Molecular Ecology Resources (2008) **8**, 387–392 doi: 10.1111/j.1471-8286.2007.01967.x

PERMANENT GENETIC RESOURCES

Isolation of 60 polymorphic microsatellite loci in EST libraries of four sibling species of the phytopathogenic fungal complex *Microbotryum*

T. GIRAUD,*¶ R. YOCKTENG,* S. MARTHEY,† H. CHIAPELLO,† O. JONOT,* M. LOPEZ-VILLAVICENCIO,* D. M. DE VIENNE,* M. E. HOOD,‡ G. REFREGIER,* A. GENDRAULT-JACQUEMARD,† P. WINCKER§ and C. DOSSAT§

*ESE, Bâtiment 360, Université Paris-Sud, 91405 Orsay, France, †MIG—INRA Bâtiment 233, Domaine de Vilvert, 78350 Jouy en Josas Cedex, France, ‡Department of Biology, Amherst College, Amherst, MA 01002, USA, §Genoscope—CNRG CP 5706, 91057 Evry Cedex, France

Abstract

We report the development of 60 microsatellite markers on four species of the fungal complex *Microbotryum*, causing anther smut of the Caryophyllaceae. Microsatellites were found in four expressed sequence tag (EST) libraries, built from isolates of *M. lychnis-dioicae*, *M. violaceum sensus stricto*, *M. lagerheimii* and *M. dianthorum*, collected, respectively, from the plants *Silene latifolia*, *S. nutans*, *S. vulgaris* and *Dianthus carthusianorum*. Intrapopulation polymorphism was investigated using 24 isolates, and cross-amplification was explored using 23 isolates belonging to at least 10 different *Microbotryum* species. This study provides numerous microsatellite markers for population genetics and mapping studies.

Keywords: basidiomycete, phytopathogen, repeats, transposable elements, *Ustilago violacea*

Received 17 June 2007; revision accepted 1 August 2007

The anther smut fungus *Microbotryum violaceum* (basidiomycete) causes disease on many plant species in the Caryophyllaceae (Thrall *et al*. 1993). Fungal spores are produced in anthers, and are transmitted by pollinators. It has recently been shown that *M. violaceum* is composed of several cryptic species, highly specialized on a single or a few plant species (Lutz *et al*. 2005; Le Gac *et al*. 2007a). All the species have, however, not been formally named yet.

Microsatellite markers have previously been isolated in the *Microbotryum* species complex, by screening enriched genomic libraries with microsatellite probes, on the sibling species *M. lychnis-dioicae* collected from *Silene latifolia* (Bucheli *et al*. 1998), *M. silenes-inflatae* collected from *Silene vulgaris*, and *M. dianthorum* collected from *Dianthus sylvestris* and *Gypsophila repens* (Giraud *et al*. 2002; Lopez-Villavicencio *et al*. 2005). Available markers are, however, not polymorphic enough for within-population studies (Giraud 2004). Furthermore, not enough markers cross-amplify to enable a genetic mapping of the hybrids that have been produced between several of the sibling species (Le Gac *et al*. 2007b). We therefore searched for new microsatellite loci in expressed sequence tag (EST) libraries, in particular on species of the complex for which no markers had been developed yet.

We constructed cDNA libraries from four isolates of the species complex, belonging to the species *M. lychnis-dioicae*, *M. violaceum sens stricto*, *M. lagerheimii* and *M. dianthorum*, respectively, collected from the plant species *S. latifolia*, *S. nutans*, *S. vulgaris* and *Dianthus carthusionarum* (Yockteng *et al*. 2007). While two distinct species have been described to infect S. *vulgaris*, we used the one called MvSv1 in Le Gac *et al*. (2007a) and *M. lagerheimii* in Denchev (2007). The previous microsatellite development on a strain collected from *S. vulgaris* involved the species called MvSv2 in Le Gac *et al*. (2007b) and *M. silenes-inflatae* in Lutz *et al*. (2005). The strain of *M. dianthorum* used here belongs to the species MvDc identified by Le Gac *et al*. (2007a). The two previous strains of *M. dianthorum* used for microsatellite development belonged to two other *Microbotryum* species: MvDsp1 for the strain on *D. sylvestris* (Le Gac *et al*.

