Population genetic structure of yellow starthistle (*Centaurea solstitialis*), a colonizing weed in the western United States

Mei Sun

Abstract: Yellow starthistle (*Centaurea solstitialis* L.), a diploid annual native to the Mediterranean region, has rapidly colonized a variety of disturbed habitats in the western United States since its accidental introduction in the mid-19th century. Population genetic attributes were investigated in association with the species' colonizing success. Allozyme electrophoretic surveys were conducted for a total of 22 populations from California, Washington, and Idaho, including the southern and northern extremes of its distribution in the Pacific states. High levels of allozyme variation exist within populations, on average 43% of loci were polymorphic, with 2.88 alleles, a heterozygosity of 0.38, and gene diversity of 0.35 per polymorphic locus. At the species level, 56% of loci were polymorphic, with 2.85 alleles per locus, and the total gene diversity was 0.172. With the exception of a marginal population San Diego, the level of genetic diversity was similarly high in all regions. Lack of interpopulation genetic divergence (G_{ST} = 0.095) suggests that majority of the colonial populations were founded by a large number of genotypes and that high levels of gene flow may exist between local populations (Nm = 2.38). Information on genetic structure of the weed populations may aid our understanding of the species' colonizing ability in North America and is important for predicting the effectiveness of biological control program.

Key words: allozyme variation, *Centaurea solstitialis*, population genetics, weed colonization, yellow starthistle.

Résumé : La centaurée jaune (*Centaurea solstitialis* L.), une plante annuelle originaire de la Méditerranée, a rapidement colonisé une variété d'habitats perturbés dans l'ouest des États-Unis, depuis son introduction accidentelle au milieu du 19ième siècle. L'auteur a étudié les caractéristiques génétiques des populations en association avec les succès de colonisation de l'espèce. Il a conduit un suivi des allozymes par électrophorèses sur un total de 22 populations de la Californie, de Washington, et de l'Idaho, incluant les extrémités sud et nord de son aire de distribution dans les états du Pacific. On retrouve de fortes variations des allozymes dans les populations, en moyenne 43% des lieux sont polymorphes, avec 2,88 allèles, une hétérozygocité de 0,38, et une diversité génétique de 0,35 par lieu polymorphe. Au niveau de l'espèce, 56% des lieux sont polymorphes, avec 2,85 allèles par lieu, et une diversité génétique totale de 0,172. Sauf pour la population marginale San Diego, le niveau de diversité génétique est semblablement élevé dans toutes les régions. Le manque de divergence entre les populations (G_{ST} = 0,095) suggère que la majorité des populations colonisatrices ont originé d'un grand nombre de génotypes et qu'un important flux génique pourrait exister entre les populations locales (Nm = 2,38). Ces informations sur la structure génétique des populations de mauvaises herbes pourraient être utiles pour comprendre la capacité colonisatrice de l'espèce en Amérique du Nord et sont importantes pour prédire l'efficacité des programmes de lutte biologique.

Mots clés : variation allozymique, *Centaurea solstitialis*, génétique des populations, colonisation adventiste, centaurée jaune.
[Traduit par la rédaction]

Introduction

Colonization is the successful founding of a new population in a region or habitat not previously occupied by that species. Many weeds are examples of good colonizers. The detrimental effects of weeds on agriculture and their population dynamics have stimulated scientists from diverse disciplines to investigate the success of weed invasions in terms of ecological and historical factors, morphological and physiological preadaptation of weeds, and chemical and biological control methods (Harper 1960; Salisbury 1961; Crafts and Robbins 1962; Baker and Stebbins 1965; King 1966; Musik 1970; Baker 1974; Holm et al. 1977). More recently, population biologists have used weed groups as experimental systems for studies on a variety of topics, including demography, life-history variation, the significance of genetic polymorphism to colonizing ability, and mating-system evolution (e.g., Law et al. 1977; Mack and Pyke 1983; Brown and Burdon 1983; Burdon et al. 1983; Barrett 1992; Sun and Corke 1992).

Understanding the genetics of weed colonization is important not only to evolutionary studies but also to applied research that aims to control or eradicate weeds. Although empirical investigations of the population genetics of weeds are still fairly new, some general patterns have emerged (see reviews by Clegg and Brown 1983; Jain 1983; Barrett and Richardson 1986; Barrett 1992). The genetic features commonly found in weed colonizers include (*i*) genetically

Received December 10, 1996.

M. Sun. Department of Botany and Zoology, University of Hong Kong, Pokfulam Road, Hong Kong. e-mail: meisun@hkucc.hku.hk

depauperate populations with strong multilocus associations that originate from a limited number of colonizing genotypes; (*ii*) substantial interpopulation differentiation owing to founder effects, genetic drift, and variable environments; and (*iii*) polyploidy with attendant fixed heterozygosity (Baker 1965; Brown and Marshall 1981; Barrett 1992). However, contrasting patterns can be found that highlight the importance of investigating particular weed species of concern (e.g., Heiser 1965; Barrett 1978; Brown and Burdon 1983; Burdon and Brown 1987; Warwick et al. 1987).

Yellow starthistle (*Centaurea solstitialis* L.) is an introduced diploid annual composite ($2n = 16$) that has become extremely successful in invading and colonizing disturbed habitats in the western United States, especially in the central valley of California. This native eastern Mediterranean species is believed to have been introduced to California as a contaminant in alfalfa seed in the mid-1800s (Maddox and Mayfield 1985). It exhibits a great ecological amplitude in the introduced range and is capable of becoming established in habitats with either deep, well-drained or shallow, rocky soils, an annual precipitation that varies from 25 to 100 cm, and elevation from sea level to 2500 m (Maddox et al. 1985). The species' invasiveness is remarkable, and it often can be found in rangelands, at the edge of crop fields, in idle or abandoned land, orchards, vineyards, dryland pastures, roadsides, and wastelands (Robbins et al. 1951; Roché and Roché 1988).

The rapid expansion of yellow starthistle in the western United States has raised interest in its ability to colonize and spread. Ecological studies have revealed significant morphological variation in plant size and leaf shape, and in phenological variation among sites in bolting and flowering time (Roché 1965; Maddox 1981). Several seed-related traits relevant to its colonizing success have been investigated, including seed longevity (Callihan et al. 1993), dispersal (Roché 1991), and germination behavior under experimental or natural conditions in different habitats (Roché 1965; Maddox 1981; M. Sun, unpublished). Recent research efforts have focused on management and biological control (e.g., Callihan et al. 1989; Thomsen et al. 1989, 1993; Maddox et al. 1991). Although population genetic information on the species would be valuable in understanding the species' colonizing ability and in predicting the effectiveness of biological control, no previous research on this aspect has been reported.

