TÜBİTAK

Turkish Journal of Botany

http://journals.tubitak.gov.tr/botany/

Research Article

Turk J Bot
(2014) 38: 1080-1094

doi:10.3906/bot-1401-97

K-means clustering and STRUCTURE analyses of genetic diversity in *Tamarix* L. accessions

Habibollah IJBARI[1], Masoud SHEIDAI[1,*], Ahmad-Reza MEHRABIAN[1], Zahra NOORMOHAMMADI[2], Somayeh GHASEMZADEH-BARAKI[1]

[1]Faculty of Biological Sciences, Shahid Beheshti University, Tehran, Iran
[2]Department of Biology, School of Sciences, Science and Research Branch, Islamic Azad University, Tehran, Iran

Received: 30.01.2014 • **Accepted:** 23.07.2014 • **Published Online:** 17.11.2014 • **Printed:** 28.11.2014

Abstract: *Tamarix* L. accessions were collected from Sistan and Baluchestan Province of Iran, covering an area of about 60 km located in the southeast of Iran. The present investigation was performed to study the genetic structure of *Tamarix* populations in this province and present some information about gene flow among these populations. Forty-two tree specimens were randomly collected and studied for molecular and morphological diversity. Based on morphological descriptions we recognized 3 species that were confirmed by detailed morphometric and genetic analysis. The studied tree specimens were separated in 3 distinct clusters by UPGMA clustering and principal coordinate analysis and multidimensional scaling ordination methods. Clustering of molecular (ISSR) data showed some degree of genetic variation within each species. Similarly, STRUCTURE analysis and K-means clustering revealed some degree of genetic admixture among the studied accessions. NeighborNet and reticulograms revealed some degree of gene exchange among the studied populations. AMOVA and the Hickory test showed significant genetic differences of the studied populations and the Mantel test showed significant correlation between genetic distance and geographical distance of these populations.

Key words: Genetic variation, species relationship, *Tamarix*

1. Introduction

Tamarix is an ancient genus in Asia and together with 2 other small Asian genera, *Myricaria* and *Reaumuria*, constitute the family Tamaricaceae (Baum, 1978).

The genus *Tamarix* L. (common name**:** tamarisk) comprises 54 species with unresolved taxonomy. Many species are morphologically very similar, probably due in part to the ability of some *Tamarix* species to interbreed (Baum, 1978).

The genus *Tamarix* is native to a zone stretching from southern Europe and North Africa through the Middle East and southern Asia to China and Japan. There are a few species in disjunct parts of Africa (Rodman, 1989). Baum (1978) considers that *Tamarix* has a major center of speciation in the Pakistan–Afghanistan–Iran–Turkmenistan–southern Kazakhstan–western China area and another in the eastern Mediterranean area.

Tamarisk is planted for erosion control and as wind breaks (Baum, 1978; Allred, 2002). These trees are not significantly grazed or browsed. The tannin substances in the vegetation parts are likely to make it unpalatable. However, tamarisk has commercial value in landscaping and the horticultural trade (Tykač, 1990). *Tamarix* species have been used for fuel and building materials by Native American tribes in the western United States (Moerman, 1998).

Hybridization is known to occur among some *Tamarix* species. For example, molecular studies suggested that *T. chinensis* (and possibly hybrids between it) and *T. ramosissima* occur in some western areas. Some authors continue to distinguish many species, while others consider *T. pentandra*, *T. tetranda*, *T. gallica*, *T. chinensis*, *T. ramosissima*, and *T. parvifolia* to be one variable species or hybridizing group best designated by the single name *T. pentandra* (Sudbrock, 1993).

Similarly, although *T. chinensis* Lour. and *T. ramosissima* Ledeb. are morphologically and genetically distinct in Asia, the North American population is dominated by their hybrids and

Friedman et al. (2012) referred to the complex of *T. ramosissima*, *T. chinensis*, and their hybrids as salt cedar.

Tamarix species are long-lived (50 to 100 years) flowering plants and are not true pines (conifers). *Tamarix* species are spreading, often multibranched trees up to 12 m

* Correspondence: msheidai@yahoo.com

tall with pendulous branches. Some taxa are evergreen (*T. aphylla*), whereas some others, including *T. ramosissima*, are deciduous. *Tamarix* species reproduce from broken stem fragments and from seeds. Vegetative reproduction is particularly successful when branches are broken up by floodwaters and carried downstream (Parsons and Cuthbertson, 1992).

According to Gaskin and Kazmer (2009), the proper approach to studying hybrids among *Tamarix* species is using multilocus molecular markers such as simple sequence repeat (SSR) markers and inter-simple sequence repeat (ISSR) markers. They suggested that using single-locus DNA sequence markers may underestimate hybrid frequency (e.g., an F1 hybrid back-crossed with one of its parental species may produce progeny with a homozygotic genotype for the single-locus marker and appear to be a parental type, when it actually contains genetic material from both species). Additionally, the single-locus marker does little to inform us about levels of introgressive hybridization (e.g., how much of a plant's genetic material comes from either parental species).

In the present investigation, we used ISSR molecular markers along with morphometric analysis to study the genetic and morphological features of *Tamarix* accessions growing in Sistan and Baluchestan Province (area of about 60 km^2). This region of Iran is one of the main areas known to contain large numbers of *Tamarix* accessions possibly from different species occurring in sympatry. Moreover, due to the known tendency of *Tamarix* species to hybridize with each other, we expected to encounter some intermediate forms or interspecific hybrids in this region. Therefore, we started morphological and genetic analyses of available accessions in this location as the first part of our study with the following aims: 1) to investigate whether or not *Tamarix* species can be differentiated by morphological and genetic features; 2) to determine whether there is gene exchange occurring among the studied taxa and if there are any intermediate or interspecific hybrids present in the area studied; and 3) if gene flow did occur among the studied accessions, to determine if it is correlated with the geographical distances of these plants.

2. Materials and methods

2.1. Plant materials

In order to identify *Tamarix* species growing in Sistan and Baluchestan Province and to study genetic structure as well as gene flow among geographical populations in this region, we randomly collected 42 tree specimens. The area covered was 60 km^2 (Table 1). The area studied is among the main distribution land of *Tamarix* in Iran. Details

Table 1. *Tamarix* specimens' localities and ecological features.

	Species	Locality	Latitude	Longitude	Altitude (m)
1	*T. tetragyna*	Haji village, Hirmand	31°12′56.34″N	61°40′37.43″E	479
2	*T. tetragyna*	Haji village, Hirmand	31°12′56.34″N	61°40′37.43″E	479
3	*T. tetragyna*	Haji village, Hirmand	31°12′56.34″N	61°40′37.43″E	479
4	*T. tetragyna*	Khaje Mountain, Hamoonshahr	30°56′16.16″N	61°15′39.82″E	479
5	*T. tetragyna*	Milag, Hirmand	30°59′54.74″N	61°47′42.73″E	489
6	*T. tetragyna*	Milag, Hirmand	30°59′54.74″N	61°47′42.73″E	489
7	*T. tetragyna*	Hirmand	31°8′12.64″N	61°47′6.97″E	483
8	*T. tetragyna*	Hirmand	31°8′12.64″N	61°47′6.97″E	483
9	*T. tetragyna*	Hirmand	31°8′12.64″N	61°47′6.97″E	483
10	*T. karkalensis*	Mohammad-abad, Zabol	30°48′19.98″N	61°25′11.05″E	481
11	*T. karkalensis*	Mohammad-abad, Zabol	30°48′19.98″N	61°25′11.05″E	481
12	*T. karkalensis*	Khaje Mountain, Hamoonshahr	30°56′16.16″N	61°15′39.82″E	479
13	*T. karkalensis*	Khaje Mountain, Hamoonshahr	30°56′16.16″N	61°15′39.82″E	479
14	*T. karkalensis*	Khemer, Hirmand	31°8′12.64″N	61°46′6.58″E	481

Table 1. (Continued).

