Genet Resour Crop Evol (2007) 54:837–848
DOI 10.1007/s10722-006-9167-4

ORIGINAL PAPER

# Capturing genetic diversity of wild populations for *ex situ* conservation: Texas wild rice (*Zizania texana*) as a model

**Christopher M. Richards · Michael F. Antolin · Ann Reilley · Jackie Poole · Christina Walters**

Received: 19 November 2005 / Accepted: 13 April 2006 / Published online: 7 December 2006


**Abstract** Genebanks complement other conservation programs because they preserve genetic diversity needed for future breeding and restoration. We evaluated efficiency of capturing genetic diversity, using endangered *Zizania texana* (Texas wild rice) as a model for plants with recalcitrant seeds. This perennial aquatic grass is restricted to 4 km of the San Marcos River in Texas. An early conservation collection included plants from stands throughout the river, based on the assumption stands would be unique genotypes. Using microsatellite markers, we found that genetic diversity was concentrated in five of 15 large, demographically stable stands; 96 stands smaller than 2 $m^2$ contributed no unique alleles. High heterozygosity and few duplicate genotypes suggested that sexual reproduction occurs more often than presumed. Simulations of stratified sampling of large stands captured all alleles in only 45 individuals, while random sampling along the river captured much less diversity. The early conservation collection captured as much diversity as expected from random sampling. Texas wild rice stands resemble a mainland-island metapopulation; our analyses suggest that stratified sampling maximizes genetic diversity for this population dynamic. Demographic and genetic information is important for validating the design of efficient *ex situ* collections and guiding *in situ* conservation.

**Key words** Stratified sampling · Allelic richness · Conservation genetics · Genetic structure · *Zizania texana*

C. M. Richards (✉) · A. Reilley · C. Walters
USDA/ARS National Center for Genetic Resources Preservation, 1111 South Mason Street, Fort Collins, CO 80521, USA
e-mail: Chris.Richards@colostate.edu

M. F. Antolin
Department of Biology, Colorado State University, Fort Collins, CO 80523-1878, USA

J. Poole
Texas Parks & Wildlife, 3000 I-35 South, Austin, TX 78704, USA

## Introduction

Assembling *ex situ* collections from wild populations is an important component of wider conservation goals (Schoen and Brown 2001). Conservation collections in botanical gardens, zoos, or genebanks are designed to safeguard genetic diversity, enhance current population size, keep useful traits accessible, and re-establish populations in restoration projects. As such, collection strategies must efficiently capture a high percentage of the extant genetic diversity, as measured directly by variation in molecular or quantitative traits (Brown and Hardner 2000) or

indirectly through a geographically stratified sampling scheme (Shands 1991). Without direct measurements of genetic variation of phenotypic traits or molecular markers, indirect methods often provide excellent first approximations of genetic variation when they take into account the target species' life history characteristics, reproductive system, and long-term population changes (Brown and Briggs 1991; Schoen and Brown 1991). For instance, if populations are abundant but are comprised mainly of a few genotypes that arise by asexual reproduction, total genetic diversity could be found in relatively small samples collected from each population.

Efficiency in collecting is important not just because risk of eminent extinction, but also because maintaining large, redundant collections may divert resources from other conservation priorities such as habitat protection. Seedbanks can provide an important conservation tool because the viability of hundreds or thousands of individuals can be maintained for decades without regeneration (Walters 2004). However, recalcitrant seeds, which cannot be stored under the standard cold and dry conditions used for most seeds, require specialized protocols to ensure their viability in genebanks. Recalcitrant seeds are common in many of the riparian and tropical species (Hong et al. 1998) that are currently under threat because of harvesting or habitat loss. Developing cryogenic methods for storage of recalcitrant seeds requires significant investment of time and resources. It is thus critical to design efficient collections that optimize the capture of genetic diversity from wild.

Population genetic analysis of the aquatic Texas wild rice (*Zizania texana* Hitchcock) provides a case study for targeted sampling for a species with a narrow geographic range: four km of the San Marcos River in south central Texas. Texas wild rice seeds are recalcitrant and do not survive in conventional storage. Monitoring of stands since the 1960s has revealed little or no flowering in the river and implied that the primary mode of reproduction of this perennial grass is asexual (Emery 1977; Power 1996). These observations suggested one approach for *ex situ* conservation was to dig plants from the river and grow them in pots in an artificial aquatic habitat. A conservation collection of 48 Texas wild rice plants from different locations throughout its range was established from 1986 to 1999 at a fish hatchery in San Marcos, Texas in raceways with a continuous flow-through of water from the river. However, this collection of plants was both costly to maintain and fraught with inherent risks: by 1999, mortality reduced the conservation collection to 35 plants (for which we have DNA samples and genotypes), and the collection continues to decline.

It is possible to establish genebanks using seeds of *Z. texana*. We have developed cryogenic methods to preserve Texas wild rice seeds and demonstrated that these methods do not reduce genetic diversity in stored samples for a set of variable microsatellite loci (Walters et al. 2002; Richards et al. 2004). Further, growth conditions that induce flowering and seed production in controlled environments are known (Power and Fonteyn 1995; C. Walters unpublished data). Although cryogenic storage protocols for Texas wild rice are time consuming and costly, developing an *ex situ* collection seems particularly useful as a complementary tool for other conservation projects for Texas wild rice because human induced pressures on this species' habitat, especially by increased diversions of water, has accelerated in the last few decades (Vaughan 1986).

In this study we examined genetic diversity and population structure of stands of Texas wild rice along its entire range in the San Marcos River, using six highly variable microsatellite markers. Sampling over three years was designed to sample among and within stands, test whether reproduction within the river is primarily asexual, and evaluate temporal changes in genetic variation and structure. We follow this analysis with a sampling simulation to identify collecting strategies that most efficiently capture the genetic diversity of Texas wild rice for *ex-situ* genebanks.