¶Correspondence: T. Giraud, Fax: +33 1 69 15 46 97; E-mail: tatiana.giraud@u-psud.fr

2007a) and to yet another species for the strain on *G. repens* (Lopez-Villavicencio *et al.* 2005; Le Gac *et al.* 2007a). Forty thousand clones were sequenced by the Genoscope Institute (Evry, France) for each library. ESTs were assembled into contigs when the criterion of a minimum identity of 95% over 50 bp was met. When an EST could not be assembled with others in a contig, it remained as a singlet. The contigs and the singlets should thus correspond to unique genomic loci and will be called hereafter 'unisequences'.

In order to find microsatellite loci, we searched the unisequences for all dinucleotide and trinucleotide motifs, repeated perfectly at least eight times, using a text editor. Primers were designed using the PRIMER 3 software (Rozen & Skaletsky 2000). Most of the microsatellites had dinucleotide repeats (83%, Table 1), which was surprising in an EST library, as coding sequences should be more robust to additions and deletions of trinucleotides than dinucleotides. The unisequences were compared to the sequences in public databases using BLASTn and tBLASTx algorithms. Eighteen of the 60 unisequences could be annotated (30%, Table 1). In addition, six unisequences had significant hits against previously cloned microsatellite-containing DNA fragments from *Microbotryum* species (Bucheli *et al.* 1998; Giraud *et al.* 2002), but only restricted parts of the sequences aligned outside of the tandem repeats.

Each microsatellite locus was screened for (i) intrapopulation polymorphism using 24 isolates of the *Microbotryum* species from which they were isolated (Table 2), and (ii) cross-amplification and species differences using a panel of 23 *M. violaceum* isolates (Table 1), belonging to at least 10 different sibling species of the *Microbotryum* complex (Le Gac *et al.* 2007a). Comparison of microsatellite loci between species were also performed on the sequences when available. Interestingly, some microsatellites were highly reduced or completely absent in some species (Table 1). DNA extraction and PCR amplifications and visualization were performed as in Giraud (2004).

Among the 73 tested loci for which primers were designed, all but six yielded PCR fragments of expected size, 60 had at least two alleles, and most of them cross-amplified in several sibling species (Table 1). The marker SVG14 appeared to be useful to distinguish the two *Microbotryum* species infecting *S. vulgaris*, MvSv1 and MvSv2, and the markers DC5 and DC8 to distinguish the two *Microbotryum* species infecting *D. carthusianorum*, MvDc and MvDsp1, which was checked on a larger number of isolates of each species.

Genetic analyses were performed using GENEPOP on the Web (http://genepop.curtin.edu.au/). Few heterozygotes were recorded, leading to significant deviations from Hardy–Weinberg equilibrium for 23 markers out of the 34 polymorphic within populations, in accordance with previous reports of the high level of homozygosity in *Microbotryum* (Giraud 2004). After Bonferroni correction for multiple statistical tests, few pairs of markers were in

Table 1 For each microsatellite locus are indicated repeat motif, GenBank Accession no., annotation when available, primer sequences, amplification size in the original sequence, additional information on the locus (location of the microsatellite in the cDNA, translation, presence or absence in other species), number of alleles in a sample of the *Microbotryum* species in which the locus was originally found, expected heterozygosity H_E, significance of deviation from Hardy-Weinberg equilibrium, number of alleles from all samples tested (from at least 10 different sibling species) and species in which an amplification was obtained, based on the host of collection