15	*T. karkalensis*	Neyatak, Hirmand	31°7′9.86″N	61°37′5.36″E	483
16	*T. karkalensis*	Neyatak, Hirmand	31°7′9.86″N	61°37′5.36″E	483
17	*T. karkalensis*	Zabol airport	31°4′46.35″N	61°32′19.23″E	483
18	*T. karkalensis*	Zabol airport	31°4′46.35″N	61°32′19.23″E	483
19	*T. karkalensis*	Milag, Hirmand	30°59′54.74″N	61°47′42.73″E	489
20	*T. karkalensis*	Hirmand	31°8′12.64″N	61°47′6.97″E	483
21	*T. karkalensis*	Hirmand	31°8′12.64″N	61°47′6.97″E	483
22	*T. karkalensis*	Hirmand	31°8′12.64″N	61°47′6.97″E	483
23	*T. kotschyi*	Haji village, Hirmand	31°12′56.34″N	61°40′37.43″E	479
24	*T. kotschyi*	Haji village, Hirmand	31°12′56.34″N	61°40′37.43″E	479
25	*T. kotschyi*	Haji village, Hirmand	31°12′56.34″N	61°40′37.43″E	479
26	*T. kotschyi*	Bibidoost	31°4′28.83″N	61°39′20.01″E	485
27	*T. kotschyi*	Bibidoost	31°4′28.83″N	61°39′20.01″E	485
28	*T. kotschyi*	Mohammad-abad, Zabol	30°48′19.98″N	61°25′11.05″E	481
29	*T. kotschyi*	Mohammad-abad, Zabol	30°48′19.98″N	61°25′11.05″E	481
30	*T. kotschyi*	Khaje Mountain, Hamoonshahr	30°56′16.16″N	61°15′39.82″E	479
31	*T. kotschyi*	Khaje Mountain, Hamoonshahr	30°56′16.16″N	61°15′39.82″E	479
32	*T. kotschyi*	Khaje Mountain, Hamoonshahr	30°56′16.16″N	61°15′39.82″E	479
33	*T. kotschyi*	Khemer, Hirmand	31°10′48.66″N	61°46′6.58″E	481
34	*T. kotschyi*	Khemer, Hirmand	31°10′48.66″N	61°46′6.58″E	481
35	*T. kotschyi*	Khemer, Hirmand	31°10′48.66″N	61°46′6.58″E	481
36	*T. kotschyi*	Neyatak, Hirmand	31°7′9.86″N	61°37′5.36″E	483
37	*T. kotschyi*	Neyatak, Hirmand	31°7′9.86″N	61°37′5.36″E	483
38	*T. kotschyi*	Zabol airport	31°4′46.35″N	61°32′19.23″E	483
39	*T. kotschyi*	Zabol airport	31°4′46.35″N	61°32′19.23″E	483
40	*T. kotschyi*	Milag, Hirmand	30°59′54.74″N	61°47′42.73″E	489
41	*T. kotschyi*	Milag, Hirmand	30°59′54.74″N	61°47′42.73″E	489
42	*T. kotschyi*	Hirmand	31°8′12.64″N	61°47′6.97″E	483

of localities of the studied trees are given in Table 1. The voucher specimens were deposited in the herbarium of Shahid Beheshti University (HSBU).

2.2. Morphometry

The morphological characters studied are presented in Table 2. The mean of quantitative morphological characters (5 readings in each case) was used. Qualitative characters were coded as binary and multistate characters. For multivariate statistical analyses, data were standardized (mean = 0, variance = 1) and used to determine Euclidean and Gower distances for clustering. Different distance measures were used to check the consistency of clustering results.

Table 2. Morphological characters and their codings.

Species	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
tetragyna	1	1	3	2	1	1	1	70	5	1	1	1	3	3	4	0.6	1	1	3	1
tetragyna	1	1	3.5	2	1	1	1	80	6	1	1	1	3	1	2.5	0.25	1	1	2	1
tetragyna	1	1	3	2	1	1	1	85	6	1	1	1	3	1	3	0.5	1	1	1	1
tetragyna	1	1	3	2	1	1	2	70	9	1	1	1	3	3	3.25	0.3	3	1	3	2
tetragyna	0	1	2.5	2	1	1	1	80	6	1	1	1	3	2	3.3	0.5	1	1	3	1
tetragyna	1	1	3	2	1	1	1	65	6	1	1	1	3	2	2.75	0.3	1	2	2	1
tetragyna	1	1	2.5	1	1	1	1	65	5	1	1	1	3	3	2.5	0.6	1	2	3	1
tetragyna	1	1	3.5	1	1	1	1	90	7	1	1	1	3	3	4	0.75	1	2	3	1
karkalensis	1	1	2.25	1	1	2	3	17.5	3	1	2	2	3	2	1.75	0.7	1	3	3	3
karkalensis	1	1	1.4	2	1	2	3	15	1.75	1	2	2	3	1	1.75	0.6	1	3	2	3
karkalensis	0	1	1.75	2	0	2	3	13	2	1	2	2	3	1	1.25	0.4	1	3	3	3
karkalensis	1	1	1.75	1	1	2	3	17.5	2.25	1	2	2	3	1	1.5	0.25	1	3	3	3
karkalensis	1	1	1.75	1	1	2	3	22.5	2.5	1	2	2	3	2	1.2	0.4	1	3	3	3
karkalensis	1	1	1.8	1	1	2	3	20	1.5	1	2	2	3	1	1.2	0.35	1	3	3	3
karkalensis	1	1	2.25	1	1	2	3	14	2.4	1	2	2	3	3	2.15	0.35	1	3	3	3
karkalensis	1	1	2.5	1	1	2	3	15	2.5	1	2	2	3	1	1.25	0.3	1	3	3	3
karkalensis	0	1	2	1	0	2	3	16	2.5	1	2	2	3	3	2.15	0.25	1	3	3	3
karkalensis	0	1	1.35	1	0	2	3	10	2.25	1	2	2	3	1	0.95	0.65	1	3	3	3
karkalensis	1	1	2.1	1	1	2	3	21	1.75	1	2	2	3	3	1.35	0.25	1	3	3	3
kotschyi	0	1	2	1	0	1	1	30	3	1	2	1	3	1	1.1	0.25	1	1	3	1
kotschyi	0	2	1.85	1	1	1	1	17.5	2.25	1	2	1	3	1	1.25	0.4	1	1	3	1
kotschyi	0	2	1.5	1	1	1	2	18	1.5	2	2	2	3	1	0.85	0.65	1	3	3	1
kotschyi	0	2	1.2	1	0	1	2	21	3	2	2	1	1	1	0.95	0.7	2	3	3	3
kotschyi	0	1	1.1	1	0	1	2	20	3	1	2	1	3	2	1.1	0.4	2	3	3	1
kotschyi	0	1	1.1	1	0	1	1	20	2	1	2	1	3	1	0.9	0.5	2	1	3	1
kotschyi	0	1	1.5	1	0	1	1	20	3	1	3	1	3	1	0.95	0.33	2	1	3	3
kotschyi	1	1	1.75	1	1	1	1	18	3	1	2	1	3	1	1.25	0.4	2	2	3	1
kotschyi	0	2	1.4	1	0	1	1	15	1.5	1	2	1	3	1	0.9	0.45	2	2	3	1
kotschyi	1	2	1.2	1	1	1	1	24	1.5	2	3	1	1	1	0.85	0.65	3	2	3	3
kotschyi	1	1	1.5	1	1	1	1	15	3	1	2	1	2	1	0.95	0.4	2	2	3	3

Character numbers 1–42 are, respectively: stem pile, leaf shape, leaf length, shape of leaf margin, leaf pile, inflorescence, flower density, length of inflorescence, width of inflorescence, leaflet attachment, shape of leaflet, shape of leaf top, ratio of leaflet size/leaf size, leaflet/calyx, length of leaflet, width of leaflet, pedicel/calyx, calyx segments, shape of internal calyx, shape of external calyx, tip of internal calyx, tip of external calyx, internal calyx naviculate, external calyx naviculate, calyx pile, length of internal calyx, length of external calyx, width of internal calyx, width of external calyx, corolla segments, corolla symmetry, corolla length, corolla width, stamen number, base of filament, attachment of stamen to lobe, place of stamen extrusion, anther tip, anther symmetry, anther length, anther width, and number of pistils.

