Middle Eastern and Russian Journal of Plant Science and Biotechnology ©2008 Global Science Books

First Comparative Genotypic Study on Khat (*Catha edulis* Forsk.) Genotypes from Yemen

M. A. H. Al-Thobhani[1] • B. N. Sathyanarayana[1] • Luke Simon[1,2*] • Suresh N. Sondur[1]

[1] Molecular Biology Laboratory, Plant Tissue Culture and Molecular Biology Unit, Department of Horticulture, University of Agriculture Sciences, Banglore-560065, India
[2] Current address: School of Medicine and Dentistry, Institute of Clinical Sciences, Queens University Belfast, Belfast, BT12 6BJ, United Kingdom

Corresponding author: * **simon_cac@yahoo.com**

ABSTRACT

In the present study, genetic relationships among 40 genotypes of khat (*Catha edulis* Forsk.) obtained from diverse locations of Yemen were analysed by RAPD (Randomly Amplified Polymorphic DNA) markers. PCR-amplifiable DNA was isolated using the CTAB method and 152 amplified fragments were obtained using 10 random primers with 78.3% polymorphism. Studies of gene diversity revealed a maximum diversity among the genotypes of the North geographical region, 'Abyadh' variant and genotypes within tree growth habits. Nei's genetic distance demonstrated the South and Centre geographical populations and 'Abyadh' and 'Azrad' variants were genetically close. The genetic dissimilarity matrix was calculated based on Euclidian Distances and revealed a maximum genetic distance (40%) between genotypes '2N' and '40N' and minimum genetic distance (11%) between the genotypes '18S' and '19S'. The Ward's method of cluster analysis grouped all the individuals on the dendrogram into three major clusters 'A', 'B' and 'C' at 46 linkage distances. The cluster analysis grouped the khat genotypes according to geographical region and growth habit but there was no correlation with pigmentation. RAPD proved to be a quick, simple and significant testing method to assess genetic diversity among khat genotypes.

Keywords: CTAB, genetic diversity, Khat (*Catha edulis* Forsk.), RAPD, POPGENE, qat, STATISTICA

INTRODUCTION

Catha edulis Forsk. (Celastraceae) commonly known as qat or khat, is an evergreen tree indigenous to East Africa, South Arabia and Yemen (Luqman and Danowski 1976). The leaves and tender twigs are chewed in certain countries of East Africa and the Arabian Peninsula, mainly Yemen (Dhaifalah and Santavy 2004; Warfa *et al*. 2007) for their central nervous system stimulating properties (Luqman and Danowski 1976). Khat chewing had become an integral part of that region's culture and represents an element of ethnic identification (Rousseau 2004). It has a deep-rooted social and cultural tradition (Kalix and Braenden 1985) and is usually chewed to enhance working capacity (Elmi 1983; Kalix and Braenden 1985; Kennedy 1987). Cathinone is believed to be the main active ingredient in fresh khat leaves (Hassan *et al*. 2007).

The habit of chewing khat has been common for many centuries. The earliest record with more factual bases showed the Arabic source Al-Biruni (973-1051 A.D.), indicated the ancient use of khat (Schopen 1978). Historically, khat has been used for medicinal purposes (Kennedy *et al*. 1987) as well as an aphrodisiac (Margetts 1967; Krikorian 1984), though it was also used for recreational purposes (Kennedy 1987). It is most valued for its stimulant effects (Baasher 1980). It is also used for avoiding sleepiness, euphoriatic effects, to boost efficiency of work and to increase sexual performance (Elmi 1983). Chewing is the most common mode of administration, it has been taken as a tea and occasionally smoked (Hodgkinson 1962; Kennedy 1987). In Ethiopia, processed leaves and roots are used to treat influenza, cough, gonorrhea, asthma and other chest problems, the roots are used for stomachache and an infusion made from them are taken to treat boils (Lemessa 2001). Recent clinical diagnostic studies reveal the evidence of mental illness (Warfa *et al*. 2007), oral squamous cell carcinoma (Fasanmade *et al*. 2007; Sawair *et al*. 2007), stroke (Vanwalleghem *et al*. 2006), acute myocardial infarction (Al-Motarreb *et al*. 2005), gastrointestinal and genito-urinary symptoms (Hassan *et al*. 2002), associated with khat chewing.

Local cultivars of khat were described based on geographical location, growth habit, and physical appearance, i.e. colour of the leaf, stem sizes, and potency of effect. It has been known for a long time that different kinds of khat have different degrees of pharmacological action. Yemeni farmers recognized four cultivars based on the colour of shoots and growing twigs as 'Abyadh' predominantly pale green in colour, 'Azraq' purplish, 'Aswad' crimson and 'Ahmar' an intermediate between 'Azraq' and 'Aswad' reddish. In Ethiopia two prominent cultivars have been described as *dimma* (red) or *ahde* (white) (Getahun and Krikorian 1973; Krikorian 1984). Barbier (1883) classified khat into three types namely 'Madness-causing', 'intoxicating-like spirit' and 'insomnia-causing' based on their effect while Getahun and Krikorian (1973) classified them into two major groups namely *kudda* and *kudda Kerti* mainly based on harvesting time. Forty kinds of khat based on geographical origin were recognized in Yemen by Ramadan *et al*. (1981). Hence, a detailed genetic study was undertaken for the first time among the cultivars in Yemen.

Diversity of khat germplasm in Yemen based on morphological characteristics has been performed (Krikorian 1984, 1985). However, due to the effects of environmental factors on these attributes, their use can be ambiguous. Therefore, markers independent of the environment are necessary for reliable identification and discrimination of cultivars. Different types of marker systems have been used for genetic analysis and genotyping, including morphological, cytological, biochemical and DNA markers. The value of markers depends on their heritability and on the level of polymorphism they can reveal (Porter and Smith 1982; Prakash *et al* 2002). DNA markers are independent from environmental interactions, unlimited in number and show high level of polymorphism. Therefore, they are considered as

Received: 20 January, 2008. Accepted: 24 March, 2008.

Original Research Paper

Table 1 Geographical details and growth characters of 40 *C. edulis* genotypes.

