

Genetic Diversity and Trait Association Analysis in the Economically Important Fiber-Producing Species *Calotropis procera* (Aiton) W.T. Aiton

Mahmod Eslami¹, Mohammadreza Kordasti^{2,3}, Somayeh Safiallah⁴, Masoud Sheidai^{*2} 

Received: 2026-06-09 Accepted: 2026-08-10

Abstract

Calotropis procera (Aiton) W.T.Aiton is an evergreen, perennial shrub of the family Apocynaceae which has been utilized in traditional medicinal systems in North Africa, the Middle East Asia, as well as Souteast Asia. We can also cite this plant importance as a fodder, and used for fuel purposes, timber and fiber production, phytoremediation, and synthesis of nanoparticles. From the economic point of view, this species is in demand for its pharmacological applications. Beside the above said applications, *C. procera* has also been used as an ornamental species. We have limited geographical distribution of this species in Iran and almost no report is available on genetic diversity, gene flow and populations' genetic structure of this precious plant in the country. We used SCoT molecular markers to conduct a population genetic study of *C. procera* and to generate novel data on the associations of genetic regions with geographical, landscape as well as important morphological characters that are related to fiber production. DCA analysis showed that SCoT loci are appropriate molecular markers for population genetic studies in *C. procera* beause these loci are scattered throughout the genome. Genetic diversity analyses revealed some private SCoT bands or loci in the geographical populations studied. A moderate level of genetic polymorphism was observed in these populations and DAPC revealed frequent and extensive gene flow among the studied populations, probably due to cross-pollination in these plants. Association studies identified several SCoT loci significantly associated with geographical and landscape variables. We also identified some genetic regions, which are closely associated with morphological characteristics related to fiber production, as well as loci associated with fiber color.

Keywords: *Calotropis procera* (Aiton) W.T. Aiton, CCA, Gene flow, LFMM, RDA, SCoT

Introduction

Calotropis procera (Aiton) W.T. Aiton, economically important plant species whose also known as the apple of Sodom, is an native range extends from Macaronesia

1-Department of Agricultural Engineering, National University of Skill(NUS), Tehran, Iran

2-Faculty of Life Sciences & Biotechnology, Shahid Beheshti University, Tehran, Iran

3-Department of Biology Education, Farhangian University, P.O. Box 14665-889, Tehran, Iran

4-Department of Agricultural Science, National University of Skill(NUS), Tehran, Iran

*Corresponding Author email address: msheidai@sbu.ac.ir

Doi: [10.48308/pae.2026.244400.1144](https://doi.org/10.48308/pae.2026.244400.1144)



Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

and tropical Africa to Indo-China (Kaur et al., 2021). This evergreen perennial shrub belongs to the family Apocynaceae and has long been utilized in traditional medicinal systems of North Africa, the Middle East, and Southeast Asia. The species contains several biologically active compounds and has been investigated for various pharmacological properties. However, the species is also recognized as a toxic plant due to the presence of cardiac glycosides and other secondary metabolites. In addition to its reported medicinal uses, *C. procera* has attracted attention because of its applications in fiber production, phytoremediation, nanoparticle synthesis, fuel production, and as an ornamental plant.

C. procera is an evergreen shrub or small tree that usually reaches a height of 2.5–6 m, and possesses a deep taproot system. Young stems are grayish-green, while mature stems develop fissured bark. The species is characterized by large opposite succulent leaves covered with a waxy layer, abundant milky latex, and attractive white-to-purple flowers arranged in umbellate inflorescences. The fruits are inflated follicles containing numerous seeds bearing silky hairs that facilitate long-distance dispersal by wind (Kaur et al., 2021). *C. procera* is mainly distributed in the southern and southeastern parts in Iran, particularly along the Persian Gulf and Oman Sea Coasts. Natural populations have been reported from Hormozgan, Bushehr, Khuzestan, Kerman, and Sistan and Baluchestan provinces, where the species occurs in arid and semi-arid habitats, coastal plains, sandy soils, and disturbed areas (Mozaffarian, 1996; Assadi

et al., 2008; Kaur et al., 2021).

In addition to its ecological and economic importance, *C. procera* has considerable ethnobotanical significance across several regions, particularly in arid and semi-arid areas of the Middle East and South Asia, including Iran. Ethnobotanical surveys have reported the traditional use of different parts of this species in folk medicine for the treatment of skin diseases, inflammation, rheumatism, as well as for topical treatment of various ailments, despite its known toxicity (Ghisalberti, 2004; Khan et al., 2015). In southern Iran, local communities have also been documented to use this plant in traditional practices, reflecting its long-standing role in local knowledge systems. However, due to the presence of potent cardiac glycosides, the application of this species requires caution and is generally limited to traditional, non-standardized applications (Kaur et al., 2021).

C. procera is well developed to arid and semi-arid climates, and can tolerate very low annual rainfall (150 mm) and dry seasons lasting up to 10 months. Mean annual monthly temperatures where *C. procera* is found are generally in the range 20–30°C. *C. procera* is not frost tolerant and can grow in a wide variety of soil types, including alkaline and saline soils, though it prefers well-drained sandy soils. It occurs across a broad altitudinal range, from exposed coastal sites to medium elevations up to 1300 m (Kaur et al., 2021).

Due to morpho-physiological adaptations and ability to tolerate various abiotic stresses, *C. procera* has naturalized beyond its introduced range and is considered

an invasive weed in several parts of the world, particularly in pasture, overgrazed rangelands, and poorly managed hay fields. This taxon can compete with desirable pasture species and form dense thickets that interfere with livestock management, particularly mustering activities. Moreover, *C. procera* is known to contain *Calotropis* which is a cardiac poison, and for this, the plant is a main suspect for death of sheep and goats in West Africa (El-Badwi et al., 1998). This species is complicated to control and eradicate, mainly because its seeds are wind-dispersed and its plants are fire resistant. In addition to wind pollination, the carpenter bees *Xylocopa pubescens* and *Xylocopa fenestrata* visit this weed and serve as the main pollinators, while, the workers of *Apis florea* act as its minor pollinators (Ali and Ali, 1989). *C. procera* produces flowers and fruits throughout the year (Little et al., 1974), and each plants may produce hundreds to thousands of seeds annually. The seeds exhibited a high germination rate (89%), with germination occurring within 7-64 days after sowing in a potting mix (Francis, 2002). These characteristics facilitate the expansion of this species within its habitat place and further invasion of new areas.