## Natural history of *Z. texana*

Texas wild rice is an aquatic perennial grass endemic to a four km stretch of the San Marcos River in Hays CO, TX (Emery 1977) (Fig. 1). The river's clear, spring-fed headwaters emerge from

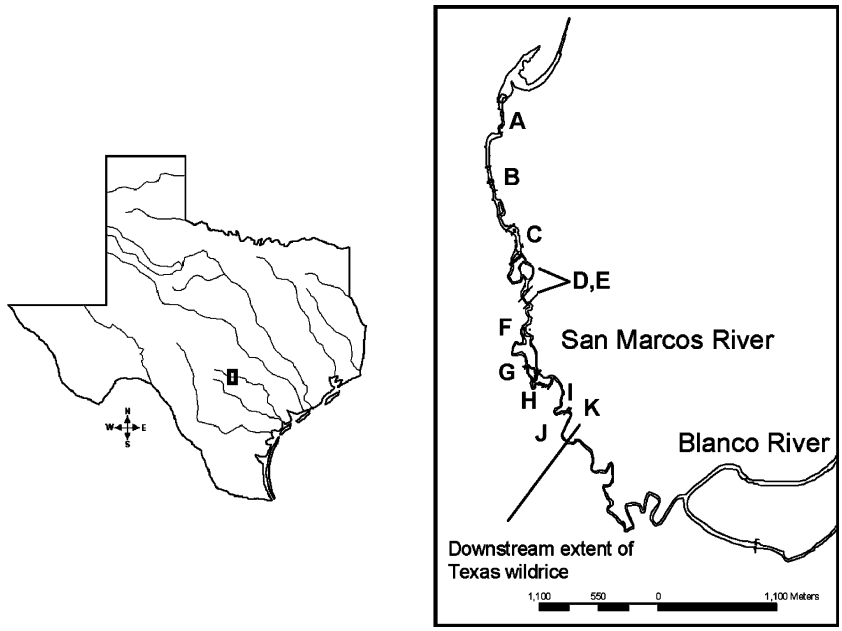


**Fig. 1** Map of stands of *Z. texana* in the San Marcos River in Texas, showing river sections in annual surveys (left)

the Edwards Aquifer through the permeable karst in the Balcones Fault of Texas at an average flow rate of 157 cu ft/s, and the river flows south to the confluence of the Blanco River (Terrell et al. 1978). Below this confluence, where the water becomes heavily silted, no *Z. texana* have been recorded. The habitat requirements of *Z. texana* appear to be confined to the upper reaches of the river, which varies from 10 to 25 m wide and is up to four m deep in some channels. Texas wild rice grows mostly in large vegetative stands that are submerged in mid-river in relatively swift and shallow water (less than one m) and rooted in sandy or gravely bottoms (Poole and Bowles 1999).

The spatial distribution of stands has been mapped annually for 15 years by the Texas Parks and Wildlife Department to monitor the abundance of this species and to describe its habitat requirements (Poole unpublished data). Occupied reaches of the river were designated as segments (Emery, Southwest Texas State University unpublished data 1978) that correspond to physical features of the river such as bridges, dams and other structures (Fig. 1). Stands, which are defined as continuous masses of *Z. texana* plants, were delineated by polygons that enclosed the irregularly-shaped stands, mapped using surveying equipment, and recorded in GIS for yearly comparisons. Estimates of the percent coverage within polygons ($m^2$) were used as relative measures of stand size. Two prominent features emerge from this long term study. First, while sizes of stands varied from 0.01 to 335 $m^2$ (see Table 3 below), most stands were small (mean size is 8.5 $m^2$). Second, the distribution and local abundance of this species was highly dynamic (Poole unpublished data 2003). Over the years, the number of stands fluctuated markedly along the river (mean = 206, SD = 45). Small stands experienced higher rates of turnover (extinction and colonization) than larger stands, with several of the largest stands being stable throughout the survey period (Poole unpublished data 2004). In addition, the census revealed that the spatial arrangement of stands was more clumped in the upper reaches of the river, in segments A and B, and progressively became more dispersed downstream of segment F.

Texas wild rice rarely flowers in the river (Emery 1977; Power 1996) but it is capable of flowering when grown in fish-rearing raceways located at the National Fish Hatchery and Technology Center in San Marcos, TX, in an old fish hatchery on the campus of Texas State University-San Marcos (Power 1996), and in greenhouses at Colorado State University (Walters unpublished

data 2004). In artificial conditions, plants display bi-modal phenology with early flowering starting in April and later flowering starting in July. Plants reproduce sexually by culms and panicles that emerge above the water and bear seed. Submerged plants can reproduce asexually by producing stolons that arise from the plant's base and form adventitious roots. These can be seen trailing downstream and may provide a means of increasing stand size by producing ramets (asexual shoots) from the mother plant. Pollen dispersal is by wind, which typical of grasses.

## Materials and methods

### Sample collection

We collected leaf tissues for DNA fingerprinting from plants within the river over three separate years. The collection strategy at first aimed to estimate overall genetic diversity within the river in 1998 and 1999. Individuals collected within stands were taken one m apart except when mapping clonal spread (see below). While the collection of multiple individuals per stand was a priority, single individuals from small stands (generally less than two $m^2$) were also collected in 1999 to capture more broadly the diversity in these ephemeral stands. Individual (singleton) samples were collected from all parts of the river from segments A through K (Fig. 1). This was followed in 2002 with more intensive sampling of three large stands (Fig. 2) to estimate the extent of asexual reproduction by obtaining leaf tissues from plants spaced less than one m from each other along defined transects. In addition, we genotyped all individuals growing in the early conservation collection housed at the National Fish Hatchery and Technology Center. All leaf tissues were stored on ice until shipment to Fort Collins, CO. Upon arrival, they were rinsed in deionized water, split into 100 mg aliquots and placed in storage at –18°C until DNA extraction.

