TÜBİTAK

Turkish Journal of Botany

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

Research Article

Turk J Bot
(2021) 45: 541-552
© TÜBİTAK
doi:10.3906/bot-2104-21

Genetic diversity analysis reveals weak population structure in invasive *Trianthema portulacastrum* L. at Fayoum depression, Egypt

Faten Y. ELLMOUNI[1,*], Dirk C. ALBACH[2], Mai Sayed FOUAD[1], Marwa A FAKHR[1,3]

[1]Botany Department, Faculty of Science, Fayoum University, Fayoum, Egypt
[2]Institute of Biology and Environmental Sciences, Carl von Ossietzky-University, Oldenburg, Germany
[3]Plant Protection and Biomolecular Diagnosis Department, Arid Lands Cultivation Research Institute, City of Scientific Research and Technological Applications, (SRTA-City), New Borg El-Arab City, Alexandria, Egypt

Received: 12.04.2021 • **Accepted/Published Online:** 04.10.2021 • **Final Version:** 30.12.2021

Abstract: *Trianthema portulacastrum* L. (Aizoaceae) is a common weed associated with cultivated crops. It is an exotic weed that originated in South Africa and is spreading all over the world. Thirty-five accessions were collected from six populations at Fayoum depression (FD), Egypt. Molecular analyses of start codon targeted (SCoT) markers were performed to identify genotypic variation among collected populations. The effectiveness of employing SCoT markers was demonstrated by the high percentage of polymorphisms. These markers revealed high genetic diversity, as well as high levels of genetic differentiation (G_{ST}), elevated gene flow (Nm) (0.195 and 2.052, respectively), high variation among a population and lower variation within populations. Linkage disequilibrium analysis supported the presence of sexual and clonal reproduction of *T. portulacastrum* in different populations. The data confirmed the weak population structure of *T. portulacastrum* demonstrated in this study via different tools such as STRUCTURE, Minimum spanning network (MSN), and discriminant analysis of principal components (DAPC) and confirmed gene flow between populations. Based on our results, we hypothesize that FD was invaded multiple times by *T. portulacastrum* facilitated by both local adaptation and phenotypic plasticity.

Key words: *Trianthema portulacastrum*, alien weed, genetic variation, population structure, SCoT analysis, Egypt, invasive plants

1. Introduction

Invading alien species threaten natural ecosystems and biological diversity (Cronk and Fuller, 2014). For potential invasiveness, invaders should have numerous characteristics that enable them to spread and proliferate once established such as large seed bank, short generation periods, environmental stress tolerance, and multiple breeding pathways (Li et al., 2019). In addition to these, local adaptation and phenotypic plasticity are considered adaptive strategies improving the establishment and spread of exotic species (Sultan, 2000). Further processes affecting the likelihood of establishment of an exotic species are the number of introductions, selfing breeding system, gene flow, and genetic variation (Tigano and Friesen, 2016; Ward et al., 2008). Phenotypic plasticity plays a role in the adaptability and invasiveness of alien species via increasing or maintaining population growth rate in various environments (Pichancourt and Van Klinken, 2012). Using population genetic analyses, we here analyze how population structure contributes to the colonization of *Trianthema portulacastrum* in the new environment at Fayoum depression (FD).

The genetic diversity and population structure play a crucial role in the success of plant invasions; the variation in a population is an essential prerequisite for the assessment of alien species in the field (Marczewski et al., 2016; Urquía et al., 2019). The marker technique based on start codon targeted (SCoT) polymorphisms introduced by Collard and Mackill (2009) involves the analysis of short, conserved nucleotide sequences that flank the start codon (ATG) for translation initiation. This technique offers several advantages compared to other molecular markers (Agarwal et al., 2019). SCoT markers exhibit high polymorphism levels and extensive, accurate genetic information (Satya et al., 2015). The low cost, reproducibility, stability, and reliable DNA amplification of the SCoT markers make it widely applicable compared to ISSR, AFLP, and RAPD (Gupta et al., 2019). Recently, SCoT markers have been extensively utilized in different molecular applications like estimation of genetic variability

* Correspondence: fyl00@fayoum.edu.eg

(Chhajer et al., 2017), population structure identification (Bhawna et al., 2017), genetic relationship examination among different species or individuals (Rajesh et al., 2015), DNA fingerprinting and molecular diversity analysis (Tabasi et al., 2020).

The genus *Trianthema* L. belongs to the family Aizoaceae (Hernández-Ledesma et al., 2015) with most species of *Trianthema* recorded globally in a broad belt between 35°N and 35°S. The study species, *Trianthema portulacastrum* L. (Carpetweed), is a prostrate, herbaceous succulent with ovate leaves and high branching capacity covering the ground by forming a green carpet (Fahad et al., 2014).

Little is known about its genetic diversity. However, analyses of genetic variation are especially important to assess plant response strategies while facing different environmental conditions (Vicente et al., 2018). The pollination system in *T. portulacastrum* is facultatively outcrossing (Branch and Sage, 2018). Low dormancy, enormous seed production, and efficient seed dispersal along with high acclimatization capacity lead to a large seed bank in the soil that enables the species to survive in harsh conditions and allows dispersion and establishment as invasive weed (Kaur and Aggarwal, 2017) .

Trianthema portulacastrum is an aggressive invasive species found natively in tropical Africa. It has been reported to be widely distributed in Egypt since 1974, but has only been scantily found before (Täckholm, 1974). In the early eighties, it became a dominant invasive especially in crop fields (Osbornová-Kosinová, 1984; Shaltout et al., 2013).

Trianthema portulacastrum is regarded as a noxious weed in Africa, Asia, and Australia (Kaur and Aggarwal, 2017) and a problematic weed in Egypt with a highly competitive growth habit (Shaltout et al., 2013). FD represents a small subsection of Egypt but constitutes an important region for agriculture. This is related to the fact that FD has a geographical landscape analogous to Egypt's topography where Qarun Lake lies on Fayoum's northern coastline, comparable to Egypt bordering the Mediterranean Sea in the north, and Bahr Yusuf canal is described as a backbone of FD similar to the Nile River for Egypt (Elgamal et al., 2017). Fayoum depression is considered to be an outlet of the Nile material through Bahr Yusuf, which has likely been the main route of dispersal to FD for several hundreds of species (Sun et al., 2019). Information on the genetics of *T. portulacastrum* is scarce; previous studies on its genetic variation and population structure were limited to plastid *rbc*L and nuclear ribosomal ITS sequence data (Hassan et al., 2005; Manhart and Rettig, 1994). To the best of our knowledge, the present work is the first attempt to analyze the genotypic variation among populations of *T. portulacastrum*. This, however, is important to understand the population structure and reproductive strategy of *T. portulacastrum* and to explore its invasion dynamics in the FD ecosystem.

2. Materials and methods

2.1. Study site and plant material

Plants were collected in all six regions of FD: Etsa, Fayoum, Senouris, Tamia, Ibshawy, and Yousef El-seddik districts, which constitute an assemblage of agricultural, desertic and coastal habitats in FD (El-Zeiny and Effat, 2017). Thirty-five accessions of *T. portulacastrum* (Tables 1 and S1) were thoroughly chosen in such a way to guarantee comprehensive coverage of *T. portulacastrum* distribution throughout FD.

2.2. Molecular and statistical analysis

2.2.1. DNA extraction, purification, and quantification

High molecular weight plant genomic DNA was extracted from 50–100 mg silica-gel dried leaf samples of *T. portulacastrum* with DNeasy Plant Mini Kits (QIAGEN, Hilden, Germany). DNA quantity and purity of extraction was verified using a NanoDrop ND-1000 spectrophotometer (NanoDrop Tech., Thermo Fisher Scientific Inc.).

2.2.2. SCoT polymorphism

The SCoT marker technique was used to analyze the genetic differentiation and diversity between

Table 1. Distribution of *Trianthema portulacastrum* samples at Fayoum depression.

Population name (district)	Site acronyms	Latitude range of represented samples	Sample size (n)	Elevation range (a.s.l)	Area (km^2)	Size of districts in Fayoum depression (% of each sector)
Etsa	Pop.E	29.7 : 29.18	6 (T1-T6)	5 to 17 m	483.75	8.35%
Fayoum	Pop.F	29.15 : 29.21	11(T7-T17)	16 to 27 m	393.94	6.8%
Senouris	Pop.S	29.22 : 29.26	4 (T18-T21)	–13 to 16 m	225.10	3.88%
Tamia	Pop.T	29. 27 : 29.32	5 (T22-T26)	–13 to 12 m	379.21	6.54%
Ibshawy	Pop.IB	29.19 : 29.21	3 (T27-T29)	14 to 18 m	165.61	2.86%
Yousef El-seddik	Pop.Y	29.17 : 29.26	6 (T30-35)	–46 to 12 m	376.60	6.50%

the studied *T. portulacastrum* accessions. *Trianthema portulacastrum* samples were assessed for genetic variation using thirteen SCoT primers as designed by (Collard and Mackill, 2009). The sequences of DNA-SCoT primers were synthesized by Macrogen (Seoul, Republic of Korea) (Table 2). Polymerase chain reaction was performed according to Ibrahim et al. (2017). A 1.5% ethidium bromide-stained polyacrylamide gel was used to visualize PCR amplicons in 1X TBE buffer. The gels were photographed and documented in a gel documentation and image analysis system according to Sambrook et al. (1989).

