Theor Appl Genet (1998) 97 : 283–292

M. Falque · J. Keurentjes · J. M. T. Bakx-Schotman
P. J. van Dijk

Development and characterization of microsatellite markers in the sexual-apomictic complex *Taraxacum officinale* (dandelion)

Received: 4 November 1997 / Accepted: 12 February 1998

Abstract Microsatellite markers were developed in *Taraxacum officinale* to study gene flow between sexual and apomictic plants and to identify clones. Twenty five thousand genomic DNA clones were hybridized with a $(CT)_{12}D$ probe. The density of $(GA/CT)_n$ repeats was estimated at one every 61 kb in the *T. officinale* genome, which translates to 13 500 repeats per haploid genome. Ninety two percent of 110 positive clones sequenced contained at least one $(GA/CT)_{n \geq 5}$ repeat. Sixteen $(CA/GT)_{n \geq 5}$ and 11 $(AT)_{n \geq 5}$ arrays were also found in these sequences, suggesting some clustering of dinucleotide repeats. Among 50 PCR primer pairs tested, 32 produced bands and 28 of them were polymorphic. Of these polymorphic markers, 15 were putatively single-locus and the other 13 produced only polymorphic fingerprints. Six loci were further characterized for polymorphism and showed between 6 and 32 alleles per locus. Among eight primer pairs used to analyze the progeny of a sexual cross, seven were co-dominant single-locus Mendelian markers, but one (MSTA10) gave a dominant pattern in accordance with the hypothesis of a null allele segregating in a Mendelian fashion. Three pairs of loci among 28 showed significant linkages of 10, 21, and 39 cM. Observed and expected heterozygosities in two sexual populations indicate that null alleles may be present at two loci, including MSTA10.

Key words Microsatellites · Null alleles · Apomixis · Silver staining · *Taraxacum*

Publication 2301, Netherlands Institute of Ecology, Center for Terrestrial Ecology, Heteren, The Netherlands

Communicated by M. A. Saghai-Maroof
M. Falque (✉)[1] · J. Keurentjes · J. M. T. Bakx-Schotman
P. J. van Dijk
Center for Terrestrial Ecology, Netherlands Institute of Ecology (NIOO-CTO)
P.O. Box 40, NL-6666ZG Heteren, The Netherlands

[1] *Present address*:
INRA-UPS-CNRS-URA 1492
Station de génétique végétale
Ferme du Moulon, F-91190 GiF-sur-Yvette,
France
Fax: + 33 (0)1 69 33 23 40
e-mail: falque@moulon.inra.fr

Introduction

Apomixis is a mode of reproduction found in some plant species in which seeds are produced asexually (Gustafsson 1946). Since the apomictic seeds are genetically identical to the mother plant, apomicts form clonal populations. However, most apomictic plants are hermaphrodite and produce reduced pollen. Therefore, they may hybridize with sexual taxa, forming a sexual-apomictic complex. To unravel the genetic structure of such a complex, and to study its dynamics, suitable model species and genetic markers are essential.

Both sexual and asexual modes of reproduction are found in *Taraxacum officinale* (dandelion). Sexual plants are diploid ($2n = 2x = 16$), most often self-incompatible, and insect-pollinated, whereas asexuals are mainly triploid and reproduce through diplosporic gametophytic apomixis of the *Taraxacum* type (Asker and Jerling 1992). Both sexuality and asexuality are obligate. Apomicts can produce seeds without any pollination, and these seeds are comparable in morphology and dispersal with those produced by sexuals. Apomicts are found either in purely clonal populations or together with sexual diploids in mixed populations where gene flow can occur between both types (Morita et al. 1990; Menken et al. 1995; Van Dijk et al., in preparation; Tas et al., in preparation). *T. officinale* is thus an outstanding model species to study the ecology and evolution of sexual and asexual reproduction.

Suitable genetic markers are required to answer three groups of questions concerning sexuality and

asexuality in the *T. officinale* complex. First, it is necessary to identify clones accurately in order to address ecological questions about clonal diversity, how different clones co-exist, whether they show local adaptation, and whether they differ in plasticity. Morphological traits used for such studies (Vavrek et al. 1996) are difficult to score, and hyper-polymorphic molecular markers would greatly improve the resolution in these studies. Second, single-locus markers are needed to study the dynamics of genetic diversity in mixed populations – including gene flow between sexuals and apomicts – for instance to infer gene flow from rare alleles (Slatkin 1985) or from allele frequencies using *F*-statistics. Third, single-locus highly polymorphic markers can be used to study the genetic origin of the offspring produced in experimental crosses between sexuals and apomicts.

Single-locus co-dominant allozyme markers have been used in *T. officinale* for population genetic studies (Menken et al. 1995), but the three polymorphic loci that can be scored consistently are only moderately variable. DNA fingerprints based on $(GATA)_4$-probed (Van Heusden et al. 1991) or rDNA- and cpDNA-probed (King 1993) RFLPs have been used to identify clones, but they do not provide single-locus data and the analysis requires a considerable amount of high-quality DNA which is difficult to extract from *T. officinale*. RAPDs have been tested but were neither very variable nor reproducible (Van Dijk and Van der Hulst, unpublished) and in any case are multilocus dominant markers. Recently, IGS-rDNA PCR-based RFLP analysis (Van Dijk et al., in preparation) and AFLPs (Van der Hulst et al., in preparation) have shown a high resolution for clonal identification, but again do not provide single-locus data.

Microsatellite markers provide interesting features matching our requirements. They are based on tandemly repeated short DNA motifs (1–5 bp) found in a wide range of eukaryotic genomes (Tautz and Renz 1984). These sequences often show a very high level of variation in the number of tandem repeats, and this variation is probably due to slipped-strand mispairing during DNA replication (Levinson and Gutman 1987). Microsatellite polymorphism can be analyzed by PCR amplification with specific primers and high-resolution electrophoresis, which gives hyper-variable single-locus co-dominant Mendelian markers (Litt and Luty 1989). In addition, microsatellites are usually considered to be selectively neutral markers (Jarne and Lagoda 1996), which is important for population studies – although one important exception involves microsatellites associated with human genetic diseases (Sutherlands and Richards 1995). Some microsatellite-based population genetic studies have been reviewed and discussed by Bruford and Wayne (1993) and by Jarne and Lagoda (1996).