### Genotyping

Extraction of DNA from leaf tissues used an anionic exchange column protocol in Qiagen DNeasy kits (Qiagen, Valencia, CA). Genetic Identification Services (Chatsworth, CA) developed microsatellite-enriched libraries from *Z. texana* genomic DNA and sequenced several clones that contained microsatellite markers. Primer sequences and repeat motifs for the markers are in Table 1. Microsatellites were amplified in a total volume of 10 μl containing 1.5 units Promega Taq DNA Polymerase and 1X Reaction Buffer (Promega US, Madison, WI), 2.5 mM $MgCl_2$, 200 μM of each dNTP, 0.1 pM of each primer (forward primers labeled with IRDye$^{TM}$ 700 or IRDye$^{TM}$ 800, MWG Biotech, Inc., High Point, NC), 0.5 ng of unlabelled reverse primers (Qiagen Operon, Alameda, CA), and 2.5 ng of template DNA. Conditions used to amplify these loci are the same as those used previously (Richards et al. 2004), products were resolved in a 6.5% KB$^{Plus}$ acrylamide Gel Matrix (LI-COR) on a LI-COR 4200 sequencer following manufacturer's recommendations (LI-COR Biosciences, Lincoln, NE). Scanned TIFF images of each gel were imported into SAGA GT$^{TM}$ software (LI-COR Biosciences, Lincoln, NE) for fragment sizing and allele scoring.

### Data analysis

Putative ramets within the data set were identified as exact genotypic matches at all six markers (or five if markers had missing values). The expected probability of identity was calculated within each stand as the probability of obtaining duplicate multilocus genotypes from the same stand, given the total probability of sampling any possible genotype twice (Hedrick 2004). Significance was assessed at an error rate of $P < 0.05$ with critical values adjusted for multiple tests by the Bonferroni correction (Sokal and Rohlf 1995) for the eight stands (or pairs of stands) from which duplicate genotypes were collected. Ramets within the 2002 collection data were mapped to specific transect sites in population A1, B1a and C5. Ramets of a single clone identified in this way were reduced to one representative genotype within a stand for population genetic analyses. Ramets identified among stands or among years within a stand were reduced to one genotype in each stand or in each year.

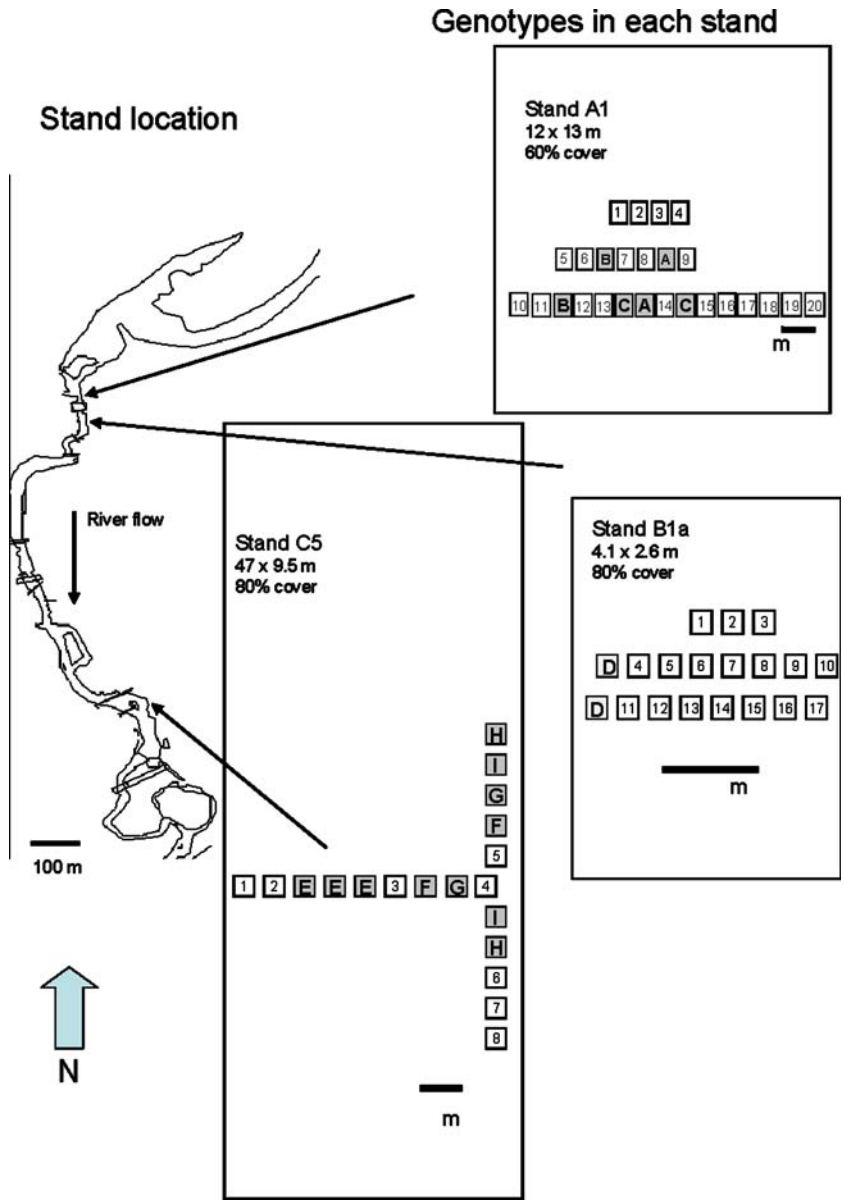


**Fig. 2** Detailed maps of genotypes within three stands (right) collected from the river (left). Each stand map (detailed schematic on right) is oriented with respect to flow direction as the river map (on left). The boxes enclosing each stand map represents the polygon that defines the stand area. Numbered squares represent unique multilocus genotypes, lettered squares represent duplicate genotypes presumed to be ramets of the same individual

Descriptive statistics, including variation between stands ($F_{st}$) and inbreeding within stands ($f$), were estimated from genotypic data using GDA (Lewis and Zaykin 2001) and FSTAT (Goudet 1995). Tests of linkage disequilibrium using GDA (Fisher's exact test) failed to identify markers in linkage disequilibrium.

