Theor Appl Genet (2000) 101:269–278

W. Choumane · P. Winter · F. Weigand · G. Kahl

Conservation and variability of sequence-tagged microsatellite sites (STMSs) from chickpea (*Cicer aerietinum* L.) within the genus *Cicer*

Received: 13 August 1999 / Accepted: 11 November 1999

Abstract The conservation of 90 microsatellite-flanking sequences from chickpea in 39 accessions of eight annual and 1 accession of a perennial species of the genus *Cicer* was investigated. All of the primer sequences successfully amplified microsatellites in related species, indicating the conservation of microsatellite-flanking sequences in chickpea's relatives. However, the degree of conservation of the primer sites varied between species depending on their known phylogenetic relationship to chickpea, ranging from 92.2% in *C. reticulatum*, chickpea's closest relative and potential ancestor, down to 50% for *C. cuneatum*. A phylogenetic tree revealed that chickpea and the other members of its crossability group were more closely related to the perennial *C. anatolicum* than to other annual species of the genus. Considerable variation in size and number of amplification products between and within species was observed. Sequence analysis of highly divergent amplification products proved that variation is either due to large differences in the number of microsatellite repeats or to the amplification of a locus unrelated to the one amplified from chickpea. Sequence information and bootstrapping using PAUP suggested that STMSs derived from chickpea may be efficiently and reliably used for synteny studies in chickpea's crossability group, including *C. anatolicum*. However, care should be taken when applying these markers to other species of the genus. Considering the data presented here and the known historical record, the age of section *Monocicer*, including chickpea, is estimated to be about 100,000 years.

Communicated by. G. Wenzel

W. Choumane
International Centre for Agricultural Research
in the Dry Areas (ICARDA), P.O. Box 5466, Aleppo, Syria,
and Faculty of Agriculture, University of Tishreen, Lattakia, Syria

P. Winter (✉) · G. Kahl
Plant Molecular Biology, Biocentre, University of Frankfurt/Main,
D-60439 Frankfurt, Germany
e-mail: P.winter@em.uni-frankfurt.de

F. Weigand
ICARDA, P.O. Box 5466, Aleppo, Syria

Key words *Cicer* species · Microsatellites · Zero-allele · Phylogenetic tree · Evolution

Introduction

Phylogeny of the genus *Cicer* has attracted considerable attention, because it harbours the third most important grain legume crop world-wide: chickpea. The genus comprises nine annual and 33 perennial species classified into four sections based on their life-cycle and morphological and geographical criteria (van der Maesen 1987). Evolutionary and genetic relationships of the eight annual species grouped into section *Monocicer* are well-described. *C. chorassanicum*, the last known of the annual species, has been classified into section *Chamaecicer*. Annual species have been subdivided into four groups on the basis of crossability (Ladizinsky and Adler 1976), karyotype (Ocampo et al. 1992; Abbo et al. 1994), isozyme polymorphisms (Tuwafe et al. 1988; Gaur and Slinkard 1990 a, b; Ahmad et al. 1992; Labdi et al. 1996) and seed storage protein characteristics (Ahmad and Slinkard 1992). The first group contains the cultigen, its presumable ancestor *C. reticulatum* and *C. echinospermum*. Both wild species produce fertile hybrids in crosses with the cultigen, though fertility barriers between chickpea and *C. echinospermum* do exist (Singh and Ocampo 1997). The perennial *C. anatolicum*, first grouped into section *Polycicer*, has also been placed into the first group based on isozyme similarities (Kazan and Muehlbauer 1991). The second group comprises *C. bijugum*, *C. pinnatifidum*, *C. judaicum* and *C. yamashitae*. The remaining two species, *C. cuneatum* and *C. chorassanicum*, can neither be crossed with each other nor with any other species and, therefore, make up the third and fourth group, respectively (Kazan and Muehlbauer 1991).

Several types of molecular markers have been used in plant breeding for a wide range of applications. These include the estimation of genetic diversity and relationships between accessions and species and map-based

cloning of agriculturally interesting genes (review in Winter and Kahl 1995). Generally, single-locus co-dominant and multi-locus dominant markers are used side by side. However, for many purposes co-dominant markers are preferred because they allow tagging and mapping of the same loci in many different populations and even species. Probably the best markers in this respect descend from small, tandemly arranged repetitive elements called microsatellites, simple-sequence repeats (SSRs) or simple tandem repeats (STRs). Microsatellites consist of 1- to 5-nucleotide-long units that are organized in loci containing hundreds or more of them. Microsatellites frequently change their length by slipped-strand mispairing and other less understood processes (Levinson and Gutman 1987; Zischler et al. 1992) that lead to variable numbers of tandem repeats (VNTRs, Nakamura et al. 1987). Mutation rates of SSRs are high: 2.5×10^{-5} to 1×10^{-2} mutations per locus per gamete per generation have been estimated (Weber and Wong 1993), which accounts for their polymorphism. The surrounding single-copy sequences are normally not affected. To convert them into markers, microsatellites are cloned, sequenced, and amplified by the polymerase chain reaction (PCR), using oligonucleotides as primers directed against the flanking monomorphic DNA (Litt and Luty 1989). These sequence-tagged microsatellite sites (STMSs, Beckmann and Soller 1990) are co-dominant markers like restriction fragment length polymorphisms (RFLPs). However, the number of alleles they detect in a population is usually much higher as compared to RFLPs. Microsatellites are very abundant. In soybean, for instance, $(CA)_n$ repeats occur every 29 kb on average (Morgante and Olivieri 1993). In chickpea, $(TAA)_n$, $(GA)_n$ and $(CA)_n$ repeats are present at more than 12,000 loci with an average spacing of around 60 kb (Huettel et al. 1999). Our efforts to improve chickpea's agricultural value by marker-assisted breeding has resulted in the generation of more than 200 STMS markers, many of them used for the generation of genetic maps of the crop (Huttel et al. 1999; Winter et al. 1999).

It has frequently been observed that SSR-flanking sequences are conserved in closely related species. For example, human STMS amplified corresponding loci from chimpanzees (Rubinstein et al. 1995; Garza et al. 1995), and bovine STMS alleles amplified those from goats and sheep (Moore et al. 1991; Forbes et al. 1995; Pepin et al. 1995). Transferability of markers was observed in *Canidae* (Roy et al. 1994), *Cetaceae* (Schloetterer and Tautz 1991), different species of pines (Smith and Devery 1995; Karhu et al. 1996; Echt and May-Marquart 1997) and *Medicago* (Diwan et al. 1997).