C. procera is highly similar to the closely related species *C. gigantea*, and the two species are often confused. However, *C. gigantea* grows taller (4.5 m and above) and has smaller leaves borne on thicker leaf-bearing stems up, which can reach to 6 cm in diameter, compared with approximately 1 cm in *C. procera* (Crothers and Newbound, 1998).

The flowers of *C. gigantea* are more open

and have petals with a uniform mauve-pink coloration. In Iran, *C. procera* is restricted to southern part of the country, where it occurs in a limited number of geographically distinct populations. Given the ecological and economic importance and invasive characteristics of *C. procera*, as well as its limited distribution in Iran, we conducted a population genetic study of this species using SCoT (Start Codon Targeted) molecular markers. SCoT markers have previously been reported as an appropriate molecular marker for population genetic studies of this species (Muriira et al., 2018). These markers preferentially target candidate functional genes and are based on short conserved sequences surrounding the start codons of plant genes. Consequently, SCoT markers offer several advantages, including no requirement for prior genomic information, high reproducibility, and high throughput (Zhang et al., 2019).

Identifying putatively adaptive genetic variation is a key objective in landscape genetic studies. Single nucleotide polymorphisms (SNPs), which represent single base-pair variations in the genome, along with other molecular markers such as SSRs and SCoTs, are widely used to detect genetically differentiated loci (outlier loci) that may be subjected to natural selection. These approaches also enable the assessment of associations between genetic variation and environmental variables while accounting for neutral evolutionary processes, including population genetic structure and demographic history (Zhang et al., 2019; Garot et al., 2021). In this study, we focused on SCoT molecular

markers to investigate genetic diversity and their associations with morphological, geographical, and environmental variables in *C. procera*. Multivariate analytical approaches, including RDA (redundancy analysis), DAPC (discriminant analysis of principal components), and LFMM (latent factor mixed models), were applied to identify potential patterns of genetic structure and association (Collins and Jombart, 2015; Zhang et al., 2019).

Material and methods

Plant sample studies

In total, 80 plants from 12 geographical populations were studied for both morphological and molecular investigations (Table 1). Plant specimens were identified based on morphological characteristics and available taxonomic references by Mr. Mohammadreza Kordasti. Voucher specimens of all studied populations were deposited in the Herbarium of Shahid Beheshti University (HSBU), Tehran, Iran. The corresponding herbarium accession

numbers are provided in Table 1.

Morphological studies

Twelve morphological characters related to leaf, stem, fruit, and seed morphology were evaluated in the studied populations. These characters were selected based on their taxonomic significance, observed variability in *C. procera*, and their potential relevance to fiber production and seed dispersal, as determined by previous morphological studies of the species (Kaur et al., 2021; Khadivi and Mirheidari, 2025). These characters included: leaf base (funnel or heart), leaf length-to-width ratio (L/W), leafstalk (visible or gone), leaf thickness (thin or thick), internode size, peduncle size, plant height, follicle size, seed color (light brown or dark brown), seed size, popup color (shiny white or creamy), and popup seed size.

Morphological observations and measurements were performed on all 80 sampled individuals representing the 12 geographical populations. The subsequent analyses were restricted to mature and

Table 1. *Calotropis procera* populations and their locality

No	POP	Habit	soil type	Humidity	Temperature	Voucher number (Herbarium: HSBU)
1	Systan, Pashmak	Shurbs	Sandy- Clay	50%	40-28	HSBU-2022101
2	Boshehr, Ab pakhsh	Tree	Sandy- Clay	25%	50-38	HSBU-2022102
3	Kerman, Kahnoj	Shurbs	Sandy- Clay	31%	45-30	HSBU-2022103
4	Hormozgan, Bandar jask	Shurbs	Sandy	70%	37-33	HSBU-2022104
5	Khozestan, Dezfol	Shurbs	Sandy	18%	45-23	HSBU-2022105
6	Hormozgan, Minab	Shurbs	Sandy	40%	45-30	HSBU-2022106
7	Kerman, Bam	Tree	Sandy- Clay	33%	44-24	HSBU-2022107
8	Khozestan, Shoahtar	Shurbs	Sandy	15%	50-28	HSBU-2022108
9	Fars, Girkarzin	Shurbs	Sandy- Clay	33%	44-24	HSBU-2022109
10	Boshehr, Samal	Tree	Sandy- Clay	20%	50-35	HSBU-2022101
11	Hormozgan, Goarzonak	Shurbs	Sandy	40%	43-30	HSBU-2022110
12	Sistan, Chabahar	Shurbs	Sandy- Clay	65%	38-35	HSBU-2022112

healthy individuals. Qualitative and quantitative characters were recorded using a stereomicroscope and graph paper following standard botanical procedures. The obtained data were subsequently used for statistical analyses of morphological variation among populations.

Molecular studies

For SCoT molecular analyses, fresh leaf samples were collected and immediately dried in silica gel powder. Genomic DNA was extracted using the CTAB-activated charcoal protocol described by Sheidai et al. (2019). The quality of the extracted DNA was assessed by electrophoresis on a 0.8% agarose gel.

A preliminary screening was carried out using several SCoT primers developed by Collard and Mackill (2009). Based on amplification quality, band sharpness, reproducibility, and the degree of polymorphism obtained, three primers (SCoT1: CAACAATGGCTACCACCA, SCoT2: CAACAATGGCTACCACCC, and SCoT41: CAATGGCTACCACTGACA) were selected for further analyses. These primers produced clear, reproducible, and polymorphic banding patterns suitable for genetic diversity and association studies in *C. procera*.

PCR amplifications were performed using a ready-mix master mix (Parts Tools Co., Iran). Each reaction contained 20 ng of genomic DNA, 1 U of Taq DNA polymerase, and other reaction components in a final volume of 25 µl.

The amplification reactions were conducted in a Bio-Rad thermocycler (USA) under the following conditions; an initial denaturation

step of 4 min at 95°C, followed by 45 cycles of 1 min at 94°C, 1 min at the primer-specific annealing temperature (49°C for SCoT1, 56°C for SCoT2, and 53°C for SCoT41), and 1 min at 72°C. A final extension step of 7 min at 72°C was performed.

The amplification products were separated on 2% agarose gel and stained with PowerLoad (PowerLoad, Kosar Co., Iran). Fragment sizes were estimated using a 100 bp DNA ladder (Fermentas, Germany).

