# Genetic diversity and structure of a rare endemic cactus and assessment of its genetic relationship with a more common congener

Niraj Rayamajhi[1], Matthew Olson[2], and Jyotsna Sharma[1*]

[1]Department of Plant and Soil Science, Texas Tech University, Lubbock, TX 79409 USA, [2]Department of Biology, Texas Tech University, Lubbock, TX 79409 USA; *jyotsna.Sharma@ttu.edu

## Abstract

Endemic plant species with narrow geographic distributions and disjunct populations are prone to loss of genetic diversity especially if they are obligate outcrossing taxa. Simultaneously, delineating clear species boundaries is important in targeted conservation and management efforts. The federally endangered cactus, *Sclerocactus brevihamatus* subsp. *tobuschii*, has a parapatric relationship with *Sclerocactus brevihamatus* subsp. *brevihamatus* but genetic distance between the two taxa is unknown. We: 1) developed taxon-specific polymorphic microsatellites, 2) assessed genetic diversity within and among nine populations of *Sclerocactus brevihamatus* subsp. *tobuschii* (SBT) and within a single population of *Sclerocactus brevihamatus* subsp. *brevihamatus* (SBB), and 3) investigated the genetic relationship between the two subspecies. We obtained allelic data from 255 individuals across both taxa using seven self-developed microsatellites. Within-population genetic diversity of subsp. *tobuschii* was moderate to moderately high (mean $H_e$ = 0.59; mean $H_o$ = 0.37). Indirect estimate of inbreeding ($F_{is}$) ranged from 0.15 to 0.63 for subsp. *tobuschii* and was relatively high for subsp. *brevihamatus* ($F_{is}$ = 0.47). Genetic differentiation among populations of subsp. *tobuschii* was low based on results from $F_{st}$ (0.08) and AMOVA ($\Phi_{PT}$ = 0.07); however, $D_{est}$ showed elevated genetic differentiation among those populations. A Mantel test and spatial autocorrelation analysis suggested no significant genetic and spatial correlation among nine populations or within each population of subsp. *tobuschii*. The Infinite Allele Model detected bottleneck ($p < 0.01$) in six populations of subsp. *tobuschii* while SMM and TPM did not. STRUCTURE analysis (K=3) and both rooted and unrooted NJ dendrograms suggested three distinct clusters across the ten populations of the two sub-taxa within *Sclerocactus brevihamatus*. Our results suggest that subsp. *tobuschii* is sufficiently differentiated from subsp. *brevihamatus* to retain its current taxonomic status.

**Figure 1.** A plant of *Sclerocactus brevihamatus* subsp. *tobuschii* (SBT) in anthesis.

**Figure 2.** A plant of *Sclerocactus brevihamatus* subsp. *brevihamatus* (SBB) in anthesis.

## Objectives

1. To develop novel and polymorphic nuclear microsatellite marker loci for *Sclerocactus brevihamatus* subsp. *tobuschii*.
2. To assess genetic diversity within and among nine populations of *Sclerocactus brevihamatus* subsp. *tobuschii* and within a single population of *Sclerocactus brevihamatus* subsp. *brevihamatus.*
3. To investigate genetic relationship between *Sclerocactus brevihamatus* subsp. *tobuschii* and *Sclerocactus brevihamatus* subsp. *brevihamatus.*

## Materials and Methods

Sample (spine) collection and DNA extraction
(225 samples from 9 populations of SBT and 30 samples from 1 population of SBB)

↓

Microsatellite, or nuclear Simple Sequence Repeat (nSSR) development
(to identify species-specific nSSR motifs in the genome of SBT)

↓

Design and optimization of nSSR primers
(to select polymorphic and reproducible nSSR loci)

↓

Polymerase Chain Reaction (PCR) using selected primers
(to amplify targeted nSSR regions in DNAs from both SBT and SBB)

↓

Genotyping of each individual DNA at seven nSSR loci
(to obtain allelic data for each locus based on fragment length)

↓

Analysis of allelic data
(population genetics parameters, fitness parameters, and genetic relations between the subspecies)

