

# High genetic diversity and population differentiation in *Boechera fecunda*, a rare relative of *Arabidopsis*

BAO-HUA SONG and THOMAS MITCHELL-OLDS

Department of Biology, Duke University, PO Box 90338, Durham, NC 27708, USA

## Abstract

Conservation of endangered species becomes a critical issue with the increasing rates of extinction. In this study, we use 13 microsatellite loci and 27 single-copy nuclear loci to investigate the population genetics of *Boechera fecunda*, a rare relative of *Arabidopsis thaliana*, known from only 21 populations in Montana. We investigated levels of genetic diversity and population structure in comparison to its widespread congener, *Boechera stricta*, which shares similar life history and mating system. Despite its rarity, *B. fecunda* had levels of genetic diversity similar to *B. stricta* for both microsatellites and nucleotide polymorphism. Populations of *B. fecunda* are highly differentiated, with a majority of genetic diversity existing among populations ( $F_{ST} = 0.57$ ). Differences in molecular diversity and allele frequencies between western and eastern population groups suggest they experienced very different evolutionary histories.

**Keywords:** adaptation, *Boechera fecunda*, conservation, genetic variation, molecular markers, population differentiation

Received 11 March 2007; revision received 7 June 2007; accepted 10 July 2007

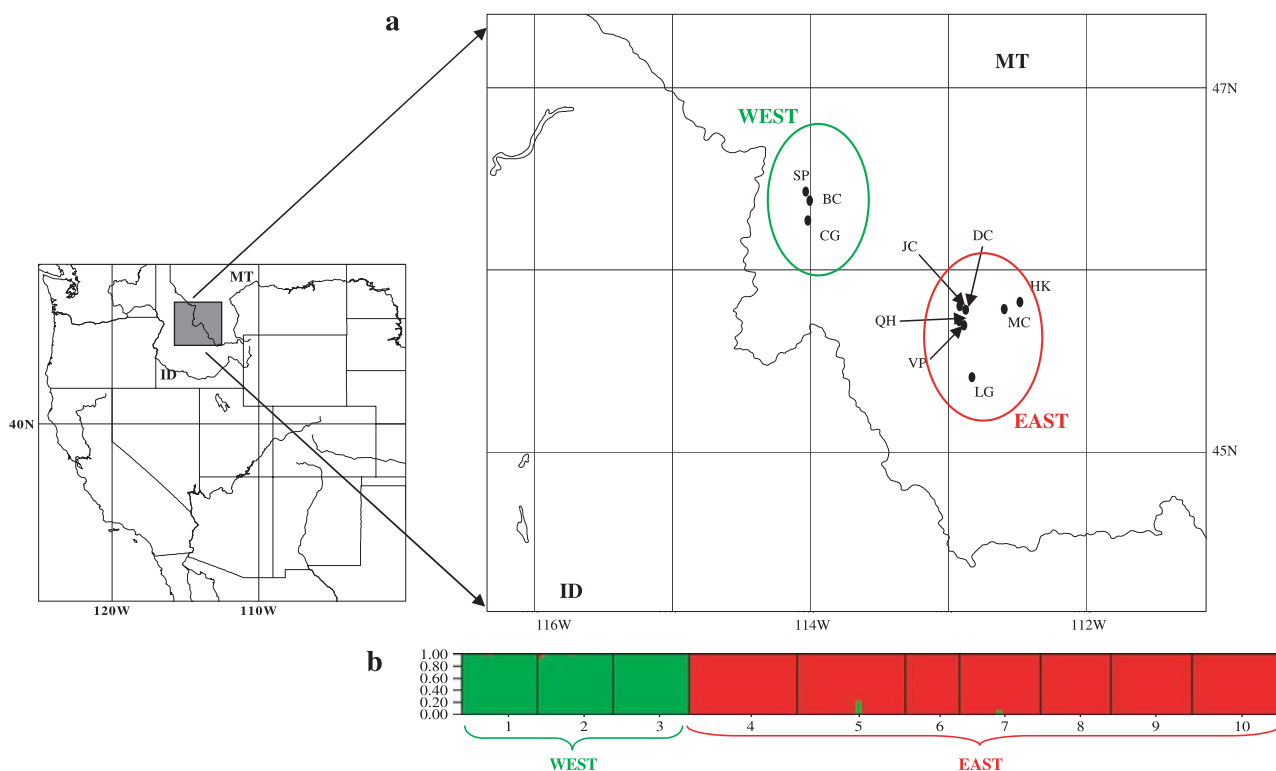
## Introduction

The continuing extinction of many species highlights the importance of conservation biology in the new century (Hedrick 2001; Brooks *et al.* 2006). A central controversy in conservation biology focuses on the levels of diversity in endangered species (Stebbins 1942; Drury 1974; Stebbins 1980; Gitzendanner & Soltis 2000). Comparisons of rare and common plant species find a significant trend for rare species to exhibit reduced genetic diversity (see review in Gitzendanner & Soltis 2000). However, several studies report rare taxa with levels of genetic variation that are similar to or greater than their common relatives (Lewis 1991; Amos 1999; Ellis *et al.* 2006), hence inferences regarding genetic variation in rare species must be made with caution. Because genetic diversity patterns can be shaped by many factors, such as mating system, demographic history, or life histories (Hamrick & Godt 1996; Charlesworth *et al.* 2003), it is advisable to compare rare species with relatives having similar life histories (Hamrick & Godt 1989; Nybom 2004) and mating system. Here, we use a set of molecular

markers to compare the endangered species *Boechera fecunda* with its widespread congener, *Boechera stricta*, both of which are predominantly inbreeding perennial herbs (Hamilton & Mitchell-Olds 1994; Schranz *et al.* 2005; Song *et al.* 2006).

The expanding information from genome projects and new applications of molecular techniques offer great promise for genetic studies of rare and endangered species. Combined with highly variable loci or genome-wide molecular markers, neutral variants used for conservation applications are becoming increasingly informative (Hedrick 2001). Genetic data from molecular markers or DNA sequences may elucidate past evolutionary events in endangered species. van Tienderen *et al.* (2002) suggested that the utility of markers in conservation biology will depend primarily on management goals, and discussed the risks of replacing genome-wide anonymous markers with a small sample of putatively important genes. In addition, recently founded populations are often genetically impoverished, making it challenging to find sufficient polymorphism for genetic analysis in rare species (Davies *et al.* 1999). Highly variable genetic markers, such as microsatellite loci, allow the quantification of patterns which may not be apparent when using less polymorphic markers (Hedrick & Parker 1998; Hedrick 1999; Parker *et al.* 1999).