## Sampling simulation

We evaluated sampling strategies that capture genetic diversity (measured by our microsatellite markers) in the river by simulating collections of individuals. The algorithm was first outlined by Schoen and Brown (1995) and was implemented in the program MSTRAT (Gouesnard et al. 2001) which was designed to assemble core collections maximally diverse for allelic richness and validate their diversity. MSTRAT uses an iterative maximization algorithm to assemble a sample group of size $r$ maximized for genotypes containing different alleles from a collection of size $N$. Briefly, the algorithm accomplishes this by selecting $r$ individual genotypes at random from a pool of $N$.

**Table 1** Primer sequences used in this study

| Marker Name | Primer sequence 5′–>3′ | Marker size range (bp) | Repeat motif |
|---|---|---|---|
| Zt1* | GCAAATCTCCTGTCTTTTTCT<br>GTTTAGCCAGCTCCCAATGTA | 259–267 | TAGA |
| Zt13 | ACGTCGTCGTCTTCCTCC<br>GCATATAATTCCGCGTGAAC | 206–250 | TC |
| Zt18* | CACCATGTCCTGCAATTC<br>TGCACTAGCTCCCTGAAA | 98–114 | TC |
| Zt21* | CTAGCTTGTTCAGACAAATGTT<br>GACTCTGCTGCATCATATCA | 170–198 | TC |
| Zt22 | CAACCCCAGAAAAACTAAATC<br>TCCAATCTCTCCACCTACAA | 200–230 | AG |
| Zt23 | GGACGTTGACATTTTCACA<br>GGATCAGTAAATCCAAATCTGT | 250–284 | AG |

* Markers described in Richards et al. (2004)

All ($r$–1) sets are assessed for allelic richness and the ($r$–1) combination with the greatest number alleles across all loci is retained. To this set, a new individual is drawn at random from the ($N$–$r$) remaining individuals in the collection and the last two steps are iterated until a convergence criterion is met (where no change in allelic richness is achieved after 30 iterations. Optimization is achieved by initiating multiple runs for each value of $r$. The relationship $r$ and allelic richness can be examined by using a plotting feature of MSTRAT. Sampling efficiency measured in this way can be interpreted as a saturation curve that identifies the minimum number of individuals necessary to capture all of the alleles in collection of genotypes. As the value of $r$ reaches $N$, the fraction of allelic richness necessarily goes to 1 but the form of the curve can vary depending on the diversity of the collection, $N$. In principle, if each individual contributes a novel allele, then the gain in allelic richness would increase linearly. In practice, however, most collections contain some inherent redundancy (due to skewed allelic frequencies) and therefore allelic richness increases asymptotically with increasing $r$ values. This implies that at some value of $r$ ($1 < r < N$) there is no additional gain in allelic diversity. The inflection point near the asymptote, therefore, represents the smallest value of $r$ for which maximal allelic richness is achieved.

We use this sampling approach not as a way of assembling a core subset for conservation *per se*, but as a way to examine the pattern of sampling for maximal allelic diversity, i.e.: what stands in the river are preferentially sampled when assembling a maximally diverse subset? In order to evaluate this, we contrasted two sampling strategies of the data set for comparison: (1) random sampling and (2) maximized sampling (M strategy) that selected only individuals that increased the number of alleles in the sample.

The first step of the analysis determined the value of $r$ needed to capture a desired proportion of genetic diversity using the M strategy. Under both maximized and random sampling, we ran 5 separate simulations for a range of values of $r$ from 10 to $N/2$. We tested differences between the two sampling strategies by a Kolmogorov–Smirnov goodness of fit procedure, under the null hypothesis that the distributions differed at one point or more (Sokal and Rohlf 1995).

The second step selected $r$ individual genotypes, determined from step one, for a maximally diverse sample (M samples). We assembled ten equally diverse M samples and ranked individual genotypes by the number of times they were selected among the ten replicates. We identified the location and characteristics of stands contributing to these M samples.

## Results

We genotyped 471 plants collected from the river, plus another 35 from the conservation collection at the fish hatchery. Of the river samples, 375 were in nested in 23 population samples from stands and 96 were collected as individual samples from small stands less than two $m^2$ in size.

Within the stands, we found identical multilocus genotypes in 17 groups (mainly pairs) of plants. Probability of identity within stands ranged from $8.24 \times 10^{-6}$ to $4.42 \times 10^{-3}$, when calculated using allele frequencies of the stands where the duplicates were found. These probabilities of identity fall below the Bonferroni-corrected critical probability of $P < 0.05$. Because of the low probability that identical genotypes could result from sexual reproduction, we interpreted the duplicates to be ramets of asexually propagating individuals The ramets were all within six of the stands (A1, B1a, B2, C5, C6, F12), and none of the 96 singleton samples collected from small stands matched any other genotypes in the river. One pair of identical genotypes was found in 1999 and 2002 in stand C5; another pair of identical genotypes was collected from two adjacent stands in section B in 1999. Some genotypes were found up to four times during the intensive sampling within stands in 2002. A map of genotypes collected at a fine scale within stands A1, B1a, and C5 in 2002 is shown in Fig. 2. While we found duplicate multilocus genotypes within each stand, duplicates were not the majority of the genotypes within the stands. Duplicate genotypes within a stand were close to each other, but not always adjacent (e.g., genotype E in population C5). Often, ramets were along the axis of water flow in the river (e.g., genotype D in stand B1a).