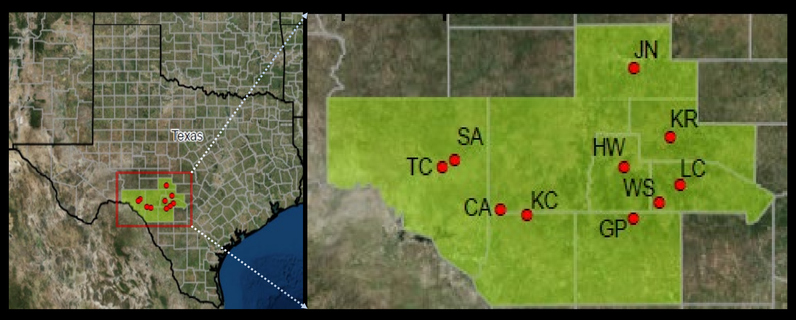


**Figure** 3. A map showing the locations of nine sampled populations (CA, GP, HW, KC, KR, LC, SA, WS, and JN) of *Sclerocactus brevihamatus* subsp. *tobuschii* and one population (TC) of *Sclerocactus brevihamatus* subsp. *brevihamatus* across eight counties of Texas.

## Results

**Table 1:** Genetic diversity statistics for seven microsatellite loci used to obtain allelic data from 9 populations of *Sclerocactus brevihamatus* subsp. *tobuschii* including number of alleles ($N_a$), mean expected heterozygosity ($H_e$), and mean observed heterozygosity ($H_o$).

| Locus | Allele size range (base pairs) | $N_a$ | $H_e$ | $H_o$ |
|---|---|---|---|---|
| NRJS5 | 347-399 | 9 | 0.66 | 0.07 |
| NRJS10 | 120-212 | 19 | 0.74 | 0.74 |
| NRJS11 | 152-168 | 5 | 0.47 | 0.26 |
| NRJS29 | 175-227 | 13 | 0.74 | 0.61 |
| NRJS32 | 146-166 | 5 | 0.60 | 0.41 |
| NRJS35 | 162-182 | 5 | 0.55 | 0.30 |
| NRJS39 | 179-191 | 4 | 0.34 | 0.20 |

**Table 2.** Population codes and Microsatellite based genetic diversity estimates of ten populations of *Sclerocactus brevihamatus* representing two subspecies from Texas. Populations representing SBT include CA, GP, HW, KC, KR, LC, SA, WS, JN, whereas TC represents SBB. N - number of individuals genotyped; $A_r$ - rarefied allelic richness; $A_n$ - mean number of alleles per locus; $A_p$ - number of private alleles; $A_e$ - effective number of alleles; $H_e$ - expected heterozygosity; $H_o$ - observed heterozygosity; $F_{is}$ - inbreeding coefficient; *populations deviated ($\alpha$ = 0.05) from Hardy-Weinberg equilibrium, HWE

| Population Code | N | $A_r$ | $A_n$ | $A_p$ | $A_e$ | $H_e$ | $H_o$ | $F_{is}$ |
|---|---|---|---|---|---|---|---|---|
| CA | 28 | 4.20 | 5.57 | 1 | 3.44 | 0.62 | 0.41 | 0.35* |
| GP | 26 | 3.51 | 4.71 | 3 | 2.41 | 0.55 | 0.34 | 0.35* |
| HW | 28 | 3.96 | 5.00 | 0 | 3.02 | 0.63 | 0.31 | 0.49* |
| KC | 25 | 4.21 | 5.57 | 3 | 3.38 | 0.61 | 0.37 | 0.33* |
| KR | 24 | 3.49 | 4.42 | 0 | 2.54 | 0.53 | 0.35 | 0.39* |
| LC | 26 | 3.97 | 4.85 | 0 | 3.02 | 0.62 | 0.44 | 0.30* |
| SA | 14 | 4.07 | 4.71 | 3 | 3.20 | 0.56 | 0.26 | 0.63* |
| WS | 26 | 3.55 | 4.28 | 2 | 2.62 | 0.56 | 0.34 | 0.44* |
| JN | 28 | 3.75 | 5.00 | 1 | 3.10 | 0.59 | 0.52 | 0.15* |
| **Mean** | **25** | **3.85** | **4.90** | **1.44** | **2.97** | **0.59** | **0.37** | **0.38*** |
| TC | 30 | 4.39 | 6.00 | 11 | 3.04 | 0.52 | 0.26 | 0.46* |

**Table 3:** Pairwise $D_{est}$ above diagonal and pairwise $F_{st}$ below diagonal for 10 populations of *Sclerocactus brevihamatus* based on seven nuclear microsatellites. Populations CA, GP, HW, KC, KR, LC, SA, WS, and JN belong to SBT whereas population TC belong to SBB; *significant $D_{est}$ values ($\alpha$ = 0.05); [a]significant $F_{st}$ values (adjusted $\alpha$ = 0.001) obtained after Bonferroni correction.