Correspondence: Thomas Mitchell-Olds, Fax: 919-660-7293; E-mail: tmo1@duke.edu



**Fig. 1** (a) Distribution of the sampled *Boechera fecunda* populations in Montana. Population identities correspond to names in Table 1. (b) Population structure of *B. fecunda*, showing the two divergent clusters identified by STRUCTURE analysis.

However, it has been suggested that the application of highly variable loci (or extremely large numbers of markers) must be used with caution because statistically significant differences might not reflect biologically important differences, or might give a different signal than other markers (Hedrick 1999). Thus, expanding neutral polymorphism surveys, especially with markers having different mutation rates and dynamics, may help resolve issues in conservation genetics (Kohn *et al.* 2006) and present an integrated picture of studied populations and species.

*Boechera fecunda* is a rare, predominantly diploid crucifer species, restricted to calc-silicate soil outcrops in western Montana, USA (Rollins 1993). It is largely self-fertilizing (Hamilton & Mitchell-Olds 1994), with only 21 existing populations isolated in two geographical groups separated by a distance of ~100 km (Hamilton & Mitchell-Olds 1994). This species is a candidate for listing as a Threatened or Endangered species by the US Fish and Wildlife Service (USDI-FWS 1993) and is considered Threatened in Montana (Lesica 1993). A previous study based on allozyme and sequence analyses found no clear population differentiation or information about population demographic history because of extremely low variation in allozymes and studied sequences (McKay *et al.* 2001). In this study, we analysed 13 microsatellite loci, together with 27 single-copy nuclear loci (SCNL) to address the following questions

for *B. fecunda*: (i) Is diversity in endangered species less than in a widespread congener? (ii) Are *B. fecunda* populations highly structured? (iii) Which populations are important for preservation?

## Materials and methods

### Sample collection

A total of 156 plants (10–20 seed families per population; one individual per seed family) were sampled from 10 *Boechera fecunda* nature populations representing the geographical distribution of this species in Montana (Table 1, Fig. 1a). Populations were typically sampled along their longest axis for approximately 100 m with a minimum distance of 1 m between samples. One hundred and fifty-three plants (three were polyploids, described below) were screened using 13 microsatellite loci to investigate microsatellite variation. In addition, 12 individuals (one individual per population except Birch Creek, from which three plants were sampled) were chosen for sequence analysis of 27 SCNL to estimate sequence polymorphism. In order to compare sequence diversity between *B. fecunda* and its common congener, 12 *Boechera stricta* accessions (one accession/population) were sampled, which represents the core species distribution range in western USA.

**Table 1** Description and sample size of 10 natural populations of *Boechera fecunda*

Ranges	Population ID	Location	Elevation (feet)	Longitude (N)	Latitude (W)	Sample size
West	BC	Birch Creek, MT	4700	46°22'	113°59'	14
	CG	Charleys Gulch, MT	5000	46°16'	114°00'	14
	SP	Spooner, MT	4350	46°25'	114°01'	14
East	JC	Jerry Creek, MT	5700	45°47'	112°54'	13
	DC	Dewey Cemetery, MT	5780	45°46'	112°51'	20
	QH	Quartz Hill, MT	8000	45°42'	112°54'	15
	VP	Vipond Park, MT	7200	45°41'	112°52'	18
	LG	Lime Gulch, MT	6200	45°24'	112°49'	20
	HK	Highwood Kiln, MT	7540	45°49'	112°28'	15
	MC	Moose Camp, MT	6800	45°46'	112°35'	10

### Nuclear loci and microsatellites

Initially, 12 individuals (one per population, except three from Birch Creek) were chosen for microsatellite primer selection. A total of 30 microsatellite loci were tested in this study (Clauss & Mitchell-Olds 2003; Dobes *et al.* 2004; Song *et al.* 2006). Finally, 13 polymorphic loci with clear polymerase chain reaction (PCR) products were selected for further analysis. Among these, 10 loci have also been assayed in *B. stricta* (Song *et al.* 2006), allowing comparison between the two species. Information from the 13 microsatellite loci, together with their inferred location in the *Arabidopsis thaliana* genome are shown in Table S1 (Supplementary material).

For sequence analysis, we screened 27 SCNL sampled across the genome (Table S2, Supplementary material). Twelve loci were from Schmid *et al.* (2005), 12 from Schranz *et al.* (2005) and three additional loci were designed for the current study. Primers were designed to amplify a 500–800-bp region using the PRIMER3 program ([http://frodo.wi.mit.edu/cgi-bin/primer3/primer3\\_www.cgi](http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi)) based on *Arabidopsis thaliana* sequence. The sequence data of 27 SCNL for the 12 *B. fecunda* plants were screened in this study, while the sequences of the same set of SCNL for the 12 *B. stricta* plants were obtained from B.-H. Song and T. Mitchell-Olds (in preparation).

### DNA extraction, genotyping and sequencing

Ten seeds per family were placed on wet filter paper in a Petri dish. After germination, one seedling was randomly chosen, transferred into a 7 × 7-cm pot and grown under short-day conditions in the growth chamber. After 2 months, one young leaf was collected for assessing genome size using flow cytometry (Partec CCA) as described in Schranz *et al.* (2005). Three families were excluded from analysis

because of polyploidy. DNA extraction, PCR amplification and allele calling for microsatellite loci were described in Song *et al.* (2006).

PCR amplification for nuclear loci was conducted in a GeneAmp PCR System 9700 (Applied Biosystems) programmed as follows: 5 min at 95 °C followed by 35 cycles of 1 min at 94 °C, 30 s at 50–55 °C, and 2 min at 72 °C, followed by a final extension for 6 min at 72 °C, and an infinite hold at 4 °C. All sequences were obtained on both strands by direct sequencing of PCR products on an ABI PRISM 3700 sequencer (Applied Biosystems) using the BigDye Terminator 2.0 kit (Applied Biosystems). In the case of heterozygous individuals, the corresponding fragment was cloned with the TOPO TA Cloning Kit for Sequencing (Invitrogen) and up to eight clones per individual were sequenced. All sequences from this article have been deposited into GenBank with accession nos. EF689898–EF690217.