Table 2. (Continued.)

21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42
3	1	1	1	0	1.4	1.75	1	1	1	1	2.25	0.9	2	2	1	1	0	2	0.75	0.2	2
3	3	0	0	0	0.9	1.4	0.9	0.6	1	1	1.9	1	1	2	1	1	0	2	0.6	0.2	2
1	1	0	1	0	1.25	1.5	0.5	0.75	1	1	2	1	1	2	1	1	0	2	0.75	0.25	3
3	3	0	0	1	1.2	1.25	0.9	1.1	1	1	2.1	0.9	1	2	1	1	0	2	0.75	0.25	3
1	1	1	1	0	1.1	1.75	1.1	1.1	1	1	2.5	1.5	1	2	1	1	0	2	0.75	0.25	3
1	1	1	1	1	1.25	1.25	0.75	0.75	2	1	2.25	0.9	2	2	1	1	0	2	0.6	0.2	3
1	1	1	1	1	1.4	1.5	0.9	1.1	2	1	2.4	1.1	2	2	1	1	0	2	0.9	0.25	3
1	1	1	1	1	1.75	1.9	0.6	0.6	2	1	2.6	1.1	2	2	1	1	0	2	0.9	0.25	3
2	1	0	1	1	0.95	1	0.6	0.7	3	2	1.1	0.95	3	1	2	2	0	2	0.3	0.2	1
2	2	0	1	0	0.6	0.9	0.5	0.6	3	1	0.6	0.5	3	1	2	2	0	2	0.4	0.25	1
2	2	1	1	0	1	1.25	0.8	0.8	3	1	0.8	0.6	3	1	2	2	0	2	0.3	0.2	1
1	1	0	1	0	0.9	0.9	0.6	0.55	3	1	1.25	0.6	3	1	2	2	0	2	0.3	0.15	1
2	2	0	1	0	1	0.9	0.65	0.7	3	1	0.95	0.7	3	1	2	2	0	2	0.35	0.17	1
2	2	1	1	0	1.1	1.1	0.6	0.55	3	1	1.1	0.6	3	1	2	2	1	2	0.35	0.18	1
2	1	1	1	1	0.9	0.95	0.55	0.65	3	1	0.6	0.5	3	1	2	2	1	2	0.3	15	1
2	2	1	1	0	0.95	1.1	0.6	0.65	3	1	1.25	0.5	3	1	2	2	0	2	0.4	0.18	1
1	1	0	1	0	1.1	1.15	0.65	0.7	3	2	1	0.65	3	1	2	2	1	2	0.35	0.2	2
1	1	0	1	0	0.85	1	0.45	0.45	3	2	1.2	0.6	3	1	2	2	1	2	0.4	0.18	1
2	2	0	1	0	0.95	0.95	0.45	0.65	3	2	0.6	0.5	3	1	2	2	1	2	0.22	0.15	1
1	1	0	1	0	0.8	0.85	0.45	0.55	1	1	1.1	0.55	1	2	1	1	1	1	0.25	0.18	3
1	1	0	1	0	1	1	0.45	0.5	1	1	1.35	0.7	1	1	1	1	1	1	0.35	0.18	3
2	1	0	1	0	0.55	0.6	0.35	0.35	3	1	0.85	0.7	3	1	2	1	1	2	0.2	0.1	1
1	1	0	0	0	0.8	0.8	0.55	0.55	3	1	2.5	0.55	3	1	1	1	1	2	0.2	0.15	1
1	1	0	1	0	0.8	0.75	0.55	0.55	3	1	1.5	0.85	3	1	2	2	0	2	0.4	0.2	1
1	1	0	0	0	0.75	0.95	0.55	0.65	1	1	1.2	0.55	1	2	1	1	1	2	0.4	0.2	1
1	1	0	0	0	0.65	0.85	0.45	0.45	1	1	1.1	0.55	1	1	1	1	0	1	0.35	0.18	1
1	1	0	1	0	0.95	0.95	0.45	0.6	3	1	1.15	0.5	3	1	1	1	0	2	0.3	0.18	1
1	1	0	1	0	0.75	0.85	0.35	0.35	2	1	1.1	0.45	2	2	1	1	0	2	0.2	0.1	1
2	2	0	1	0	0.85	1	0.45	0.55	2	1	1.15	0.55	2	1	2	2	1	2	0.25	0.12	2
1	1	0	1	0	0.65	0.6	0.3	0.45	2	1	1.25	0.55	2	2	1	1	0	2	0.2	0.1	3

Multivariate analysis of variance (MANOVA) testing was performed to show significant differences in quantitative morphological characters among the studied species.

The unweighted pair group method with arithmetic mean (UPGMA) and the Ward method (minimum spherical cluster method) were used for grouping of the accessions after 100 bootstrapping runs (Podani,

2000). Different ordination methods were applied for standardized data like principal components analysis (PCA), principal coordinate analysis (PCoA), and multidimensional scaling (MDS) (Podani, 2000). Data analyses were performed using PAST ver. 2.17 (Hamer et al., 2012). Different clustering and ordination methods were used to check the consistency of results, and when results are similar, only one of them is presented here.

2.3. ISSR assay and genetic analyses

For molecular studies, the fresh leaves were randomly collected from 42 randomly selected plants in the studied area and were dried in silica gel powder. The genomic DNA was extracted using CTAB-activated charcoal protocol (Križman et al., 2006). The extraction procedure was based on activated charcoal and polyvinylpyrrolidone (PVP) for binding of polyphenolics during extraction and under mild extraction and precipitation conditions. This promoted high-molecular-weight DNA isolation without interfering contaminants. Quality of extracted DNA was examined by running on 0.8% agarose gel.

Ten ISSR primers, UBC 807, UBC 810, UBC 811, UBC 834, CAG(GA)7, (CA)7AC, (CA)7AT, (CA)7GT (GA)9A, and (GA)9T, commercialized by the University of British Columbia, were used. PCR reactions were performed in a 25-µL volume containing 10 mM Tris-HCl buffer at pH 8, 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM of each dNTP (Bioron, Germany), 0.2 µM of a single primer, 20 ng of genomic DNA, and 3 U of Taq DNA polymerase (Bioron). Amplification reactions were performed in a Techne thermocycler (Germany) with the following program: 5 min for initial denaturation step at 94 °C, 30 s at 94 °C, 1 min at 52 °C, and 1 min at 72 °C. The reaction was completed by a final extension step of 7 min at 72 °C. The amplification products were visualized by running on 2% agarose gel, followed by ethidium bromide staining. The fragment sizes was estimated using a 100-bp molecular size ladder (Fermentas, Germany). The experiment was replicated 3 times and constant ISSR bands were used for further analyses.

The ISSR bands obtained were treated as binary characters and coded accordingly (presence = 1, absence = 0). Genetic diversity parameters were determined in each species. These parameters were the percentage of allelic polymorphism, allele diversity (Weising, 2005), Nei's gene diversity (H), the Shannon information index (I) (Weising, 2005), the number of effective alleles, and percentage of polymorphism. The genetic divergence of the studied populations was checked by PCoA and after 999 permutations. The Jaccard and Nei genetic distances (Weising, 2005) were determined among the studied trees and used for the grouping of the genotypes. Two different similarity and distance measures were used to check the consistency of the results.

Neighbor joining (NJ) trees followed by 100 bootstrapping runs, PCoA, and MDS were used for the grouping of the studied trees (Podani, 2000). PAST ver. 2.17 (Hamer et al., 2012) and DARwin ver. 5 (2012) were used for these analyses.