Locus*	Repeat motif identified	GenBank accession numbers of the EST included in the contigs	Annotation†	Additional information when available	Primer sequences (5′–3′)	T_m (°C)	Size (bp)	Number of genotyped strains	Nb of detected alleles in the species from which they were isolated‡	H_E	H_O	HW	Total number of detected alleles including sibling species	Cross-amplification§
SL2	$(TG)_{10}$	CU388401			GACTTATCGGTCGCAATGGT GCAAGTTGTTTCATTCAGTCC	58	226	22	3	0.16	0	***	7	sv2, sl, dm, sd, lf, dc
SL3	$(CT)_{16}$	CU370511	*Microbotryum violaceum* Sv6 microsatellite (Giraud *et al.* 2002)	3′ UTR	TTCTGTGAGGCATCCCCTTA GAAATCAATGTGTTCTCCCAGA	60	198	19	3	0.28	0.10	***	4	sd, sl
SL4	$(TC)_{11}T$	CU380976, CU372524, CU383264, CU382751		3′ UTR	CGGCAAAATTAGGAAAGTCG TACAAAACACTCCCGTGCAG	59	175	24	4	0.35	0	***	3	sl
SL7	$(CT)_{15}$	CU388504	*Microbotryum violaceum* L14 microsatellite (Bucheli *et al.* 1998)		ACACTCTAATCACATTATTTCAATCG AGAGGGTCGTCGAAAAGGTT	57	150	24	3	0.16	0.12	*	9	sv1, sv2, sl, dm, ds, lf, sd, sn, sa, sc
SL8	$(GA)_{15}$	CU387006			GCGGTTCTACGAGGTTTCAC TTCTTCTCCTCCCCGAGATT	59	180	19	5	0.38	0.10	***	4	gr, sv1, sl, dm, ds, dc, lf, sd, sa, sc
SL9	$(CT)_{11}$	CU384826, CU381153			CAAGAACAACCCCAAGAGGA CATCACACCTCGTTCGACAC	60	152	24	2	0. 14	0	***	8	gr, sv1, sl, dm, ds, dc, lf, sd, sc
SL10	$(CT)_{11}$	CU392005, CU375936		3′ UTR	TAAGATCGTCCCCAGTCGAC TTCACTCCCTCTCCCTTTCTC	59	181	24	1	0	0	—	7	gr, sv2, sl, dm, ds, dc, lf, sd, sn, sa, sc
SL11	$(GA)_6T(AG)_{13}T_3$	CU368079, CU390547	Csf1p *Saccharomyces cerevisiae*	3′ UTR $(GA)_{18}$ in Sn, $(GA)_{10}$ in Sv, $(GA)_{14}$ in Dc	GGCTAGGTCTCATCACGTGA CAATGCACCCGAGTTCTCTC	58	158	22	3	0.28	0.09	***	11	sv1, sv2, sl, dm, ds, dc, lf, sd, sn, sa, sc