Code	Variant Name (Local Name)	Colour of young twigs	Growth habits	Location Name	Sub-location Name	Latitude (°N)	Longitude (°E)	Altitude (MSL)
1N	Ahmar	Purplish red	Tree	Hamden	Wadi Dhulla'a	15 27 05	44 08 21	2200
2N	Abyadh	Pale green	Tree	Hamden	Al-Qarya	15 25 23	44 07 18	2400
3N	Azraq (Thehabi)	Purplish	Tree	Hamden	Wadi Dhaher	15 26 45	44 05 30	2200
4N	Azraq	Purplish	Tree	Hamden	Wadi Dhaher	15 26 45	44 05 40	2200
5N	Aswad	Crimson	Tree	Hamden	Wadi Dhaher	15 26 45	44 05 20	2200
6N	Abyadh	Pale green	Tree	Hamden	Bait Athafeef	15 31 47	44 07 13	2200
7N	Azraq (Tuffahi)	Purplish	Tree	Bani Matter	Khusha'an	15 12 03	44 04 18	2560
8S	Abyadh (Saffi)	Pale green	Tree	Sabir	Mawa'dim	13 33 18	44 01 20	2200
9S	Abyadh (Saffi Habashi)	Pale green	Tree	Sabir	Mawa'dim	13 33 08	44 01 10	2200
10S	Azraq (Se-ni areedh Alawraq)	Purplish	Tree	Sabir	Mawa'dim	13 33 08	44 01 20	2200
11S	Azraq (Se-ni Raffa Alawraq)	Purplish	Tree	Sabir	Mawa'dim	13 33 00	44 01 20	2200
12S	Ahmar (Ashrag Shebh Aswad)	Purplish red	Tree	Sabir	Mawa'dim	13 32 45	44 01 12	2300
13S	Azraq (Ashrag Hindi)	Purplish	Tree	Sabir	Mawa'dim	13 32 45	44 01 02	2300
14S	Ahmar	Purplish red	Tree	Sabir	Mawa'dim	13 32 45	44 01 25	2400
15S	Aswad (Ashrag Hindi)	Crimson	Tree	Sabir	Mawa'dim	13 32 45	44 01 35	2400
16S	Azraq	Purplish	Tree	Qadas	Mattaran	13 16 29	44 07 53	2000
17S	Ahmar	Purplish red	Tree	Qadas	`Halaqan	13 17 28	44 09 00	1800
18S	Azraq (Aswad Maqttart)	Purplish	Tree	Al-Maqattera	Adahamesha	13 12 18	44 11 16	1900
19S	Azraq (Sag)	Purplish	Tree	Al-Maqattera	Adahamesha	13 11 53	44 11 16	1840
20S	Azraq (Mushattitah)	Purplish	Tree	Al-Maqattera	Al-how-ishah	13 11 34	44 11 01	1800
21S	Ahmar (Rome)	Purplish red	Shrub	Mawyah	Wadi Alesaab	13 34 05	44 21 17	1200
22C	Abyadh (Rome)	Pale green	Shrub	Assyani	Wadi Nakhlan	13 48 23	44 10 25	1720
23C	Azraq (Rome)	Purplish	Shrub	Assyani	Wadi Nakhian	13 48 06	44 10 17	1720
24C	Abyadh	Pale green	Tree	Thee Suffai	Al-Khuryff	13 49 28	44 01 57	2000
25C	Aswad	Crimson	Tree	Thee Suffai	Al-Khuryff	13 49 28	44 01 57	2000
26C	Aswad (Rome)	Crimson	Shrub	Al- `Hada`	Al-Magel	14 45 41	44 27 47	2270
27C	Ahmar (Rome)	Purplish red	Shrub	Al- `Hada`	Wadi Lubb	14 47 02	44 29 27	2280
28C	Abyadh (Kahrabi)	Pale green	Shrub	`Anns	Thee Guzab	14 26 02	44 26 33	2550
29C	Azraq (Rome)	Purplish	Shrub	`Anns	Wadi Al-Har	14 27 05	44 24 01	2400
30N	Aswad (Hrami)	Crimson	Shrub	Al-Mahabisha	Al-Gediah, Mhabishah	15 57 03	43 26 59	1200
31N	Azraq (Shami)	Purplish	Shrub	Al-Mahabisha	Al-Gediah, Mhabishah	15 57 21	43 26 03	1200
32N	Abyadh (Raffa Alawraq)	Pale green	Tree	Al Ahjer	Azeelah	15 26 36	43 55 03	2080
33C	Ahmar (Qat Hiss)	Purplish red	Tree	Yaff`a	Wadi H`ass	13 46 54	45 26 46	1300
34C	Azraq (Yaff`ai)	Purplish	Tree	Yaff`a	Wadi Al-Qarnine	13 44 29	45 26 23	1600
35C	Abyadh (Dhal`ai)	Pale green	Shrub	Adhal`ai	Wadi Nasham	13 42 11	44 45 33	1500
36C	Ahmar (Dhal`ai)	Purplish red	Shrub	Adhal`ai	Wadi Nasham	13 41 54	44 45 25	1500
37C	Azraq (Moraisi)	Purplish	Shrub	Adhal`ai	Muraiss	14 01 05	44 40 39	2000
38N	Azraq (Sawtti)	Purplish	Tree	`Amran	Ah-noom	15 55 40	43 47 45	1960
39N	Ahmar (Sawtti)	Purplish red	Tree	`Amran	Ah-noom	15 55 40	43 47 29	1900
40N	Abyadh (Swatti)	Pale green	Tree	`Amran	Ah-noom	16 13 42	43 38 17	1300

valuable tools for determining genetic relationships/diversity. The most frequently used DNA markers include restriction fragment length polymorphisms (Sambrook *et al*. 1989), random amplified polymorphic DNAs (Williams *et al*. 1990) or RAPDs, simple sequence repeats (Gupta *et al*. 1996), amplified fragment length polymorphism (Vos *et al*. 1995) and inter simple sequence repeats (Provost and Wilkinson 1999). These molecular markers are based on different principles and are obtained by using procedures of varying complexity and generate different amounts of polymorphic data. RAPD markers gained importance due to their simplicity, efficiency, the relative ease to perform the assay and non-requirement of DNA sequence information (Khanuja *et al*. 1998). They have been used in studies of genetic diversity (Chalmers *et al*. 1994), phylogeny and systematics (Sun *et al*. 1998), genetic linkage mapping (Cheung *et al*. 1997) and gene tagging (Tiwari *et al*. 1998). In the present study, genetic diversity among 40 khat genotypes classified based on their physical appearance and geographical location were analysed by RAPD markers.

Fig. 1 Variants. Four types of variants among *Catha edulis* based on their pigmentation.

MATERIALS AND METHODS

Plant materials

Forty genotypes of khat were collected from different geographical locations of Yemen and were used for the analysis (**Table 1**). The genotypes were classified into three categories based on their main geographical region (13 - North, 14 - Central and 13 - South), pigmentation (10 - Abyadh, 9 - Ahmar, 16 - Azraq and 5 - Aswad) and growth habits (28 tree and 12 shrub types). The four variants based on the pigmentation of leaves and twigs are shown in **Fig. 1**. Approximately 50 g of recently matured leaves (10-12 days old) were collected, washed using distilled water, wiped with 70% (v/v) ethanol, then dried in oven at 30-35°C for 20 h and powered

by using a 'Remi' mixer for 45 to 60 sec, prior to storage at room temperature in sealed plastic bags.