Data analysis

Morphological analyses

Standardized data (mean= 0, variance= 1) were used for multivariate statistical analyses. Morphological data analyses were performed in PAST software version 4.03 (Hammer et al., 2001). A heat-map based on UPGMA (Unweighted Pair Group Method with Arithmetic Mean) was constructed to study morphological divergence among individuals and their relationships with geographical and landscape features.

All statistical analyses were performed using individual-based morphological data, whereas population-level summaries were used exclusively for descriptive purposes and were excluded in multivariate analyses. A Mantel test was conducted to examine whether morphological distance is associated with geographical distance among the studied populations (Mantel, 1967; Sokal and Rohlf, 1995). Pearson correlation coefficients were also calculated among variables.

Genetic diversity analyses

The SCoT bands obtained were treated as binary characters (Podani, 2000) and used for further analyses. Detrended correspondence

analysis (DCA) was performed to evaluate the distribution pattern of SCoT bands obtained for population genetic analysis (Hill and Gauch, 1980). Similarly, discriminant power of bands was determined by the DAPC method, which is appropriate for binary data (Jompard et al., 2010), as well as Nei's G_{st} value as determined in the POPGENE program. Genetic diversity parameters, including the number of effective alleles (Ne), Shannon index (I), Nei's genetic diversity (H_e), unbiased H_e (U_{He}), and percentage of polymorphism (P%), were determined in the studied populations. Genetic differentiation of the populations was studied by two different approaches. First, we used UPGMA (Unweighted Paired Group using average) cluster analysis to show if the replicates of each population are separated from the others. AMOVA (Analysis of molecular variance) was then performed using GeneAlex software to partition molecular variation among and within populations. Nei's genetic distance was also being determined among the studied populations, and populations grouping was subsequently performed based on DAPC analysis as well as heat map construction based on the Jaccard similarity index in R package 4.1. DAPC analysis was used to identify discriminating SCoT loci that differentiate the studied populations of *C. procera*, and an assignment test was performed to assess gene flow among the studied plants.

Association studies

Association analysis was performed using multiple approaches, from Chi-square and Fisher's exact test to DAPC and Bayesian-

based LFMM (Collins and Jombart, 2015; Frichot et al., 2013), combined to study associations between genetic data and both morphological and landscape data. LFMM algorithms correct for unobserved confounders, improving power and controlling false discoveries, with LFMM_Ridge and LFMM_Lasso functions based on regularized least-squares solutions. The R package LFMM fits latent factor mixed models for genome-wide (GW), genome-environment (GE), and epigenome-wide (EW) association studies. Due to complex interactions between genetic structure and environmental variables, LFMM analysis used PCA followed by Bonferroni and false discovery rate (FDR) corrections to avoid type I and II errors, in R 4.1. DAPC accuracy was assessed by cross-validation in R 4.1. CCA and RDA were used to identify the contribution of environmental variables to population genetic structure. CCA, the canonical form of correspondence analysis (CA), is a direct gradient analysis where explanatory variables intervene in the CA solution, representing only explainable correspondence. It suits unimodal response variables, preserves chi-squared (χ^2) distances, and identifies response variables maximally related to linear combinations of explanatory variables, their importance determined by correlation to the final ordination (Greenacre, 2007). RDA, used in landscape genomics, identifies how spatial and environmental factors structure genetic variation among populations, modeling linear relationships among environment predictors and genomic variation. Integrating ordination and multiple regression, it

represents data matrix variation in reduced dimensions via multiple regressions through *principal components analysis* (Legendre and Legendre, 2012; Segovia et al., 2020).

Dispersal versus vicariance

Reconstruction of ancestral states (RASP) analysis was performed to investigate dispersal versus vicariance of SCoT loci in a population geographical context as implemented in RASP software. Statistical Dispersal–Vicariance Analysis (S-DIVA), as implemented in RASP software, was used to infer the historical biogeographic relationships among the populations based on a likelihood-based approach. (Yu et al., 2015).

Results

Morphological variation and population relationships

UPGMA dendrogram (Fig. 1) based on morphological data revealed that the studied populations have a high degree of morphological similarity which formed different clusters at approximately 0.15 Euclidean distance. However, some of the studied populations

show a higher degree of within-population similarity as the plant samples from these populations were closely grouped. This holds true for populations 9 and 10. The other populations show an admixture of plant samples. PCA analysis of morphological features, revealed that morphological similarity among populations 9 and 10 was mainly due to leaf thickness, while similarity of populations 1, 4, and 6, is due to the size of internode.

Figure 2 illustrates clustering patterns based on combined datasets, showing morphological similarity and admixture among populations, with weak and non-significant association with geographical and environmental variables. Heatmap constructed from combined morphological, geographical, and landscape data revealed that morphological characters such as internode size, plant height, and leaf length-to-width ratio showed similar patterns in grouping of the studied populations. Similar patterns were also observed for peduncle size, temperature, leaf thickness, leaf base, and leaf stalk (Fig. 2). However, other geographical and landscape variables showed weak and non-significant

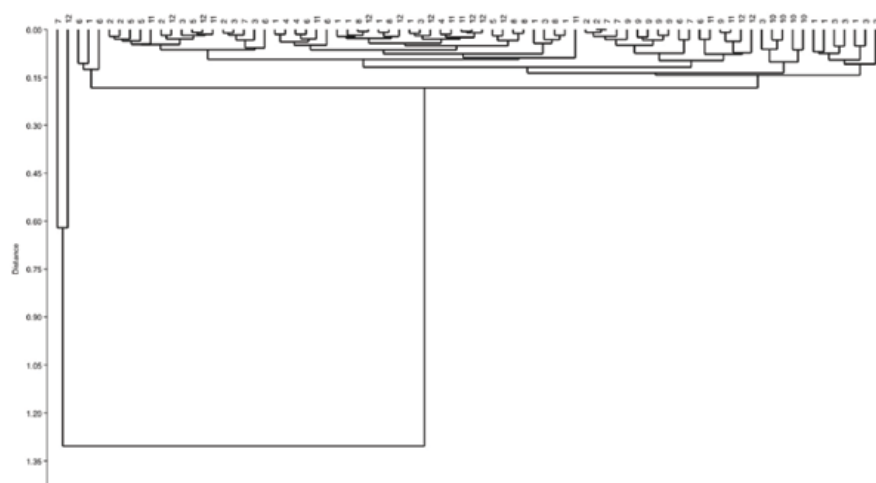


Fig. 1. UPGMA dendrogram of *Calotropis procera* populations based on morphological characteristics

associations with morphological grouping, indicating a generally mixed pattern of relationships. These findings were further supported by the non-significant Mantel test between morphological and ecogeographical distances among populations ($R = 0.038$, $p = 0.33$).