### Data analysis

**Microsatellite analysis Diversity analysis.** For each microsatellite locus, genetic polymorphisms were assessed by calculating the observed number of alleles ( $n_a$ ), effective number of alleles ( $n_e$ ), the average observed heterozygosity ( $H_O$ ) and gene diversity ( $H_S$ ) across populations, and total gene diversity ( $H_T$ ). Wright's  $F_{IS}$  statistic was computed to infer mating system. For each population, genetic diversity was estimated across all loci using  $n_a$ ,  $n_e$ ,  $H_O$ ,  $H_S$ , percentage of polymorphic loci ( $P$ ), and allele richness ( $R_S$ ).  $P$ ,  $n_a$  and  $n_e$  were calculated using POPGENE 1.32 (Yeh & Boyle 1997), and other values were calculated using FSTAT version 2.9.3 (Goudet 2001). The level of genetic differentiation among populations was quantified using  $F_{ST}$  (Weir & Cockerham 1984) and  $R_{ST}$  (Slatkin 1995). Using FSTAT, we calculated  $F_{ST}$  and  $R_{ST}$ , and then conducted a

randomization test of pairwise  $F_{ST}$  differentiation with 1000 iterations (Goudet 2001). All these analyses were performed using 13 microsatellite loci.

In order to test whether *B. fecunda* has lower genetic diversity than its widespread congener, *B. stricta*, we compared these two species using data from 10 shared microsatellite loci (*B. stricta* data from Song *et al.* 2006). A permutation test implemented in FSTAT version 2.9.3 (Goudet 2001) was employed to test the difference in genetic diversity between the two close relatives.

**Population structure.** We used the program STRUCTURE (Pritchard *et al.* 2000), which implements a Bayesian clustering method to infer population structure, assign individuals to populations and identify migrants and admixed individuals. As recommended for inbreeding species, we sampled one allele at random from each heterozygous locus (Song *et al.* 2006; J. Pritchard, personal communication). In order to fit the assumption of marker independence, linkage disequilibrium was analysed for these 13 loci in each population with FSTAT. The number of populations (clusters),  $K$ , was set from 1 to 12. Each  $K$  was replicated five times for 100 000 iterations after a burn-in period of 50 000 without prior information on the population of origin of each sampled individual. We used the admixture model in which the fraction of ancestry from each cluster is estimated for each individual. For population comparisons, the fractions of ancestry were averaged over individuals within each of the 10 populations. Our structure analyses repeatedly identified two clusters, for which we analysed patterns of diversity based on  $R_S$ ,  $H_S$  and  $F_{IS}$  with FSTAT with 1000 permutations using data from all microsatellite loci. The principal coordinates analysis (PCA) is a multivariate technique that summarizes the major patterns of variation in multilocus data set. Here, we use PCA analysis implemented in the program GENALEX 6 (Peakall & Smouse 2006) to plot the relationship between distance matrix elements based on their first two principal coordinates. Spatial genetic structure was also investigated by testing for isolation by distance. Geographical distances between pairs of populations were calculated from linear distances based on latitude and longitude (<http://jan.ucc.nau.edu/cvm/latlongdist.html>). A Mantel test with 10 000 random permutations was performed between the matrix of pairwise genetic differentiation between populations ( $F_{ST}/(1 - F_{ST})$ ) (Rousset 1997), and the matrix of the natural logarithm of geographical distances. The analyses were performed with IBWVS 3.02 (Jensen *et al.* 2005) on the seven eastern populations.

**Bottleneck analysis.** To infer the population history of the 10 *B. fecunda* populations, we tested for equilibrium patterns of genetic variation using the BOTTLENECK version 1.2.02 program of Cornuet & Luikart (1996) and Piry *et al.* (1999).

This approach compares observed and expected gene diversities based on observed number of alleles under mutation–drift equilibrium to test for compatibility with an equilibrium model.

### Sequence analysis

Sequences were assembled with SEQMAN 5.0 (DNASTAR) and all variable sites were checked manually during the construction of a consensus sequence from two primers. The alignments were conducted using CLUSTALW (Thompson *et al.* 1994) in BIOEDIT (<http://www.mbio.ncsu.edu/BioEdit/page2.html>). The DNASP program version 4.10.7 (Rozas *et al.* 2003) was used for both intra- and interspecific analyses of nucleotide polymorphism. Nucleotide diversity of a multiple alignment was calculated as  $\pi$ , the average number of nucleotide substitutions per site between two sequences (Lynch & Crease 1990), and  $S$ , the number of segregating sites (Watterson 1975). The diversity comparison between western and eastern regions of *B. fecunda*, as well as statistical differences in nucleotide diversity between *B. fecunda* and its common congener, *B. stricta*, was examined by  $t$ -test in SYSTAT 10 (SPSS Inc. 2000). Finally, we inferred the effective population size for *B. fecunda* based on  $N_e = \theta_s/4\mu$  (where  $\theta_s$  is the level of silent nucleotide diversity), based on a previously estimated rate of synonymous nucleotide substitution,  $\mu = 1.5 \times 10^{-8}$  from Koch *et al.* (2000).

## Results

### Microsatellite variation in *Boechera fecunda*

The number of alleles and gene diversity at microsatellite loci is shown in Table 2. The total number of alleles per locus in our sample of 153 individuals ranged from 2 to 13, with a total of 92 alleles scored over the 13 loci. The overall gene diversities ( $H_T$ ) varied widely among loci; from 0.05 to 0.80, with an average of 0.46. The average gene diversity ( $H_S$ ) over 13 loci was 0.20, with a range of 0.05–0.41. Diversity estimates varied among populations, with gene diversity ( $H_S$ ) ranging from 0.03 to 0.56, observed heterozygosity ( $H_O$ ) ranging from 0 to 0.27, and allelic richness ( $R_S$ ) ranging from 1.08 to 4.06 (Table 3). The  $F_{IS}$  values for all populations ranged from 0.52 to 1.0, with average value of 0.82. Genetic differentiation among populations was measured with several statistics (Table 2). The mean  $F_{ST}$  and  $R_{ST}$  were 0.57 and 0.51, respectively, indicating that more than half of the total genetic variation resided among populations under both infinite allele and stepwise-mutation models. Across all loci, most of the comparisons among populations showed significant differentiation, even after Bonferroni corrections for multiple tests (Table S3, Supplementary material).