DNA extraction and purification

All the reagents and chemicals were obtained form Bangalore Genei, Bangalore, India of molecular biology grade. DNA was extracted from the dried leaf powder of khat by the cetyl trimethyl ammonium bromide (CTAB) method according to a modified protocol of Porebski *et al*. (1997). Half a gram of leaf powder was mixed with 20 ml extraction buffer, preheated to 65°C, which contained 100 mM Tris-Base, pH 8.0, 20 mM EDTA, pH 8.0, 1.4 M NaCl, 3% (w/v) CTAB, 2% polyvinyl pyrrolidone and 1% β-mercaptoethanol, then incubated at 65°C for 30 min with gentle shaking. The mixture was cooled to room temperature, 10 ml cold 24:1 (v/v) chloroform: isoamylalcohol was added, and the contents were mixed well. After centrifugation at 6000 × *g* for 20 min at 4°C, the supernatant was transferred to a fresh tube and the chloroform: isoamylalcohol step was repeated until a clear supernatant was obtained. 5 M NaCl was added to the supernatant (0.5 v/v) and mixed gently followed by the addition of 0.8 vol of cold isopropanol to precipitate the DNA. The mixture was incubated overnight at 4°C, and then centrifuged at 8,000 × *g* for 20 min. The resulting pellet was washed with 70% (v/v) ethanol, air-dried, and dissolved in TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0). Two µg RNase (bovine pancreatic ribonuclease, Bangalore Genei, Bangalore, India) was added to each sample which was incubated for 3 h at 37°C, mixed with an equal volume of phenol and centrifuged at 6000 × *g* for 20 min at room temperature. This step was followed by a wash with an equal volume (1:1 (v/v)) of phenol: chloroform then with chloroform alone. DNA was precipitated overnight at 4°C with 0.5 vol of 5 M NaCl and 1 vol cold isopropanol, and the resulting pellet obtained after centrifugation was dissolved in TE buffer, analysed on an agarose gel and quantified using a spectrophotometer (NanoDrop Technologies, Wilmington, USA).

PCR amplification

PCR amplification followed the protocol of Williams *et al*. (1990) with minor modifications. Of the 71 primers screened using the genotype '40N', 10 showing clear and distinguishable bands were selected for RAPD-PCR analysis (**Table 2**). Reproducibility of the primers was tested by repeating the PCR amplification twice under similar conditions. PCR reactions were carried out in a volume of 25 µl containing 50-75 ng template DNA, 250 µM each dNTP, 1.5 mM $MgCl_2$, 1 unit *Taq* DNA polymerase (Bangalore GENEI, Bangalore, India), 5 pmol primer (G, U, V, W and X series in addition to primer numbers 1, 2, 3, 4 and 5, Metabion Gmb H Technologies Inc, Germany) in PCR buffer (50 mM KCl, 10 mM Tris-HCl pH 8.3, 0.001% (v/v) gelatin, 0.05% (v/v) NP40 and 0.05% (v/v) Triton X-100). Amplifications were performed in a MJ Research PTC-100 Thermocycler (Bio-Rad Laboratories, Bangalore, India), programmed for an initial denaturation at 93°C for 3 min, followed by 50 cycles of denaturation at 93°C for 1 min, annealing at 35°C for 30 sec, primer extension at 72°C for 2 min, and a final extension at 72°C for 8 min. PCR products were resolved in a 1.5% (w/v) agarose gel, visualised and documented using an Alpha Digidoc system (Alpha Innotech, San Leandro, CA, US). The sizes of the fragments were estimated using 500 bp standard DNA markers (Sigma Aldrich Chemicals, Bangalore, India), co-electrophoresised with the PCR products.

Statistical analysis

Amplified fragments from each primer were manually scored for their presence (1) or absence (0) and a matrix of the different RAPD phenotypes of all ten primers were assembled for statistical analysis. A genetic dissimilarity matrix (**Table 3**) was developed using Euclidean Distances, which estimates all pair-wise differences in the amplification products (Sokal and Sneath 1973) and a cluster analysis was based on Ward's method using a minimum variance algorithm (Ward 1963) by STATISTICA programme. Based on the presence and absence data of the amplified products, genetic diversity parameters such as gene diversity, Shannon's information index and percent polymorphism, mean similarity index, frequency distribution of similarity index and Nei's genetic distance were analyses using the population genetics software (POPGENE version 1.32; Yeh and Boyle 1997).

RESULTS AND DISCUSSION

The work presented here is the first to describe genetic variability and molecular differentiation among cultivated populations of *Catha edulis* from Yemen. Genetic diversity among crop plant species is important for efficient utilization of plant genetic resources since geographical isolation of a population may cause its genome to drift away from other populations of the same species (Biron *et al*. 2002). Hence, authentic identification of taxa is necessary both for breeders to ensure protection of intellectual property right and also for propagators and consumers. The traditional method of identifying species by morphological characters is now been replaced by DNA profiling which are more reliable due to several limitations of their morphological data (Nayak *et al*. 2003). Many papers have reported the effectiveness of RAPD markers in discriminating between species, subspecies and populations in a wide range of organisms, which were confirmed by results of the present study, since considerable levels of genetic divergence were detected between the *C. edulis* genotypes analysed.

Recently matured leaves preferably, 10-12 days old were used to extract DNA, as mature leaves were rich in polyphenols that hindered the extraction of PCR quality DNA. The pre-treatment of the leaves removed dust particles and external microbial contaminations. The CTAB method for DNA extraction was found optimal to release the nucleic acid from the cell and to remove RNA and proteinaceous contamination rendering the DNA suitable for PCR amplifications. PCR amplification was followed by a standard protocol (Williams *et al*. 1990) with minor modifications, which produced good amplifications with 50-75 ng of template DNA. The amplifications using 1 unit of *Taq* DNA polymerase and 1.5 mM $MgCl_2$ produced intense and clear

Table 2 RAPD-PCR Primers. The sequence and level of polymorphism of primers.

Primer name	Sequence (5'-3')	Total № of bands	№ of polymorphic shared bands	№ of polymorphic unique bands	№ of monomorphic bands
Primer 02	TGCCGAGCTG	18	10	1	7
KIT V 02	CATCGCCGCA	17	15	2	0
KIT V 08	AGGCAGAGCA	10	05	2	3
KIT V 09	AGCAGCGCAA	20	11	2	7
KIT V 11	GAGCGATGGG	18	18	0	0
KIT V 18	GTGGAGTCAG	12	10	0	2
KIT V 19	TGAGGGTGGC	16	10	1	5
KIT W 06	AGGCCCGATG	14	09	1	4
KIT W 13	CACAGCGACA	15	13	0	2
KIT W 16	CAGCCTACCA	12	08	1	3
Total		152	109	10	33
Percentage (%)			71.7	6.6	21.7

Table 3 Genetic dissimilarity matrix of 40 khat genotypes based on polymorphism of RAPD markers.

