Heredity (2003) 90, 326–335

www.nature.com/hdy

Genetic structure of a population sample of apomictic dandelions

RGM Van Der Hulst[1,6], THM Mes[2], M Falque[3], P Stam[4], JCM Den Nijs[1] and K Bachmann[5]

[1]*Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Kruislaan 318 NL-1098 SM Amsterdam, The Netherlands;* [2]*Department of Parasitology and Tropical Veterinary Medicine Institute of Infectious Diseases and Immunology, Utrecht University, The Netherlands;* [3]*The Netherlands Institute of Ecology (NIOO), Center for Terrestrial Ecology, Heteren, The Netherlands;* [4]*Laboratory of Plant Breeding, Wageningen University, Wageningen, The Netherlands;* [5]*Institut für Pflanzengenetik und Kulturpflanzenforschung, IPK Gatersleben, Gatersleben, Germany*

In Northern Europe, dandelion populations consist solely of triploid or higher polyploid apomicts. Without a regular sexual cycle or lateral gene transmission, a clonal structure is expected for *Taraxacum* apomicts, although this was not found by compatibility analysis. In this study, we investigate whether this observation could be suported by performing independent tests based on data from hypervariable microsatellite markers as well as more conservative data based on allozymes and matrilinear cpDNA markers. In addition, population genetic methods were used to test departure from panmictic expectations, which is expected for clonal populations. Results indicated that many data sets, again, did not agree with expectations from clonal evolution because only small groups of genotypes exhibit no marker incompatibility. Population genetic analysis revealed that virtually all genotypes, but not individuals, agreed with random segregation and genotypic equilibria. Exceptions were genotypes with rare allozyme alleles or nearly identical microsatellite genotypes. Consequently, a population sample of apomictic dandelions essentially harbours genotypes that resulted from segregation and/or recombination and only a few genotypes that may have differentiated by somatic mutations.
Heredity (2003) **90,** 326–335. doi:10.1038/sj.hdy.6800248

Keywords: *Taraxacum*; apomixis; AFLPs; character incompatibility

Introduction

The extensive amount of genotypic diversity in populations of apomictic plants remains largely unexplained. Since Ellstrand and Roose (1987) reviewed clonal diversity in apomictic plant populations, numerous studies have confirmed the existence of many multilocus genotypes in these populations (eg Assienan and Noirot, 1995; Noyes and Soltis, 1996; Van der Hulst *et al*, 2000). This contrasted with previous ideas that such populations were relatively depauperate with respect to genotypic diversity (reviewed by Silander, 1986). Since there appears to be no paucity of genetic variation in asexual plant populations, the main question is whether somatic mutations provide a significant source of genetic variation in populations comprising obligate apomicts (Silander, 1986).

The question whether and on what geographic scale, clone-mates differing by somatic mutations, can be detected certainly applies to apomictic dandelions, '*Taraxacum*' agamosperms. The genus *Taraxacum* comprises about 2000 named microspecies and the genus is distributed over large areas of the world (Richards, 1973). Outside areas in Central Europe (Den Nijs and Sterk, 1980) and Asia (Richards, 1973), where diploid sexuals are found, dandelions are almost exclusively polyploid and obligate apomicts. Despite the lack of sexuals in these areas, it is not exceptional to find tens of different genotypes co-occurring in meadows, pastures or verges (Lyman and Ellstrand, 1984; Van der Hulst *et al*, 2000, and unpublished results). The wealth of genotypes present in regions where sexuals are lacking could be the result of immigration from regions where sexuals and apomicts co-occur and hybridise. Richards (1973) postulated that migration of a hybrid swarm of apomicts occurred after the last ice age and that the present polyploid apomictic genotypes are the remnants of this event. Mogie and Ford (1988) extended this hypothesis by indicating that the historical dominance of triploid apomicts in Northern Europe exists because self-incompatible diploid sexuals currently cannot or only marginally penetrate areas where triploids prevail. In large areas that are absolutely dominated by triploid apomicts, the maintenance of genotypic diversity through sexual recruitment will be limited if not absent. As a result genotypic differentiation will mostly depend on accumulation of mutations in apomicts. From this point of view, it is interesting to infer to what extent somatic mutations contribute to genotypic differentiation in stands of dandelions. This question was first investigated by King (1993) who analysed chloroplast and nuclear ribosomal DNA to establish that the variation among apomicts in North America was the result of hybridisation rather than mutation. Similar conclusions were drawn by Van der Hulst *et al* (2000) who used AFLP markers, analysed for character compatibility,

Correspondence: GM Van der Hulst, PO Box 386, 6700 AJ Wageningen, The Netherlands. E-mail: Ron.vanderHulst@wvz.nl
[6]*Current address: PO Box 386, 6700 AJ Wageningen, The Netherlands.*

on apomictic dandelions from The Netherlands and Denmark.

Our present study extends the work of Van der Hulst *et al* (2000) in understanding the genetic structure of triploid apomictic dandelions. This study should further answer our question concerning the relative contribution of segregation, recombination, and mutation to the genotypic variability in a triploid apomictic population. We reanalysed a population ('Viborg'), previously analysed by Van der Hulst *et al* (2000) using AFLP markers, with additional markers. These markers include allozymes, microsatellites and chloroplast DNA markers. The aim of using additional markers was threefold. (1) By adding locus-specific markers (microsatellites and allozymes) data were obtained that are suitable for additional tests such as testing departure from random segregation and from genotypic equilibria. These methods are commonly used to study organisms with 'cryptic sex' or 'clonality' (Burt *et al*, 1996; Gräser *et al*, 1996; Xu *et al*, 1999; Simon *et al*, 1999; Anderson and Kohn, 1998). (2) The aim of including chloroplast DNA markers was to delimit the nuclear marker data (AFLPs and microsatellites) to matrilineal groups (genotypes with identical cpDNA haplotypes), which should enhance the detection of clonal lineages by character compatibility (also see Clepitis, 2001). (3) The aim of using hypervariable microsatellite markers was to increase the probability of detecting genetic differentiation by somatic mutation in clonal lineages.

Our intention was to use population genetic and phylogenetic methods to find out whether variation in a stand of genetically variable apomicts is the result of sexuality or that evidence can be found for genotypic differentiation by somatic mutations. Our second goal was to evaluate and to discuss the markers and methods in the light of their ability to answer questions on clonality in polyploid apomicts.

Materials and methods

Plant material

Dandelions were sampled nonselectively from a park lawn in the city of Viborg, Jutland, Denmark by taking single individuals every metre apart along lines separated by 1 m. Plants were grown in pots in our greenhouse in Amsterdam. All plants represented 'modern' *Taraxacum* sections (*sensu* Richards, 1973) including *Ruderalia (=Vulgaria), Hamata* and *Celtica* (and one member of either Sect. *Borealia* or *Ruderalia*) as determined by P. Oosterveld. As our aim was to characterise the genotypic variability in this population, sampling was independent of taxonomic determination.