In this study, the genetic structure of colonizing populations of yellow starthistle in western North America was investigated. The amount of within-population variation and degree of interpopulation divergence following invasions in the species were compared with the general features of other plant colonizers reported in the literature, in order to examine the factors affecting population genetic structure of yellow starthistle in North America. The relevance of genetic studies to biological control of yellow starthistle is discussed.

Materials and methods

Study populations

A total of 22 populations (Fig. 1) were sampled during July and August 1989, 1990, and 1995. Fifteen were from California, four from Washington, and three from Idaho (see Table 1 for population name and site information). Sites from the Sacramento Valley of California may represent the center of the species' density in North America (Robbins et al. 1951; Maddox and Mayfield 1985). Thus, 10 populations were sampled from this area. In Washington, the species was concentrated in the southeastern corner of the State, with nearly half of the total infestation occurring in Walla Walla County alone. This area was represented by samples from Walla Walla and Asotin. The species was also relatively common in southern Klickitat County. For assessing whether population bottlenecks or founder effects are important during secondary colonization, populations currently occupying geographically marginal sites, such as Gold Hill in the North and San Diego in the South, were included for comparison with those of central distribution representing the earliest colonization. In Idaho, yellow starthistle is primarily distributed in the west and north. The three Idaho populations studied were sampled from the Clearwater Canyon area,

Fig. 1. Outline map of the western United States indicating approximate sites of population collecting locations in California, Washington, and Idaho (see Table 1 for details).

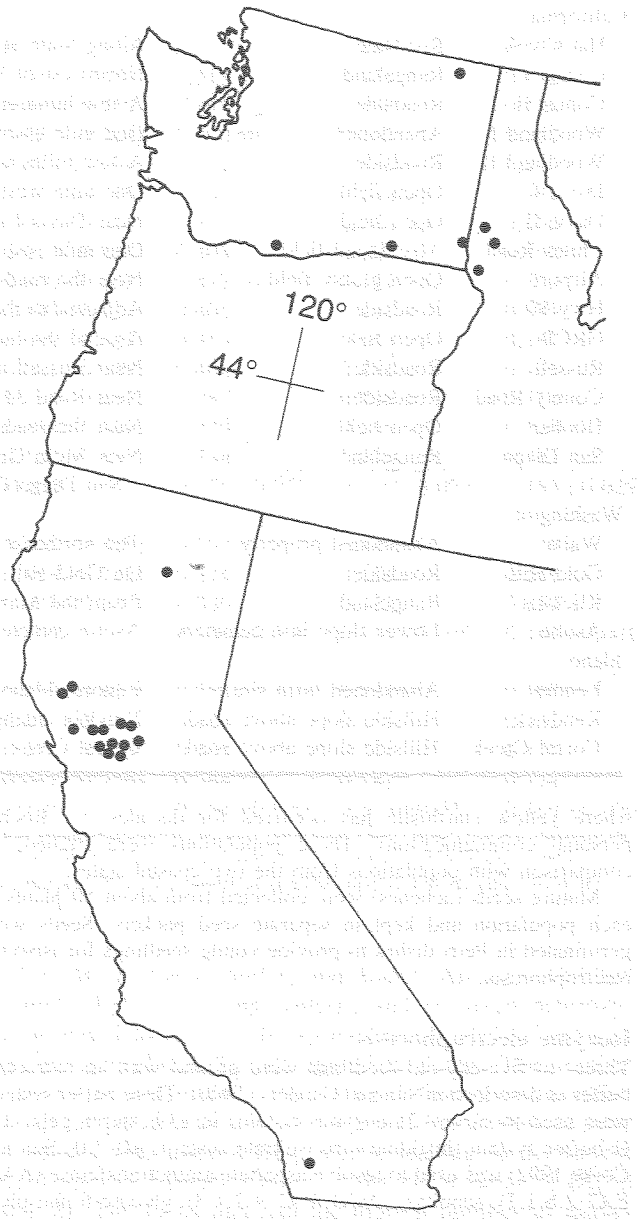

Table 1. Population names, habitats, and site locations of 22 populations of yellow starthistle studied.

Population	Habitat	Location
California		
Hat Creek	Roadside	Along State Highway 299, near Hat Creek
Colusa I	Rangeland	Northwest of Williams, 2 miles west of Lodoga, Colusa County
Colusa II	Roadside	A few hundred meters away from the Colusa I site, Colusa County
Woodland I	Abandoned farm site	One mile east of Woodland, Yolo County
Woodland II	Roadside	A few miles west of Woodland, Yolo County
Davis I	Open field	One mile west of the University of California-Davis Campus, Davis
Davis II	Open field	Near Davis I site, one mile west of the University of California-Davis Campus, Davis
Chiles Road	Abandoned field	One mile east of Davis
Airport	Open grassy field	Near the roadside of Sacramento Airport, Sacramento County
Hwy 80	Roadside	Adjacent to the Interstate Highway 80, Solano County
GRCP	Open field	Around the building of the Extension Services, University of California-Davis, Davis
Russell	Roadside	Near Russell and Fedrik Drive, Yolo County
County Road	Roadside	Near Road 31 and Road 96, Yolo County
Hoobar	Open field	Near the roadside of State Highway 128, one mile west of Hoobar
San Diego	Rangeland	Near Mesa Grande Road, west of Santa Ysabel, about 40 miles northeast of San Diego, San Diego County
Washington		
Walla	Abandoned property	The northeast edge of the city of Walla Walla, Walla Walla County
Gold Hill	Roadside	On Gold Hill, east of Kettle Falls, Stevens County
Klickitat	Rangeland	From the Mary Hill – Stone Henge area, Klickitat County
Asotin	Lower slope in a cemetery	Asotin cemetery, two miles south of Asotin on State Highway 129, Asotin County
Idaho		
Lenore	Abandoned farm site	Lenore, Idaho, T37N, R2W, sec. 28
Kendrick	Hillside slope above road	Kendrik, Idaho, T38N, R3W, sec. 35
Corral Creek	Hillside slope above road	Corral Creek, Idaho, T31N, R4W, sec. 7

where yellow starthistle has occurred for decades (C. Roché, personal communication). These populations were studied for comparison with populations from the two coastal states.

Mature seeds (achenes) were collected from about 50 plants in each population and kept in separate seed packets. Seeds were germinated in Petri dishes to provide young seedlings for isozyme electrophoresis.