**Table 2** Comparison of the genetic diversity found at 13 microsatellite loci in *Boechera fecunda*

Locus	$n_a$	$n_e$	$H_O$	$H_S$	$H_T$	$F_{ST}$	$R_{ST}$
ADH1	4	1.2	0.03	0.10	0.17	0.42	0.53
Bd266	11	4.8	0.10	0.41	0.80	0.49	0.63
BD495	12	2.2	0.03	0.24	0.55	0.57	0.33
BF3	13	1.8	0.04	0.17	0.48	0.65	0.62
BF9	8	1.7	0.05	0.13	0.43	0.69	0.52
BF18	2	1.5	0.04	0.10	0.35	0.71	0.74
BF20	7	2.2	0.03	0.40	0.57	0.30	0.29
c8	6	2.1	0.09	0.18	0.50	0.65	0.15
d3	5	2.2	0.04	0.11	0.52	0.79	0.91
e9	7	2.9	0.05	0.23	0.68	0.66	0.94
H5	3	1.0	0.01	0.05	0.05	0.10	0.14
ICE3	12	3.8	0.11	0.39	0.74	0.48	0.39
GC67	2	1.2	0.05	0.09	0.15	0.44	0.47
over all	7	2.2	0.05	0.20	0.46	0.57	0.51

Number of alleles over all populations ( $n_a$ ); effect allele number ( $n_e$ ); observed heterozygosity ( $H_O$ ); average gene diversity within populations ( $H_S$ ); overall gene diversity ( $H_T$ ); among population differentiation ( $F_{ST}$ ); among population differentiation ( $R_{ST}$ ).

**Table 3** Genetic variability estimates from populations of *Boechera fecunda*

POP ID	$P$	$n_a$	$n_e$	$H_O$	$H_S$	$R_S$	$F_{IS}$
BC	92.3	4.38	2.82	0.269	0.56	4.06	0.52
CG	92.3	3.31	2.40	0.102	0.54	3.21	0.81
DC	46.2	1.92	1.40	0.043	0.20	1.83	0.78
HK	38.3	1.62	1.26	0.005	0.15	1.59	0.97
JC	7.7	1.08	1.05	0.006	0.03	1.08	0.81
LG	15.4	1.15	1.11	0.008	0.07	1.15	0.89
MC	7.7	1.15	1.04	0.000	0.03	1.15	1.00
QH	23.1	1.23	1.12	0.005	0.08	1.23	0.94
SP	61.5	1.77	1.40	0.038	0.24	1.76	0.84
VP	30.8	1.31	1.13	0.030	0.09	1.30	0.66
Mean	41.5	1.90	1.47	0.051	0.20	1.83	0.82

Per cent of polymorphic loci ( $P$ ); observed allele number ( $n_a$ ); effective allele number ( $n_e$ ); observed heterozygosity ( $H_O$ ); gene diversity ( $H_S$ ); allelic richness ( $R_S$ ); inbreeding coefficient ( $F_{IS}$ ).

*Sequence diversity in B. fecunda*

Sequence diversity estimates for the 16 nuclear loci in *B. fecunda* are shown in Table 5. (No polymorphism in the other 11 loci was observed.) Polymorphic sites per locus ranged from 1 to 8, with average of 3. Diversity per locus ranged from  $\pi = 0.0004$ – $0.004$  with a mean of 0.0016.

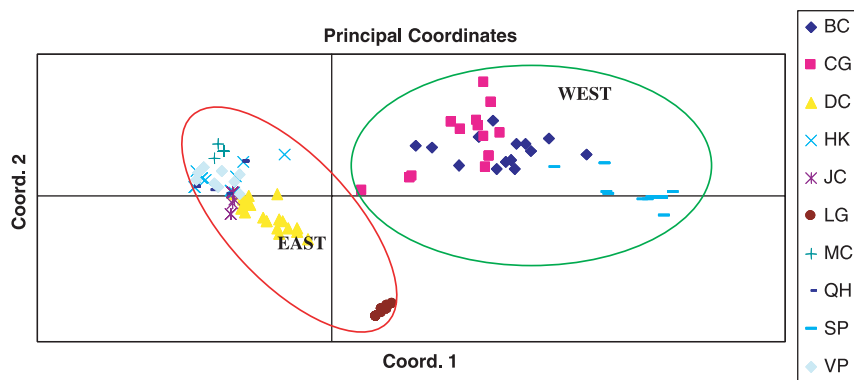
*Diversity comparison between B. fecunda and B. stricta*

Gene diversity for the 10 microsatellite loci shared between these two species ranged from 0.32 to 0.79 with mean value of  $0.55 \pm 0.14$  in *B. fecunda* and from 0.30 to 0.91, with mean value of  $0.66 \pm 0.20$  in *B. stricta*. We found no significant differences between these two species for gene diversity  $H_S$  ( $P = 0.29$ ) or allelic richness  $R_S$  ( $P = 0.5$ ) (Fig. 4a). Finally,

*B. stricta* showed significant higher  $F_{IS}$  than *B. fecunda* ( $F_{IS-Bs} = 0.95$ ;  $F_{IS-Bf} = 0.76$ ,  $P = 0.001$ ). The comparison of sequence diversity at 27 SCNL found that *B. stricta* showed slightly higher gene diversity than *B. fecunda*, but this difference was not significant (Fig. 4b:  $\pi_{Bs} = 0.0015$ ,  $\pi_{Bf} = 0.00095$ ,  $P = 0.2$ ).