2.2.3. Population diversity

Based on SCoT marker analysis, genetic diversity and distance-based relationships were analyzed for the 35 *T. portulacastrum* accessions. Consequently, polymorphic bands in the SCoT profiles were scored as 0 and 1, according to Collard and Mackill (2009). The SCoT amplicons that were steadily scored with fixed size compared to a ladder were considered a unique locus corresponding to a targeted genome's distinctive position. We used an online program, Marker Efficiency Calculator (iMEC)[1] (Amiryousefi et al., 2018) to calculate polymorphism indices.

The estimation of population genetic (PG) parameters such as allele number (Na), effective alleles (Ne), Nei's expected heterozygosity (h), Shannon's diversity index (I), percent polymorphism (Pp), total genetic diversity (Ht), population genetic diversity (Hs), population genetic differentiation (G_{ST}), and gene flow (Nm) were analyzed using POPGENE software version1.31 (Yeh et al., 1999).

The PG parameter assessment was followed and confirmed by using R (version 3.5.1; (R_Core_Team, 2018). The binary data were clone corrected to eliminate identical multilocus genotypes (MLGs) from each collection region. By utilizing the same package, we calculated the association (*IA*) index and used 100,000 permutations to provide a *p*-value to employ it in the linkage disequilibrium test (LD). This test is used to infer whether populations are clonal or sexual based on the significant disequilibrium (Grünwald et al., 2017). A cluster tree was constructed based on "Nei's genetic distance" and plotted using the R-package "Poppr" (Kamvar et al., 2014).

2.2.4. Genetic differentiation and population structure

A Mantel test for correlation between genetic and geographic distances seeking a spatial pattern of genetic variation and analysis of molecular variance (AMOVA) was performed to analyze the distribution of genetic variation among and within populations using GenAlEx version 6.5 (Peakall and Smouse, 2012).

For analyzing population genetic structure, STRUCTURE v2.3.4 was utilized in a Bayesian clustering approach to analyze population genetic structure (Pritchard et al., 2000). The parameter was set for MCMC (Markov Chain Monte Carlo), 100,000 repetitions, and 20 replicates run of K= 2 - 7 (Evanno et al., 2005). To determine the optimum K for the data, we used Structure Harvester v6.0 (Earl and vonHoldt, 2012). The program BOTTLENECK (V.1.2.02) was used to detect potential bottlenecks for SCoT data, aiming to explore population dynamics (Piry et al., 1999).

The two R packages "*magrittr*" and "*Poppr*" (Kamvar et al., 2014) were used to create a minimum spanning network (MSN) for visualizing the relationships among accessions. Depending on Bruvo's distance, MSN approximates the genetic distance between accessions rather than between collection regions (Bruvo et al., 2004). We used the "adegenet" package (Jombart, 2008) to construct the discriminating analysis of principal components (DAPC), which is considered appropriate for populations that are clonal or partially clonal (Grünwald et al., 2017). An agglomerative hierarchical clustering was generated by scoring bands from the data (Kolde and Kolde, 2015) in the R-package "*pheatmap*".

3. Results

3.1. SCoT marker analysis

The 13 SCoT primers amplified 193 amplicons with a range of 13 to 18 bands per primer, exhibiting 100% polymorphic bands (Table 2; Figure S1). The lengths of the products varied from 150 bp to 1700 bp. The mean values of polymorphism indices such as heterozygosity index (H), polymorphism information content (PIC), effective multiplex ratio (E), arithmetic mean heterozygosity (Havp), marker index (MI), discriminating power (D), and resolving power (Rp) were 0.453, 0.35, 5.2, 0.0008, 0.004, 0.87, and 8.01, respectively. The maxima of PIC (0.368), H (0.488), E (6.34), and MI (0.005) were found for SCoT 12, and the highest and lowest (Rp) values of 10.2 and 5.94 are shown by SCoT 1 and SCoT 28, respectively (Table 2).

3.2. Population genetic diversity analysis

The observed and effective number of alleles ranged between 1.51–1.81 and 1.34–1.44, respectively. Correspondingly, Nei's gene diversity (h) and Shannon's Information index (I) ranged between 0.2–0.27 with an overall diversity of 0.29 and 0.29–0.42 with an average value of 0.45, respectively. The percentage of polymorphic loci (Pp) is estimated in the range of 51.81% to 91.19 % (Table 3). Mean total genetic diversity (Ht) and genetic diversity within populations (Hs) in *T. portulacastrum* samples gathered from six ecogeographic regions of FD were found to be high (0.29 and 0.23, respectively).

We observed significant support for linkage disequilibrium with *p-value* ($\bar{r}_d$) =1e^{-05} and detected a

[1] available online at https://irscope.shinyapps.io/iMEC/

Table 2. Genetic polymorphisms generated by 13 SCoT markers in *T. portulacastrum*. The conserved start codon is underlined.

Primer name	Sequence (5′ › 3′)	% GC	Scored bands	NPB	PPB	H_0	PIC_0	E_0	H.av_0	MI_0	D_0	R_0	Product size (bp)
SCOT-1	CAACAATGGCTACCACCA	50%	18	18	100%	0.4545	0.3512	6.2857	0.000721	0.004535	0.8784	10.2285	170-1400
SCOT-11	AAGCAATGGCTACCACCA	50%	15	15	100%	0.4679	0.3584	5.6	0.000891	0.004991	0.8610	8.9142	158-1370
SCOT-12	ACGACATGGCGACCAACG	61%	15	15	100%	0.4880	0.3689	6.3428	0.00093	0.005897	0.8216	9.2	150-1700
SCOT-14	ACGACATGGCGACCACGC	67%	15	15	100%	0.4365	0.3412	4.8285	0.000832	0.004015	0.8967	6.8	230-1200
SCOT-15	ACGACATGGCGACCGCGA	67%	13	13	100%	0.4598	0.3540	4.6571	0.001011	0.004706	0.8721	8.4571	270-1050
SCOT-16	ACCATGGCTACCACCGAC	56%	16	16	100%	0.4323	0.3388	5.0571	0.000772	0.003904	0.9004	9.1428	220-1500
SCOT-18	ACCATGGCTACCACCGCC	67%	14	14	100%	0.4658	0.3573	5.1714	0.000951	0.004917	0.8640	8.5714	240-1200
SCOT-23	ACCATGGCTACCACCGCC	61%	15	15	100%	0.4698	0.3594	5.6571	0.000895	0.005062	0.8582	8.1714	250-1600
SCOT-25	ACCATGGCTACCACCGGG	67%	15	15	100%	0.4564	0.3522	5.2857	0.000869	0.004595	0.8762	7.7714	240-1200
SCOT-28	CCATGGCTACCACCGCCA	67%	14	14	100%	0.4480	0.3476	4.7428	0.000914	0.004336	0.8856	5.9428	210-1400
SCOT-29	CCATGGCTACCACCGGCC	72%	14	14	100%	0.3992	0.3195	3.8571	0.000815	0.003142	0.9245	6.7428	200-1400
SCOT-33	CCATGGCTACCACCGCAG	67%	15	15	100%	0.4444	0.3456	5	0.000847	0.004233	0.8893	7.7714	170-1100
SCOT-35	CATGGCTACCACCGGCCC	72%	14	14	100%	0.4679	0.3584	5.2285	0.000955	0.004994	0.8609	6.5142	250-1392
Mean			193	193	100%	0.4531	0.350	5.2087	0.000877	0.004564	0.8761	8.0175	

NPB: number of polymorphic bands, PPB: percentage of polymorphic bands, H: expected heterozygosity, PIC: polymorphism information content, E: effective multiplex ratio, Havp: mean heterozygosity, MI: marker index, D: discriminating power, R: resolving power.

maximum value of the standardized index of association ($\bar{r}_d$) (0.0285 and 0.0156) at Tamia and Fayoum regions, respectively, which falls outside of the distribution expected under no linkage (Figure S2a and S2b). Etsa and Yousef El-seddik regions had *p-values* ($\bar{r}_d$) =0.02; thus, the null hypothesis was rejected and suggested no linkage among markers; however, moderate ($\bar{r}_d$) values (0.00675 and 0.00596), respectively, appeared on the right end of the resampled distribution (Figure S2c and S2d). Finally, Senouris and Ibshawy regions failed to reject the linkage disequilibrium hypothesis with *p-values* ($\bar{r}_d$) = 0.794 and 0.502 and negative values of ($\bar{r}_d$) (–0.00469 and –0.00273), respectively (Figure S2e and S2f). The average value was ($\bar{r}_d$) =0.00831 with a *p-value* =1e^{-04} for all populations. The $\bar{r}_d$ values were significantly ≥ zero, indicating linkage equilibrium and the significance of *p-values*.

