

# Excess heterozygosity and scarce genetic differentiation in the populations of *Phoenix dactylifera* L.: human impact or ecological determinants

Sakina Elshibli\* and Helena Korpelainen

Department of Applied Biology, PO Box 27, FI-00014 University of Helsinki, Finland

Received 9 May 2008; Revised 21 July 2008; Accepted 23 July 2008 – First published online 20 October 2008

## Abstract

Extensive research has been conducted on the characterization of hundreds of date palm (*Phoenix dactylifera* L.) cultivars worldwide. However, the population genetics of date palms has never been studied. In this study, we collected 200 individuals from 19 populations of different geographical locations in Sudan. The collection sites were grouped according to the type of dates (date palm fruits) that dominates in the area. Ten microsatellite markers were used to investigate the genetic diversity within and among populations and the correlation between the genetic and geographical distances. The tested microsatellite markers showed a high level of polymorphism. A total of 261 alleles were detected at the ten loci. The overall mean value of fixation indices equalled  $-0.163$ , which shows the presence of excess heterozygosity. However, the  $\chi^2$  tests conducted for every locus in each population indicated no significant deviation from the Hardy–Weinberg equilibrium. The analyses of molecular variance exhibited that about 95% of the total genetic variation existed within populations, while significant differentiation within the type groups could be detected. Although significant isolation by distance ( $r^2 = 0.552$ ,  $P < 0.05$ ) was detected by a Mantel test, it seems that the spatial effect has become complicated as a result from the exchange and introduction of different kinds of plant materials by date palm growers and traders, as well as seed dispersal. This complexity was clearly apparent in the weak clustering relationships among most of the tested populations.

**Keywords:** *Phoenix dactylifera*; heterozygosity; genetic; differentiation; microsatellite markers; populations

## Introduction

Thousands of date palm cultivars exist in different growing countries. These cultivars have been developed by continuous selection performed by date palm growers all over the world to improve crop yield and quality. Extensive research has been conducted on the characterization of date palm cultivars regarding a wide range of morphological and chemical characters, and many studies have focused on investigating the genetic diversity and relationships among date palm cultivars using different molecular markers (Sedra *et al.*, 1998; Cao and Chao,

2002; Al-Khalifah and Askari, 2003; Soliman *et al.*, 2003; Zehdi *et al.*, 2004a,b; El-Assar *et al.*, 2005; Adawy *et al.*, 2006; Elshibli and Korpelainen, 2008). While high genetic diversities have been reported among date palm cultivars, cluster analyses based on different genetic markers have been unable to discriminate cultivars from different geographical locations (Cao and Chao, 2002; Zehdi *et al.*, 2004a,b; El-Assar *et al.*, 2005; Elshibli and Korpelainen, 2008). Besides genetic markers, cultivar characterization also includes, for example, crown structure, thorn arrangement, leaf shape, leaflet arrangement and seed shape, but the fruit characters remain as the main features that differentiate the cultivars (Nixon, 1950; Elhoumaizi *et al.*, 2002; Zaid and de Wet, 2002a; Jaradat and Zaid, 2004).

\*Corresponding author. E-mail: sakina.elshibli@helsinki.fi

Date palm cultivars are classified, according to the dry matter content, and the concentration and type of sugars, into soft and dry types of dates. In Sudan, the dry type dominates in the area from the Egyptian border in the north of Sudan to the south around Abo-Hamad (19° 32'N, 33° 21'E; Osman, 1984). The geographical change from dry to soft types of dates is gradual. For example, in Sudan, there are some areas in between where the commercial production of both types is successful. The grouping of date cultivars according to the chemical and morphological characters is not always supported by specific genetic characteristics (Sedra *et al.*, 1998; Trifi *et al.*, 2000; Elshibli and Korpelainen, 2008). As cultivar selection and propagation should follow cloning through offshoots, the absence of specific clones in most tested cultivars was surprising, suggesting that the propagation cycles performed by farmers have been interfered by a mixture of plants derived from seeds. Such unexpected observations may support the view that the environment has a strong effect on the array of genotypes, which is then reflected as a wide range of ecotypes and agronomic diversity (Elshibli and Korpelainen, 2008).

In many production areas in the world, serious stresses resulting in considerable losses have been reported on date palm culture. These stresses include, for example, diseases, drought, desertification and floods (Baaziz *et al.*, 2000; Zaid *et al.*, 2002; United Nations Environment Programme, 2007). Millions of date palm trees were lost within one decade in north Africa due to such natural disasters (Baaziz *et al.*, 2000). These facts suggest that large numbers of traditional cultivars in many countries may be diminishing, and it has been suggested that urgent conservation efforts are needed worldwide (Chao and Krueger, 2007).

Date palm is a multipurpose tree, with a history of cultivation and utilization in north Africa and Middle East for at least 5000 years. However, its exact origin is unknown (Wrigley, 1995). Date palm is a diploid, perennial, monocotyledonous plant with separate male and female trees. The tree has two apparently independent propagation strategies: sexual and asexual (Zaid and de Wet, 2002c). Asexual propagation (offshoot propagation) is used for the cloning of selected material. On the other hand, sexual propagation results in large numbers of seeds, and the offspring can possess a range of character combinations, including also traits other than fruit and yield quality. For instance, the characters may include resistance and/or tolerance to biotic and abiotic stresses, and, thus, they may be useful attributes for scientific selection in long-term breeding strategies.

The geographical distribution of date palm covers a wide range of environmental conditions; it grows and flourishes from 392 m below to 1500 m above sea level with an altitude range of 1892 m (Zaid and de Wet, 2002b). Besides the wide range of geographical distribution, also the

biological nature of date palm is expected to affect the genetic structure and culture of date palm worldwide. Yet, the population genetics of date palms has never been studied. In this study, we aim to study the genetic structure and genetic diversity of date palm populations collected from different geographical locations in Sudan. This work is expected to provide new and valuable information for better understanding of the adaptation and ecological roles of date palms.

