Invasion of *Lepidium draba* (Brassicaceae) in the western United States: distributions and origins of chloroplast DNA haplotypes

JOHN F. GASKIN,* DAO-YUAN ZHANG† and MARIE-CLAUDE BON‡

*Northern Plains Agricultural Research Laboratory, USDA-ARS, P.O. Box 463, Sidney, MT 59270 USA, †South Beijing Road 40-3, Xinjiang Research Institute of Ecology and Geography, Chinese Academy of Science, 830011 Urmqii, Xinjiang, China, ‡European Biological Control Laboratory, USDA-ARS, Campus International de Baillarguet, 34980 Montferrier le Lez, France

Abstract

Advances in phylogeography are of great value for understanding the population structure and origins of invasive genotypes. Such insights provide constructive information for current or future biological control research efforts. In this study, we investigated a highly variable chloroplast DNA (cpDNA) marker for populations of the weed *Lepidium draba* (Brassicaceae) in its native Eurasian and invasive US ranges. We sequenced DNA from 684 individuals from Eurasia and the US and found 41 different haplotypes. Our comparative study between the native and invasive ranges showed a 33% reduction in allelic richness (*A*) and a 7% reduction in haplotype diversity (*h*) since introduction into the US. Most genetic variation in the native range was observed within geographical regions and populations, not between regions, and this result was similar for the invasive range. Assignment tests indicated the most likely origins of many invasive haplotypes. Some of these occurred in western Europe, supporting an expanded native range that had been proposed for the species. Exact locations were identified for a diverse set of invasive haplotypes which can be used in ongoing host-specificity tests of potential biological control agents.

Keywords: hoary cress, invasion, *Lepidium draba*, origins, population structure, whitetop

Received 24 December 2004; revision accepted 1 April 2005

Introduction

*Lepidium draba* L. [= *Cardaria draba* (L.) Desv.] is a perennial weed native to Eurasia. It was first collected in North America on the east coast in 1862 (Mulligan & Frankton 1962) and is now found in most US states (except the southeastern region) (USDA 2004), most Canadian provinces (Corns & Frankton 1952), and also in regions of Mexico (Rios & Garcia 1998). *Lepidium draba* invades pastures and crops such as alfalfa and grains, rangelands and riparian areas (Lyons 1998), and is considered noxious in most central and western US states (USDA 2004) and some western Canadian provinces (Rice 2004).

*Lepidium draba* encompasses two subspecies: *Lepidium draba* ssp. *draba* (commonly named whitetop) and *Lepidium draba* ssp. *chalepense* (L.) Thellung [= *Cardaria chalepensis* (L.) Hand.-Mazz.] (commonly named lenspod whitetop). Both subspecies are native to the Balkan Peninsula, Turkey, Georgia, Armenia, Azerbaijan, Israel, Syria, Iraq, Iran, Turkmenistan, Uzbekistan, and from Kazakhstan to the Irkutsk region of Russia (Mulligan & Frankton 1962). Others suggest that *L. draba* ssp. *draba* is additionally native to southern Europe (e.g. Ball 1964). *Lepidium draba* is most closely related to another species listed as noxious in some US states, *Lepidium appelianum* Al-Shehbaz [= *Cardaria pubescens* (C.A. Mey.) Jarm.] (Mummenhoff *et al.* 2001) (commonly named hairy whitetop), which is from central Asia and was first collected in North America in 1919 (Mulligan & Frankton 1962). Differences in morphology (such as pubescence and the shape of the fruit pods), chromosome number (Mulligan & Frankton 1962) and assumed ploidy level (*L. draba* ssp. *draba* 2n = 8x = 64; *L. draba* ssp. *chalepense* 2n = 10x = 80; *Lepidium appelianum* 2n = 2x = 16) delineate the taxa. There is evidence of hybridization between the two subspecies of *L. draba* (Mulligan & Frankton 1962; Rollins 1993) and neither *L. draba* nor
L. appelianum are able to self-pollinate (Mulligan & Frankton 1962). At times, these species have been separated from the genus Lepidium, due mostly to their indehiscent and usually inflated fruit pods, but molecular phylogenetic analysis has recently helped clarify the taxonomy, placing them back within Lepidium (Mummenhoff et al. 2001; Al-Shehbaz et al. 2002). Both species can form dense stands and spread by rhizomes or seeds (Mulligan & Findlay 1974), although Larson et al. (2000) showed that increase of plant density in study populations was mostly due to clonal propagation.

