Molecular Ecology Notes (2007) **7**, 428–432 doi: 10.1111/j.1471-8286.2006.01608.x

PRIMER NOTE

Fifty-three polymorphic microsatellite loci in the chestnut blight fungus, *Cryphonectria parasitica*

T. L. KUBISIAK,* C. DUTECH† and M. G. MILGROOM‡

**USDA Forest Service, Southern Institute of Forest Genetics, 23332 Highway 67, Saucier, MS 39574, USA, †Institut National de la Recherche Agronomique, Unité Mixte de Recherche BIOGECO, Equipe de Pathologie Forestière, Domaine de la Grande Ferrade, 71 avenue Edouard Bourlaux, BP 81, 33883, Villenave d'Ornon Cedex, France, ‡Department of Plant Pathology, Cornell University, Ithaca, NY 14853, USA*

Abstract

We report on 53 microsatellite loci for use in population genetic or linkage mapping studies in *Cryphonectria parasitica*. In 40 isolates collected from throughout the Northern Hemisphere, the number of alleles per locus ranged from two to 14 (mean 5.17) with gene diversity values ranging from 0.049 to 0.859 (mean 0.437). Samples from Asia were more diverse than those from Europe and North America. Most of the markers (48 of 53) were developed from an expressed sequence tag library, and hence, offer the opportunity to examine population structure or provide genome location information for specific expressed genes vs. anonymous genomic regions.

Keywords: chestnut blight fungus, *Cryphonectria parasitica*, microsatellite, polymorphism, simple sequence repeat, SSR

Received 31 July 2006; revision accepted 8 October 2006

The availability of genetic markers for studying the chestnut blight fungus, *Cryphonectria parasitica*, has increased markedly in the last 2 years. A set of 11 codominant sequence characterized amplified region markers (Davis *et al*. 2005) and 13 microsatellite markers (Breuillin *et al*. 2006) were recently developed and used for population genetic analyses. In addition, 140 random amplified polymorphic DNA (RAPD) markers have been used for recombinational linkage mapping in *C. parasitica* (Kubisiak & Milgroom 2006). We report here an additional 53 polymorphic microsatellite markers available for population genetic analyses and recombinational linkage mapping studies. These new markers are derived as an extension of the marker development done by Breuillin *et al*. (2006) from a combination of enriching for microsatellite motifs from a genomic library and mining an expressed sequence tag (EST) library from *C. parasitica* (Dawe *et al*. 2003).

We identified sequences containing di- through pentanucleotide repeat motifs from approximately 4200 ORFs in an EST library available in GenBank (Dawe *et al*. 2003). Using the software program SPUTNIK (Chris Abajian unpublished; program available at http://espressosoftware.com), 152 EST clones were identified that had repeat motifs that were at least 18 bp or larger and for which the repeat motif represented at least 80% of the sequence. Primer pairs designed for eight of these clones were reported previously in Breuillin *et al*. (2006). For the other 144 repeat-containing EST clones, 114 were found to be unique sequences based on multiple alignments using the software SEQUENCHER version 5.1.1 (Gene Codes). Microsatellite primer pairs were developed for these unique repeat-containing clones using the web-based software program PRIMER 3 (Rozen & Skaletsky 1998; http//cbr-rbc.nrc-cnrc.gc.ca/cgi-bin/primer3_www.cgi). The general primer picking conditions were set for an optimal primer size of 20 bases and optimal thermal melting temperature (Tm) of 60 °C. To maximize our post-PCR (polymerase chain reaction) multiplexing capacity, primer pairs were developed to produce expected amplicons 100–400 bp, in size increments of 50 bases (assuming no introns). Five additional primer pairs were developed from a genomic library enriched for $(GA)_n$ and $(CA)_n$ repeats as described in Breuillin *et al*. (2006). To reduce the costs associated with primer screening and increase our post-PCR multiplexing flexibility and capacity, an M13-specific sequence (5′-CACGACGTTGTAAAACGAC-3′)

Correspondence: T.L. Kubisiak, Fax: +1228 8320130, E-mail: tkubisiak@fs.fed.us

was added on to the 5′ end of each forward primer as similarly described by Schuelke (2000). To favour 3′ adenylation of the forward amplified strand, all reverse primers were PIG-tailed with a 7-base sequence (5′-GTTTCTT-3′) on the 5′ end (Brownstein *et al*. 1996). For fluorescent detection, three-primer PCR was performed which included a 5′ dye-labelled M13-specific primer (same sequence as the M13 'tail' described above).