## Material and methods

### Study sites

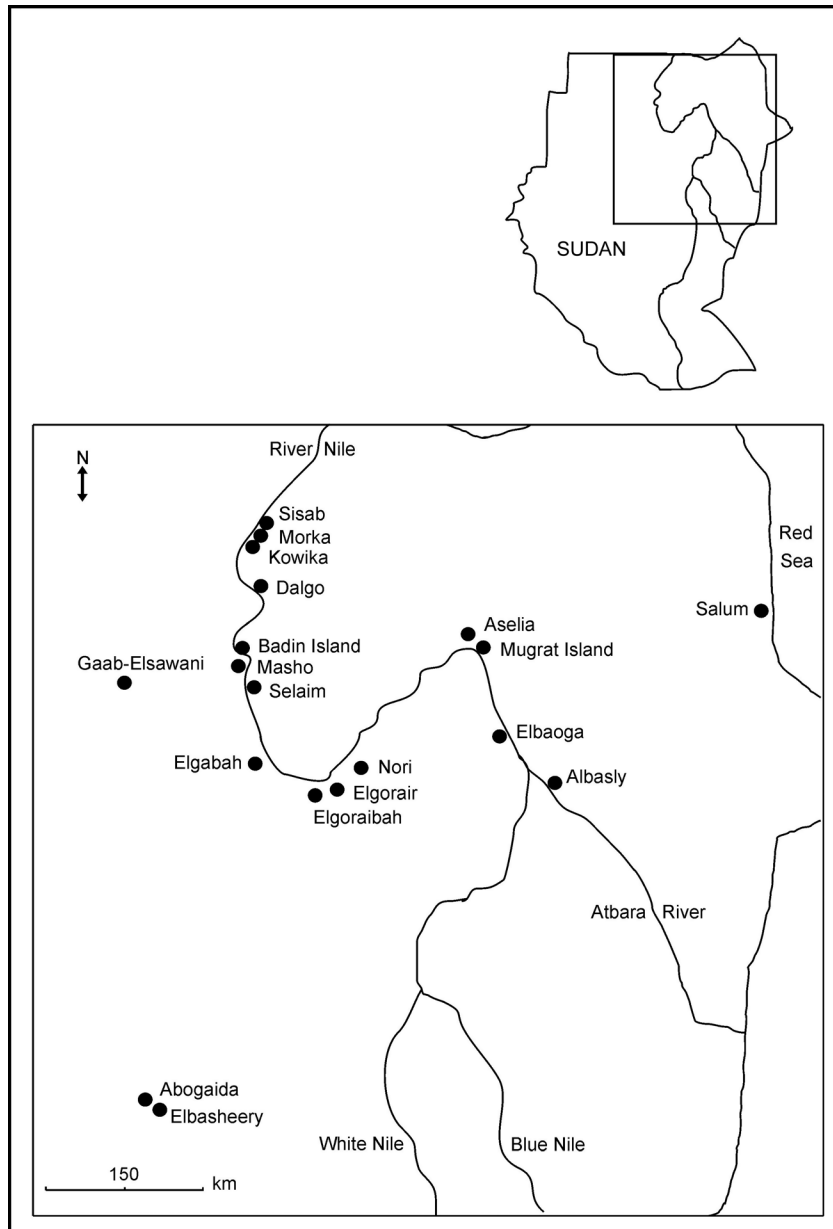
Field work was conducted in Sudan in different locations situated between latitudes 14° and 20° 45'N and longitudes 30° 19' and 35° 7'E. According to rainfall characterization, our research sites belong to three different subregions regarding annual rainfall trends (Hulme, 1990): the arid north (16°N north to the Egyptian border), the semi-arid central region (between 12°N and 16°N) and the winter rainfall region along the Red Sea coast (Fig. 1). To classify different sites in relation to the type of dates that dominate in the area, we used the categorization given by Osman (1984; Table 1).

### Plant material

The plant materials collected from 200 individual trees represented 19 populations of date palms. A single orchard or farm was considered as one population. Ten trees were assigned randomly to each population in all locations except for Nori, Elbaoga and Elbasheery (20, 12 and 8 trees, respectively). Seeds were germinated in sand soil in an incubator for 2 months under a temperature condition of 27°C. Young, healthy leaves were harvested for direct DNA extraction, where each tree was represented by a single seedling.

### Microsatellite analysis

Total genomic DNA was extracted from fresh leaves, using a DNAeasy Plant Mini Kit (Qiagen, Hilden, Germany). Ten pairs of primers were used to amplify (GA)<sub>n</sub> microsatellite loci using markers developed by Billotte *et al.* (2004) for *Phoenix dactylifera*. The amplification reactions were performed in a volume of 5 µl. The 5-µl reaction mixture contained about 3 ng of genomic DNA per each microlitre of reaction volume, 0.1 µl of 10 mM dNTP mix, 0.3 U of DyNAzyme™ II DNA polymerase with a reaction buffer containing 10 mM Tris-HCl (pH 8.8), 1.5 mM MgCl<sub>2</sub>, 50 mM KCl and 0.1% Triton X-100 (Finnzymes) and 0.5 µl of each of



**Fig. 1.** The locations of 19 populations of date palms collected from Sudan and genotyped with ten microsatellite markers.

two 5  $\mu$ M primers. The reverse primers were fluorescently FAM (6-Carboxyfluorescein) labelled at the 5' end. The thermocycler (model PTC-200; MJ Research, Inc., Watertown, MA, USA) was used for amplifications with the following programme: a denaturation step of 4 min at 94°C followed by 35 cycles of denaturation for 45 s at 94°C, annealing at 52°C for 45 s, elongation for 1 min at 72°C and a final elongation step of 8 min at 72°C. The amplification products were first detected using electrophoresis with 1.4% agarose gels and by staining with ethidium bromide. For final analyses, 2  $\mu$ l of amplified DNA (diluted to 1:50) and 12  $\mu$ l of MegaBACE™ ET400-R DNA Size Standard (GE Healthcare Company, Slough, Berkshire,

UK) were loaded onto 96-well plates. Genotyping was carried out using an automatic DNA analyser, MegaBACE 1000 (Amersham Biosciences, Buckinghamshire, UK). Allele size scoring was performed by MegaBACE software, MegaBACE Genetic Profiler 2.2 (Amersham Biosciences, 2003).