Genetic diversity analyses

DCA plot (Fig. 3), revealed that SCot molecular markers are scattered along the plot and therefore are distributed throughout the genome and are suitable for population genetic studies. We obtained 39 distinct and stable SCoT loci for population genetic analyses.

Genetic diversity parameters determined in

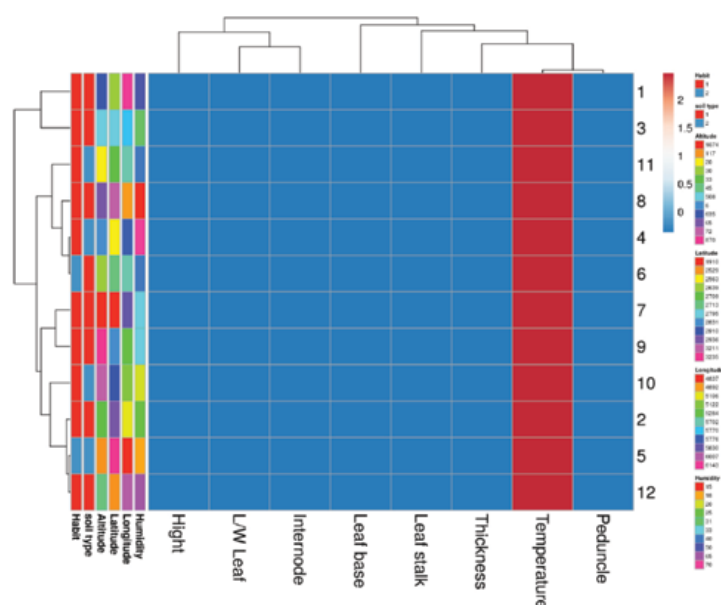


Fig. 2. Heatmap of morphological, geographical, and landscape variables in studied *Calotropis procera* populations.

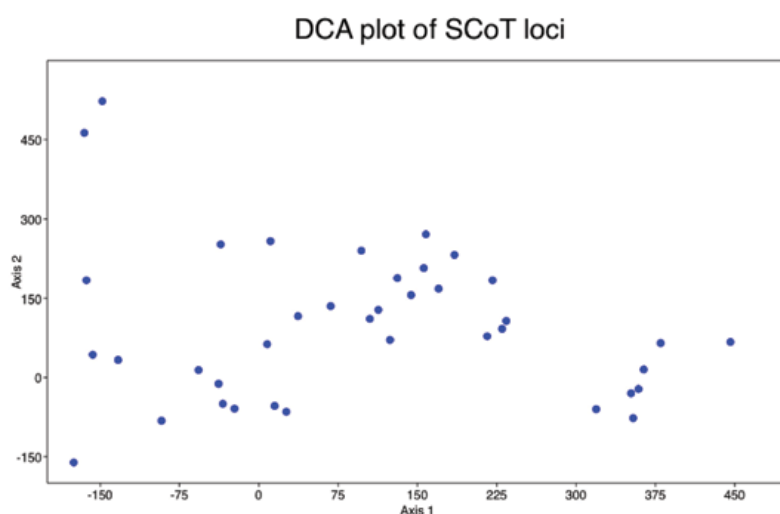


Fig. 3. DCA plot of SCoT loci obtained in *Calotropis procera* populations

the studied populations are provided in Table 2. Polymorphism percentages varied from 25% in population 4 to 64% in population 9. The moderate genetic polymorphism observed in these populations may be due to cross pollination behavior of this species. The mean gene diversity (H_e), varied from 0.09 to 0.191 in the same populations. This low gene diversity may be related to high magnitude of gene flow among geographical populations, as will be covered in coming paragraphs. AMOVA-paired revealed significant genetic difference ($P < 0.01$), in most of the studied populations, which shows inspire of frequent gene flow among populations, they still have a mean genetic distinctness which differentiate them from each other. This is further supported by

overall AMOVA result which revealed that 70% of genetic variance is due to within population genetic variability and 30% due to among population genetic difference.

Nei genetic distance obtained among populations ranged from 0.02 to 0.22 only (Table 3). This is due to high level of genetic similarity in the studied populations (> 0.87). We obtained a significant correlation between genetic distance and geographical distance of the studied populations ($r = 0.15$, $p < 0.01$), by Mantel test.

Genetic grouping and gene flow analysis

The mean G_{st} value of SCoT loci versus number of migration (N_m value), obtained by pop gene, produced $G_{st} = 0.3688$, and $N_m = 0.85$, which show a high magnitude of gene flow for SCoT loci. This is further

Table 2. Genetic diversity parameters in *Calotropis procera* populations

Pop	Na	Ne	I	He	uHe	%P
Pop1	1.333	1.368	0.339	0.223	0.233	66.67%
Pop2	1.103	1.342	0.283	0.192	0.209	51.28%
Pop3	1.179	1.310	0.277	0.184	0.194	53.85%
Pop4	0.795	1.150	0.138	0.091	0.110	25.64%
Pop5	0.897	1.248	0.227	0.150	0.171	43.59%
Pop6	1.154	1.372	0.312	0.212	0.228	56.41%
Pop7	0.744	1.257	0.208	0.143	0.159	35.90%
Pop8	1.051	1.279	0.255	0.167	0.186	51.28%
Pop9	1.282	1.316	0.296	0.191	0.209	64.10%
Pop10	0.513	1.155	0.130	0.088	0.101	23.08%
Pop11	0.718	1.185	0.173	0.113	0.120	35.90%
Pop12	1.128	1.300	0.266	0.175	0.184	56.41%

a. N: No of plants studied, b. Na: No. of alleles, c. Ne: Effective No. of alleles, d. He: Gene diversity, e. UHe: Unbiased gene diversity, f. %P: Polymorphism percentage

Nei genetic distance obtained among populations ranged from 0.02 to 0.22 only (Table 3). This is due to high level of genetic similarity in the studied populations (> 0.87). We obtained a significant correlation between genetic distance and geographical distance of the studied populations ($r = 0.15$, $p < 0.01$), by Mantel test.