All 119 microsatellite primer pairs were screened against DNA from 40 haploid isolates of *C. parasitica* from various populations throughout the Northern Hemisphere (Table 1). Microsatellite PCR mixtures consisted of the following in 24 μL total volume: 1.25 ng of template DNA, 0.208 μM 5′-dye-labelled M13 primer DNA, 0.052 μM of 5′-M13-tailed forward primer, 0.208 μM of reverse PIG-tailed primer, 600 μM of total dNTPs (150 μM each), 2.4 μL 10× *Taq* DNA polymerase reaction buffer (500 mM KCl, 100 mM Tris-HCl, 1.0% Triton X-100, 15 mM $MgCl_2$), and 0.8 U of *Taq* DNA polymerase. Reactions were loaded in flexible microtitre plates and overlaid with 25 μL mineral oil. Microtitre plates were placed in preheated (85 °C), programmable temperature cyclers (MJ Research PTC-100) and covered with mylar film. DNA samples were immediately amplified using the following thermal profile: 4 min at 95 °C; 35 cycles of 20 s at 92 °C, 20 s at 55 °C, 20 s at 72 °C; 7 min at 72 °C; indefinite hold at 4 °C. Completed reactions were diluted 1:10 with distilled water and 1 μL was electrophoresed on an ABI PRISM 3100 Genetic

Table 1 Origin of *Cryphonectria parasitica* isolates collected from throughout the Northern Hemisphere

Isolate	Country of origin	Collection site
BF93	USA	Danby, New York
BF110	USA	Danby, New York
DU21	USA	Depot Hill, New York
EP155	USA	Bethany, Connecticut
RH10-4	USA	Rocky Hill, Connecticut
KY1	USA	Powell County, Kentucky
MI5	USA	County Line, Michigan
NJ22	USA	Five Points, New Jersey
V5	USA	Mountain Lake, Virgina
WV75-m1	USA	Parsons, West Virgina
ONT02	Canada	Wardsville, Ontario
P17-8	Italy	Laboratory strain
RC15	Italy	Cittanova, Reggio di Calabria
FI1	Italy	Pomino, Firenze
VO1	Italy	Crevoladossola, Verbania
CT2	Italy	Zafferana, Catania
GS10	Greece	Stagira, Region of Macedonia
G31	Republic of Macedonia	Glogi, Tetovo
FR21	Republic of Macedonia	Frangovo, Struga
OS11	Republic of Macedonia	Osoj, Kichevo
09207	China	Ninghua, Fujian
09321	China	Heyuan, Guandong
09331	China	Heyuan, Guandong
09371	China	Xiuning, Anhui
09373	China	Xiuning, Anhui
09383	China	Guixi, Jiangxi
09520	China	Guixi, Jiangxi
09530	China	Guixi, Jiangxi
CD2	Japan	Chudai, Kyoto
CD18	Japan	Chudai, Kyoto
CD20	Japan	Chudai, Kyoto
JA17	Japan	Chiyoda, Ibaraki
JA19	Japan	Iwama, Ibaraki
OK1	Japan	Okoba, Kumamoto
OK27	Japan	Okoba, Kumamoto
OK28	Japan	Okoba, Kumamoto
OB3-5	Japan	Obuse, Nagano
KB4	Japan	Kobuchizawa, Yamanashi
KB8	Japan	Kobuchizawa, Yamanashi
IF2	Japan	Iwate Town, Iwate

Table 2 Characterization of 53 microsatellite loci isolated from *Cryphonectria parasitica* and their diversity in 40 isolates collected from throughout the Northern Hemisphere