To estimate measures of diversity and population structure of the 23 samples, we removed duplicate genotypes from the data set, and reduced total sample size by 27 to 348 individuals. The genetic variation revealed by these six loci was high (Table 2), with an average gene diversity (Nei 1987) of 0.662 and a total of 70 different alleles sampled. Estimates of $F_{st}$ for individual markers among the 23 samples ranged from 0.093 to 0.288, with standard errors of 0.04 or less. Within sampling years, average genetic differentiation among stands was high (average $F_{st} > 0.102$), but was not significantly different among sampling years (all 95% confidence intervals overlap). The correlation between pairwise physical distances between stands and genetic distance ($F_{st}/[1\text{-}F_{st}]$) was only 0.17 and was not significant using a standard Mantel test. In addition, the pattern of allelic richness and differentiation between populations upstream vs. downstream showed no significant differences.

Within stands, average gene diversity was 0.507 and observed heterozygosity was 0.551 (Table 3). Sample size was correlated with stand size ($r = 0.615$), the number of alleles ($r = 0.771$), and gene diversity within the stand ($r = 0.598$) (Table 3, Fig. 3A). There was no evidence of inbreeding within stands, as none of the inbreeding coefficients ($f$) differed from zero (permutation tests in FSTAT, Goudet 1995). If anything, we detected a tendency toward an excess of heterozygotes in 17 of 23 samples, even after duplicate genotypes were removed (average $f = -0.091$). About 90% of the genetic diversity was found in stands where we collected more than 15 individuals; median area of these stands was 99 $m^2$ (Table 3). A greater contribution of the larger stands to allelic diversity is consistent with the observation that the 96 singleton samples did not increase overall allelic diversity. All unique alleles were found in samples from large stands (F12, C5 B7, B2).

## Sampling simulation

Simulations of sampling showed that capture of allelic diversity depended on sample size and whether sampling was random or maximized (Fig. 3B). At first, we only included the 348 genotypes identified from the large stands (i.e., not the 96 singleton genotypes collected from small stands). Random sampling did not capture all the alleles, even if samples were half the size of the total data set ($N = 174$) (Fig. 3B, open circles). By contrast, sampling using the maximization algorithm in MSTRAT captured all 70 alleles within a sample of only 45 individuals (Fig. 3B, solid circles). Using a Kolmogorov–Smirnov test, we found that the curves generated by random and maximized sampling differed (KS = 0.693, $P = 0.0029$). In a second simulation, we included the 96 singleton samples to test whether they increased the minimum number of individuals needed in a collection. The 96 additional genotypes did not change the saturation curves given in Fig. 3B because the small stands did not harbor unique alleles and most of the allelic richness of *Z. texana* resided in the larger stands.

**Table 2** Diversity measured within 23 samples of stands of *Z. texana* from 1998, 1999, and 2002

| Marker name | Number genotyped | Alleles per marker | Gene diversity ($H_e$) | Observed heterozygosity ($H_o$) | Overall (23 stands) | | 1998 (5 stands) | | 1999 (15 stands) | | 2002 (3 stands) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | $F_{st}$ | s.e. | $F_{st}$ | s.e. | $F_{st}$ | s.e. | $F_{st}$ | s.e. |
| Zt-1 | 346 | 4 | 0.252 | 0.243 | 0.093 | 0.029 | 0.018 | 0.034 | 0.121 | 0.039 | 0.029 | 0.005 |
| Zt-13 | 313 | 20 | 0.787 | 0.623 | 0.133 | 0.026 | 0.055 | 0.012 | 0.165 | 0.046 | 0.087 | 0.116 |
| Zt-18 | 298 | 11 | 0.713 | 0.715 | 0.155 | 0.039 | 0.101 | 0.084 | 0.119 | 0.022 | 0.547 | 0.270 |
| Zt-21 | 307 | 15 | 0.825 | 0.805 | 0.171 | 0.040 | 0.096 | 0.089 | 0.215 | 0.065 | 0.180 | 0.133 |
| Zt-22 | 306 | 7 | 0.713 | 0.523 | 0.288 | 0.038 | 0.255 | 0.192 | 0.256 | 0.040 | 0.422 | 0.407 |
| Zt-23 | 307 | 13 | 0.684 | 0.609 | 0.135 | 0.042 | 0.053 | 0.020 | 0.203 | 0.058 | 0.096 | 0.078 |
| Average | 312.8 | 11.7 | 0.662 | 0.586 | 0.171 | (0.13,0.22) | 0.102 | (0.05,0.17) | 0.188 | (0.15, 0.23) | 0.260 | (0.12,0.43) |

Standard errors (s.e.) of the between-stand variance of allele frequencies ($F_{st}$) for each marker were estimated by jackknifing over samples; confidence intervals (n parentheses) of estimates from all six markers were calculated by bootstrapping over markers

We compared our sampling simulations to the fish hatchery conservation collection. Allelic diversity of the 35 plants taken from the river is marked as a black square (Fig. 3B) and falls within a region consistent with random sampling of genotypes. The 35 individuals making up this collection captured only as much allelic diversity at these loci as a random sample of similar size.

Finally, sampling was simulated to generate an M sample of 45 individuals that maximized allelic diversity. The 45 individuals chosen for each of ten replicates consistently had rare alleles. Stands contributing individuals to the M samples are indicated by solid circles (five or more individuals) and half-filled circles (between one and four individuals); stands not contributing to the M samples are represented by open circles (Fig. 3A). Five stands contributed 5 or more individuals to M samples are represented by solid circles (Fig. 3A). These populations contributed 35 of the 45 individuals in the M sample and all were larger than 52 $m^2$ (median stand size was 149 $m^2$, Table 3).