Isozyme electrophoresis

Three- to five-day-old seedlings were ground with an extraction buffer as described in Sun and Ganders (1990). Three buffer systems were used to survey 21 enzyme systems in 13% starch gels. The H buffer system (histidine–tris–citrate system, pH 7.0; Sun and Corke 1992) was used to resolve aspartate aminotransferase (AAT; E.C. 2.6.1.1), aconitase (ACO; E.C. 4.2.1.3), glucose-6-phosphate dehydrogenase (G6PD; E.C. 1.1.1.49), isocitrate dehydrogenase (IDH; E.C. 1.1.1.42), malate dehydrogenase (MDH; E.C. 1.1.1.40), phosphoglucomutase (PGM; E.C. 2.7.5.1), 6-phosphogluconate dehydrogenase (6PGD; E.C. 1.1.1.44), and shikimate dehydrogenase (SKDH; E.C. 1.1.1.25). The L buffer system (pH 7.8; Shields et al. 1983) was used to resolve acid phosphatase (ACP; E.C. 3.1.3.2), alcohol dehydrogenase (ADH; E.C. 1.1.1.1), alkaline phosphatase (ALP; E.C. 3.1.3.1), diaphorase (DIA; E.C. 1.6.4.3), fluorescent esterase (F-EST; E.C. 3.1.1.1), (β-glucosidase (GLU; E.C. 3.2.1.21), leucine aminopeptidase (LAP; E.C. 3.4.11.1), and phosphoglucoisomerase (PGI; E.C. 5.3.1.9). The R buffer system (lithium–borate and tris–citrate, pH 8.3; Ashton and Braden 1961) was used to resolve esterase (EST; E.C. 3.1.1.1), glutamate dehydrogenase (GDH; E.C. 1.4.1.2), (α-glycerophosphate dehydrogenase (α-GPDH; E.C. 1.1.1.8), malic enzyme (ME; 1.1.1.40), and superoxide dismutase (SOD; E.C. 1.15.1.1). The protocols of Wendel and Weeden (1989) were used for enzyme activity staining. Genetic interpretation of band patterns followed standard principles (Weeden and Wendel 1989; Wendel and Weeden 1989).

Data analysis

Since some of the isozyme loci surveyed were not consistently recordable for every population, these loci were used to compute polymorphism and allelic diversity at the species level but were excluded in the population data analysis. Twenty loci (including both polymorphic and monomorphic loci) were used for computing parameters of genetic diversity at the population level: *Aat-1*, *Aat-2*, *Aco-1*, *Aco-2*, *Alp-2*, *Dia-2*, *Idh-1*, *Gdh-1*, *G6p-1*, *Lap-1*, *Mdh-1*, *Mdh-2*, *Mdh-3*, *Me-1*, *Pgi-1*, *Pgi-2*, *Pgm-1*, *Pgm-2*, *Skd-1*, and *6pg-1*.

Measures of within-population genetic variation include polymorphism, allelic diversity, and heterozygosity. The percentage of polymorphic loci, P, was calculated based on the 99% criterion. Since there were few rare alleles in the populations studied, the P values were not significantly different when the 95% criterion was used. Allelic diversity includes two measurements: the number of alleles per locus (A) and the number of alleles per polymorphic locus (A_p). Gene diversity (H), the mean expected heterozygosity assuming Hardy–Weinberg equilibrium, was calculated as

$$[1] \quad H = 1 - \frac{1}{n}\Sigma_j\Sigma_i p_{ij}^2$$

where p_{ij} is the frequency of the ith allele at the jth locus, and n is the total number of isozyme loci surveyed. Another measure of heterozygosity is the observed frequency of heterozygotes per locus per individual (H_o). Heterozygosities observed (H_{op}) and expected (H_{ep}) at polymorphic loci were calculated in the same way as for H_o and H but excluding monomorphic loci in the samples.

The number of multilocus genotypes present in a population is related to the number of polymorphic loci and number of alleles at each locus. In this study, six polymorphic loci were used to estimate the number of multilocus genotypes, including the *Lap-1* locus plus any other five polymorphic loci in each population. Data of the first 30 individuals for each population were used to count the number

Table 2. The mean within-population allozyme variation in 22 populations of yellow starthistle.

Population	N	P (%)	A	A_p	H_{op} (range)	H_{ep}	H_o	H
California								
Airport	54	45.0	1.95	3.11	0.415 (0.058–0.870)	0.388	0.187	0.194
Chiles Road	68	50.0	1.80	2.60	0.312 (0.029–0.670)	0.318	0.156	0.142
Colusa I	60	50.0	1.80	2.60	0.376 (0.043–0.750)	0.378	0.188	0.178
Colusa II	57	40.0	1.60	2.75	0.408 (0.083–0.780)	0.300	0.163	0.112
County Road	44	45.0	1.85	2.89	0.369 (0.091–0.698)	0.331	0.166	0.147
Davis I	64	50.0	1.75	2.50	0.363 (0.047–0.719)	0.351	0.182	0.145
Davis II	71	45.0	1.90	3.00	0.355 (0.085–0.614)	0.352	0.160	0.156
GRCP	52	45.0	1.90	3.00	0.399 (0.019–0.820)	0.405	0.180	0.192
Hat Creek	68	50.0	1.75	2.50	0.381 (0.060–0.662)	0.417	0.191	0.142
Hoobar	47	45.0	1.95	3.11	0.425 (0.170–0.723)	0.404	0.191	0.179
Hwy 80	41	45.0	1.95	3.11	0.363 (0.049–0.683)	0.353	0.163	0.167
Russell	49	45.0	1.90	3.00	0.408 (0.061–0.735)	0.408	0.184	0.193
San Diego	47	35.0	1.35	2.00	0.301 (0.152–0.457)	0.257	0.105	0.114
Woodland I	62	50.0	1.80	2.60	0.329 (0.038–0.690)	0.339	0.165	0.127
Woodland II	84	45.0	1.70	2.67	0.386 (0.102–0.633)	0.338	0.174	0.113
Mean ± SD	58	45.7±4.2	1.80±0.16	2.76±0.31	0.372±0.037	0.356±0.045	0.170±0.022	0.153±0.029
Idaho								
Lenore	43	30.0	1.75	3.50	0.498 (0.205–0.825)	0.297	0.149	0.149
Corral Creek	72	35.0	1.75	2.88	0.454 (0.113–0.718)	0.314	0.159	0.149
Kendrick	53	40.0	1.75	3.14	0.321 (0.038–0.547)	0.300	0.128	0.159
Mean ± SD	56	35.0±5.0	1.75±0.00	0.32±0.31	0.424±0.092	0.304±0.009	0.145±0.016	0.149±0.001
Washington								
Asotin	67	35.0	1.90	3.57	0.375 (0.030–0.612)	0.353	0.131	0.167
Gold Hill	64	40.0	1.75	2.88	0.363 (0.094–0.613)	0.326	0.145	0.162
Klickitat	64	40.0	1.75	2.67	0.342 (0.078–0.623)	0.368	0.154	0.174
Walla	43	45.0	2.00	3.22	0.403 (0.047–0.756)	0.368	0.181	0.174
Mean ± SD	60	40.0±4.1	1.85±0.12	3.09±0.40	0.371±0.025	0.354±0.020	0.153±0.021	0.169±0.006