Locus	Library*	GenBank Accession no.	Core motif	Primer sequence†	*n*	Size range (bp)‡	N_a§	*H*¶
CPG6	Genomic	DQ834941	$(GA)_{17}$	F-ATCATCACGACGCAATGGTA R-TCCGGGCATTCAGCAMAT	37	263–289	7	0.520
CPG7	Genomic	DQ834942	$(GGT)_7(GTT)_4$	F-GGGCATGACAACTGAGAAGG R-GCTGTCGTCTCTCTCGATCAC	40	251–257	4	0.468
CPG8	Genomic	DQ834943	$(GT)_7GA(GT)_8$	F-GTCCGACTGGCTGTTCGTC R-AAAAGGTGGCAGGATGAGG	40	186–214	4	0.546
CPG12	Genomic	DQ834944	$(GT)_3A(GT)_6$ $A(TG)_3C(TG)_2$	F-GTTGGCTCCAGAACCGTTAC R-AGTGTCCGTCCATCGAAGAG	39	277–292	4	0.233
CPG14	Genomic	DQ834945	$(AGGAAG)_6$	F-TTCTGAAGGTGGTTGTGGTG R-GGTCCGAACCATCAAAAGAC	40	257–305	11	0.763
CpSI001	EST	CB690636	$(ACT)_8$	F-CAGAGTCCATCTCGCCTCAT R-CCGGAGCAGTTTCTCAAAGA	40	129–156	5	0.271
CpSI002	EST	CB690609	$(AGG)_4TGG(AGG)_3$	F-TTGGATAGACCCAGGTGTCC R-GAGGTCTTCGAGGGCGTAG	37	474–514	7	0.665
CpSI006	EST	CB690297	$(AC)_{13}$	F-ATGTCGAGTTTACCCGATGG R-GAGATGTGTGGAATGCAACG	40	160–164	3	0.634
CpSI007	EST	CB690258	$(AGC)_7$	F-CAACATCATGAGGCAAGACG R-GGTTGCCAGTCTTCACTTGG	39	253–256	2	0.484
CpSI008	EST	CB690203	$(AGC)_3AAC(AGC)_7$	F-CACAGGCACAAGCACAGG R-TGTGTTGGACCATTGGTGAG	39	321–323	2	0.142
CpSI009	EST	CB690185	$(AT)_{14}$	F-ATCATCCATCCTGTCCGAGT R-TGGGGTTGGCATAATCTTCT	40	164–206	9	0.771
CpSI011	EST	CB690082	$(AGG)_5TGG$ $(AGG)_3$**$(A)_{27}$††	F-TCAAGACCTTGTGGTCATTCC R-CCTGTTGTGAAGCCCATCTT	37	253–271	14	0.773
CpSI012	EST	CB690071	$(AAC)_6$	F-CGCGTCCGTAGACTACTCATC R-AGAGGCAGACATGGCAGAAG	38	171–174	2	0.051
CpSI014	EST	CB690056	$(AAG)_9$	F-TCGGAGGCTTTATTGTCGTT R-TGGGTGTATTTGCTCGGTAA	40	263–299	7	0.554