One of the aims of our continuing work is the marker-assisted utilisation of the primary and secondary gene pool of chickpea for the improvement of the crop. Therefore, in the investigation reported here we examined whether and to which extent STMS primers designed for the cultigen could also be applied to genome analysis of wild *Cicer* species. We explored the possibility that the conservation of microsatellite-flanking sequences reflects the known evolutionary relationship between these species and estimated the age of section *Monocicer* based on both these findings and on the known historical record of chickpea culture in the Middle East (Zohary and Hopf 1993). Further, we investigated what underlies the differences in the number and size of amplification products derived from the same or different species observed in preliminary experiments.

Table 1 Accession numbers and geographical origin of wild *Cicer* species investigated in this study

Species	Accession	Origin	Species	Accession	Origin
C. anatolicum		Turkey	*C. judaicum*	ILWC7	Lebanon
C. bijugum	ILWC7	Turkey		ILWC31	Jordan
	ILWC32	Turkey		ILWC43	Syria
	ILWC79	Turkey		ILWC44	Syria
	ILWC195	Syria		ILWC273	Lebanon
	ILWC240	Turkey	*C. pinnatifidum*	ILWC9	Turkey
C. cuneatum	ILWC37	Ethiopia		ILWC29	Turkey
	ILWC40	Ethiopia		ILWC49	Syria
	ILWC185	Ethiopia		ILWC171	Turkey
	ILWC187	Ethiopia		ILWC226	Turkey
	ILWC232	Ethiopia	*C. reticulatum*	ILWC105	Turkey
C. chorassanicum	ILWC233	Afghanistan		ILWC109	Turkey
	ILWC90	Afghanistan		ILWC123	Turkey
	ILWC146	Afghanistan		ILWC242	Turkey
	ILWC147	Afghanistan		ILWC247	Turkey
C. echinospermum	ILWC180	Turkey	*C. yamashitae*	ILWC3	Afghanistan
	ILWC181	Turkey		ILWC53	Afghanistan
	ILWC235	Turkey		ILWC55	Afghanistan
	ILWC238	Turkey		ILWC214	Afghanistan
	ILWC239	Turkey		ILWC215	Afghanistan

Materials and Methods

Plant material

Seeds from annual wild *Cicer* species and chickpea accession ILC 3279, for which the primers were designed (Huettel et al. 1999; Winter et al. 1999) and used as reference here, were obtained from the Germplasm Resource Unit at ICARDA, Aleppo, Syria, and germinated in the greenhouse. Species and accession numbers are listed in Table 1 together with their geographical origin. Seeds of the perennial *C. anatolicum* were a kind gift of Prof. Fred J. Muehlbauer, Washington State University, Pullman, USA.

DNA extraction, STMS markers, PCR, separation of amplification products and data analysis

Total DNA was extracted from young leaflets of a single plant of each accession using the CTAB method (Rogers and Bendich 1985). The 90 primer pairs used in this study have been described either by Huettel et al. (1999, CaSTS markers) or by Winter et al. (1999, all other markers). PCR was performed in a Perking Elmer 9600 thermocycler. Reaction mixtures of 20 µl contained 75–100 ng DNA, 100 µ*M* of each dNTP (Roth, Germany), 1.5 m*M* $MgCl_2$, 45 pmol of each primer and 0.4 U Silverstar *Taq* DNA polymerase (Eurogentec, Belgium). Initial denaturation of DNA at 96°C for 2 min was followed by 35 cycles of denaturation at 96°C (20 s) and annealing of primers at 55°C (50 s) and elongation at 60°C (50 s). After a final extension at 60°C for 5 min, amplification products were separated on 2% agarose gels in TBE buffer (Sambrook et al. 1989). The DNA was stained with ethidium bromide, bands documented using a CCD camera and scored as either present or absent. A phylogenetic relationship based on the information of zero-alleles was established using the PAUP program (D.L. Swofford, Washington, DC 20 560, USA).

Cloning and sequencing of PCR fragments

DNAs from *C. bijugum*, *C. reticulatum* and *C. cuneatum* were amplified with TA18, TA14 and TA37 primer pairs, respectively, and DNA from *C. arietinum* was used as reference. The amplification reactions were performed as described above. Amplified DNA from *C. reticulatum* was purified using a Qiaquick column (Qiagen). For the other species, the amplified fragments were separated on agarose gels, cut out and purified using the QiaexII Gel Extraction Kit (Qiagen). The purified fragments were inserted into the pGEM T-Easy Vector (Promega) and transformed into *E. coli* DH10B (Biorad) by electroporation. Clones were selected on LB medium containing 100 µg ampicillin/ml. Enzymatic digestion and PCR amplification with the appropriate primers were used to select interesting clones. Plasmid DNA was extracted from 10-ml overnight cultures by alkaline lysis (Sambrook et al. 1989), and 0.5–1 µg of purified plasmid DNA was used for sequencing. Sequencing was performed with a thermal cycle sequencing kit using an ABI PRISM Model 377 version3 autosequencer. Both strands were sequenced for each fragment.

Results

Conservation of chickpea primer binding sites in wild *Cicer* species

The DNAs from 1 accession each of the eight annual and one perennial wild *Cicer* species were amplified with the 90 STMS primer pairs to determine whether and with which primer pairs amplification products could be obtained. DNA from chickpea accession ILC 3279 served as the control. After the separation of amplification products on agarose gels, bands were scored as present or absent without determining their exact molecular weights. Results of these experiments are summarised in Table 2. Several STMS primers created bands in all species. For other loci, only 1 or 2 STMS-species combinations produced amplification products. For example, STMS Ta72 produced bands only in *C. anatolicum* and *C. yamashitae*, and Ta45 did so solely in *C. reticulatum and C. echinospermum*, whereas control DNA was consistently amplified. Such loci were scored as zero-alleles (Table 2). A zero-allele was determined after several repetitions of the assay to exclude an artefact. The same template DNA allowed to amplify distinct products at other loci.