Table 3. Nei genetic distance among *Calotropis procera* populations (populations 1-12 are according to Table 1)

Pop1	Pop2	Pop3	Pop4	Pop5	Pop6	Pop7	Pop8	Pop9	Pop10	Pop11	Pop12	
0.000											Pop1	
0.106	0.000										Pop2	
0.102	0.163	0.000									Pop3	
0.223	0.222	0.285	0.000								Pop4	
0.066	0.140	0.147	0.310	0.000							Pop5	
0.069	0.107	0.105	0.259	0.146	0.000						Pop6	
0.055	0.154	0.197	0.367	0.086	0.131	0.000					Pop7	
0.068	0.115	0.188	0.332	0.066	0.079	0.088	0.000				Pop8	
0.035	0.132	0.108	0.335	0.060	0.085	0.069	0.046	0.000			Pop9	
0.107	0.147	0.184	0.296	0.068	0.167	0.135	0.144	0.123	0.000		Pop10	
0.029	0.145	0.141	0.270	0.046	0.110	0.063	0.079	0.038	0.069	0.000	Pop11	
0.044	0.150	0.154	0.303	0.045	0.143	0.061	0.087	0.057	0.051	0.037	0.000	Pop12

supported by DAPC assignment analysis. Genetic grouping of the studied populations (Figure 4.I), differentiated them in three main groups. Assignment test of plants by

DAPC placed them in the three main genetic groups and also revealed a high level of gene flow and admixture among the studied populations (Figure 4.II).

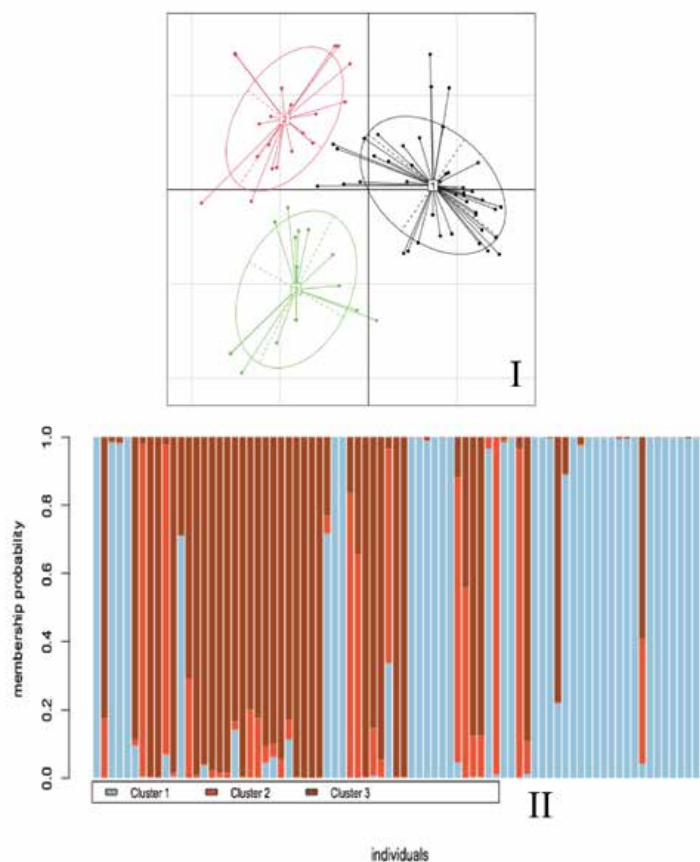


Fig. 4. I. DAPC plot of populations genetic groups in *Calotropis procera* plants studied, II. Assignment and admixture plot of *Calotropis procera* plants in three main genetic groups (clusters)

Dispersal versus vicariance

RASP analysis revealed that some of the populations were less involved in migration and gene flow as most of their plants were grouped together in an almost distinct clade. For example, populations 2, 3, 10 and 12 (Figure 5), show less degree of vicariance compared to the other studied populations. The reason for this is not known to us and it may be due to either presence of more severe ecological and environmental barriers they encounter, or shortage of frequent pollinating visitors in their geographical area.

*Association studies**SCoT loci and geographical variables*

Canonical correspondence analysis and RDA plots produced similar results and both produced significant association ($p < 0.01$) between SCoT loci and geographical variables. Here we provide only RDA plot to be brief (Figure 6). Some of the SCoT loci were closely associated with geographical variables. For example, SCoT loci. 24 and 31 are close to both longitude and altitude, while SCoT loci 2, 4, and 6, are closely associated to humidity. Loci 21 and 22