CpSI015	EST	CB690009	$(CCG)_6$	F-GTAAAGACGGGTGCCAACAT R-CGTCTGGTGTAGGAGCATGA	40	349–360	4	0.143
CpSI018	EST	CB689747	$(AGG)_9$	F-AATTACCCACCTGCATACCTG R-TCTTGAGCATGTCAACGACAG	40	175–196	5	0.348
CpSI019	EST	CB689690	$(AG)_{12}$	F-GGCATCATGACCAAGGATCT R-GCCGCTTGTTGAATACATGA	39	320–322	2	0.460
CpSI020	EST	CB689642	$(AGG)_5G(AGG)_7$	F-CCCTCCCTCCTAGACGACTT R-GTGTAGGGGCTTCCTTGGAG	38	332–347	6	0.529
CpSI022	EST	CB689616	$(AGC)_5$	F-CCGCAGTGGTGGCGGCGG R-CTCACGATCTCCCAGTCCAC	40	264–269	3	0.096
CpSI028	EST	CB689427	$(AATG)_5$	F-TCTTCCTCCTCGTCACCATC R-GTTGTGAAACGTGAGCGGTA	38	274–277	4	0.611
CpSI031	EST	CB689403	$(AGG)_7$	F-TACACCACCGAACTCACCAA R-TCTTCGTCCCTCCTATGTGC	40	345–357	5	0.415
CpSI032	EST	CB689367	$(AAC)_{11}$	F-GCCAGCACTTGTTACGACAG R-CGGAGCAGACATCAGGTTCT	40	437–477	9	0.794
CpSI034	EST	CB689322	$(GAA)_6$	F-CAGGGAAGACGAAGACGAAC R-GAGGTGCAGCTGTGCTCTTA	40	246–254	5	0.188
CpSI038	EST	CB689033	$(AGC)_8$	F-ACCACCAGCCTCCTTCCTAC R-CTCGCGACGAGATTGAGATT	40	183–204	5	0.628
CpSI040	EST	CB689009	$(AGC)_7$	F-TGCTGCTTGGGCTTCTTATT R-GCGAGCAGTGCTTGAATCT	35	221–239	4	0.488
CpSI042	EST	CB688954	$(AAC)_6$	F-CAGCACTAGCACTGGTCGAT R-GAACGAGGACAGGTCAGGAG	39	200–208	4	0.487
CpSI044	EST	CB688849	$(AAC)_9$	F-CCTTCGTATCCATTCCCTTG R-GTAGGCCTCCTTGAGCTCCT	39	287–296	4	0.588
CpSI045	EST	CB688846	$(AAC)_7CACCAC$ $(AGC)_5$	F-GCAATTCAAGTGCGAGCATA R-GCGTGCTTGACCAGGATAAC	34	323–334	5	0.659
CpSI048	EST	CB688615	$(AAC)_5$	F-GCTCTGCAAAACTCCTCTGC R-GCTCCCGACTCTTCACAGAA	40	232–235	2	0.049