| | CA | GP | HW | KC | KR | LC | SA | WS | JN | TC |
|---|---|---|---|---|---|---|---|---|---|---|
| CA | - | 0.11* | 0.07* | 0.06 | 0.20* | 0.14* | 0.20* | 0.23* | 0.04* | 0.35* |
| GP | 0.05 | - | 0.04 | 0.11* | 0.10* | 0.14* | 0.20* | 0.23* | 0.11* | 0.42* |
| HW | 0.03 | 0.02 | - | 0.05 | 0.06* | 0.09* | 0.19* | 0.20* | 0.07* | 0.39* |
| KC | 0.03 | 0.06 | 0.02 | - | 0.13* | 0.07* | 0.24* | 0.15* | 0.07* | 0.40* |
| KR | 0.11 | 0.06 | 0.05 | 0.08 | - | 0.10* | 0.18* | 0.16* | 0.18* | 0.40* |
| LC | 0.07 | 0.07 | 0.03 | 0.03 | 0.06 | - | 0.18* | 0.16* | 0.14* | 0.34* |
| SA | 0.10 | 0.12 | 0.09 | 0.12 | 0.11 | 0.09 | - | 0.19* | 0.29* | 0.06* |
| WS | 0.12 | 0.14 | 0.11 | 0.09 | 0.10 | 0.09 | 0.11 | - | 0.15* | 0.37* |
| JN | 0.02 | 0.06 | 0.04 | 0.04 | 0.11[a] | 0.07 | 0.16[a] | 0.09[a] | - | 0.44* |
| TC | 0.19 | 0.24 | 0.20 | 0.22 | 0.24 | 0.18[a] | 0.05 | 0.22[a] | 0.24[a] | - |

**Table 4.** Results of AMOVA based on seven nuclear microsatellites across nine populations of SBT. df, degree of freedom; SS, sum of squares; MS, Mean sum of squares; E.Var., Estimated variation; %, variation percentage; $Phi_{PT}$ ($\Phi_{PT}$) , percentage of genetic difference among populations; and p, p-value. $Phi_{PT}$ ($\Phi_{PT}$) does not consider genetic variation within populations.

| Source of variation | df | SS | MS | E. var | % | $\Phi_{PT}$ | p |
|---|---|---|---|---|---|---|---|
| **Among Populations** | 8 | 172 | 21 | 0.57 | 7 | 0.07 | 0.00 |
| **Within populations** | 216 | 1587 | 7 | 7.35 | 93 | | |
| **Total** | 224 | 1760 | | 7.92 | 100 | | |

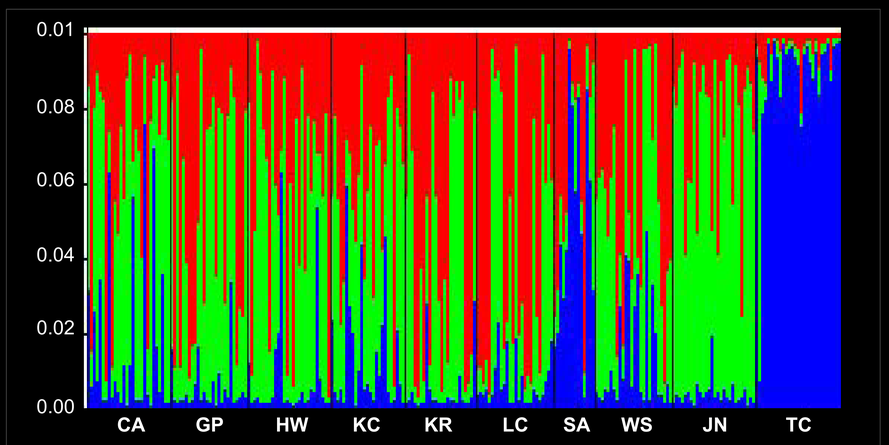


**Figure 4.** Bar plot from STRUCTURE analysis. Colors in the bar graph represent clustering of populations at K = 3, the optimal value of K. Populations CA, GP, HW, KC, KR, LC, SA, WS, and JN belong to *Sclerocactus brevihamatus* subsp. *tobuschii* whereas population TC belongs to *Sclerocactus brevihamatus* subsp. *brevihamatus*. Each thin line in the graph represents a single individual.