In order to study the dynamics of genetic diversity in a sexual-asexual complex, microsatellite isolation has already been carried out in the flat worm *Dugesia polychroa* (Ramachandran et al. 1997), which is a model species very close to *T. officinale* with respect to its system of reproduction. Similarly, microsatellite development in *T. officinale* is expected to provide hyper-polymorphic Mendelian single-locus co-dominant neutral markers, which would be powerful tools to study questions related to the ecology and evolution of sexual and asexual reproduction. This paper reports on the density of microsatellites in the genome of *T. officinale*, their repeat structure, the polymorphism and linkage relationships of loci, and the inheritance of alleles.

Materials and methods

Library construction

A young flower bud of a triploid plant was dissected to remove sepals and all latex-rich green parts, and the flowerlets were used for CTAB-based DNA extraction according to Rogstad (1992). After RNAse treatment, 10 μg of DNA was digested with 20 units of *Mbo*I for 3 h at 37°C and purified in four QIAquick columns (QIAquick PCR Purification Kit, Qiagen). The eluate was electrophoresed on a 1.2% (w/v) agarose gel (1.25 μg per lane) with 0.1 μg/ml of ethidium bromide. The gel was then cut between 450 and 1000 bp under UV light, and the DNA was extracted by using the QIAEXII Gel Extraction Kit (Qiagen), measured by spectrophotometry, and further purified and concentrated with a QIAquick procedure to a final concentration of 25 ng/μl.

Ligation was performed in a total volume of 20 μl, with 200 ng (56 fmol) of pUC19 vector cut with *Bam*HI and de-phosphorylated (Appligene), 275 ng (about 600 fmol) of insert DNA and 5 units of T4 ligase (Promega), and was incubated overnight at 16°C. The ligation product was purified by using the QIAquick procedure with elution in 50 μl of ultra-pure water. This eluate was either kept at − 20°C or directly used to transform MAX Efficiency DH5α Competent Cells (GibcoBRL Life Technologies) according to the manufacturer's instructions, except that the DNA was added to the transformation tubes without dilution. The bacteria were spread in large plates (135 mm in diameter; 200 μl/plate) containing 70 ml of LB medium with 0.01% (w/v) ampicillin, 0.0032% (w/v) X-Gal, and 0.0048% (w/v) IPTG.

Library screening for microsatellites

The plates were incubated overnight at 37°C, until the diameter of the colonies was about 2 mm. The colonies were then lifted on Hybond N + membrane discs (gridded, 132 mm in diameter, Amersham), previously marked with a HB graphite pencil, and the position of the membranes was marked on the plates. The membranes were piled up for 10 min with filter paper soaked with 0.4 N NaOH in between, removed, and piled up again with filter paper soaked with 2 × SSPE in between, gently rubbed with a gloved hand to remove the colony debris from the surface, piled up in a tray with fresh 2 × SSPE for 10 min on a shaker, air-dried colony side up on filter paper, and finally cross-linked with 120 MJ of UV light.

The plates were placed at 4°C immediately after colony lift when some traces of the colonies were still visible on the agar, or first incubated at 37°C for 2 h when the colonies seemed to be completely removed. Up to 40 membranes were then soaked in 2 × SSPE and piled up with meshes in between in a plastic box with 400 ml of pre-hybridisation solution [5 × SSPE, 5 × Denhart, 1% blocking reagent (Boehringer), 0.1% SDS], and the sealed box was incubated at 57°C for 4 h on a shaker. Eight-hundred picomoles of HPLC-purified 5′-digoxygenine-labelled $(CT)_{12}D$ oligonucleotide probe (Isogen Bioscience, Inc.) were then heated for 5 min at 95°C and slowly added to the pre-hybridisation solution. Hybridisation was

carried out overnight at 57°C, and the membranes were washed three times for 30 min with 400 ml of 6 × SSC at room temperature, and once for 7 min with 400 ml of 6 × SSC at 60°C. Antibody mediated CSPD (Tropix) chemiluminescence detection of the probe was then performed according to the manufacturer's instructions (Boehringer) and the films were exposed for 15 h at room temperature. The positive colonies were located in the plates by using the marks and the grid visible on the background of the film. They were then picked up for overnight culture and plasmid extraction by using the QIAprep Spin Miniprep Kit (Qiagen). In addition, all positive colonies, as well as 96 randomly picked-up white colonies, were used as templates for standard PCR amplification with M13 universal primers, to size the products on agarose gels. The size of the inserts of the positive clones was also confirmed on agarose gels after double restriction of the plasmids with *Eco*RI and *Pst*I. Automatic sequencing of the inserts was performed by a subcontractor (Genome Express, Inc.). After editing and alignment of the sequences using the program Sequence Navigator 1.0 (Applied Biosystems), PCR primers were designed with the help of the OLIGO 5.0 program (National Biosciences).

Microsatellite genotyping

DNA extractions were performed using either bud tissue and CTAB extraction, as described above, or by punching and grinding two leaf discs in an Eppendorf tube with 200 µl of 0.5 N NaOH and diluting 5 µl of this extract with 495 µl of 100 mM Tris-HCl pH 8.0, according to Wang et al. (1993). This very quick NaOH method gave somewhat weaker bands than the CTAB extraction, but was sufficient for routine typing. Ten nanograms of CTAB-extracted or 2 µl of diluted NaOH-extracted DNA were used as a template for 25 µl PCR reactions with 2.5 mM of $MgCl_2$, 200 nM of each primer, 200 µM of each dNTP, 0.5 units of Expand™ High Fidelity polymerase (Boehringer), and one drop of mineral oil. The PCR program was 95°C for 2 min, 40 cycles of (95°C, annealing temperature estimated with OLIGO, and 72°C, 1 min each), and 72°C for 4 min. When the bands were too weak, the PCR products were concentrated before electrophoresis by adding 75 µl of 96% (v/v) ethanol, incubating for 10 min at −80°C, centrifuging for 10 min at 13 000 rpm, air-drying the pellet, and re-dissolving it with 6 µl of TE.