Molecular markers and genotyping

DNA was isolated using a modified protocol of Doyle (1991). AFLP analysis was as previously performed by Van der Hulst *et al* (2000). Plants were analysed for six additional loci. These included three allozyme loci, *6-Pgdh-1, 6-Pgdh-2,* and *Mdh* following electrophoretic techniques that were essentially those described in Menken *et al* (1989) and three microsatellite loci MSTA72, MSTA78 and MSTA64 following the protocol of Falque *et al* (1998).

cpDNA haplotyping

Chloroplast fragments of specific regions, the *trnL* intron, the *trnL–trnF* intergenic spacer (Taberlet *et al*, 1991), and the *psbA* and *trnH* intergenic spacer (Sang *et al*, 1997) were amplified using PCR, and fragment sizes were determined by agarose gel electrophoresis. Based on cpDNA sequence data obtained from a survey of many Eurasian dandelions (approximately 100 *trnL–trnF* sequences and 50 *psbA* sequences; THM Mes, unpublished data), enzymes *BsteII*, *DraI* and *SspI* were chosen that uncover polymorphic restriction sites and length variants that could be interpreted in terms of known insertions or deletions. Cladistic analysis (using PAUP 3.1.1, Swofford, 1993) indicated that chloroplast characters are not homoplasious and reflect ancestor-derivative relationship (unpublished data). The chloroplast haplotypes were used to sort the plants into identical matrilineal groups.

Testing random segregation

To test deviations from Hardy–Weinberg proportions, the permutation version of the exact test (Guo and Thompson, 1992), adapted for triploidy was used. In this test, random triplets of alleles were taken from the observed data without replacement and repeated one million times. By arranging these sample compositions according to their frequency of occurrence, significance was calculated as the proportion of samples that is identical to or greater than the observed array of genotypes.

Testing random segregation of diploid data (for MSTA78) was performed by a permutation test using the program MLD (Zaykin *et al*, 1995) such that the association of alleles within a single locus (deviation from random segregation) was tested.

Genotypic disequilibrium

A test for association of alleles across pairs of loci based on permutation was employed using the program MLD (Zaykin *et al*, 1995). In this permutation test, observed genotypes were compared with the expectation of freely recombining populations, estimated by resampling, without replacement of genotypes within each locus. In this special case of triploid data, genotypes at individual loci were not broken up; thus resampled di-locus genotypes are combinations of observed single locus genotypes. The lack of distinction of the observed data from the data calculated by permutation is interpreted as evidence for random segregation. For a concise explanation of the rationale, see Tibayrenc (1997).

UPGMA clustering

A dendrogram was made using the program NTSYS-pc (Rohlf, 1993). Dice similarity was calculated for a data set including both AFLP and microsatellite data using SIMQUAL, and UPGMA clustering was performed using SAHN.

Character incompatibility

Mes (1998) and Van der Hulst *et al* (2000) explain the rationale for using character incompatibility. The rationale is also clearly outlined in 'Box 1' in Anderson and Kohn (1988). The presence of four genotypes that are different combinations of two binary coded markers (0/0, 1/0, 0/1, and 1/1) can be explained by two

mechanisms. The first mechanism is by parallel or reverse mutation. Starting with genotype 0/0, three others can arise by either three mutations of which two are parallel (eg 0/0→1/0 and 0/0→0/1, and 0/1→1/1), or three mutations, including one reverse mutation (eg 0/0→0/1 and 0/1→1/1, and 1/1→1/0). The second mechanism is a combination of two mutations followed by recombination. For example, 0/0→0/1 and 0/0→1/0 and 0/1×1/0 resulting in 1/1. The basic assumption, using incompatibility as a measure of recombination, is that presence of four genotypes that are different combinations of two binary coded markers (0/0, 1/0, 0/1, and 1/1) is more 'parsimoniously' explained by recombination than by three mutation events, especially because the parallel and reverse mutations are likely to be infrequent events. The term 'matrix incompatibility' refers to the presence of 0/0, 1/0, 0/1, and 1/1 for any two marker pairs present in the data and the word 'incompatibility' itself means that it is considered 'incompatible' with clonal evolution. In a population that is strictly clonal there should be no incompatibility between pairs of markers. Our working hypothesis is that if the genotypes that contribute the most to incompatibility are successively deleted from the data, this should eventually reveal a subgroup of genotypes that have strict clonal relationships and could be arranged in a phylogenetic tree without homoplasy (see Mes, 1998; Van der Hulst *et al*, 2000; Clepitis, 2001). If successive deletion does not result in a sudden decrease of incompatibilities and the deletion of virtually all genotypes is required to remove matrix incompatibility, it is concluded that relationships among all of the genotypes are essentially the result of recombination. The contribution of each genotype to the matrix incompatibility was calculated using a slightly modified version of PICA95 (Wilkinson, 1995). In this modified version, the interdependence of markers can be specified. This option prevents the calculation in character incompatibility between alleles of the same microsatellite locus. In contrast to Mes (1998), Van der Hulst *et al* (2000) and Clepitis (2001), the total matrix incompatibility 'counts' are not specified. Instead, the number of marker pairs that are incompatible (incompatibility counts) relative to the number of marker pairs that are left in the data set were counted. This allows a better comparison of the relative incompatibility across our data sets (Figure 2). In this study, character incompatibility is analysed in both the data sets comprising all multilocus genotypes and the data sets comprising genotypes that have identical cpDNA haplotypes.

Results

Genotyping

The results of genotyping using allozymes, microsatellites, AFLP and cpDNA markers are shown in Table 1, with a more detailed characterisation of chloroplast haplotypes in Table 2. Using three allozyme loci, 18 genotypes were detected. A much greater number of genotypes was detected using either microsatellite loci or AFLPs, which were similar and consistent in resolution, and detected 32 and 33 genotypes, respectively. AFLP genotypes are represented by codes (1–33) and their marker information is summarised by a UPGMA dendrogram (Figure 1). In some cases, AFLP markers detected genetic differences among individuals, that were not detected by microsatellites (AFLP genotypes 4 & 5, 7 & 8, 15 & 16, 24 & 25 & 26; Table 1). Conversely, some individuals shared AFLP genotypes (9, 11, and 24) that were separated by microsatellites. Compared to allozymes, AFLPs and microsatellites show a higher genotypic resolution. The high genotypic resolution of microsatellites is the result of a large number of alleles per locus. At microsatellite loci MSTA72, MSTA78 and MSTA64, 28, 13, and six alleles were found, respectively. The allozyme loci *6-pgdh-2*, *6-pgdh-1*, and *mdh* only have three or four alleles. The dominant AFLP markers have only two alleles per locus. However, because of the large number of bi-allelic markers (60 markers in total), the AFLP fingerprinting, like microsatellites, detected twice as many genotypes as were found using allozymes. Clearly, the marker resolution of allozymes is not high enough to tell all genotypes apart, as identical multilocus allozyme genotype were found that are associated with different cpDNA haplotypes (Table 1). In contrast, there are no identical multilocus microsatellite or AFLP genotypes that differ for cpDNA haplotypes (Table 1). It is therefore likely that microsatellites and AFLP reach a resolution, high enough to detect genetic variation within matrilineal groups (nearly identical clone-mates or 'satellite clones' as referred to by Menken *et al*, 1995.

