Isolation of Nuclear Microsatellite Markers for Cyperus fuscus (Cyperaceae)

Authors: Böckelmann, Jörg, Wieser, David, Tremetsberger, Karin, Šumberová, Kateřina, and Bernhardt, Karl-Georg

Source: Applications in Plant Sciences, 3(11)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1500071

Applications in Plant Sciences 2015 3(11): 1500071

PRIMER NOTE

ISOLATION OF NUCLEAR MICROSATELLITE MARKERS FOR *CYPERUS FUSCUS* (CYPERACEAE)[1]

JÖRG BÖCKELMANN[2,4], DAVID WIESER[2], KARIN TREMETSBERGER[2], KATEŘINA ŠUMBEROVÁ[3], AND KARL-GEORG BERNHARDT[2]

[2]Institute of Botany, Department of Integrative Biology and Biodiversity Research, University of Natural Resources and Life Sciences, Vienna, Gregor-Mendel-Straße 33, 1180 Vienna, Austria; and [3]Department of Vegetation Ecology, Institute of Botany, Czech Academy of Sciences, Lidická 25/27, 602 00 Brno, Czech Republic

- *Premise of the study:* Microsatellite markers were characterized in the extremely specialized ephemeral wetland plant species *Cyperus fuscus* (Cyperaceae). The markers will be used for studying population genetics in natural vs. anthropogenic habitats, on a European scale, and the role of the soil seed bank in the life cycle of this ephemeral species.
- *Methods and Results:* Twenty-one microsatellite loci were established and scored in two populations, with mean number of alleles of 2.6 and 2.9 and mean expected heterozygosity of 0.405 and 0.470, respectively. Forty-four additional loci with the number of alleles ranging from one to four (mean = 2.1) were successfully amplified in seven individuals.
- *Conclusions:* The novel microsatellite markers will be useful for studying the genetic structure of populations of this ephemeral plant as well as their seed bank.

Key words: 454 sequencing; Cyperaceae; *Cyperus fuscus*; Isoëto-Nanojuncetea; microsatellites.

Cyperus fuscus L. (Cyperaceae) is an annual herb that is native in the Mediterranean region and temperate Eurasia and introduced in North America. It grows on muddy, sandy, or gravelly substrata, on shores of rivers or lakes, and is also found in anthropogenic habitats like gravel pits, wet fields, and traditionally used fish ponds. It has a short life cycle, taking just two to three months from seedling to ripe fruits (von Lampe, 1996).

Cyperus fuscus is anemophilous and self-compatible. With 0.24 pg/1C (or 234.72 Mbp; Doležel et al., 2003), the genome is relatively small (Tremetsberger et al., unpublished data). Plants with $2n = 36$ and 72 chromosomes are known (Krahulcová, 2003), most probably corresponding to diploid and tetraploid cytotypes (Roalson, 2008). The large amounts of seeds produced build up a persistent soil seed bank, which can also function as a "genetic memory" by storing the genetic variability in viable seeds (Leck, 1989). We developed 21 microsatellite markers to compare the genetic variation in the seed bank of various natural and manmade habitats.