Note: Variables are as follows: P, percentage of polymorphic loci; A, number of alleles per locus; A_p, number of alleles per polymorphic locus; H_{op} and H_{ep}, observed and expected heterozygosity, respectively, averaged over polymorphic loci; H_o, observed heterozygosity; and H, gene diversity averaged over all loci.

of different multilocus genotypes. Based on the allelic composition at each of the polymorphic loci, the theoretical upper limit to the number of multilocus genotypes that can be generated was calculated for each population as the product of the number of expected genotypes at each locus, assuming random recombination of genotypes between loci.

The proportion of genetic variation distributed within versus between populations was measured using Nei's (1973) gene diversity statistics. The coefficient of genetic differentiation among populations, G_{ST}, was used to estimate the level of gene flow, Nm (the number of migrants exchanged between local populations per generation), based on the following relationship:

$$[2] \quad G_{ST} = \frac{1}{4Nm + 1}$$

where G_{ST} is Nei's (1973) estimator of F_{ST} (Wright 1951). The UPGMA dendrogram of Nei's (1972) genetic distance was constructed using the GDD computer program provided by K. Ritland.

Results

Of the 21 enzyme systems investigated, a total of 36 isozyme loci were detected. A high number of polymorphic loci exist in the species, including *Acp-1*, *Acp-2*, *Aco-1*, *Aco-2*, *Adh-1*, *Alp-1*, *Alp-2*, *Alp-3*, *Dia-1*, *Dia-2*, *Est-1*, *Est-2*, *Lap-1*, *Pgi-2*, *Pgm-1*, *Pgm-2*, *Pgm-3*, *6pgd-1*, *6pgd-2*, and *Skdh-1*. The monomorphic loci include *Aat-1*, *Aat-2*, *Adh-2*, *F-est -1*, *β-Glu-1*, *α-Gpdh-1*, *Idh-1*, *Gdh-1*, *G6p-1*, *Mdh-1*, *Mdh-2*, *Mdh-3*, *Me-1*, *Pgi-1*, *Sod-1*, and *Sod-2*. All loci exhibited a simple diploid banding pattern, and no fixed heterozygosity was detected. At the species level, 56% of the loci were polymorphic with 2.85 alleles per locus. The number of alleles at polymorphic loci ranged from two to eight. Two of the polymorphic loci, *Aco-2* and *Pgi-2*, showed only two alleles with uneven frequencies, leading to the lowest heterozygosity estimate (shown as the lower range of H_{op} in Table 2). The locus that had the highest number of alleles was *Lap-1*, which gave the highest heterozygosity estimate (shown as the high range of H_{op} in Table 2). All populations exhibited similarly high levels of genetic variation, except the San Diego population, which had the lowest values for all parameters calculated (Table 2). The average values across all populations were $P = 43.2 \pm 5.7\%$ (mean ± SD), $A = 1.80 \pm 0.14$; $A_p = 2.88 \pm 0.36$; $H_{op} = 0.379 \pm 0.046$, $H_{ep} = 0.348 \pm 0.042$, $H_o = 0.164 \pm 0.022$, and $H = 0.156 \pm 0.025$.

The number of multilocus genotypes in each population was found to be limited only by the sample or population size since every individual sampled in the population possessed a different six-locus genotype. The maximum number of six-locus genotypes expected, assuming populations were infinitely large, ranged from 729 (San Diego) to 69 984 (Walla and Hwy 80) and averaged 32 677 ± 20 201 across populations. The observed variation in the theoretical upper limit to

Fig. 2. Dendrogram (UPGMA method) of Nei's genetic distances between 22 populations of yellow starthistle.

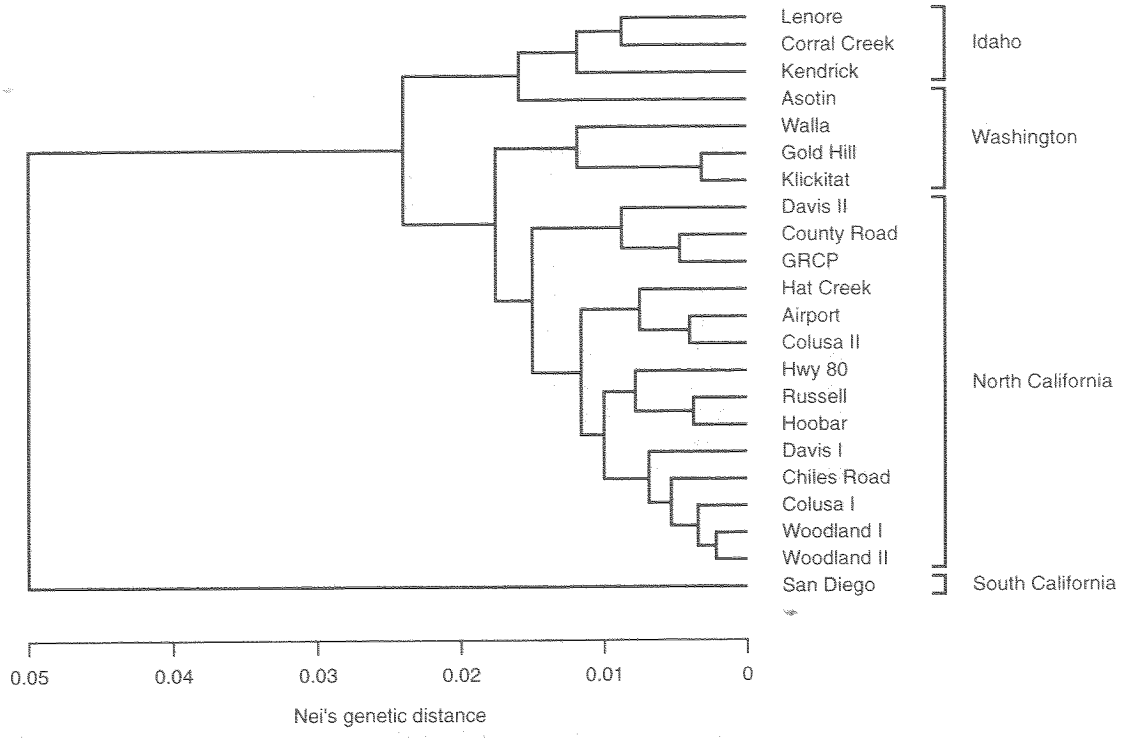

Table 3. Genetic diversity parameters and estimates of gene flow (*Nm*) in populations of yellow starthistle.