Testing random segregation

Virtually all single locus genotypes share alleles (Table 1), which gives the impression of a population 'segregating' for allozyme alleles, rather than a population in which reproductively isolated genotypes differentiated by somatic mutation. The allelic distribution at one of the three individual loci (*6-Pgdh-2*) does not differ significantly from random segregation (including either all individuals or multilocus genotypes; Table 3). For the other two allozyme loci, *6-Pgdh-1* and *Mdh*, random segregation is rejected for data including all individuals. By using genotypes (either multilocus allozyme or microsatellite genotypes) instead of individuals for *6-Pgdh-1* and *Mdh* the probability was increased by a factor of 10^4, but the null hypothesis of random segregation is still rejected (see Table 3). The deviation from random segregation for the two loci *6-Pgdh1* and *Mdh* is primarily caused by the presence of rare alleles found in one or two single locus allozyme genotypes 'BCC' at *6-Pgdh-1*, and 'EKK' and 'AKO' at *Mdh* (see Table 1; frequency $B = 0.015$, frequency $K = 0.025$, $O = 0.005$). By deleting these rare alleles (*B* at *6pghdh1* and *O* and *K* at *mdh*), a data set remained that was not significantly different from random segregation at all three individual allozyme loci (Table 2).

For the microsatellite loci, a pattern similar to that of allozymes is observed. Again, genotypes are alternative combinations of alleles found in other genotypes and the number of alleles unique to a certain genotype is low (Table 1). For locus *MSTA78*, all genotypes in the data were found to have a '164 bp' allele at locus *MSTA78*. This *MSTA78-164bp* allele is known to be linked to diplospory and because of its typical inheritance (van Dijk *et al*, in preparation), all apomictic genotypes were expected to be simplex for this allele. As a result, the data (excluding *MSTA78-164bp*) could be interpreted in a

Table 1 Multilocus genotypes using microsatellites, AFLP markers (also see Figure 1) and allozymes and cpDNA markers (also see Table 2)

Section	*Microspecies*	*MSTA72*	*MSTA78*	*MSTA64*	*AFLP genotype code*	6-Pgdh-2	6-Pgdh-1	Mdh	*cpDNA - haplotype*	*Number*
Borea or Rududeralia		184-189	157-162-164	185-207	1	BBC	AAC	AAA	2	1
Ruderalia		179-180-187	ND	191-211	2	BCC	AAA	AAA	1	1
Ruderalia	*(gelertiformae)*	176-185-196	150-164-168	191-213	3	BCC	AAA	AAA	2	1
Ruderalia		185-203-210	164-168-170	185-211	4	BCC	AAA	AAE	1	1
Ruderalia		185-203-210	164-168-170	185-211	5	BCC	AAA	AAE	1	1
Ruderalia/Celtica	*rhodopodum*	181-185-196	164-168-174	185-191-207	6	BCC	AAA	AAE	1	2
Ruderalia/Celtica	*rhodopodum*	181-185-196	164-168-174	185-191-207	ND	BCC	AAA	AAE	1	1
Ruderalia		187-190	149-164-174	191-197-213	7	BCC	AAA	AAE	3	3
Ruderalia		187-190	149-164-174	191-197-213	8	BCC	AAA	AAE	3	1
Ruderalia	*valens*	170-190-**195**	164-166-174	207-211-213	9	BCC	AAA	AEE	2	4
Ruderalia	*valens*	170-190-**197**	164-166-174	207-211-213	9	BCC	AAA	AEE	2	1
Ruderalia	*valens*	170-190-**199**	164-166-174	207-211-213	9	BCC	AAA	AEE	2	1
Ruderalia		178-182-191	162-164-168	191-213	10	BCC	AAC	AAA	1	2
Ruderalia		174-**182**-184	162-164-168	207-211-213	11	BCC	AAC	AAA	2	1
Ruderalia		174-184-**190**	162-164-168	207-211-213	11	BCC	AAC	AAA	2	1
Ruderalia	*(subundulatum)*	186-189-196	153-164-168	185-207-213	12	BCC	AAC	AAE	1	1
Ruderalia		184-185-190	149-162-164	185-207	ND	BCC	AAC	AAE	2	1
Ruderalia		173-186-197	164-166-172	185-213	13	BCC	AAC	AAE	2	1
nd		176-181-182	164-170-176	185-191-207	ND	BCC	AAC	AEE	2	1
Ruderalia	*(fagerstroemii)*	176-181-182	164-170-176	185-191-207	14	BCC	AAC	AEE	2	2
Ruderalia		175-180-184	162-164	191-213	15	BCC	AAC	EKK	4	1
Ruderalia		175-180-184	162-164	191-213	16	BCC	AAC	EKK	4	1
Ruderalia	*(dilatatum)*	176-182-188	155-164-**174**	191-213	17	BCC	BCC	AAE	4	2
Ruderalia	*(dilatatum)*	176-182-188	155-164-**176**	191-213	18	BCC	BCC	AAE	4	1
Ruderalia		186-187-188	162-164-176	185-211	19	BCE	AAC	AAA	1	1
Hamata	*hamatiforme*	170-194	149-164-168	185-210	20	CCD	AAA	AAA	4	2
not possible		176-180-183	153-164-170	185-211	21	CCC	AAA	AAA	1	1
Ruderalia	*(angustiquameum)*	178-197	149-164-166	211-213	22	CCC	AAA	AAE	2	7
nd		182-184-186	164-166	191-207-211	23	CCC	AAA	AAE	6	2
Hamata		164-170-**178**	164-170-174	185-191-207	24	CCC	AAA	AEE	2	1
Hamata		164-170-**180**	164-170-174	185-191-207	24	CCC	AAA	AEE	2	9
Hamata		164-170-**178**	164-170-174	185-191-207	25	CCC	AAA	AEE	2	1
Hamata		164-170-**180**	164-170-174	185-191-207	26	CCC	AAA	AEE	2	1
Hamata		164-170-**180**	164-170-174	185-191-207	ND	CCC	AAA	AEE	2	1
Ruderalia	*acroglossum*	184-196-203	164-168-176	197-211	27	CCC	AAA	ND	1	1
Ruderalia	*ancistrolobum*	191-193-197	162-164-166	185-213	28	CCC	AAC	AAA	2	2
nd		191-193-197	162-164-166	185-213	ND	CCC	AAC	AAA	2	1
Ruderalia		174-175-180	164-166-172	185-213	29	CCC	AAC	AAA	4	1
nd		176-182-199	162-164	185-213	30	CCC	AAC	AAE	2	1
Ruderalia		175-180-181	149-164-166	211-213	31	CCC	AAC	AAE	2	1
Ruderalia		170-180	162-164	185-213	32	CCC	AAC	AKO	2	1
Ruderalia		174-186-199	149-164-168	185-211	33	CCE	AAC	AAA	5	2
nd		184-185-190	149-162-164	185-207	ND	ND	ND	ND	2	1

ND=not determined. Alleles that are underlined are putative somatic mutations by which alleged clone-mates have differentiated.