### Data analysis

Basic descriptive statistics, including genetic diversity indices [number of alleles, observed heterozygosity ( $H_{obs}$ ), expected heterozygosity ( $H_{exp}$ )] in each population, as well as the estimator  $F_{ST}$  (Wright, 1969) and

**Table 1.** Geographical description, type of dates (TD), number of polymorphic loci (*P*), mean number of alleles (*A*), expected heterozygosity ( $H_{exp}$ ), observed heterozygosity ( $H_{obs}$ ) and fixation index (*F*) for the populations of *Phoenix dactylifera*

Location	Latitude	Longitude	TD	<i>P</i>	<i>A</i>	$H_{exp}$	$H_{obs}$	<i>F</i>
Sisab (Sai Island)	20° 45'	30° 19'	Dry	10	6.5	0.781	1.000	-0.280
Morka (Sai Island)	20° 43'	30° 20' 45"	Dry	9	6.7	0.810	1.000	-0.235
Kowika	20° 40' 30"	30° 21' 15"	Dry	10	5.9	0.773	1.000	-0.294
Dalگو	20° 7'	30° 40 45"	Dry	7	6.9	0.802	0.990	-0.234
Badin Island	19° 39' 30"	30° 23' 30"	Dry	10	5.8	0.796	0.880	-0.106
Masho	19° 33'	30° 36'	Dry	10	6.5	0.797	0.990	-0.242
Selaim	19° 8' 54"	30° 30'	Dry	10	7.8	0.827	0.910	-0.100
Gaab-Elsawani	19° 3' 30"	30° 6' 15"	Dry	8	7.0	0.801	0.852	-0.064
Elgabab	18° 9'	30° 45'	Dry	8	7.9	0.842	0.927	-0.101
Nori	18° 32' 45"	31° 54' 15"	Mixed	10	11.3	0.845	0.888	-0.051
Elgorair	18° 16' 10"	31° 41' 20"	Mixed	10	6.9	0.822	1.000	-0.217
Elgoraibab	18° 7' 30"	31° 35' 30"	Mixed	9	5.8	0.789	0.960	-0.217
Aselia	19° 33'	33° 9' 15"	Soft	9	6.4	0.793	0.940	-0.185
Mugrat Island	19° 30' 45"	33° 15'	Soft	10	7.4	0.826	0.900	-0.090
Elbaoga	18° 17'	33° 54' 30"	Soft	10	8.8	0.849	0.867	-0.021
Albasly	17° 28'	34° 10' 30"	Soft	10	7.0	0.819	0.940	-0.148
Abogaida	14° 30'	30° 1'	-	9	7.5	0.832	0.869	-0.044
Elbasheery	14°	30° 10'	-	9	5.3	0.773	0.975	-0.261
Salum	19° 23' 22"	37° 6' 39"	-	10	6.5	0.823	1.000	-0.215

All *F* values were not significantly deviated from the Hardy–Weinberg equilibrium ( $P > 0.05$ ).

the test for the Hardy–Weinberg equilibrium, were calculated using Arlequin 3.11 (Excoffier *et al.*, 2005). The sequential Bonferroni corrections were applied when necessary (Rice, 1989). The fixation indices were calculated using the formula  $F = 1 - H_{obs}/H_{exp}$ . The significance of the deviation of the fixation index from zero was tested using a  $\chi^2$  test (Li and Horwitz, 1953). Allele frequencies were used in the parsimony analysis conducted using PHYLIP 3.65 (Felsenstein, 2005); Gendist was used to compute Nei's genetic distances (Nei, 1972), Kitch to estimate phylogenies from distance matrix data and Consense to find the consensus tree applying the extended majority rule. Support for the inferred tree topology was evaluated using the bootstrap approach (Felsenstein, 1985) with 1000 replicates.

Analyses of molecular variance AMOVA; Excoffier *et al.*, 1992) were performed at different hierarchical levels. We classified the collection sites according to the type of dates that dominates in the area (Fig. 1) as follows: the dry type production area includes the sites from Elgabab to the north, the soft type production area ranges from Aselia south to the Albasly along the Atbara River, and the area between Elgabab and Aselia, which is regarded as a separate group because there is considerable production of both types of dates. The sites in the Elbasheery oases (the Abogaida and Elbasheery populations) in the central semi-arid region of the country in north Kordofan are considered as a separate group apart from the River Nile side, as also the collections from Salum along the Red Sea coast in the east side of the country. Analyses of molecular variance were also performed separately for the dry, soft

and mixed types of date populations, as well as for the populations in Elbasheery oases. The Salum samples representing only one population were excluded in these analyses. For each analysis, 1000 permutations were performed to obtain significance levels of variance using Arlequin 3.11 (Excoffier *et al.*, 2005). To test the presence of isolation by distance, a Mantel test (Mantel, 1967) was applied for data including pairwise geographical distances and  $F_{ST}/(1 - F_{ST})$  values between populations (Rousset, 1997). The coordinates of different locations (Table 1) were defined and given by the General Administration for Survey in the River Nile State in Sudan (2007). The Aselia and Abogaida locations were approximated using the travelling distance from Mugrat Island and Elbasheery locations, respectively. The geographical distances were calculated using the Longitude Latitude Distance Calculator version 4.2.3 (<http://www.csgnetwork.com/latlongdistcalc.html>, 2007).