3.3. Genetic differentiation and population structure

Among different *T. portulacastrum* populations, we found high levels of genetic differentiation (G_{ST}: 0.195; G_{ST} > 0.15 is considered high (Hamrick et al., 1991; Nei, 1978)) and a high value of gene flow (Nm=2.052; Nm > 1 is considered high (Shekhawat et al., 2018)). The AMOVA demonstrated that a large amount of genetic variation (35%) was observed within the populations, but the variance among populations contributed even more (65%) and, thus, the highest genetic variance (PhiPT = 0.654, P = 0.001) (Table 4). The data showed a significant correlation between the genetic and geographic distances among populations analyzed using a Mantel test (r = 0.36, p < 0.05).

Based on the highest ΔK value generated by STRUCTURE HARVESTER software, the optimal number of clusters was inferred to be four (Figure 1a). Population Ibshawy mainly consisted of the green cluster individuals; half of the individuals belonging to the Etsa population were distinct by forming the blue cluster. The rest of the populations were mixed, indicating admixture among all clusters. The MSN supported the STRUCTURE results, in which the admixture between populations with each other appear evident (Figure 1b).

DAPC and the cluster tree findings supported the STRUCTURE and MSN results clustering all individuals into four main groups, those from Ibshawy in a single supported branch. These individuals were also grouped together by DAPC. Based on genetic distance, Ibshawy is most distant from the rest (76.7% bootstrap support (BS)), followed by Senouris (56.8% BS) (Figures 2a and 2b). STRUCTURE based on individual ancestry proportions (Q values) expressed genetic relationships

Table 3. Genetic diversity statistics and differentiation parameters for six populations of *T. portulacastrum*.

Population/group	N	Na ± SD	*Ne ± SD*	*h ± SD*	*I ± SD*		*Pp*	*Ht*	*Hs*	G_{ST}	*Nm*	PhiPT
Pop.1.Etsa	6	1.69 ± 0.46	1.40 ± 0.36	0.23 ± 0.19	0.36 ± 0.27	134	69.43%					
pop.2.Fayoum	11	1.91 ± 0.28	1.44 ± 0.32	0.27 ± 0.15	0.42 ± 0.21	176	91.19%					
pop.3. Senouris	4	1.55 ± 0.49	1.34 ± 0.36	0.20 ± 0.19	0.30 ± 0.28	108	55.96%					
pop.4.Tamia	5	1.74 ± 0.43	1.43 ± 0.35	0.26 ± 0.18	0.39 ± 0.25	143	74.09%					
pop.5.Ibshawy	3	1.51 ± 0.50	1.34 ± 0.38	0.20 ± 0.20	0.29 ± 0.29	100	51.81%					
pop.6.Yousef El-seddik	6	1.81 ± 0.38	1.39 ± 0.31	0.24 ± 0.16	0.38 ± 0.22	158	81.87%					
Mean		2.00 ± 0.00	1.47 ± 0.31	0.29 ± 0.14	0.45 ± 0.18	193	70.73	0.29 ± 0.02	0.23 ± 0.01	0.196	2.052	0.65

N: No. of samples, Na: Observed no of alleles, Ne: Effective no of alleles, h: Nei's gene diversity, I: Shannon's information index, Pp: Percentage of polymorphic loci, Ht: Total genetic diversity, Hs: population diversity G_{ST}: Diversity among populations; Nm: Gene flow (0.5 (1– G_{ST})/G_{ST}); phiPT: population differentiation

Table 4. Analysis of molecular variance (AMOVA) of 35 *T. portulacastrum* accessions belonging to six different populations.

Source of variation	df	SS	MS	Est. Var.	% variation	*P*
Among populations	5	2.16	0.432	0.070	65%	0.001
Within populations	29	1.082	0.037	0.037	35%	
Total	34	3.242		0.108	100%	

df: degree of freedom, SS: sum of squares, MS: mean squares.

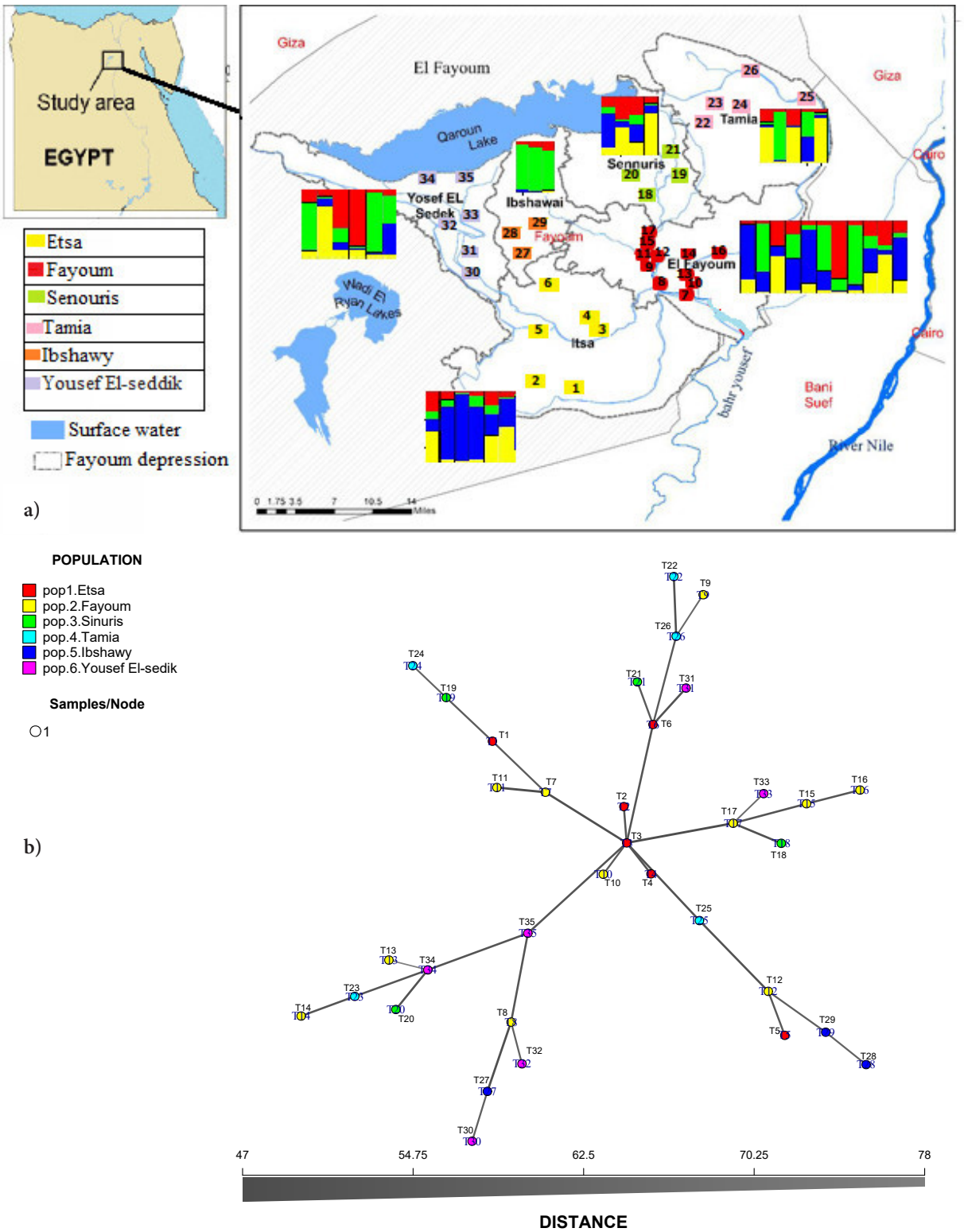

Figure 1. a) Geographical distribution of the studied *T. portulacastrum* populations in the Fayoum depression in Egypt and the results of genetic assignment of individuals analysis based on the Bayesian method implemented in STRUCTURE assuming correlated frequencies and admixed origin of populations for K = 4. b) Minimum spanning network (MSN) of *T. portulacastrum* based on Bruvo's genetic distance for 13 SCOT loci. The nodes of the MSN represent individual multilocus genotypes (MLGs) with the color and size representing population. Lines between nodes represent genetic distance between MLG.