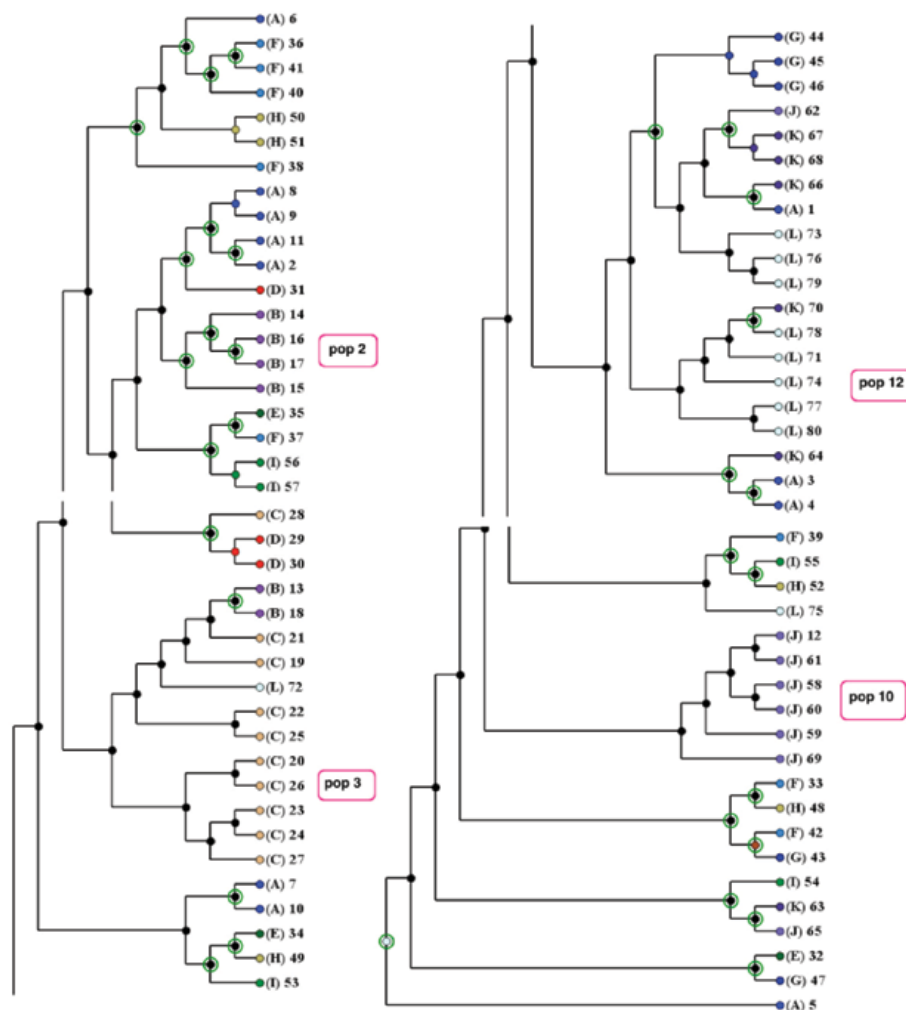


Fig. 5. RASP phylogenetic tree of *Calotropis procera* populations constructed by S-DIVA approach

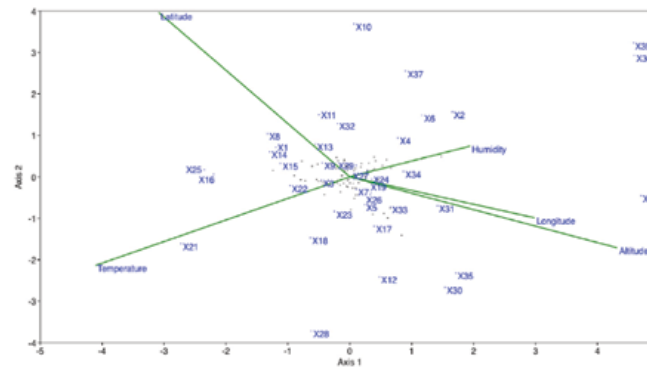


Fig. 6. RDA plot of SCoT loci against geographical variables showing close association between some of the loci and variables studied.

showed close association with temperature.

SCoT loci and soil type

Chi square analysis followed by Fishers' exact test produced significant association between SCoT loci and soil type (X-squared = 7.1, df = 2, p-value = 0.02) (Table 1). Similarly, DAPC analysis cross-validation (Figure 7-a), indicated that the model of association produced is efficient and

identified three SCoT loci number 1, 3, 10, and 30, are significantly associated with the soil type habitat of plants. Ridge and Lasso models of LFMM analysis of the same data (Figure 7-b), followed by Bonferroni correction and false discovery rate test, also identified SCoT loci. No. 3, and 10, as well as 11 and 15, are significantly associated with the soil type.

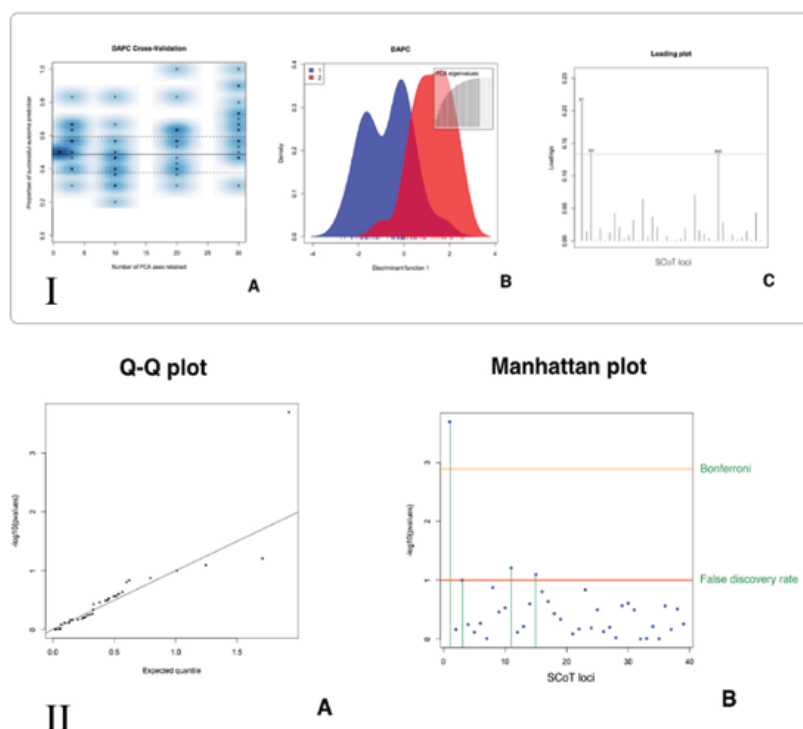


Fig. 7. I. DAPC association analyses of SCoT loci and the soil type, II. Q-Q plot and Manhattan plot of LFMM analysis showing significantly associated SCoT loci with the soil type.

SCoT loci and habit

Plants studied had two types of growth forms (Table 1). Although Fishers' exact test did not produce significant result, both DAPC and LFMM analyses showed significant association between some of the SCoT loci No. 7, 9, 11, 14, 15, 16, and 24, and habit form of the studied plants.

Association studies between SCoT loci and morphological characters

Some of morphological characters in *C. procera* are associated with fiber production. These are the size of internode, peduncle, seed size, and seed pod (follicle) which when dehiscent releases seeds with silky white popups. Moreover, detailed observation of silky papas showed two types of colors, 1- shiny white and 2-creamy, which may be of economic importance. Significant association between genetic regions and these important morphological

characteristics are of importance in future selection and breeding of this precious fiber producing plant species. Different methods used showed significant association between SCoT loci and significant morphological characters assumed to be related to fiber production. Details of SCoT loci associated with these characteristics are summarized and presented in Table 4.