*Population structure and demographic history in B. fecunda*

The *B. fecunda* populations were highly differentiated in terms of both microsatellite ( $F_{ST} = 0.57$  for 10 populations,  $F_{ST} = 0.34$  between eastern and western groups) data and sequence data ( $G_{ST} = 0.33$  between western and eastern groups). For the 13 microsatellite loci used for structure analysis, no tests for linkage disequilibrium were significant



**Fig. 2** Principle coordinate analysis based on 13 microsatellite loci among the 10 *Boechera fecunda* populations.

Source of variation	d.f.	Variance components	Percentage of variation	P value
Among regions	1	4.677	32%	$P < 0.001$
Among populations/regions	8	5.445	38%	$P < 0.001$
Within populations	143	4.439	30%	$P < 0.001$
Total	152	14.561		

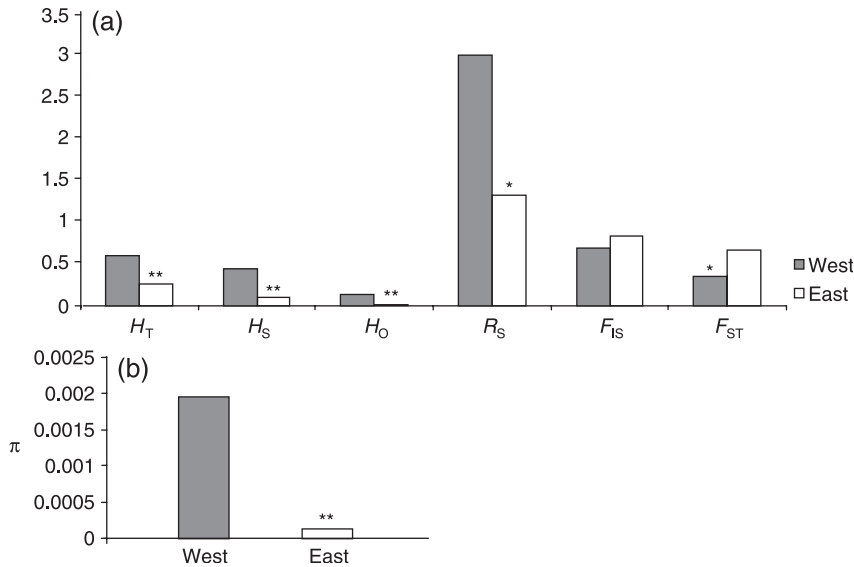
**Table 4** Analysis of molecular variance (AMOVA) for 10 *Boechera fecunda* populations

after Bonferroni correction, indicating that these loci are appropriate for STRUCTURE analysis. The inferred number of population clusters,  $K$ , was based on the protocols suggested by Pritchard and colleagues (<http://pritch.bsd.uchicago.edu/structure.html>; Pritchard *et al.* 2000). Based on these criteria,  $K = 2$  is the smallest value that captures the major structure in the data. Also, with  $K = 2$ , we consistently got the same population structure from five independent runs. The two clusters detected in the STRUCTURE analysis correspond to the eastern and western geographical groups (Fig. 1b). The PCA and STRUCTURE analyses also showed comparable patterns (Figs 1b and 2), as did distance-based analysis (not shown). The analysis of molecular variance (AMOVA) implemented in GENALEX confirmed significant differentiation between the western and eastern clusters (32% of total variation occurs between western and eastern clusters; Table 4). The Mantel test found no significant isolation by distance within the eastern cluster ( $r = 0.43$ ,  $P = 0.15$ , 10 000 randomization; Fig. S1, Supplementary material).

As described above, the population structure analysis based on molecular markers showed genetic differentiation between western and eastern groups, which was concordant with the distribution of *B. fecunda* populations, which are isolated in two geographical and elevational groups separated by a distance of 100 km (Lesica 1993). Interestingly, the two groups showed different levels of genetic variation. The eastern group showed significantly lower genetic diversity than the western group (Fig. 3a:  $R_{SE} = 1.33$ ,  $R_{SW} = 3.01$ ,  $P = 0.012$ ;  $H_{SE} = 0.10$ ,  $H_{SW} = 0.45$ ,

$P = 0.002$ ). In addition, the eastern group showed higher genetic differentiation among populations than the western group (Fig. 3a:  $F_{ST-e} = 0.67$ ,  $F_{ST-w} = 0.36$ ,  $P = 0.027$ ) but no significant difference in  $F_{IS}$  ( $F_{IS-e} = 0.84$ ,  $F_{IS-w} = 0.70$ ,  $P = 0.53$ ). After investigating the alleles in each group, we found the western group has more private alleles than the eastern group (55 vs. 13), and these two groups share 23 alleles. Finally, patterns of sequence variation are consistent with microsatellite data: the western group showed significantly higher gene diversity than the eastern group (Fig. 3b:  $\pi_w = 0.002$ ,  $\pi_e = 0.0001$ ,  $P < 0.001$ ).

Bottleneck analysis showed contrasting evolutionary histories between western and eastern groups. In two out of three sampled populations from the western group, the BOTTLENECK comparison between observed and expected heterozygosities under a mutation–drift model showed no significant deviation from mutation–drift equilibrium under the infinite allele model (IAM), the stepwise-mutation model (SMM) and the intermediate two-phase model (TPM). In contrast, most populations (five of seven) in the eastern group showed significant deviation from mutation–drift equilibrium under these models. Finally, estimates of  $\theta_s$  in the western and eastern groups were 0.00217 and 0.00013, respectively. From this, we can roughly estimate  $N_e$  as  $\sim 3 \times 10^4$  and  $\sim 2 \times 10^3$  in the west and east, respectively. Because of the high levels of uncertainty inherent in estimates of  $\mu$ , as well as deviation of these *B. fecunda* populations from models of ideal populations, these estimates of  $N_e$  must be regarded with considerable caution.