METHODS AND RESULTS

Plants were grown in the greenhouse from ripe seeds collected in the field (Appendix 1). Genomic DNA of fresh leaves from one plant was extracted with the DNeasy Plant Mini Kit (QIAGEN, Hilden, Germany) following the manufacturer's instructions and sent to LGC Genomics (Berlin, Germany) for next-generation sequencing (NGS) on a Genome Sequencer FLX Titanium Instrument (454 Life Sciences, a Roche Company, Branford, Connecticut, USA). In this first run, 143,027 sequence reads with an average length of 238 bp were obtained (Table 1). NGS data are deposited in the GenBank Sequence Read Archive (BioProject no. PRJNA275048). MSATCOMMANDER version 0.8.2 (Faircloth, 2008) was used to detect 520 sequences with simple sequence repeat (SSR) motifs (options: dinucleotide repeats ≥10 repeat units, tri- and tetranucleotide repeats ≥6 repeat units, combine multiple arrays within a sequence if within 50 bp distance). Primers for microsatellite-containing sequences were also designed in MSATCOMMANDER using Primer3 (Rozen and Skaletsky, 1999), with a GTTT PIG-tail (Brownstein et al., 1996) added to the 5′ end of one primer and a CAG or M13R tail (CAG: 5′-CAGTCGGGCGTCATCA-3′; M13R: 5′-GGAAACAGCTATGACCAT-3′) added to the 5′ end of the other primer (Schuelke, 2000). Due to the shortness of the sequences (range = 7–762 bp, mean = 238 bp), only 101 out of the 520 SSR-containing sequences were suitable for primer design. PCR amplifications were performed in a 25-µL final volume of REDTaq ReadyMix PCR Reaction Mix (Sigma-Aldrich, St. Louis, Missouri, USA) with 0.40 µM 5′ FAM-labeled universal CAG or M13R primer, 0.40 µM GTTT-tailed primer, 0.04 µM CAG- or M13R-tailed primer, and 1 µL diluted DNA extract (2–20 ng DNA). Reactions were performed using a touchdown PCR protocol in an Eppendorf Mastercycler gradient (Eppendorf, Hamburg, Germany), with an initial 5 min of denaturation at 95°C; 24 cycles with denaturation at 95°C for 45 s, annealing at 63–48.6°C (0.6°C decrease per cycle) for 90 s, and extension at 72°C for 60 s; 19 cycles with denaturation at 95°C for 45 s, annealing at 50°C for 90 s, and extension at 72°C for 60 s; and a final extension at 72°C for 5 min and 60°C for 30 min. Amplified fragments were analyzed on a 3500 Genetic Analyzer (Applied Biosystems, Foster City, California, USA) and sized using GeneMarker 2.4 (SoftGenetics, State College, Pennsylvania, USA). The markers were tested on seven individuals from different localities (Appendix 1). Seven loci could be unambiguously scored in all seven test individuals. Four of these were applied to a larger number of individuals (primers with the prefix Cf in Table 2; remaining loci are shown in Appendix 2).

A second NGS run of an SSR-enriched library was performed at ecogenics (Balgach, Switzerland), starting from a mix of genomic DNA of two individuals (Appendix 1). Size-selected fragments from genomic DNA were enriched for SSR content by using magnetic streptavidin beads and biotin-labeled CT, GT, AAG, and ATGT repeat oligonucleotides. The SSR-enriched library

[1]Manuscript received 17 June 2015; revision accepted 15 July 2015.

The authors thank K. Bubíková (Brno), Z. Hroudová, P. Kúr, S. Píšová (Prague), and Z. Kącki (Wrocław) for assistance collecting in the field; K. Takayama (Tokyo) and A. Buser (Balgach) for advice in marker development; and K. Fohringer and G. Kohl (Vienna) for help in the laboratory. Financial support was provided by the Austrian Science Fund (FWF, project P 24558-B16), the Czech Science Foundation (14-36079G), and the Institute of Botany of the Czech Academy of Sciences (project RVO 67985939).

[4]Author for correspondence: joerg.boeckelmann@boku.ac.at

doi:10.3732/apps.1500071

Table 1. Characteristics of the two 454 GS FLX Titanium sequencing runs.[a]

Sequencing run	Total no. of reads	Range of read lengths (bp)	Average read length (±SD; bp)	GC content (%)	SSR-containing sequences (total no. of SSRs encountered)	No. of reads useful for primer design
First run	143,027	7–762	238 (±130)	40.2	520 (539)	101
Second run	4877	34–801	415 (±165)	40.7	967 (990)	494

Note: SD = standard deviation.

[a] In the first run, a crude extract of genomic DNA of a single *Cyperus fuscus* individual was used. In the second run, an enriched library, generated from genomic extracts of two *C. fuscus* individuals, was used. See Appendix 1 for origin of sequenced individuals.