Genetic differentiation of the studied populations were studied by 2 different approaches. First, we used AMOVA with 1000 permutations as performed in GenAlex 6.4 (Peakall and Smouse, 2006).

Second, we used Hickory ver. 1.0 (http://www.eeb.uconn.edu), a Bayesian program that calculates the θB value. This is the estimate parameter related to population genetic structure. In this way we overcome the potential problems caused by the dominance of ISSR markers. The Bayesian method used here does not assume that genotypes are in Hardy–Weinberg proportions within populations, and it does not treat multilocus ISSR phenotypes as haplotypes. It takes full advantage of the information provided by dominant markers, allowing us to incorporate uncertainty about the magnitude of the within-population inbreeding coefficient into estimates of F_{ST} (http://www.eeb.uconn.edu).

Moreover, new parameters of genetic differentiation such as G_{ST} est = standardized measure of genetic differentiation [(G_{ST} est (n – 1 + Hs est))((n – 1) (1 – Hs est)] (Hedrick 2005) and D est = Jost measure of differentiation (Jost, 2008) were determined.

Since we had no idea about the extent of gene flow among the studied taxa or whether the *Tamarix* trees studied cross-pollinate freely throughout the studied area, the Mantel test (Podani, 2000) was performed to study the association between molecular distance and geographical distance of the studied populations.

Genetic structure of the populations was studied by model-based clustering as performed by STRUCTURE software ver. 2.3 (Pritchard et al., 2000). We used the admixture ancestry model under the correlated allele frequency model. A Markov chain Monte Carlo simulation was run 20 times for each value of *K* (1–4) after a burn-in period of 10^5. Data were scored as dominant markers and analysis followed the method suggested by Falush et al. (2007).

For the optimal value of *K* in the population studied, we used 2 methods:

1) The STRUCTURE Harvester website (Earl and von Holdt, 2012) was used to perform the Evanno method to identify the proper value of *K* (Evanno et al., 2005). The choice of the most likely number of clusters (*K*) was carried out comparing log probabilities of data [Pr (X|K)] for each value of *K* (Pritchard et al. 2000), as well as by calculating an ad hoc statistic Δ*K* based on the rate of change in the log probability of data between successive *K* values, as described by Evanno et al. (2005).

2) K-means clustering was performed with Genodive ver. 2. (2012), based on the original work of Meirmans and Van Tienderen (2004). In K-means clustering, the optimal clustering is the one with the smallest amount of variation within clusters, which is calculated using the within-clusters sum of squares. The minimization of the within-groups sum of squares that is used in K-means clustering is, in the context of hierarchical AMOVA, equivalent to minimizing the among-populations-within-groups sum of squares, SSDAP/WG. The hierarchical population structure in AMOVA then consists of different hierarchical levels: individuals, populations, and clusters of populations. Different F-statistics can be calculated based on the variance components for the different hierarchical levels. In terms of F-statistics, the minimization of SSDAP/WG comes down to a maximization of FCT, the variance among clusters (C) relative to the total variance (T) (Meirmans and Van Tienderen, 2004).

We used 2 summary statistics to present K-means clustering, pseudo-F (Caliński and Harabasz, 1974) and the Bayesian information criterion (BIC; Schwarz, 1978). Pseudo-F (Caliński and Harabasz, 1974) relates r^2, the fraction of the total variance that is explained by the clustering, to the number of clusters *K* and the number of populations n: $F_k = r^2 / (1 - r^2) (n - k)$, where $r^2 = (SSDT - SSDAP/WG)/(SSDT - SSDWP)$. The clustering with the highest value for pseudo-F is regarded to provide the best fit (Meirmans and Van Tienderen, 2004).

The BIC is calculated as:

$BIC_k = n \times \ln (SSE) + k \times \ln (n)$.

In this case, the clustering with the lowest value for the BIC is regarded to provide the best fit (Meirmans and Van Tienderen, 2004).

In order to identify gene exchange among the studied accessions, we performed both reticulation analysis by DARwin ver. 5 and the distance-based NeighborNet (Bryant and Moulton, 2004) as implemented in SplitsTree4 (Huson and Bryant, 2006). The DARwin program infers the reticulogram from a distance matrix. For reticulation, we first built a supporting tree using NJ, followed by a reticulation branch that minimizes the least-squares at each step of the algorithm (Legendre and Makarenkov, 2002).

Due to the occurrence of a high degree of gene exchange and genetic admixture in the studied *Tamarix* trees, the 'hybrid index' as implemented in Genodive ver. 2 (2012) was calculated. A hybrid index is a quantitative estimate of the genetic contribution of 2 parental species or populations to an individual of unknown provenance. Genodive uses the method of Buerkle (2005) to calculate a maximum likelihood estimate of such a hybrid index. The analysis requires 3 datasets, which should be coded as populations in a genetic data file. Two populations should contain the genotypes for the 2 parental gene pools, referred to as the reference population and the alternative population; usually these are 2 species. The third population should contain the genotypes for the putatively hybrid individuals. The analysis returns the maximum likelihood estimate of the hybrid index, the likelihood value, and the upper and lower limits of the 95% confidence interval.

3. Results

3.1. Morphometry

MANOVA showed significant difference ($F = 11.24$, $P < 0.01$) in all quantitative morphological characters among the studied species. Pairwise comparisons (post hoc) by pairwise Hotelling tests also showed significant differences for morphological characters among all 3 studied species ($P < 0.01$).

PCA analysis of morphological data revealed that the first 3 components account for about 60% of total variance. The most variable characters of the first PCA axis with about 32% of total variance were shape of leaf margin, width of inflorescence, leaflet attachment, shape of leaf top, width of leaflet, length of external calyx, width of internal calyx, stamen number, and anther width. These characters had the highest correlation value (>0.7) with this axis.

The UPGMA dendrogram of morphological characters after 100 bootstraps is presented in Figure 1. Tree specimens of *T. tetragyna* were separated from the other 2 species with 100% bootstrap value. Tree specimens of *T. karkalensis* and *T. kotschyi* were placed closer to each other but still formed separate clusters with a 77% bootstrap value.

The MDS plot (Figure 2) also separated the tree specimens of 3 studied species into distinct groups, supporting the UPGMA dendrogram. In these plots the 3 species were placed in distinct positions, indicating their morphological differences and supporting the MANOVA results.

3.2. Genetic diversity

Genetic analyses are provided in Table 3 and Figures 3–8. All ISSR primers produced polymorphic bands. A data matrix of 42 × 103 was formed for genetic analyses. Genetic diversity parameters determined in the 3 studied species are presented in Table 3. The highest value for genetic polymorphism (92.59%) was observed in *T. karkalensis*. The same species had the highest values of gene diversity and Shannon information index (0.385 and 0.239, respectively). *T. tetragyna* had the lowest values of these genetic parameters while *T. kotschyi* had values in between those of the other 2 studied species.