## Results

For the pooled samples, the ten microsatellite primer pairs produced 261 alleles with high levels of expected (0.771–0.906) and observed heterozygosities (0.813–0.995; Table 2) per locus. High levels of genetic diversity were observed in all tested populations at the multi-locus level, as indicated by gene diversity estimates ranging from 0.765 to 0.849 with an average of 0.809. Over the ten loci examined, the mean observed heterozygosity ( $H_{obs}$ ) was higher than the expected heterozygosity

**Table 2.** Genetic diversity of pooled samples, including observed and expected heterozygosities and fixation index ( $F$ ) for each locus

Locus	Repeat motif	Total number of alleles	$H_{exp}$	$H_{obs}$	$F$
mpdCIR010	(GA) <sub>22</sub>	30	0.810	0.985	-0.216
mpdCIR015	(GA) <sub>15</sub>	25	0.835	0.980	-0.174
mpdCIR016	(GA) <sub>14</sub>	16	0.771	0.995	-0.291
mpdCIR032	(GA) <sub>19</sub>	28	0.906	0.813	0.103
mpdCIR035	(GA) <sub>15</sub>	26	0.877	0.863	0.016
mpdCIR050	(GA) <sub>21</sub>	33	0.920	0.824	0.104
mpdCIR057	(GA) <sub>20</sub>	35	0.861	0.991	-0.151
mpdCIR070	(GA) <sub>17</sub>	24	0.839	0.985	-0.174
mpdCIR078	(GA) <sub>13</sub>	22	0.869	0.99	-0.139
mpdCIR093	(GA) <sub>16</sub>	22	0.821	0.954	-0.162

All  $F$  values were not significantly deviated from the Hardy–Weinberg equilibrium ( $P > 0.05$ ).

( $H_{exp}$ ) in every date palm population, resulting in a negative fixation index ( $F$ ) with an overall mean value of  $-0.163$  (Table 1). However, the  $\chi^2$  tests conducted for every locus in each population indicated no significant deviation from the Hardy–Weinberg equilibrium.

The  $F_{ST}$  values within populations varied from 0.000 to 0.146 (negative  $F_{ST}$  values considered zero; Table 3). The Salum and Badin Island populations showed significant differentiation compared with each of the other 18 populations. The Aselia population, which belongs to the soft type of dates, was significantly differentiated from all populations belonging to the dry type and from the Nori population, while not significantly differentiated from the Elgoraibah and Elgorair populations. Nori, Elgorair and Elgoraibah are considered as locations for the mixed type of dates (Fig. 1). A similar trend was observed in the Mugrat Island population. Other populations showed either significant or non-significant levels of pairwise differentiation without any relation to the location and/or type of dates produced in the area. This trend was clearly observed in the populations from the Elbaoga, Elbasheery and Abogaida locations.

The Elbasheery population was significantly differentiated from the Abogaida population. Both populations are located in north Kordofan within the Elbasheery oases area, completely isolated from the Salum population on the Red Sea coast and all other populations along the River Nile (Fig. 1). The Elbasheery population showed non-significant differentiation from the Sisab population in the Sai Island in the far northern part of the Sudan, while the Abogaida population possessed the same trend related to the Sai Island (Sisab and Morka), Selaim, Gaab-Elsawani and Elgabah populations along the River Nile in the area of the dry type of dates. The average value of  $F_{ST}$  equalled 0.049, which shows that only about 5% of the genetic variation was present among populations.

When the whole dataset was studied using AMOVA, it was observed that most genetic variability was present among individuals within populations (Table 4a).

However, significant differentiation within the type groups could be detected. A similar trend was observed when populations of different types of dates were analysed separately (Table 4b): most of the genetic variation was observed within populations. Yet, significant differentiation among the populations representing each type of dates was also detected.

A tree was created to test the genetic relationships between different populations (Fig. 2). Bootstrap values do not support most of the clusters. However, each of Morka, Sisab, Masho, Slaim and Kowika constitute a distinct group with bootstrap values ranging from 50 to 57%. On the other hand, the following groups of populations showed stronger genetic relationships compared with other groups: the Nori and Elbaoga populations (86%); the Badin Island and Salum populations (86%); the Elgoraibah and Aselia populations (63%); the Elgoraibah, Aselia and Elgorair populations (54%) and the Elgoraibah, Aselia, Elgorair and Mugrat Island populations (51%).

The geographical distances between population pairs in this study ranged between 4.8 and 2533 km. According to the Mantel test, a significant isolation-by-distance effect was observed when analysing data for all the 19 populations ( $r^2 = 0.552$ ,  $P < 0.05$ ).

## Discussion

The presence of high genetic diversity among date palm cultivars has been reported in Tunisia and Sudan (Zehdi *et al.*, 2004a; Elshibli and Korpelainen, 2008). However, these studies were based on individual collections to investigate the genetic relationships between cultivars; almost all studied samples were references to cultivars (females). The high intrapopulation genetic diversities (mean 0.809) detected in this study indicate that the studied populations possess an abundant source of variation that is needed for evolutionary responses in a changing environment and for breeding purposes to