was analyzed on a Roche 454 platform using the GS FLX Titanium reagents (454 Life Sciences, a Roche Company). In total, 4877 reads with a mean length of 415 bp were obtained and deposited in the GenBank Sequence Read Archive (BioProject no. PRJNA275048), of which 967 contained SSR motifs (MSATCOMMANDER search and primer design settings same as above; Table 1). Four hundred ninety-four reads were suitable for primer design. Ecogenics sent 80 primer pairs also designed with Primer3, containing an M13 tail at the 5′ end of the forward primer (5′-TGTAAAACGACGGCCAGT-3′; Schuelke, 2000) and no PIG-tail. For primer testing, the concentrations and volumes for PCR were the same as above, but we used JumpStart REDTaq ReadyMix Reaction Mix (Sigma-Aldrich) and a regular PCR protocol, with an initial 5 min of denaturation at 95°C; 38 cycles of denaturation at 95°C for 45 s, annealing at 56°C for 60 s, and extension at 72°C for 1 min; and a final extension at 72°C for 5 min and 60°C for 30 min. Of these 80 markers, 22 showed no PCR product or had a weak signal, failures, or were unspecific. The remaining 58 markers showed clear peaks. Ten of these were monomorphic and 48 polymorphic. Seventeen polymorphic markers were selected for further analysis and combined into four multiplex PCRs with Multiplex Manager version 1.0 (Holleley and Geerts, 2009; PCR multiplex sets 1–4 in Table 2). The remaining loci are shown in Appendix 2. For application of PCR multiplex sets 1–4 to a larger number of individuals, a GTTT PIG-tail was added to the reverse primers (as for primers with the prefix Cf). For multiplex PCR reactions, the forward primers were directly labeled with a fluorescent dye at the 5′ end (Table 2).

The 21 newly developed microsatellite markers were applied to 25 individuals from each of two fish pond populations in the Czech Republic (Appendix 1). Interpretation of electropherograms in all loci and all individuals is compatible with a diploid cytotype. The number of alleles, observed (H_o) and expected heterozygosity (H_e), fixation index, and exact test for Hardy–Weinberg equilibrium (HWE) were calculated with Arlequin version 3.5.1.3 (Excoffier and Lischer, 2010). The mean number of alleles per locus is 2.6 in Zahrádky and 2.9 in Libohošt (Table 3). H_o ranges from 0 to 0.32 (mean = 0.109) in Zahrádky and from 0 to 0.24 (mean = 0.135) in Libohošt. H_e ranges from 0.078 to 0.706 (mean = 0.405) in Zahrádky and from 0.040 to 0.667 (mean = 0.470) in Libohošt. Deviation from HWE is very high in most loci in both populations, with fixation indices ranging from 0.341 to 1 (mean = 0.756) in Zahrádky and from 0 to 1 (mean = 0.687) in Libohošt.

CONCLUSIONS

The 21 polymorphic loci developed in this study will be useful for studying genetic diversity of *C. fuscus* and the role of the soil seed bank in the life cycle of this ephemeral plant in natural and anthropogenic habitats. The inbreeding coefficients of the two tested populations attest to the very high selfing rate of this species.

LITERATURE CITED

Brownstein, M. J., J. D. Carpten, and J. R. Smith. 1996. Modulation of non-templated nucleotide addition by *Taq* DNA polymerase: Primer modifications that facilitate genotyping. *BioTechniques* 20: 1004–1010.

Doležel, J., J. Bartoš, H. Voglmayr, and J. Greilhuber. 2003. Nuclear DNA content and genome size of trout and human. *Cytometry Part A* 51A: 127–128.

Excoffier, L., and H. E. L. Lischer. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10: 564–567.

Faircloth, B. C. 2008. MSATCOMMANDER: Detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources* 8: 92–94.

Holleley, C. E., and P. Geerts. 2009. Multiplex Manager 1.0: A cross-platform computer program that plans and optimizes multiplex PCR. *BioTechniques* 46: 511–517.

Krahulcová, A. 2003. Chromosome numbers in selected monocotyledons (Czech Republic, Hungary, and Slovakia). *Preslia* 75: 97–113.

Leck, M. A. 1989. Wetland seed banks. *In* M. A. Leck, V. T. Parker, and R. L. Simpson [eds.], Ecology of soil seed banks, 283–305. Academic Press, San Diego, California, USA.

Roalson, E. H. 2008. A synopsis of chromosome number variation in the Cyperaceae. *Botanical Review* 74: 209–393.

Rozen, S., and H. Skaletsky. 1999. Primer3 on the WWW for general users and for biologist programmers. *In* S. Misener and S. A. Krawetz [eds.], Methods in molecular biology, vol. 132: Bioinformatics methods and protocols, 365–386. Humana Press, Totowa, New Jersey, USA.

Schuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18: 233–234.

Von Lampe, M. 1996. Wuchsform, Wuchsrhythmus und Verbreitung der Arten der Zwergbinsengesellschaften. Dissertationes Botanicae, vol. 266. J. Cramer, Berlin, Germany.