Similar analysis for fiber color also produced significant association between SCoT loci and this interesting characteristics. As shown in Table 4, several significant associations were identified between SCoT loci and seed-related traits. Two SCoT loci (No. 2 and 26) were significantly associated with seed cob traits, whereas loci No. 14 and 23 showed significant associations with seed pod size. These characters are directly related to the fiber production and few SCoT loci are associated with them which can

Table 4. Association results of SCoT loci with morphological characters, geographical and landscape variables

[illegible]

be characterized and sequenced for future breeding and selection purposes. Some other interesting associations are between SCoT loci No. 18 and size of leaf stalk, SCoT loci No. 20 and 21 are related to the size of internode. These characters are also related to the size of plant and fiber production via more seed production. SCoT loci associated with the soil type and temperature, would also be important as they may contribute to the ecological and landscape tolerance for this plant species.

Discussion

Genetic and morphological diversity

This study reports morphological and genetic variability of *C. procera* in different localities of Iran. The studied populations showed both morphological and genetic admixture due to extensive gene flow, probably caused by cross-pollination. SCoT loci were efficient molecular markers for population genetic studies and germplasm evaluation of *C. procera*.

A moderate level of genetic polymorphism was obtained. Genetic diversity in *C. procera* has been analyzed using SSR and ISSR markers (Majeed et al. 2020; Navabpour et al., 2021), both showing moderate polymorphism (mean PIC = 0.575; Na = 3.625, Ho = 0.58 for SSR; ~36% for ISSRs), agreeing with our results. A significant correlation was found between genetic and geographical distance ($r = 0.15$, $p < 0.01$) by Mantel test, consistent with Majeed et al. (2020) ($r = 0.147$, $P = 0.010$).

AMOVA also showed significant genetic differences among populations. Genetically diverse populations can serve

as an important resource for germplasm collection in breeding programs (Majeed et al., 2020; Navabpour et al., 2021). Ecological grouping based on combined genetic and environmental data revealed that populations tend to cluster partly according to geographical origin and environmental conditions. Although gene flow resulted in considerable genetic admixture, certain populations showed partial structuring reflecting adaptation to local conditions such as temperature, humidity, and habitat type. The weak but significant correlation between genetic and geographic distances supports isolation by distance in shaping population structure.

Overall, while *C. procera* populations in Iran are not strictly ecologically segregated, environmental heterogeneity may still contribute to subtle local differentiation. Both neutral processes (e.g., gene flow) and selective forces associated with habitat variability likely act simultaneously on the genetic structure of this species.

Association studies

This study identified genetic regions by SCoT loci significantly correlated with geographical and landscape variables as well as with important morphological traits related to fiber production, including fiber color in *C. procera*. These are novel findings. Different approaches used to identify genetic loci associated with geographical and ecological variables were all in agreement, improving understanding of associated SNPs or genetic regions and giving insights into contemporary evolutionary processes (Segovia et al., 2020).

Heterogeneous environmental conditions

bring about changes in genetic diversity, resulting in local adaptations (Zhang et al., 2019; Garot et al. 2021, Segovia et al. 2020). Studies on the genetic basis of local adaptation and identification of adaptive genetic loci or SNPs can improve knowledge of the genetic mechanism of local adaptation and species diversification within a genus (Zhang et al. 2019). *C. procera* is emerging as a new, yet undomesticated, fiber resource comparable to cotton and kapok (Majeed et al., 2020). Therefore, screening efficient genotypes from wild populations and identifying genetic regions associated with fiber production are important in the pre-domestication process. Desired genotypes can then be improved through conventional breeding. Identification of important genes is an important task in genetic studies of *C. procera*. For example, Ramadan et al. (2014) identified the Δ -pyrroline-5-carboxylate synthase (P5CS) gene, which synthesizes proline, a compatible solute that accumulates widely in plants to tolerate unfavorable conditions. Similarly, Shokry et al. (2014) detected a Usp-like gene in *C. procera*. The universal stress protein (USP) domain provides organisms with the ability to respond to environmental stresses. Mutwakil et al. (2017) studied *C. procera* response to drought and osmotic stress using polyethylene glycol and NaCl, with transcriptomic and metabolomics measurements across a five-day time-course, identifying a dynamic transcriptional response including changes in inositol signaling, stress response genes, and cytokinins, along with sharp increases of myo-inositol and elevated amino acid

metabolites at later times.

Based on evaluated morphological traits related to fiber production, certain populations exhibited superior performance in key economically important characteristics such as internode length, plant height, and fiber-related seed and leaf traits. These populations may be considered potential candidate genetic resources for future breeding and domestication of *C. procera*. However, given the observed genetic admixture and absence of strong population-level structuring, these superior traits should be interpreted cautiously and further validated through multi-environmental and controlled breeding experiments before formal selection of elite genotypes.

In conclusion, we provided novel findings on genetic structure, gene flow as well as association of genetic regions to the geographical, land scape and important morphological characters related to fiber production in *C. procera*.

Funding

None

Conflicts of interest:

All authors have no conflict of interest

Availability of data and material

All data are available from the corresponding author on reasonable request.

Code availability

NONE

References

- Ali, T. and Ali, S.I. (1989) 'Pollination biology of *Calotropis procera* subsp. *hamiltonii* (Asclepiadaceae)', *Phyton (Horn)*, 29(2), pp. 175–188.