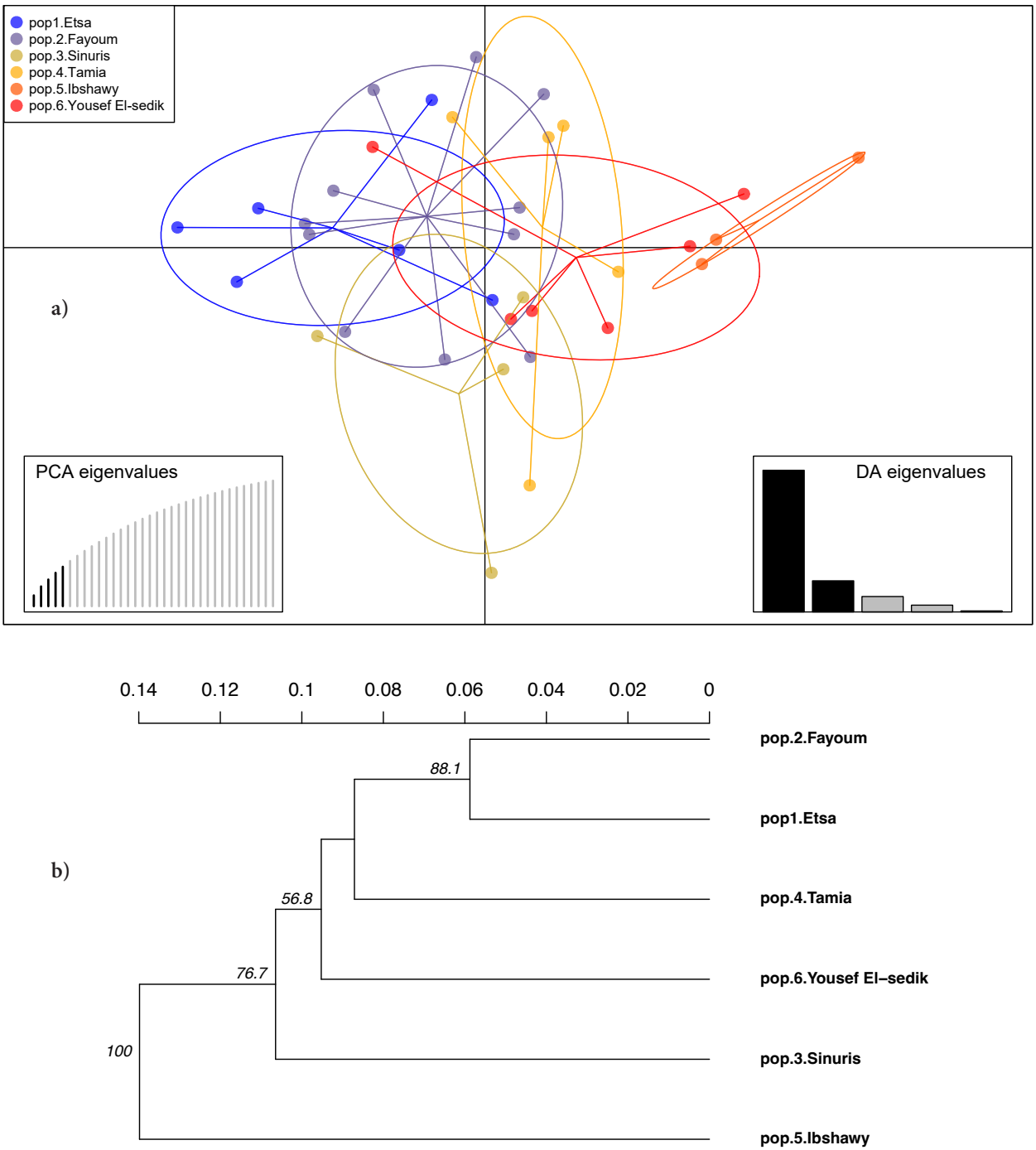

Figure 2. a) Ordination plot for the first two principal component axes using discriminant analysis of principal components (DAPC) method among 6 populations for each individual, ellipses indicate their assignment to the genetic clusters inferred. The low-right graph indicates the variance explained by the principal component axes used for DAPC (dark grey). b) Distance-based tree for populations divergence based on Nei's genetic distance.

and emphasized a high genetic variance among the 35 *T. portulacastrum* accessions (Figure 3a). Agglomerative hierarchical clustering (Heatmap) divided the samples into two clusters, each one separated into two subclusters. Subcluster 1a is the smallest one and contains individuals from different populations with blue, green and yellow clusters represented. Cluster 1b has individuals found in the green cluster. Cluster 2a includes individuals from

the yellow cluster whereas cluster 2b groups individuals from the blue cluster and two individuals from the red one (Figure 3).

4. Discussion

4.1. Genetic diversity and differentiation

In agreement with earlier investigations, e.g., Etminan et al. (2018) on *Triticum turgidum* var. *durum* and Yang et al. (2019), SCoT markers showed high percentage of polymorphisms (100%) and moderate PIC values of (0.368), indicating the high information potential of the markers (Table 2). PIC values were previously categorized into three categories, high (PIC > 0.5), medium (0.25 < PIC < 0.5), and low (PIC < 0.25) (Yadav et al., 2011). Based on these criteria, the SCoT markers developed for *T. portulacastrum* exhibited a moderately informative level of PIC. *Trianthema portulacastrum* displayed a moderate level of genetic diversity and Shannon's information index, averaging 0.29 and 0.45, respectively (Table 3). Similar results were observed with SCoT and ISSR markers in *Dendrobium nobile* (0.28 and 0.43; (Bhattacharyya et al., 2013) and watermelon ecotypes (0.29 and 0.41; (Soghani et al., 2018), both also able to reproduce sexually and clonally. The current study revealed high genetic differentiation levels (G_{ST}) and an elevated gene flow: 0.195 and 2.052, respectively.

A possible explanation for the high gene flow observed in *T. portulacastrum* may be its strong reproductive thermotolerance allowing flower production in high midday temperature conditions (Branch and Sage,