1N	0																																							
2N	24	0																																						
3N	19	23	0																																					
4N	21	20	19	0																																				
5N	29	29	25	25	0																																			
6N	23	25	17	23	29	0																																		
7N	21	24	17	17	20	20	0																																	
8S	21	14	19	21	19	25	21	0																																
9S	17	23	14	21	24	17	16	18	0																															
10S	23	28	21	21	21	24	12	22	19	0																														
11S	24	21	15	12	24	21	12	18	17	16	0																													
12S	25	24	22	24	27	21	27	21	21	29	25	0																												
13S	22	18	16	21	25	23	22	19	15	18	18	19	0																											
14S	24	24	18	25	27	24	24	21	19	24	23	23	17	0																										
15S	23	25	21	27	29	21	20	25	16	20	23	23	18	20	0																									
16S	24	25	22	16	25	28	16	23	24	21	16	32	25	25	23	0																								
17S	27	24	19	21	27	23	24	25	20	25	23	21	17	16	23	27	0																							
18S	28	27	23	29	31	23	23	26	21	24	25	29	26	24	29	21	23	0																						
19S	25	25	23	23	27	25	21	28	23	18	25	33	25	21	21	18	29	11	0																					
20S	29	34	25	26	29	26	21	29	30	23	22	30	31	27	26	17	33	19	20	0																				
21S	23	27	27	26	37	22	27	29	27	27	27	26	23	23	23	30	23	27	32	31	0																			
22C	22	29	17	27	29	19	21	20	19	22	21	25	23	19	19	25	25	22	21	24	27	0																		
23C	27	28	27	28	36	24	29	30	31	27	27	28	27	29	24	27	29	24	29	26	22	26	0																	
24C	27	26	24	21	27	23	23	23	27	25	17	30	29	33	27	23	30	30	31	23	36	25	29	0																
25C	33	37	31	31	36	28	27	35	31	31	29	33	31	32	31	31	35	33	33	30	30	31	28	33	0															
26C	28	36	29	28	29	27	25	30	28	28	23	27	27	32	27	27	32	31	29	29	31	26	25	29	25	0														
27C	28	39	27	28	29	28	28	34	32	27	31	29	30	32	27	31	28	33	31	34	29	29	29	33	28	25	0													
28C	34	39	35	33	31	26	26	35	31	29	29	37	35	34	29	33	31	31	33	33	29	29	27	33	25	26	21	0												
29C	29	31	26	27	32	21	28	34	24	25	25	31	26	25	21	28	28	27	26	31	26	29	24	31	33	28	28	29	0											
30N	31	37	25	31	37	22	29	35	27	33	30	29	28	29	26	33	26	26	28	31	29	27	25	33	31	23	25	28	26	0										
31N	30	37	27	31	34	26	33	33	27	31	34	29	31	31	25	30	26	25	28	31	25	25	27	36	30	26	26	27	23	23	0									
32N	31	31	29	26	34	27	25	27	30	29	25	34	29	33	30	26	37	37	35	28	29	29	27	27	25	23	30	27	29	31	31	0								
33C	28	33	23	31	31	23	23	29	24	27	20	31	29	25	29	29	31	27	27	23	30	21	24	25	32	24	28	27	31	22	29	26	0							
34C	28	33	23	24	24	29	19	30	27	24	17	32	30	29	28	25	29	32	31	22	29	27	32	25	31	28	25	30	31	30	35	26	16	0						
35C	35	31	29	27	31	29	29	31	31	31	26	34	33	29	29	30	29	33	32	33	32	25	30	29	38	30	27	27	30	28	29	31	22	25	0					
36C	33	35	29	30	29	27	31	31	31	31	25	29	33	34	29	33	35	33	32	31	33	27	27	23	29	23	27	29	25	32	29	24	25	23	25	0				
37C	25	29	21	17	23	25	21	27	25	25	17	26	25	29	29	19	25	30	24	25	32	27	29	21	29	26	29	35	27	33	32	32	25	23	29	27	0			
38N	29	29	27	25	33	25	25	29	30	29	23	29	32	30	27	26	33	30	28	25	28	23	23	23	29	27	30	28	31	28	31	31	19	25	19	27	25	0		
39N	27	27	22	27	28	23	23	25	25	27	21	31	29	28	27	29	28	27	33	26	29	19	27	23	31	29	32	27	31	33	29	27	19	23	22	27	27	18	0	
40N	33	40	31	29	31	30	26	37	30	30	29	37	35	33	34	29	33	35	31	29	32	33	30	32	27	25	26	29	29	29	28	27	23	25	29	25	25	31	30	0

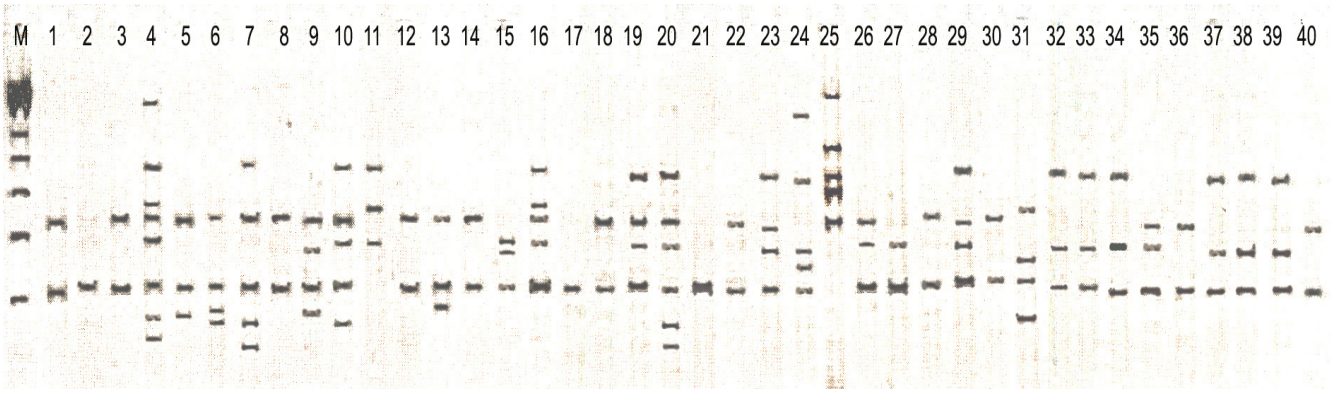

Fig. 2 Gel profile. Amplification profile by RAPD-PCR primers KIT V2 among *Catha edulis* genotypes. Lanes 1-46, contain the amplification profile obtained using the genotypes '1N', '2N', '3N', '4N', '5N', '6N', '7N', '8S', '9S', '10S', '11S', '12S', '13S', '14S', '15S', '16S', '17S', '18S', '19S', '20S', '21S', '22C', '23C', '24C', '25C', '26C', '27C', '28C', '29C', '30N', '31N', '32N', '33C', '34C', '35C', '36C', '37C', '38N', '39N' and '40N'. Lane M, standard 100 bp DNA markers.

banding patterns. A primary screening of 71 RAPD primers resulted in the selection of 10 primers that produced clear fragment patterns. Screening is essential to save time and cost, and to reject primers not informative for the analysis (Prakash *et al*. 2002).

About 152 readable and reproducible RAPD markers were produced using the 10 primers selected. The number of bands obtained per primer (**Table 2**) varied from 10 to 20 with an average of 15.2 bands per primer and the size ranged from 150 bp to 4.2 kbp. Of the 152 bands, 109 (71.7%) were polymorphic and shared between at least two individuals, 33 (21.7%) was monomorphic and common to all the individuals. Only ten (6.6%) were polymorphic and unique. The representative polymorphic gel profiles of primer V2 is shown in **Fig. 2**. A pair-wise genetic dissimilarity matrix (**Table 3**) was calculated using Euclidian Distance (Sokal and Sneath 1973). The highest genetic dissimilarity of 40% was between genotypes '2N' and '40N', while the least genetic dissimilarity (11%) was between the genotypes '18S' and '19S', belonging to the same regions of collections.