The samples were added to an equal volume of denaturing loading buffer (12.5% (w/v) sucrose, 0.05% (w/v) xylene-cyanol blue, 0.05% (w/v) bromo-phenol blue, 0.1 N NaOH, 95% (v/v) formamide), denatured for 2 min at 95°C just before loading the gel, and kept on ice while loading. Electrophoresis was performed under denaturing conditions in 6% urea-polyacrylamide gels (Ultra Pure Sequagel-6, Biozym). Sequencing gels (40 × 40 cm, 0.4-mm thick) were pre-run at 80 W constants to reach 50°C, and 4 µl of samples were loaded per well. The gels were loaded up to three times separated by 800 V.h of migration at 45 W constants, and followed by a last migration step until the first xylene-cyanol blue dye reached the bottom of the gel. The gels were then silver-stained with a fully automatic gel stainer (NIOO-CTO, Heteren, The Netherlands) by using the protocol of the DNA Silver Staining System (Promega) with half of the quantities of chemicals advised for manual staining. The gels were dried for 30 min at 70°C and permanent copies were produced using Silver Sequence APC film (Promega).

Results and discussion

Cloning results

Among 25 000 colonies screened with the $(CT)_{12}D$ probe, 158 positive clones were found. The mean insert size of the library was 385 bp. This corresponds to one $(GA/CT)_n$ repeat every 61 kb in the genome. We estimated the genome size of *T. officinale* to be 820 megabases per haploid genome (PARTEC ploidy analyzer with DAPI staining), which indicates that about 13 500 $(GA/CT)_n$ repeats may be present in the *T. officinale* haploid genome, of which we have screened about 1.2% with our library. However, our estimate of the repeat density may be biased by the cloning method (Jarne and Lagoda 1996). Indeed, Condit and Hubbell (1991) found up to 4-fold differences in microsatellite densities between libraries of the same species built with different restriction enzymes. In our case, we may have overestimated the density because the recognition sequence of *Mbo*I (GATC) is more likely to be found in $(GA/CT)_n$-rich regions. Indeed, most of the similar cloning-based studies in plants indicate lower densities, with distances between two adjacent $(GA/CT)_n$ repeats ranging between e.g. 100 kb for rapeseed (Kresovich et al. 1995) to 1176 kb for tomato (Broun and Tanksley 1996). However, this value in the tropical tree *Virola sebifera* was only 17 kb (Condit and Hubbell 1991).

Sequencing results

Sequences were obtained from 116 clones. We found six times two clones giving the same sequence, so that we ended up with 110 distinct sequences. In 101 sequences at least one perfect $(GA/CT)_{n \geq 5}$ array was present. Fourteen of these microsatellites were immediately followed by another type of dinucleotide array of five or more repeats, forming six $(GA/CT)_n(CA/GT)_m$, five $(GA/CT)_n(AT)_m$, two $(CA/GT)_n(GA/CT)_m(CA/GT)_p$, and one $(GA/CT)_n(AT)_m(GA/CT)_p$ compound microsatellites. Additional $(GA/CT)_{n \geq 5}$, $(CA/GT)_{n \geq 5}$ or $(AT)_{n \geq 5}$ arrays separated from the first $(GA/CT)_{n \geq 5}$ were also found in respectively 13, six and five sequences, but no $(GC)_{n \geq 5}$ array was observed. The frequency of at least one $(CA/GT)_{n \geq 5}$ or $(AT)_{n \geq 5}$ repeat being found in clones already containing a $(GA/CT)_{n \geq 5}$ microsatellite seems much too high (25%) to reflect only the genomic density of $(CA/GT)_n$ or $(AT)_n$ repeats, although $(AT)_n$ microsatellites are more frequent than $(GA/CT)_n$ ones in plant sequences from databases (Morgante and Olivieri 1993; Wang et al. 1994). These $(CA/GT)_n$ or $(AT)_n$ repeats were probably not selected by our hybridization, because they were never found alone in positive clones. Thus, there seems to be some association between the different types of dinucleotide repeats in the *T. officinale* genome, as has been reported before in tropical tree species (Condit and Hubbell 1991). The combination of point mutations and slipped-strand mispairing may cause the evolution of pure microsatellites towards more complex repeat regions including compound arrays (Wright 1994), which could explain such clustering.

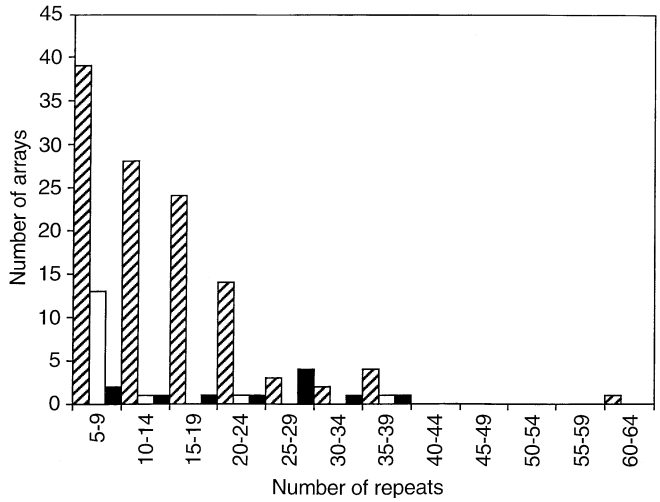

Fig. 1 Size distribution of perfect dinucleotide repeats in 101 genomic clones that hybridized to the $(CT)_{12}D$ probe. *Dashed bars* $(GA/CT)_n$; *empty bars* $(CA/GT)_n$; *solid bars* $(AT)_n$ repeats.

In our sequences, the different types of dinucleotide microsatellites showed length distributions significantly different from each other (Fig. 1, Kolmogorov-Smirnov test $P < 0.05$), and the average length of the $(GA/CT)_{n \geq 5}$, $(CA/GT)_{n \geq 5}$, and $(AT)_{n \geq 5}$ arrays were respectively 14.5, 10.1, and 22.3 perfect repeats, or 17.1, 11.7, and 22.3 repeats when allowing for one single base-pair imperfection. This is consistent with the findings in plant sequence databases that $(AT)_n$ arrays reach longer sizes than $(GA/CT)_n$ ones (Morgante and Olivieri 1993; Wang et al. 1994). Our sequences also showed 20 $(A/T)_{n \geq 10}$ and two $(G/C)_{n \geq 10}$ mononucleotide, as well as one $(TCC/AGG)_5$, and one $(TGG/ACC)_5$ trinucleotide repeats.