The process of invasion includes one to many founding events perhaps with subsequent bottlenecks, such that an invasion should result in a reduction in the genetic diversity when compared to the total diversity found across the native range (Tsutsui et al. 2000; Sakai et al. 2001), although individual populations in the invaded region can exhibit more diversity than individual populations in the native range (e.g. Novak & Mack 1993; Kolbe et al. 2004). Life history, reproduction and the number and genetic content of introductions can also impact both the genetic diversity and structure of an invasion, and thus may affect biological control efficacy (Burdon & Marshall 1981; Nissen et al. 1995). Intraspecific genetic variation in plant defense to biological control agents (both pathogens and arthropods) can exist, and in cases has been detected with molecular markers (e.g. Jarosz & Burdon 1990; Bruckart et al. 2004; Goolsby et al. 2004). A mixture of both resistant and non-resistant genotypes within an invasion may hinder biological control efforts, as plants resistant to the biological control agent may become predominant [as has happened in the Chondrilla/Puccinia system in Australia (Burdon et al. 1981, 1984)].

The utility of highly polymorphic cpDNA as a phylogeographical marker has been established in studies ranging from analyses of glacial cycles (e.g. King & Ferris 1998; Tremblay & Schoen 1999) to recent invasions (e.g. Saltonstall 2002; Trewick et al. 2004). The typical maternal inheritance of cpDNA in angiosperms, and subsequent movement of the marker in seeds only (not pollen), help to retain any genetic structure that may have originated with introductions, when compared to nuclear markers that are subject to gene flow via both pollen and seeds.

Lepidium draba patches are difficult to eradicate with herbicides (Stougaard et al. 1999), grazing, or manual control (Scurfeld 1962). All three taxa are targets of a relatively new biological control feasibility study that involves many universities, government agencies, and other groups from North America and Europe (USDA 2002). Knowledge of genotype origins, genetic structure across an invasion, and native centres of diversity (Roderick & Navajas 2003; Briese 2004), along with climatic and other factors, are important clues for determining locations of potential biological control agents. The distribution and frequency of invasive plant genotypes, especially if there is preliminary evidence of differential preference of targets by an agent, should be investigated. Hence, a selection of plant genotypes that represents the diversity found in the invasion should be available to biological control researchers for host-specificity testing. In order to do this, and to better understand the origins and invasion process of Lepidium draba, we have analysed a variable cpDNA sequence region for over 700 plant specimens from across the US and from Spain to China. Specifically, we are interested in knowing: (i) the amount of native diversity that exists in the US invasion; (ii) which invasive haplotypes are the most common; (iii) how the haplotypes are distributed across the invasion and within populations; and (iv) the origins of US invasive haplotypes.

Materials and methods

Sampling

Young leaf tissue was collected and silica dried from 734 plants in Asia \((n = 186)\), Europe \((n = 188)\) and the US \((n = 360)\) (Fig. 1). Herbarium and/or DNA voucher specimen and population details are listed in Table S1 (supplementary material). We made an effort to collect plants that were at least far enough apart that they should be genetically separate individuals, although the maximum size of a clonal patch is not known. Collections focused on Lepidium draba \((n = 684)\) plants, but also included the closely related species Lepidium appelianum \((n = 29)\) and Lepidium propinquum Fisch. & C.A. Mey. \(\approx \) Cardaria propinqua (Fisch. & C.A. Mey.) N. Busch \((n = 21)\) when these were growing in or near patches of L. draba. Many L. draba collections were made when the plants were not producing flowers or fruits (e.g. rosette stage), thus preventing their accurate determination to subspecies.

Between 1 and 11 plants were collected per population from 139 locations worldwide, and the greatest distance between plants in a population varied from 5 m to 1.6 km (Table S1, supplementary material). We followed the suggestion of Pons & Petit (1995) for surveying genetic variation by collecting fewer plants per population, but collecting more populations. The US was divided into 12 state regions, and alternatively, eight regions delineated by major mountain ranges and riverways. The Eurasian samples were divided into four regions going from west to east, using political boundaries of countries (Fig. 1).