TABLE 2. Characteristics of 21 SSR loci developed in *Cyperus fuscus*.

Locus[a]	Primer sequences (5′–3′)[b]	PCR multiplex set	Fluorescent dye[c]	Repeat motif	*A*	Allele size range (bp)[d]	EMBL accession no.
Cf_008	F: GGAAACAGCTATGACCATAGATAATTAACGGATCAGGGACG R: GTTTGAGACAGATTACTCACCTCTCAAG M13R: GGAAACAGCTATGACCAT	NA	ATTO 565	$(AG)_{11}$	4	312–344	LN848930
Cf_017	F: GGAAACAGCTATGACCATGAGGCAATAGAAATTGTTGGAG R: GTTTACGAAATGAGGAGCCATAACTG M13R: GGAAACAGCTATGACCAT	NA	ATTO 550	$(CTTT)_{13}$	3	218–242	LN848931
Cf_019	F: GTTTAATTGTCAGGCCACATGCC R: GGAAACAGCTATGACCATACAGGGAGCAACCTGAGC M13R: GGAAACAGCTATGACCAT	NA	FAM	$(CTT)_7 + (CTT)_6$	2	184–205	LN848932
Cf_104	F: GGAAACAGCTATGACCATGACAGAAGATGAATTAAGGCCAC R: GTTTCGATGACAGTTTAAAGGTCCAG M13R: GGAAACAGCTATGACCAT	NA	Yakima Yellow	$(GT)_{14}$	2	180–184	LN848934
Cypfus_0173	F: CGCCAAAGGAGAATGAGGTG R: GTTTATCGAACAATCCGATCTCGC	1	ATTO 532	$(GAA)_9$	3	189–201*	LN848937
Cypfus_0551	F: TTGCCACATTGACGCACAC R: GTTTAGCGTGCTATTTACAACCTTGG	1	ATTO 565	$(TGTA)_9$	2	205–229*	LN848938
Cypfus_1207	F: ATCTCTTCACTCCCGCCATC R: GTTTGGAGTAAACCACGGACTCG	1	FAM	$(CAG)_7$	3	138–150*	LN848946
Cypfus_2257	F: AACCAGAGAAGTCCAGGTGC R: GTTTGGGTCCCAGTCTCTGACATC	4	FAM	$(CT)_{13}$	3	230–236*	LN848952
Cypfus_2506	F: ACCCTAACGACTGCATCACC R: GTTTAAATCTTGCCGTCTTCACCG	1	ATTO 550	$(TTC)_{12}$	4	218–245*	LN848954
Cypfus_2663	F: TGCAATTAAAGCCGTCCCAG R: GTTTACCTCCCTATGAGGTTCTTTAGC	3	Yakima Yellow	$(CATA)_7$	3	230–242*	LN848957
Cypfus_2987	F: ACGGATTCCTTCTCACACCC R: GTTTGCACGATGCTGCCTATACTTG	3	ATTO 550	$(CTT)_9$	4	249–264*	LN848964
Cypfus_2993	F: ATCGACTGCAAAGCATAGGG R: GTTTGGCCTCGGTCAGTTCTAC	4	ATTO 550	$(GAA)_8$	3	141–162*	LN848965
Cypfus_3114	F: TCCCGACTTCCTCCCAATTC R: GTTTAGCTCGCAGCATACCTAGAC	2	ATTO 565	$(CT)_{15}$	4	160–180*	LN848967
Cypfus_3212	F: ACACCTAAAAGCGAAAGCGG R: GTTTGACCGAAAGACGCTTGGAAC	3	ATTO 565	$(AAG)_8$	3	209–227*	LN848969
Cypfus_3218	F: TGTCCTCCTCTCCAACAAGC R: GTTTGAAATTCAACGGAGAGCGGG	4	ATTO 565	$(CTT)_9$	3	163–193*	LN848970
Cypfus_3300	F: TTTTGTTCTGGTTCCACGGG R: GTTTAGGTCCTCATTCTCTTCACCG	2	ATTO 550	$(GTAT)_{13}$	3	232–248*	LN848971
Cypfus_3921	F: ATGGATGACGAGGAGGTTGG R: GTTTGTAGAGGGAGGTTGGTAGCG	3	FAM	$(CGC)_8$	3	261–270*	LN848982
Cypfus_4093	F: TGTAAAACGACGGCCAGTGTCTCTCCAAACAGGAGGGC R: GTTTGTACAGGTAAGCGCAAGAGC M13: TGTAAAACGACGGCCAGT	2	FAM	$(GA)_{13}$	2	94–98*	LN848986
Cypfus_4216	F: GTTGTGAAAACCCTAGGCGG R: GTTTATTGAGGCCAGCACAACAAC	2	FAM	$(TTC)_{20}$	5	183–213*	LN848989
Cypfus_4236	F: GCTGTACGTGGAGAGAGGAG R: GTTTAAATCCACCGTCGCAAATCC	4	Yakima Yellow	$(AG)_{12}$	3	176–184*	LN848990
Cypfus_4666	F: GGGTGTTTGCATGACTGTAGC R: GTTTCGTAAGGGTACATAAGTCGATCC	2	Yakima Yellow	$(TATG)_7$	3	189–221*	LN848995