Table 2 cpDNA haplotypes and the diagnostic fragment sizes and restriction enzymes used for haplotyping

	TrnL *intron length (bp)*	trnL-trnF *spacer length (bp)*	Bste*II site positon*	psbA-trnH *spacer length (bp)*	Ssp*I site position*	Dra*I site position(s)*
cpDNA1	525	290	—	525	170	—
cpDNA2	525	290	—	525	130	—
cpDNA3	480	290	—	525	130	—
cpDNA4	480	290	—	525	130	350
cpDNA5	480	290	250	500	130	225+350
cpDNA6	480	400	250	500	130	225+350

'diploid context', which allowed testing departure from Hardy–Weinberg equilibrium (HWE) using a conventional population genetic package (such as MLD, Zaykin *et al*, 1995). Random segregation at *MSTA78* was rejected for data from all individuals ($P<0.0001$). For data including genotypes ($n=32$), however, no significant deviation from random segregation was found ($P=0.233$; Table 3). This result suggested that the pattern of random segregation at locus *MSTA78* is mainly disrupted as the result of apomixis, which amplifies

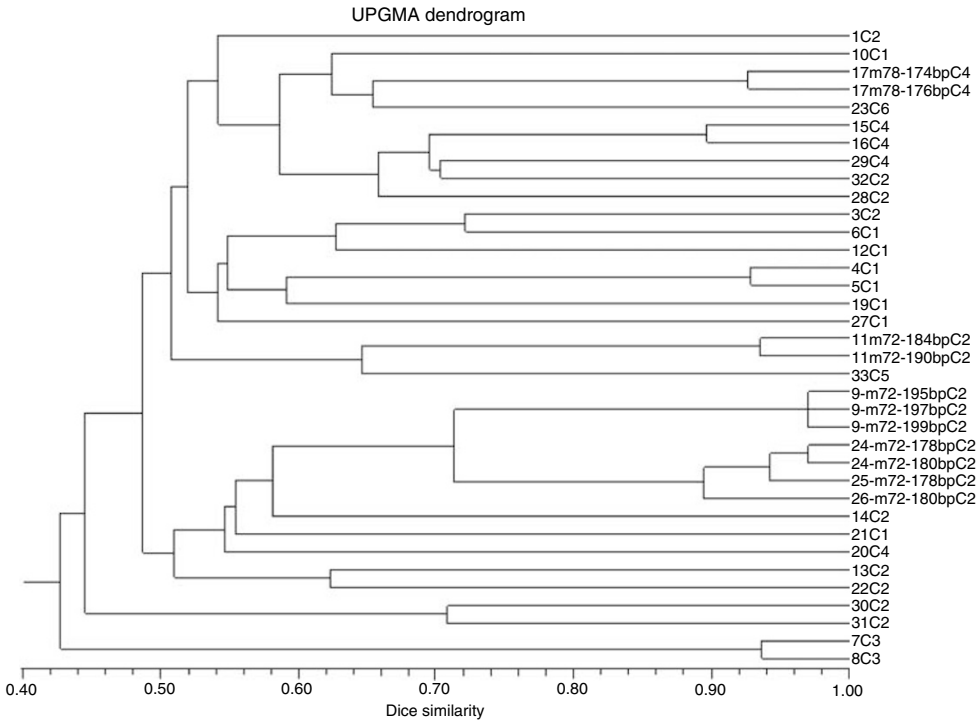

Figure 1 UPGMA dendrogram for genotypes characterised by AFLP markers and three microsatellite loci (also see Table 1). The first number of each genotype refers to its given AFLP code. The number following an 'm' refers to a microsatellite locus at which single allele differences (usually of 2bp differences) occur. These variants are indicated because it is assumed that these are somatic mutations by which the genotypes clustering together have differentiated. Numbers following a 'C' refer to chloroplast haplotypes. Note that chloroplast haplotypes are not restricted to single clades.

Table 3 Probability that data agree with random segregation, calculated for the loci *6-Pgdh-1*, *6-Pgdh-2*, *Mdh* and *MSTA78* for different data sets

Data including	6-Pgdh-1	6-Pgdh-2	Mdh	msta78	6-Pgdh-1 *without rare alleles*	Mdh *without rare alleles*
Individuals (68)	10^{-7}	0.15	4×10^{-6}	$<10^{-5}$	0.24 (65)	0.045 (65)
Allozyme genotypes (18)	0.02	0.24	0.016	NT	0.64 (17)	0.64 (16)
Microsatellite genotypes (32)	8.4×10^{-4}	0.48	0.0082	0.233	0.095 (30)	0.83 (30)
Microsatellite genotypes, without 'mutants' that differ for one allele (27)	7.6×10^{-3}	0.57	0.0147	0.544	0.09 (26)	1 (25)

Data sets include individuals, genotypes based on three allozyme loci, genotypes based on three microsatellite loci, or microsatellite genotypes for which the 'mutant' types that differ for only one allele were deleted. Rare alleles have frequencies <0.05. NT=not tested. Numbers indicated in parenthesis refer to the number of genotypes in the particular data set tested.

genotypes rather than because of clonal mutation accumulation. Some individuals, differing only for a single microsatellite allele, however, appeared to be good candidate representatives of genotypes resulting from somatic mutation (underlined in Table 1). To test the effect of the presence of these 'mutant' genotypes on the probability that the data resulted from random segregation, a data set without nearly identical genotypes (the minority type of genotypes differing for only one allele deleted, $n=27$) was analysed. The probability that this data set resulted from random segregation is two times higher ($P=0.544$ instead of $P=0.252$, Table 3) than observed for the data that included all genotypes ($n=32$). Although it is not significant, this may be interpreted as support for the hypothesis that some genotypic differentiation by somatic mutation has occurred.