**Table 3.** Pairwise  $F_{ST}$  values of the 19 populations of date palms based on ten microsatellite loci

Population	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. Sisab	0.000																		
2. Morka	0.000	0.000																	
3. Kowika	0.024	0.027	0.000																
4. Dalgo	0.023	0.028	<b>0.047</b>	0.000															
5. Selaim	0.008	0.011	0.028	0.000	0.000														
6. Badin	<b>0.087</b>	<b>0.085</b>	<b>0.103</b>	<b>0.073</b>	<b>0.054</b>	0.000													
7. Masho	0.002	0.000	0.025	0.029	0.000	<b>0.086</b>	0.000												
8. Gaab-Elsawani	<b>0.050</b>	<b>0.051</b>	<b>0.072</b>	<b>0.052</b>	0.016	<b>0.050</b>	0.021	0.000											
9. Elgabab	0.033	0.030	<b>0.043</b>	<b>0.036</b>	0.026	<b>0.066</b>	<b>0.026</b>	0.037	0.000										
10. Nori	<b>0.044</b>	<b>0.043</b>	<b>0.071</b>	<b>0.031</b>	0.021	<b>0.037</b>	<b>0.044</b>	0.015	0.018	0.000									
11. Elgorair	<b>0.058</b>	<b>0.044</b>	<b>0.067</b>	<b>0.064</b>	0.044	<b>0.110</b>	<b>0.047</b>	<b>0.087</b>	<b>0.043</b>	<b>0.057</b>	0.000								
12. Elgoraibab	0.038	<b>0.045</b>	<b>0.058</b>	<b>0.047</b>	0.034	<b>0.108</b>	<b>0.042</b>	<b>0.071</b>	0.026	<b>0.040</b>	0.000	0.000							
13. Aselia	<b>0.075</b>	<b>0.076</b>	<b>0.098</b>	<b>0.088</b>	<b>0.078</b>	<b>0.099</b>	<b>0.081</b>	<b>0.092</b>	<b>0.036</b>	<b>0.047</b>	0.006	0.000	0.000						
14. Mugrat Island	<b>0.067</b>	<b>0.059</b>	<b>0.080</b>	<b>0.071</b>	0.008	<b>0.054</b>	0.041	<b>0.054</b>	0.019	<b>0.041</b>	0.021	0.013	0.009	0.000					
15. Elbaoga	0.031	<b>0.033</b>	<b>0.051</b>	0.008	0.011	<b>0.037</b>	0.027	0.028	0.003	0.000	<b>0.051</b>	0.024	<b>0.039</b>	0.023	0.000				
16. Albasly	0.025	0.025	<b>0.044</b>	<b>0.026</b>	0.016	<b>0.083</b>	0.011	<b>0.048</b>	0.008	<b>0.027</b>	<b>0.033</b>	0.005	0.023	0.003	0.009	0.000			
17. Elbasheery	0.021	<b>0.038</b>	<b>0.053</b>	0.029	0.030	<b>0.096</b>	0.038	<b>0.069</b>	0.034	<b>0.046</b>	<b>0.067</b>	0.045	<b>0.073</b>	<b>0.070</b>	0.036	0.006	0.000		
18. Abogaitda	0.030	0.036	<b>0.051</b>	<b>0.037</b>	0.009	<b>0.044</b>	0.029	0.017	0.018	0.018	<b>0.059</b>	0.036	<b>0.049</b>	0.034	0.012	0.023	<b>0.049</b>	0.000	
19. Salum	<b>0.112</b>	<b>0.115</b>	<b>0.132</b>	<b>0.109</b>	<b>0.101</b>	<b>0.123</b>	<b>0.118</b>	<b>0.116</b>	<b>0.095</b>	<b>0.099</b>	<b>0.128</b>	<b>0.137</b>	<b>0.146</b>	<b>0.129</b>	<b>0.093</b>	<b>0.109</b>	<b>0.109</b>	<b>0.110</b>	0.000

Significant values are shown in bold.

**Table 4.** Analyses of molecular variance (AMOVA) for 19 populations of *Phoenix dactylifera* L. based on ten microsatellite markers: (a) populations were structured into five groups based on geographical location and type of dates, (b) AMOVA performed separately for each group of populations from production areas representing different types of dates

Samples	Source of variation	df	Variance components	Variance (%)	Fixation index	<i>P</i> value
(a)						
All groups	Among groups	4	0.083	1.97	0.020 ( $F_{CT}$ )	0.002
	Among populations within groups	14	0.141	3.33	0.034 ( $F_{SC}$ )	0.000
	Within populations	381	4.005	94.70	0.053 ( $F_{ST}$ )	0.000
(b)						
Dry type	Among populations	8	0.154	3.76	0.038 ( $F_{ST}$ )	0.000
	Within populations	171	3.946	96.24		
Soft type	Among populations	3	0.077	1.85	0.019 ( $F_{ST}$ )	0.001
	Within populations	80	4.100	98.15		
Mixed type	Among populations	2	0.167	3.97	0.040 ( $F_{ST}$ )	0.000
	Within populations	77	4.039	96.03		
Elbasheery oases	Among populations	1	0.165	4.90	0.049 ( $F_{ST}$ )	0.004
	Within populations	34	3.207	95.10		