PCR primer testing

Fifty primer pairs were designed and tested for PCR amplification and polymorphism on six different triploid clones from a purely clonal meadow in central Netherlands. Eighteen gave smears or no amplification and 32 gave the banding patterns described in Table 1. Only four primer pairs were monomorphic, but some highly polymorphic markers produced more than three bands, or else intensively "stuttering" bands, making scoring unreliable. These markers can still be useful as DNA fingerprints for clonal identification. Indeed, MSTA1 gave identical complex patterns for 23 offspring of an apomictic mother plant, while it is extremely polymorphic, allowing the distinction of 27 different genotypes in a sample of 64 apomictic plants from a single meadow (data not shown). Such complex patterns have also been reported in other plant species, e.g. apple (Guilford et al. 1997). On our six triploid plants, 15 polymorphic markers produced three or fewer bands which could be sized with a 1-bp accuracy. Because the putative alleles have been checked for single-locus Mendelian inheritance for some of these markers (see below), they are hereafter referred to as alleles, and the corresponding markers as loci. Three of those polymorphic markers gave clear bands without a "stuttering" pattern (Fig. 2 a, Table 1), whereas in the others, the bands were followed by typical less intense "stuttering" bands (Fig. 2 b), usually 2- and 4-bp smaller than the main band, and sometimes also 2-bp larger. We found no effect of magnesium concentration (1.5 or 2.5 mM), the amount of enzyme (0.5 or 0.75 units), or the number of PCR cycles (25 or 40) on this phenomenon, which is consistent with previous results on $(CA/GT)_n$ microsatellites (Hite et al. 1996). On the other hand, the "non-stuttering" microsatellites were on average more complex than the other ones, and their mean repeat length was significantly lower (Table 1), suggesting that the pure and/or long repeats might be the most prone to slipped-strand mispairing during PCR amplification.

Polymorphism

The polymorphism at six loci was studied in more detail on 40 diploid sexuals and 18 triploid apomicts originating respectively from three different mixed populations and one purely asexual population of central Netherlands. Large differences in allele diversity were observed between loci (Table 2), which is a common finding with microsatellites, e.g. in maize (Taramino and Tingey 1996). In partially heterozygous triploid plants showing only two bands (e.g. Fig. 2 b sample 10), it was usually not possible to infer the exact genotype from the relative intensity of the bands. Indeed, differences in intensity between alleles were frequent even in three-banded patterns, and would mask any differences in intensity due to a gene-dosage effect. The distributions of allele frequencies at each locus (Fig. 3) show that most of the allele lengths at a given locus differ by even numbers of base pairs, which is expected in the case of mutation involving changes in the number of dinucleotide repeats. However, MSTA78 and particularly MSTA72, which are perfect dinucleotide microsatellites, show single base-pair differences between alleles (Fig. 3; e.g. in Fig. 2 b between samples 6 and 7, 18 and 19, or 19 and 20). This indicates that other mutation mechanisms have operated, producing insertions or deletions of an odd number of base pairs.

Null alleles

In the study of polymorphism (134 haploid genomes), four non-amplifying samples (two diploids and two triploids) were observed with the microsatellite

Table 1 Microsatellite primers tested on six triploid clones: description of the markers that gave banding patterns