	All populations	California	Washington	Idaho
No. of populations	22	15	4	3
Total gene diversity (H_T)	0.172	0.166	0.179	0.156
Within populations (H_S)	0.156±0.057	0.153±0.057	0.170±0.058	0.149±0.055
Between populations (D_{ST})	0.016±0.011	0.013±0.012	0.010±0.006	0.007±0.007
Coefficient of genetic differentiation (G_{ST})	0.095	0.075	0.053	0.042
Estimates of *Nm*	2.38	3.08	4.47	5.70
Mean genetic distance	0.027±0.014	0.023±0.014	0.019±0.008	0.011±0.008

the number of six-locus genotypes was caused by between-population variation in allelic diversity at the loci sampled. When the multilocus genotypes were based on three- or four-locus polymorphism, 73 or 93% of the individuals, respectively, could be genotypically distinguished in the populations. No multilocus association or linkage disequilibrium was detectable in the samples.

Nei's total gene diversity (H_T) was estimated to be 0.172 in the 22 populations studied, of which the within population diversity (H_s) was 0.156 ± 0.057 and between population diversity was 0.016 ± 0.011. A low coefficient of population genetic differentiation was obtained, G_{ST} = 0.095. Based on this value, the level of gene flow between populations, *Nm*, was estimated to be 2.38 per generation. Populations from each state were separately analyzed to compare regional genetic diversity (Table 3). Levels of genetic diversity were found to be similarly high in all regions.

To explore whether a microgeographic pattern of secondary colonization exists, pairwise comparisons of genetic relationship between populations were made based on the matrix of Nei's genetic distances (Table 4). The most similar pairs of populations were Woodland I and Woodland II (I = 0.998 ± 0.014), Woodland I and Airport (I = 0.997 ± 0.014), and Woodland II and Davis I (I = 0.997 ± 0.005). The most distant pair was Walla and San Diego (D = 0.091 ± 0.050). The average genetic distance among the 22 populations was 0.027 ± 0.014, and genetic similarity was 0.973. The highest average genetic distance was found between the San Diego population and the rest of the populations studied (0.05 ± 0.02). As shown in Fig. 2, grouping of the populations was largely in agreement with their geographical distribution on a regional scale.

Discussion

Genetic variation within populations

Colonizing species as a group often are expected to be markedly depauperate in genetic variation within populations owing to founder effects and genetic drift, and many empirical studies have provided supporting evidence for this (e.g., Brown and Marshall 1981, Table V). However, nearly all the species listed in Brown and Marshall were either known

or presumed to be self-pollinating. It is often difficult to separate founder effects from the effects of uniparental mating on population genetic structure during colonization. Comparative studies of actively colonizing plant species with contrasting mating systems may help to identify genetic consequences of colonization (e.g., Sun, 1997).

High levels of genetic variation exist in the populations of yellow starthistle in North America, with average values of P, A, H, and H_o being 15, 2, 12, and 82 times higher, respectively, than the corresponding means in the inbreeding colonizers listed in Brown and Marshall (1981). Comparing the yellow starthistle data with those reported in 163 studies of allozyme variation (reviewed in Loveless and Hamrick 1984), levels of observed heterozygosity and gene diversity at polymorphic loci were also markedly higher than the corresponding mean values in other plant species. However, the level of genetic variation in yellow starthistle is very similar to an annual outbreeding weed, *Echium plantagineum* L., one of the most successful invaders of temperate Australia (Burdon and Brown 1987). On average, 2.7 alleles, a gene diversity of 0.34, and observed heterozygosity of 0.32 per locus were reported at the polymorphic loci in colonial populations of *E. plantagineum*, which are comparable with the corresponding values in the populations of yellow starthistle studied (2.9, 0.35, and 0.38, respectively).

The "general-purpose" genotype (Baker 1965) that is believed to occur frequently in colonizers and confer wide environmental tolerance does not exist in colonial populations of yellow starthistle. The species exhibited an extremely high level of multilocus genotypic diversity, with virtually every individual possessing a unique multilocus genotype. The wide ecological amplitude of yellow starthistle in North America may arise at least in part from this high genotypic variability. A similar level of multilocus genotypic diversity was reported in the outbreeding weed *E. plantagineum* (Brown and Burdon 1983), whereas few multilocus genotypes were observed in self-fertilizing weeds (e.g., Warwick 1990). A study of mating system parameters of yellow starthistle from California, Idaho, and Washington found an average population outcrossing rate of 0.975 ± 0.023 (M. Sun, unpublished). Thus, the pattern of within-population genetic diversity in weed colonizers appears to be primarily associated with their breeding systems.

Comparisons of allelic diversity between marginal and central populations can reveal whether genetic bottlenecks or founder effects are prominent during secondary or local colonization. Theoretically, when an outbred population passes through a bottleneck, genetic variation should decline in direct proportion to the severity of the bottleneck (Nei et al. 1975). Thus, losses of genetic variation can be used to detect bottlenecks or founder effects in colonial populations. In comparison with heterozygosity, allelic diversity is the most sensitive indicator of a bottleneck or founder effects (e.g., Sun 1996). Most populations of yellow starthistle studied apparently experienced no significant bottlenecks during colonization, likely owing to presence of a large number of genotypes at founding (e.g., as a contaminant in commercial seed stocks). A marked reduction in genetic variation was found in only one marginal population, San Diego. Despite its large population size and high density, 16 alleles at the polymorphic loci (nearly 50% of the total)

Table 4. Matrix of Nei's genetic distances between 22 populations of yellow starthistle.