AMOVA showed significant genetic differences among the studied species' φPT value (0.151, $P = 0.01$). This analysis revealed that 15% of total genetic variation

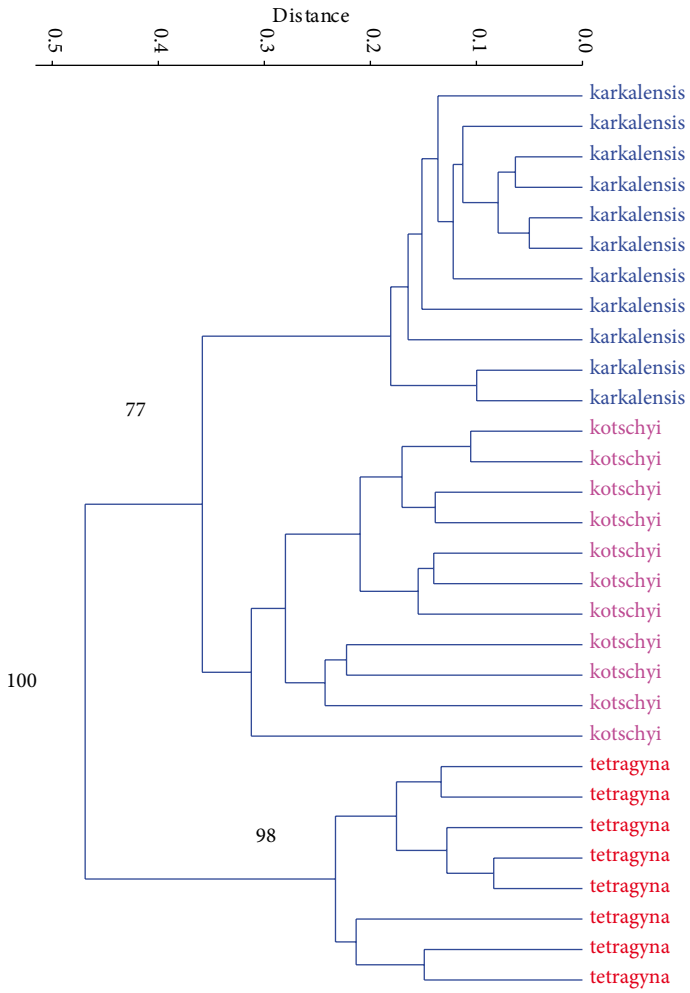


Figure 1. UPGMA dendrogram of morphological characters (numbers below branches are bootstrap values).

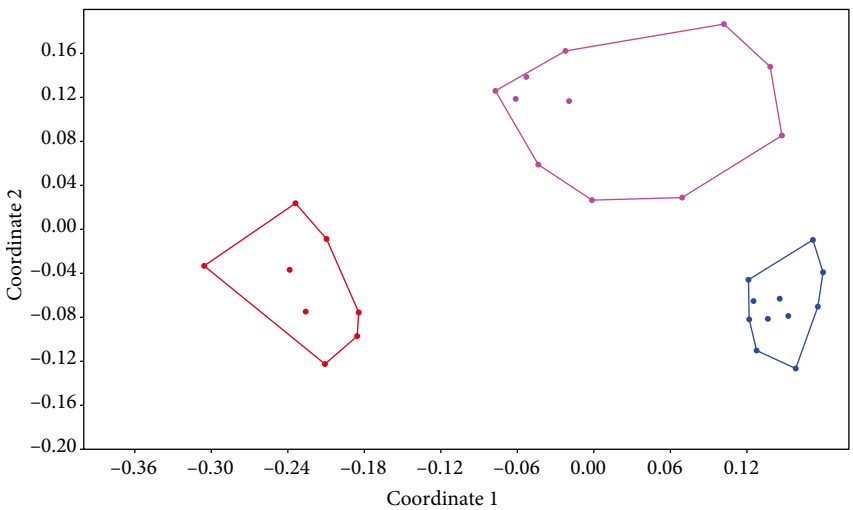


Figure 2. MDS plot of morphological characters.

Table 3. Genetic diversity parameters in 3 studied *Tamarix* species.

Sp.	Na	Ne	I	He	P%	Hs	Hs/Ht
T. tetragyna	1.816	1.432	0.172	0.104	40.74	0.282	0.311
T. karkalensis	1.932	1.504	0.385	0.239	92.59	0.316	0.349
T. kotschyi	1.990	1.482	0.224	0.155	37.04	0.307	0.339
Mean	1.912	1.472	0.260	0.166	56.79		

Na = Mean number of alleles, Ne = no. of effective alleles, I = Shannon's information index, He = gene diversity, Hs = expected heterozygosity within populations, Ht = total expected heterozygosity.

occurred among the species and 85% within the studied species. Pairwise φPT values obtained among species showed significant difference between *T. tetragyna* and *T. kotschyi*, as well as between *T. karkalensis* and *T. kotschyi* ($P < 0.01$). However, *T. tetragyna* and *T. karkalensis* did not differ significantly from each other. Similarly, the Hickory test produced a theta-II value of 0.160342 among the studied species, which is considered to represent significant genetic difference.

Different genetic differentiation parameters determined by permutation test in the Genodive program also showed significant difference among the species and indicated that they genetically differentiated. It produced the following values: $F_{ST} = 0.151$ ($P = 0.01$), G_{ST} (fixation index) = 0.05 ($P = 0.01$), D est (Jost differentiation index) = 0.087 ($P = 0.01$).

The NJ tree and PCoA grouping of the tree specimens based on ISSR data produced similar results. Therefore, only the PCoA plot is presented and discussed here (Figure 3). Tree specimens collected from *T. tetragyna* were almost all grouped together and separated from the other 2 studied species. Tree specimens of *T. karkalensis* and *T. kotschyi* were intermixed. In our morphological results, we observed closer affinity between *T. karkalensis* and *T. kotschyi*. The Mantel test with 5000 permutations, performed between genetic distance and morphological distance of randomly selected trees of the 3 studied species, produced $r = 0.37$ and $P < 0.01$. This indicates that *T. tetragyna*, which differed morphologically from the others, also has higher genetic distance/difference from the other 2 species.

The Mantel test with 5000 permutations was also performed between genetic distance and geographical distance of the studied trees. It produced $r = 0.21$, $P = 0.03$, which is a high r value in permutation testing, indicating significant relationship between the 2 investigated distances. Therefore, with increase in geographical distance of the studied *Tamarix* trees, a lower gene flow occurs among them.

Therefore, the results showed that *Tamarix* trees located closer to each other are genetically more alike and those placed far from each other become genetically more differentiated.

The reticulation tree of the studied *Tamarix* species is presented in Figure 4. It shows the occurrence of gene exchange among trees of all 3 species. This figure, however, also reveals that gene exchange occurred between trees located far from each other, too. For example, gene exchange occurred between tree no. 2 of *T. tetragyna* collected from Haji village, Hirmand, and trees no. 30 and 31 of *T. karkalensis* collected from Khaje Mountain, Hamoonshahr. Similarly, tree no. 24 of *T. karkalensis* collected from Haji village, Hirmand, and tree no. 28 of the same species collected from Mohammad-abad, Zabol, also had gene exchange.

NeighborNet results are provided in Figure 5. This figure reveals more refined details of gene exchange among *Tamarix* accessions. It shows gene exchange between trees of each species and also between all 3 studied species. Moreover, NeighborNet showed close genetic relationships between the studied species. Figure 5 also shows that the specimens collected from *T. tetragyna* (coded as A1 in the figure) were almost all grouped together, while tree specimens of *T. karkalensis* and *T. kotschyi* were intermixed and close to each other (coded A2 and A3 in the figure).

STRUCTURE analysis triangle and plot are presented in Figures 6 and 7. The triangle (Figure 6) that is based on the Bayesian approach supports the PCoA plot results. It shows genetic admixture of the studied species, while the STRUCTURE plot (Figure 7), reveals the allelic composition and their frequency in the studied trees.

The triangle (Figure 6) reveals a continuous placement of tree specimens of *T. karkalensis* and *T. kotschyi* close to each other due to genetic affinity. It is interesting to note that the tree specimens of these 2 species occurred close to each other almost throughout the entire studied area. Therefore, the close genetic affinity of these 2 species is possibly due to gene exchange between them, as also revealed by NeighborNet.

