2.50 | 0.167 | 0.366 | 0.564 | 5 | 3 | cg |
| Europe | Germany | Salziger See | 51.485 | 11.736 | 10 | 2.17 | 0.191 | 0.307 | 0.424 | 4 | 2 | cg |
| Europe | Germany | Lettewitz | 51.577 | 11.842 | 10 | 2.17 | 0.067 | 0.280 | 0.783 | 4 | 2 | cg |
| Europe | Germany | Hohenerxleben 2 | 51.861 | 11.634 | 5 | 1.67 | 0.033 | 0.226 | 0.882 | 4 | 1 | cg |
| Europe | Germany | Hohenerxleben 1 | 51.864 | 11.643 | 20 | 2.50 | 0.088 | 0.299 | 0.719 | 4 | 4 | cg |
| Europe | Netherlands | Bierlap | 52.143 | 4.361 | 5 | 2.50 | 0.167 | 0.400 | 0.652 | 5 | 2 | cg |
| Europe | Netherlands | Meijendel Dunes | 52.162 | 4.350 | 7 | 2.83 | 0.480 | 0.552 | 0.209 | 6 | 1 | cg |
| Europe | Germany | Neustrelitz | 53.317 | 13.460 | 16 | 1.67 | 0.036 | 0.098 | 0.658 | 1 | 1 | cg |
| North America | Wyoming | Afton | 42.716 | −110.930 | 6 | 2.00 | 0.250 | 0.194 | 0.200 | 4 | 0 | cg |
| North America | Montana | Boulder River | 45.658 | −110.108 | 9 | 2.67 | 0.229 | 0.338 | 0.375 | 5 | 0 | cg |
| North America | Montana | Yellowstone River, Livingston | 45.719 | −110.468 | 20 | 2.67 | 0.231 | 0.455 | 0.513 | 5 | 4 | cg |
| North America | Montana | Upper Deer Creek, Big Timber | 45.772 | −109.847 | 14 | 2.33 | 0.262 | 0.365 | 0.321 | 5 | 1 | |
| North America | Montana | Jennings Camp, Sula | 45.907 | −113.841 | 7 | 1.83 | 0.159 | 0.248 | 0.433 | 3 | 0 | |
| North America | Montana | Crazy Mountains | 46.030 | −110.168 | 17 | 2.50 | 0.316 | 0.457 | 0.337 | 5 | 1 | |
| North America | Washington | Clarkston | 46.410 | −117.050 | 12 | 1.67 | 0.050 | 0.090 | 0.486 | 1 | 1 | cg |
| North America | Montana | Lee Metcalf WR, Florence | 46.520 | −114.054 | 12 | 2.33 | 0.217 | 0.309 | 0.341 | 4 | 1 | |
| North America | Idaho | Dworshak Reservoir | 46.544 | −116.288 | 11 | 1.33 | 0.019 | 0.073 | 0.771 | 1 | 1 | cg |
| North America | Montana | Blue Mountain, Missoula | 46.828 | −114.101 | 19 | 1.50 | 0.039 | 0.101 | 0.631 | 1 | 1 | |
| North America | Montana | Lubrecht | 46.868 | −113.421 | 7 | 1.33 | 0.028 | 0.090 | 0.737 | 1 | 1 | |
| North America | Montana | Ninemile Prairie | 46.950 | −113.558 | 17 | 2.33 | 0.088 | 0.215 | 0.609 | 5 | 2 | cg |
| North America | Montana | Lavalle Creek | 46.977 | −114.067 | 17 | 1.83 | 0.020 | 0.116 | 0.841 | 2 | 1 | cg |

**Table 1** continued

| Continent | State/Country | Collection site | Latitude | Longitude | N | $N_a$ | $H_o$ | $H_e$ | $F_{IS}$ | Polymorphic loci | #Loci different from H–W | Phenotype |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| North America | Montana | Tamarack Creek | 47.359 | −115.060 | 14 | 2.00 | 0.038 | 0.126 | 0.715 | 1 | 1 | cg |
| North America | British Columbia | Fenwick Road | 49.557 | −115.545 | 6 | 1.50 | 0.067 | 0.097 | 0.407 | 1 | 0 | cg |
| North America | Alberta | Pincher Creek | 49.748 | −114.034 | 10 | 1.50 | 0.017 | 0.113 | 0.866 | 1 | 1 | cg |