Population	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1 Colusa I	0.004	0.007	0.008	0.009	0.009	0.004	0.032	0.011	0.029	0.042	0.029	0.028	0.018	0.014	0.017	0.020	0.015	0.019	0.026	0.027	0.040
2 Woodland II		0.010	0.004	0.003	0.011	0.002	0.008	0.005	0.006	0.016	0.009	0.016	0.007	0.009	0.009	0.009	0.006	0.022	0.012	0.011	0.024
3 Colusa II			0.011	0.016	0.009	0.010	0.025	0.004	0.009	0.018	0.014	0.030	0.014	0.015	0.017	0.021	0.011	0.039	0.023	0.027	0.033
4 Chiles Road				0.010	0.012	0.004	0.025	0.012	0.007	0.021	0.006	0.022	0.012	0.010	0.007	0.009	0.011	0.035	0.013	0.013	0.030
5 Davis I					0.018	0.007	0.018	0.012	0.021	0.031	0.021	0.017	0.009	0.012	0.015	0.011	0.011	0.025	0.019	0.020	0.035
6 Hat Creek						0.010	0.028	0.007	0.017	0.020	0.021	0.023	0.015	0.018	0.019	0.022	0.013	0.047	0.021	0.029	0.035
7 Woodland I							0.011	0.003	0.004	0.011	0.006	0.016	0.004	0.009	0.005	0.007	0.006	0.020	0.008	0.012	0.019
8 Asotin								0.024	0.021	0.018	0.025	0.039	0.036	0.020	0.011	0.015	0.035	0.058	0.021	0.013	0.015
9 Airport									0.016	0.020	0.020	0.029	0.016	0.011	0.011	0.017	0.016	0.042	0.034	0.027	0.029
10 Klickitat										0.011	0.004	0.027	0.019	0.019	0.011	0.014	0.020	0.067	0.039	0.027	0.032
11 Walla											0.014	0.030	0.028	0.022	0.014	0.021	0.031	0.091	0.025	0.024	0.020
12 Gold Hill												0.020	0.014	0.017	0.009	0.009	0.015	0.067	0.034	0.023	0.027
13 Davis II													0.012	0.019	0.018	0.005	0.017	0.062	0.023	0.028	0.026
14 GRCP														0.015	0.012	0.005	0.009	0.050	0.041	0.029	0.029
15 Hwy 80															0.010	0.013	0.006	0.044	0.028	0.016	0.026
16 Hoobar																0.011	0.004	0.051	0.028	0.016	0.018
17 County Road																	0.008	0.054	0.024	0.016	0.015
18 Russell																		0.044	0.039	0.024	0.031
19 San Diego																			0.069	0.058	0.087
20 Lenore																				0.012	0.009
21 Kendrick																					0.014
22 Corral Creek																					

were absent in the population, indicating a severe genetic bottleneck in its establishment. However, despite the impact of the initial bottleneck on genetic variation in the population, subsequent evolution can lead to high levels of heterozygosity through rapid population growth (Nei et al. 1975). Maddox (1981) documented high reproductive output in yellow starthistle populations, as more than 10 000 seeds can be produced per plant in some sites. Thus, a rapid recovery in population size following invasion could be expected for the species as a whole, and genetic drift is apparently insignificant in the population following invasion.

Populations of yellow starthistle from the native range were not available to compare with North American populations in this study. However, the level of within-population genetic diversity likely is similar between the two continents. Similar to yellow starthistle, *E. plantagineum* is of Mediterranean origin. It was introduced to Australia during the middle of the 19th century as a garden ornamental and as a contaminant of seed and stock feed. A comparative study of the species in Australia and Europe found an equally high level of genetic diversity in the putative source populations as in the colonial populations (Burdon and Brown 1987). Another outbreeding weed, *Apera spica-venti* L., also exhibited similar levels of variability in native European and introduced Canadian populations (Warwick et al. 1987).

Genetic divergence between populations

Several factors contribute to population genetic differentiation in a colonizing species, such as founder effects, isolation from the source population, and novel selection pressures in the new habitats (Clegg and Brown 1983). In many cases, invasion of new territories by weedy colonizers has led to genetic divergence between populations ($G_{ST} > 0.40$; see review in Loveless and Hamrick 1984). Again, the observed pattern of population differentiation could be related to the prevalence of uniparental modes of reproduction in the weedy species surveyed, which in itself can lead to a high level of population differentiation (mean $G_{ST} = 0.523$; Loveless and Hamrick 1984).

Populations of yellow starthistle in western United States span a range of habitats such that climatic and edaphic differences among environments probably exert selection for phenotypic divergence. Although morphological and phenological differentiation are apparent among populations of yellow starthistle (Maddox 1981), there is a general lack of interpopulation divergence at the allozyme level. This pattern could result from the combination of an outbreeding system with large founder sizes. In addition, the relatively recent spread of yellow starthistle in North America may not have allowed sufficient time for population divergence in response to local selection pressures.

The southernmost marginal population, San Diego, was the only one strongly differentiated from all others, supporting predictions of the central–marginal model of colonization of Barrett and Husband (1989). This population apparently is of recent origin, and it occurs at only one of the two sites reported from San Diego County (Beauchamp 1986). The San Diego population also was more genetically substructured than the others, leading to a much smaller mating neighborhood and thus to substantial inbreeding (M. Sun, unpublished). In contrast, no significant founder effects were observed in Gold Hill, the northernmost marginal population. The level of genetic variation in this population is similar to the average in other more centrally distributed populations. The Gold Hill population was sampled from an area with a recorded history of yellow starthistle invasion since 1928, and the site is not geographically isolated from other populations in Stevens County.

The colonizing history of yellow starthistle in North America

Roché (1991) reported that the dispersal distance of yellow starthistle seeds under natural conditions is very limited. However, its rapid spread in western North America suggests effective seed dispersal, such as may occur through anthropogenic movement.

It generally is believed that there have been multiple introductions of yellow starthistle to the United States (e.g., Maddox et al. 1985). According to Roché and Talbott (1986), the species was first reported near western U.S. seaports. Early California records indicate its occurrence at Oakland in 1869 and at Vacaville in 1887. The earliest collection in Washington was in 1898. Similar levels of genetic diversity in California and Washington seem to support the multiple introduction hypothesis. However, lack of genetic divergence among populations suggests that if multiple introduction events occurred, seeds probably came from the same source region or from different populations that were not genetically differentiated.

On a local scale, microgeographical analysis of population genetic composition may aid in constructing the historical processes of colonization of yellow starthistle in the western states. For example, regional distributions of yellow starthistle were well documented in the state of Washington (see information in Talbott 1987). Populations were located primarily inland in the State, and the highest frequency of occurrence and population density were in Walla Walla County, with a recorded history from 1900. Populations from this area contain a high level of allelic diversity (Table 2). The species was reported to occur in Klickitat County from 1924 and in Stevens County from 1928. Populations from these two counties could be the products of secondary migration from the Walla Walla area, as shown by the loss of several alleles in Gold Hill and Klickitat in comparison with Walla and Asotin. Although the species was not recorded in Asotin County until 1970, Asotin contains nearly all the alleles present in Walla. This supports a history of sequential movement of yellow starthistle from Walla Walla to Columbia then to Garfield and finally to Asotin in southeastern Washington (C. Roché, personal communication). In contrast, populations from adjacent Idaho contain fewer alleles than Asotin, but it is unlikely that they were derived through this sequential dispersal route as they represent older invasions than populations in Asotin County.