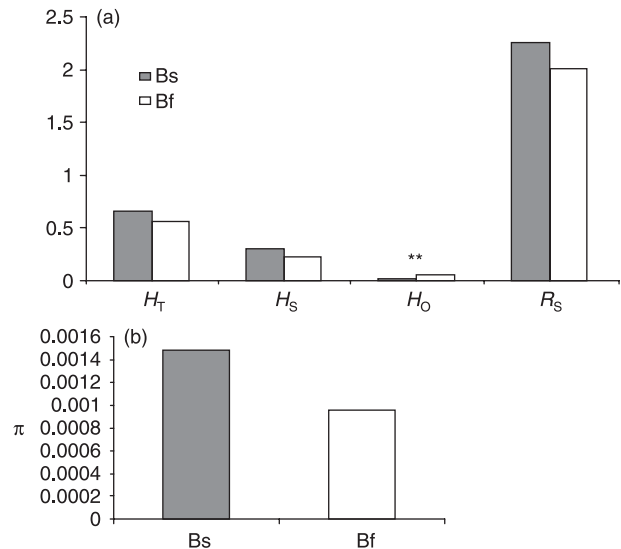
**Fig. 3** (a) Comparison of microsatellite diversity between the western and eastern groups in *Boechera fecunda* (\*\* $P < 0.01$ , \* $P < 0.1$ ). (b) Comparison of sequence diversity between the western and eastern groups (\*\* $P < 0.01$ ).

## Discussion

### Rare does not mean depauperate in *Boechera fecunda*

Despite the general expectation of reduced genetic variation in rare species, in comparison to its widespread congener, *Boechera fecunda* does not show reduced diversity at either the population or the species level (Fig. 4a). Interestingly, *B. fecunda* showed significantly higher observed gene diversity than *Boechera stricta* ( $H_O$ , see Fig. 4a). At the sequence level, however, *B. fecunda* and *B. stricta* have similar levels of nucleotide polymorphism (Fig. 4b:  $\pi_{Bf} = 0.0010$ ;  $\pi_{Bs} = 0.0015$ ,  $P = 0.2$ ). Although there is an overall tendency for rare species to have lower levels of genetic variation than their widespread relatives (Gitzendanner & Soltis 2000), there are, nevertheless, important exceptions to this trend (Vogelmann & Gastony 1987; Ranker 1994; Young & Brown 1996; Gitzendanner & Soltis 2000; Ellis *et al.* 2006). For example, Ellis *et al.* (2006) studied an endangered sunflower species and found that it displayed higher levels of genetic diversity than its common congener. Thus, it is clear that the expectation of reduced genetic diversity in rare species is not always borne out.

Factors influencing the maintenance of genetic variation in rare species include large population size (Ellstrand & Elam 1993), antiquity of populations (Lewis 1991), outcrossing mating system (Ranker 1994) and hybridization (Rieseberg 1997). For *B. fecunda*, the observed levels of genetic variation may reflect relatively large, long-term populations on persistent, calc-silicate soils with little competition. Ellstrand & Elam (1993) argued that rare species with large local populations may exhibit high levels of genetic variation. Based on estimates of nucleotide polymorphism and mutation rates, we estimate the effective



**Fig. 4** (a) Comparison of genetic diversity between *Boechera stricta* (Bs) and *Boechera fecunda* (Bf) for the microsatellite data (\* $P < 0.1$ ). (b) Comparison of genetic diversity ( $\pi$ ) between *B. stricta* (Bs) and *B. fecunda* (Bf) for the sequence data.

population size of *B. fecunda* as in the order of  $3 \times 10^4$ , which is substantially lower than the census population size of  $\sim 10^5$  (Lesica 1993). Such differences have also been observed in other species (Frankham 1995). Patterns of microsatellite variation also were compatible with mutation–drift equilibrium in western populations. In contrast, in the eastern group of populations, levels of genetic variation are much lower, and estimated effective population size is only  $\sim 2000$ .

Several statistics showed high levels of genetic differentiation among populations, which can contribute to

**Table 5** Polymorphism comparison between western and eastern groups of *Boechera fecunda* based on sequences from 16 single-copy nuclear loci.

Abbreviation	Locus	West + East			West			East			$G_{ST}$
		#Seq	S	$\pi$	#Seq	S	$\pi$	#Seq	S	$\pi$	
Abi3	At3g24650	12	1	0.0004	5	1	0.0007	7	0	0.0000	0.16
Ang	At1g01510	11	1	0.0006	5	1	0.0012	6	0	0.0000	0.15
AtG1	At3g51790	12	1	0.0006	5	1	0.0011	7	0	0.0000	0.16
AtIII8x5	At3g21470	13	3	0.0033	6	0	0.0000	7	0	0.0000	1.00
Cip7	At4g27430	11	5	0.0013	5	5	0.0028	6	0	0.0000	0.15
Dex1	At3g09090	12	4	0.0014	5	2	0.0008	7	0	0.0000	0.52
Dpa1	At5g67100	13	3	0.0020	6	2	0.0014	7	1	0.0008	0.30
Fnr	At2g15620	13	4	0.0040	6	4	0.0026	7	0	0.0000	0.68
GOLM19	noncoding	11	2	0.0017	4	2	0.0026	7	0	0.0000	0.47
GOLM23	noncoding	14	3	0.0011	7	3	0.0021	7	0	0.0000	0.12
GOLM54	noncoding	10	3	0.0008	3	3	0.0028	7	0	0.0000	0.20
GOLM66	At3g49900	13	1	0.0004	6	2	0.0015	7	0	0.0000	0.06
GOLM73	At4g12485	11	8	0.0038	4	8	0.0072	7	2	0.0008	0.16
GOLM80	At4g13020	13	2	0.0016	6	2	0.0016	7	0	0.0000	0.49
MPS35	At2g32950	12	3	0.0020	5	2	0.0017	7	1	0.0004	0.37
Nuo	At5g08530	12	1	0.0008	5	1	0.0009	7	0	0.0000	0.31
Average		12	3	0.0016	5	2	0.0020	7	0.3	0.0001	0.33

maintenance of genetic variation at the species level. Overall estimates of differentiation among populations ranged from 51% ( $R_{ST}$ , Table 2) to 70% (AMOVA, Table 4). On a larger scale, about 33% of molecular variation differentiates the eastern and western population groups (Tables 4 and 5). Evidently, long-term differentiation among *B. fecunda* populations contributes to maintenance of genetic variation at the species level. Indeed, a rare species capable of surviving in diverse habitats may maintain higher levels of genetic diversity than even a widespread species confined to one uniform habitat (Gitzendanner & Soltis 2000).