Note: *A* = number of alleles sampled; EMBL = European Molecular Biology Laboratory.

[a] Primers with the prefix Cf are from an NGS run from raw genomic DNA libraries; primers with the prefix Cypfus are from an NGS run from an enriched library.

[b] GTTT PIG-tails (Brownstein et al., 1996), M13R tails (5′-GGAAACAGCTATGACCAT-3′; Cf-primers), and M13 tails (5′-TGTAAAACGACGGCCAGT-3′; Cypfus_4093) added to the 5′ ends of primers are underlined.

[c] Fluorescent dye at the 5′ ends of M13R and M13 primers (Cf-primers and Cypfus_4093) and forward primers (remaining loci).

[d] The allele range is based on seven test individuals (Appendix 1).

* Length of PCR products is without PIG-tail, but with M13 tail (as for other loci resulting from the second NGS run in Appendix 2).

TABLE 3. Genetic diversity of 21 newly developed SSR markers in two fish pond populations of *Cyperus fuscus*.[a]

	Zahrádky (N = 25)				Libohošt (N = 25)			
Locus	A	H_o	H_e	F_{IS}[b]	A	H_o	H_e	F_{IS}[b]
Cf_008	4	0.240	0.577	0.589***	4	0.160	0.565	0.721***
Cf_017	2	0.160	0.470	0.664**	3	0.240	0.528	0.551**
Cf_019	3	0.040	0.365	0.892***	2	0.120	0.497	0.762***
Cf_104	2	0.040	0.301	0.870***	2	0.200	0.301	0.341
Cypfus_0173	3	0.120	0.541	0.782***	2	0.200	0.510	0.613**
Cypfus_0551	2	0.080	0.509	0.846***	2	0.080	0.509	0.846***
Cypfus_1207	3	0.080	0.223	0.646*	3	0.160	0.496	0.682***
Cypfus_2257	2	0.200	0.301	0.341	3	0.160	0.545	0.711***
Cypfus_2506	2	0.160	0.509	0.690***	3	0.120	0.667	0.823***
Cypfus_2663	2	0.080	0.444	0.823***	3	0.160	0.562	0.719***
Cypfus_2987	4	0.120	0.381	0.690***	3	0.120	0.548	0.784***
Cypfus_2993	3	0.080	0.401	0.804***	2	0.040	0.184	0.786**
Cypfus_3114	3	0.120	0.541	0.782***	3	0.200	0.601	0.672***
Cypfus_3212	2	0.040	0.510	0.923***	2	0.120	0.507	0.767***
Cypfus_3218	2	0.120	0.510	0.768***	4	0.240	0.584	0.594***
Cypfus_3300	4	0.320	0.706	0.552***	5	0.200	0.579	0.569***
Cypfus_3921	2	0.000	0.078	1.000*	2	0.160	0.490	0.678***
Cypfus_4093	2	0.120	0.301	0.607*	3	0.000	0.153	1.000***
Cypfus_4216	3	0.120	0.411	0.712***	4	0.000	0.584	1.000***
Cypfus_4236	2	0.040	0.350	0.888***	3	0.120	0.411	0.712***
Cypfus_4666	2	0.000	0.078	1.000*	2	0.040	0.040	0.000
Mean ± SD	2.6 ± 0.7	0.109 ± 0.078	0.405 ± 0.157	0.756 ± 0.159	2.9 ± 0.9	0.135 ± 0.071	0.470 ± 0.163	0.687 ± 0.212

Note: A = number of alleles sampled; F_{IS} = fixation index; H_e = expected heterozygosity; H_o = observed heterozygosity; N = number of individuals sampled; SD = standard deviation.