Table 2 *Continued*

Locus	Library*	GenBank Accession no.	Core motif	Primer sequence†	*n*	Size range (bp)‡	N_a§	*H*¶
CpSI050	EST	CB688563	$(AAG)_6(TTTC)_3$	F-GGTCTTGCCCAACACATCA R-CTCGACGGGTTTTTGACC	40	197–203	3	0.096
CpSI052	EST	CB688417	$(AG)_{14}G(AG)_4$	F-CTCACTGGCCTGAGCTTTTT R-AACGCCTCTCGTGACACAC	39	341–420	11	0.656
CpSI053	EST	CB688400	$(AAG)_3AAA(AAG)_6$	F-TGATGTTGAGGAGCTTGCAG R-TTATGCCTGTCCAAGGTTCC	40	289–292	2	0.095
CpSI054	EST	CB688333	$(AG)_{29}GG(AG)_3$	F-CTTGATAAACGAGGGCCAAG R-GAAGGTGACGTTTTGCCAGT	37	290–343	13	0.859
CpSI055	EST	CB688332	$(ATC)_8$	F-CGGCTATCGTTGTCGATTTT R-ATGCCGAGAAGCTGGAGTAT	40	354–361	3	0.184
CpSI056	EST	CB688286	$(AG)_{13}$	F-GTCCGCGTTTAGGACCATC R-TCACATTTCCTTCCTTTCCAA	38	209–215	4	0.632
CpSI058	EST	CB688253	$(AGAGG)_3(AG)_4$	F-CCGAACATATCGACCAAACC R-TACATGGCGGAAATGTTTGA	37	132–137	2	0.053
CpSI061	EST	CB688116	$(AGC)_9(AAC)_4(AGC)_7$	F-GGCTATGGCGTGACGACTAC R-CTGACCTCGACATTCAAGCA	40	230–297	10	0.726
CpSI062	EST	CB688052	$(CCG)_6$	F-AGCAACAGAGGCCAGCTAAG R-AGGGACGCCAGGACATAGAC	37	330–342	3	0.104
CpSI063	EST	CB688003	$(AGC)_7$	F-GAAGCAGGTCAGGAGACTGG R-AGGATGCCGGATTCCTTTAG	40	175–184	3	0.096
CpSI068	EST	CB687797	$(AAC)_5$	F-CGCCAATGAAGCCATACTTT R-GACAGCCCCATCATGAGC	40	555–567	6	0.483
CpSI069	EST	CB687759	$(AGC)_8(AAC)_8$	F-GGTGGACATCATGAGGGAAG R-GAGTGCTTCTTCCGGAGTCA	40	251–267	4	0.344
CpSI072	EST	CB687613	$(AAC)_{13}$	F-CTACGACCGAGAGGTTGAGG R-GTCATCCCAGTACAGCAGCA	40	324–359	11	0.795
CpSI073	EST	CB687597	$(GTT)_7$	F-GCCATAATCGTCCTCGTCAT R-GGGAACGAGTGAACGACCTA	40	155–164	5	0.586
CpSI074	EST	CB687592	$(ACC)_2C(ACC)_8$	F-CGCAACACTGCAAAATTCAA R-GAGGTCCACTTGCTGGTCAC	40	219–224	3	0.486
CpSI079	EST	CB687453	$(AG)_5AC(AG)_5$	F-CATTCGATTTCACGAAGCAG R-CAAGTCGTCCAACAGATGGATA	40	197–200	2	0.049
CpSI080	EST	CB687451	$(AG)_{12}$	F-AACAACGGATCCAAGAGCAT R-ACTCGTGCTGCCTGAATCAT	40	398–412	4	0.431
CpSI083	EST	CB687434	$(AGG)_7$	F-TGGGCATGTTCATACTGCTC R-GCCAGCTCTCATCATGACAAT	40	192–213	6	0.385
CpSI085	EST	CB687363	$(AC)_{11}$	F-AGGCCTGCTTCTTTTTGGAT R-CGGGTCTATATGGTGGCTTC	39	248–252	3	0.441
CpSI090	EST	CB687187	$(AG)_9$	F-ATGGATACAGGGGGAAAAGG R-CGCCCCCAATATATACCTTC	40	219–222	3	0.486
CpSI101	EST	CB686785	$(AGC)_6$	F-GGCAATCTCAATAGAGGACCTG R-GTTGAAGGCTCCCAAGTTCA	40	174–183	2	0.095
CpSI102	EST	CB686746	$(ACG)_{12}$	F-GCTCCGAGGACTTTGATGAG R-TCATCACCACCAACACCATT	39	270–282	4	0.548
CpSI107	EST	CB686545	$(AGC)_{14}$	F-CAGGGAACTCGAAAAGGTTG R-CCCTCCAGGTCCTTGTGATA	40	264–302	10	0.691
CpSI108	EST	CB686531	$(AAG)_{10}$	F-CGGAACTACCTGCTCTTTGC R-GCGATCCGCATTCCTGTAT	40	221–265	9	0.429
Mean							5.17	0.437

*Markers from the genomic library were identified by enrichment for repetitive motifs $(GA)_n$ and $(CA)_n$ as described by Breuillin *et al*. (2006). Markers from the expressed sequence tag (EST) library were identified by searching for microsatellite motifs among *C. parasitica* EST sequences in GenBank (Dawe *et al*. 2003) as described by Breuillin *et al*. (2006) and in the text.
†M = A, C.
‡Observed sizes include additional bases due to the forward and reverse primer tails.
§Observed number of alleles.
¶Nei's (1978) gene diversity.
††**Stretch of nonrepeat-containing sequence.

Analyser as described by the manufacturer (Applied Biosystems). For markers less than 500 bp ABI PRISM LIZ500 was used as an internal size standard. Allele sizes were determined using the local Southern algorithm implemented by ABI PRISM GENEMAPPER software version 3.7. Due to mobility shift problems associated with the 250-bp and 340-bp bands, these sizes were excluded from all sizing calls. For markers greater than 500 bp LIZ600 was used as an internal standard.