Among the main geographical regions, the North region showed maximum gene diversity of (0.186 ± 0.175) while the lowest diversity was revealed in the Southern region (0.131±0.171). This might probably be due to the greater

Table 4 Genetic diversity parameters based on RAPD data of *C. edulis* grouped under three categories.

Category	Sample size	Gene diversity		Shannon's diversity index		Percent polymorphism (%)
		Mean	SD±	Mean	SD±	
A. Main geographical regions						
North	13	0.186	0.175	0.291	0.252	64.29
Centre	13	0.174	0.174	0.274	0.251	62.14
South	14	0.131	0.171	0.205	0.251	45.71
B. Variants						
Abyadh	10	0.202	0.172	0.315	0.250	66.43
Azraq	16	0.164	0.178	0.257	0.257	57.86
Ahmar	9	0.126	0.179	0.192	0.263	37.14
Aswad	5	0.162	0.195	0.242	0.285	42.86
C. Growth habits						
Tree	28	0.173	0.167	0.279	0.236	73.57
Shrub	12	0.161	0.175	0.251	0.256	54.29

Table 5 Differences in the mean gene diversity among categories of *C. edulis*

	Main geographical regions		
	North	**Centre**	**South**
North	0	0.567 (0.571)	2.651 (p<0.008)
Centre		0	2.086 (p<0.038)
South			0
	Variants		
	Abyadh	**Azraq**	**Ahmar**	**Aswad**
Abyadh	0	1.801 (0.073)	3.609 (p<0.000)	1.803 (0.072)
Azraq		0	1.778 (0.078)	0.085 (0.932)
Ahmar			0	-1.612 (0.108)
Aswad				0

Table 6 Differences in mean of similarity indices among categories of *C. edulis*

	Main geographical regions		
	North	**Centre**	**South**
North	0	1.564(0.120)	8.388(p<0.00)
Centre		0	7.855(p<0.00)
South			0
	Variant s		
	Abyadh	**Azraq**	**Ahmar**	**Aswad**
Abyadh	0	-6.073(p<0.00)	-6.7(p<0.00)	-0.941(0.351)
Azraq		0	-5.529(p<0.00)	2.441(p<0.016)
Ahmar			0	5.389(p<0.000)
Aswad				0

expansion of khat cultivation areas in the North and Centre (Muharram *et al*. 2002) leading to the establishment of new plantations with collections from different parts of the country. Among the four variants recorded the mean gene diversity of 'Abyadh' (0.202 ± 0.172) was significantly higher than other variants. The genotypes of tree habitat (0.173 ± 0.167) revealed the maximum gene diversity, more than the shrub types. Correspondingly, the Shannon's diversity index and the percent polymorphism also confirmed a similar outcome (**Table 4**). The comparison of gene diversity confirmed previous results of maximum difference (2.651) between the Northern and Southern regions. Among the variants the lowest difference (-1.612) was noticed between 'Ahmar' and 'Aswad', while a maximum difference (3.609) between 'Abyadh' and 'Ahmar' (**Table 5**).

The mean similarity indices among the categories revealed that North populations had the least (0.799 ± 0.058), which was significantly different from the South population (0.859 ± 0.035) indicating the individuals of North populations were diverse than those of South region. The higher similarity of South populations may be due to less geographic distance separating individuals (**Table 1**) and *C. edulis* was first domesticated in southern parts of Yemen (Kennedy 1987). Among the variants the highest mean similarity index was computed for 'Ahmar' (0.860 ± 0.048), while 'Abyadh' recorded the lowest index (0.775 ± 0.071) which was on par with the 'Aswad' variant (0.797 ± 0.042). The histogram representing the mean similarity indices are shown in **Fig. 3**. However, tree genotypes had lower mean similarity index as compared to the shrub types. The differences among the mean similarity indices showed a maximum between North and South regions and 'Ahmar' and 'Aswad' variants (**Table 6**).

The frequency distribution of similarity index was analysed using Kolmogorov-Smirnov (KS) test revealed the frequencies between the Centre and each of North and South were significant. Among the variants 'Abyadh' and 'Azraq' showed wider distributions as compared to 'Aswad' and 'Ahmar', indicating the individuals of later variants were less diverse among themselves (**Table 7**). It was also observed that the trees and shrubs were not considerably different. The histogram representing the frequencies of genotypes based on main geographical regions and variants are shown in **Fig. 4**. The Nei's genetic distance between South and Centre populations (0.0317) were higher whereas, the distances between North and South populations was least (0.0172), indicating that the later populations were genetically close (**Table 8**). Comparison among the variants revealed that 'Abyadh' and 'Azrad' were genetically close with the least distance (0.0174) in contrast to 'Abyadh' and 'Aswad' (0.0452).

In the dendrogram (**Fig. 5**), all 40 genotypes were clearly clustered into three major clusters 'A', 'B' and 'C' at 46 linkage distance. Cluster 'A' consisted of one genotype '40N' identified from the Amran location of Northern region. This genotype was characterized as a tree type commonly called 'Abyadh' (Swatti) with pale-green twigs. Two polymorphic and unique bands specific to the genotype '40N' sized 3,850 and 3,200 bp were produced by primers V09 and W16, respectively. Cluster 'B' consisted of 15