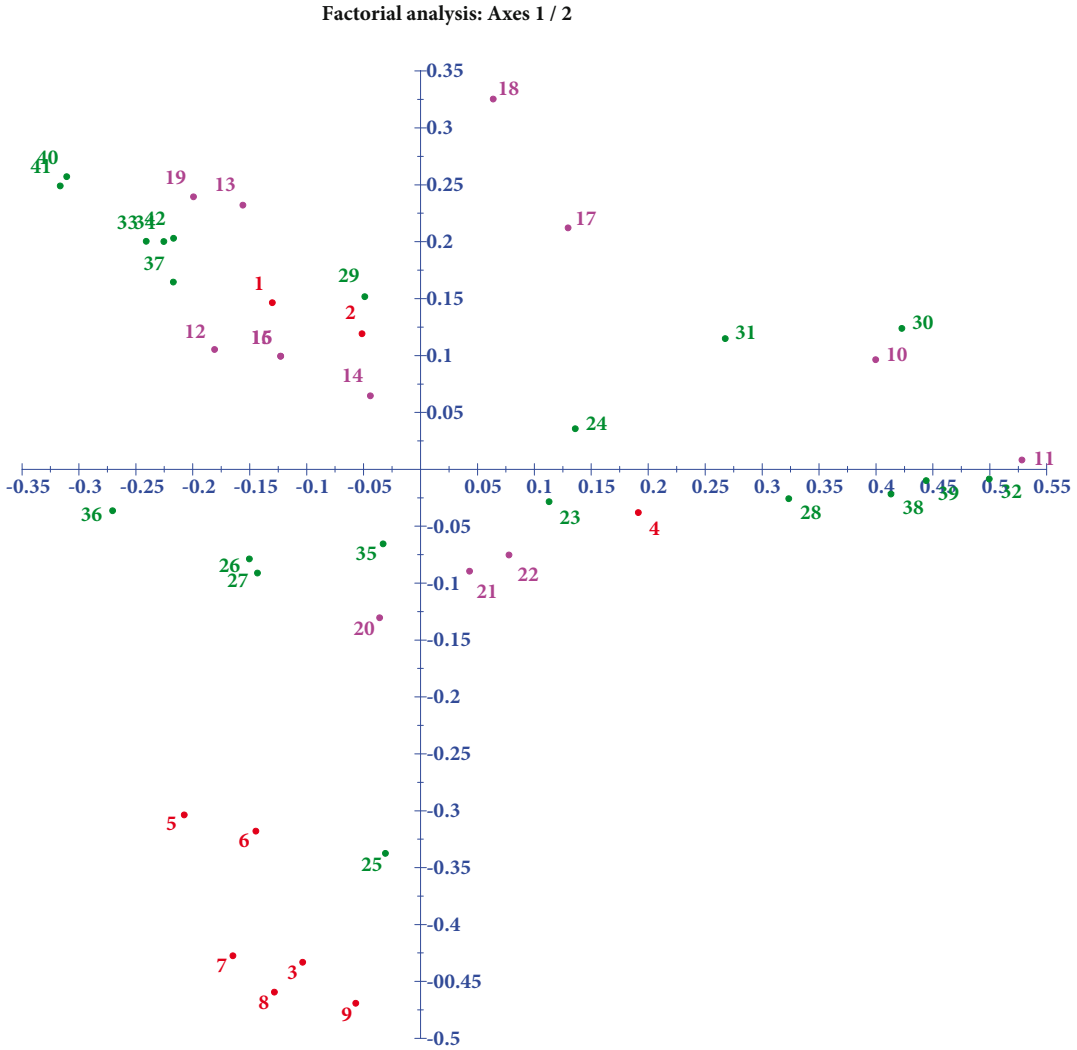


Figure 3. PCoA plot of ISSR data (trees no. 1–8 = *T. tetragyna*, 9–22 = *T. karkalensis*, 23–42 = *T. kotschyi*).

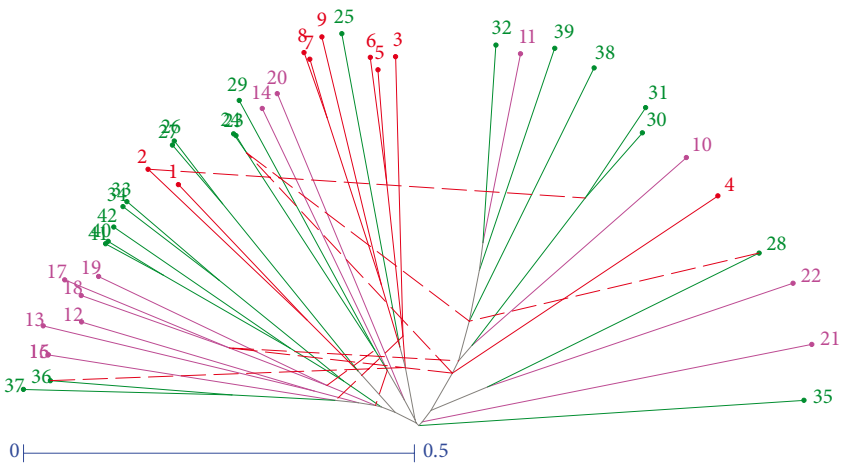


Figure 4. Reticulation tree of *Tamarix* trees (trees no. 1–8 = *T. tetragyna*, 9–22 = *T. karkalensis*, 23–42 = *T. kotschyi*; dashed lines indicate gene exchange).

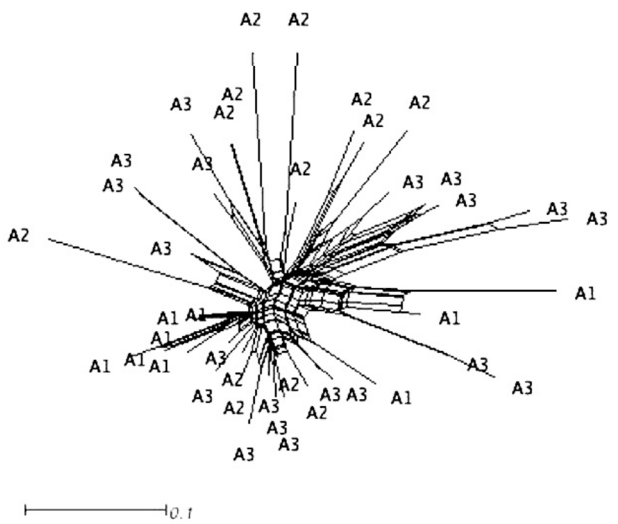


Figure 5. NeighborNet results of the *Tamarix* accessions obtained by the SplitsTree4 program. A1–A3: *T. tetragyna*, *T. karkalensis*, and *T. kotschyi*, respectively.

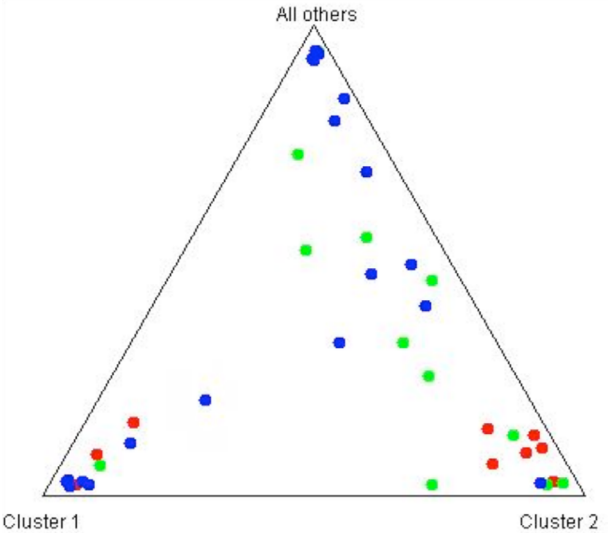
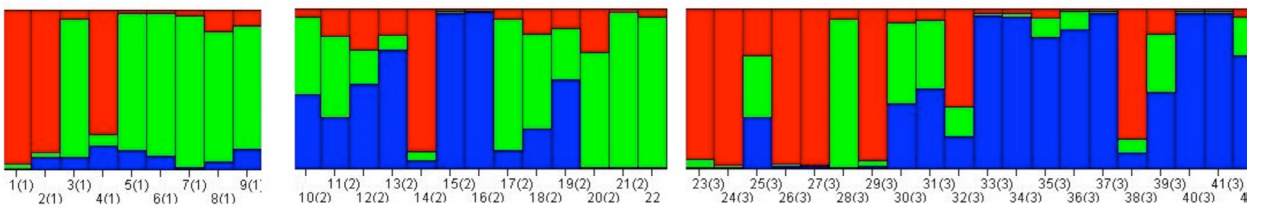


Figure 6. Triangle plot of STRUCTURE (red dots = *T. tetragyna*, blue dots = *T. karkalensis*, green dots = *T. kotschyi*).