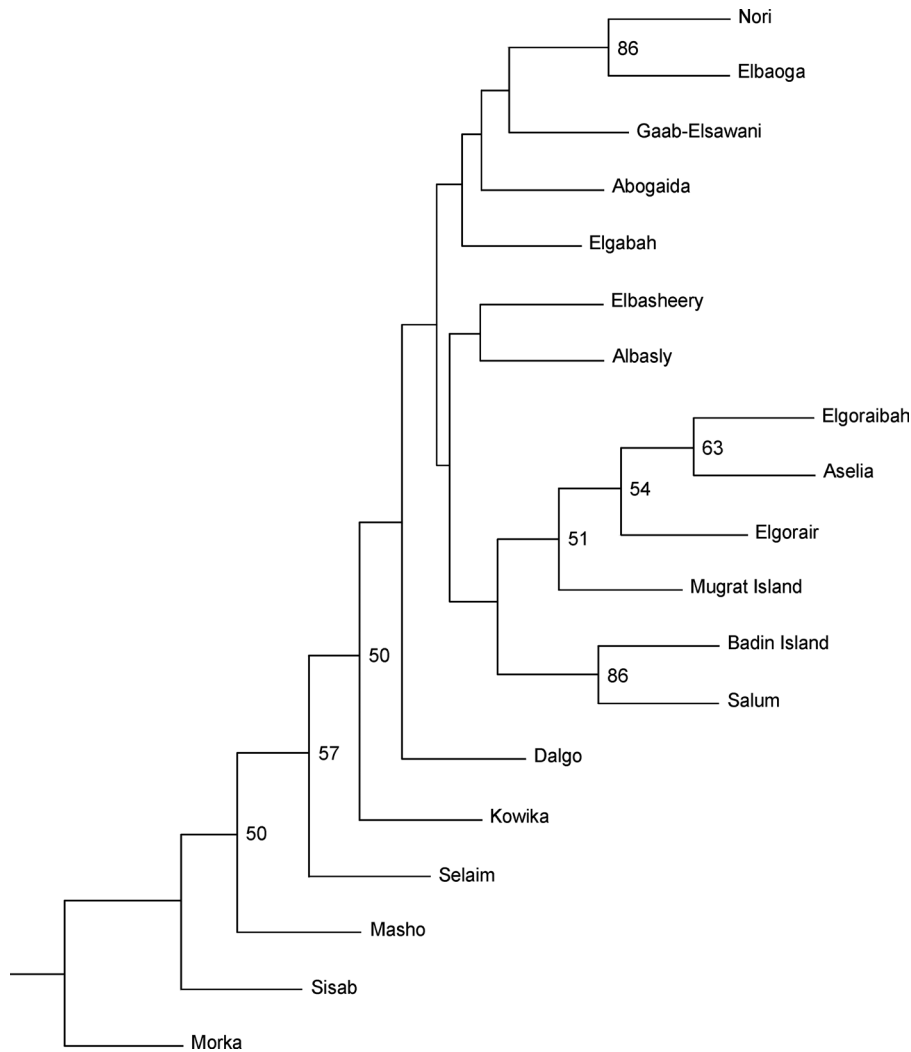
improve the agronomical and commercial characters of date palm. The observed levels of genetic diversity are in line with those reported in other woody, wind-pollinated tree species. Zhang *et al.* (2006) reported an average expected heterozygosity equalling 0.779 among five natural populations of *Quercus aquifolioides* examined with microsatellite markers. Moreover, tree species with both sexual and asexual modes of reproduction may maintain higher levels of genetic diversity (Hamrick *et al.*, 1992), as it was observed in this study on date palm with the mean  $H_{exp}$  equalling 0.851.

The excess heterozygosity of the 19 tested populations, as indicated by the negative mean values of fixation indices, is not surprising, as 100% cross-pollination of date palm trees results in the hybridization of genotypes. This result is in line with our previous study on the genetic diversity of a wide range of date palm cultivars (Elshibli and Korpelainen, 2008), in which we suggested that the exchange of offshoots (propagules) of date palm cultivars may result in heterozygous populations. However, the distribution of genetic constituents may result also from the introduction of seedlings after fruit consumption. Yet, the deviations of the observed heterozygosity from the Hardy–Weinberg equilibrium were generally non-significant.

In general, the microsatellite loci exhibited low  $F_{ST}$  values among different populations of date palms, which show that the observed high genetic variability is maintained locally within populations. These values agree with those observed by Zehdi *et al.* (2004a) and Elshibli and Korpelainen (2008) in different groups of date palm cultivars collected from Tunisia and Sudan. According to the measured  $F_{ST}$  values, generally there is no clear trend in the differentiation of populations originating from different locations. However, the distinct differentiation of the Salum population (mean

$F_{ST} = 0.116$ ,  $P < 0.001$ ) from all other 18 populations is not surprising and can be mainly attributed to the effect of geographical distance (2530–2015 km). Other characteristic features of the Salum location on the Red Sea coast include the rainy winter season and the altitude of the Red Sea Hills (up to 2000 m; Google Earth, 2008; Babiker and Gudmundsson, 2004) situated between the Salum location (173 m) and the River Nile banks.

Pollination of date palm naturally takes place by wind, bees and insects, and is found to yield a fair fruit set in various areas within the date palm growing countries (Zaid and de Wet, 2002d). Natural pollination of date palms is characterized by 100% seedling composition. However, commercial date production necessitates artificial pollination. Date growers are aware of the importance of artificial pollination and practise it according to different traditional methods or by using a mechanized device. Pollen grains selected from male trees according to the well-verified effects of xenia (effect of pollen grains on the size of fruits and seeds) and metaxenia (effect of pollen grains on the time of fruit ripening; Ahmed and Ali, 1960; Nixon and Carpenter, 1978; Zaid and de Wet, 2002d). It is worth mentioning that pollen grains are usable as fresh or dried and can be stored for the next season as well. The use of dry stored pollen grains is essential when early female blooms take place before adequate numbers of male spathes open, a situation well experienced by date palm growers (Zaid and de Wet, 2002d). On the other hand, pollen grains are stored and exchanged for their favourable metaxenia effect. In this regard, the dispersal of pollen grains is expected to be more extensive, and the exchange process between farmers is expected to be more effective along the River Nile rather than in the Salum location, where the dispersal of pollen grains may remain effective only within the location.



**Fig. 2.** Neighbour-joining tree illustrating the relationships among 19 populations of date palms collected from Sudan. The tree is based on Nei's, (1972) pairwise genetic distances. The primary allele frequency data were bootstrapped 1000 times. Only bootstrap values of at least 50% are shown.

The significant differentiation between the two populations from the Elbasheery oases, accompanied by non-significant differentiation of one of these populations from date palms in the Sai Island (Sisab population; Fig. 1) was surprising, especially if we consider the distance between the two sites (750 km) and the location of the Elbasheery oases apart from the River Nile and other date palm production areas. It is clear that some factors other than geographical distances play an important role in affecting the gene flow and genetic structure of date palm populations worldwide. These factors may include the exchange of propagation and/or pollination material, and seed dispersal. Within Sudan, it has been reported that in the 1960s the Horticultural Administration in Sudan collected 86,000 offshoots of date palm from the Halfa Valley (21° 47'N, 31° 22'E), where they were under the risk of flooding during the construction

of the High Dam. These offshoots were redistributed to farmers in other regions of the country (Nixon, 1967). On the other hand, at some time, the Horticultural Administration in Sudan served the farmers by collecting date palm offshoots from their orchards for rooting, and redistributed them to interested farmers, while some cultivars were introduced from Tunisia and Saudi Arabia (Osman, 2001).