Marker	Repeat	Primer sequences		T_a[a]	EL[b]	POL[c]	PC[d]
MSTA1	$(CT)_{18}$	AAAGTCCATCCACCTAATTC	TGAAACTTCCTCATCTTCTG	50	190	Y	2–3
MSTA10	$(TC)_9TTCT(TC)_5C(TC)_3CC(TC)_2T$	TATGTGCTTGTGCTACAGTGT	TGGGTTTTCTTCCTCTTTTA	51	205	Y	1
MSTA11	$(CT)_{14}$	GATCTAAATCTCCTCTCT	TGTCACCTGCCTACT	48	200	Y	2–3
MSTA16	$(CT)_{10}$–$(CT)_8$–$(CT)_7$	CTTGACCAATCGTTTAGAATG	TTTTAATACCCAATGTCTTTTGA	48	227	Y	3
MSTA20B	$TTTCTTT(CTT)_4$	GATGAATGTCAACAGGAGTA	AATTAGCTTAATCCAATTCTATTA	49	172	N	1
MSTA21	$(CT)_{10}$	TTGGTACTGAGATGTTTGAAC	CCACCAACCTTTTGAAC	51	277	Y	3
MSTA23	$(CT)_{10}$–$(CT)_{13}$	GGTAAAGAAAAATAGCAAGAAT	AGAGAGGAAGGGTTTGG	51	312	Y	3
MSTA29A	$(CT)_{20}$	GATCAAAGATCTAGCTAGAGAG	GCCATCATATTTCGTGTT	49	98	Y	2–3
MSTA31	$(CT)_{17}$	CCTCAAAGCCCGAACTT	ACGACCCCAACTGATTTTTAC	51	240	Y	2
MSTA41A	$(TGG)_2C(TGG)_5$	CTGTTCGGAGGGTGGAAA	TCGCACGAGTTGATTATGGT	57	132	N	1
MSTA41B	$(CT)_{11}$	CAGCCGAAATTCTTGAAAAT	GGGTCACACCGTCAGTAAC	51	197	Y	2–3
MSTA42	$(CT)_3CC(CT)_6$	CCACCCTCTCCAACCGTC	GATCTAAGCCTTTGATTTGCTCA	54	279	N	2
MSTA44B	$(CT)_{19}$	AGTTTCTCTAAAATGGGAAGAT	TGTCAGGTATATTCAAAAGATTC	51	191	Y	2
MSTA45	$(CT)_{23}$	TTTAATGCTGGGAAACCTCT	CTACATCCACGGGCAAAC	51	233	Y	3
MSTA53	$(TC)_{12}(GT)_8$	CAATTATTATGGTCTCGTCCTT	CCAGTTGAAGCAAAAACAGT	55	203	Y	2
MSTA54	$(CT)_{23}$	GTAACGGTCATGAAAGTGGC	TCTTCTAATGAGGCAGTATGGC	56	251	Y	2–3
MSTA55	(CT)13	GCCCGTTCTTCCACAAT	AACAACAGATTCCGTTTTTGAA	55	274	Y	3
MSTA58	$(CT)_{17}$	CCACTTTGAACGACATTGTTA	TCTCTGATTCAACCTGTGTTTC	53	125	Y	2
MSTA59	$(CT)_{24}$	GCTGGGAAACCTCTGCTGAAAT	ACATCCACGGGCAAACGA	57	226	N	2
MSTA60	$(CT)_9$–$(CT)_{10}$	TTTGCCAAATAAATGCTACA	CGTTGGAATGGATAAGAAAG	54	317	Y	1
MSTA61	$(CT)_9(GT)_8$	CTACGCTCTTCCCCATTGAT	TGGTGGTTGATGCTCTGTTT	55	126	Y	2
MSTA64	$(TC)_4TT(TC)_2TT(TC)_2TT(TC)_5A(TC)_4$–$(A)_{16}$	TGCTTTTTGAACGACAGTG	TTTGCTTGGTTATTAGTGAACAT	55	191	Y	1
MSTA67	$(TC)_{22}T(CA)_{12}$	TTCGGATATGACCCTTCACT	GACATCTTGCACCTAAAAACAAT	56	219	Y	2
MSTA68	$(TA)_{28}$	TTCTTGTTTAGCCCAATCTA	CCATAGTGGATGCTCTTATC	48	185	Y	2
MSTA72	$(CT)_{18}$	GATGAATTTGTCTTTGCT	CTATATCATTTTAGGTGTCGT	49	181	Y	2
MSTA73	$(TC)_{21}CTG(TC)_8$	CCGCATGAGGTTGTCT	TGGGCTGTTTAATAGAACTTA	53	216	Y	2
MSTA74	$(CT)_{10}$	GAGGTCTTTTTATTCGGTTTT	GGATGCCTTTACAGTTACAAT	49	223	Y	2
MSTA78	$(CT)_9$	TGATTGATTCTGCCCTAAACC	TGCCAAGACATCCGAAAAG	52	151	Y	2
MSTA81	$(CT)_{16}$	TTGAATCCGAATGCTTGTG	CCAGCAAAAATCAAAGTCGTA	54	235	Y	2–3
MSTA83A	$(TA)_{24}$	GATCCGAACCGTGATAT	GATTAGGGTTTCCGATTT	47	162	Y	3
MSTA83B	$(TC)_{18}A(CT)_5$	CGCCTCCATCTCACTCAAACT	ACAAAGGAACGCGAAGGTG	58	177	Y	2–3
MSTA85	$(CT)_{20}$	TGCATGTTCGTTCTACTGGT	ACGTAATAAAATTGGAAGTCAGG	55	196	Y	2

[a] T_a: PCR annealing temperature
[b] EL: expected product length
[c] POL: presence (Y) or absence (N) of polymorphism
[d] PC: pattern clearness within the expected range of size: three or fewer clear bands with little or no "stuttering" (1), three or fewer bands with "stuttering", but main bands distinguishable (2), more than three bands or intense "stuttering" pattern making scoring unreliable (3)

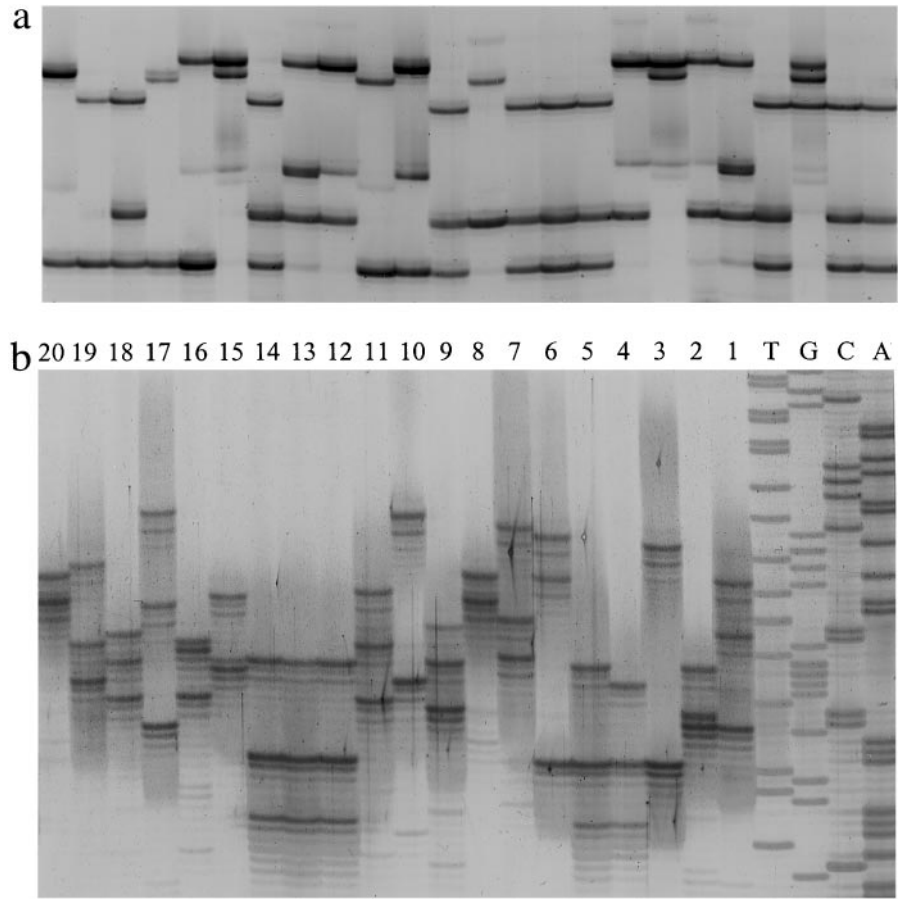

Fig. 2a, b Microsatellite genotyping of triploid *T. officinale* apomicts from a single meadow. **a** locus MSTA64. **b** locus MSTA72. A, C, G, T: sequence reaction of the pUC19 plasmid used as a molecular-weight marker. Sequencing gels with automated silver staining