Table 1 *Continued*

Locus*	Repeat motif identified	GenBank accession numbers of the EST included in the contigs	Annotation†	Additional information when available	Primer sequences (5′–3′)	T_m (°C)	Size (bp)	Number of genotyped strains	Nb of detected alleles in the species from which they were isolated‡	H_E	H_O	HW	Total number of detected alleles including sibling species	Cross-amplification§
SL12	$(GT)_{10}$	CU385694, CU387385, CU380921, CU380563, CU374335, CU375617	Q5KNS5_CRYNE VpsA, putative	3′ UTR	CCTTTCTACACCCCGACACT TTATCTATTGGTCAAGATCATGCT	57	199	22	3	0.25	0.14	***	5	gr, sv1, sv2, sl, dm, ds, dc, lf, sd, sn, sa, sc
SL13	$(TG)_{14}C(GT)_4$	CU369126, CU391972		3′ UTR	GACTTGCTTCTCAACCCACC TCTTGGTGGCATGAATATTCC	59	199	24	1	0	0	—	5	sl, ds, lf, sd,
SL14	$(GGT)_8$	CU377528, CU380886, CU389514, CU380883, CU385280, CU391406	Q2CQG5_ECTHAc Chorismate mutase, gamma-, beta- and epsilon-proteobacteria	Translated in GGGGG (-COOH term region, after the region similar to chorismutase)	GAATCGACTTGTGCTCGTGA AGGAAGGGGTTCAGCAAACT	60	197	24	1	0	0	—	9	gr, sv1, sv2, sl, dm, ds, dc, lf, sd, sn, sa, sc
SL15	$G_7(AAG)_8AGGA$	CU372347, CU387139		3′ UTR	TAGGGACGCAAGATTAGCTTG CGAAGAGCGAGGCTTTTGTA	60	241	24	1	0	0	—	7	gr, sv1, sv2, sl, dm, ds, dc, lf, sd, sn, sa, sc
SL16	$(AAG)_{11}$	CU385505, CU370268			CCTCTCCCGTTGAGATTGTC TGTTGCTCATGCTTTTGCAT	59	197	20	2	0.13	0	***	12	sv1, sv2, sl, dm, ds, dc, lf, sd, sn, sa, sc
SL19	$(AAC)_3AAA(AAC)_{12}$	CU376837		Translated in NNNNNNNNNNNN	TTGCTACGCCTCGCAAAG ATGAGCCTTCCCGGTCTT	59	250	24	2	0.23	0.20	***	5	sv1, sv2, sl, dc, sd, sc
DC1	$(CA)_{12}$	CU457645, CU457678, CU457619, CU457526, CU457698, CU457514			GTTCCTCGGTAGTGCTCTCG CACAAAAATGAAGCCATGAGA	58	199	20	2	0.28	0	***	2	sv1, sv2
DC2	$(TG)_{17}$	CU457702		3′ UTR	TTTGATGCAGCCTGTGGTAG CTCCCCACATTGTAGGTTCC	59	173	22	2	0.43	0.09	***	4	gr, sv1, sv2, sl, dm, ds, dc, lf, sd, sn,
DC5	$(CT)_{11}$	CU457516, CU457691		3′ UTR	AGCATAGCCCTCAGCACACT AAGTGATTGGCAGTATCGGACT	60	107	22	4	0.43	0.59	NS	5	gr, sv2, sl, dm, ds, dc, lf, sd, sn