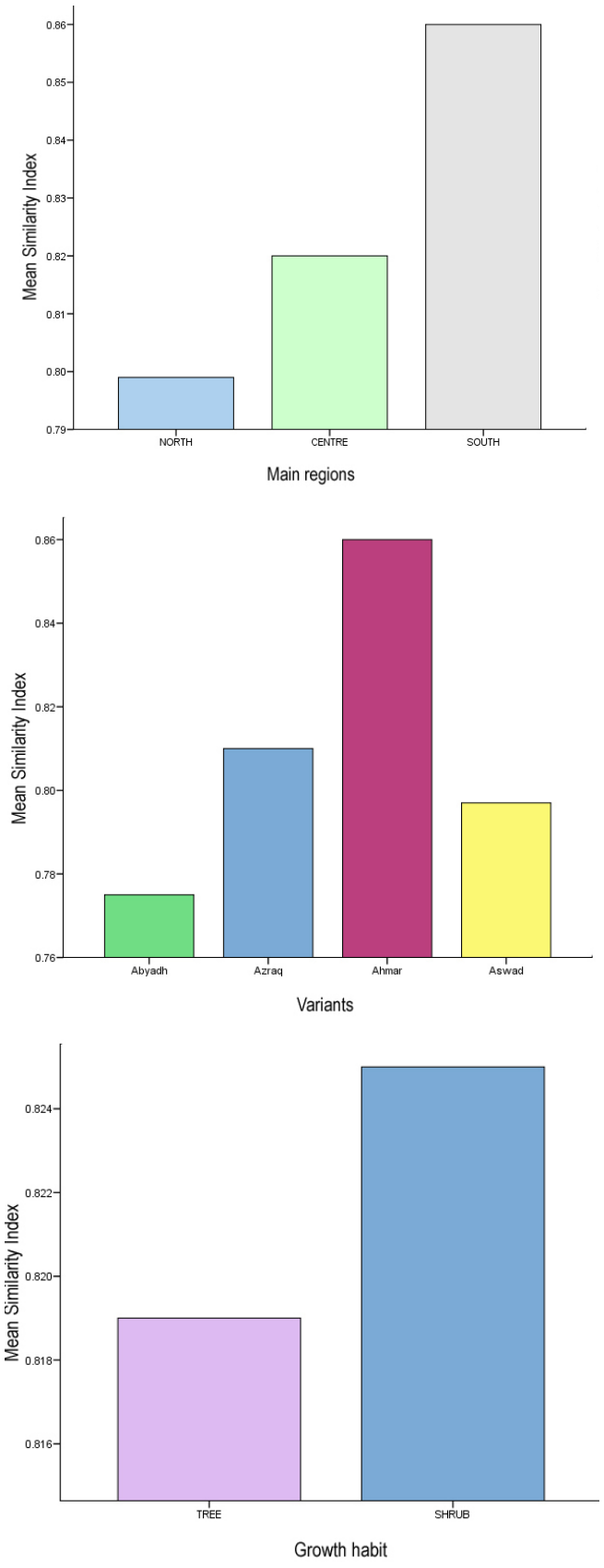

Fig. 3 Histogram. Mean Similarity Indices among 40 genotypes of *C. edulis* based on three categories.

genotypes clustered at 44 distances and further classified into two prominent sub-clusters 'B_1' and 'B_2'. Sub-cluster 'B_1', comprising of 11 genotypes, was divided into two groups (I and II) at 42 linkage distance. Group I consisted of eight genotypes belonging to North and Central geographical regions. Genotypes '38N' and '39N', and '33C' and '34C' were closely clustered among the group at 28 and 25 linkage distances, respectively. Genotypes '30N' and '31N' were closely linked and clustered with genotype '29C' at 39 linkage distances in group II of 'B_1'. A polymorphic band specific to sub-cluster 'B_1' was obtained using primer V18. Sub-cluster 'B_2', consisting of four genotypes ('25C', '26C', '27C', '28C') from the Central region, clustered at 39 linkage distances. The genotypes grouped in group II of 'B_1' and sub-cluster 'B_2' predominantly characterized a shrub habitat. The RAPD clustering of cluster 'B' revealed that all the genotypes of the cluster belong to the North and Central geographical regions of Yemen.

Cluster 'C' consists of 24 khat genotypes segregated into two sub-clusters 'C_1' and 'C_2' at 42 linkage distance. Sub-cluster 'C_1' consists of a single genotype '24C' obtained from Thee Suffai in the Central region of Yemen. The genotype was classified under variant 'Abyadh' with pale-green twigs. Sub-cluster 'C_2' was segregated into three groups (I, II and III) linked at 40 distances. Group I had two genotypes '21S' and '23C' associated at 37 distance. Both genotypes showed a shrub habitat but were collected from different places and identified as variants 'Ahmar' and 'Azraq', respectively. Genotypes '21S' and '23C' were characterized by a RAPD marker of 1,300 bp in size produced by primer W06. Group II of 'C_2' consists of one genotype '22C' collected from Assyani in the Central region. The genotype was identified as a shrub of the variant 'Abyadh'. Group III consisted of 21 genotypes clustered at 35 linkage distance. Genotypes '18S' and '19S', '13S' and '14S', and '10S' and '11S' were closely related members of the group. The RAPD clustering of cluster 'C' predominantly showed that the genotypes belong to North and South geographical regions of Yemen.

It appears from the results that the four variants of *C. edulis* are not strongly differentiated from each other genetically (**Fig. 5**). This could be related to the collection of a small number 'Abyadh' (10), 'Aswad' (5), 'Azraq' (16) and 'Ahmar' (9). However, a small sample size can lead to a significant bias in the population genetics (Fischer *et al.* 2000) and a sample size of 10-15 individuals are adequate for largely unbiased results (Isabel *et al.* 1999). Low to moderate polymorphism (11-40%) was obtained among the

Table 7 Kolmogorov-Smirnov test for frequency distribution of similarity values among categories of *C. edulis*.

Combinations	D max	Table value (5%)	P value
North vs Centre	0.167	0.218	0.115
North vs South	0.383	0.210	0.000
Centre vs South	-0.427	0.210	0.000
Abyadh vs Azraq	0.347	0.238	0.000
Abyadh vs Ahmar	0.600	0.304	0.000
Abyadh vs Aswad	0.356	0.475	0.126
Azraq vs Ahmar	0.292	0.258	0.009
Azraq vs Aswad	0.525	0.448	0.006
Ahmar vs Aswad	0.800	0.486	0.000

Table 8 Nei's unbiased measures of genetic distance among the categories of *C. edulis*.

	Main geographical regions			
	North	**Centre**	**South**	
North	0			
Centre	0.0216	0		
South	0.0172	0.0317	0	
	Variants			
	Abyadh	**Azraq**	**Ahmar**	**Aswad**
Abyadh	0			
Azraq	0.0174	0		
Ahmar	0.0181	0.0203	0	
Aswad	0.0452	0.0420	0.0334	0
	Growth form			
	Tree		**Shrub**	
1 Tree	0			
2 Shrub	0.0241		0	

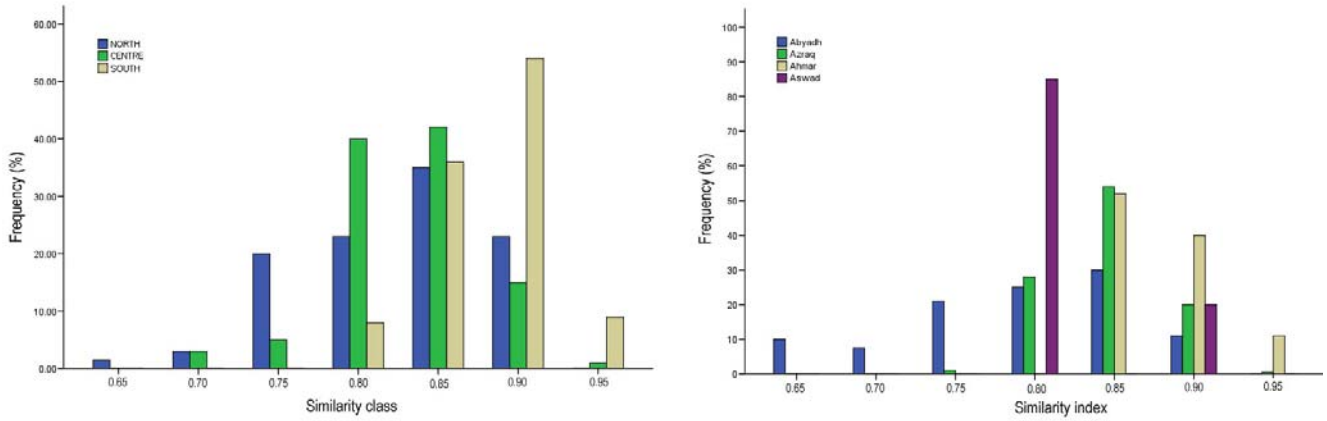

Fig. 4 Histogram. Frequency distributions among 40 genotypes of *C. edulis* based on three categories.