T. tetragyna *T. karakalensis* *T. kotschyi*

Figure 7. STRUCTURE plot of *Tamarix* trees showing allelic combinations and their frequencies.

The STRUCTURE plot revealed that, in general, 3 types of allelic combinations (red-, green-, and blue-colored segments) were present in the studied *Tamarix* accessions. The tree specimens of *T. tetragyna* had higher frequency of green and red segments. Even individual trees of this species differed greatly in their allelic composition from each other. High intraspecific genetic variation was shown by AMOVA as stated before.

The tree specimens of *T. karkalensis* had higher frequency of green and blue segments. One individual tree of this species (tree no. 14) had more of the red segments. The trees of this species also differed in their allelic composition.

The tree specimens of *T. kotschyi* had all 3 colors of segments, but slightly more of the blue segments. This result is in agreement with our previous statement that *T. kotschyi* has high genetic affinity to *T. karkalensis*, due to continuous contact and gene exchange. The results of 2 different approaches presented here support such an assumption. First, the results of the Evanno test based on ΔK (Figure 8) revealed that the optimum number of K is 2. Second, K-means clustering also showed that the best fit is for $K = 2$. It produced the highest value of pseudo-F (2.99; Table 4) and the lowest value of BIC (322.442) for $K = 2$.

In another attempt, we used the hybrid index that is based on the maximum likelihood method to calculate genetic affinity between *T. kotschyi* and either *T. tetragyna* or *T. karkalensis* (Table 5). We obtained a higher mean value of the hybrid index (0.5793) when *T. karkalensis* was considered as the reference population. We even had cases for *T. kotschyi* that showed 100% (h = 1) resemblance to members of *T. karkalensis* (tree no. 32 and tree no. 39 in Table 4). These results strongly support the STRUCTURE triangle presented earlier and also indicate the occurrence of gene flow between *T. karkalensis* and *T. kotschyi*.

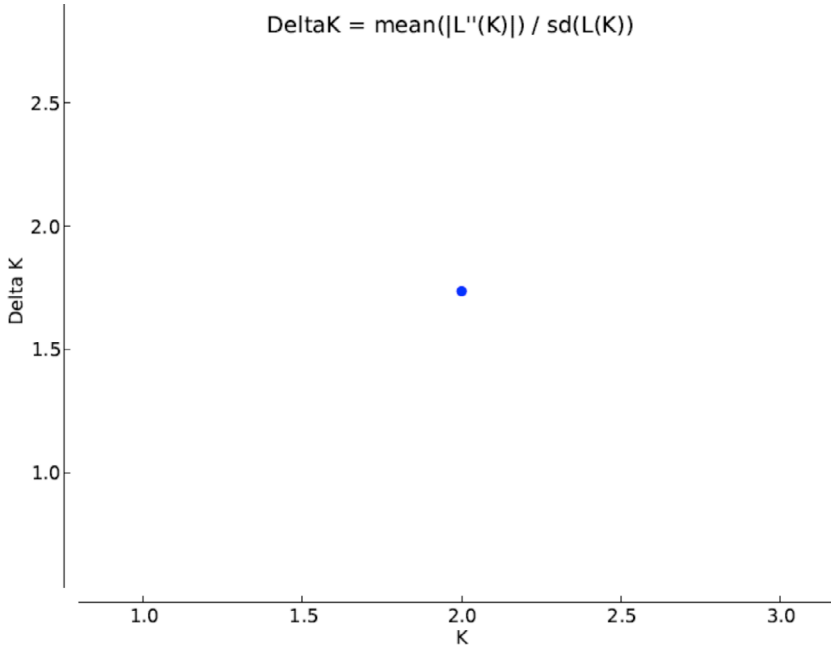


Figure 8. Delta *K* plot of Evanno test.

Table 4. K-means clustering statistics from K = 2 to K = 3.

K	SSD(T)	SSD(AC)	SSD(WC)	r-squared	pseudo-F	BIC	Rho
2&*	1941.643	135.032	1806.611	0.07	2.99	322.442	0.091
3	1941.643	244.451	1697.192	0.126	2.809	323.556	0.127

&: Best clustering according to Bayesian information criterion: K = 2.
*: Best clustering according to Caliński and Harabasz's pseudo-F: K = 2.
Best BIC clustering has been stored as clones.

4. Discussion

There is probably not another genus of plants as well known as the tamarisks in which the species are so poorly understood or separated based on more obscure characters (McClintock, 1951). For example, the distinction between *T. gallica*, *T. ramosissima*, and *T. chinensis* is based on differences in morphology of the nectary disk and staminal filaments that are often difficult to separate, as the traits are not clearly or unequivocally expressed (Allred, 2002).

In the present study we used a few morphometric characters that can be used in *Tamarix* species' delimitation. Morphological results obtained here not only differentiate the studied species but also show an affinity between *T. karkalensis* and *T. kotschyi*.

Tamarisk can tolerate an extreme range of environmental conditions, and Brotherson and Winkel (1986) suggested a general purpose genotype in salt cedar that gives it the capability to exploit a wide spectrum of habitats. Phenotypic plasticity, ecotypic differentiation, and high genetic variation suggest a high invasive potential.

Sexton et al. (2002) found no genetic differences between regions for most functional traits sampled in *Tamarix*. An exception was a regional genetic divergence (likely a result of multiple introductions) for root biomass investment in cold environments, indicating ecotypic differentiation and perhaps local adaptation in seedlings. Their results showed plasticity for all morphological and gas exchange traits sampled in *Tamarix* (Sexton et al., 2002).

Tamarisk flowers are mainly bisexual and Brotherson and Winkel (1986) suggested that *Tamarix* is cross-pollinated by wind. However, experiments by Stevens (1989) showed that no seed development occurs without insect visitation. In both conditions, this kind of cross-pollination brings about high genetic diversity and may lead to interspecific hybrid formation in *Tamarix*. For

Table 5. Hybrid index values for *T. kotschyi* individuals.

T. kotschyi individuals	Reference population: *T. karkalensis*	Reference population: *T. tetragyna*
	H	H
23	0.374	0.795
24	0.441	0.491
25	0.207	0.485
26	0.323	0.461
27	0.396	0.436
28	0.852	0.59
29	0.903	0.602
30	0.733	0.409
31	0.747	0.284
32	1	0.503
33	0.466	0.189
34	0.41	0.434
35	0.178	0.779
36	0.677	0.336
37	0.689	0.467
38	0.583	0.409
39	1	0.533
40	0.453	0.443
41	0.479	0.32
42	0.675	0.437
Mean	0.5793	0.47015

example, *Tamarix ramosissima* and *T. chinensis* differ slightly in sepal margin, petal shape, and filament insertion (Baum, 1967, 1978), and they have been considered to be synonyms (Allred, 2002). Recent molecular work has shown that although these species are genetically distinct in their native range, they form hybrids in the United States (Gaskin and Schaal, 2002).

Similarly, Gaskin and Shafroth (2005) reported relatively rare and localized hybrids between *T. aphylla* (L.) Karst. and both *Tamarix ramosissima* and *T. chinensis* (Gaskin and Shafroth, 2005). Close affinity observed between *T. karkalensis* and *T. kotschyi* both in morphological and genetic features may be due to gene exchange among them. This is particularly evidenced by the reticulation results, STRUCTURE analyses, and hybrid index values presented above.