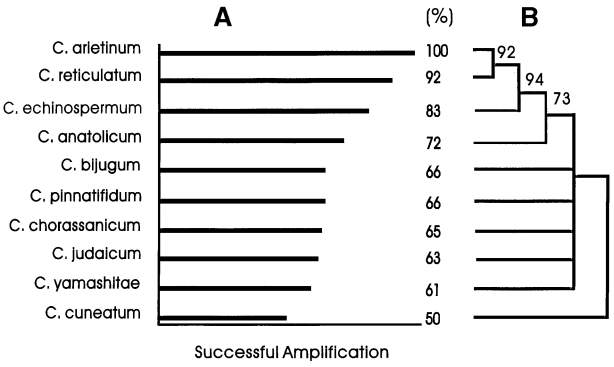

Fig. 1 Percentage of successful amplification of STMS loci from *Cicer species* (**A**) and the resulting phylogenetic tree (**B**). *Numbers* in **B** indicate bootstrapping values obtained after 500 times of bootstrapping

The percentage of successful amplification is summarised in Fig. 1. At a first glance these data seem to reflect phylogenetic relationships between the species as determined by isozymes and other criteria (e.g. Kazan and Muehlbauer 1991). However, to verify that STMS markers can in fact be used for the determination of relationships within the genus *Cicer*, we analysed the data by PAUP and scrutinised it by 500 times bootstrapping. PAUP produced a phylogenetic tree (Fig. 1B) that confirms the tree derived from isozyme analysis (Kazan and Muehlbauer 1991). Though the two trees are similar, high bootstrapping values, indicative of the high reliability of the data, could only be obtained for members of the first crossability group and, astonishingly, also for the *perennial C. anatolicum*. Bootstrapping values for the annual species crossability groups 2, 3 and 4 were below 50 and not included in the tree.

Detection of STMS polymorphism within Cicer species

In another set of experiments, 31 STMS primer pairs that successfully amplified DNA from at least one wild species were used to amplify DNA from all other available accessions (see Table 1) to assess whether they detect polymorphism within the species. The results of these

Table 2 Conservation of microsatellite-flanking loci between different wild *Cicer* species

Loci	*C. urietinum*	C. anatolicum	C. reticulatum	C. echinospermum	C. judaicum	C. pinnatifidum	C. bijugum	C. chorassanicum	C. yamashitae	C. cuneatum
Ta2	1[a]	0	1	1	1	1	1	0	0	1
Tal1	1	1	1	1	1	1	1	1	1	0
Tal4s	1	0	1	1	1	0	1	1	1	1
Ta21	1	0	1	0	0	0	0	1	1	0
Ta34	1	1	1	1	1	1	1	1	1	0
Ta34s	1	1	1	1	1	1	1	1	1	0
Ta45	1	0	1	1	0	0	0	0	0	0
Ta52	1	0	1	0	0	0	0	0	0	1
Ta54	1	0	1	1	0	0	1	0	0	0
Ta59	1	1	1	1	1	1	1	1	0	0
Ta64	1	1	1	1	0	0	0	0	0	0
Ta72	1	1	0	0	0	0	0	0	1	0
Ta78	1	1	0	0	0	0	0	0	0	0
Ta80s	1	0	1	1	0	0	1	1	1	1
Ta89	1	0	1	1	0	0	0	0	0	0
Ta96s	1	0	1	1	0	1	0	0	1	0
Tal04	1	1	1	1	1	1	1	1	0	0
Tal06	1	0	1	1	0	0	0	1	1	0
Tal10	1	1	1	1	1	1	1	1	1	0
Tal17	1	1	1	0	0	0	0	1	1	0
Tal25	1	1	1	1	1	1	1	1	1	0
Tal46s	1	0	1	1	1	1	0	0	0	0
Tal80	1	1	1	1	0	1	0	0	1	1
Ta203	1	0	0	0	0	0	1	0	0	0
Tr1	1	1	1	1	1	1	1	1	1	0
Tr7	1	0	1	0	0	1	1	1	0	0
Rr19	1	0	1	0	0	1	1	0	0	0
Tr20s	1	0	0	1	0	0	0	0	0	0
Tr23	1	1	1	0	1	0	0	0	0	0
Tr29	1	1	0	0	0	0	0	0	0	0
Tr43	1	1	1	1	1	1	1	1	0	0
Tr44	1	1	1	1	1	1	1	0	1	0
Tr56	1	1	1	1	0	0	0	0	0	0
Tr58s	1	1	1	1	0	0	0	0	0	1
TS12	1	1	1	0	0	0	0	1	0	0
ITS19	1	0	1	1	0	0	0	0	0	0
TS23	1	1	1	1	0	0	0	0	0	0
TS45	1	0	1	1	0	0	0	1	0	0
TS52	1	0	1	0	0	0	0	0	0	0
TS53	1	0	1	0	0	0	0	0	0	0
TS54	1	0	1	1	0	0	0	0	0	1
TS57	1	1	1	1	0	1	1	1	1	0
TS72	1	1	1	1	0	1	0	0	0	1
TS83	1	0	1	1	1	1	1	1	1	1
TS104	1	0	1	1	0	0	0	0	0	0
TS105	1	1	1	1	1	1	1	1	0	0
TS129	1	0	1	1	1	1	1	0	0	1
CaSTS2	1	1	0	0	0	0	0	1	0	0
CaSTS5	1	1	1	1	1	1	1	1	1	0
CaSTS7	1	0	1	1	1	1	1	1	1	1
CaSTS10	1	1	1	1	0	0	0	0	0	0
CaSTS16	1	1	1	1	1	1	1	1	1	0
CaSTS19	1	1	1	1	1	1	1	0	0	0
CaSTS20	1	1	0	1	0	0	0	0	0	0
CaSTS23	1	1	1	1	1	1	0	1	1	0
CaSTS24	1	1	1	0	1	1	1	0	0	0
CaSTS28	1	0	1	1	1	0	1	1	1	1
Cicer loci[b]	1	1	1	1	1	1	1	1	1	1
No. of loci tested	90	90	90	90	90	90	90	90	90	90
No. of conserved	90	65	83	75	57	60	60	59	55	45
% of conserved	100	72.2	92.2	83.3	63.3	66.6	66.6	65.6	61.1	50

[a] 1, Presence of amplification products; 0, zero-allele

[b] *Cicer loci*: represent the 33 primer pairs that give amplification products in all investigated species (Ta1, Ta3, Ta5,„Ta8, Ta18, Ta20, Ta22, Ta27, Ta37, Ta71, Ta76, Tar2 s, Tr9, Tr43 s, Tr59, Ts17, Ts53, Ts47, CaSTS1, CaSTS3, CaSTS4, CaSTS7, CaSTS8, CaSTS9, CaSTS11, CaSTS12, CaSTS17, CaSTS18, CaSTS21, CaSTS22, CaSTS25, CaSTS26, CaSTS27). Note that all markers designated CaSTS and a number are described in Huettel et al. (1999). All other markers are published in Winter et al. (1999)

Table 3 STMS polymorphisms[a] between and within species of the genus *Cicer*. (*C.arie C. arietinum · C.ana. C. anatolicum · C.ret. C. reticulatum · C.echino. C. echinospermum · C.judai. C. judaicum · C.pinn. C. pinnatifidum · C.bij. C.bijugum · C.choras. C. chorassanicum · C.yam. C. yamashitae · C.cun. C. cuneatum*)