DC6	$(TC)_2CT(TC)_{19}$	CU457651		3′ UTR $(TC)_4$ in Sl, $(TC)_4$ in Sv, $(TC)_7$ in Sn	TGTTTCCGGTGACTTTCCTC TCAACATTGCAGCTTTTTGG	59	136	20	2	0.35	0.10	***	6	gr, sv1, sv2, sl, dm, ds, dc, lf, sd, sn, sc
DC7	$T_{4(}CT)_{14}$	CU457657	Q6DRC1_BRARE Small nuclear ribonucleoprotein F	3′ UTR	GGAAGGGGAACGAAGTAAGG GGGTTTGAGAGAACTGTGTT	54	138	20	3	0.41	0	***	5	gr, sv1, dm, ds, dc, lf, sd, sn
DC8	$(TC)_{12}A_8G_7A_3CA_7$	CU457657		Translated in SLSLSLSQKKGGETKK	CCCTTCACAAAACTCCCAAA ATCGCTGGAACCGTTCATT	59	138	20	3	0.48	0	***	2	gr, sv1, sv2, sl, dm, ds, dc, lf, sd, sn, sa, sc
DC9	$T_4+C_4(TC)_2C(TC)_8T_2(TC)_2$	CU457618		3′ UTR	TCATGAGCACGGAGTTTCTG ATGTGAACACCCACCATCCT	60	128	20	3	0.22	0	***	9	sv, sl, dm, ds, lf, sd, sa, sc
DC10	$(TC)_9$	CU457530	*Microbotryum violaceum* microsatellite sequence L17 (Bucheli *et al.* 1998)	3′ UTR	GTGGGCCTTCTGTCGTTAGT AGAAACGGGCCAAATTCAC	59	138	24	5	0.73	1.00	***	8	gr, sv1, sv2, sl, dm, ds, dc, lf, sd, sn, sc
DC11	$TTT(TC)_{15}TT$	CU457537, CU457699, CU457641, CU457556, CU457571, CU457673, CU457670	Q6DRC1_BRARE Small nuclear ribonucleoprotein F		CCTTACTTCGTTCCCCTTCC ATACAAAAGGTCACCCCTCG	58	160	24	3	0.46	0.71	***	7	gr, sl, dm, ds, dc, lf, sd, sn
DC12	$(TG)_{10}$	CU457536, CU457548	EAL85594 phosphatidyl synthase	3′ UTR	TGACATTCCATCATAACCTG TGACATTCCATCATAACCTG	54	168	24	1	0	0	—	2	gr, sv1, sv2, sl, dm, ds, dc, lf, sd, sn, sa, sc
SVG1	$(TG)_8CG(TG)_5$	CU457635, CU457644, CU457587		3′ UTR	AGGCTAGCAAGCTTTGAGGA TTCGCGACTGATGTCAACTT	59	143	24	2	0.19	0.37	NS	3	sv1, sv2, sl, ds, dc, sd, sn, sc
SVG2	$(TG)_7$	CU457668, CU457542, CU457652			CAAGCTGTGGGACGCTATCT CGACACGATGGGAGAAAAAT	60	144	22	3	0.27	0.31	NS	5	gr, sv1, sv2, sl, dm, ds, dc, lf, sd, sn, sa, sc
SVG3	$(GT)_{10}$	CU457653			TTCGGATACTCTGCGCTTTT TTCTTCTTCAAGCACGCAAA	59	149	24	1	0	0	—	11	sv1, sv2, sl, dm, ds, dc, lf, sd, sn, sa, sc
SVG4	$(GT)_{11}$	CU457607, CU457637	Q59P43_CANAL RAN-like GTP binding protein	3′ UTR	CGTTGCCTGATGAAGAGGAT TTGGGCTCATCATACGAGGT	60	146	24	1	0	0	—	6	gr, sv1, sv2, sl, dc, lf, sa