## Discussion

The genetic diversity measured in *Z. texana* by this set of microsatellite markers is greater than one would predict in a species that appears to be entirely asexual in its native habitat. Stands were comprised of multiple genotypes that most likely arose from seed (Table 3), and because relatively few duplicate genotypes were detected, the presumption that stands arose predominantly from asexual reproduction must be rejected. Furthermore, the data in Tables 2 and 3 portray a population with high heterozygosity and allelic richness. Estimates of the within-stand inbreeding coefficient ($f \leq 0$, Table 3) suggest random mating or perhaps outcrossing, as has been reported in the closely related *Z. palustris* (Hayes et al. 1989). The spatial distribution of genotypes within stands implies that Texas wild rice undergoes rounds of sexual reproduction, with some spread within stands by asexual reproduction via rooted stolons. In this way heterozygosity can persist in this population even though sexual reproduction is episodic. Specific biological cues

**Table 3** Diversity measured within 23 samples of stands in 1998, 1999, and 2002

| Location | Year | Stand area ($m^2$) | Average number of plants genotyped | Number of alleles | Gene diversity ($H_e$) | Observed hetero-zygosity ($H_o$) | Inbreeding/ Outcrossing ($f$) | Number in core collection |
|---|---|---|---|---|---|---|---|---|
| A1 | 1998 | 20.5 | 9.3 | 22 | 0.507 | 0.556 | –0.103 | 1 |
| | 1999 | 33.5 | 5.2 | 12 | 0.428 | 0.45 | –0.059 | |
| | 2002 | 41 | 18.2 | 14 | 0.383 | 0.413 | –0.082 | |
| A3 | 1998 | 5.1 | 4.8 | 19 | 0.519 | 0.542 | –0.049 | 1 |
| | 1999 | 3 | 4 | 16 | 0.542 | 0.625 | –0.184 | |
| B1a | 1999 | 4.8 | 4.3 | 16 | 0.456 | 0.542 | –0.226 | |
| | 2002 | 10.2 | 15.2 | 12 | 0.421 | 0.6 | –0.449 | |
| B1f | 1999 | 27.7 | 3.8 | 10 | 0.292 | 0.333 | –0.171 | |
| B1h | 1999 | 99.8 | 4 | 13 | 0.464 | 0.625 | –0.429 | |
| B2 | 1998 | 52.4 | 22.8 | 45 | 0.668 | 0.61 | 0.089 | 10 |
| | 1999 | 43.1 | 5.8 | 14 | 0.461 | 0.489 | –0.069 | |
| B4b | 1999 | 10.8 | 4 | 14 | 0.536 | 0.75 | –0.5 | 1 |
| B7 | 1999 | 99.3 | 21.8 | 36 | 0.675 | 0.546 | 0.194 | 10 |
| C2 | 1999 | 83.3 | 14.5 | 25 | 0.533 | 0.629 | –0.19 | |
| C5 | 1998 | 284.5 | 28.2 | 26 | 0.586 | 0.639 | –0.092 | |
| | 1999 | 335.6 | 36.8 | 25 | 0.518 | 0.639 | –0.239 | 5 |
| | 2002 | 333.4 | 9 | 14 | 0.467 | 0.718 | –0.589 | |
| C6 | 1999 | 24.5 | 4 | 12 | 0.435 | 0.292 | 0.364 | 1 |
| F11 | 1999 | 38.6 | 4 | 15 | 0.393 | 0.25 | 0.4 | 1 |
| F12 | 1998 | 154.5 | 34.7 | 42 | 0.656 | 0.646 | 0.015 | 5 |
| | 1999 | 149.3 | 37.7 | 39 | 0.627 | 0.564 | 0.102 | 5 |
| F6 | 1999 | 84.1 | 11 | 25 | 0.593 | 0.651 | –0.103 | 4 |
| H2 | 1999 | 3.0 | 9.7 | 20 | 0.511 | 0.554 | –0.09 | 1 |
| Average | | 84.4 | 13.6 | 21.1 | 0.507 | 0.551 | –0.091 | |
| Conservation collection | | | 35 | 41 | 0.582 | 0.519 | | |

The early conservation collection is included for comparisons only. Column on far right shows collecting locations of plants in the early conservation collection

required for flowering in the river are not fully known, but early findings indicate that water depth or flow rate is important (Power 1997; Walters unpublished data).

Stands in the river were genetically differentiated (Table 2), and genetic structure did not significantly change in the three yearly samples in 1998, 1999, and 2002. The extent of differentiation among stands did not correlate with geographic distance. Apparently, idiosyncratic movement of seeds, rooted stolons, or pollen creates admixtures of genotypes in the river that cannot be explained by distance alone.

Much of the allelic diversity of *Z. texana* in these six SSR loci is concentrated in samples from large demographically stable stands (Fig. 3); samples from small stands, which are often ephemeral and do not contain unique SSR alleles. Consolidation of allelic diversity in the larger stands is consistent with unidirectional gene flow perhaps even a mainland-island structure (Hastings and Harrison 1994), with several large subpopulations acting as the source for colonizing seed or rooted stolons. This type of population structure is also consistent with the highly skewed distribution of stand areas: most of the stands were small (i.e., the 96 individual samples from unique stands from 1999) and even the largest stands show a skewed size distribution (Table 3).