DNA extraction, PCR, and sequencing

Genomic DNA was extracted from approximately 20 mg of dried material using a modified cetyltrimethyl ammonium bromide (CTAB) method (Hillis et al. 1996). The chloroplast intergenic region between the trn S (GCU)
and trn G (UCC) genes was amplified using the primer pair trn S (GCU) and trn G (UCC) of Hamilton (1999). The adjoining region between trn G (UCC) and trn R was amplified using the primer pair trn G (UCC) rev/comp (UCC) \(5'\)-GTGGTAAAAGTGTGATTCGTTC\(3'\) [which is a reverse complement of Hamilton’s (1999) trn G] and a trn R primer from Doyle et al. (1992) \(5'\)-GTCCTATCCATTAGACAATGG\(3'\) (see Fig. 2). The thermal cycling program was as follows for these two fragments: one cycle of 95 °C (2 min); 30 cycles of 95 °C

© 2005 Blackwell Publishing Ltd, Molecular Ecology, 14, 2331–2341

**Fig. 1** Map of populations (dots) collected from Eurasia and the US. Dotted lines indicate regions used in hierarchical analysis of molecular variance.
(1 min), 55 °C (50 °C for the second primer pair) (1 min), 72 °C (2 min); and then 32 °C (5 min). A 30 µL reaction containing 3 µL of genomic extract, 3 µL 1X NH₄ poly-
merase chain reaction (PCR) buffer (Bioline), 1.5 µL 50 mM MgCl₂, 0.24 µL 100 mM dNTP mix, 3 µL 2 µM of each primer, 0.15 µL (0.75 unit) of Biolase DNA polymerase (Bioline), and H₂O was performed for each region for each individual. PCR products were purified using QIAquick PCR Purification Kit (QIAGEN) prior to sequencing in a Beckman CEQ 2000XL automated sequencer using standard protocols including the LFR-1 method of injection time and voltage. Both strands of the PCR products were sequenced using the primer pair trn S (GCU) and trn G (UCC) for the first product, and primer pair trn G (UCC) rev/comp and an internal primer interior3reverse (5′-TGGTCGAGTTCTCCTGATCT-3′) for the second product (see Fig. 2) (the region between the trn R Doyle and interior3reverse primers was not used due to lack of variation found in initial sequences).

Data analysis

Sequences generated in this study have been deposited in the GenBank database, and accession numbers are listed in Table S2 (supplementary material). Sequences were aligned manually using the software se-al (Rambaut 1996). The alignment is available upon request from the first author. Insertion/deletion events (indels), no matter what size, were treated as one mutational event (one evolutionary step). In the case of nucleotide substitutions and increase or decrease of mono- and dinucleotide repeat number, we assumed that any state was equally likely to evolve to any other existing state. All base substitutions and indels were used in the analysis. The complete set of haplotype sequences from L. draba, L. appelianum, and L. propinquum was used to create a haplotype network. This was done with the network building software TCS (Clement et al. 2000) which uses statistical parsimony (95%) and the genealogical reconstruction algorithms of Templeton et al. (1992). Portions of the haplotype network that could not be connected by TCS were enclosed in loops on Fig. 3. These loops were connected manually with the observed, nonhomo-
plausible, mutational differences (this is not intended to connect the network, but is used to demonstrate minimum distances between unconnected portions). Targeted DNA of L. draba plants with sequences that were tip haplotype singletons (haplotypes that appeared only once and were external, not internal, on the network) was re-amplified and re-sequenced in order to distinguish true haplotypes from PCR error. Neutrality tests based on Tajima’s D (Tajima 1989) and Fu and Li’s D* (Fu & Li 1993) statistics were used to determine if the sequenced regions are evolving in a neutral manner and therefore appropriate for use in a phylogeographical study. Allelic richness (A) was calculated using fstat (Goudet 1995), standardized for sample size using the rarefaction method (El Mousadik & Petit 1996; equation 1), and was based on the smallest group size per analysis. ARLEQUIN version 2.000 software (Schneider et al. 2000) was used to measure haplotype diversity (h) (Nei 1987; equation 8.5) and its variance (not shown), and is defined as the probability that two randomly chosen haplotypes are different in the sample. A Z test was used to determine if differences in h values were significant. Analysis of molecular variance (AMOVA) (Excoffier et al. 1992), as implemented by ARLEQUIN, was used to examine the distribution of genetic variation among and within regions and populations. Measures of regional subdivision are expressed as percentage of variation, and utilize the frequency of haplotypes as well as the number of mutations between them. GENECLASS version 2.0 (Piry et al. 2004) was used to generate scores for assignment to regions, using Bayesian methods (Rannala & Mountain 1997) and the simulation algorithm of Cornuet et al. (1999) with 100 000 simulated individuals. That analysis was trivial in cases where the haplotype to be assigned was only found in one Eurasian region, but in cases where a haplotype was found in mul-
tiple regions, we were able to compare assignment scores.