Table 1 *Continued*

Locus*	Repeat motif identified	GenBank accession numbers of the EST included in the contigs	Annotation†	Additional information when available	Primer sequences (5′–3′)	T_m (°C)	Size (bp)	Number of genotyped strains	Nb of detected alleles in the species from which they were isolated‡	H_E	H_O	HW	Total number of detected alleles including sibling species	Cross-amplification§
SVG5	$(TG)_8$	CU457654, CU457665, CU457610, CU457600, CU457557, CU457565, CU457630, CU457578, CU457640, CU457636, CU457527, CU457689, CU457551	XP_756025 60S ribosomal protein L37a	3′ UTR $(TG)_{11}$ in Dc, $(TG)_7$ in Sl, $(TG)_6CG(TG)_2$ in Sn	GCACTCCCATCACGTAGTCA CGTCCTCCCCATCTTCTTCT	59	148	22	2	0.04	0	***	9	gr, sv1, sv2, sl, dm, ds, dc, sd, sn, sa, sc
SVG6	$(GT)_9$	CU457543, CU457532	*Microbotryum violaceum* microsatellites GR22, GR11 and GR14 (Giraud *et al.* 2002)	3′ UTR	GGTCAAAGCAGGAGAAGCAG GGCCTATGATCATTTCATCCA	59	130	24	1	0	0	—	6	sv2, sl, dc, lf, sd
SVG7	$(TG)_{10}$	CU457554			ATACGACCTTCGGCTTGAAA GCGGGGTCTATTCAATTCAT	58	106	24	1	0	0	—	2	sl, dc, sd, sa
SVG8	$(GT)_{12}$	CU457553		3′ UTR $(GT)_4$ in Sn	TTTCATCATGTGCTGCTTCC ACGCAAAGGCACAAGAAAAC	59	123	22	3	0.1	0.09	NS	13	gr, sv1, sv2, sl, dm, ds, dc, lf, sd, sn, sa, sc
SVG9	$(TG)_8$	CU457544		translated in VCVCVC	CGGAATAGCGATGATGAGGT AAAAAGCACACCACACACCA	60	142	24	1	0	0	—	2	sv2, sl, lf, sd
SVG10	$(TC)_8$	CU457683, CU457503, CU457605, CU457597, CU457539, CU457692, CU457520, CU457522		3′ UTR	GGATTTTCCTCTGGTTGTGG GGGATTCGGACAAGGTGTAG	59	121	24	1	0	0	—	2	gr, sv1, dm, ds, lf, sa
SVG11	$(TC)_8$	CU457666, CU457575, CU457693, CU457569	*Microbotryum violaceum* microsatellite sequence	3′ UTR	GGGTGGGTCTCTTGTCGTTA GGTGTCCACCAATCCGATAC	59	145	24	2	0.11	0.04	***	10	gr, sv1, sv2, sl, dm, ds, dc, lf, sd, sn, sa, sc
SVG13	$(AC)_8$	CU457531, CU457510	L17 (Bucheli *et al.* 1998)		GCTGCCTCGACATGATCTTC GCCGAAGGCGATACAAATAC	59	141	22	2	0.08	0.09	NS	3	sv1, sv2, sl, dc, lf, sa, sc
SVG14	$(CTC)_2(TTC)_{10}$	CU457540	CAA45132 acetylglutamate kinase	3′ UTR $(CTC)_3(TTC)_6$ in Sl, $(CTC)_3(TTC)_4$ in Sn, $(CTC)_3(TTC)_5$ in Dc	ATCCACCACCAACAAAGCAT GCGTTGCAAGTTTCTACGAT	58	108	24	1	0	0	—	8	sv1, sv2, sl, dm, ds, lf, sd, sn, sa, sc
SVG15	$G_7(GAA)_{14}$	CU457577, CU457621		translated in E_{14}	CACGAGGAGGAAGAAGAAGG CTCATACGTGGCAGGACCTC	59	141	24	1	0	0	—	6	sv1, sv2, sl, ds, dc, lf, sd, sn, sc
SVG16	$(GAG)_8GAA(GAG)_4$	CU457593, CU457659, CU457611, CU457583	EAZ63041 ubiquinol-cytochrome c oxidoreductase subunit 6	translated in E_{13}, $(GAG)_5GAA$ $(GAG)_2GAA(GAG)_2$ in Sn	GACGAGGACAAGGCTGAAG TGGTCGAAATGGTGCTTGTA	58	140	24	2	0.18	0	***	4	sv1, sv2, sl, dm, ds, dc, lf, sd, sa, sc
SN1	$(TG)_8$	CU457662, CU457680, CU457679		3′ UTR	GAGCGACGATGAAGGGTAGA GCCTGTCCGAAAAGTTTTGT	59	143	24	2	0.23	0	***	6	sv1, sv2, sl, ds, lf, sn
SN2	$(TG)_8(TGGG)_3$ TGT	CU457669			CTCGTCTCCTTTACCCACCA GCGCAAAATCAATCATCAGA	59	106	24	1	0	0	—	4	sv1, sv2, sl, ds, lf, sc
SN3	$(GT)_{12}G$	CU457617, CU457546, CU457620, CU457549, CU457576, CU457675, CU457590	EDK45713 protein translation factor SUI1	3′ UTR $(GT)_6G$ in Sl, Sv, and Dc	CGAAAGGAATGTCTGCCTGT GGAGCGAACAATTGATGACA	59	136	23	2	0.12	0	***	2	sv1, ds, sc
SN4	$(GT)_{10}G$	CU457534, CU457696, CU457690	Q5KIG3_CRYNE Cytoplasm protein, putative	3′ UTR	AAATCGGGTTTATGTGGAAGG GAGGGGTTTTTGTATAGCCTCT	57	116	24	1	0	0	—	7	sv1, sv2, sl, dc, lf, sn
SN5	$(GT)_{10}G$	CU457594, CU457626, CU457509		translated in VCVCVC	TTTGCATTCCGAATTGTTCA AAAGTGGGGTAGGCGAAGAT	60	113	24	2	0.12	0.08	NS	6	gr, sv1, sv2, sl, dm, ds, lf, sd, sa
SN6	$(GT)_8$	CU457625, CU457572, CU457592, CU457566	Q5KNS5_CRYNE VpsA, putative, 694 bp	3′ UTR	AACGGGTTGAATGGGTTCTT CGGTTTGAGAAGAGGATA	50	132	24	1	0	0	—	5	sv1, sl, dm, ds, lf
SN7	$(GT)_5CT(GT)_9+$ $(T)_8$	CU457646			TGTGTGTGTGATTTTGACTTTTCA GATCTCTGAACAAGCTCTCACAA	58	148	24	1	0	0	—	11	sv1, sl, dm, ds, sd, sa, sc