- Assadi, M., Maassoumi, A.A., Khatamsaz, M. and Mozaffarian, V. (eds.) (2008) *Flora of Iran*. Tehran: Research Institute of Forests and Rangelands.
- Collins, C. and Jombart, T. (2015) *A tutorial for discriminant analysis of principal components (DAPC) using adegenet*. London: Imperial College London.
- Crothers, M. and Newbound, S. (1998) *Rubber Bush (Calotropis procera)*. Northern Territory of Australia, Australia.
- El-Badwi, S.M.A., Adam, S.E.I., Shigidi, M.T. and Hapke, H.J. (1998) ‘Studies on laticiferous plants: toxic effects in goats on *Calotropis procera* latex given by different routes of administration’, *Deutsche Tierärztliche Wochenschrift*, 105(11), pp. 425–427.
- Francis, J.K. (2002) ‘*Calotropis procera* (Ait.) Ait.’, in Francis, J.K. (ed.) *Wildland shrubs of the United States and its territories: thamnic descriptions*. General Technical Report IITF-WB-1. U.S. Department of Agriculture, Forest Service.
- Frichot, E., Schoville, S.D., Bouchard, G. and François, O. (2013) ‘Testing for associations between loci and environmental gradients using latent factor mixed models’, *Molecular Biology and Evolution*, 30(7), pp. 1687–1699. <https://doi.org/10.1093/molbev/mst063>.
- Garot, E., Dussert, S., Domergue, F., Joët, T., Fock-Bastide, I., Combes, M.C. and Lashermes, P. (2021) ‘Multi-approach analysis reveals local adaptation in a widespread forest tree of Reunion Island’, *Plant and Cell Physiology*, 62(2), pp. 280–292. <https://doi.org/10.1093/pcp/pcaa160>.
- Ghisalberti, E.L. (2004) ‘The medicinal plant *Calotropis procera*’, *Fitoterapia*, 75(4), pp. 310–337.
- Greenacre, M. (2007) *Correspondence Analysis in Practice*. 2nd edn. Boca Raton, FL: Chapman & Hall/CRC.
- Hammer, Ø., Harper, D.A.T. and Ryan, P.D. (2001) ‘PAST: paleontological statistics software package for education and data analysis’, *Palaeontologia Electronica*, 4(1), pp. 1–9.
- Hill, M.O. and Gauch, H.G. (1980) ‘Detrended correspondence analysis: an improved ordination technique’, *Vegetatio*, 42, pp. 47–58. <https://doi.org/10.1007/BF00048870>.
- Jombart, T., Devillard, S. and Balloux, F. (2010) ‘Discriminant analysis of principal components: a new method for the analysis of genetically structured populations’, *BMC Genetics*, 11, 94. <https://doi.org/10.1186/1471-2156-11-94>.
- Kaur, A., Batish, D.R., Kaur, S. and Chauhan, B.S. (2021) ‘An overview of the characteristics and potential of *Calotropis procera* from botanical, ecological, and economic perspectives’, *Frontiers in Plant Science*, 12, 690806. <https://doi.org/10.3389/fpls.2021.690806>.
- Khadivi, A., Mirheidari, F., Tunç, Y., Saeidifar, A., Moradi, Y., Rabiei, V. and Razavi, F. (2026) ‘Morphological characterization of *Calotropis procera* (Aiton) W.T. Aiton, a neglected medicinal plant’, *PLOS ONE*, 21(2), e0341425. <https://doi.org/10.1371/journal.pone.0341425>.
- Khan, M.A. et al. (2015) ‘Ethnobotanical uses of medicinal plants in arid regions’,

- Journal of Ethnopharmacology*, 173, pp. 1–15.
- Legendre, P. and Legendre, L. (2012) *Numerical Ecology*. 3rd edn. Amsterdam: Elsevier.
- Little, E.L. Jr, Woodbury, R.O. and Wadsworth, F.H. (1974) *Trees of Puerto Rico and the Virgin Islands*. Washington, DC: U.S. Department of Agriculture, Forest Service.
- Majeed, A., Goel, B., Mishra, V., Kohli, R. and Bhardwaj, P. (2020) ‘Elucidation of genetic diversity base in *Calotropis procera* – a potentially emerging new fibre resource’, *Plant Genetic Resources*, 18(3), pp. 159–167. <https://doi.org/10.1017/S1479262120000167>.
- Mantel, N. (1967) ‘The detection of disease clustering and a generalized regression approach’, *Cancer Research*, 27(2), pp. 209–220.
- Mozaffarian, V. (1996) *A Dictionary of Iranian Plant Names*. Tehran: Farhang Moaser Publishers.
- Muriira, N.G., Muchugi, A., Yu, A., Xu, J. and Liu, A. (2018) ‘Genetic diversity analysis reveals genetic differentiation and strong population structure in *Calotropis* plants’, *Scientific Reports*, 8, 7832. <https://doi.org/10.1038/s41598-018-26275-x>.
- Mutwakil, M.Z., Hajrah, N.H., Atef, A. et al. (2017) ‘Transcriptomic and metabolic responses of *Calotropis procera* to salt and drought stress’, *BMC Plant Biology*, 17, 231. <https://doi.org/10.1186/s12870-017-1155-7>.
- Navabpour, S., Yamchi, A. and Golcheshmeh, S. (2021) ‘Assessment of genetic diversity between different accessions of *Calotropis procera* with ISSR molecular markers’, *Plant Genetic Resources*, 8(1), pp. 17–28. <https://doi.org/10.52547/pgr.8.1.2>.
- Podani, J. (2000) *Introduction to the Exploration of Multivariate Data*. Leiden: Backhuys Publishers.
- Ramadan, A.M. and Hassanein, S.E. (2014) ‘Characterization of P5CS gene in *Calotropis procera* plant from the de novo assembled transcriptome contigs’, *Comptes Rendus Biologies*, 337(12), pp. 683–690. <https://doi.org/10.1016/j.crv.2014.09.002>
- Segovia, N.I., González-Wevar, C.A. and Haye, P.A. (2020) ‘Signatures of local adaptation in the spatial genetic structure of the ascidian *Pyura chilensis*’, *Scientific Reports*, 10, 14098. <https://doi.org/10.1038/s41598-020-70798-1>.
- Sheidai, M., Shaghali, T., Keshavarzi, M., Koohdar, F. and Ijbari, H. (2019) ‘Species delimitation and inter-specific gene flow in *Tamarix* L.’, *Hacquetia*, 18(2), pp. 1–10. <https://doi.org/10.2478/hacq-2019-0001>.
- Shokry, A.M., Al-Karim, S., Ramadan, A. et al. (2014) ‘Detection of a Usp-like gene in *Calotropis procera* plant from genome contigs’, *Comptes Rendus Biologies*, 337(2), pp. 86–94. <https://doi.org/10.1016/j.crv.2013.12.008>.
- Sokal, R.R. and Rohlf, F.J. (1995) *Biometry: The Principles and Practice of Statistics in Biological Research*. 3rd edn. New York: W.H. Freeman.
- Yu, Y., Harris, A.J., Blair, C. and He, X. (2015) ‘RASP (Reconstruct Ancestral

State in Phylogenies): a tool for historical biogeography', *Molecular Phylogenetics and Evolution*, 87, pp. 46–49. <https://doi.org/10.1016/j.ympev.2015.03.008>.

Zhang, X.X., Liu, B.G., Li, Y., Liu, Y., He, Y.X., Qian, Z.H. and Li, J.X. (2019) 'Landscape genetics reveals that adaptive genetic divergence in *Pinus bungeana* is driven by environmental variables relating to ecological habitats', *BMC Evolutionary Biology*, 19, 160. <https://doi.org/10.1186/s12862-019-1489-x>.