Testing genotypic disequilibrium

Genotypic disequilibrium was tested for pairs of genotypes across allozyme and/or microsatellite loci. This analysis was restricted to those loci for which it is possible to determine unambiguously single-locus-genotype frequencies. In addition to *MSTA78*, also locus *MSTA72* could be used for this analysis, because each *MSTA72* genotype that is not fully heterozygous (marker phenotype) was unique in the data set. The results of the genotypic disequilibrium tests are shown in Table 4. All 10 pairs of loci tested showed a significant deviation

Table 4 Test for genotypic disequilibrium between pairs of loci for *6-Pgdh-2*, *6-Pgdh-1*, *Mdh*, *MSTA78*, and *MSTA72* for data including individuals, genotypes or microsatellite genotypes without 'mutants'; that differ for only one microsatellite allele, determined by an exact permutation test (MLD; Zaykin *et al*, 1995)

Probability of disequilibrium between loci		*For data including*		
		Individuals	*Genotypes*	*Data without 'mutants' differing at one allele*
6-Pgdh-2	*6-Pgdh-1*	0.0225	0.741	0.937
6-Pgdh-2	*Mdh*	0.032	0.7714	0.7879
6-Pgdh-1	*Mdh*	$\sim 10^{-4}$	0.1415	0.7772
6-Pgdh-2	*MSTA78*	$< 10^{-4}$	0.0356	0.4025
6-Pgdh-1	*MSTA78*	$< 10^{-4}$	0.0056	0.1004
Mdh	*MSTA78*	$< 10^{-4}$	0.0031	0.1812
6-Pgdh-2	*MSTA72*	$< 10^{-4}$	0.0764	1
6-Pgdh-1	*MSTA72*	$< 10^{-4}$	0.0023	1
Mdh	*MSTA72*	$< 10^{-4}$	0.1418	1
MSTA78	*MSTA72*	$< 10^{-4}$	1	1

For each pairs of markers, the probability that the markers agree with random association (genotypic equilibrium) are indicated.

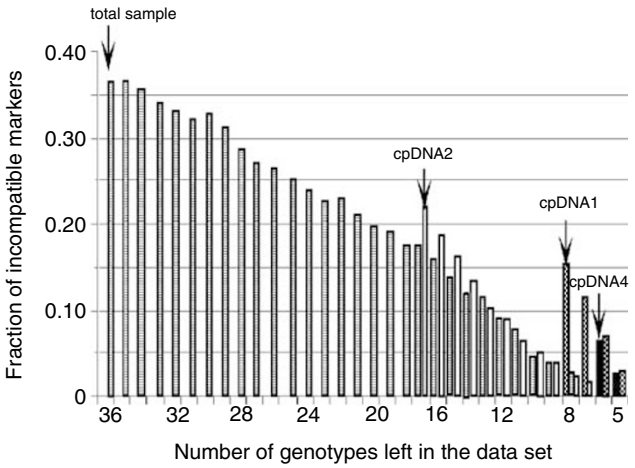

Figure 2 The fraction of incompatible marker pairs left in the data after successive deletion of genotypes that showed the highest contribution to incompatibilities. The total sample starts with 36 genotypes. The tests are also performed for genotypes that share particular chloroplast haplotypes (cpDNA2, cpDNA1 and cpDNA4). These sets of genotypes start with smaller numbers of genotypes (also see Table 1) but these exhibit similar fractions of incompatible markers and show similar gradual decrease of incompatibilities in the test.

from a genotypic equilibrium for data using individuals. For data including all genotypes instead of individuals (correction for clonal copies), six of 10 pairs of loci fit to a genotypic equilibrium. If, however, data are used in which genotypes differing by a single microsatellite allele are counted as originally one, then all 10 pairs of loci agree with a genotypic equilibrium (Table 4). This genotypic disequilibrium test suggests that the genotypes essentially represent a recombining population but only after correction for identical and nearly identical genotypes.

Incompatibility tests (all genotypes)

As mentioned in the Materials and methods section, the incompatibility test performed in this study is similar to the one performed by Van der Hulst *et al* (2000). Microsatellite data related to the former study were also used for compatibility analysis as binary coded data. The results obtained for this new data set (Figure 2) are essentially similar to previously obtained results. It takes the deletion of many genotypes to get rid of marker pairs that are incompatible with genetic differentiation by mutation in clones. Compared to the AFLP data, analysis of data combining both AFLPs and microsatellites did not identify larger groups of genotypes that have strictly clonal relationships. There is a gradual decrease of incompatibility upon successive deletion of genotypes contributing most to incompatibility (Figure 2). Consistent with the analysis of AFLP data (Van der Hulst *et al*, 2000), only eight genotypes remained in the data set that are free from incompatible marker pairs, which is close to the minimum number of four genotypes required to identify incompatibilities. Among these genotypes, the nearly identical genotypes were found that differ for only one microsatellite marker (specified by AFLP genotype code 9, AFLP genotype codes 17 and 18, and AFLP genotype code 24; see Table 1). In these remaining eight genotypes, three different cpDNA haplotypes (cpDNA-2, cpDNA-3 and cpDNA-4) were represented. In conclusion, the incompatibility analysis and deletion procedure provided three groups of nearly identical genotypes not linked by recombination and one separate genotype (AFLP genotype code 7). However, the number of genotypes that have a clonal relationship is lower. This is because only five genotypes, belonging to two groups of nearly identical genotypes (members of genotypes specified by AFLP codes 9 and 24), also have a matrilineal relationship as inferred by their shared cpDNA haplotype 2 (Table 1).

Incompatibility tests (data sets delimited by haplotypes)

Results for the incompatibility tests for data delimited by cpDNA haplotypes are shown in Figure 2. A gradual decrease of incompatibility upon deletion of genotypes contributing most to matrix incompatibility was found for genotypes of three cpDNA subsets (cpDNA1, cpDNA2, and cpDNA4). In every case, four to six genotypes were free from incompatibilities, which is close to the minimum number of four genotypes required. In contrast to what was expected, the fraction of incompatibilities as a function of the number of genotypes left in the data is even higher within cpDNA haplotype groups than for the total data (Figure 2). This indicates that there is more incompatibility present within cpDNA haplotypes than within the total data set. For the remaining three cpDNA groups, there are no large groups of genotypes that are differentiated by somatic mutation because these harbour only one (cpDNA3) or two nuclear marker genotypes (cpDNA5, and cpDNA6). In conclusion, the results are in contrast to our hypothesis of enhanced probability of finding clonemates within data sets that comprise genotypes with identical cpDNA haplotypes.