Phenotypes measured for a subset of populations in common gardens (cg)

models were fit to populations from each range. In these analyses, each population was represented by the mean for that trait. We ran identical models with average allele number ($N_a$) as the predictor variable. We were unable to use the average inbreeding coefficient ($F_{IS}$) as a more direct measure of inbreeding due to the number of markers that were fixed for a single allele, particularly in the introduced range (see Table 1), leading to an observed heterozygosity of zero.

## Results

At each of the six microsatellite markers for *C. officinale*, we found fewer alleles in introduced populations compared to native populations (Table 2). At four out of the six loci, two-thirds or more of the alleles were restricted to populations in the native range (Table 2), and we found only one allele (at marker 79) that was restricted to the region we sampled in the introduced range. On average, rarefied allelic richness was lower in introduced populations compared to native populations, a proportional loss of 13 % (Fig. 2a; $t_{29.97} = 2.05$, $P = 0.049$). Expected heterozygosity ($H_e$) was also lower (32 %) in introduced populations (Fig. 2b; $t_{29.90} = 2.21$, $P = 0.035$), but the average inbreeding coefficient ($F_{IS}$) did not differ between ranges (Fig. 2c; $t_{28.66} = 1.25$, $P = 0.22$). However, the latter comparison is biased by the fact that many introduced populations, and presumably those with the greatest history of inbreeding, lacked the polymorphism necessary to calculate $F_{IS}$ for more than one locus. Thus for nearly half of the introduced populations, the value represents $F_{IS}$ for only the most variable marker (79). Most populations contained at least one locus that was fixed for a particular allele (Table 1), and within populations, the proportion of polymorphic loci was significantly lower in the introduced range compared to in the native range (Fig. 2d; $Z = 3.034$, $P = 0.002$).

Continent of origin (North America vs. Europe) explained ~20 % of the overall variation in genetic diversity (AMOVA results in Table 3). Within the native range, more than half (51.1 %) of the variation in genetic diversity could be attributed to variation within populations, whereas differences among regions explained only 20.6 % (Table 3). In contrast, in the introduced range, the variation in genetic diversity was distributed nearly equally among

regions, among populations within regions, and within populations (Table 3). While there was no relationship between geographic and genetic distance in the introduced range ($R^2 = 0.0002$, $P = 0.46$), genetic distance increased significantly but weakly with geographic distance in the native range ($R^2 = 0.061$, $P = 0.033$, increase in $F_{ST}$ of 0.009 units for every 100 km).

Consistent with the AMOVA results, Bayesian clustering analyses indicated only moderate population genetic structure associated with geographical location. At K = 16, where mean values of L(K) asymptoted in STRUCTURE (Pritchard et al. 2000), a few populations were composed of individuals assigned to a single cluster, but many individuals (particularly in European populations) were assigned to multiple clusters, and there was little correspondence between cluster assignment and geography (data not shown). Similarly, the InSTRUCT analysis (which is similar to STRUCTURE, but relaxes assumptions of Hardy–Weinberg equilibrium within clusters), indicated that K = 13 was best by the Deviance Information Criterion (DIC); however, at K = 13, most individuals were again assigned partially to 2–4 clusters each (data not shown), indicating that these assignments may not be robust. In contrast, the ΔK method of Evanno et al. (Evanno et al. 2005) indicated that K = 2 had the highest support in the STRUCTURE analysis. At K = 2, individual assignments (Fig. S1) and the proportional membership of each population (Fig. 1) reveal an interesting pattern potentially reflecting the history of the invasion. At this level, the majority of North American populations cluster with a single European population (EHX, near Staßfurt, central Germany), indicating that populations in the invaded range represent a subset of the genetic and phenotypic diversity present in the native range. However, populations from southern Canada, Idaho and Washington belong to the dominant European cluster, potentially indicating multiple introductions.