Sampling models that assume random mating but no prior knowledge of population structure have been designed with the goal of capturing alleles with frequencies above 5% across as many populations and environments as possible (reviewed by Lockwood et al. in press Horticultural Reviews). These models recommend distributed sampling across multiple locations and populations of different size, especially when populations are

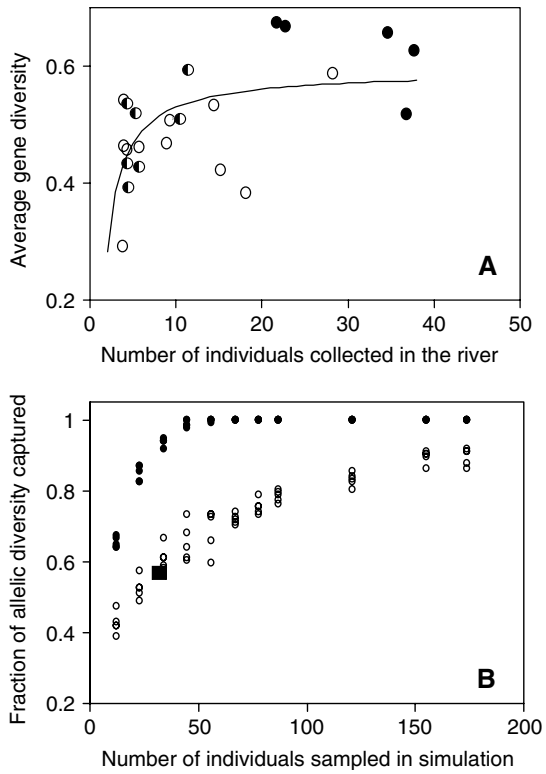


**Fig. 3** The relationship between number of individuals in a sample and the amount of genetic diversity captured. In the upper panel (**A**), solid circles represent stands that were identified through simulation as contributing five or more individuals to core collections; half-filled circles represent stands contributing one to four individuals; open circles represent stands that did not contribute to the core sample of 45 individuals. In the lower panel (**B**), the proportion of total genetic diversity captured by simulations that drew from the 348 individuals collected from stands listed in Table 3. One sampling algorithm drew samples at random (open circles), another that drew samples to maximize allelic richness (solid circles). The black square represents the genetic diversity measured from the 35 individuals of the conservation collection at the National Fish Hatchery and Technology Center in San Marcos, TX

genetically differentiated (Neel and Cummings 2003; Jin et al. 2003). On the other hand, it is often possible to identify populations that constitute sampling foci for maximizing allelic richness. Previous work has purposed that molecular surveys are useful in identifying genetic hot spots for diversity and that these initial surveys may be helpful in prioritizing collection efforts (Schoen and Brown 1991; Schoen and Brown 1993). With sufficient knowledge of how genetic diversity is distributed across populations or habitats, effective collections can be quite small (Ceska et al. 1997; Jin et al. 2003; Caujape-Castells and Pedrola-Monfort 2004; McKhann et al. 2004). While the MSTRAT sampling pattern suggest focusing on large populations, it is critical that cross validation of this diversity is undertaken with quantitative trait variation using an approach which controls for genotype by environment interactions by planting in a common environment. Allelic diversity at microsatellite loci is not indicative of adaptive differentiation *per se*, but the differentiation and diversity at these marker loci are indicative of dispersal. Therefore, basing conservation collection on microsatellite diversity may not be appropriate. In this species, with its unusually small and uniform spring-fed habitat, understanding the patterns of gene dispersal may be particularly useful when obvious selection for ecotypic variation among stands appears to be minimal.

It is generally agreed that all available ecological and molecular marker data should be used when identifying populations of particular conservation importance (Schemske et al. 1994; Petit et al. 1998; Crandall et al. 2000; Moritz 2002). Neutral variation is useful in estimating differentiation and diversity, but it is only one of several approaches that help to define adaptive markers do not define adaptive differentiation differences in ecologically important traits. A combination of reproductive, genetic and demographic information provides a clearer guide for maximizing diversity within conservation collections (e.g. Young et al. 2002; Cavers et al. 2004). Given the restricted and homogeneous aquatic habitat of *Z. texana*, knowledge of the history of stands in the San Marcos River combined with our population genetic analysis was a key to identifying optimum sampling of genetic diversity for *ex situ* preservation.

**Acknowledgments** We acknowledge Kathyrn Kennedy (formerly of US Fish and Wildlife Service, now Director, Center for Plant Conservation, St. Louis, MO) for inspiring research on this species and initiating agreements with CW and MFA that helped fund the project. P. Power helped with field collections, Dr. D. Touchell aided in sample collection and marker development. This manuscript was improved with comments from Drs. L. Panella and K. McCluskey. This work was funded by an ARS postdoctoral fellowship grant to CW in 1998, and through

Cooperative Agreements between US Fish and Wildlife Service and CSU (1448-20181-99-J810) in 1999 and USDA-ARS (58-5402-1-0153) in 2000 and between USDA-ARS and CSU (58-5402-9-113) in 2000.

## References

Brown AHD, Briggs JD (1991) Sampling strategies for genetic variation in *ex situ* collections of endangered plant species. In: Falk DA, Holsinger KE (eds) Genetics and conservation of rare plants. Oxford University Press, New York, pp 99–119

Brown AHD, Hardner CM (2000) Sampling the gene pools of forest trees for *ex situ* conservation. In: Young A, Boyle TJB, Boshier D (eds) Forest conservation genetics: principles and practice. CABI Publishing, Wallingford, pp 185–196

Caujape-Castells J, Pedrola-Monfort J (2004) Designing *ex-situ* conservation strategies through the assessment of neutral genetic markers: application to the endangered *Androcymbium gramineum*. Conserv Genet 5:131–144

Cavers S, Navarro C, Lowe AJ (2004) Targeting genetic resource conservation in widespread species: a case study of *Cedrela odorata* L. Forest Ecol Manage 197:285–294