Loci	Number of bands Accession	*C.arie.*	*C.ana.*	*C.retic.* I2345	*C.echino.* I2345	*C.judai.* I2345	*C.pinn.* I2345	*C.bij.* I2345	*C.choras.* I234	*C.yam.* I2345	*C.cun.* I2345
Ta1	Number of bands	1	2	11212	11111	12111	22332	22222	1111	22222	33333
	polymorphic			Yes	No	Yes	Yes	No	No	Yes	No
Ta2	Number of bands	1	0	12111	11111	11222	21132	22444	0000	00000	33333
	polymorphic			Yes	Yes	Yes	Yes	Yes	No	No	No
Ta3	Number of bands	1	1	11111	11111	11111	11111	11111	1111	11111	11111
	polymorphic			No	No	Yes	Yes	No	No	No	No
Ta5	Number of bands	1	1	22211	11111	12111	21121	22111	3333	11112	11111
	polymorphic			Yes	Yes	Yes	Yes	Yes	No	Yes	No
Ta18	Number of bands	1	2	11111	11111	23333	25344	23233	1111	32332	22112
	polymorphic			Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Ta14s	Number of bands	1	0	11111	11111	11111	00000	11111	1111	11111	00000
	polymorphic			Yes	Yes	No	No	No	No	No	No
Ta20	Number of bands	1	1	22221	22322	12214	44411	34211	1222	11111	11111
	polymorphic			Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ta22	Number of bands	2	1	22222	22222	11111	31312	11111	2222	22222	22222
	polymorphic			Yes	Yes	Yes	Yes	No	No	Yes	Yes
Ta27	Number of bands	1	1	12222	11111	11111	22222	33333	3333	33333	11111
	polymorphic			Yes	Yes	No	No	No	No	No	No
Ta34	Number of bands	1	1	21211	11222	12222	32121	12212	1111	11111	00000
	polymorphic			Yes	Yes	Yes	Yes	Yes	No	Yes	No
Ta37	Number of bands	1	1	21221	22222	11111	11201	22222	3333	11111	33333
	polymorphic			Yes	No	Yes	Yes	No	No	Yes	No
Ta59	Number of bands	2	1	21111	11111	12111	11001	11211	1211	00000	10222
	polymorphic			Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
TA71	Number of bands	1	1	12121	11111	22222	22222	22222	3221	12113	22332
	polymorphic			Yes	No	No	No	No	Yes	Yes	Yes
Ta76	Number of bands	1	1	11111	11111	11111	11000	11111	1111	00001	01101
	polymorphic			Yes	Yes	No	Yes	No	No	Yes	Yes
Ta80s	Number of bands	1	0	12111	11111	00000	00000	00112	2112	22222	11111
	polymorphic			Yes	Yes	No	No	Yes	Yes	Yes	Yes
Ta96s	Number of bands	1	0	11111	11111	00000	00100	00000	0000	11110	00000
	polymorphic			Yes	Yes	No	Yes	No	No	Yes	No
Ta106	Number of bands	1	0	21111	11100	00000	00000	00000	1111	11111	00000
	polymorphic			Yes	Yes	No	No	No	Yes	Yes	No
Ta180	Number of bands	1	2	10110	11111	00000	00100	00000	1000	01110	11110
	polymorphic			Yes	Yes	No	YYes	No	Yes	Yes	Yes
Ta203	Number of bands	1	0	00000	00000	00000	10010	11111	0000	00000	00000
	polymorphic			No	No	No	Yes	No	No	No	No
Tr1	Number of bands	1	1	11111	11111	12111	11111	11000	1001	02012	00000
	polymorphic			Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Tr43	Number of bands	1	1	22222	22222	11222	11111	22222	1221	00000	00000
	polymorphic			Yes	No	Yes	Yes	No	Yes	No	No
Tr58s	Number of bands	1	1	12213	22111	00000	01000	00000	0001	00000	11111
	polymorphic			Yes	Yes	No	Yes	No	Yes	No	No
TS12	Number of bands	1	0	11101	00000	00000	00000	00000	1011	00000	00000
	polymorphic			Yes	No	No	No	No	Yes	No	No
TS45	Number of bands	1	0	11111	11100	00000	00000	00000	1111	00000	00000
	polymorphic			Yes	Yes	No	No	No		No	No
CaSTS7	Number of bands	1	0	11111	11111	11111	11111	11111	1111	22122	11111
	polymorphic			Yes	Yes	No	Yes	Yes	Yes	Yes	No
CaSTS5	Number of bands	1	1	11111	11111	01111	01010	10000	1110	10000	01010
	polymorphic			Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
CaSTS12	Number of bands	1	1	11111	11111	11121	21122	11111	1111	11111	11111
	polymorphic			Yes	No	Yes	Yes	No	No	Yes	No
CaSTS19	Number of bands	1	1	11111	11111	01000	11011	11011	0000	00000	00000
	polymorphic			No	No	Yes	Yes	Yes	No	No	No
CaSTS21	Number of bands	1	1	11111	11111	11111	11111	11111	1111	11111	11111
	polymorphism			Yes	No	Yes	Yes	No	Yes	Yes	No
CaSTS22	Number of bands	1	1	11111	11111	11111	11111	11111	1111	11111	11111
	polymorphic			Yes	Yes	No	Yes	Yes	No	Yes	No
CaSTS28	Number of bands	1	0	11101	11122	11222	00000	22222	1111	1221	12222
	polymorphic			Yes	Yes	Yes	No	Yes	Yes	Yes	Yes

[a] The numbers 1,2,3,4 and 5 in the same boxes below the different species represent numbers analysed for each species, where the accession numbers correspond to the accessions coded in Table 1. In the case where all accessions of a species have an identical number of bands (e.g. 11111) and polymorphism is ”yes”, the polymorphism affects only the molecular weight of the fragments