[a] See Appendix 1 for locality information for each population.

[b] Significant departures from Hardy–Weinberg equilibrium: *P < 0.05, **P < 0.01, ***P < 0.001.

APPENDIX 1. Voucher information for *Cyperus fuscus* populations used in this study. All vouchers are deposited at the Institute of Botany, University of Natural Resources and Life Sciences, Vienna (WHB). Individuals were grown from seeds in the greenhouse.

Voucher no.	Collection locality	Geographic coordinates	N
62957[a]	Czech Republic, Záryby	50°13.424′N, 14°37.717′E	1
62959[b]	Czech Republic, Semovice	49°45.067′N, 14°39.655′E	1
62987[b]	Czech Republic, Tchořovice	49°26.115′N, 13°48.442′E	1
62963[c]	Czech Republic, Mšec	50°11.815′N, 13°54.651′E	1
62960[c]	Czech Republic, Hluboká nad Vltavou	49°02.624′N, 14°25.952′E	1
62996[c]	Czech Republic, Smrkovec	49°26.078′N, 13°54.699′E	1
62982[c]	Czech Republic, Břeclav	48°42.710′N, 16°54.169′E	1
62979[c]	Czech Republic, Velké Němčice	48°59.056′N, 16°39.894′E	1
62973[c]	Poland, Borków	51°40.477′N, 16°12.239′E	1
62955[c]	Poland, Cigacice	48°18.739′N, 16°54.224′E	1
62968[d]	Czech Republic, Zahrádky	50°37.687′N, 14°32.595′E	25
62964[d]	Czech Republic, Libohošt	49°42.057′N, 14°35.398′E	25

Note: N = number of individuals sampled.

[a] Used for first NGS run at LGC Genomics (Berlin, Germany).

[b] Used for second NGS run at ecogenics (Balgach, Switzerland).

[c] Test individuals for screening of primer pairs.

[d] Test populations for assessment of genetic diversity.

Appendix 2. Characteristics of 44 additional SSR loci with flanking regions useful for primer design in *Cyperus fuscus*.