Fifty-three of the 119 (44.5%) microsatellite primer pairs screened were found to be polymorphic (> 1 allele) across the 40 isolates analysed (Table 2; monomorphic loci are not shown). Allele data for the 53 polymorphic loci were imported into the software program POWERMARKER version 3.25 (http://www.powermarker.net) for further analysis. The number of alleles ranged from two to 14 (mean 5.17). Nei's gene diversity ranged from 0.049 to 0.859 (mean 0.437). Pairwise linkage disequilibrium (LD) was tested with GENEPOP 3.4 (Raymond & Rousset 1995). After correcting for multiple comparisons, significant LD was observed for only two locus pairs (CpSI069:CpSI073 and CpSI072:CpSI073). Although physical linkage might explain the LD observed for locus pair CpSI069:CpSI073, we can rule out physical linkage between locus pair CpSI072:CpSI073 (unpublished data).

Based on hierarchical analysis, the sample from Asia was found to be more diverse than the sample from Europe and North America. The mean number of alleles observed in Asia was 4.98 vs. 1.87 in Europe and North America. Similarly, the mean gene diversity observed in Asia was 0.522 vs. 0.246 in Europe and North America. These results are similar to those found using restriction fragment length polymorphisms (RFLPs) where greater diversity was observed in Asia than in Europe or North America (Milgroom *et al*. 1996), and are consistent with *C. parasitica* originating in Asia, with a recent introduction into Europe and North America.

These additional markers have expanded the set of useful microsatellites now available for fine-scale population genetic analyses and recombinational linkage mapping studies. Thirty-one of these 53 markers are segregating in a mapping population (results not shown) and will be placed on a linkage map (Kubisiak & Milgroom 2006) in the near future. A majority of the polymorphic loci (48 of 53) described here were developed from an EST library, and hence, offer the opportunity to examine population structure or provide genome location information for specific expressed genes as opposed to anonymous genomic regions.

Acknowledgements

We thank Charles Burdine and Dennis Deemer for their skilled technical assistance with data collection using an ABI 3100 Genetic Analyser and GENEMAPPER, and Joanne Davis and Cristina McGuire for DNA preparations.

References

Breuillin F, Dutech C, Robin C (2006) Genetic diversity of the chestnut blight fungus *Cryphonectria parasitica* in four French populations assessed by microsatellite markers. *Mycological Research*, **110**, 288–296.

Brownstein MJ, Carpten JD, Smith JR (1996) Modulation of non-templated nucleotide addition by *Taq* polymerase: primer modifications that facilitate genotyping. *BioTechniques*, **20**, 1004–1010.

Davis JE, Kubisiak TL, Milgroom MG (2005) Polymorphic sequence-characterized codominant loci in the chestnut blight fungus, *Cryphonectria parasitica*. *Molecular Ecology Notes*, **5**, 195–197.

Dawe AL, McMains VC, Panglao M, Kasahara S, Chen BS, Nuss DL (2003) An ordered collection of expressed sequences from *Cryphonectria parasitica* and evidence of genomic microsynteny with *Neurospora crassa* and *Magnaporthe grisea*. *Microbiology*, **149**, 2373–2384.

Kubisiak TL, Milgroom MG (2006) Markers linked to vegetative incompatibility (*vic*) loci and a region of high heterogeneity and reduced recombination near the mating type locus (*MAT*) in *Cryphonectria parasitica*. *Fungal Genetics and Biology*, **43**, 453–463.

Milgroom MG, Wang K, Zhou Y, Lipari SE, Kaneko S (1996) Intercontinental population structure of the chestnut blight fungus, *Cryphonectria parasitica*. *Mycologia*, **88**, 179–190.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Rozen S, Skaletsky HJ (1998) *PRIMER 3*. Code available at http://www-genome.wi.mit.edu/genome_software/other/primer3.html.

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