In comparison to *B. stricta*, the lower selfing rate in *B. fecunda* may facilitate the maintenance of genetic diversity in this rare species. Based on 10 microsatellite loci screened in both species, the  $F_{IS}$  value is 0.76 for all sampled populations, which is significantly smaller than that of *B. stricta* ( $F_{IS} = 0.94$ ), and  $F_{IS}$  for the western *B. fecunda* group is even lower ( $F_{IS} = 0.70$ ). Lastly, although hybridization is common among some *Boechera* taxa (Schranz *et al.* 2005), so far, we find no evidence of hybridization involving *B. fecunda*.

#### Conclusions and conservation implications

Most programs in conservation genetics seek to preserve significant pools of heritable variation, while simultaneously preventing the fixation of deleterious alleles (e.g. Hedrick 1992; Avise & Hamrick 1995). Comparisons of both microsatellite variation and sequence polymorphism revealed a marked difference between western and eastern groups (Fig. 3). The western group exhibited greater variation

than the eastern group (Fig. 3). Variation in most western populations approximated mutation–drift equilibrium, while eastern populations showed nonequilibrium patterns suggesting historical population bottlenecks within the relatively recent time frame reflected by these highly mutable markers. The high differentiation among eastern populations ( $F_{ST} = 0.67$ ) and lack of isolation by distance (Fig S1) also suggest founder effects during their establishment. In contrast, higher variation in the western region suggests a longer history of *B. fecunda* in this area, where this species may have originated.

Maintaining adaptive genetic variation is an important conservation strategy, which requires identification of populations that show local adaptation, as well as differences in ecology (Kohn *et al.* 2006). Despite controversies regarding the relative importance of within- vs. between-population genetic diversity in conservation (Notter 1999; Reist-Marti *et al.* 2003; Toro & Caballero 2005), it is clear that management priorities should emphasize preservation of adaptive diversity within the whole species (Crandall *et al.* 2000). STRUCTURE analyses based on microsatellite variation clearly showed that western and eastern populations clustered in two different groups (Fig. 1b) and AMOVA analysis showed 32% ( $P < 0.001$ ) of total variation occurred between these two groups. In addition, high differentiation between the western and the eastern groups in *B. fecunda* was detected based on 13 microsatellite loci ( $F_{ST} = 0.34$ ). Despite differing mutation rates between microsatellite and sequence polymorphisms, 16 nuclear loci gave very similar estimates of population differentiation ( $G_{ST} = 0.33$ ). Management of rare species will be more successful if

variation in neutral markers is compared to genetic variation in morphological and physiological traits that are expected to confer adaptation to habitat heterogeneity (McKay *et al.* 2001). The patterns of diversity and differentiation in the two regions of *B. fecunda* shows that both molecular markers (this study) and putatively adaptive traits (McKay *et al.* 2001) are consistently differentiated between the two regions. Thus, both groups should be managed and conserved separately, and transplantation is not recommended.

### Acknowledgements

We thank S. Gebauer-Jung for help with data format transformation, K. Weniger for genome size measurements, D. Schnabelrauch and Jana Sitte for technical assistance. We also thank M. J. Clauss, the subject editor D. McCauley and three anonymous reviewers for helpful comments. Partial support was provided by Duke University and the Max-Planck Society.

### References

- Amos W (1999) Two problems with the measurement of genetic diversity and genetic distance. In: *Genetics and the Extinction of Species* (eds Landweber LF, Dobson AP), pp. 75–100. Princeton University Press, Princeton, New Jersey.
- Avise J, Hamrick JL (1995) *Conservation Genetics: Case Histories from Nature*. Chapman & Hall, New York.
- Brooks TM, Mittermeier RA, da Fonseca GAB *et al.* (2006) Global biodiversity conservation priorities. *Science*, **313**, 58–61.
- Charlesworth B, Charlesworth D, Barton NH (2003) The effects of genetic and geographic structure on neutral variation. *Annual Review of Ecology, Evolution, and Systematics*, **34**, 99–125.
- Clauss MJ, Mitchell-Olds T (2003) Population genetics of tandem trypsin inhibitor genes in *Arabidopsis* species with contrasting ecology and life history. *Molecular Ecology*, **12**, 1287–1299.
- Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–2014.
- Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology & Evolution*, **15**, 290–295.
- Davies N, Villablanca FX, Roderick GK (1999) Determining the source of individuals: multilocus genotyping in nonequilibrium population genetics. *Trends in Ecology & Evolution*, **14**, 17–21.
- Dobes C, Mitchell-Olds T, Koch MA (2004) Intraspecific diversification in North American *Boechera stricta* (= *Arabidopsis drummondii*), *Boechera xdivaricarpa*, and *Boechera holboellii* (Brassicaceae) inferred from nuclear and chloroplast molecular markers – an integrative approach. *American Journal of Botany*, **91**, 2087–2101.
- Drury WH (1974) Rare species. *Biological Conservation*, **6**, 162–169.
- Ellis JR, Pashley CH, Burke JM, McCauley DE (2006) High genetic diversity in a rare and endangered sunflower as compared to a common congener. *Molecular Ecology*, **15**, 2345–2355.
- Ellstrand NC, Elam DR (1993) Population genetic consequences of small population-size – implications for plant conservation. *Annual Review of Ecology and Systematics*, **24**, 217–242.
- Frankham R (1995) Effective population size/adult population size ratios in wildlife: a review. *Genetical Research*, **66**, 95–107.
- Gitzendanner MA, Soltis PS (2000) Patterns of genetic variation in rare and widespread plant congeners. *American Journal of Botany*, **87**, 783–792.
- Goudet J (2001) *FSTAT, a Program to Estimate and Test Gene Diversities and Fixation Indices* (v. 2.9. 3), <http://www2.unil.ch/popgen/softwares/fstat.htm>.
- Hamilton MB, Mitchell-Olds T (1994) The mating system and relative performance of selfed and outcrossed progeny in *Arabidopsis fecunda* (Brassicaceae). *American Journal of Botany*, **81**, 1252–1256.
- Hamrick JL, Godt MJW (1989) *Allozyme Diversity in Plant Species*. Sinauer & Associates, Sunderland, Massachusetts.
- Hamrick JL, Godt MJW (1996) Conservation genetics of endemic plant species. *Conservation Genetics – Case Histories from Nature*, 281–304.
- Hedrick P (1992) Population genetics – shooting the RAPDs. *Nature*, **355**, 679–680.
- Hedrick PW (1999) Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution*, **53**, 313–318.
- Hedrick PW (2001) Conservation genetics: where are we now? *Trends in Ecology & Evolution*, **16**, 629–636.
- Hedrick PW, Parker KM (1998) MHC variation in the endangered Gila topminnow. *Evolution*, **52**, 194–199.
- Jensen JL, Bohonak AJ, Kelley ST (2005) Isolation by distance, web service. v.3.0.1. *BMC Genetics*, **6**, 13. Available from URL: <http://ibdws.sdsu.edu/>.
- Koch MA, Haubold B, Mitchell-Olds T (2000) Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabidopsis*, and related genera (Brassicaceae). *Molecular Biology and Evolution*, **17**, 1483–1498.
- Kohn MH, Murphy WJ, Ostrander EA, Wayne RK (2006) Genomics and conservation genetics. *Trends in Ecology & Evolution*, **21**, 629–637.
- Lesica P (1993) *Report on the conservation status of Arabis fecunda, a candidate threatened species*. Montana Natural Heritage Program, Helena, MT.
- Lewis PO (1991) Allozyme variation in the rare Gulf Coast endemic *Polygonella macrophylla* Small (Polygonaceae). *Plant Species Biology*, **6**, 1–10.
- Lynch M, Crease TJ (1990) The analysis of population survey data on DNA sequence variation. *Molecular Biology and Evolution*, **7**, 377–394.
- McKay JK, Bishop JG, Lin JZ *et al.* (2001) Local adaptation across a climatic gradient despite small effective population size in the rare sapphire rockcress. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **268**, 1715–1721.
- Notter DR (1999) The importance of genetic diversity in livestock populations of the future. *Journal of Animal Science*, **77**, 61–69.
- Nybom H (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology*, **13**, 1143–1155.
- Parker KM, Sheffer RJ, Hedrick PW (1999) Molecular variation and evolutionarily significant units in the endangered Gila topminnow. *Conservation Biology*, **13**, 108–116.
- Peakall ROD, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.
- Piry S, Luikart G, Cornuet J-M (1999) BOTTLENECK: a computer program for detecting recent reductions in effective population size from allele frequency data. *Journal of Heredity*, **90**, 502–503.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