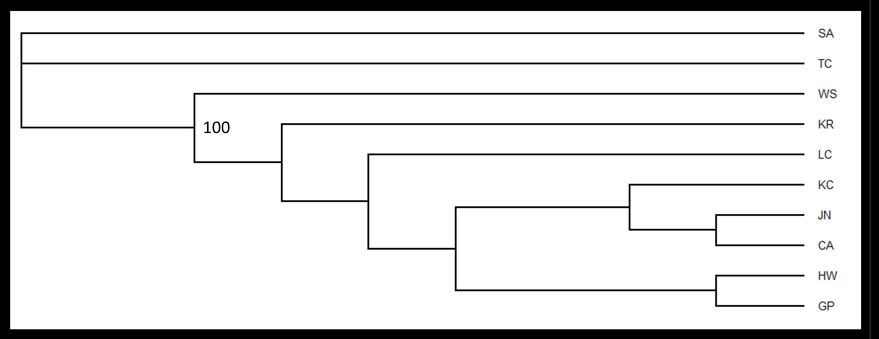


**Figure 5.** An unrooted majority rule consensus neighbor joining dendrogram of ten *Sclerocactus brevihamatus* populations constructed using Nei's genetic distances (Nei 1972) estimated based on seven nuclear microsatellites. Values at the branches are the number of bootstraps support for the split of branches. Only the bootstrap values ≥ 50 are shown.

**Table 5.** Summary of bottleneck test results in 10 *Sclerocactus brevihamatus* populations based on seven nuclear microsatellites using infinite allele model (IAM), two-phase mutation model (TPM), and stepwise mutation model (SMM). The p values were estimated based on 10,000 replications using Wilcoxon test and the values in bold red color are statistically significant ($\alpha$ = 0.05).

| Population | IAM | TPM | SMM |
|---|---|---|---|
| CA | **0.019** | 0.53 | 0.59 |
| GP | 0.460 | 0.85 | 0.97 |
| HW | **0.007** | 0.28 | 0.53 |
| KC | **0.030** | 0.59 | 0.76 |
| KR | 0.054 | 0.53 | 0.81 |
| LC | **0.007** | 0.18 | 0.65 |
| SA | 0.054 | 0.59 | 0.85 |
| WS | **0.003** | 0.18 | 0.71 |
| JN | **0.039** | 0.46 | 0.59 |
| TC (SBB) | 0.590 | 0.46 | 0.59 |

## Conclusions

- We conclude that *Sclerocactus brevihamatus* subsp. *tobuschii* has moderate (mean $H_o$ = 0.37) to moderately high (mean $H_e$ = 0.59) microsatellite based within-population genetic diversity (Table 2). These values are also lower than the estimates in a threatened congener *Sclerocactus glaucus (*mean $H_o$ = 0.47, $H_e$ = 0.66).
- We infer that populations HW, SA, WS, and TC with an inbreeding coefficient ($F_{is}$) close to or higher than 0.50 (Table 2) are at a higher risk of deleterious effect of inbreeding than the other sampled populations (Table 2).
- Population differentiation measures such as mean $F_{st}$ (0.08, Table 3) and $\Phi_{PT}$ (0.07, Table 4) disagree with the mean $D_{est}$ (0.18, Table 3). The difference between $F_{st}$ and $D_{est}$ could be due to high mutation rate of microsatellites or high within-population genetic diversity.
- The Mantel correlation between unbiased pairwise Nei's genetic distances ($r^2$ = 0.054, p = 0.135) and Wright's Fst ($r^2$ = 0.073, p = 0.113) against pairwise geographic distances (ranging from 17 km to 145 km amongst the 9 sampled SBT populations) was not significant.
- Clustering analyses including STRUCTURE (Figure 5), rooted (data not shown) and unrooted neighbor joining dendrograms (Figure 6) consistently showed three genetically distinct groups.
- Although bottleneck effect on CA, HW, KC, LC, WS, and JN under IAM model is evident (Table 5), the other models did not detect bottlenecks in any of the sampled populations.

## Acknowledgments

We gratefully acknowledge U.S. Fish and Wildlife Service and the University of Texas Lady Bird Johnson Wildflower Center Endangered Species Conservation Grant Program for funding. Collectively, Dr. Karen Clary (UT Lady Bird Johnson Wildflower Center), Jackie Poole (retired from Texas Parks and Wildlife Department), Chris Best (U.S. Fish and Wildlife Service), and Charlotte Reemts (The Nature Conservancy) provided access to plant material, populations, geo-coordinate data, and / or assisted with the collection of research materials. We also acknowledge Dr. Shannon Fehlberg and Jessica Allen (Desert Botanical Garden, Phoenix, AZ) for providing insights into extraction protocols. Drs. Mark Burow, Madhav Pandey, and Ratan Chopra assisted with some laboratory and statistical protocols.