A

```
             1                                                                   36
C.arietinum  ACTTACATGAATTATC(T)3CTTGGT(C)3TT G C(T)3 A--------------------------
C.cuneatum2  ACTTACATGAATTATC(T)3CTTGGT(C)3TT A C(T)3 ATCTGATAATATC(T)3ATT(G)3TGG

             37                                                                  73
C.arietinum  ------------------------------ACGAT(A)4TC(ATT)2(G)3TGTCC(T)4GAAT(A)3G
C.cuneatum2  C(T)7AATACAAGTGGAGCAA(T)4ATATCC-----(A)4TC(ATT)2(G)3TG-C--------------

             74                                                                  186
C.arietinum  TGGAGC(A)3(T)3(A)4(TAA)20TAAAT(A)4T(A)5TC(A)5C(A)7AG(A)4----- GTAG(A)4T
C.cuneatum2  --------------------------A-(A)4-(A)5TC--------AG(A)4AATTA ----------

             195                                                                 242
C.arietinum  (A)3TGTTCTTAGT-CCTAG(T)3G(A)4GA--------------- TACAATAAGGAGCTCCACTC-
C.cuneatum2  ------- GTTAGTTCC-AG(T)3G(A)4GACACTATC(A)4GTGT --------------TCCACTCC

             243                                        283
C.arietinum  ----GGATTAACACA(T)3GATTGATG(A)3GATTA(T)3GAATACG
C.cuneatum2  CACC-----AACACA(T)3GATTGATG(A)3GATTA(T)3GAATACG
```

B

```
             1                                                                   57
C.arietinum  ACTTACATGAATTATC(T)3CTTGGTCC-CTTGC(T)3A-----ACGAT(A)4TCATT A TT(G)3
C.cuneatum3  ACTTACATGAATTATC(T)3CTTGGTCCCCT-GC(T)3ATTGAG----T(A)4TCATT C TT(G)3

             58                                                                  63
C.arietinum  TGT---------------------------------------------------------------CCT
C.cuneatum3  TGC(ATT)2GATTATAGTGGAGCAACTT(A)7CAT(A)4TC(A)6TAA(TGT)2TCATTCCGG-----

             64                                                                  188
C.arietinum  (T)3GAAT(A)3GTGGAGC(A)3(T)3(A)4(TAA)20T(A)3T(A)4T(A)5TC(A)5C(A)8G(A)4GT
C.cuneatum3  ----------------------------------------------------------------------

             189                                                                 254
C.arietinum  AG(A)4T(A)3TGTTCTTAGTCCTAG(T)3G(A)4GATACAATAAGGAGCTCCACTCGGATTAACACA
C.cuneatum3  ----------------------------------------------------------------------

             255                         283
C.arietinum  (T)3GATTGATG(A)3GATTA(T)3GAATACG
C.cuneatum3  (T)3GATTGATG(A)3GATTA(T)3GAATACG
```

Fig. 2a, b Comparison of sequences of the amplification products derived from STMS Ta37 from *C. arietinum* with *C. cuneatum* fragment cun2 (**a**) and *C. cuneatum* fragment cun3 (**b**). – Sequences not present in one or the other amplification product

experiments are summarised in Table 3. They revealed extensive polymorphism not only based on the size of the amplification products but also on their number. Whereas in *C. arietinum* only one or, exceptionally, two bands were visible, the same primer pair produced up to five bands in the wild species. The number of bands did not only vary between species but also between accessions of a species. For example, the number of amplification products at locus Ta18 in *C. pinnatifidum* varied between two (accession ILWC9) and five (accession ILWC29). In some cases, as for example in the combination of STMS Ta37 and *C. cuneatum*, one band was amplified from the control, but three similarly sized bands from each *C. cuneatum* accession. In 15 cases, primers reproducibly created products only from some but not all accessions of a species. This was especially pronounced for CaSTS5 which was present in six of the nine species analysed.

Considerable size differences of amplification products between and within the species were detected. For example, Ta18 produced a band of only 150 bp from DNA of ILC 3279 but a 1,094 bp fragment from all accessions of *C. judaicum*. STMS Ta14 generated a 249-bp fragment in the control and a 274-bp fragment in *C. reticulatum* accession ILWC247 but a larger fragment of 726 bp in accession ILWC105 of the same species. In contrast to the large amount of polymorphism detected in several cases, sometimes only one amplification product was obtained from all species, including the control (Table 2).

Sequence analysis of selected amplification products

To explore the cause of the three most frequently observed differences in number and size of amplification products from different species and different accessions of the same species, we cloned a number of specific fragments in *E. coli*, sequenced them and then compared their sequence with the corresponding product from chickpea. The following cases were analysed:

1) differences in the number of bands between different species (locus Ta37)
2) large size differences between products from different species (locus Ta18)

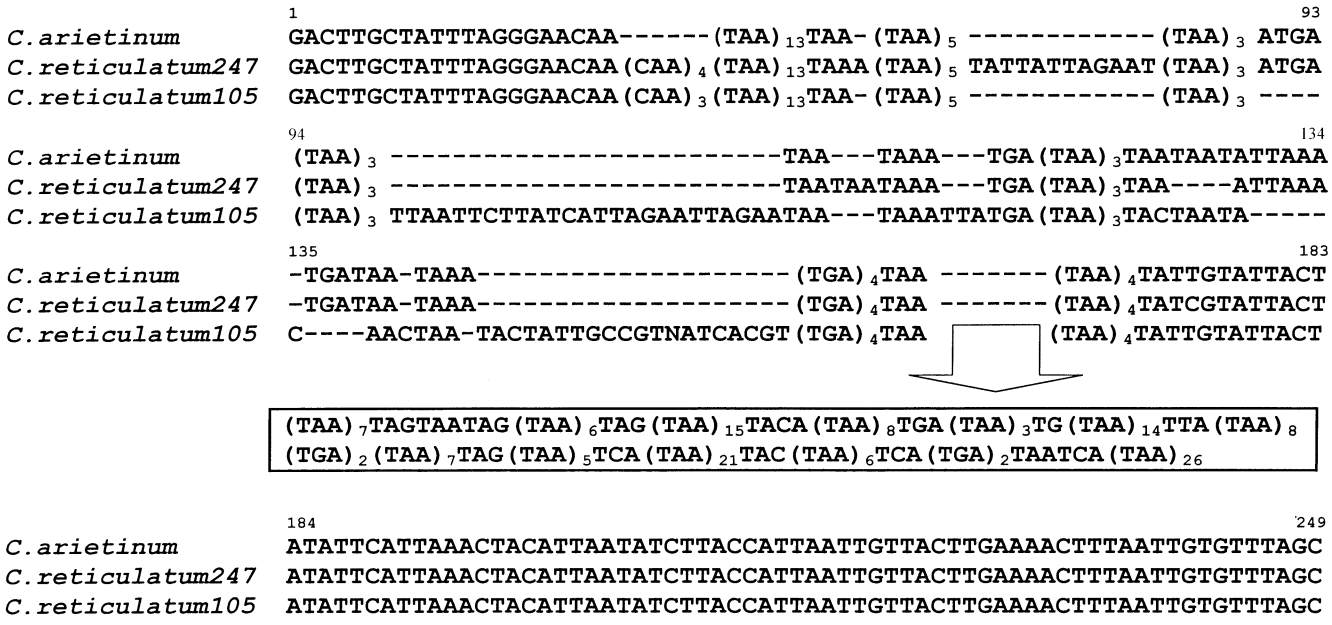

Fig. 3 Comparison between the sequences of amplification products obtained by STMS Ta14 from *C. arietinum* and from *C. reticulatum* accessions ILWC247 and ILWC105, respectively. – Insertions and deletions of sequences

3) large variation in the size of products among accessions of the same species (locus Ta14).