The populations in the Elbasheery oases showed moderate gene flow ( $F_{ST} = 0.05$ ) compared with other groups of populations. The isolation by distance from other populations present along the River Nile may cause difficulties in exchanging and dispersing plant materials to those locations. This situation may lead to the development of small isolated populations of date palms in the Elbasheery oases and highlights the need for better management of date palm culture in that region.



The observed high degree of genetic variability within populations compared with relatively low levels of genetic differentiation among the 19 populations, detected also when the groups were analysed separately, may confirm the idea of effective dispersal, which decreases population structuring. In date palm, the dispersal agents include seed, pollen grains and offshoots. Offshoots and pollen grains are mainly distributed between farmers within a village, province or country, while seed dispersal takes place by travellers and traders across geographical borders. The dispersal effect on the effective population size is a function of the variance in dispersal distances (Loveless and Hamrick, 1984). Long-distance seed dispersal and propagule establishment prevent population divergence. On the other hand, date palm trees have a characteristic long-lived cycle in nature. Although the average economic life of a date palm tree is estimated to be up to 50 years, the tree may stay productive up to 150 years (Chao and Krueger, 2007). Such long generation times also increase the effective population size and prevent population divergence (Loveless and Hamrick, 1984).

The presence of considerable variability ( $P < 0.01$ ; Table 4a) between date palm populations as related to the type of date production area agrees with the general classification of these areas according to the morphological and chemical characters of fruits and cultivars. When moving from northern Sudan at the Egyptian border to the south, the type of date palm fruits tends to change from the dry to semi-soft type and to the soft type. This trend seems to follow the degree of aridity, varying from the arid region in the north to the semi-arid region in the central part of the country (Hulme, 1990), including the Khartoum area where all commercial date production comprises soft type dates. Unfortunately, the Khartoum area was not included in this study because of a failure in seed germination. Although our results significantly contribute to the knowledge of date palm, it is clear that the classification of different types of dates related to the production area and the characteristic environmental features need to be investigated further in detail in future studies.

Significant ( $P < 0.05$ ) isolation by distance, revealed by a Mantel test, was observed. This observation was detectable for the Salum location. However, it seems that the spatial effect has become complicated as a result from the exchange and introduction of different kinds of plant materials by date palm growers and traders. This complexity was apparent in the weak clustering relationships among most of the tested populations, while a close relationship was detected, for example, between the populations from the Badin Island and Salum on the Red Sea coast. This kind of relationship can be attributed mainly to the human activity in exchanging disparate material. However, there is no specific statistical

data of the activity of date palm growers in these locations, which could further explain the observed spatial genetic pattern of the populations. Comparable relationships have also been reported by Sedra *et al.* (1998); Cao and Chao (2002); Zehdi *et al.* (2004b) and Elshibli and Korpelainen (2008), who independently tested cultivar samples from Tunisia, California, Morocco and Sudan with different molecular markers.

The available evidence indicates that the genetic structure of date palms is governed by multiple factors. Three main factors, which indicate the involvement of a natural mechanism, were apparent as affecting the genetic differentiation of the trees in the form of populations and/or cultivar groups. These factors include: the geographical isolation by distance, the biological nature of the tree and the environmental characters of the specific region, e.g. the degree of aridity is represented strongly in fruit characters and slightly in population differentiation. However, the history of cultivation and cultural practices are difficult to trace but may constitute a stronger effect in the structuring of date palms. The complex exchange processes of date palm germplasm throughout the world following cultivar and/or pollen grain selection, along with the random distribution of dispersal agents point out that, among other factors, the human impact possibly creates the highest influence on structuring date palm genome worldwide.

## Acknowledgements

We are sincerely grateful to Ms Amal Elshibli of the General Administration for Survey in the River Nile State in Sudan for technical help in providing and calculating the coordinates of different locations used in this study when needed. This work was financially supported by the Emil Aaltonen Foundation.

## References

- Adawy SS, Hussein EHA, Ismail SEME and El-Itriby HA (2006) Genomic diversity in date palm (*Phoenix dactylifera* L.) as revealed by AFLPs in comparison to RAPDs and ISSRs. In: *Abstracts Of 3rd International Date Palm Conference*, Abu Dhabi, United Arab Emirates, 19–21 February 2006.
- Ahmed M and Ali N (1960) Effect of different pollens on the physical and chemical characters and ripening of date fruit. *Punjab Fruit Journal* 23: 10–11.
- Al-Khalifah NS and Askari E (2003) Molecular phylogeny of date palm (*Phoenix dactylifera* L.) cultivars from Saudi Arabia by DNA fingerprinting. *Theoretical and Applied Genetics* 107: 1266–1270.
- Baaziz M, Majourhat K and Bendiab K (2000) Date palm culture in the Maghreb countries: constraints and scientific research. *Proceedings of the Date Palm International Symposium*, Windhoek, Namibia, 22–25 February, pp. 306–311.