Inbreeding and genetic drift can affect population fitness through the increased expression of recessive deleterious alleles as homozygosity increases in small populations. Genetic drift is expected to randomly reduce variation within small populations, causing loss of low frequency alleles, which can be associated with population fitness (Lande, 1999).

According to Gilpin and Soule (1986), these genetic factors, combined with demographic stochasticity, may result in extinction vortices in small populations. This process reduces the number of individuals until populations become extinct. However, other researchers indicated that, although the above discussion is true, not all small populations are genetically going through extinction vortices (e.g., Godt and Hamrick, 1998; Gitzendanner and Soltis, 2000). There are other factors such as species' life-history, biogeography, and gene flow into the population

that could also play critical roles in determining the current genetic composition of populations (Hamrick and Godt, 1996a, 1996b; Holsinger et al., 1999).

The results obtained by AMOVA and STRUCTURE plot clearly indicated high intraspecific genetic variation in the studied *Tamarix* species even within the limited area of investigation (60 km^2). This high genetic diversity may be used for local adaptation and also prevents homozygosity and genetic extinction of the studied *Tamarix* taxa.

It is interesting to mention that the studied *Tamarix* species could be separated from each other by morphometric analyses but were partially intermixed in ISSR analyses. This may be due to the fact that high genetic exchange occurs among these species and that the studied loci of the genome are not affecting morphological characters.

High degree of gene flow between neighboring trees even between different species prevents genetic distinction of species and, if this genetic admixture is accompanied with morphological admixture, brings about difficulties in *Tamarix* species delimitation and identification. This is the problem that we are facing today.

References

Allred KW (2002). Identification and taxonomy of *Tamarix* (Tamaricaceae) in New Mexico. Desert Plants 18: 26–29.

Baum BR (1967). Introduced and naturalized tamarisks in the United States and Canada [Tamaricaceae]. Baileya 15: 19–25.

Baum BR (1978). The Genus *Tamarix*. Jerusalem, Israel: Israel Academy of Sciences and Humanities.

Brotherson JD, Winkel V (1986). Habitat relationships of saltcedar (*Tamarix ramosissima*). The Great Basin Natural 46: 535–541.

Bryant D, Moulton V (2004). NeighborNet: an agglomerative algorithm for the construction of planar phylogenetic networks. Mol Biol Evol 21: 255–265.

Buerkle CA (2005). Maximum-likelihood estimation of a hybrid index based on molecular markers. Mol Ecol Notes 5: 684–687.

Caliński R, Harabasz J (1974). A dendrite method for cluster analysis. Commun Stat 3: 1–27.

Earl DA, von Holdt BM (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resources 4: 359–361.

Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14: 2611–2620.

Falush D, Stephens M, Pritchard JK (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol Ecol Notes 7: 574–578.

Friedman JM, Roelle JE, Cade BS (2012). Genetic and environmental influences on cold hardiness of native and introduced riparian trees. In: Haase DL, Pinto JR, Riley LE, technical coordinators. National Proceedings: Forest and Conservation Nursery Associations, 2011. Fort Collins, CO, USA: USDA Forest Service, Rocky Mountains Research Station, pp. 82–85.

Gaskin JF, Kazmer DJ (2009). Introgression between invasive saltcedars (*Tamarix chinensis* and *T. ramosissima*) in the USA. Biol Invasions 11: 1121–1130.

Gaskin JF, Schaal BA (2002) Hybrid *Tamarix* widespread in US invasion and undetected in native Asian range. P Natl Acad Sci USA 99:11256–11259.

Gaskin JF, Shafroth PB (2005) Hybridization of *Tamarix ramosissima* and *T. chinensis* (saltcedars) with *T. aphylla* (athel) (family Tamaricaceae) in the southwestern USA determined from DNA sequence data. Madrono 52: 1–10.

Gilpin ME, Soule ME (1986). Minimum variable populations: process of species extinction. In: Soule ME, editor. Conservation Biology: The Science of Scarcity and Diversity. Sunderland, MA, USA: Sinauer Associates, pp. 19–34.

Gitzendanner MA, Soltis PS (2000). Patterns of genetic variation in rare and widespread plant congeners. Am J Bot 87: 783–792.

Godt MJW, Hamrick JL (1998). Allozyme diversity in the endangered pitcher plant *Sarracenia rubra* ssp. *alabamensis* (Sarraceniaceae) and its close relative *S. rubra* ssp. *rubra*. Am J Bot 85: 802–810.

Hamrick JL, Godt MJW (1996a). Conservation genetics of endemic plant species. In: Avise JL, Hamrick JL, editors. Conservation Genetics Case Histories from Nature. London, UK: Chapman and Hall, pp. 281–304.

Hamrick JL, Godt MJW (1996b). Effects of life history traits on genetic diversity in plant species. Philos Trans R Soc Lond B 351: 1291–1298.

Hedrick PW (2005). A standardized genetic differentiation measure. Evol 59: 1633–1658.

Jost L (2008). GST and its relatives do not measure differentiation. Mol Ecol 17: 4015–4026.

Holsinger KE, Mason-Gamer RJ, Whitton J (1999). Genes, demes, and plant conservation. In: Landweber LF, Dobson AP, editors. Genetics and the Extinction of Species: DNA and the Conservation of Biodiversity. Princeton, NJ, USA: Princeton University Press, pp. 23–46.

Huson DH, Bryant D (2006). Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 23: 254-267.

Križman M, Jakse J, Baricevic D, Javornik B, Prosek M (2006) Robust CTAB-activated charcoal protocol for plant DNA extraction. Acta Agri Slov 87: 427–433.

Lande R (1999). Extinction risks from anthropogenic, ecological, and genetic factors. In: Landweber LF, Dobson AP. editors. Genetics and the Extinction of Species: DNA and the Conservation of Biodiversity. Princeton, NJ, USA: Princeton University Press, pp. 1–22.

Legendre P, Makarenkov V (2002). Reconstruction of biogeographic and evolutionary networks using reticulograms. Syst Biol 51: 199–216.

McClintock E (1951). Studies in California ornamental plants: 3. The tamarisks. J California Hort Soci 12: 76–83.

Meirmans PG, Van Tienderen PH (2004). GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. Mol Ecol Notes 4: 792–794.

Moerman DE (1998). Native American Ethno-Botany. Portland, OR, USA: Timber Press.

Parsons WT, Cuthbertson EG (1992). Noxious Weeds of Australia. Melbourne, Australia: Inkata Press, pp. 620–622.

Peakall R, Smouse PE (2006). GENALEX 6: Genetic Analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6: 288–295.

Podani J (2000). Introduction to the Exploration of Multivariate Data (English translation). Leiden, the Netherlands: Backhuys.

Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype Data. Genetics 155: 945–959.

Rodman J (1989). Reflections on tamarisk bashing. In: Hughes HG, Bonnicksen TM, editors. Papers from the First Annual Meeting of the Society for Ecological Restoration. Madison, WI, USA: Society for Ecological Restoration, pp. 59–68.

Schwarz GE (1978). Estimating the dimension of a model. Ann Stat 6: 461–464.

Sexton JP, McKay JK, Sala A (2002). Plasticity and genetic diversity may allow saltcedar to invade cold climates in North America. Ecol Appl 12: 1652–1660.

Stevens LE (1989) The status of ecological research on tamarisk (Tamaricaceae: *Tamarix ramosissima*) in Arizona. In: Kunzmann MR, Johnson RR, Bennett P, technical coordinators. Tamarisk Control in Southwestern United States: Proceedings, 2–3 September 1987. Tucson, AZ, USA: National Park Service, pp. 99–105.

Sudbrock A (1993). Tamarisk control-fighting back: an overview of the invasion, and a low-impact way of fighting it., Restor Manage Notes 11: 31–34.

Tykač J (1990). The Illustrated Guide to Ornamental Shrubs. London, UK: Treasure Press.

Weising K, Nybom H, Wolf K, Kahl G (2005). DNA Finger Printing in Plants. 2nd ed. Boca Raton, FL, USA: CRC Press.