As an example of the first case, the three fragments amplified with Ta37 from all *C. cuneatum* accessions with Ta37 – named Cun1, Cun2 and Cun3 – were sequenced. Cun1 carried only one of the two primer binding and insertion sites of the vector. We hypothesise that during cloning and propagation of the amplification product these sequences were lost due to recombination within *E. coli*. Most of the sequence was readable and shared some small homologies with the fragment from chickpea but without the $(TAA)_{20}$ tract present there. Because the sequence was probably mutated by *E. coli*, we did not analyse it further. Cun2 was 62 bp shorter than the 283-bp amplicon from chickpea, and here primer binding and vector insertion sites were present. Part of the sequence of the 221-bp fragment was similar to the chickpea fragment. As shown in Fig. 2a, again no (TAA) microsatellite was found. Cun3 was 155 bp long, partly homologous to the chickpea product and again lacked the microsatellite. Instead, it contained ten small (A) repeats ranging in size from $(A)_4$ to $(A)_8$. In Fig. 2b this sequence is compared to the one from chickpea.

The second case – large size differences between amplicons from different species – was analysed using the 1,094-bp fragment from *C. judaicum* amplified by Ta18. This fragment did not contain the $(TAA)_{22}$ microsatellite found in the reference amplicon from chickpea. However, the (TAA) motif appeared to be dispersed throughout the whole fragment along with stretches of A and T. These were considerably longer than in the chickpea fragment, which also contained A-rich regions (data not shown). The rest of the sequence was not homologous to that from chickpea.

Large size differences in amplification products obtained with Ta14 from different accessions of the same species represent the third case. It was analysed by sequencing the 274-bp fragment of *C. reticulatum* accession ILWC247 and the 724-bp fragment of ILWC105. A comparison of these sequences with the 249-bp fragment from *C. arietinum* revealed that they were probably derived from the same locus. The 25-bp difference between the amplification products of *C. arietinum* and *C. reticulatum* accession ILWC 247 was caused by an insertion of $(CAA)_4$ at position 22, an insertion of $TA(TTA)_2GAAT$ at position 79 and other less structured insertions, deletions and substitutions along the sequence (Fig. 3). The 724-bp amplification product of ILWC105 differed from that of *C. arietinum* by 475 bp. This extensive variation was mainly caused by amplification of the 35 (TAA) repeats present in *C. arietinum* (position 157) to 161 units in ILWC105. Insertions of 25 bp (position 101, $TTAATTCTTATCA[TTAGAA]_2$) and of 20 bp (at position 143) as well as smaller modifications were visible (Fig. 3).

Discussion

STMS markers and phylogeny of the genus *Cicer*

The possibility of using microsatellite-flanking primers in more than one species is desirable in terms of the extensive efforts and expenses involved in designing each one of them. The conservation of microsatellite-flanking regions has actually been demonstrated in *Brassica*, *Vitis* and *Cucurbita* species (Kresovich et al. 1995; Thomas and Scott 1993; Katzir et al. 1996). Limited conservation among related species has also been reported (wheat, barley and rye, Röder et al. 1995).

In an effort to test the applicability of chickpea-derived STMS markers for improvement of the crop via utilisation of the gene pool of its wild relatives, we amplified the DNA of several of these species with 90 STMS primer pairs. Amplification either resulted in the presence or absence of products. Absence was interpreted as being caused by mutations in at least 1 primer binding site. The percentage of such mutations in the different species was related to their phylogenetic distance from the cultigen (Fig. 1). It ranged from 92% for *C. reticulatum*, chickpea's presumable ancestor, to 50% for *C. cuneatum*. Based on these data a parsimony tree was constructed using *C. cuneatum* as an out-grouper. This decision is justified because this species, formerly grouped into the section *Monocicer* together with the other annual species (except *C. chorassanicum*, van der Maesen 1987), lacks two characteristics: the specific isozyme pattern (Kazan and Muehlbauer 1991) and the abundant satellite DNA families present in all other analysed species (Staginnus et al. 1999). We compared the resulting tree with the isozyme tree of Kazan and Muehlbauer (1991), and both were similar. The first crossability group, including the cultigen, *C. reticulatum and C. echinospermum*, is most closely related to the perennial *C. anatolicum* and most distantly related to *C. cuneatum*. The trees differ in that *C. judaicum* and *C. yamashitae* of the second crossability group change place. On the basis of our results and those obtained by Kazan and Muehlbauer (1991), we doubt that the present classification of *C. cuneatum* into section *Monocicer* is justified.

Grouping of C. anatolicum close to the first crossability group

Several lines of evidence suggest that *C. anatolicum*, though the only perennial species, should be grouped into section *Monocicer*. Moreover, of all the known *Cicer* species it is most closely related to the members of the first crossability group. First, its isozyme patterns are so similar to those of the first group that Kazan and Muehlbauer (1991) regarded it as a potential ancestor of the species of this group. Secondly, it contains both satellite families already mentioned (Staginnus et al. 1999). Third, it possesses a high level of conserved, chickpea-derived STMSs. Both the isozyme and STMS-based parsimonous trees directly relate this species to the first group (Fig. 1). It is also the only species outside the first crossability group for which reliable bootstrap values could be obtained. This indicates that the loci amplified with chickpea-derived primers have been inherited from an ancestor also common to chickpea. Since valuable genes (e.g. *Ascochyta* blight resistance genes) have already been identified in this species (Nene and Reddy 1987), its close genetic relationship to the cultigen will allow their transfer into the cultigen – either by crossing or somatic hybridisation. Therefore, *C. anatolicum* as well as other perennial species of the genus *Cicer* will certainly attract more attention in future.