To determine whether genetic diversity within populations was associated with population mean performance, as might be the case if bottlenecks and inbreeding had also increased the frequency of deleterious alleles, we examined the relationship between expected heterozygosity ($H_e$) and two fitness traits in two different garden environments. In the German garden (within the native range), all plants were larger and more fecund compared to in the Montana garden (within the introduced range) (Williams et al. 2008), indicating growing conditions were relatively favorable. In the German garden, plants from introduced populations were larger and more fecund on average than those from native populations, but neither trait was associated with $H_e$ (Fig. 3a; $H_e$ (introduced) on plant size: $F_{1,8} = 0.049$, $P = 0.83$; seeds: $H_e$: $F_{1,16} = 1.02$, $P = 0.33$). In contrast, plants that came from native populations with higher $H_e$ were larger (Fig. 3a; $H_e \times$ range: $F_{1,16} = 5.31$, $P = 0.035$; $H_e$ (native): $F_{1,8} = 6.62$, $P = 0.033$), but the effect did not persist through seed production ($H_e$: $F_{1,16} = 1.02$, $P = 0.33$). In the less favorable environment in the garden in Montana, $H_e$ was not related to plant size or seed production for either native or introduced populations (Fig. 3b; size: $F_{1,16} = 0.0052$, $P = 0.94$; seed production: $F_{1,16} = 1.52$, $P = 0.23$). We found qualitatively similar relationships between genotype and phenotype when average allelic diversity ($N_a$) was used as the predictor variable.

## Discussion

*Cynoglossum officinale* has lost substantial genetic diversity during its introduction to North America, as indicated by declines in allelic diversity and expected heterozygosity between the native and introduced ranges. However, reductions in genetic diversity have not prevented *C. officinale* from spreading across the introduced range, where it is now considered a noxious weed in several states, nor from shifting its life history from almost exclusive semelparity in the native range to a high frequency of iteroparity in some introduced populations. Here we discuss the factors contributing to the observed genotypic patterns and their potential influences on phenotypic variation.

The life history and habit of *C. officinale* shape the patterns of genetic variation in microsatellite markers that we observed. In both ranges, new populations of *C. officinale* are often established by a few seeds that may be dispersed by mammals into recently disturbed areas or by recolonization of old populations from the short-lived seed bank following local extinction events (van der Meijden et al. 1992). The frequent colonization and extinction of populations on the landscape results in new populations that have a high probability of containing a subset of the regional

genetic diversity, and leads to populations that are nearby in space but genetically quite different. This differentiation is strongest in the native range, where more than half of the total genetic variation can be explained by differences among populations (Table 3). In contrast, we observed much weaker patterns in differentiation across space in the introduced range, both in the amount of genetic variation observed among populations and in the absence of a relationship between geographic distance and genetic differentiation. We also found that within introduced populations, at least one locus, and frequently several, were fixed for a single allele. Unless a new population was colonized by seeds from several different populations, the loss of alleles would persist following introduction. The differences in genetic structure between continents reflect both bottlenecks associated with introduction to North America (Genton et al. 2005; Gladieux et al. 2011), which reduced the total pool of variation, and rapid human-mediated spread to disturbed/cultivated habitats within the invaded range. A shallower spatial genetic structure in the introduced range compared to the native range has also been demonstrated for other invasive plants (Durka et al. 2005; Genton et al. 2005).

**Table 2** Total number of alleles for each microsatellite marker in the native and introduced ranges of *Cynoglossum officinale*, and percent of alleles that were unique to the native range

| Marker | Total native range | Total introduced range | % restricted to native range |
|---|---|---|---|
| 19 | 6 | 2 | 66.6 |
| 41 | 5 | 1 | 80.0 |
| 43 | 5 | 2 | 60.0 |
| 62 | 5 | 4 | 20.0 |
| 72 | 6 | 2 | 66.7 |
| 79 | 10 | 8 | 18.2 |
| All | 37 | 19 | 47.4 |

The exact geographic origin of *C. officinale* populations in North America is unknown. The results of the