Relevance to biological control

Information on the genetic diversity present in a particular area, as well as knowledge of the source region for a particular invasion, can be of value in devising effective methods of biological control of a weed (Barrett 1992). Scientists interested in biological control of yellow starthistle need information on the number of biotypes and (or) genotypes

in the introduced populations. If multiple introductions of yellow starthistle to North America occurred historically and the genetically diverse colonizers came from different source regions, attempts to search for the species' natural enemies from a particular source population within the native range would prove to be futile. On the other hand, if only a single introduction or multiple introductions from a single source actually occurred, the use of allozyme information can help to identify the specific source population that may provide the best adapted pathogens or insects that could attack yellow starthistle in North America.

The success of a biological control program also depends on the likelihood of weed populations evolving resistance to pest or pathogen attack. In a genetically more variable species, the level of genetic variation in resistance to the control agent is likely to be high. A negative relationship has been suggested between the level of genetic variation of the target species and the degree of biological control that can be achieved (Burdon and Marshall 1981; Burdon and Brown 1987; Barrett 1992). The population genetic structure of yellow starthistle in North America points to the potential for rapid evolution of resistance to its introduced pathogens. Thus, biological control programs alone may not be effective in the long-term suppression of yellow starthistle populations. However, a combination of biological control with other control methods, such as controlled livestock grazing (e.g., Thomsen et al. 1993), plant competition (e.g., Brown et al. 1993), herbicides, timed mowing, and prescribed burning, would be able to achieve more effective long-term control of this invasive weed.

Conclusions

High levels of genetic variation exist within the colonizing populations of yellow starthistle in western North America, and little genetic divergence has occurred among populations. The most important contributors to this pattern of genetic structure appear to be its outbreeding system and anthropogenic factors in seed dispersal. The high levels of genetic variability could contribute further to the species' colonizing success in North America and add difficulties to its biological control. It is important for weed control specialists to be familiar with the extent of genetic variation present in the North American populations of yellow starthistle. A synthesis of information on the ecology, physiology, and genetics of yellow starthistle is needed to develop an effective management program.

Acknowledgments

The author thanks H. Corke, C.T. Roché, B.F. Roché, J. Connett, J.P. McCaffery, C.D. Thomsen, and D. Allen for assistance with seed collections; C.D. Thomsen for his work on yellow starthistle management that started my research curiosity on the species; A. Lai for help with laboratory work; and H. Corke, C. Roché, and two reviewers for helpful comments and suggestions that improved the manuscript. This research was supported by a CRCG grant from the University of Hong Kong.

References

Ashton, G.C., and Braden, A.W.H. 1961. Serum β-globulin polymorphism in mice. Aust. J. Biol. Sci. **14**: 248–254.

Baker, H.G. 1965. Characteristics and modes of origins of weeds. *In* The genetics of colonizing species. *Edited by* H.G. Baker and G.L. Stebbins. Academic Press, London, U.K.. pp. 141–172.

Baker, H.G. 1974. The evolution of weeds. Annu. Rev. Ecol. Syst. **5**: 1–24.

Baker, H.G., and Stebbins, G.L. 1965. The genetics of colonizing species. Academic Press, London, U.K.

Barrett, S.C.H. 1978. Heterostyly in a tropical weed: the reproductive biology of the *Turnera ulmifolia* complex (Turneraceae). Can. J. Bot. **56**: 1713–1725.

Barrett, S.C.H. 1992. Genetics of weed invasions. *In* Applied population biology. *Edited by* S.K. Jain and L.W. Botsford. Kluwer Academic Publishers, Dordrecth, the Netherlands. pp. 91–119.

Barrett, S.C.H., and Husband, B.C. 1989. The genetics of plant migration and colonization. *In* Plant population genetics, breeding, and genetic resources. *Edited by* A.H.D. Brown, M.T. Clegg, A.L. Kahler, and B.S. Weir. Sinauer, Sunderland, Mass. pp. 254–277.

Barrett, S.C.H., and Richardson, B.J. 1986. Genetic attributes of invading species. *In* Ecology of biological invasions: an Australian perspective. *Edited by* R.H. Groves and J.J. Burdon. Australian Academy of Science, Canberra, Australia. pp. 21–30.

Beauchamp, R.M. 1986. A flora of San Diego County, California. Sweet Water River Press, National City, Calif.

Brown, A.H.D., and Burdon, J.J. 1983. Multilocus diversity in an outbreeding weed, *Echium plantagineum* L. Aust. J. Biol. Sci. **36**: 503–509.

Brown, A.H.D., and Marshall, D.R. 1981. Evolutionary changes accompanying colonization in plants. *In* Evolution today. *Edited by* G.E.C. Scudder and J.L. Reveal. Hunt Institute for Botanical Documentation, Carnegie-Mellon University, Pittsburgh, Pa. pp. 351–363.

Brown, C.S., Bugg, R.L., and Anderson, J.H. 1993. Roadside plantings of perennial grasses. Grasslands, **3**: 1–3.

Burdon, J.J., and Brown, A.H.D. 1987. Population genetics of *Echium plantagineum*, a target weed for biological control. Aust. J. Biol. Sci. **39**: 369–378.

Burdon, J.J., and Marshall, D.R. 1981. Biological control and the reproductive mode of weeds. J. Appl. Ecol. **18**: 649–658.

Burdon, J.J., Marshall, D.R., and Brown, A.H.D. 1983. Demographic and genetic changes in populations of *Echium plantagineum*. J. Ecol. **71**: 667–679.

Callihan, R.H., Northam, F.E., Johnson, J.B., Michalson, E.L., and Prather, T.S. 1989. Yellow starthistle biology and management in pasture and rangeland. Curr. Inf. Ser. No. 634. University of Idaho, College of Agriculture, Moscow, Russia.

Callihan, R.H., Prather, T.S., and Northam, F.E. 1993. Longevity of yellow starthistle (*Centaurea solstitialis*) achenes in soil. Weed Technol. **7**: 33–35.