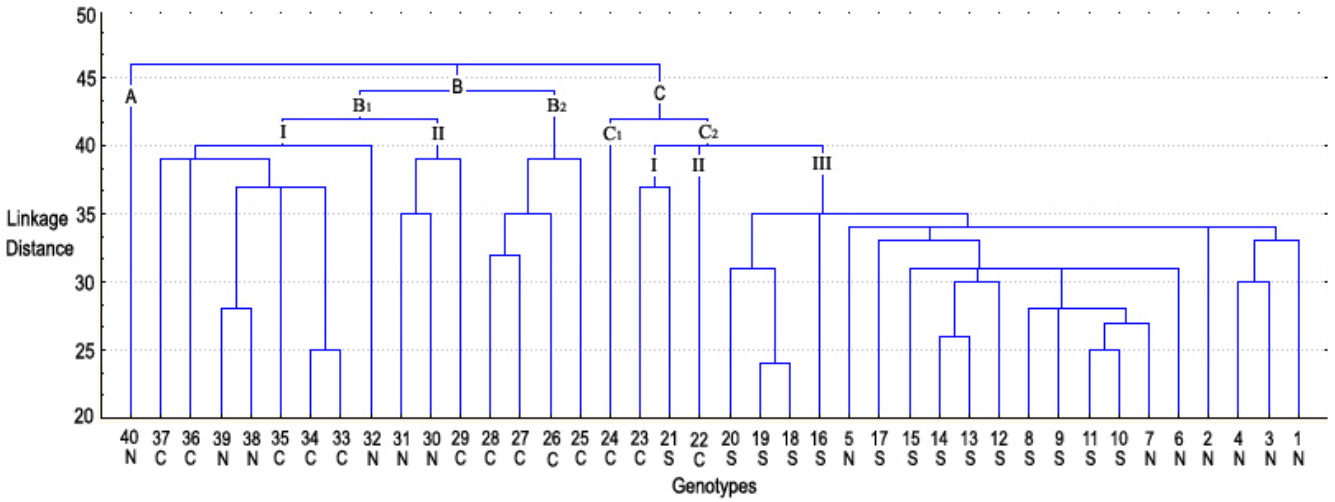

Fig. 5 Dendrogram. RAPD-marker-based genetic relationships among 40 *Catha edulis* genotypes.

genotypes which could be the result of vegetative propagation of khat by cuttings and suckers, since seeds of *C. edulis* show a poor germination rate (Lemessa 2001). Although, somatic mutations could occur in populations where asexual propagation is common (Lamote *et al*. 2002) recent studies reveal that genetic diversity will decrease at a constant rate with an increase in the rate of asexual reproduction (Balloux *et al*. 2003; Bengtsson *et al*. 2003).

The results indicate that populations from different geographical regions and growth habits in Yemen showed a clear genetic differentiation. The variants classified based on pigmentation are independent from geographical regions indicating their random distribution across the study area. Similar results were obtained by Zhayi *et al.* (2003) among *Loropetalum chinense* samples resulting in no correlation between RAPDs and the colour of leaves and flowers. However, analysis in *Kaempferia parviflora* by Pojanagaroon *et al.* (2004) showed cultivars segregated into clusters based on leaf colour. Further studies on chemical analysis of the chemotypes and their correlation with morphological and molecular markers would help to understand the variations among khat. Understanding the spatial organization of genetic diversity within plant populations is of critical importance for the development of strategies designed to preserve genetic variation (Brown and Briggs 1991; Hamrick *et al.* 1991). Hence, the present study could help in the conservation of genetically distinct *C. edulis* genotypes collected from Yemen.

REFERENCES

Al-Motarreb A, Briancin S, Al-Jaber N, Al-Adhi B, Al-Jailani F, Salek MS, Broadley KJ (2005) Khat chewing is a risk factor for acute myocardial infarction: a case-control study. *British Journal of Clinical Pharmacology* **59 (5)**, 574-581

Baasher TA (1980) The use of Qat: a stimulant with regional distribution. In: *Drug Problems in the Sociocultural Context – A Basis for Policies and Programme Planning*, World Health Organization, Geneva, pp 86-93

Balloux F, Lehmann L, de Meeûs T (2003) The population genetics of clonal and partially clonal diploids. *Genetics* **164**, 1635-1644

Barbier DMAC (1883) Notice sur l'Arabie méridionale d'aprés un document turc. In: *Mélanges Orientaux* (2ième série), Textes et Traductions publiés par Les Professur de l'école special des langues orientales vivantes, Ernest Leroux Editeur, Paris, pp 87-123

Bengtsson BO (2003) Genetic variations in organisms with sexual and asexual reproduction. *Journal of Evolutionary Biology* **16**, 189-199

Biron DG, Landry BS, Nenon JP, Coderre1 D, Boivin G (2002) Geographical origin of an introduced pest species, *Delia radicum* (Diptera: Anthomyiidae), determined by RAPD analysis and egg micromorphology. *Bulletin of Entomological Research* **90**, 23-32

Brown AH, Briggs JD (1991) Sampling strategies for genetic variation in *ex situ* collections of endangered plant species. In: Falk DA, Holsinger KF (Eds) *Genetics and Conservation of Rare Plants*, Oxford University Press, New York, pp 99-119

Chalmers KJ, Newton AC, Waugh R, Wilson J, Powell W (1994) Evaluation of the extent of genetic diversity in mahoganies (Meliaceae) using RAPD markers. *Theoretical and Applied Genetics* **89**, 504-508

Cheung WY, Champagne G, Hubert N, Landry BS (1997) Comparison of the genetic maps of *Brassica napus* and *Brassica oleracea*. *Theoretical and Applied Genetics* **94**, 569-582

Dhaifalah I, Šantavy J (2004) Khat habit and its health effect. A natural amphetamine. *Biomedical papers of the Medical Faculty of the University Palacky, Olomouc, Czechoslovakia* **148 (1)**, 11-15

Elmi AS (1983) Khat: consumption and problems in Somalia. In: *International Conference on Qat - The Health and Socioeconomic Aspects of Qat Use*, International Council on Alcohol and Addictions, Madagascar, pp 78-86

Fasanmade A, Kwok E, Newman L (2007) Oral squamous cell carcinoma associated with khat chewing. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontics* **104 (1)**, 53-55

Fischer M, Husi R, Prati D, Peintinger M, van Kleunen M, Schmid B (2000) RAPD variations among and within small and large populations of the

rare clonal plant *Ranunculus reptans* (Rananculaceae). *American Journal of Botany* **87 (8)**, 1128-1137

Forskal P (1775) Flora Aegyptico-Arabica. *Sive Descryptiones Plantarum, quas per Aegyptum Inferiorem et Arabiam Felicem Detexit*, Ex Officina Molleri, Aulae Typographi, Copenhagen, 63 pp

Getahun A, Krikorian AD (1973) Chat: Coffee's rival from Harar, Ethiopia. I. Botany, cultivation and use. *Economic Botany* **27 (4)**, 353-377

Gupta PK, Balyan HS, Sharma PC, Ramesh B (1996). Microsatellites in Plants: A new class of molecular markers. *Current Science* **70 (1)**, 45-54