Synteny of STMS loci in the genus Cicer

One of the aims of this study was to assess the suitability of chickpea-derived STMS markers for synteny studies within the genus *Cicer* in order to exploit the extensive mapping information available from chickpea for the map-based cloning of agronomically valuable genes from wild *Cicer* species. The transfer of markers from one species to an other is possible if these markers detect the same alleles in both species as is the case, for instance, if cDNAs are used as restriction fragment length polymorphism (RFLP) markers. For chickpea STMS markers such information is not yet available. The similarity of the parsimony tree presented here to relationships inferred from isozyme and other data is promising in this respect because it shows that many loci within the section *Monocicer* are inherited from common ancestors. This is especially true for the members of the first crossability group (high bootstrap values). Nevertheless, two lines of evidence suggest that STMSs from chickpea are only of limited use as syntenic markers beyond the first group. First, reliable bootstrap values could only be obtained from species of the first crossability group. Second, the sequences of amplicons derived from species of other crossability groups differed so much from the chickpea sequences that they most probably represent different loci. Only the amplicons from 2 accessions of *C. reticulatum* were homologous enough to assume that they were alleles of the *C. arietinum* locus.

What could be the reason that heterologous STMS primers produced amplification products even though they obviously amplified different loci? One explanation is that most of the microsatellite loci amplified here reside in phylogenetically old repetitive elements that are present in all *Cicer* species. More or less random mutations in these elements would then lead to the generation of new primer-binding sides at new loci, a hypothesis that is supported by the sequences of the amplification fragments from the wild species. There, only the fragments derived from *C. reticulatum* were homologous to the Ta14 locus from *C. arietinum*, contained a microsatellite and thus most probably represent the same locus. The fragments derived from *C. cuneatum* locus Ta37 and *C. judaicum* locus Ta18 did not contain the expected microsatellite and were unrelated to the respective sequences from *C. arietinum*. Consequently, STMS markers derived from chickpea will reliably detect syntenic loci only in the first crossability group. Between members of the second crossability group they may detect polymorphisms, but care must be taken to prove them to be alleles of chickpea loci.

Comparison of STMS loci from chickpea and wild species

Two types of STMS polymorphisms have been detected in wild *Cicer* species: (1) variation in the number and (2) variation in the size of bands. The differences in band

numbers can be explained essentially the same way as zero-alleles: either by the generation or loss of primer binding sites by mutations in already quite similar loci. However, the differences in fragment sizes between loci from different accessions of one species or between species deserve a closer inspection. Size differences in amplicons from different species could theoretically be caused by variable numbers of tandem repeats (VNTRs). However, this was the case only for the 475-bp expansion of the Ta14 allele from 249 bp in *C. arietinum* to 724 bp in *C. reticulatum* ILWC105. There, an extended, complex microsatellite intermingled with A/T-rich non-microsatellite sequences caused the large size variation, but other, non-VNTR mutations were also present. Also, the 25-bp difference between the 274-bp allele of *C. reticulatum* accession ILWC247 and the *C. arietinum* allele was only partly due to a VNTR mutation. The insertion of a $(CAA)_4$ was the only microsatellite-related variation, whereas the less-structured insertions, deletions and substitutions cannot easily be explained by slipped-strand mispairing thought to cause polymorphism at STMS loci (Levinson and Gutman 1987; Zischler et al. 1992). Similar complex mutations have been reported by Geistlinger et al. (1997) for a fungal STMS locus.

Evolution of the section Monocicer

Together with lentils, peas and the progenitors of wheat, chickpea is one of the most ancient crops. It provided the basis for the agricultural revolution that accompanied the conversion of human communities from hunters and gatherers to agriculture-based societies about 10,000 years ago in the so-called Fertile Crescent, comprising parts of present-day Turkey, Syria, Iraq and Iran. The relatively precise dating of the controlled growth of chickpea rests on findings of chickpea, pea and lentil together with Einkorn (*T. monococcum*) seeds, pottery and other remnants of the first farmers (Zohary and Hopf 1993). Based on known historical records and the data provided in this study, the minimum average mutation rates for microsatellite-flanking sequences of the chickpea genome and the minimal time for the evolution of the section *Monocicer* from a common ancestor can tentatively be estimated.

However, several assumptions are necessary:

1) Chickpea was selected from *C. reticulatum* by the first farmers, and the changes in morphology and chromosomal rearrangements, visible today, occurred at that time.
2) Mutations affect only one of the two primer binding sites for each primer.
3) The percentage of zero-alleles measured here results from mutations in microsatellite-flanking sequences (i.e. primer binding sites) and not from the loss of the entire locus.
4) Mutation rates of different primer binding sites are almost the same in different species.

In the present study, 8% zero-alleles were detected in *C. reticulatum* using primers designed to fit microsatellite-flanking sequences in *C. arietinum*. Since both flanking sequences of a microsatellite are necessary for successful amplification, the average mutation rate for a single primer site is 4% within 10,000 years or 4×10^{-6} per year. This is lower than the mutation rates estimated for microsatellite sequences themselves (Weber and Wong 1993). Accordingly, *C. echinospermum* (12% zero-alleles) would have evolved from a common ancestor with *C. reticulatum* around 30,000 years ago, and *C. anatolicum* (26% zero-alleles) around 65,000 years ago. The second crossability group with similar figures of zero-alleles (approx. 34–37%) would then have deviated from the ancestor it had in common with the members of the first group about 100,000 years ago. These two groups would have diverged from their common ancestor with *C. chorassanicum* (39% zero-alleles) a short time before (110,000 years ago). The line leading to *C. cuneatum* [largest number of zero-alleles (50%) of the species investigated here] would have deviated from a common ancestor of groups I, II and III 125,000 years ago. It may have happened at that time that the two satellite families characteristic of the section *Monocicer* evolved because *C. cuneatum* does not show even traces of these sequences on Southern blots (Staginnus *et al*. 1999).

Acknowledgements The research of the authors was supported by BMZ, Bonn (Grant 89.7860.3–01.130).

References

Abbo S, Miller TE, Reader SM, Dunford RP, King IP (1994) Ribosomal DNA sites in lentil and chickpea by fluorescent in situ hybridisation. Genome 37:713–716

Ahmad F, Slinkard AE (1992) Genetic relationship in the genus *Cicer* L. as revealed by polyacrylamide gel electrophoresis of seed storage proteins. Theor Appl Genet 84:688–692

Ahmad F, Gaur PM, Slinkard AE (1992) Isozyme polymorphism and phylogenetic interpretations in the genus *Cicer* L. Theor Appl Genet 83:620–627

Beckman JS, Soller M (1990) Toward a unified approach to genetic mapping of eukarotes based on sequence tagged microsatellite sites. Biotechnology 8:30–32

Diwan N, Bhagwat AA, Bauchan GB, Cregan PB (1997) Simple sequence repeat DNA markers in alfalfa and perennial and annual *Medicago* species. Genome 40:887–895

Echt CS, May-Marquardt P (1997) Survey of microsatellite DNA in pine. Genome 40:9–17