Results

Variation in the cpDNA sequences and measures of diversity

The region sequenced between trn S (GCU) and trn G (UCC) varies from 503 to 580 base pairs (bp) in length (aligned length 586 bp), and the region sequenced between trn G (UCC) and trn R varies from 386 to 403 bp (aligned length 410 bp), for a total aligned length of 996 bp. There are 53 single base substitution sites (four of these have three or four possible character states), eight indel events ranging from 6 to 29 bp (all of these are found at different loci and none of the loci overlap), seven indels derived from two dinucleotide repeat regions, and seven indels derived from four mononucleotide repeat regions. In all,
counting indels as a single event, there are 74 variable sites, and 16 (22%) of these are homoplasious (mutations found on more than one place on the haplotype network). Of the 16 homoplasious mutations, 12 are from single base substitutions, none are from 6 to 29 bp indels, two are from dinucleotide repeats, and two are from mononucleotide repeats. Because levels of homoplasy were relatively similar in ≤ 2 bp repeats compared to single base substitutions (28% vs. 23% respectively), we retained all indels in the analyses.

When considering *Lepidium draba* only, we identified 30 single base substitution sites (seven of these are homoplasious), one indel of 9 bp (not homoplasious), five indels derived from two dinucleotide repeat regions (one of these indels is homoplasious), and two indels derived from two single-nucleotide repeat regions (one of these indels is homoplasious). In all, there are 38 (3.8%) variable sites within *L. draba*.

Neither Tajima’s *D* nor Fu and Li’s *D* * statistics rejected the null hypothesis of neutral evolution (*D* = −0.7137, *D* * = −1.1077; *P* > 0.1). Based on 734 individuals sampled, 55 haplotypes were distinguished. Six of these belong to *Lepidium propinquum*, eight to *Lepidium appelianum*, and the remaining 41 to *L. draba* (Fig. 3). Of the eight *L. appelianum* haplotypes, three were found in Asia and five in the USA.
and none were shared between Asia and the US. We did not detect the *L. propinuum* haplotypes in the US, nor has that species ever been listed as present there (Mulligan & Frankton 1962; USDA 2004). *Lepidium draba* haplotypes that were singleton tip haplotypes reproduced the same sequences when analysis was repeated.

The haplotype network could not be connected between the three species using statistical parsimony (95%) (Fig. 3). The *L. draba* and *L. appelianum* networks were separated by 14 mutations, including two indels (6 bp and 29 bp), and the *L. draba* and *L. propinuum* networks were separated by 11 mutations, including two indels (14 bp and 20 bp).

Since many collections were lacking flower and fruit, and only one herbarium voucher per population was collected, distinguishing between *L. draba ssp. draba* and *L. draba ssp. chalepense* was not a focus of this research. Two populations (4188, 4309) of restricted size (100–200 m diameter) in the US contained both *L. draba* and *L. appelianum* haplotypes. Three populations (4487, 4489, 4490) (200–300 m diameter) in Asia were mixes of *L. draba* and *L. propinuum* haplotypes. Populations that were continuous (with no obvious barrier to clonal growth such as a paved road or a canal) and relatively small (5–300 m between furthest plants) were either polymorphic (populations 4036, 4192, 4297, 4441, 4442) or monomorphic (4024, 4440, 4482, 4484) for this molecular marker.

Twenty haplotypes were found in the 341 US *L. draba* plants and 31 were found in the 343 Eurasian plants [A (allelic richness) = 20.0 and 30.1, respectively] (Table 1). Measurements of A in the four Eurasian regions varied from 9.7 to 15, and within the eight US regions A varied from 1.0 to 4.9. Haplotype diversity for *L. draba* in the US (*h* = 0.82) was lower (*P* = 0.0005) than in Eurasia (*h* = 0.88) and haplotype diversity in Europe (*h* = 0.82) was lower (*P* = 0.0011) than in Asia (*h* = 0.90) (Table 1). Within the US, haplotype diversity (*h*) varied from 0 to 0.85.