Table 1 *Continued*

Locus*	Repeat motif identified	GenBank accession numbers of the EST included in the contigs	Annotation†	Additional information when available	Primer sequences (5′–3′)	T_m (°C)	Size (bp)	Number of genotyped strains	Nb of detected alleles in the species from which they were isolated‡	H_E	H_O	HW	Total number of detected alleles including sibling species	Cross-amplification§
SN8	$(GT)_7GC(GT)_4$	CU457581		3′ UTR	CACCTATCCGGACTTTTTGC AAATTTCACGGAAACTATTCA	54	121	24	1	0	0	—	5	sv1, sv2, sl, dm, ds, dc, lf, sd, sc
SN9	$(TG)_8$	CU457588	Q875L4_USTMA Small G-protein Ras1	3′ UTR	CCTTTTCTTGTTGCGTCGTT AAAAGAGGAAGGACGAGCTG	58	128	24	1	0	0	—	6	sv1, sv2, dm, ds, lf, sn, sa, sc
SN10a	$(GT)_4GC(GT)_{13}(GT)_5$	CU457588		3′ UTR	GAGCGAAGAAAAAGGGTGTG ATGCCACACGCACATAATTG	59	131	24	1	0	0	—	4	sv1, sv2, sl, lf, sd, sc
SN10b	$(CTT)_3TCTTC(CTT)_3C(CTT)_2$	CU457588			CGGGGCTGTTTCTCTTCTTC CGGATTGATAGTTTCATATGTCG	58	146	24	1	0	0	—	5	sv1, sv2, sl, ds, dc, lf, sd, sa,
SN11	$(TC)_8$	CU457579	*Microbotryum violaceum* microsatellites GR4		AAGTCAGCCCTCAGCACACT AGTGACTGGCAGTATGGGATG	60	102	24	1	0	0	—	2	sv1, sv2, sl, dc, sd,
SN12	$(TC)_8$	CU457580, CU457622, CU457672, CU457704	Q764D2_LENED Putative S-phase specific ribosomal protein cyc07	3′ UTR $(TC)_2$ in Sl	GCGAGACAGAAGGCAACCTA ACTACACACGCCAAGGAACC	60	138	24	1	0	0	—	2	sv1, dm, lf
SN13	$(CT)_{12}$	CU457550, CU457552, CU457507, CU457582, CU457684, CU457525, CU457568, CU457664, CU457560, CU457545, CU457688	BAF56680 developmental regulator Le-CDC5 partner	$(CT)_7$ in Sl and Dc, nothing in Sv	TAGGAGAGGATCGGGGTTTC AGCTATGGCCCAACCTCTTT	60	142	24	1	0	0	—	2	sv1, ds
SN15	$T_3(CT)_9GT(CT)_4CGT(CT)_2$	CU457564, CU457561		$(CT)GT(CT)_{16}CGT(CT)$ in Dc, $(CT)_3$ in Sl	TGCCAAACTGCATCTTTCTC GCAAAGCGAACGAGCTTCTA	59	138	24	3	0.24	0	***	5	sv1, sv2, sl, dm, ds, lf, sd, sa
SN16	$(TC)_8$	CU457676		3′ UTR	TCGGACGAACTCTGGTAGC CGGTATTTCTACAGAAAGGACTGC	59	150	24	1	0	0	—	14	sv1, sv2, sl, dm, ds, lf, sd, sa
SN17	$(TC)_8$	CU457658	Q5K7L0_CRYNE Heat-shock protein, putative	3′ UTR	AGGTCAAGCCAAAATGAACG GCGTCGATGTAAGAGCACAA	60	145	24	1	0	0	—	7	sv1, sl, dm, ds, sd, sc
SN18	$(TC)_8$	CU457649			CGTCCAGGTCTTCATTGCTT TGTAAGGTGGTGACGGATTG	59	142	24	1	0	0	—	3	sv1, dm, ds, dc
SN19	$(AC)_{15}A$	CU457529			TGCGACACTCGATAGACAGC ACTCGAACTACAACGCACGA	59	110	24	2	0.25	0.12	NS	7	gr, sv1, sv2, dm, ds, lf, sa, sc
SN20	$(AT)_7$	CU457512, CU457612	Q6PW77_ACRST Glucooligosaccharide oxidase	3′ UTR	TTGACGATATCGCATCTTCG GGATGATCCCATCGATTCC	59	139	24	1	0	0	—	2	sv2, dc, sa, sc
SN21	$(TTC)_{14}$	CU457586			GGAAAGGTCGTTTCTGC AACCCCGACATCGTAAAGAA	54	146	24	5	0.25	0.04	***	5	sv1, sv2, sl, dc, sd, sa, sc
Mean ± SD									1.95 ± 1.26	0.14 ± 0.17			5.90 ± 3.00	