- Babiker M and Gudmundsson A (2004) The effects of dykes faults on groundwater flow in an arid land: The Red Sea Hills, Sudan. *Journal of Hydrology* 279: 256–273.
- Billotte N, Marseilla CN, Brottier P, Noyer JL, Jacquemoud-Collet JP, Moreau C, *et al.*, (2004) Nuclear microsatellite markers for the date palm (*Phoenix dactylifera* L.): characterisation, utility across the genus *Phoenix* and in other palm genera. *Molecular Ecology Notes* 4: 256–258.
- Cao BR and Chao CT (2002) Identification of date cultivars in California using AFLP markers. *HortScience* 37: 966–968.
- Chao CT and Krueger RR (2007) The date palm (*Phoenix dactylifera* L.): overview of biology, uses and cultivation. *HortScience* 42: 1077–1082.
- El-Assar AM, Krueger RR, Devanad PS and Chao CT (2005) Genetic analysis of Egyptian date (*Phoenix dactylifera* L.) accessions using AFLP markers. *Genetic Resources and Crop Evolution* 52: 601–607.
- Elhoumaizi MA, Saaïdi M, Oihabi A and Cilas C (2002) Phenotypic diversity of date-palm cultivars (*Phoenix dactylifera* L.) from Morocco. *Genetic Resources and Crop Evolution* 49: 483–490.
- Elshibli S and Korpelainen H (2008) Microsatellite markers reveal high genetic diversity in date palm (*Phoenix dactylifera* L.) germplasm from Sudan. *Genetica* (in press). DOI 10.1007/s10709-007-9232-8.
- Excoffier L, Laval G and Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics* 1: 47–50.
- Excoffier L, Smouse PE and Quattro M (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human Mitochondrial DNA restriction data. *Genetics* 131: 479–491.
- Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783–791.
- Felsenstein J (2005) PHYLIP (Phylogeny Interference Package) version. 3.6. Distributed by the author. Department of Genome Sciences, University Washington, Seattle.
- Hamrick JL, Godt MJW and Sherman-Broyles SL (1992) Factors influencing levels of genetic diversity in woody plant species. *New Forests* 6: 95–124.
- Hulme M (1990) The changing rainfall resources of Sudan. *Transactions of the Institute of British Geographers* 15: 21–34.
- Jaradat AA and Zaid A (2004) Quality traits of date palm fruits in a center of origin and center of diversity. *Food, Agriculture and Environment* 2: 208–217.
- Loveless MD and Hamrick JL (1984) Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics* 15: 65–95.
- Li CC and Horwitz DG (1953) Some methods of estimating the inbreeding coefficient. *American Journal of Human Genetics* 5: 107–117.
- Rousset F (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics* 145: 1219–1228.
- Rice WR (1989) Analysing tables of statistical tests. *Evolution* 43: 223–225.
- Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research* 27: 209–220.
- Nei M (1972) Genetic distance between populations. *American Naturalist* 106: 283–292.
- Nixon RW (1950) Imported cultivars of dates in the United States USDA Circular No. 834: 144.
- Nixon RW (1967) Date culture in Sudan, Date Growers' Institute Report 44: 9–14.
- Nixon RW and Carpenter JB (1978) Growing dates in the United States US Department of Agriculture, Agriculture Information Bulletin No. 207: USDA. Technical Document, p. 63.
- Osman AMA (1984) The performance of date palms in the Sudan. *Acta Horticulturae* 143: 231–237.
- Osman AMA (2001) Development of date palm culture in Republic of Sudan Paper presented at a workshop on date palm culture and dates production in Republic of Sudan. Date Palm Research and Development Network, Khartoum, Sudan, 22–17 August.
- Sedra MH, Lashermes P, Trouslot P, Combes M and Hamon S (1998) Identification and genetic diversity analysis of date palm (*Phoenix dactylifera* L.) varieties from morocco using RAPD markers. *Euphytica* 103: 75–82.
- Soliman SS, Ali BA and Ahmed MMM (2003) Genetic comparisons of Egyptian date palm cultivars (*Phoenix dactylifera* L.) by RAPD-PCR. *African Journal of Biotechnology* 2: 86–87.
- Trifi M, Rhouma A and Marrakchi M (2000) Phylogenetic relationships in Tunisian date-palm (*Phoenix dactylifera* L.) germplasm collection using DNA amplification fingerprinting. *Agronomie* 20: 665–671.
- United Nations Environment Programme (2007) Natural disasters and desertification. In: *Sudan – Post – Conflict Environmental Assessment, Nairobi, Kenya*. ISBN: 978-92-807-2702-9, pp. 56–69.
- Wright S (1969) Evolution and genetics of populations. *The Theory of Gene Frequencies*, vol. 2. Chicago: University of Chicago Press.
- Wrigley G (1995) Date palm (*Phoenix dactylifera* L.). In: Smartt J and Simmonds NW (eds) *The Evolution of Crop Plants*. London: Longman, pp. 399–403.
- Zaid A and de Wet PF (2002a) Botanical and systematic description of the date palm. In: Zaid A (ed.) *Date Palm Cultivation*. FAO Plant Production and Protection Paper No. 156. Rome: Food and Agriculture Organisation of the United Nations, pp. 1–28.
- Zaid A and de Wet PF (2002b) Origin, geographical distribution and nutritional values of date palm. In: Zaid A (ed.) *Date Palm Cultivation*. FAO Plant Production and Protection Paper no. 156, Rome: Food and Agriculture Organisation of the United Nations, pp. 29–44.
- Zaid A and de Wet PF (2002c) Date palm propagation. In: Zaid A (ed.) *Date Palm cultivation*. FAO Plant Production and Protection Paper No. 156. Rome: Food and Agriculture Organisation of the United Nations, pp. 73–89.
- Zaid A and de Wet PF (2002d) Pollination and bunch management. In: Zaid A (ed.) *Date Palm Cultivation*. FAO Plant Production and Protection Paper No. 156. Rome: Food and Agriculture Organisation of the United Nations, pp. 145–175.
- Zaid A, de Wet PF, Djebri M and Oihabi A (2002) Diseases and pests of date palm. In: Zaid A (ed.) *Date Palm Cultivation*. FAO Plant Production and Protection Paper No. 156. Rome: Food and Agriculture Organisation of the United Nations, pp. 227–242.
- Zehdi S, Trifi M, Billotte N, Marrakchi M and Pintaud JC (2004a) Genetic diversity of Tunisian date palms (*Phoenix dactylifera* L.) revealed by nuclear microsatellite polymorphism. *Hereditas* 141: 278–287.
- Zehdi S, Sakka H, Rhouma A, Ould Mohamed Salem A, Marrakchi M and Trifi M (2004b) Analysis of Tunisian date palm germplasm using simple sequence repeat primers. *African Journal of Biotechnology* 3: 215–219.
- Zhang X, Korpelainen H and Li C (2006) Microsatellite variation of *Quercus aquifolioides* populations at varying altitudes in the Wolong natural reserve of China. *Silva Fennica* 40: 407–415.