**Population structure and assignment**

Analysis of molecular variance of *L. draba* in the US showed that only 12.2% (*P* = 0.002) of the total genetic variation was between the 12 states while 37.0% (*P* < 0.001) was among populations within states (Table 2). The remaining variation was within populations (50.8%, *P* < 0.001). When considering the US as eight regions defined by natural, rather than political, boundaries (Fig. 1), only 7.2% (*P* = 0.020) of the variation was between the eight regions, with 42.0% (*P* < 0.001) of the variation among populations within regions, and 50.8% (*P* < 0.001) of the variation within populations.

Likewise, there was little variation between the four geographical regions we defined in Eurasia (Fig. 1) (7.7%, *P* = 0.003), with most variation occurring among populations within regions (50.2%, *P* < 0.001) and within populations (42.1%, *P* < 0.001).

We were unable to assign to Eurasian regions the 10 *L. draba* haplotypes that were found in the US but not in Eurasia (7, 12, 14, 19, 20, 27, 28, 43, 46, 51) and these accounted for 20% (69/341) of the US *L. draba* plants. Ten haplotypes of *L. draba* were found within both the US and in Eurasia (1, 2, 4, 5, 9, 10, 11, 18, 21, 52), representing 80% (272/341) of the US *L. draba* plants, and their assignment scores for the four regions of Eurasia are shown in Fig. 4.

**Discussion**

**Genetic variation of invasive and native populations**

The amount of haplotypes or alleles existing in an invasion is a function of many factors, including the number of introductions, the size of each introduction, the population structuring in the native range, and subsequent drift and selection pressures that occur after introduction. Previous studies have shown variation in the amount of native diversity that can exist in an invasion, ranging from relatively high (e.g. Novak & Mack 1993; Meekins et al. 2001) to relatively low (e.g. Gaskin & Schaal 2002; Saltonstall 2002). Higher allelic richness would be expected of introductions from seed contaminants involving high numbers of seeds, compared to deliberate horticultural introductions of few individuals. The US *L. draba* invasion, when compared to that in Eurasia, retained 66% of its allelic richness (*A*) (Table 1), which is very high when compared to a deliberate horticultural introduction (*saltecedar — Tamarix ramosissima* Ledeb./*Tamarix chinensis* Lour., most likely imported as cuttings) which retained only 16% of its allelic richness (*A* = 49.0 in the native Asian range compared to 7.8 in the US (Gaskin & Schaal 2002)). Likewise, the high retention of *L. draba* haplotype diversity (*h*) in the US compared to Eurasia (93%) suggests multiple, or few but diverse, introductions. Historical seed contamination evidence supports a scenario of both multiple and diverse introductions into the US. The increase in alfalfa production in the early 1900s led to a demand for alfalfa seed importation [6 million pounds in 1913 (Brown 1914) from areas such as Europe and central Asia]. Low-grade clover and imported alfalfa seed lots surveyed in 1906 often contained *Lepidium* species (peppergrass) (Brown & Crosby 1906). A 1928 survey found that 90% of alfalfa seed batches from Spain (n = 25), and 40% from Turkestan (n = 100), contained *L. draba* (Hillman & Henry 1928).

Our results indicate that across a small patch (e.g. pop 4440: 10 samples within a 5-m diameter) there can be multiple haplotypes (1, 21, and 9), while widespread populations (e.g. pop 4290: 10 samples within a 100 m diameter) with discontinuous patches of *L. draba* may be monotypic for this marker (all plants sharing haplotype 1). Even though a single *L. draba* plant can produce over 400 shoots per year in the absence of competition (Kirk et al. 2000), each shoot is generally monomorphic.

**References**

Our results indicate that across a small patch (e.g. pop 4440: 10 samples within a 5-m diameter) there can be multiple haplotypes (1, 21, and 9), while widespread populations (e.g. pop 4290: 10 samples within a 100 m diameter) with discontinuous patches of *L. draba* may be monotypic for this marker (all plants sharing haplotype 1). Even though a single *L. draba* plant can produce over 400 shoots per year in the absence of competition (Kirk et al. 2000), each shoot is generally monomorphic. **References**