Hamrick JL, Godt MJ, Murawski DA, Loveless MD (1991) Correlations between species traits and allozyme diversity: Implications for conservation biology. In: Falk DA, Holsinger KF (Eds) *Genetics and Conservation of Rare Plants*, Oxford University Press, New York, pp 75-86

Hassan NAGM, Gunaid AA, El Khally FMY, Murray-Lyon IM (2002) The subjective effects of chewing Qat leaves in human volunteers. *Annals of Saudi Medicine* **22 (1-2)**, 34-37

Hassan NAGM, Gunaid AA, Murrar-Lyon IM (2007) Khat (*Catha edulis*): health aspects of khat chewing. *Eastern Mediterranean Health Journal* **13 (3)**, 706-718

Hodgkinson R (1962) A study of the clinical effects of *Catha edulis* (Qat, miraa) in Kenya. *The Medical Journal of Australia* **49**, 884-886

Isabel N, Beaulieu J, Theriault P, Bousquet J (1999) Direct evidence for biased gene diversity estimates from dominant random amplified polymorphic DNA (RAPD) fingerprints. *Molecular Ecology* **8 (3)**, 477-483

Kalix P, Braenden O (1985) Pharmacological aspects of the chewing of Qat leaves. *Pharmacological Reviews* **37**, 149-164

Kennedy JG (1987) The flower of paradise. In: Reidel D (Ed) *The Institutionalized Use of the Drug Qat in North Yemen*, Kluwer Academic Publishers, Dordrecht, The Netherlands, p 284

Khanuja SPS, Shasany AK, Daroker MP, Kumar S (1998) DNA fingerprinting of plant genetic resources: the need of time. *Journal of Medicinal and Aromatic Plant Sciences* **20**, 348-351

Krikorian AD (1984) Qat and its use: an historical perspective. *Journal of Ethnopharmacology* **12**, 115-178

Krikorian AD (1985) Growth mode and leaf arrangements in *Catha edulis* (khat). *Economic Botany* **39 (4)**, 514-521

Lamote V, Roldan-Ruiz I, Coart E, de Loose M, Bockstaele VE (2002) A study of genetic variation in *Iris pseudacorus* populations using amplified fragment length polymorphism (AFLPs). *Aquatic Botany* **73 (1)**, 19-31

Lemessa D (2001) Khat (*Catha edulis*): Botany, distribution, cultivation, usage and economics in Ethiopia. Report UN emergencies unit for Ethiopia, 14 pp

Luqman W, Danowski TS (1976) The use of khat (*Catha edulis*) in Yemen. Social and medical observations. *Annals of Internal Medicine* **85**, 246-249

Margetts EL (1967) Miraa and myrrh in East Africa - clinical notes about *Catha edulis. Economic Botany* **21**, 358-362

Muharram IA, Mukred A, Noman A (2002) Qat and agriculture. Conference Discussion Materials, *National Conference on Qat*, 6-7 April 2002, Sana'a, Republic of Yemen, pp 54-68

Nayak S, Rout GR, Das P (2003) Evaluation of the genetic variability in bamboo using RAPD markers. *Plant Soil Environment* **49 (1)**, 24-28

Pojanagaroon S, Praphet R, Kaewrak C, Yotdi A (2004) Characterization of Krachai-Dam (*Kaempferia parviflora*) cultivars using RAPD markers, morphological traits and chemical components of essential oil from rhizomes. *Proceedings of the 42nd Kasetsart University Annual Conference*, Kasetsart, Thailand, pp 56-65

Porebski S, Bailer G, Baun BR (1997) Modification of a CTAB DNA extraction protocol for plant containing high polysaccharides and polyphenolic compounds, *Plant Molecular Biology Reporter* **15 (1)**, 8-15

Porter WM, Smith DH (1982) Detection of identification errors in germplasm collection. *Crop Science* **22**, 701-703

Prakash DP, Narayanaswamy P, Sondur SN (2002) Analysis of molecular diversity in guava using RAPD markers. *Journal of Horticultural Science and Biotechnology* **77**, 287-293

Provost A, Wilkinson MJ (1999) A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. *Theoretical and Applied Genetics* **98**, 107-112

Ramadan MA, Tash FM, Qirbi AA, Yehia HL (1981) Evaluation of cathine and cathinone in khat all over different areas of Yemen and factors affecting it. Khat Research biochemistry department, Sana'a University, Sana'a, 63 pp

Rousseau C, Said TM, Gagné M, Bibeau G (2004) Between myth and madness: the premigration dream of leaving among young Somali refugees. *Culture, Medicine and Psychiatry* **22 (4)**, 385-411

Sambrook J, Fristch EF, Maniatis T (1989) *Molecular Cloning. A laboratory Manual* (2nd Edn), Cold Spring Harbor Laboratory Press, NY, 1659 pp

Sawair FA, Al-Mutwakel A, Al-Eryani K, Al-Surhy A, Maruyama S, Cheng J, Al-Sharabi A, Saku T (2007) High relative frequency of oral squamous cell carcinoma in Yemen: qat and tobacco chewing as its aetiological background. *International Journal of Environmental Health Research* **17 (3)**, 185-195

Schopen A (1978) Das quat: geschichte und gebrauch des genussmittels *Catha edulis* Forsk., In Der Arabischen Republic Jemen, Wiesbaden, Franz Steiner Verlag GMBH, 97 pp

Sokal RR, Sneath PHA (1973) *Principles of Numerical Taxonomy*, W. H. Freeman and Co., San Francisco, CA, USA, 359 pp

Sun Q, Ni Z, Liu Z, Gao J, Huan T (1998) Genetic relationship and diversity among Tibetan wheat, common wheat and European spelt wheat revealed by RAPD markers. *Euphytica* **99**, 205-211

Tiwari KR, Penner GA, Warkentin TD (1998) Identification of coupling and repulsion phase RAPD markers for powdery mildew resistance gene earl in pea. *Genome* **41**, 440-444

Vanwalleghem IE, **Vanwalleghem PW**, **de Bleecker JL** (2006) Khat chewing can cause stroke. *Cerebrovascular Diseases* **22 (2-3)**, 198-200

Vos P, Hogers R, Bleeker M, Reijians M, Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP, a new technique for DNA fingerprinting. *Nucleic Acid Research* **23 (21)**, 4407-4414

Ward JH (1963) Hierarchic grouping to optimize an objective function. *Journal of the American Statistical Association* **58**, 236-239

Warfa N, Klein A, Bhui K, Leavey G, Craig T, Alfred SS (2007) Khat use and mental illness: a critical review. *Social Science and Medicine* **65 (2)**, 309-318

Williams JG, Kubelki AR, Livak KJ, Rafaleski JA, Tingy SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* **18**, 6531-6535

Yeh FC, Boyle TJB (1997) POPGENE Version 1.32 Microsoft window-based software for population genetics analysis, Department of Renewable Resources, University of Alberta, Alberta

Zhiyi B, Hua Z, Jinsong B, Zhe H (2003) Analysis of genetic diversity among *Loropetalum chinense* var. *rubrum* by RAPD. *Journal of Zhejiang University Agriculture and Life Sciences* **29 (6)**, 665-670