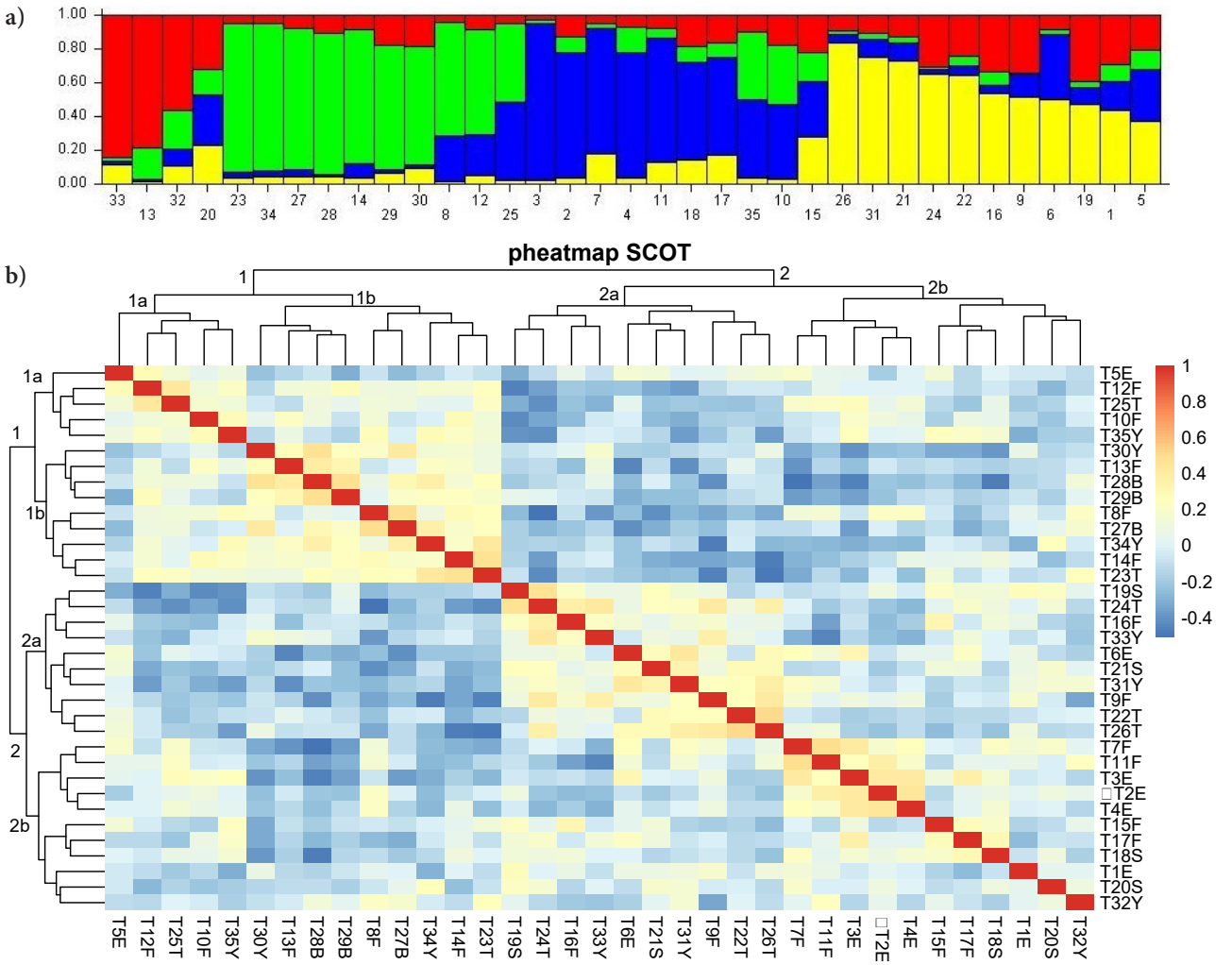

Figure 3. a) Population structure of different *T. portulacastrum* accessions in FD based on STRUCTURE software and Structure Harvester, the Bayesian analysis results indicated for K = 4 (SORT BY Q), the values of K corresponding to the number of clusters (represented by different colors) summarizing the samples at six populations. b) Agglomerative hierarchical clustering (Heatmap) generated by scored bands data of SCOT marker.

2018), when other flowers are scarce. In these times, it is considered an important subsistence food for honeybees and other insects (Dalio, 2015). High gene flow by seeds and vegetative parts is likely based on human agricultural practices and by irrigation channels. High levels of gene flow in genetically diverse species potentially introduce locally adaptive alleles to new populations and allow natural selection to aid in local adaptation to drought climates (Shekhawat et al., 2018).

Whereas at first sight counter-intuitive, high gene flow is accompanied by strong genetic differentiation. However, we consider these results to be caused by multiple introductions of *T. portulacastrum* to FD and incomplete mixing of the populations (e.g., Ibshawy) as demonstrated by the analyses of population structure. Results by Wu et al. (2020) are consistent with our results for the occurrence of genetic differentiation in parallel with high gene flow, which suggests that situations of high gene flow and genetic differentiation exist in cases of high gene flow in species with strong population structure. Such population structure may be caused by independent origins but also local adaptation or strong bottlenecks in a formerly widespread species. We cannot exclude either of these explanations but consider multiple introductions to different parts of FD the most likely explanation for genetic differentiation in FD. Larger scale analyses of intraspecific variation in *T. portulacastrum* would be necessary to distinguish between the alternatives.

SCoT data on intraspecific population genetic structure is currently unavailable for most invasive plants, although these are essential for understanding adaptation and evolution of invasive species (Colautti et al., 2017). In addition, the genetic variation of plants is affected by biological features of the species, such as mating systems, dispersal syndrome, and gene flow (Avise and Hamrick, 1996).

In our data, higher genetic variability was noted among populations (65%) than within the populations (35%) in *T. portulacastrum*. Compared to other systems, these numbers indicate a rather high between-population differentiation. However, one should bear in mind the multiple origins of FD *T. portulacastrum*. Thus, the numbers are easily explained by a mixed mating system characteristic for *T. portulacastrum* and/or frequent dispersal between populations, and some degree of population differentiation due to independent introductions. Normal L-shaped distribution demonstrates an absence of bottlenecks in *T. portulacastrum* supporting that genetic variation has increased attributable to gene flow, outbreeding nature, possibly high numbers of introduced seeds in multiple events and admixture of different genetic sources among invasive populations (Li et al., 2019).

4.2. Population structure and multiple introduction

Analysis of linkage disequilibrium (LD) is important to estimate if the observed alleles at different loci are linked (asexual reproduction) or are not linked allowing alleles to recombine freely into a new genotype (sexual reproduction) (Grünwald et al., 2017). Significant linkage disequilibrium was observed at Fayoum, Tamia, Yousef El-sedik, and Etsa regions indicating that *T. portulacastrum* reproduced in these regions by clonal reproduction. Abd-Elgawad et al. (2013) mentioned that these regions have an especially arid climate due to high temperature, evaporation, low humidity, and wind action. Soil salinization resulting from irrigation is higher here than within the Nile River path, and this may limit pollinator activity. Nevertheless, genetic diversity is high in these regions and varies among populations (Table 3). Thus, different amounts of linkage disequilibrium as a consequence of differences in recombination and genetic drift are expected (Slatkin, 2008).

Whereas nonsignificant linkage disequilibrium was observed at Ibshawy and Senouris regions indicating that *T. portulacastrum* reproduced in these regions by sexual reproduction, significant disequilibrium was found in the other populations either indicating clonal reproduction or other factors simulating the same effect. Differences in linkage disequilibrium are important in invasive species, since linkage disequilibrium interacts with selection and genetic drift in ways that are difficult to predict. Thus, strong selection on linked loci can cause high amounts of LD, whereas high genetic drift likewise increases LD (Slatkin, 2008). Thus, small, isolated populations with low genetic diversity and low selection pressure but some sexual reproduction may have lower linkage disequilibrium than large populations of diverse origin and strong selection pressure but predominantly clonal reproduction.

Trianthema portulacastrum seeds are dispersed by wind and water flow due to the small and lightweight seeds (Fahmy et al., 2019; Shaltout et al., 2013). Given that the Fayoum region is the main entry gate of Nile material through Bahr Yusuf, the life artery of FD, it is likely that genetic diversity is elevated here through multiple introductions from the Nile River and other parts of Egypt. According to the aforementioned results, we are implying a weak population structure of *T. portulacastrum* (Figure 1a) that might be caused by most of the populations being admixed and consisting of a dominant allele from more than one founder event (Li et al., 2019). According to our suggestion, the Fayoum region is probably the ancestral population from which other populations derived.

5. Conclusions

Our results suggest that in ways the invasion of *T. portulacastrum* is favored by multiple introductions, outcrossing pollination, high genetic diversity, and highly

dynamic gene flow, which facilitates local adaptation. Future studies should investigate genetic diversity of *T. portulacastrum* of FD in relation to genetic diversity in other parts of Egypt and the extent of local adaptation by common garden experiments. However, it would also be interesting to estimate the importance of the species for the survival of insect pollinator populations.

Acknowledgments
The authors would like to extend special thanks to Genetics team members Botany Department, Faculty of Science, Ain Shams University, Cairo, Egypt, in appreciation of their suggestion to use SCoT markers as a technique choice and for his valuable comments and time.

Author contributions
All authors contributed to the study equally.

Conflict of interest
The authors declare that they have no conflicts of interest.

Ethical approval
This article does not contain any studies with human participants or animals performed by any of the authors.

References

Abd-Elgawad M, Shendi MM, Sofi DM, Abdurrahman HA, AM A (2013). Geographical distribution of soil salinity, alkalinity, and calcicity within Fayoum and Tamia districts, Fayoum Governorate, Egypt. In: Shahid SA, Abdelfattah, MA, Taha, FK, (editors) Developments in soil salinity assessment and reclamation: innovative thinking and use of marginal soil and water resources in irrigated agriculture. Dordrecht, The Netherlands: Springer pp. 219-236. doi:10.1007/978-94-007-5684-7_14

Agarwal A, Gupta V, Haq SU, Jatav PK, Kothari S et al. (2019). Assessment of genetic diversity in 29 rose germplasms using SCoT marker. Journal of King Saud University Science 31 (4):780-788. doi:10.1016/j.jksus.2018.04.022

Amiryousefi A, Hyvönen J, Poczai P (2018). iMEC: online marker efficiency calculator. Appl Plant sciences 6 (6):e01159. doi.org/10.1002/aps3.1159

Avise JC, Hamrick JL (1996). Conservation genetics: case histories from nature. New York, New York, USA: Columbia University Press.