© 2005 Blackwell Publishing Ltd, Molecular Ecology, 14, 2331–2341
Table 1 Lepidium draba haplotype distributions, frequencies and measures of diversity

| Region          | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 14 | 16 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 36 | 39 | 40 | 41 | 42 | 43 | 46 | 49 | 51 | 52 | 57 | All | $A^*$ | $h$ |
|-----------------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| USA             | 122 | 7 | 25 | 11 | 17 | 13 | 65 | 10 | 15 | 9  | 14 | 5  | 2  | 3  | 7  | 3  | 8  | 1  | 2  | 2  | 341 | 20.0 | 0.82 |
| Eurasia         | 79  | 16 | 6  | 23 | 4  | 6  | 8  | 1  | 71 | 6  | 22 | 1  | 15 | 2  | 1  | 16 | 3  | 9  | 1  | 1  | 17 | 2  | 5  | 1  | 1  | 1  | 4  | 4  | 1  | 7  | 9  | 343 | 30.1 | 0.88 |
| Total           | 201 | 23 | 6  | 48 | 15 | 6  | 17 | 8  | 14 | 136| 16 | 15 | 9  | 22 | 15 | 5  | 2  | 18 | 2  | 1  | 16 | 3  | 9  | 7  | 3  | 1  | 1  | 17 | 2  | 5  | 1  | 1  | 1  | 4  | 4  | 8  | 1  | 1  | 2  | 9  | 9  | 684|
| Europe          | 40  | 15 | 1  | 6  | 8  | 1  | 65 | 6  | 10 | 2  | 1  | 1  | 5  | 1  | 1  | 14 | 1  | 5  | 1  | 1  | 14 | 1  | 5  | 3  | 1  | 1  | 1  | 1  | 1  | 4  | 4  | 1  | 2  | 9  | 155 | 23.0 | 0.90 |
| Asia            | 39  | 16 | 6  | 8  | 3  | 6  | 12 | 1  | 15 | 1  | 15 | 3  | 4  | 3  | 1  | 1  | 1  | 1  | 4  | 4  | 1  | 1  | 1  | 1  | 4  | 4  | 1  | 1  | 1  | 3  | 8  | 66  | 10.5 | 0.85 |
| Western Europe  | 16  | 9  | 8  | 1  | 32 | 5  | 2  | 1  | 4  | 3  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    | 81  | 9.7  | 0.78 |
| Eastern Europe  | 24  | 6  | 1  | 6  | 33 | 10 | 1  | 1  | 1  | 1  | 1  | 11 | 1  | 5  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    | 107 | 12.5 | 0.83 |
| Western Asia    | 16  | 6  | 8  | 3  | 6  | 7  | 7  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    | 69  | 15.0 | 0.90 |
| Central Asia    | 23  | 16 | 5  | 1  | 8  | 15 | 3  | 4  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    | 86  | 10.5 | 0.85 |
| California      | 11  | 1  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    | 15  | 2.8  | 0.47 |
| Great Basin     | 25  |    | 1  | 1  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    | 31  | 2.5  | 0.35 |
| Intermountain   | 11  | 5  | 6  | 10 | 1  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    | 46  | 4.9  | 0.85 |
| Northern Plains | 25  | 4  | 5  | 1  | 15 | 1  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    | 69  | 4.8  | 0.80 |
| Northwest       | 20  | 1  | 12 |    | 1  | 7  | 2  | 7  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    | 51  | 4.1  | 0.77 |
| Snake River     | 15  | 1  | 6  | 9  | 1  | 16 | 4  | 3  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    | 63  | 4.9  | 0.84 |
| Southern Plains | 15  | 1  | 3  | 19 | 3  | 3  | 6  | 6  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    | 58  | 4.6  | 0.81 |
| St. Louis       | 8   |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    | 8   | 1.0  | 0   |

$A$, allelic richness.

$h$, haplotype diversity.

*Based on a minimum population size of 341 individuals.
†Based on a minimum population size of 155 individuals.
‡Based on a minimum population size of 69 individuals.
§Based on a minimum population size of 8 individuals.
we find that individual populations of *L. draba* are not always genetically uniform [this result was also seen for the predominantly clonal invasive Brassicaceae species *Rorippa austriaca* (Crantz) Besser (Bleeker 2003)], nor are small, continuous patches always clonal. In cases, the two subspecies of *L. draba*, and even *L. appelianum*, were intermixed in populations. Lower than expected occurrence of purely clonal populations indicates that collecting more samples per population is advised in any future studies of *L. draba*. The recent development of microsatellite markers for these species (Bon *et al.* 2005) will be extremely useful for investigating the contributions of clonality, seed production, and hybridization within patches.