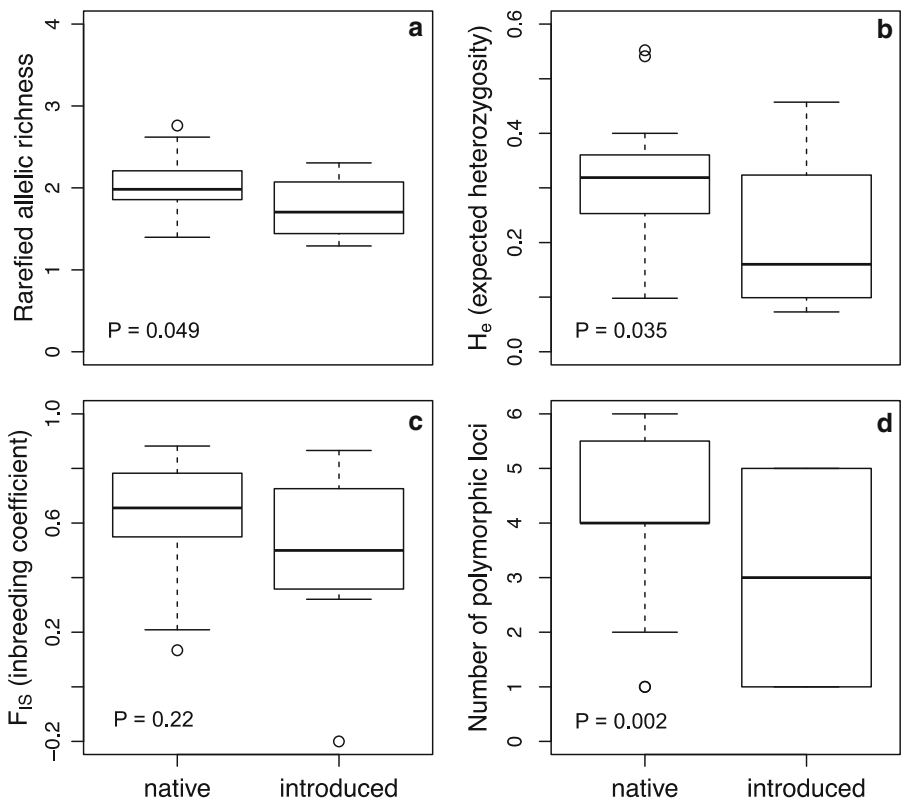


**Fig. 2** *Box plots* comparing measures of genetic diversity between native and introduced populations: **a** rarefied allelic richness, **b** expected heterozygosity ($H_e$), **c** inbreeding coefficient ($F_{IS}$), and **d** number of polymorphic loci (out of 6)

**Table 3** Analysis of molecular variance (AMOVA) results showing the distribution of variation in genetic diversity of microsatellite markers among continents, regions, and populations

| Source of variation | Df | Sum of squares | Variance components | Percentage of variation |
|---|---|---|---|---|
| Native and introduced ranges | | | | |
| Among ranges | 1 | 87.1 | 0.202*** | 19.68 |
| Among populations within ranges | 31 | 352.1 | 0.504*** | 49.14 |
| Within populations | 699 | 223.6 | 0.320*** | 31.18 |
| Total | 731 | 662.8 | 1.026 | |
| Native range | | | | |
| Among regions | 3 | 83.0 | 0.261*** | 20.64 |
| Among populations within regions | 13 | 94.1 | 0.357*** | 28.28 |
| Within populations | 319 | 205.8 | 0.645** | 51.07 |
| Total | 335 | 382.9 | 1.263 | |
| Introduced range | | | | |
| Among regions | 4 | 91.4 | 0.252*** | 31.69 |
| Among populations within regions | 11 | 81.1 | 0.277*** | 34.82 |
| Within populations | 380 | 101.2 | 0.266** | 33.48 |
| Total | 395 | 273.75 | 0.795 | |