Discussion

There is little tradition in using population genetic and/or phylogenetic criteria to test clonality versus

recombination and segregation as sources of genetic variation in clonal plants. Exceptions are studies of King (1993), who implicitly and Mes (1998), Van der Hulst *et al* (2000), and Clepitis (2001), who explicitly used incompatibility as criterion to determine sexual origin, and Menken *et al* (1995), who used nondeparture from HWE as a criterion for a sexual origin of triploid dandelions. The lack of a tradition in studying higher plant clonality using statistics or phylogenetic criteria strongly contrasts with the wealth of population genetic and phylogenetic studies of microorganisms that focus on clonality (eg Gräser *et al*, 1996; Xu *et al*, 1999; Tibayrenc *et al*, 1991, Tibayrenc, 1997). In this discipline, a population genetic and phylogenetic framework has been developed to explore the potential impact of genetic exchange. The central null hypothesis in this framework is that, as opposed to clonality, genetic exchange occurs randomly. Tests of genetic exchange may include: (1) testing departure from HWE (Gräser *et al*, 1996); (2) linkage disequilibrium by testing the index of association (Maynard-Smith *et al*, 1993; Burt *et al*, 1996; Xu *et al*, 1999); (3) testing genotypic disequilibrium between pairs of loci (eg Gräser *et al*, 1996); (4) testing the length of a phylogenetic tree against tree lengths obtained from random segregation of the data (haploid data; Burt *et al*, 1996); and (5) determining the incompatibilities (Achtman *et al*, 1999) and testing significance (Rosendahl and Taylor, 1997).

It has been argued by Tibayrenc *et al* (1991) 'that the mode of reproduction of microorganisms in nature can only be decided by population genetic information'. We think that this idea also holds for clonal plants, such as dandelions. Therefore, we tested (1) departure from random mating; (2) linkage disequilibrium for allelic markers (allozymes, microsatellites), and (3) character compatibility for DNA markers (microsatellites, AFLP markers).

The aim of this study was to analyse the molecular marker data obtained from a stand of triploid apomictic dandelions to decide whether the genetic variation in this population was more likely to be the result of somatic mutation within relatively few clonal lines or the result of (sexual) recombination in a population, producing very many apomictic lines. The allozymes, microsatellites, and AFLPs markers were subjected to population genetic analyses and analysis of character compatibility. We conclude that these methods illustrate that a stand of apomictic dandelion genotypes essentially resembles a population that is segregating and recombining for the markers analysed, with a few indications for genotypic differentiation by somatic mutations. These putative mutations occurred predominantly in nearly identical genotypes.

Population genetic tests

Allelic data were analysed using tests for random segregation and genotypic linkage disequilibrium. Both tests rejected the hypothesis that individuals segregate or recombine randomly. This result is expected for populations in which apomixis amplifies fixed combinations of alleles. To correct for apomixis, data sets that included a single representative of each genotype were analysed (for similar approaches in *Sitobion avenae* and *Candida albicans*, see Simon *et al*, 1999; Xu *et al*, 1999). By correcting for apomixis, it was found that departure from random segregation or recombination is essentially the result of the presence of only two genotypes that combine two rare alleles at loci *Mdh* and *6-Pgdh*-1 and that departure is not the result of marker skewness in many genotypes. In the light of the hypothesis that mutations disrupt a panmictic population structure, these rare alleles possibly resulted from mutation. Deleting the rare alleles from the data provides a set of genotypes that essentially resembles a panmictic population. A similar approach was performed for microsatellite data. Among the sampled genotypes, nearly identical genotypes are found that differ by only one microsatellite allele. Intuitively, one would suspect that these are clone-mates that have differentiated by somatic mutation. Statistical support was sought by determining the probabilities of random segregation for two data sets: one data set with all genotypes and another for which genotypes, that differed for one microsatellite allele, are counted as originally one. We interpreted the higher probabilities of both random segregation and recombination in data sets with putative clone mates present only once as evidence for differentiation by somatic mutation, although alternative hypotheses such as a low statistical power must be considered (see below). To date, DNA variation in nearly identical *Taraxacum* genotypes collected from populations has only been reported by Van Heusden *et al* (1991). These authors also suggested that it probably resulted from (somatic) mutations within clonal lineages.

Character incompatibility

Attempts to find evidence for genotypes differing by somatic mutation, initiated efforts to improve the data for the analysis of character compatibility. By adding hypervariable microsatellite markers, we intended to uncover mutations, and by delimiting our data to nuclear marker genotypes with identical cpDNA haplotypes, we aimed at analysing groups of genotypes that have close matrilineal relationships (also see Clepitis, 2001). In the initial analyses of character compatibility performed by Van der Hulst *et al* (2000), deletion of genotypes contributing most to character incompatibility should reveal the presence of clonal lineages. Intuitively, this method seems appropriate, but it cannot be excluded that larger, or different groups of genotypes without incompatibilities, may be obtained by alternative ways of deletion. This could be tested by exhaustive deletion series, which, however, is impractical for large data sets. Alternatively, detection of clonal lineages may be enhanced by preselection of the data into groups that are more likely to be matrilineal, as was carried out in this study. The incompatibility test applied to data sets delimited in this manner did not detect large clonal lineages in our collection of dandelions. Genotypes without incompatibilities were detected, but their number nearly equals the minimal number to detect character incompatibilities. This resembles the results obtained by Clepitis (2001), who also found recombination within chlorotypes of *Allium vineale*, a species with a capacity for mixed sexual and asexual reproduction. Interestingly, our incompatibility analysis provided a set of genotypes with clonal relationships that includes the same, nearly identical genotypes that showed up in the population genetic test. We conclude from both population genetic tests and incompatibility tests that this stand of dandelions represents a collection of recombinant genotypes of which only few have further differentiated by somatic mutation.

General interpretation

It is not unusual to find evidence for both clonality and sexual reproduction. Such results have also been obtained for the yeast *C. albicans* (Gräser *et al*, 1996), which caused Tibayrenc (1997) to question the validity of methods to determine clonality. The points raised by Tibayrenc (1997) were that: (1) 'any statistical departures from panmictic expectations show that gene exchange is inhibited to some extent and are taken as circumstantial evidence of clonality, though alternative explanations must also be considered'; (2) the 'lack of linkage at several pairs of loci can be explained … by the lack of power of a statistical test.' (this also holds for the lack of departure from random segregation); and (3) 'the presence of all possible recombinants' (incompatibility) 'could be generated by mutation only'. We generally agree with these points but stress aspects of the life history of apomictic dandelions in Northern Europe. Only triploid apomictics and no sexual diploids have been found in this area (see Menken *et al*, 1995). Given a considerable time of reproductive isolation, a clonal structure is expected for *Taraxacum* apomicts that do not have a regular sexual cycle or other means of lateral gene transmission. In the light of this latter fact, it is surprising that a predominant genetic signature in the population sample results from random segregation and recombination when correcting for clonal copies. It is important, however, to realise that in the interpretation of population genetic results of organisms with cryptic sex one can support a particular view by stressing the devitation from an extreme scenario. Under strict clonality, one would expect departure from HWE at all loci, genotypic disequilibrium across genotypes at all loci, and not a single incident of all four possible pairs of loci, typical for recombinant genotypes (disregarding parallel mutation). Under panmixis and no physical linkage of markers one would expect exactly the opposite. It could be argued that dandelions probably do not meet either of these two extreme scenarios. Apomictic dandelions have been regarded as a hybrid swarm, fossilised by apomixis, which existed since the last ice age (see Richards, 1973). The time of reproductive isolation may not be long enough to obtain a detectable clonal structure (by somatic mutations), at least not when looking at a single populations.