Clegg, M.T., and Brown, A.H.D. 1983. The founding of plant populations. *In* Genetics and conservation. *Edited by* C.M. Schonewald-Cox, S.M. Chambers, B. MacBryde, and W.L. Thomas. Benjamin/Cummings, Redwood City, Calif. pp. 216–228.

Crafts, A.S., and Robbins, W.W. 1962. Weed control. McGraw-Hill, New York.

Harper, J.L. 1960. Biology of weeds. Br. Ecol. Soc. Symp. No. 1. Blackwell, Oxford, U.K.

Heiser, C.B. 1965. Sunflowers, weeds and cultivated plants. *In* The genetics of colonizing species. *Edited by* H.G. Baker and G.L. Stebbins. Academic Press, London, U.K. pp. 391–401.

Holm, L.G., Plucknett, D.L., Pancho, J.V., and Herberger, J.P. 1977. The world's worst weeds: distribution and biology. University of Hawaii Press, Honolulu, Hawaii.

Jain, S.K. 1983. Genetic characteristics of populations. *In* Ecological

studies: analysis and synthesis. *Edited by* H.A. Mooney and M. Gordon. Springer-Verlag, Berlin, Germany. pp. 240–256.
King, L.J. 1966. Weeds of the world: biological control. Hill, London, U.K.
Law, R., Bradshaw, A.D., and Putwain, P.D. 1977. Life history variation in *Poa annua*. Evolution, **31**: 233–246.
Loveless, M.D., and Hamrick, J.L. 1984. Ecological determinants of genetic structure in plant populations. Annu. Rev. Ecol. Syst. **15**: 65–95.
Mack, R.N., and Pyke, D.A. 1983. The demography of *Bromus tectorum*: variation in time and space. J. Ecol. **71**: 69–93.
Maddox, D.M. 1981. Introduction, phenology, and density of yellow starthistle in coastal, intercoastal, and central valley situations in California. Agric. Res. Results No. ARR-W-20/July. U.S. Department of Agriculture, Agriculture Research Service, Washington, D.C.
Maddox, D.M., and Mayfield, A. 1985. Yellow starthistle infestations are on the increase. Calif. Agric. **39**: 10–12.
Maddox, D.M., Joley, D.B., Mayfield, A., and Mackey, B.E. 1991. Impact of *Bangasternus orientalis* (Coleoptera: Curculionidae) on achene production of *Centaurea solstitialis* (Asterales: Asteraceae) at a low and high elevation site in California. Environ. Entomol. **20**: 335–337.
Maddox, D.M., Mayfield, A., and Poritz, N.H. 1985. Distribution of yellow starthistle (*Centaurea solstitialis*) and Russian knapweed (*Centaurea repens*). Weed Sci. **33**: 315–327.
Musik, T.J. 1970. Weed biology and control. McGraw-Hill, New York.
Nei, M. 1972. Genetic distance between populations. Am. Nat. **106**: 283–292.
Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. U.S.A. **70**: 3321–3323.
Nei, M., Maruyama, T., and Chakraborty, R. 1975. The bottleneck effect and genetic variability in populations. Evolution, **29**: 1–10.
Robbins, W.W., Bellue, M.K., and Ball, W.S. 1951. Weeds of California. Printing Division (Documents Section), California State Department of Agriculture, Sacramento. Calif.
Roché, B.F. 1965. Ecologic studies of yellow starthistle (*Centaurea solstitialis* L.). Ph.D. dissertation, University of Idaho, Moscow, Russia.
Roché, B.F. 1991. Achene dispersal in yellow starthistle (*Centaurea solstitialis* L.). Northwest Sci. **66**: 62–65.
Roché, B.F., and Roché, C.T. 1988. Distribution and amount of four knapweed (*Centaurea* L.) species in eastern Washington. Northwest Sci. **62**: 242–253.
Roché, B.F., and Talbott, C.J. 1986. The collection history of *Centaureas* found in Washington state. Res. Bull. No. XB 0978. Agriculture Research Center, College of Agriculture and Home Economics, Washington State University, Pullman, Wash.
Salisbury, E.J. 1961. Weeds and aliens. Collins, London, U.K.
Shields, C.R., Orton, T.J., and Stuber, C.W. 1983. An outline of general resource needs and procedures for the electrophoretic separation of active enzymes for plant tissue. *In* Isozymes in plant genetics and breeding. Part A. *Edited by* S.D. Tanksley and J.J. Orton. Elsevier Science, Amsterdam, the Netherlands. pp. 443–468.
Sun, M. 1996. Genetic diversity in *Spiranthes sinensis* and *S. hongkongensis*: the effects of population size, mating system, and evolutionary origin. Conserv. Biol. **10**: 785–795.
Sun, M. 1997. Genetic diversity in three colonizing orchids with contrasting mating systems. Am. J. Bot. **84**: 224–232.
Sun, M., and Corke, H. 1992. Population genetics of colonizing success of weedy rye in northern California. Theor. Appl. Genet. **83**: 321–329.
Sun, M., and Ganders, F.R. 1990. Outcrossing rates and allozyme variation in rayed and rayless morphs of *Bidens pilosa*. Heredity, **64**: 139–143.
Talbott, C.J. 1987. Distribution and ecologic amplitude of selected *Centaurea* species in eastern Washington. M.S. thesis, Washington State University, Pullman, Wash.
Thomsen, C.D., Williams, W.A., George, M.R., McHenry, W.B., Bell, F.L., and Knight, R.S. 1989. Managing yellow starthistle on rangeland. Calif. Agric. **43**: 4–6.
Thomsen, C.D., Williams, W.A., Vayssieres, M., Bell, F.L., and George, M.R. 1993. Controlled grazing on annual grassland decreases yellow starthistle. Calif. Agric. **47**: 36–40.
Warwick, S.I. 1990. Allozyme and life history variation in five northwardly colonizing North American weed species. Plant Syst. Evol. **169**: 41–54.
Warwick, S.I., Thompson, B.K., and Black, L.D. 1987. Genetic variation in Canadian and European populations of the colonizing weed species *Apera spica-venti*. New Phytol. **106**: 301–317.
Weeden, N.F., and Wendel, J.F. 1989. Genetics of plant isozymes. *In* Isozymes in plant biology. *Edited by* D.E. Soltis and P.S. Soltis. Dioscorides Press, Portland, Oreg. pp. 46–72.
Wendel, J.F., and Weeden, N.F. 1989. Visualization and interpretation of plant isozymes. *In* Isozymes in plant biology. *Edited by* D.E. Soltis and P.S. Soltis. Dioscorides Press, Portland, Oreg. pp. 5–45.
Wright, S. 1951. The genetical structure of populations. Ann. Eugen. **15**: 323–354.