*** $P < 0.001$; ** $P < 0.01$

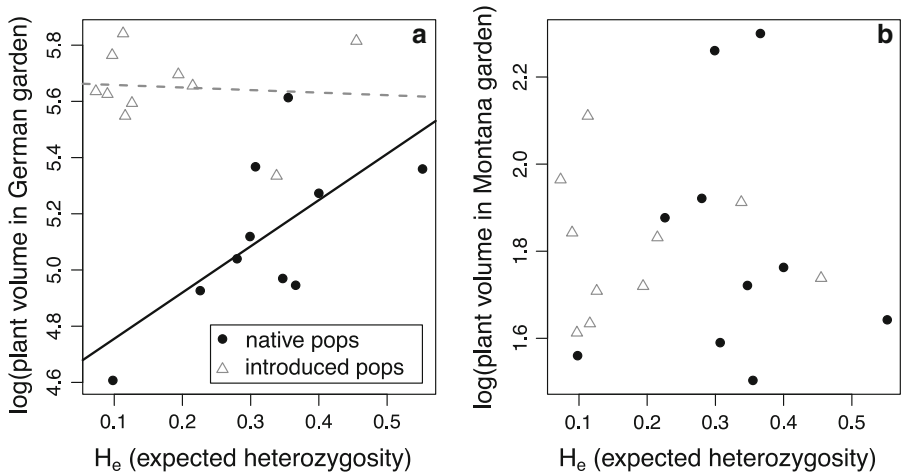


**Fig. 3** Relationships between average expected heterozygosity ($H_e$) and average measure of plant size for each population for plants grown in a common garden in (**a**) Germany (within the native range) and (**b**) in Montana (within the introduced range). When linear models were statistically significant, best *fit lines* are shown (significance of all linear models is presented in "Results")

STRUCTURE analysis do suggest that invasive western North American populations likely came from at least two sources. Historical records indicate that seeds were introduced as a contaminant of feed being shipped from Europe to North America (de Jong et al. 1990; Upadhyaya et al. 1988). Since feed entering North America was not required to be cleaned until the early 1800s and this rule was rarely enforced even after that, many weeds that arrived as contaminants arrived multiple times (Gaskin et al. 2005; Mack 2003; Novak and Mack 2005). This history, along with the patterns of genetic diversity in *C. officinale* that we observed, suggest that more than one introduction occurred. The large proportion of introduced populations that were

assigned to the same cluster (at K = 2) as one of the German populations we sampled could indicate that these populations originated in Germany (Fig. 1). Alternatively, the true populations of origin were not included in our sample of the native range, or populations in western North America were secondarily introduced from eastern North America.

The pathway of migration as a feed contaminant makes it likely that introduced genotypes came from crop field margins in the native range and thus may not represent genotypes from the full range of native habitats. Our genotypic and phenotypic data are consistent with this scenario. In particular, in the STRUCTURE analysis, many populations in the introduced range cluster with the single European population (EHX) sampled from crop-adjacent habitat (Fig. 1). With data from only six polymorphic loci and a restricted sampling of populations, we do not want to over-interpret this finding. However, introduction of genotypes primarily from a subset of native populations that perform well in high nutrient areas such as crop margins could contribute to the non-intuitive patterns of phenotypic variation observed for houndstongue. Specifically, plants from the introduced range were able to achieve much higher fitness than plants from native populations when grown in lush growing conditions in a common garden in Germany, despite having lower genetic diversity, consistent with a history of bottlenecks and/or inbreeding (Fig. 3a; Williams et al. 2008). In contrast, plants from native populations were smaller and their size was positively correlated with population heterozygosity in the same garden (Fig. 3a). In the Montana garden, where growing conditions were harsher, plants from native populations, which were sampled from a variety of habitats, were able to outperform plants from introduced populations. More targeted sampling would be necessary to test for a consistent differences among native populations from different habitats, but biased probability of introduction, rather than selection within the introduced range, may (in part) account for patterns of phenotypic variation in this invasive species. This result emphasizes that the genotypes which are best suited to a new range are not necessarily the ones which are introduced (Novak and Mack 2005) and yet they may still do well for a variety of other reasons.