Bhattacharyya P, Kumaria S, Kumar S, Tandon P (2013). Start Codon Targeted (SCoT) marker reveals genetic diversity of *Dendrobium nobile* Lindl., an endangered medicinal orchid species. Genetic Resources and Crop Evolution 529 (1):21-26. doi:10.1016/j.gene.2013.07.096

Bhawna, Abdin M, Arya L, Verma M (2017). Use of SCoT markers to assess the gene flow and population structure among two different populations of bottle gourd. Plant Gene (9):80-86. doi:10.1016/j.plgene.2016.09.001

Branch HA, Sage RF (2018). Reproductive heat tolerance in a Mojave Desert annual plant, *Trianthema portulacastrum*. American Journal of Botany 105 (12):20182024.doi:10.1002/ajb2.1199

Bruvo R, Michiels NK, D'souza TG, Schulenburg H (2004). A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. Molecular Ecology 13 (7):2101-2106. doi:10.1111/j.1365-294X.2004.02209.x

Chhajer S, Jukanti AK, Kalia RK (2017). Start codon targeted (SCoT) polymorphism-based genetic relationships and diversity among populations of *Tecomella undulata* (Sm.) Seem—an endangered timber tree of hot arid regions. Tree Genet Genomes 13 (4):84. doi:10.1007/s11295-017-1169-1

Colautti RI, Alexander JM, Dlugosch KM, Keller SR, Sultan SE (2017). Invasions and extinctions through the looking glass of evolutionary ecology. Philosophical Transactions of the Royal Society B 372 (1712):20160031. doi:10.1098/rstb.2016.0031

Collard BC, Mackill DJ (2009). Start codon targeted (SCoT) polymorphism: a simple, novel DNA marker technique for generating gene-targeted markers in plants. Plant molecular biology reporter 27 (1):86. doi:10.1007/s11105-008-0060-5

Cronk QC, Fuller JL (2014). Plant invaders:the threat to natural ecosystems.London, UK: Routledge, Earthscan Publications Limited.

Dalio JS (2015). Foraging behaviour of Apis mellifera on *Trianthema portulacastrum*. Journal of Entomology Zoology Studies 3 (2):105-108.

Earl D, vonHoldt B (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4 (2):359-361. doi:10.1007/s12686-011-9548-7

El-Zeiny AM, Effat HA (2017). Environmental monitoring of spatiotemporal change in land use/land cover and its impact on land surface temperature in El-Fayoum governorate, Egypt. Remote Sensing Applications: Society and Environment (8):266-277. doi:10.1016/j.rsase.2017.10.003

Elgamal MM, El-Alfy KS, Abdallah MG, Abdelhaleem FS, Elhamrawy AM (2017). Restoring Water and Salt Balance of Qarun Lake, Fayoum, Egypt. Mansoura Engineering Journal 42 (4):1-14. doi:10.21608/BFEMU.2020.97674

Etminan A, Pour-Aboughadareh A, Noori A, Ahmadi-Rad A, Shooshtari L, Mahdavian Z, Yousefiazar-Khanian M (2018). Genetic relationships and diversity among wild *Salvia* accessions revealed by ISSR and SCoT markers. Biotechnol Biotechnol Equip 32 (3):610-617. doi: 10.1080/13102818.2018.1447397

Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14 (8):2611-2620.

Fahad S, Hussain S, Saud S, Hassan S, Muhammad H et al.(2014). Consequences of narrow crop row spacing and delayed *Echinochloa colona* and *Trianthema portulacastrum* emergence for weed growth and crop yield loss in maize. Weed Research 54 (5):475-483. doi:10.1111/wre.12104

Fahmy GM, Moussa SA, Farrag HF, Rehem RAAE (2019). Seed and germination traits of the summer weed *Trianthema portulacastrum* L. Egyptian Journal Experimental Biology (Bot) 15 (2):235-242. doi:10.5455/egyjebb.20190618082647

Grünwald NJ, Everhart SE, Knaus B, Kamvar ZN (2017). Best practices for population genetic analyses. Phytopathology 107 (9):1000-1010. doi:10.1094/PHYTO-12-16-0425-RVW

Gupta V, Jatav PK, Haq SU, Verma KS, Kaul VK et al. (2019). Translation initiation codon (ATG) or SCoT markers-based polymorphism study within and across various *Capsicum* accessions: insight from their amplification, cross-transferability and genetic diversity. Journal of Genetics 98 (2):1-12. doi:10.1007/s12041-019-1095-0

Hamrick J, Godt M, Murawski D, Loveless M (1991). Correlations between species traits and allozyme diversity: implications for conservation biology. In: Falk D, Holsinger, KE., (editors) Genetics conservation of rare plants.New York, New York, USA: Oxford University Press, pp. 75– 86.

Hassan N, Thiede J, Liede-Schumann S (2005). Phylogenetic analysis of Sesuvioideae (Aizoaceae) inferred from nrDNA internal transcribed spacer (ITS) sequences and morphological data. Plant Systematics and Evolution 255 (3-4):121-143. doi:10.1007/s00606-005-0357x

Hernández-Ledesma P, Berendsohn WG, Borsch T, Von Mering S, Akhani H et al. (2015). A taxonomic backbone for the global synthesis of species diversity in the angiosperm order Caryophyllales. Willdenowia 45 (3):281-383.doi:10.3372/wi.45.45301

Ibrahim M, Mohamed AS, Teleb S, Ibrahim S, Tantawy M (2017). Taxonomic and molecular study on some Asian cultivars of *Triticum aestivum* L. Taeckholmia 37 (1):16-29. doi: 10.21608/TAEC.2017.11932

Jombart T (2008). adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24 (11):1403-1405.

Kamvar Z, Tabima J, Grünwald N (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2013 1–14. doi:10.7717/peerj.281

Kaur M, Aggarwal N (2017). *Trianthema portulacastrum* L.- the noxious weed and its control. Advances in plants agriculture research 6 (3):00213.

Kolde R, Kolde MR (2015). Package 'pheatmap'. R package 1 (7):790.

Li F, van Kleunen M, Li J, Liu X, Gao K et al (2019). Patterns of genetic variation reflect multiple introductions and pre-admixture sources of common ragweed (*Ambrosia artemisiifolia*) in China. Biological Invasions 21 (6):2191-2209. doi:10.1007/s10530-019-01966-2

Manhart JR, Rettig JH (1994). Gene sequence data. In: Behnke HD. MTJ (editors) Caryophyllales.Berlin, Heidelberg: Springer, pp. 235-246. doi:10.1007/978-3-642-78220-6_10

Marczewski T, Ma Y-P, Zhang X-M, Sun W-B, Marczewski AJ (2016). Why is population information crucial for taxonomy? A case study involving a hybrid swarm and related varieties. AoB Plants 8 1-17. doi:10.1093/aobpla/plw070

Nei M (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89 (3):583-590. doi:10.1093/genetics/89.3.583

Osbornová-Kosinová J (1984). Notes on the synanthropic plants of Egypt. Folia Geobotanica et Phytotaxonomica 19 (3):279-285. doi:10.1007/BF02853091

Peakall R, Smouse P (2012). GenAlEx Tutorials-Part 2: Genetic distance and analysis of molecular variance (AMOVA). http://www.anu.edu.au/BoZo/GenAlEx/

Pichancourt JB, Van Klinken RD (2012). Phenotypic plasticity influences the size, shape and dynamics of the geographic distribution of an invasive plant. PloS one 7 (2):e32323. doi:10.1371/journal.pone.0032323

Piry S, Luikart G, Cornuet JM (1999). Computer note. BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. Journal of Heredity 90 (4):502-503. doi:10.1093/jhered/90.4.502

Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. Genetics 155 (2):945-959.

R_Core_Team (2018). R: A Language and Environment for Statistical Computing. In. Vienna, Austria: R Foundation for Statistical Computing.

Rajesh M, Sabana A, Rachana K, Rahman S, Jerard B et al. (2015). Genetic relationship and diversity among coconut (*Cocos nucifera* L.) accessions revealed through SCoT analysis. Biotech 5 (3):999-1006. doi: 10.1007/s13205-015-0304-7

Sambrook J, Fritsch F, Maniatis T (1989). Molecular Cloning – A laboratory manual, Second edition. edn.New York: Cold Spring Harbor Laboratory Press.

Satya P, Karan M, Jana S, Mitra S, Sharma A et al. (2015). Start codon targeted (SCoT) polymorphism reveals genetic diversity in wild and domesticated populations of ramie (*Boehmeria nivea* L. Gaudich.), a premium textile fiber producing species. Meta gene 3 62-70. doi:10.1016/j.mgene.2015.01.003

Shaltout K, Baraka DM, Shehata MN, Ahmed D, Arief OM (2013). Distributional behavior and growth performance of *Trianthema portulacastrum* L.(Aizoaceae) in Nile Delta. Egyptian Journal of Botany 183-199.

Shekhawat JK, Rai MK, Shekhawat N, Kataria V (2018). Start codon targeted (SCoT) polymorphism for evaluation of genetic diversity of wild population of *Maytenus emarginata*. Industrial Crops and Products 122 202-208. doi:10.1016/j.indcrop.2018.05.074

Slatkin M (2008). Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. Nat Rev Genet 9 (6):477-485. doi:10.1038/nrg2361

Soghani ZN, Rahimi M, Nasab MA, Maleki M (2018). Grouping and genetic diversity of different watermelon ecotypes based on agro-morphological traits and ISSR marker. Iheringia Série Botânica 73 (1):53-59.