Locus	Primer sequences (5′–3′)[a]	Repeat motif	*A*	Allele size range (bp)[b]	EMBL accession no.
First NGS run					
Cf_007	F: CAGTCGGGCGTCATCAGAAGTGTATATTGAGATTAGGAGCC R: GTTTGGCTAGATCCAAATGGCGG	$(AT)_{11}$	3	274–286	LN848929
Cf_020	F: GGAAACAGCTATGACCATCTGCTGCCACCATTTCGAG R: GTTTAGGCTCAACCCTATGCACC	$(GGT)_5 + (GGT)_5$	1	273	LN848933
Cf_112	F: GTTTGTGGTGTGGCAGGAAGGG R: CAGTCGGGCGTCATCAGTCAGCTGTCAATCTGCACC	$(AATG)_7$	1	203	LN848935
Second NGS run					
Cypfus_0023	F: TGTAAAACGACGGCCAGTCTGCCTTCGATGAACTCCTG R: TCTTGTTCGGCGTCTAACCC	$(AGA)_7$	2	180–183	LN848936
Cypfus_0563	F: TGTAAAACGACGGCCAGTGAGAAGCGGGCATTCATCAG R: TATCCTCAGCTCCGTGTGTG	$(TC)_{12}$	3	139–143	LN848939
Cypfus_0568	F: TGTAAAACGACGGCCAGTCTGAGTCCCATGTCTCCTCC R: TGGTAATGCTCCATGCAAAGAC	$(CT)_{13}$	3	153–175	LN848940
Cypfus_0604	F: TGTAAAACGACGGCCAGTCACAGCTAGTGCAGTCAACG R: TGAGAAGTCGAGAGGAACGG	$(GA)_{18}$	4	160–170	LN848941
Cypfus_0785	F: TGTAAAACGACGGCCAGTAGGCGAGCTAGAGAAATGGG R: GAGGCGCCATCGATTCTTTC	$(AGA)_8$	2	152–155	LN848943
Cypfus_1174	F: TGTAAAACGACGGCCAGTCCCAACTGGAGCAAAGAAGC R: GCGGAAGTAGTTCAGGCAAC	$(TC)_{12}$	2	226–228	LN848945
Cypfus_1302	F: TGTAAAACGACGGCCAGTTTAACCAGGTCTCGTGGTCG R: ACAAAAGAGGCCGGATAGGC	$(TACA)_{12}$	2	162–166	LN848947
Cypfus_1319	F: TGTAAAACGACGGCCAGTAGAGGTTATTTGGCCCCAGC R: AGTGTTTGGCATGGGCTTTC	$(TATG)_8$	1	154	LN848948
Cypfus_1818	F: TGTAAAACGACGGCCAGTTCGCAGTTACGATAGGTACTC R: CATGGACGTGTCAAACAAAGC	$(CA)_{12}$	2	109–121	LN848949
Cypfus_1819	F: TGTAAAACGACGGCCAGTAGTGGACAAGGTCAAGAGGG R: CCATTGGGAGTCAAAGCCAC	$(GAA)_8$	2	207–210	LN848950
Cypfus_1966	F: TGTAAAACGACGGCCAGTATGGCATCGCAATCAACCAG R: GATGCGAGGTTTAAGCAGGG	$(GAA)_8$	3	216–222	LN848951
Cypfus_2381	F: TGTAAAACGACGGCCAGTGCACGTAACTTCCTTCTAGTGG R: TGGAAATAACTAGCTCACCACAC	$(TATG)_{18}$	3	191–263	LN848953
Cypfus_2517	F: TGTAAAACGACGGCCAGTTGAGCTGCAACCAATCAAGC R: TGTGCTGCCAGTTTTCCAAG	$(GAA)_7$	2	213–216	LN848955
Cypfus_2640	F: TGTAAAACGACGGCCAGTATCAAAACCCATCGCACTCC R: CGCTTATGCGCAAACAAACC	$(AGA)_7$	1	122	LN848956
Cypfus_2806	F: TGTAAAACGACGGCCAGTGCCTGATAAAGCATGTGACCG R: TCGAATTGACACCATGCCTC	$(AG)_{12}$	3	187–193	LN848958
Cypfus_2832	F: TGTAAAACGACGGCCAGTAGCACAAGTTGGGTCTCCTC R: TTGATCACCCCCACTAAGGC	$(GAA)_7$	1	173	LN848959
Cypfus_2855	F: TGTAAAACGACGGCCAGTCAGCGGAAGGGAAGATTTCG R: CTCAGCCATCTCAATCACCG	$(CTT)_7$	3	202–215	LN848960
Cypfus_2888	F: TGTAAAACGACGGCCAGTTGCTCCGCTTCTATTTTGCTC R: GACCGAAGCTGCTGATTTCC	$(CTT)_{12}$	2	151–154	LN848961
Cypfus_2891	F: TGTAAAACGACGGCCAGTGGACTGGTTTAGAAATGTGTGC R: TTTTGGCAACGTGAAAGTGC	$(CT)_{12}$	1	255	LN848962
Cypfus_2898	F: TGTAAAACGACGGCCAGTAAAAACAGCTGAATCGGGGC R: CTGCAGACCCATCTCTCTCC	$(TTC)_9$	3	226–247	LN848963
Cypfus_3033	F: TGTAAAACGACGGCCAGTTATGGCGCTGGAGGAGAAAG R: CGTGTCGTAAGGCAGAAAATAAAATC	$(CTT)_{10}$	2	220–226	LN848966
Cypfus_3195	F: TGTAAAACGACGGCCAGTGGGGAGGAGAGTTCCTTGAC R: CTTCAGTGATCCCATGTGGC	$(AG)_{12}$	1	197	LN848968
Cypfus_3323	F: TGTAAAACGACGGCCAGTCCTAGCACTTGCAAAGGGTG R: CGCCCCTTTTCGTATTGTCC	$(AAG)_7$	1	217	LN848972
Cypfus_3372	F: TGTAAAACGACGGCCAGTCAGCTCCACGATACTCGATTG R: AAGGGACTCAATATCGCCCC	$(GAA)_8$	2	255–258	LN848974
Cypfus_3416	F: TGTAAAACGACGGCCAGTTCTTCAAAACTTGCCTATGGGTC R: TGTGCAGACATTTGAGGAAGC	$(GAGT)_7$	3	238–244	LN848975
Cypfus_3423	F: TGTAAAACGACGGCCAGTTCTGTCCTCCTCGCTCAATC R: TCAAACCAAGTAATTTTCCAAAGAG	$(GAA)_7$	2	193–202	LN848976
Cypfus_3542	F: TGTAAAACGACGGCCAGTGCGGGACTTCCATTCCATTC R: GGTAGACGGCGCTTTTTGAG	$(TCT)_{15}$	3	241–253	LN848977
Cypfus_3597	F: TGTAAAACGACGGCCAGTTGCATGTTCACTTCTGGTGC R: CACCTTCTGCTGCTCAATCG	$(GAA)_7$	1	181	LN848978
Cypfus_3776	F: TGTAAAACGACGGCCAGTTCGGTAATATACTTTGGGTCAGC R: GAACGGGAAACAAGACGCTC	$(CT)_{12}$	2	245–249	LN848979
Cypfus_3864	F: TGTAAAACGACGGCCAGTCGAAGAATTTTCCCACCCCG R: CCGTTAAACAGGTCCGAAGC	$(CT)_{14}$	3	210–218	LN848980
Cypfus_3873	F: TGTAAAACGACGGCCAGTGAAAAGACAGATGCCTCCGC R: CCGCCTCTACCAGATACTGC	$(GAA)_8$	3	172–211	LN848981
Cypfus_3898	F: TGTAAAACGACGGCCAGTTCAGGCCACATGCCTTTTTC R: GGGAGCAAACCTGAGCAATC	$(TTC)_8$	2	174–195	LN868257
Cypfus_4041	F: TGTAAAACGACGGCCAGTAGGTGGAAGTAGGAAGCCAG R: CATTTGCAGCCCCATCCTTC	$(GAA)_9$	1	176	LN848983