Table 2 Microsatellite polymorphism and heterozygote deficiency. Populations 1 and 2 include only sexuals from two mixed populations

Locus	A[a]	NA[a]	Population 1 ($n = 12$)			Population 2 ($n = 18$)		
			NHobs[b]	NHexp[c]	H-W[d]	NHobs	NHexp	H-W
MSAT10	11	4	1	7.0	0.00	4	13.3	0.00
MSTA31	14	0	8	9.6	0.22	10	12.4	0.20
MSTA61	18	0	12	10.6	1.00	13	13.7	0.28
MSTA64	6	0	4	5.9	0.13	2	5.4	0.01
MSTA72	32	0	12	11.2	1.00	16	16.7	0.34
MSTA78	16	0	9	8.9	0.69	18	14.4	1.00

[a]A: number of different alleles observed, and NA: number of individuals giving no amplification, in 40 diploid sexuals and 18 triploid apomicts from central Netherlands
[b]NHobs: number of heterozygous genotypes observed
[c]NHexp: number of heterozygous genotypes expected from the observed allele frequencies, assuming Hardy-Weinberg equilibrium
[d]H-W: probability of the exact test for Hardy-Weinberg equilibrium, considering the alternative hypothesis of a heterozygote deficiency (Rousset and Raymond 1995), computed with GENEPOP3 (Raymond and Rousset 1995).

MSTA10, but none with the other loci (Table 2). To investigate the possibility of null alleles, the observed numbers of heterozygotes were compared with the numbers expected under Hardy-Weinberg equilibrium in two small samples of sexual diploid plants (Table 2). No significant deviations from Hardy-Weinberg equilibrium were observed for MSTA31, MSTA61, MSTA72, and MSTA78, but in spite of the small sample sizes, a significant deficit of heterozygotes was observed for MSTA10 in both populations, and for MSTA64 in population 2 (Table 2). Null alleles are one of the factors that may explain these deficits. For MSTA10, the presence of a null allele was confirmed in the progeny of a controlled cross ($n = 92$, see segrega-

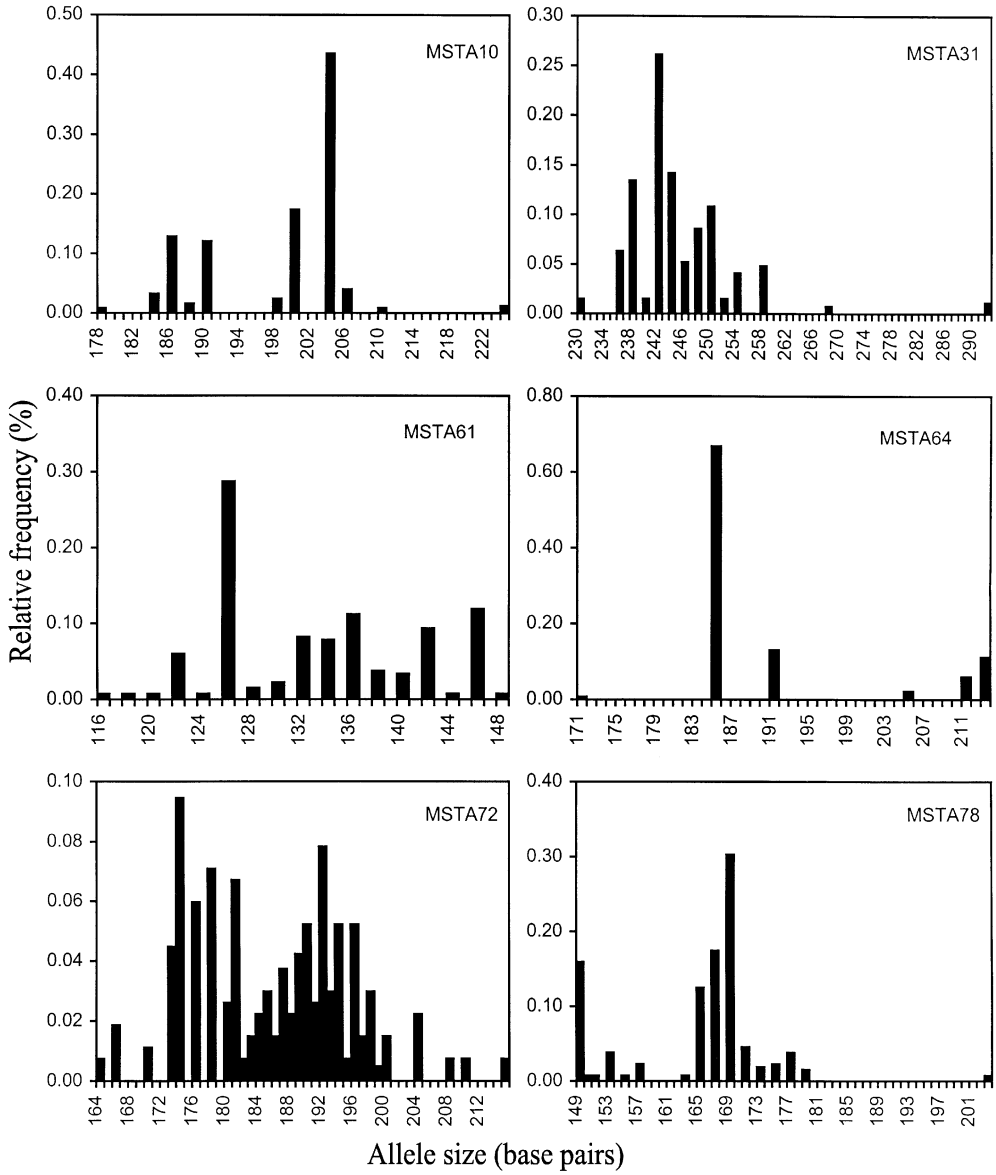

Fig. 3 Distributions of allele frequencies at six microsatellite loci, analyzed in 40 diploid sexuals and 18 triploid apomicts from central Netherlands