Forbes SH, Hogg JT, Buchanan FC, Crawford AM, Allendrof FW (1995) Microsatellite evolution in congeneric mammals: domestic and bighorn sheep. Mol Biol Evol 12:1106–1113

Garza JC, Slatkin M, Freimer NB (1995) Microsatellite allele frequencies in humans and chimpanzees, with implications for constraints on allele size. Mol Biol Evol 12:594–603

Gaur PM, Slinkard AE (1990a) Inheritance and linkage of isozyme coding genes in chickpea. J Hered 81:455–461

Gaur PM, Slinkard AE (1990b) Genetic control and linkage relations of additional isozyme markers in chickpea. Theor Appl Genet 80:648–656

Geistlinger J, Weising K, Kaiser WJ, Kahl G (1997) Allelic variation at a hypervariable compound microsatellite locus in the ascomycete *Ascochyta rabiei*. Mol Gen Genet 256:298–305

Huettel B, Winter P, Weising K, Choumane W, Weigand F, Kahl G (1999) Sequence-tagged microsatellite site markers for chickpea (*Cicer arietinum* L.). Genome 42:1–8

Karhu A, Hurme P, Karjalainen M, Karvonen P, Karkkainen K, Neale DB, Savolainen O (1996) Molecular markers reflect patterns of differentiation in adaptive traits in conifers. Theor Appl Genet 93:215–221

Katzir N, Danin-Poleg Y, Tzuri G, Karchi Z, Lavi U, Cregan PB (1996) Length polymorphism and homologies of microsatellites in several *Cucurbitaceae* species. Theor Appl Genet 93:1282–1290

Kazan K, Muehlbauer FJ (1991) Allozyme variation and phylogeny in annual species of *Cicer* (*Leguminosae*). Plant Syst Evol 175:11–21

Kresovich S, Szewc-McFadden AK, Bliek SM (1995) Abundance and characterization of simple-sequence repeats (SSRs) isolated from a size-fractionated genomic library of *Brassica napus* L. (rapeseed). Theor Appl Genet 91:206–211

Labdi M, Robertson LD, Singh KB, Charrier A (1996) Genetic diversity and phylogenetic relationships among the annual *Cicer* species as revealed by isozyme polymorphism. Euphytica 88:181–188

Ladizinsky G, Adler A (1976) Genetic relationships among the annual species of *Cicer*. Theor Appl Genet 48:197–203

Levinson G, Gutman GA (1987) Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol Biol Evol 4:203–221

Litt M, Luty JA (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am J Hum Genet 44:397–401

Moore SS, Sargeant LL, King TJ, Mattick JS, Georges M, Hetzel DJS (1991) The conservation of dinucleotide microsatellites among mammalian genomes allows the use of heterologous PCR primer pairs in closely related species. Genomics 10:654–660

Morgante M, Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. Plant J 3:175–182

Nakamura Y, Leppert M, O'Connel P, Wolff R, Holm T, Culver M, Martin C, Fujimoto E, Hoff M, Kumlin E, White R (1987) Variable number of tandem repeat (VNTR) markers for human gene mapping. Science 235:1616–1622

Ocampo B, Venora G, Errico A, Singh KB, Saccardo F (1992) Karyotype analysis in the genus *Cicer*. J Genet Breed 46:229–240

Pepin L, Amigues Y, Lepingle A, Berthier J-L, Bensaids A, Vaiman D (1995) Sequence conservation of microsatellites between *Bos taurus* (cattle), *Capra hircus* (goat) and related species. Example for use in parentage testing and phylogeny analysis. Heredity 74:53–61

Nene YL, Reddy MV (1987) Chickpea diseases and their control. In: Saxena MC, Sing KB (eds) The chickpea CAB Int, Wallington, UK, pp 230–270

Röder MS, Plaschke J, Konig SU, Borner A, Sorrells ME, Tanksley SD, Ganal MW (1995) Abundance, variability and chromosomal location of microsatellites in wheat. Mol Gen Genet 246:327–333

Rogers SO, Bendich AJ (1985) Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. Plant Mol Biol 5:69–76

Roy MS, Geffen E, Smith D, Ostrander EA, Wayne RK (1994) Patterns of differentiation and hybridization in North American wolflike canids, revealed by analysis of microsatellite loci. Mol Biol Evol 11:553–570

Rubinstein DC, Amos W, Leggo J, Goodburn S, Jain S, Li SH, Margolis RL, Ross CA, Ferguson-Smith MA (1995) Microsatellite evolution – evidence for directionality and variation in rate between species. Nat Genet 10:337–343

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, New York

Schlötterer C, Tautz D (1991) Slippage synthesis of simple sequence DNA. Nucleic Acids Res 20:211–215

Singh KB, Ocampo B (1997) Exploitation of wild *Cicer* species for yield improvement in chickpea. Theor Appl Genet 95:418–423

Smith DN, Devery ME (1995) Occurrence and inheritance of microsatellites in *Pinus radiata*. Genome 37:977–983

Staginnus C, Winter P, Desel C, Schmidt T, Kahl G (1999) molecular structure and chromosomal localization of major repetitive DNA families in the chickpea (*Cicer arietinum* L.) genome. Plant Mol Biol 39:1037–1050

Thomas MR, Scott NS (1993) Microsatellites in grapevine reveal DNA polymorphisms when analysed as sequence-tagged sites (STSs). Theor Appl Genet 86:985–990

Tuwafe S, Kahler AL, Boe A, Ferguson M (1988) Inheritance and geographical distribution of allozyme polymorphisms in chickpea (*Cicer arietinum*). Heredity 79:170–174

van der Maesen LJG (1987) Origin, history and taxonomy of chickpea. In: Saxena MC, Sing KB (eds) The chickpea CAB Int, Wallington, UK, pp 11–34

Weber JL, Wong C (1993) Mutation of human short tandem repeats. Hum Mol Genet 2:1123–1128

Winter P, Kahl G (1995) Molecular marker technologies for plant improvement. World J Microbiol Biotechnol 11:438–448

Winter P, Pfaff T, Udupa SM, Hüttel B, Sharma PC, Sahi S, Arreguin-Espinoza R, Weigand F, Muehlbauer FJ, Kahl G (1999) Characterization and mapping of sequence tagged microsatellite sites in the chickpea (*Cicer arietinum* L.) genome. Mol Gen Genet 262:90–101

Zischler H, Kammerbauer C, Studer R, Grzeschik KH, Epplen JT (1992) Dissecting (CAC)5/(GTG)5 multilocus fingerprints from man into individual locus-specific, hypervariable components. Genomics 13:983–990

Zohary D, Hopf M (1993) Domestication of plants in the old world. Clarendon Press, Oxford, UK