Our analyses, as well as a recent study of native German and botanical garden populations (Enßlin et al. 2011) indicate that *C. officinale* harbors substantial among-population variability in life history and fitness traits despite relatively low levels of genetic diversity range-wide. Although average genetic diversity is further reduced in North American populations (Fig. 2), due to bottlenecks during introduction and drift in the introduced range, this may or may not inhibit its success in North America. On the one hand, *C. officinale* is a successful invader (De Clerck-Floate and Wikeem 2009; Upadhyaya et al. 1988) and introduced genotypes harbored sufficient variation to undergo a shift to iteroparity despite the rarity of this trait the native range. On the other hand, it is clear that some native genotypes not present in the sampled introduced range actually outperform plants from invasive populations in drier North American habitats. Coupled with the observation that only the most diverse native populations perform as well as plants from the introduced range in a lush (native range) common garden, this suggests that the overall loss of genetic diversity maybe much less important than historical events (such as biased introduction from agricultural habitats) and selection within the introduced range in shaping the ecology and evolution of invasive plant populations.

**Acknowledgments** We thank Jennifer Andreas, Harald Auge, Tom de Jong, and Stefan Toepfer for collecting seeds, and Samantha Campbell for assisting in the greenhouse. We are grateful to Camille Barr for invaluable advice in the lab and to Mike Schwartz for statistical advice. Jake Alexander, Darren Johnson and Joanna Kelley provided helpful comments on the manuscript. Support for this study was provided by NSF grant DBI-0321329 to LF and to JW by an NSF Graduate Research Fellowship, NSF Doctoral Dissertation Improvement Grant DEB 05-08102 and from the National Center for Ecological Analysis and Synthesis, a Center Funded by NSF (Grant #EF-0553768), the University of California, Santa Barbara, and the State of California. Seeds were imported into the US under USDA-APHIS permit 37-86531.

## References

Barrett SCH, Husband BC (1990) The genetics of plant migration and colonization. In: Brown AHD, Clegg MT, Kahler AL, Weir BS (eds) Plant population genetics, breeding, and genetic resources. Sinauer Associates, Sunderland, MA, pp 254–277

Blair AC, Wolfe LM (2004) The evolution of an invasive plant: an experimental study with *Silene latifolia*. Ecology 85:3035–3042

Boorman LA, Fuller RM (1984) The comparative ecology of two sand dune biennials: *Lactuca virosa* L. and *Cynoglossum officinale* L. New Phytol 69:609–629

Bossdorf O, Auge H, Lafuma L et al (2005) Phenotypic and genetic differentiation between native and introduced plant populations. Oecologia 144:1–11

Chun YJ, Nason JD, Moloney KA (2009) Comparison of quantitative and molecular genetic variation of native versus invasive populations of purple loosestrife (*Lythrum salicaria* L., Lythraceae). Mol Ecol 18:3020–3035

Chun YJ, Le Corre V, Bretagnolle F (2011) Adaptive divergence for a fitness-related trait among invasive *Ambrosia artemisiifolia* in France. Mol Ecol 20:1378–1388

De Clerck-Floate R, Wikeem B (2009) Influence of release size on establishment and impact of a root weevil for the biocontrol of houndstongue (*Cynoglossum officinale*). Biocontrol Sci Tech 19:169–183

de Jong TJ, Klinkhamer PGL, Boorman LA (1990) Biological flora of the British Isles: *Cynoglossum officinale* L. J Ecol 78:1123–1144

Dlugosch KM, Parker IM (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. Mol Ecol 17:431–449

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Durka W, Bossdorf O, Prati D et al (2005) Molecular evidence for multiple introductions of garlic mustard (*Alliaria petiolata*, Brassicaceae) to North America. Mol Ecol 14:1697–1706

Enßlin A, Sandner TM, Matthies D (2011) Consequences of ex situ cultivation of plants: genetic diversity, fitness and adaptation of the monocarpic *Cynoglossum officinale* L. in botanic gardens. Biol Conserv 144:272–278

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50

Fishman L, Willis JH (2005) A novel meiotic drive locus almost completely distorts segregation in *Mimulus* (monkeyflower) hybrids. Genetics 169:347–353

Gao H, Williamson S, Bustamante CD (2007) A Markov chain Monte Carlo approach for joint inference of population structure and inbreeding rates from multilocus genotype data. Genetics 176:1635–1651