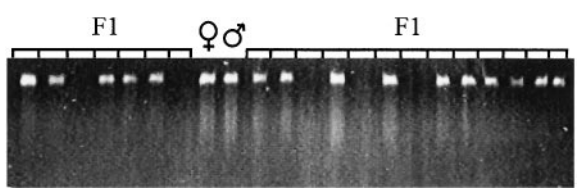

Fig. 4 Segregation of the null allele at locus MSTA10 in the progeny of a cross between diploid sexuals (see text). Data shown: genotypes of the parents and of 20 plants from the progeny; 20 × 20 cm polyacrylamide gel stained with ethidium bromide

tion analysis). Both parents produced one band, and there was a segregation for the presence:absence of the band in the offspring (Fig. 4). Both parents were then assumed to be heterozygous for the null allele. Assuming Hardy-Weinberg equilibrium, the expected frequency of the null alleles at locus MSTA10 in populations 1 and 2 (Table 2) was estimated to be 0.40 and 0.34, respectively [calculation based on observed and expected heterozygosities according to Brookfield (1996), considering the non-amplifying individuals (one in population 1 and one in population 2) as valid null homozygotes]. Null alleles at microsatellite loci have already been frequently reported (e.g. by Pemberton et al. 1995), and special care must therefore be taken in population studies.

Segregation analysis

Mendelian segregation of the putative alleles was checked for eight polymorphic markers (MSTA10, 11, 31, 53, 64, 67, 72, and 78) in the progeny of a controlled

Table 3 Single-locus segregation analysis of microsatellite markers in the progeny of a cross between diploid sexuals (see text)

Locus	Expected segregation ratios	Observed segregation ratios	Chi-square	Degrees of freedom	Population size	$P(X > \chi^2)$
MSAT10	3:1[a]	68:24[a]	0.06	1	92	0.81
MSTA11	2:1:1	43:25:23	0.36	2	91	0.83
MSTA31	1:1:1:1	24:27:21:21	1.07	3	93	0.79
MSTA53	1:1	46:43	0.10	1	89	0.75
MSTA64	1:1	46:45	0.01	1	91	0.92
MSTA67	1:1	47:44	0.10	1	91	0.75
MSTA72	1:1	39:43	0.20	1	82	0.66
MSTA78	1:1:1:1	21:20:27:22	1.29	3	90	0.73

[a]Segration for presence:absence of a band

Table 4 Test of independance of microsatellite markers in the progeny of a cross between diploid sexuals

Pair of loci	Chi-square	Degrees of freedom	Population size	$P(X > \chi^2)$
MSTA10[a] – MSTA11	0.35	2	91	0.84
MSTA10[a] – MSTA31	3.15	3	87	0.37
MSTA10[a] – MSTA53	0.19	1	89	0.67
MSTA10[a] – MSTA64	0.14	1	85	0.71
MSTA10[a] – MSTA67	0.04	1	91	0.84
MSTA10[a] – MSTA72	0.77	1	77	0.38
MSTA10[a] – MSTA78	0.58	3	84	0.90
MSTA11 – MSTA31	8.47	6	86	0.21
MSTA11 – MSTA53	0.36	2	88	0.84
MSTA11 – MSTA64	19.31	2	84	0.00
MSTA11 – MSTA67	3.06	2	90	0.22
MSTA11 – MSTA72	0.61	2	77	0.74
MSTA11 – MSTA78	3.71	6	83	0.72
MSTA31 – MSTA53	5.42	3	85	0.14
MSTA31 – MSTA64	0.69	3	87	0.88
MSTA31 – MSTA67	5.80	3	86	0.12
MSTA31 – MSTA72	5.91	3	78	0.12
MSTA31 – MSTA78	4.53	9	86	0.87
MSTA53 – MSTA64	0.93	1	82	0.34
MSTA53 – MSTA67	0.06	1	88	0.80
MSTA53 – MSTA72	0.09	1	74	0.76
MSTA53 – MSTA78	56.16	3	81	0.00
MSTA64 – MSTA67	3.95	1	84	0.05
MSTA64 – MSTA72	2.12	1	81	0.15
MSTA64 – MSTA78	2.70	3	88	0.44
MSTA67 – MSTA72	0.48	1	77	0.49
MSTA67 – MSTA78	4.63	3	83	0.20
MSTA72 – MSTA78	5.03	3	80	0.17

[a]Segration of presence:absence of a band for MSTA10

cross between two diploid sexual plants from distant geographical origins (female: Lex2x-10 from central Netherlands, male: TJX4–15 from southern France). For all markers, except MSTA10, we found no significant deviation from co-dominant single-locus Mendelian segregation patterns (Table 3). For MSTA10, a segregation for the presence and absence of a band was observed (Fig. 4) and the frequencies were consistent with the hypothesis of a null allele heterozygous in both parents and segregating in a Mendelian fashion (Table 3).

Significant co-segregations were observed between MSTA11 and MSTA64, between MSTA53 and MSTA78 as well as between MSTA64 and MSTA67 (Table 4), corresponding to genetic distances of respectively 21, 10, and 39 cM. Apart from these three pairs, no significant linkage was observed between the eight loci (Table 4). Although the observation of three pairs of linked loci among eight provides insufficient data to conclude about the genomic distribution of the loci, there seems to be no strong clustering of our microsatellites among and within the eight chromosomes of the *T. officinale* haploid genome. This is consistent with the common finding that microsatellites are evenly distributed throughout plant genomes (Wu and Tanksley 1993; Ma et al. 1996). Random genomic

distribution of loci is important for population studies, because linked loci decrease the information content of the data. In addition, disturbed chromosome segregation can occur during male meiosis in triploid *T. officinale* apomicts. Tagging different chromosomes with microsatellites is then useful to analyze crosses between sexuals and asexuals, in order to understand which meiotic events lead to the generation of new clones in natural mixed populations.