Ceska JF, Affolter JM, Hamrick JL (1997) Developing a sampling strategy for *Baptisia arachnifera* based on allozyme diversity. Conserv Biol 11:1133–1139

Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. Trends Ecol Evol 15:290–295

Emery WHP (1977) Current status of Texas wild rice. The Southwestern Nat 22:393–394

Goudet J (1995). FSTAT (vers. 1.2): A computer program to calculate F-statistics. J Heredity 86:485–486

Gouesnard B, Bataillon TM, Decoux G, Rozale C, Schoen DJ, David JL (2001) MSTRAT: An algorithm for building germplasm core collections by maximizing allelic or phenotypic richness. J Heredity 92:93–94

Hastings A, Harrison S (1994) Metapopulation dynamics and genetics. Ann Rev Ecol Syst 25:167–188

Hayes PM, Stucker RE, Wandrey GG (1989) The domestication of American wild-rice (*Zizania palustris, Poaceae*). Econ Bot 43:203–214

Hedrick PW (2004) Genetics of populations, 3rd edn. Jones and Bartlett Publishers, Sudbury, MA

Hong TD, Linington SH, Ellis RH (1998) Compendium of Information on Seed Storage Behavior, two volumes. Royal Botanic Gardens, Kew, UK

Jin Y, Zhang W-J, Fu D-X, Lu B-R (2003) Sampling strategy within a wild soybean population based on its genetic variation detected by ISSR markers. Acta Bot Sinica 45:995–1002

Lewis PO, Zaykin D (2001). GDA User's manual. Department of Ecology and Evolutionary Biology, University of Connecticut. http://www.lewis.eeb.uconn.edu/lewishome

McKhann HI, Camilleri C, Berard A, Bataillon T, David JL, Reboud X, Le Corre V, Caloustian C, Gut IG, Brunel D (2004) Nested core collections maximizing genetic diversity in *Arabidopsis thaliana*. Plant J 38:193–202

Moritz C (2002) Strategies to protect biological diversity in the evolutionary processes that sustain it. Syst Biol 51:238–254

Neel MC, Cummings MP (2003) Effectiveness of conservation targets in capturing genetic diversity. Conserv Biol 17:219–229

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, NY

Petit RJ, El Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. Conserv Biol 12:844–855

Poole J, Bowles DE (1999) Habitat characterization of Texas wild-rice (*Zizania texana* Hitchcock), an endangered aquatic macrophyte from the San Marcos River, TX, USA. Aquat Conserv: Marshes Freshwater Ecosyst 9:291–302

Power P, Fonteyn PJ (1995) Effects of oxygen concentration and substrate on seed germination and seedling growth of *Zizania texana* (Texas wildrice). The Southwestern Nat 40:1–4

Power P (1996) Reintroduction of Texas wildrice (*Zizania texana*) in Spring Lake: some important environmental and biotic considerations. In: Proceedings of the Second Conference on Southwestern rare and endangered plants on 1995 September 11–14 in Flagstaff, AZ. General Technical Report RM-GTR-283. USDA Forest Service Rocky Mountain Forest and Range Experiment Station, Fort Collins CO

Power P (1997) Moisture, seeds and reproductive failure in Texas wildrice (*Zizania texana*). The Southwestern Nat 42:435–440

Richards CM, Reilley A, Touchell DH, Antolin MF, Walters C (2004) Microsatellite primers for Texas wild rice (*Zizania texana*), and a preliminary test of the impact of cryogenic storage on allele frequency at these loci. Conserv Genet 5:853–859

Schemske DW, Husband BC, Ruckelshaus MH, Goodwillie C, Parker IM, Bishop JM (1994) Evaluating approaches to the conservation of endangered plants. Ecology 75:584–606

Schoen DJ, Brown AHD (1991) Intraspecific variation in population gene diversity and effective population size correlates with the mating system in plants. Proc Natl Acad Sci USA 88:4494–4497

Schoen DJ, Brown AHD (1993) Conservation of allelic richness in wild crop relatives is aided by assessment of genetic markers. Proc Natl Acad Sci USA 90:10623–10627

Schoen DJ, Brown AHD (1995) Maximising genetic diversity in core collections of wild relatives of crop species. In: Hodgkin T, Brown AHD, van Hintum ThJL, Morales EAV (eds) Core Collections of Plant Genetic Resources. John Wiley & Sons, Chichester, UK, pp. 55–76

Schoen DJ, Brown AHD (2001) The conservation of wild plant species in seed banks. BioScience 51:960–966

Shands HL (1991) Complementarity of *in situ* and *ex situ* germplasm conservation from the standpoint of the future user. Israel J Bot 40:521–528

Sokal RR, Rohlf FJ (1995) Biometry, 3rd edn. W.H. Freeman and Company, New York

Terrell EE, Emery WHP, Beaty HE (1978) Observations on *Zizania texana* (Texas Wild Rice), an endangered species. Bull Torrey Botanical Club 105:50–57

Vaughan JE (1986) Populations and autecological assessment of *Zizania texana* Hitchc. (*Poaceae*) in the San Marcos River. M.Sc. Thesis, Southwest Texas State University, San Marcos, TX

Walters C, Touchell DH, Power P, Wesley-Smith J, Antolin MF (2002) A cryopreservation protocol for embryos of the endangered species *Zizania texana*. CryoLetters 23:291–298

Walters C (2004) Principles for preserving germplasm in genebanks. In: Guerrant E, Havens K, Maunder M (eds) Ex situ plant conservation: supporting species survival in the wild. Island Press, Covelo, CA, pp 442–453

Young AG, Hill JH, Murray BG, Peakall R (2002) Breeding system, genetic diversity and clonal structure in the sub-alpine forb *Rutidosis leioepis* F. Muell. (Asteraceae). Biol Conserv 106:71–78