What causes a stand of apomictic genotypes to exhibit a more or less panmictic population structure? There are several mechanisms that may contribute to the observed panmictic structure. It could be that at least a small fraction of the apomicts have the capacity to reproduce sexually. It is known that when apomicts lose particular chromosomes, harbouring the components of apomixis, these may become (at least partially) sexual (Sørensen and Gudjonsson, 1948; Sørensen, 1958; Van Dijk personal communication). Another possibility is that apomicts produce genetically diverse siblings as the result of recombination occurring at restitutional meiosis (van Baarlen *et al*, 2000) or as the result of mitotic recombination in somatic tissues and/or of transposon activity (Richards, 1989, 1996). It can also be that apomicts do not have any form of sexual reproduction, in particular geographical areas, but have migrated from source populations in which apomicts and sexual diploids co-occur (Menken *et al*, 1995). One way to find out which of these mechanisms contribute to the panmictic population structure is to screen pedigrees obtained from apomicts taken from natural populations for a large number of molecular markers. It is highly preferred to use genetically mapped markers for this purpose, because a more accurate description of different types of mutations is possible. For example when markers, located on a single chromosome, are lost in siblings, this points to chromosome loss. When only one or a few (unlinked markers) are lost or gained, this points to mutation. When a group of markers is gained or lost, limited to a particular part of a chromosome, this points to translocation. In this way, parallel mutations, which may underlie the AFLP banding patterns, can be identified and used to improve the analysis of genetically diverged clone-mates. The temporal aspect, involved in the assessment of the population genetic structure of dandelions, can be determined by establishing whether the observed panmicitic structure is relictual or resulted from current genetic exchange for example, by the construction of phylogenetic trees on the basis of DNA sequences. If apomictic genotypes combine alleles of genes that have diverged since particular evolutionary time span, this indicates that genetic exchange has taken place since the existence of the most recently evolved allele. Unfortunately, strongly supported organismal phylogenies for *Taraxacum* lack.

Although the methods applied in the present study are suitable to study clonality, these can be improved. For example, one of our conclusions from the population genetic tests was that the distribution of markers in the data, corrected for clonal copies, does not significantly deviate from random expectation. To find out whether this nonsignificance results from type II errors that are related to a small sample size only, or can be ascribed to a biologically meaningful correction for clonal copies, simulation studies may be performed. Another possible refinement concerns the incompatibility test. In the present study, cpDNA data were used to delimit data to matrilineal groups that should enhance the detection of clonal lineages. An alternative method is to analyse the data exhaustively, rather than heuristically, to find the largest set of clonally related genotypes. An exhaustive search is, however, not practical considering the size of most data sets in population genetic studies. Further, the observed incompatibilities could be compared with random expectations in permutation tests to determine significance as a measure of recombination (eg Rosendahl and Taylor, 1997). The refinements may contribute to a more accurate assessment of clonality in dandelions. It should, however, be noted that other factors will also influence the outcome of population genetic or incompatibility tests. The first is the type of population that is studied. In the literature of apomictic dandelions there are essentially two ways of interpreting populations. A population sample may comprise, randomly sampled apomictic dandelions (eg Lyman and Ellstrand, 1984; Menken *et al*, 1995; Van der Hulst *et al*, 2000). Alternatively, a population sample may correspond to a set of apomicts, selected on the basis of certain morphological features that are diagnostic for microspecies (eg Hughes and Richards, 1988; Menken and Morita, 1989). As our aim was to analyse genetic variation in a stand of apomicts and to infer whether apomicts have clonal relationships, we applied the first type of sampling. Another aspect is whether tests of

population differentiation are valid when migration is at hand. Since some genotypes are shared among populations (Van der Hulst *et al*, 2000), genotypes from different source populations have been included in a single test for random segregation, which is principally not correct. This argument seems most relevant, however, when random segregation is rejected, which is the opposite of our results. Also, the incompatibility tests apparently has not been hampered by the inclusion of genotypes from different populations. The presence of genotypes that originate from different source populations would have shown up as a sudden decrease of incompatibility upon deletion of genotypes that contribute most to incompatibility. This was not observed.

A second important factor, which influences population genetic and compatibility tests, is the interpretation of marker phenotypes for statistical tests. It was assumed that homozygous allozyme marker phenotypes have three copies of the particular allele (no null-alleles). Further, it was assumed that the absence or presence of an AFLP markers represents two allelic states, although the absence or presence may have different causes (see Vos *et al*, 1995). As an alternative to our current markers, more microsatellite markers or DNA sequences can be used. But every marker technique would only partially allow genotyping as a result of the fact that (a) many individuals are not fully heterozygous, and (b) generally, the number of allelic copies is unknown. As a result of polyploidy, marker dominance, migration, and the available analytical methods, assumptions have to be made that render statistical tests of clonality or sexuality of population of apomictic plants subject to debate. Those who, nevertheless, analysed stands of apomicts found a substantial number of genotypes (reviewed by Ellstrand and Roose, 1987) and found that the distribution of genetic markers agrees with random segregation and/or recombination (Menken *et al*, 1995; Mes, 1998; Van der Hulst *et al*, 2000; Clepitis, 2001). Our study indicates that application of multiple and highly polymorphic marker types simply substantiates these results.

References

Achtman M, Azuma T, berg DE, Ito Y, Morelli G, Pan ZJ *et al* (1999). Recombination and clonal groupings within *Helicobacter pylori* from different geographical regions. *Mol Microbiol* **32**: 459–470.

Anderson JB, Kohn LM (1998). Genotyping, gene genealogies and genomics bring fungal population genetics above ground. *Trends Ecol Evol* **13**: 444–449.

Assienan B, Noirot M (1995). Isozyme polymorphism and organization of the agamic complex of the Maximae (*Panicum maximum* Jacq., *P. infestum* Anders and *P. trichocladum* K. Schum.) in Tanzania. *Theor Appl Genet* **91**: 672–680.

Burt A, Carter DA, Koenig GL, White TJ, Taylor JW (1996). Molecular markers reveal cryptic sex in the human pathogen *Coccidioides immitis*. *Proc Natl Acad Sci* **93**: 770–773.