The low level of regional structuring for this *L. draba* chloroplast marker in Eurasia suggests movement of seeds subsequent to the evolution of many of the haplotypes found in the species. The seeds and fruits of *L. draba* have no obvious mechanism or vector for dispersal. The inflated indehiscent siliques may float, or potentially roll for a short distance in the wind, but movement of seed across long distances through natural means seems unlikely, and must be primarily through accidental movement by humans, as would be expected of a seed contaminant or a weed found in crops such as alfalfa that are baled and transported for feed (Scurfield 1962). The low level of structuring across the native range regions may have had an effect on *L. draba* in the US, allowing a diverse invasion to arise from potentially few introduction events, whereas it would take more introductions if the genotypes were spatially structured in their native range. Our analysis of molecular variance also determined that there was very little genetic structure across either natural regions or states within the US, which is evidence of multiple introductions, few introductions of high diversity, or few introductions of low diversity and subsequent movement of seeds between states.

Even though the AMOVA suggests that there is little structure at the 12 US state level, some of the 10 most common haplotypes show up in certain states only (Table 1). The eight plants from the one population in Missouri (population 4481) were all haplotype 43. This haplotype was not found anywhere else in the world, and differs by six muta-

<table>
<thead>
<tr>
<th>Regional structure (no. of regions)</th>
<th>Between regions or states</th>
<th>Among populations/within regions or states</th>
<th>Within populations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>P%</td>
<td>%</td>
</tr>
<tr>
<td>US states (12)</td>
<td>12.2</td>
<td>0.002</td>
<td>37.0</td>
</tr>
<tr>
<td>US regions (8)</td>
<td>7.2</td>
<td>0.020</td>
<td>42.0</td>
</tr>
<tr>
<td>Eurasian regions (4)</td>
<td>7.7</td>
<td>0.003</td>
<td>50.2</td>
</tr>
</tbody>
</table>

%, percentage molecular variation explained by hierarchical level; P, level of significance for the distribution of variation for that hierarchical level being different from random.

Even though the AMOVA suggests that there is little structure at the 12 US state level, some of the 10 most common haplotypes show up in certain states only (Table 1). The eight plants from the one population in Missouri (population 4481) were all haplotype 43. This haplotype was not found anywhere else in the world, and differs by six muta-

![Fig. 4](image-url) Assignment scores for 10 haplotypes shared between the US and four regions of Eurasia. The regional divisions can be seen in Fig. 1. Assignment scores of < 0.1 are not shown. Haplotypes shown with assignment scores for more than one region are connected with a line to help compare them visually.

from the most closely related haplotype (4), even though these plants were morphologically determined to be *L. draba*. This divergent haplotype lets us hypothesize that collections from central and eastern US, where
L. draba does exist, but not in the numbers found in the western US, may uncover additional haplotypes. However, we found L. draba difficult to locate between Missouri and the western Great Plains, even in places that had been collected in the past, so our focus was on the more heavily populated regions in the western US.

The lack of Eurasian matches for 10 of the US haplotypes is interesting. It is clear that L. draba is not native to the US, as the species name had been in use in Eurasia for more than a century before the first plant was detected in North America. It is possible that the haplotypes found only in the US were relatively rare in Eurasia, but the evolutionary processes during the introduction and subsequent invasion have increased their frequencies relative to those found in the native range.

Assignment scores trivially suggest that haplotypes 2 and 18 originated in central Asia, while haplotypes 9 and 11 originated in western Europe. The other haplotypes in Fig. 4 were assigned to multiple origins, such as haplotype 1, which could not be excluded from any region. Other assignments and their scores can be viewed in Fig. 4, but of interest are haplotype 21, excluded from Europe and with a highest score for originating in western Asia; and haplotype 10, excluded from central Asia and with a highest score for originating in Europe. The assignment results are additional evidence that there were introductions from different regions of Eurasia.