- Ranker TA (1994) Evolution of high genetic variability in the rare Hawaiian fern *Adenophorus periens* and implications for conservation management. *Biological Conservation*, **70**, 19–24.
- Reist-Marti SB, Simianer H, Gibson J, Hanotte O, Rege JEO (2003) Weitzman's approach and conservation of breed diversity: an application to African cattle breeds. *Conservation Biology*, **17**, 1299–1311.
- Rieseberg LH (1997) Hybrid origins of plant species. *Annual Review of Ecology and Systematics*, **28**, 359–389.
- Rollins RC (1993) *The Cruciferae of Continental North America*. Stanford University Press, Stanford, California.
- Rousset F (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.
- Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DNASP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.
- Schmid KJ, Ramos-Onsins S, Ringys-Beckstein H, Weisshaar B, Mitchell-Olds T (2005) A multilocus sequence survey in *Arabidopsis thaliana* reveals a genome-wide departure from a neutral model of DNA sequence polymorphism. *Genetics*, **169**, 1601–1615.
- Schranz ME, Dobes C, Koch MA, Mitchell-Olds T (2005) Sexual reproduction, hybridization, apomixis, and polyploidization in the genus *Boechera* (Brassicaceae). *American Journal of Botany*, **92**, 1797–1810.
- Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457–462.
- Song BH, Clauss MJ, Pepper A, Mitchell-Olds T (2006) Geographic patterns of microsatellite variation in *Boechera stricta*, a close relative of *Arabidopsis*. *Molecular Ecology*, **15**, 357–369.
- Stebbins GL (1942) The genetic approach to problems of rare and endemic species. *Madroño*, **6**, 241–258.
- Stebbins GL (1980) Rarity of plant species: a synthetic viewpoint. *Rhodora*, **82**, 77–86.
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**, 4673–4680.
- van Tienderen PH, de Haan AA, van der Linden CG, Vosman B (2002) Biodiversity assessment using markers for ecologically important traits. *Trends in Ecology & Evolution*, **17**, 577–582.
- Toro MA, Caballero A (2005) Characterization and conservation of genetic diversity in subdivided populations. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **360**, 1367–1378.
- USDI-FWS (1993) 50 CFR Part 17. Endangered and threatened wildlife and plants; review of plant taxa for listing as endangered or threatened species; notice of review. *Federal Register*, **58**, 51144–51190.
- Vogelmann JE, Gastony GJ (1987) Electrophoretic enzyme analysis of North American and Eastern Asian populations of *Agastache* sect. *Agastache* (Labiatae). *American Journal of Botany*, **74**, 385–393.
- Watterson GA (1975) On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, **7**, 256–276.
- Weir B, Cockerham C (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.
- Yeh F, Boyle T (1997) Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belgian Journal of Botany*, **129**, 157.
- Young AG, Brown AHD (1996) Comparative population genetic structure of the rare woodland shrub *Daviesia suaveolens* and its common congener *D. Mimosoides*. *Conservation Biology*, **10**, 1220–1228.

---

Dr Bao-Hua Song is working on natural genetic variation and ecologically important adaptation using molecular genetic and genomic approaches. Professor Thomas Mitchell-Olds studies the functional basis of evolutionary forces influencing ecologically important genetic variation.

---

## Supplementary material

The following supplementary material is available for this article:

**Fig. S1** Plot of pairwise genetic distance (Nei) against the geographical distance between populations from eastern group of *Boechera fecunda*.

**Table S1** Name, location, repeat unit, range in allele size, and primer sequences for 13 microsatellite loci.

**Table S2** Information for the 27 single-copy nuclear loci analyzed in this study.

**Table S3** Pairwise  $F_{ST}$  among populations of *Boechera fecunda*. \* indicates significance at the 5% level with Bonferroni correction.

This material is available as part of the online article from:  
<http://www.blackwell-synergy.com/doi/abs/10.1111/j.1365-294X.2007.03500.x>  
 (This link will take you to the article abstract).

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.