Sultan SE (2000). Phenotypic plasticity for plant development, function and life history. Trends Plant Sci 5 (12):537-542. doi:10.1016/S1360-1385(00)01797-0

Sun Q, Liu Y, Salem A, Marks L, Welc F et al. (2019). Climate-induced discharge variations of the Nile during the Holocene: Evidence from the sediment provenance of Faiyum Basin, north Egypt. Glob Planet 172 200-210. doi:10.1016/j.gloplacha.2018.10.005

Tabasi M, Sheidai M, Hassani D, Koohdar F (2020). DNA fingerprinting and genetic diversity analysis with SCoT markers of Persian walnut populations (*Juglans regia* L.) in Iran. Genetic Resources and Crop Evolution 1-11. doi:10.1007/s10722-020-00914-7

Täckholm V (1974). Students' flora of Egypt, 2nd.

Tigano A, Friesen VL (2016). Genomics of local adaptation with gene flow. Molecular Ecology 25 (10):2144-2164. doi:10.1111/mec.13606

Urquía D, Gutierrez B, Pozo G, Pozo MJ, Espín A et al. (2019). *Psidium guajava* in the Galapagos Islands: population genetics and history of an invasive species. PloS one 14 (3):e0203737. doi:10.1371/journal.pone.0203737

Vicente S, Máguas C, Trindade H (2018). Genetic diversity and differentiation of invasive *Acacia longifolia* in Portugal. Web Ecology 18 (1):91-103. doi:10.5194/we-18-91-2018

Ward SM, Gaskin JF, Wilson LM (2008). Ecological genetics of plant invasion: what do we know? Invasive Plant Sci Manag 1 (1):98-109. doi :10.1614/IPSM-07-022.1

Wu Q, Zang F, Ma Y, Zheng Y, Zang D (2020). Analysis of genetic diversity and population structure in endangered *Populus wulianensis* based on 18 newly developed EST-SSR markers. Global Ecology and Conservation 24 e01329.doi:10.1016/j.gecco.2020.e01329

Yadav HK, Ranjan A, Asif MH, Mantri S, Sawant SV et al.(2011). EST-derived SSR markers in *Jatropha curcas* L.: development, characterization, polymorphism, and transferability across the species/genera. Tree Genet Genomes 7 (1):207-219. doi:10.1007/s11295-010-0326-6

Yang S, Xue S, Kang W, Qian Z, Yi Z (2019). Genetic diversity and population structure of *Miscanthus lutarioriparius*, an endemic plant of China. PloS one 14 (2):e0211471. doi :10.1371/journal.pone.0211471

Yeh F, Boyle T, Rongcai Y, YE Z (1999). POPGENE: Microsoft window-based freeware for population genetic analysis In. Edmonton, Alberta,Canada: University of Alberta Press.

Supporting Information

Table S1. Voucher information for taxa used in the analyses, representing a total of 35 *Trianthema portulacastrum* L. individuals collected at Fayoum depression.

Sample code	Locality	Latitude	Longitude	Elevation (m.a.s.l.)
T1 E	Daniel, Etsa, Fayoum Governorate	29°7′0.9″ N	30°45′6.9″E	17
T2 E	Al Gharq-Qebli, Etsa, Fayoum Governorate	29°7′51″ N	30°42′18.3″E	8
T3 E	Al Gharq-Fayoum way, Izbat Ad Daw, Minya, Etsa	29°12′44.4″ N	30°45′50.1″E	10
T4 E	Al Gaafra - Minya Al Hayt, Minya, Etsa, Faiyum Governorate	29°13′32.8″ N	30°45′29.5″E	12
T5 E	Minya Al Hayt -Abu Jandir way,Al Awfi, Etsa, Fayoum Governorate	29°14′8.6″ N	30°42′3.5″E	5
T6 E	Madinat Al Fayoum - Ibshawy way, Gerdo, Etsa, Faiyum Governorate	29°18′01.8″N	30°43′12.6″E	15
T7 F	Hawaret Al Maqtaa, Ezbet Ali Farag - Al Hadqa, Fayoum Governorate	29°15′8.2″ N	30°53′32.3″E	23
T8 F	Ezbet Ali Farag - Al Hadqa,Al Hadeqah, Al Fayoum, Fayoum Governorate	29°16′2.9″ N	30°50′42.5″E	22
T9 F	Abgig, Al Fayoum, Fayoum Governorate	29°16′56.351″ N	30°48′57.67″E	16
T10 F	Senofar, Al Fayoum, Fayoum Governorate	29°17′12.6″N	30°52′41.2″E	27
T11 F	Madinat Al Fayoum - Ibshawy gate on ring road, Abgig-Al Fayoum, Fayoum Governorate	29°17′55.397″ N	30°48′14.153″E	19
T12 F	Qesm Al Fayoum, Al Fayoum, Fayoum Governorate	29°18′01.0″N	30°49′03.1″E	20
T13 F	Kofour an Nil,Al Fayoum, Fayoum Governorate	29°18′ 46.63″ N	30°53′23.14″E	20
T14 F	Al Eelam, Al Fayoum, Fayoum Governorate	29°19′38.2″N	30°52′07.5″E	20
T15 F	El-Mandara, Al Fayoum, Fayoum Governorate	29°19′54.9″N	30°48′34.8″E	20
T16 F	Al Edwah, Al Fayoum, Fayoum Governorate	29°19′58.7″ N	30°55′52″E	18
T17 F	Zawyet Al Kerdaseya, Bani Saleh, Al Fayoum, Fayoum Governorate	29°21′18.8″N	30°48′18.9″E	21
T18 S	Behmo, Senoures, Fayoum Govern	29°22′27.1″ N	30°50′51.4″E	16
T19S	Madinat Al Fayoum - Kafr Mahzooz way, Matar Tares, Senoures, Fayoum Governorate	29°22′46.1″ N	30°54′17.1″E	14
T20S	Madinat Al Fayoum - Tersa way, Naqalifah, Senoures, Fayoum Governorate	29°24′39.8″ N	30°49′31.9″E	–2
T21S	Ezbet Mohammed Mahfouz ,Madinet Senouris, Senoures, Fayoum Governorate	29°26′08.3″N	30°53′03.4″E	–13
T22 T	Qasr Rashwan, Tamia, Fayoum Governorate	29°27′30.4″N	30°55′22.5″E	–12
T23T	Madinet Tameyah, Tamia, Fayoum Governorate	29°29′02.2″N	30°56′28.1″E	–13
T24T	Madinet Tamia, Tamia, Fayoum Governorate	29°29′25.5″N	30°58′41.9″E	–8
T25 T	Monshaat Doctor El-Gammal, Tamia, Fayoum Governorate	29°30′32.1″N	31°03′21.3″E	8
T26T	Kafr Al Maslat - Tamya, Fanous, Tameyah, Faiyum Govern	29°32′52.5″N	30°58′55.9″E	12
T27 B	Madinat Al Fayoum - Tohbar way, Al Agameyin, Ibshawy, Fayoum Governorate	29°19′48.2″N	30°42′58.6″E	14
T28B	Al Agameyin-Ibsheway way,Zaid, Ibshawy, Fayoum Governorate	29°20′57.9″N	30°41′26.0″E	17
T29 B	Qasr Bayad, Ibshawy, Fayoum Governorate	29°21′30.2″N	30°44′12.1″E	18
T30 Y	Al Hamouli, Youssef El-Seddik, Fayoum Governorate	29°17′10″N	30°37′15.4″E	0
T31 Y	An Nazlah, Youssef El-Seddik, Fayoum Gove	29°17′59.8″N	30°38′14.4″E	–3
T32 Y	Qasr Al Gabali, Youssef El-Seddik, Fayoum Governorate	29°19′57.8″N	30°37′33.2″E	6
T33 Y	Al Shawashna - Ezbet Gabal Saed,Qasr Al Gabali, Youssef El-Seddik, Fayoum Governorate	29°21′04.9″N	30°38′38.4″E	–46
T34 Y	Izbat Burish Ash Sharqiyyah, Qarun Lake Touristic Road, Al Mashrak, Youssef El Seddik, Fayoum Governorate	29°24′40.0″N	30°33′30.6″E	–33
T35 Y	Kahk, Youssef El-Seddik, Fayoum Governorate	29°26′02.0″N	30°38′39.4″E	–42 m