Conclusion

Microsatellite markers are now available to study the ecology and evolution of sexual and asexual *T. officinale*. The high level of polymorphism that we observed for these microsatellites makes them very powerful tools for genotype identification. Because mutation polymorphism can also be detected within old asexual lineages (Avise et al. 1992), identifying members of the same asexual lineage may require a study of the distribution of genetic distance between pairs of plants (Brookfield 1992) or allele divergence (Birky 1996). Microsatellites may provide suitable data for both types of studies. Concerning the analysis of gene exchange between sexuals and apomicts, our microsatellite loci are well-suited for paternity analyses of controlled crosses involving apomicts (Tas et al., in preparation), as well as for studying gene flow in natural *T. officinale* mixed populations. Such studies are presently being undertaken in our laboratory.

Acknowledgments This research was funded by the Life Science Foundation (SLW) of the Netherlands Organization for the Advancement of Scientific Research (NWO) and the Royal Dutch Academy of Arts and Sciences (KNAW). We are grateful to Hans den Nijs from the University of Amsterdam, who provided us with the parent TJX4-15. Special thanks are also due to all our other colleagues from the SLW *Taraxacum* network for regular discussions about the progress of our work, and to Arjen Biere, Jos van Damme, Hans Koelewijn, Tracey Parrish, and Peter van Tienderen for useful comments on the manuscript.

References

Asker SE, Jerling L (1992) Apomixis in plants. CRC Press, London.

Avise JC, Quattro JM, Vrijenhoek RC (1992) Molecular clones within organismal clones. Mitochondrial DNA phylogenies and the evolutionary histories of unisexual vertebrates. Evol Biol 26:225–245

Birky CW (1996) Heterozygosity, heteromorphy, and phylogenetic trees in asexual eukaryotes. Genetics 144:427–437

Brookfield JFY (1992) DNA fingerprinting in clonal organisms. Mol Ecol 1:21–26

Brookfield JFY (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. Mol Ecol 5:453–455

Broun P, Tanksley SD (1996) Characterization and genetic mapping of simple repeat sequences in the tomato genome. Mol Gen Genet 250:39–49

Bruford MW, Wayne RK (1993) Microsatellites and their application to population genetic studies. Curr Opin Genet Dev 3:939–943

Condit R, Hubbell SP (1991) Abundance and DNA sequence of two-base repeat regions in tropical tree genomes. Genome 34:66–71

Guilford P, Prakash S, Zhu JM, Rikkerink E, Gardiner S, Bassett H, Forster R (1997) Microsatellites in *Malus* × *domestica* (apple): abundance, polymorphism and cultivar identification. Theor Appl Genet 94:249–254

Gustafsson A (1946) Apomixis in higher plants. I. The mechanism of apomixis. Lunds Univ Arsskr 42:1–67

Hite JM, Eckert KA, Cheng KC (1996) Factors affecting fidelity of DNA synthesis during PCR amplification of $d(C\text{-}A)_n \cdot d(G\text{-}T)_n$ microsatellite repeats. Nucleic Acids Res 24:2429–2434

Jarne P, Lagoda PJL (1996) Microsatellites, from molecules to populations and back. Trends Ecol Evol 11:424–429

King LM (1993) Origins of genotypic variation in North-American dandelions inferred from ribosomal DNA and chloroplast DNA restriction enzyme analysis. Evolution 47:136–151

Kresovich S, Szewc-McFadden AK, Bliek SM, McFerson JR (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

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) Hypervariable microsatellite revealed by in vitro amplification of dinucleotide repeats within the cardiac muscle actin gene. Am J Hum Genet 44:397–401

Ma ZQ, Röder M, Sorrells ME (1996) Frequencies and sequence characteristics of di-, tri-, and tetra-nucleotide microsatellites in wheat. Genome 39:123–130

Menken SBJ, Smit E, Den Nijs JCM (1995) Genetical population structure in plants: gene flow between diploid sexual and triploid asexual dandelions (*Taraxacum* section *Rudelaria*) Evolution 49:1108–1118

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

Morita T, Menken SBJ, Sterk AA (1990) Hybridization between European and Asian dandelions (*Taraxacum* section *Rudelaria* and section *Mongolica*). 1. Crossability and breakdown of self-incompatibility. New Phytol 114:519–529

Pemberton JM, Slate J, Bancroft DR, Barrett JA (1995) Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. Mol Ecol 4:249–252

Ramachandran S, Beukeboom LW, Gerace L, Pavlovic N, Carranza S, Michiels NK (1997) Isolation and characterization of microsatellite loci from the planarian *Dugesia polychroa* (Schmidt) (Platyhelminthes: Tricladida). Mol Ecol 6:389–391

Raymond F, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86:248–249

Rogstad SH (1992) Saturated NaCl-CTAB solution as a means of field preservation of leaves for DNA analyses. Taxon 41:701–708

Rousset F, Raymond F (1995) Testing heterozygote excess and deficiency. Genetics 140:1413–1419

Sutherlands GR, Richards RI (1995) Simple tandem DNA repeats and human genetic disease. Proc Natl Acad Sci USA 92:3636–3641

Slatkin M (1985) Rare alleles as indicators of gene flow. Evolution 39:53–65

Taramino G, Tingey S (1996) Simple sequence repeats for germplasm analysis and mapping in maize. Genome 39:277–287

Tautz D, Renz M (1984) Simple sequences are ubiquitous repetitive components of eukaryote genomes. Nucleic Acids Res 12:4127–4138

Van Heusden AW, Rouppe van der Voort J, Bachmann, K (1991) Oligo-(GATA) fingerprints identify clones in asexual dandelions (*Taraxacum*, Asteraceae). Fingerprints News 3:13–15

Vavrek MC, McGraw JB, Yang HS (1996) Within-population variation in demography of *Taraxacum officinale*: maintenance of genetic diversity. Ecology 77 : 2098–2107

Wang H, Qi M, Cutler AJ (1993) A simple method of preparing plant samples for PCR. Nucleic Acids Res 21 : 4153–4154

Wang Z, Weber JL, Zhong G, Tanksley SD (1994) Survey of plant short tandem DNA repeats. Theor Appl Genet 88 : 1–6

Wright JM (1994) Mutation at VNTRs: are minisatellites the evolutionary progeny of microsatellites? Genome 37 : 345–347

Wu KS, Tanksley SD (1993) Abundance, polymorphism and genetic mapping of microsatellites in rice. Mol Gen Genet 241 : 225–235