Gaskin JF, Zhang D-Y, Bon M-C (2005) Invasion of *Lepidium draba* (Brassicaceae) in the western United States: distributions and origins of chloroplast DNA haplotypes. Mol Ecol 14:2331–2341

Genton BJ, Shykoff J, Giraud T (2005) High genetic diversity in French invasive populations of common ragweed, *Ambrosia artemisiifolia*, as a result of multiple sources of introduction. Mol Ecol 14:4275–4285

Gladieux P, Giraud T, Kiss L et al (2011) Distinct invasion sources of common ragweed (*Ambrosia artemisiifolia*) in Eastern and Western Europe. Biol Invasions 13:933–944

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://www.unil.ch/izea/softwares/fstat.html

Goudet J (2005) Hierfstat, a package for R to compute and test hierarchical F-statistics. Mol Ecol Notes 5:184–186

Hardesty BD, Le Roux JJ, Rocha OJ et al (2012) Getting here from there: testing the genetic paradigm underpinning introduction histories and invasion success. Divers Distrib 18:147–157

Keller SR, Taylor DR (2008) History, chance and adaptation during biological invasion: separating stochastic phenotypic evolution from response to selection. Ecol Lett 11:852–866

Kolar CS, Lodge DM (2001) Progress in invasion biology: predicting invaders. Trends Ecol Evol 16:199–204

Korbecka G, Wolff K (2004) Characterization of nine microsatellite loci in *Cynoglossum officinale* (Boraginaceae). Mol Ecol Notes 4:229–230

Lavergne S, Molofsky J (2007) Increased genetic variation and evolutionary potential drive the success of an invasive grass. Proc Nat Acad Sci 104:3883–3888

Mack RN (2003) Global plant dispersal, naturalization and invasion: pathways, modes and circumstances. In: Ruiz GM, Carlton JT (eds) Invasive species: vectors and management strategies. Island Press, Washington, DC, pp 3–30

Maron JL, Vilà M, Bommarco R et al (2004) Rapid evolution of an invasive plant. Ecol Monogr 74:261–280

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

Novak SJ, Mack RN (2005) Genetic bottlenecks in alien plant species: influence of mating systems and introduction dynamics. In: Sax DF, Stachowicz JJ, Gaines SD (eds) Species invasions: insights into ecology, evolution, and biogeography. Sinauer, Sunderland, MA, pp 201–228

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295

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

R Core Development Team (2013) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Richards CL, Bossdorf O, Muth NZ et al (2006) Jack of all trades, master of some? On the role of phenotypic plasticity in plant invasions. Ecol Lett 9:981–993

Upadhyaya MK, Tilsner HR, Pitt MD (1988) The biology of Canadian weeds. 87. *Cynoglossum officinale* L. Can J Plant Sci 68:763–774

van der Meijden E, Klinkhamer PGL, de Jong TJ et al (1992) Meta-population dynamics of biennial plants: how to exploit temporary habitats. Acta Botanica Neerlandica 41:249–270

Vrieling K, Saumitou-Laprade P, Cuguen J et al (1999) Direct and indirect estimates of the selfing rate in small and large individuals of the bumblebee pollinated *Cynoglossum officinale* L (Boraginaceae). Ecol Lett 2:331–337

Wares JP, Hughes AR, Grosberg RK (2005) Mechanisms that drive evolutionary change: insights from species introductions and invasions. In: Sax DF, Stachowicz JJ, Gaines SD (eds) Species invasions: insights into ecology, evolution, and biogeography. Sinauer, Sunderland, MA, pp 229–257

Williams JL (2009) Flowering life-history strategies differ between the native and introduced ranges of a monocarpic perennial. Am Nat 174:660–672

Williams JL, Auge H, Maron JL (2008) Different gardens, different results: native and introduced populations exhibit contrasting phenotypes across common gardens. Oecologia 157:239–248

Xu C-Y, Julien MH, Fatemi M et al (2010) Phenotypic divergence during the invasion of *Phyla canescens* in Australia and France: evidence for selection-driven evolution. Ecol Lett 13:32–44