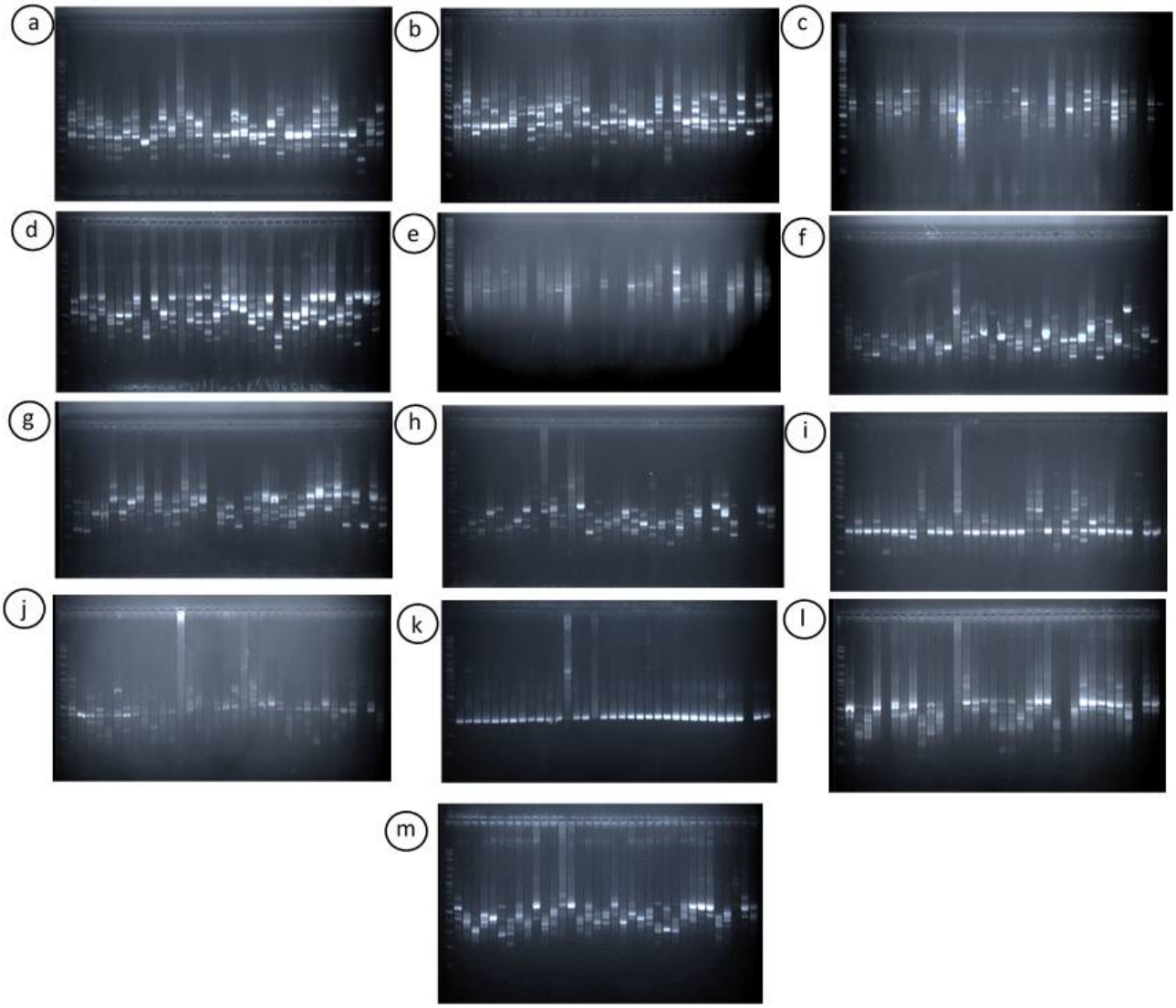

Figure S1. SCOT profiles by PCR amplification using 13 SCoT primers in the 35 *T. portulacastrum* individuals. The first column refers to the molecular marker used; a) SCOT1, b) SCOT11, c) SCOT12, d) SCOT14, e) SCOT15, f) SCOT16, g) SCOT23, h) SCOT25, h) SCOT25, i) SCOT28, j) SCOT29, k) SCOT31, l) SCOT33, m) SCOT33.

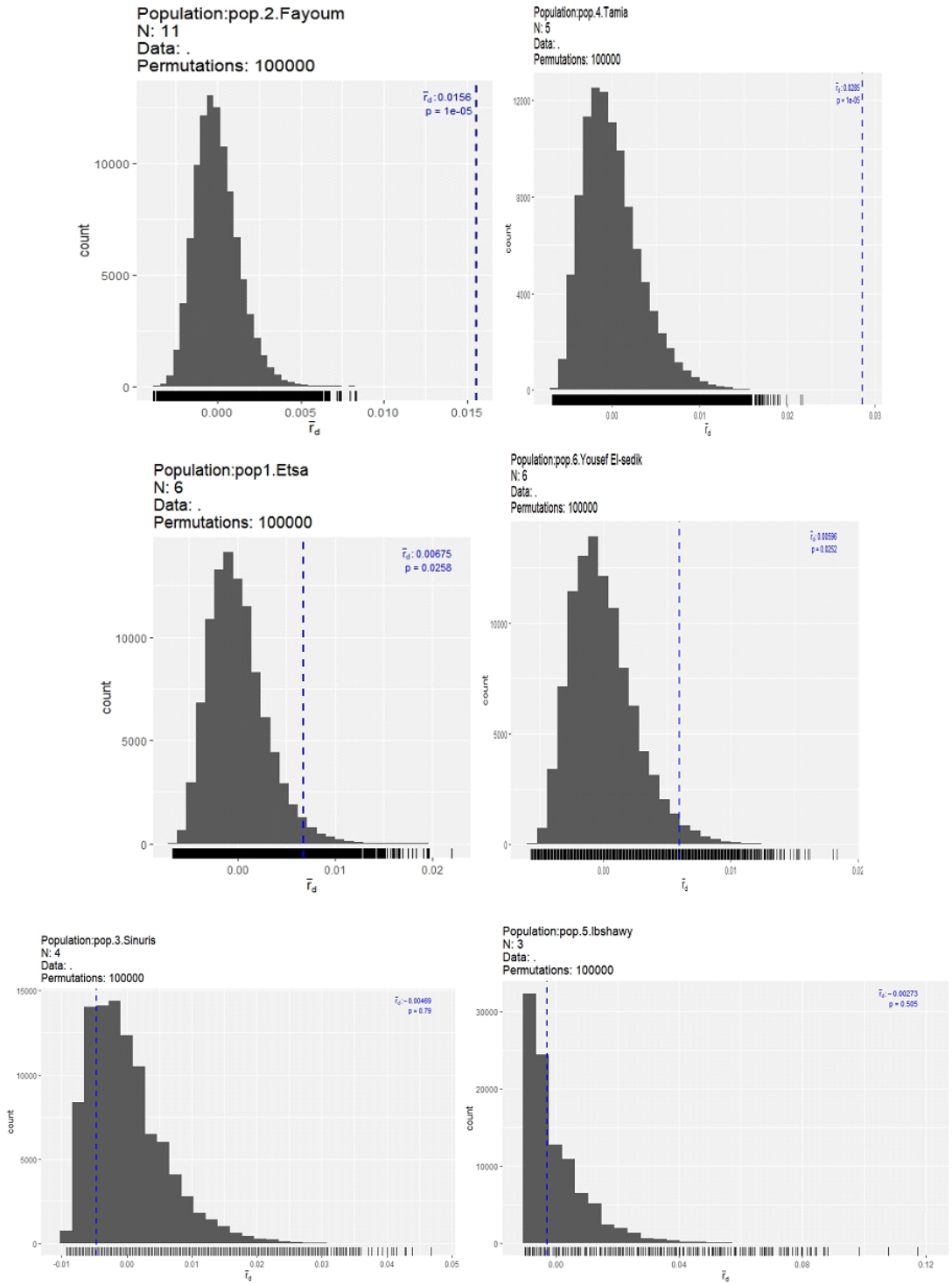

Figure S2. Linkage disequilibrium test visualization, in which observed values (blue dashed lines) of ($\bar{r}_d$) are compared to histograms showing results of 100.000 permutations. a) & b) clone-corrected data rejects the hypothesis of no linkage among markers at Tamia and Fayoum regions with *p-value* ($\bar{r}_d$) =1e-05 and standardized index of association ($\bar{r}_d$) (0.0285 and 0.0156) respectively, and falls outside of the distribution expected under no linkage. c) &d) clone-corrected data rejects the hypothesis of no linkage among markers at Etsa and Yousef El-seddik regions with *p-value* ($\bar{r}_d$) =0.02. Thus, the null hypothesis was rejected and suggested no linkage among markers, however, ($\bar{r}_d$) falls on the right tail of the resampled distribution (0.00675 and 0.00596), values expected under no linkage. e) &f) clone-corrected data failed to reject the linkage disequilibrium hypothesis at Senouris and Ibshawy populations with *p-values* ($\bar{r}_d$) = 0.794 and 0.502 and negative values of ($\bar{r}_d$) (–0.00469 and –0.00273), located inside the distribution expected from unlinked loci.