Clepitis A (2001). The importance of sexual and asexual reproduction in the recent evolution of *Allium vineale*. *Evolution* **55**: 1581–1591.

Den Nijs JCM, Sterk AA (1980). Cytogeographical studies of *Taraxacum* sect. *Taraxacum* (=sect. *Vulgaria*) in central Europe. *Bot Jahrb Syst* **101**: 527–544.

Doyle JJ (1991). DNA protocols for plants. In: Hewitt GM, Johnston AWB, Young JPW (eds) *Molecular Techniques in Taxonomy*, Springer-Verlag: Berlin. pp 283–285.

Ellstrand NC, Roose ML (1987). Patterns of genotypic diversity in clonal plant species. *Am J Bot* **74**: 121–131.

Falque M, Keurentjes J, Bakx-Schotman JMT, Van Dijk PJ (1998). Development and characterisation of microsatellite markers in the sexual apomictic complex *Taraxacum officinale* (dandelion). *Theor Appl Genet* **97**: 283–292.

Gräser Y, Volovsek M, Arrington J, Schönian G, Presber W, Mitchell TG *et al* (1996). Molecular markers reveal that population structure of the human pathogen *Candida albicans* exhibits both clonality and recombination. *Proc Natl Acad Sci USA* **93**: 12473–12477.

Guo SW, Thompson EA (1992). Performing the exact test of Hardy–Weinberg proportions for multiple alleles. *Biometrics* **48**: 361–372.

Hughes J, Richards AJ (1988). The genetic structure of populations of sexual and asexual *Taraxacum* (dandelions). *Heredity* **60**: 161–171.

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.

Lyman JC, Ellstrand NC (1984). Clonal diversity in *Taraxacum officinale* (Compositae), an apomict. *Heredity* **53**, 1–10.

Maynard-Smith J, Smith NH, O'-Rourke M, Spratt bg (1993). How clonal are bacteria? *Proc Natl Acad Sci USA* **90**: 4384–4388.

Menken SBJ, Morita T, Waardenaar ECP, Boersma A (1989). Genetic interpretation of enzyme variation in sexual and agamospermous taxa of *Taraxacum* section *Ruderalia* and *Mongolica*. *Genetica* **78**: 111–119.

Menken SBJ, Morita T (1989). Uniclonal population structure in the pentaploid obligate agamosperm. *Taraxacum albidum* Dahlst. *Plant Species Biol* **4**: 29–36.

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

Mes THM (1998). Character compatibility of molecular markers to distinguish asexual and sexual reproduction. *Mol Ecol* **7**: 1719–1727.

Mogie M, Ford H (1988). Sexual and asexual *Taraxacum* species. *Biol J Linn Soc* **35**: 155–168.

Noyes RD, Soltis DE (1996). Genotypic variation in agamospermous *Erigeron compositus* (Asteraceae). *Am J Bot* **83**: 1292–1303.

Richards AJ (1973). The origin of *Taraxacum* agamospecies. *Bot J Linn Soc* **66**: 189–211.

Richards AJ (1989). A comparison of within-plant karyological heterogeneity between agomopsperms and sexuals *Taraxacum* (Compositae) as asessed by the nucleolar organizer chromsome. *Plant Syst Evol* **163**: 177–185.

Richards AJ (1996). Genetic variability in obligate apomicts of the genus *Taraxacum*. *Folia Geobot Phytotax* **31**: 405–414.

Rohlf FJ (1993). NTSYS-pc. *Numerical Taxonomy and Multivariate Analysis System*. Exeter Software: New York.

Rosendahl S, Taylor JW (1997). Development of multiple genetic markers for studies of genetic variation in arbuscular mycorrhizal fungi using AFLP. *Mol Ecol* **6**: 821–829.

Sang T, Crawford DJ, Stuessy TF (1997). Chloroplast DNA phylogeny, reticulate evolution and biogeography of *Paeonia* (Paeoniaceae). *Am J Bot* **84**: 1120–1136.

Silander Jr, JA (1986). Microevolution in clonal plants. In: Jackson JBC, Buss LW. Cook RA (eds) *Population Biology and Evolution of Clonal Organims*. Yale University Press: New Haven and London. pp 107–152.

Simon JC, Baumann S, Sunnucks P, Herbert PDN, Pierre JS, Le Gallic JF *et al* (1999). Reproductive mode and populations genetic structure of the cereal aphid *Sitobion avenae* studied using phenotypic and microsatellite markers. *Mol Ecol* **8**: 531–545.

Sørensen Th, Gudjonsson G (1948). Spontaneous chromosome-aberrants in apomictic *Taraxaca*. Morphological and cytogenetical investigations. *Biol Skr* **4**: 1–48.

Sørensen Th (1958). Sexual chromosome-aberrants in triploid apomictic *Taraxaca*. *Bot Tidsskr* **54**: 1–22.

Swofford DL (1993). PAUP. *Phylogenetic Analysis Using Parsimony, Version 3.1.1* Illinois Natural History Survey: Champaign.

Taberlet P, Gielly L, Pautou G, Bouvet J (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Mol Biol* **17**: 1105–1109.

Tibayrenc M, Kjellberg F, Arnaud J, Oury B, Breniere SF, Darde ML *et al* (1991). Are eukaryotic microorganisms clonal or sexual? A population genetic vantage. *Proc Natl Acad Sci USA* **88**: 5129–5133.

Tibayrenc M (1997). Are *Candida albicans* natural populations subdevided? *Trends Microbiol* **5**: 253–254.

Van Baarlen P, Van Dijk PJ, Hoekstra RF, De Jong J (2000). Meiotic recombination in sexual diploid and apomictic triploid dandelions (*Taraxacum officinale* L.). *Genome* **43**: 827–835.

Van Heusden AW, Van Der Voort JR, Bachmann K (1991). Oligo (GATA) fingerprints identify clones in asexual dandelions (*Taraxacum*, Asteraceae). *Fingerprint News* **3**: 13–15.

Van Der Hulst RGM, Mes THM, Den Nijs JCM, Bachmann K (2000). AFLP markers reveal that population structure of triploid dandelions (*Taraxacum officinale*) exhibits both clonality and recombination. *Mol Ecol* **9**: 1–9.

Wilkinson M (1995). PICA95: *Software and Documentation*. School of Biological Sciences: University of Bristol.

Vos P, Hogers R, Bleeker M, Reijans M, Vandelee T, Hornes M *et al* (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* **23**: 4407–4414.

Xu J, Mitchell TG, Vilgalys R (1999). PCR-restriction fragment length polymorphism (RFLP) analyses reveal both extensive clonality and local genetic differences in *Candida albicans*. *Mol Ecol* **8**: 59–73.

Zaykin D, Zhivotovsky L, Weir BS (1995). Exact tests for association between alleles at arbitrary numbers of loci. *Genetica* **96**: 169–178.