Appendix 2. Continued.

Locus	Primer sequences (5′–3′)[a]	Repeat motif	*A*	Allele size range (bp)[b]	EMBL accession no.
Cypfus_4074	F: TGTAAAACGACGGCCAGTTTGCAAATGGGCACAGGAAG R: CCTAATAAAGGTAGGACAGAGCG	$(TC)_{13}$	4	184–198	LN848985
Cypfus_4102	F: TGTAAAACGACGGCCAGTTGGGCGTTCTCAAATCAAAGAG R: GGGGCCCACTGAAGAAAAAG	$(GA)_{13}$	1	260	LN848987
Cypfus_4240	F: TGTAAAACGACGGCCAGTCTTCTTCATTTCCCGCACCC R: GCCACCTGCATTCATCATCC	$(TACA)_7$	2	251–255	LN848991
Cypfus_4347	F: TGTAAAACGACGGCCAGTTCATTTCAACTCGGAATCCTCTAC R: CAACTACAACCGGCACCTTC	$(TGTA)_7$	2	252–256	LN848992
Cypfus_4468	F: TGTAAAACGACGGCCAGTCGAATCTGAGAAGCGCTGTG R: ACTCATCGCTTGAGAGGCAG	$(CT)_{12}$	3	259–275	LN848993
Cypfus_4479	F: TGTAAAACGACGGCCAGTTGGGTGCCAAACAAAAATTGG R: AGATATCAAAAGCAACCGACCC	$(AAG)_8$	2	158–233	LN848994
Cypfus_4799	F: TGTAAAACGACGGCCAGTTATGGGCTTCCCGTCTTCTG R: CTGTCATGCTCGACACCAAG	$(AAG)_7$	2	248–251	LN848996
Cypfus_4849	F: TGTAAAACGACGGCCAGTAATGAAGAGCGCACCAATCG R: ACAATACATTCCTCGGTTAGACAG	$(GA)_{12}$	1	158	LN848997

Note: *A* = number of alleles sampled; EMBL = European Molecular Biology Laboratory.

[a]GTTT PIG-tails (Brownstein et al., 1996), CAG and M13R tails (CAG: 5′-CAGTCGGGCGTCATCA-3′; M13R: 5′-GGAAACAGCTATGACCAT-3′; only in Cf_007, Cf_020, and Cf_112), and M13 tails (5′-TGTAAAACGACGGCCAGT-3′) added to the 5′ ends of primers are underlined.

[b]The allele range is based on seven test individuals (Appendix 1).