*Loci named SL, DC, SVG and SN were cloned respectively in the libraries from *S. latifolia*, *D. carthusianorum*, *S. vulgaris* and *S. nutans*. The loci SN10a and SN10b were located in the same unisequence and are therefore tightly linked.
†Microsatellite loci indicated in this column mean that part of the sequence was significantly similar to a previously developed marker, but not the whole sequence, meaning that they represent different loci.
‡Except for the SN markers, tested on the populations collected from *Silene vulgaris*.
§Sibling species for which amplification products were obtained are indicated by the initials of their host plant:gr = *Gypsophila repens*, sv = *Silene vulgaris* (sv1 correspond to the species MvSv1 and sv2 to the species MvSv2), sl = *Silene latifolia*, dm = *Dianthus monspessulanus*, ds = *Dianthus sylvestris*, dc = *Dianthus carthusianorum*, lf = *Lychnis flos-cuculi*, sd = *Silene dioicea*, sn = *Silene nutans*, sa = *Silene acaulis*, sc = *Silene caroliniana*.

Table 2 Populations used for assessing within-population polymorphism for each of the four *Microbotryum* species

Species	Population	Host plant	Location	Number of individuals	
M. lychnis-dioicae	4	*Silene latifolia*	Essone region, France	12	24
M. lychnis-dioicae	7			12	
M. violaceo-irregulare	302	*Silene vulgaris*	French Pyrenees	13	24
M. violaceo-irregulare	431		Swiss Alps	11	
M. dianthorum	432	*Dianthus carthusianorum*	Swiss Alps	24	24

linkage disequilibrium: SL8 and SL12, SL11 and SL19, DC2 and DC8, SN5 and SN19, and SN3 and SN21.

The alleles of the markers containing dinucleotide motifs generally spanned a wide range of sizes and were not restricted to repeat number variation divisible by three. The BLASTx alignments or the detection of a long coding frame (more than 110 amino acids) showed that 34 microsatellites were located in the 3′ untranslated regions (Table 1), but seven were translated (Table 1).

Acknowledgements

This work was funded by the 'Consortium National de Recherche en Génomique' for sequencing the libraries. GR acknowledges a post-doc grant from the Foundation des Treilles.

References

Bucheli E, Gautschi B, Shykoff JA (1998) Isolation and characterization of microsatellite loci in the anther smut fungus *Microbotryum violaceum*. *Molecular Ecology*, **7**, 665–666.

Denchev CM (2007) Microbotryum lagerheimii sp. nov. (microbotryaceae). *Mycologia Balcanica*, **4**, 61–67.

Giraud T (2004) Patterns of within population dispersion and mating of the fungus *Microbotryum violaceum* parasitizing the plant *Silene latifolia*. *Heredity*, **93**, 559–565.

Giraud T, Fournier E, Vautrin D *et al*. (2002) Isolation of 44 polymorphic microsatellite loci in three host races of the phytopathogenic fungus *Microbotryum violaceum*. *Molecular Ecology Notes*, **2**, 142–146.

Le Gac M, Hood ME, Fournier E, Giraud T (2007a) Phylogenetic evidence of host-specific cryptic species in the anther smut fungus. *Evolution*, **67**, 15–26.

Le Gac M, Hood ME, Giraud T (2007b) Evolution of reproductive isolation within a parasitic fungal complex. *Evolution*, **61**, 1781–1787.

Lopez-Villavicencio M, Enjalbert J, Hood ME *et al*. (2005) The anther smut disease on *Gypsophila repens*: a case of parasite sub-optimal performance following a recent host shift? *Journal of Evolutionary Biology*, **18**, 1293–1303.

Lutz M, Goker M, Piatek M *et al*. (2005) Anther smuts of Caryophyllaceae: molecular characters indicate host-dependent species delimitation. *Mycological Progress*, **4**, 225–238.

Rozen S, Skaletsky HJ (2000) Primer 3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (eds by Krawetz S, Misener S), pp. 365–386. Humana Press, Totowa, New Jersey.

Thrall PH, Biere A, Antonovics J (1993) Plant-life history and disease susceptibility—the occurrence of *Ustilago violacea* on different species within the Caryophyllaceae. *Journal of Ecology*, **81**, 489–498.

Yockteng R, Marthey S, Chiapello H *et al*. (2007) Expressed sequence tags of the anther smut fungus, *Microbotryum violaceum*, identify mating and pathogenicity genes. *BMC Genomics*, **8**, 272.