Implications for biological control agent research

It is notable that seven of the 31 Eurasian haplotypes (23%) were found exclusively west of the putative native range of Mulligan & Frankton (1962). This result supports an expanded native range of the L. draba that includes southern Europe [as suggested by Ball (1964)]. There is strong evidence that L. draba is not native to the British Isles (Scurfield 1962). Also, initial studies using amplified fragment length polymorphisms (AFLPs) on 190 plants of L. draba ssp. draba, mostly from Eurasia, indicated two major clades: one ranging from northern Spain to Italy and the other containing plants from across Eurasia (Fumanal et al. 2004). This also suggests that the true native range of L. draba may include southern Europe, and if so, it is possible that some diffusely co-evolved strains of biological control agents could exist exclusively in southern Europe. Alternatively, since alfalfa was introduced from Mesopotamia to western Europe some 2000 years BP (Hendry 1923), it seems possible that L. draba, also native to Mesopotamia and a common weed of alfalfa, may have also been introduced to western Europe at that time, and not have a such long evolutionary history in this region.

At this point there is no knowledge of how the haplotypes we found correlate with phenotypic characters that may be relevant to invasion success [see cautions in Reed & Frankham (2001)], such as habit, phenology, and, especially in the Brassicaceae, secondary metabolites such as glucosinolates (Kliebenstein 2004). Correlation between molecular markers and biological control efficacy has been shown in some systems, but that should not be assumed a priori. Even so, since this chloroplast DNA region is extremely variable within L. draba, it does allow us to investigate how the many different haplotypes are structured across invasive and native regions, and within populations. A frequently used alternative method of attempting to obtain genetically diverse invasive specimens for host-specificity testing is to collect from varied geographical regions. That method may not have worked in the case of L. draba, as the two most common haplotypes (1 and 10) account for over 50% of the US plants, and are found in almost all western US regions. Our AMOVA results show little variation between regions, thus, any future collections of a few plants from geographically diverse areas (e.g. for host-specificity testing) that do not utilize genotyping may represent little of the invasion’s genetic diversity.

This analysis provides US locations (Table S1, supplementary material) for a diversity of haplotypes that are represented in the invasion, which should assist in the current biological control feasibility studies. Our data specify which haplotypes are dominant in the invasion, and where all US haplotypes may be collected, occasionally in monotypic stands, which should yield a higher probability of collecting the desired haplotype. Additional sequencing of the chloroplast marker would be needed to confirm haplotype identity but, once that is done, the plants can be propagated clonally from root cuttings. Published microsatellite markers (Bon et al. 2005) should allow an even more precise analysis of the genetic structure of US populations.

Conclusion

The US L. draba invasion contains a high level of the genetic diversity found in its origins, and neither native nor introduced regions show evidence of strong population structuring across large regions. Certain invasive haplotypes most likely originated from different populations structuring across large regions. Certain invasive haplotypes most likely originated from different regions of Eurasia, and some of these are assigned to western Europe, supporting an expanded range for the species, and the potential for some biological control agents to exist exclusively in western Europe. Patches of L. draba are not exclusively clonal, and multiple haplotypes can be found within small populations. These data suggest that the L. draba invasion is not a genetically uniform entity originating from few introduction events, and any biological control effort should be prepared to contend with a genetically diverse invasion derived from multiple Eurasian origins.
Acknowledgements

Collections were generously provided by A. Baitulin, J. Beck, M. Cripps, B. Fumanal, M. Gaffri, F. Gahremanei-nejad, H. Hinze, D. Jewett, M. Khutishivili, D. Kurbanov, M. Mayer, J. McKenney, V. Samoilova and R. Sohbian. Laboratory work was performed by A. Goss, J. Lassey, C. Knudsen, K. Mann, and M. Mayer. Thanks to D. Kazmer and anonymous reviewers for useful comments on the manuscript. This research was made possible in part by financial support from the USA Department of the Interior, Bureau of Land Management and Bureau of Indian Affairs, and by the Key Knowledge Innovation Project of the Chinese Academy of Science (KZCX3-SW-343).

Supplementary material

The following materials are available from http://www.blackwellpublishing.com/products/journals/ supp/mat/MEC/MEC2589/MEC2589.sm.htm

Table S1 Population collection information for Lepidium spp.

Table S2 Haplotype vouchers and GenBank Accession numbers.

References


Rios JLV, García JFE (1998) Catálogo de Malezas de Mexico. Universidad Nacional Autónoma de Mexico y el Fondo de Cultura Económica, Mexico DF.


Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN, Version 2.000: a software program for population genetics data analysis. Genetics and Biometry Laboratory, Department of Anthropology, University of Geneva, Switzerland.


