Biotechnol Lett (2011) 33:2217–2223
DOI 10.1007/s10529-011-0682-9

ORIGINAL RESEARCH PAPER

Comparison of three microsatellite analysis methods for detecting genetic diversity in *Phytophthora sojae* (Stramenopila: Oomycete)

Silvina Stewart · Damitha Wickramasinghe · Anne E. Dorrance · Alison E. Robertson

Received: 21 April 2011 / Accepted: 22 June 2011 / Published online: 9 July 2011

Abstract Analysis of an organism's genetic diversity requires a method that gives reliable, reproducible results. Microsatellites are robust markers, however, detection of allele sizes can be difficult with some systems as well as consistency among laboratories. In this study, our two laboratories used 219 isolates of *Phytophthora sojae* to compare three microsatellite methods. Two capillary electrophoresis methods, the Applied Biosystems 3730 Genetic Analyzer and the CEQ 8000 Genetic Analysis system, detected an average of 2.4-fold more alleles compared to gel electrophoresis with a mean of 8.8 and 3.6 alleles per locus using capillary and gel methods, respectively. The two capillary methods were comparable, although allele sizes differed consistently by an average of 3.2 bp across isolates. Differences between capillary methods could be overcome if reference standard DNA genotypes are shared between collaborating laboratories.

Keywords Alleles · Capillary electrophoresis · Microsatellites · Pathogenic diversity · *Phytophora sojae*

Electronic supplementary material The online version of this article (doi:10.1007/s10529-011-0682-9) contains supplementary material, which is available to authorized users.

S. Stewart · A. E. Robertson (✉)
Department of Plant Pathology, Iowa State University, Ames, IA 50011, USA
e-mail: alisonr@iastate.edu

S. Stewart
e-mail: silvina@iastate.edu

D. Wickramasinghe · A. E. Dorrance
Department of Plant Pathology, The Ohio State University, Wooster, OH 4469, USA

D. Wickramasinghe
e-mail: wickramasinghe.2@osu.edu

A. E. Dorrance
e-mail: dorrance.1@osu.edu

Introduction

Phytophthora sojae Kauf. & Gerd. is an important pathogen that causes Phytophthora root and stem rot on soybeans worldwide (Hartman et al. 1999). High levels of pathogenic variation within the species occurs and more than 200 pathotypes of this pathogen have been reported and more continue to emerge (Dorrance and Grunwald 2009). Interestingly, little is known about how this variation occurs and the diversity within endemic populations. Oomycetes are diploid organisms whose life cycle includes both asexual and sexual reproduction. Organisms that reproduce asexually tend to exhibit a high degree of clonality, with few genotypes present at high frequencies, while sexually reproducing organisms usually have a higher degree of genotypic diversity (Chen and McDonald 1995). Due to its homothallic nature, *P. sojae* is considered an essentially clonally propagating

organism (Gijzen and Qutob 2009). Previous studies have indicated that little, if any, heterozygosity is present in populations (Förster et al. 1994).

As with many soil borne pathogens *P. sojae* has limited means of dispersal, thus gene flow is thought to be limited (McDonald and Linde 2002). It has been suggested however, that a large reservoir of genetic diversity exists in *P. sojae* populations (Hobe 1981), albeit, only a few studies have attempted to characterize this diversity using genetic markers (Dorrance and Grunwald 2009; Drenth et al. 1996; Förster et al. 1994; Gally et al. 2007; Meng et al. 1999). Co-dominant microsatellites or simple sequence repeats (SSRs) are suited for population-genetic studies, since they enable quantification of putative heterozygotes which enables estimation of naturally occurring outcrossing. SSRs for *P. sojae* were previously identified from transcript sequences (Garnica et al. 2006), as well as from genome sequences (Tyler et al. 2006). Schena et al. (2008) identified 12 SSRs that could be used on a restricted number of *Phytophthora* species related to *P. sojae*. In another study, 21 SSRs developed from *P. sojae* race 2 sequences, were used in a preliminary study on 33 isolates from Ohio (Dorrance and Grunwald 2009). An average of 2.5 alleles per locus and 0.015 observed heterozygosity was found, as well as, 100% of loci deviated from Hardy–Weinberg equilibrium (Dorrance and Grunwald 2009).

Reproducibility of molecular markers has been tested in laboratory networks (Jones et al. 1997). Random amplified polymorphic DNAs (RAPDs) have proven difficult to reproduce from one laboratory to the next. Amplified fragment length polymorphisms (AFLPs), although reproducible, result in single-band differences between labs. While SSRs are considered robust markers, differences in allele sizing can appear across laboratories depending on the analysis system used (Jones et al. 1997; Weeks et al. 2002; Widmark et al. 2011). The estimated allele size is not only dependent on the number of nucleotides but also on the mobility of the fragment in the electrophoresis (Weeks et al. 2002; Widmark et al. 2011), the type of fluorescent label used, the distance of the allele from the standard used (Jones et al. 1997), and the use of different instruments using different software (Weeks et al. 2002). Nevertheless, these discrepancies could be minimized if reference standard DNA genotypes were shared between collaborating laboratories. Our objective was to compare three microsatellite methods across two laboratories, standardize measurements and name the alleles detected.

Materials and methods

A total of 219 isolates of *P. sojae* were evaluated in this study. Genomic DNA was extracted from mycelium using either a modification of the cetyltrimethylammonium bromide (CTAB) procedure (Dorrance et al. 1999), or a rapid extraction protocol (Zelaya-Molina et al. 2011). Twenty-five microsatellite primer pairs were identified (Dorrance and Grunwald 2009; Schena et al. 2008) and amplicons were separated on 4% agarose gels (Supplementary Table 1).

Alleles which differ in many base pairs of length can be readily resolved on agarose gels but single repeat differences are difficult to separate, especially in SSRs with small size repeats (Jones et al. 1997). Eight SSRs were selected for further comparisons using two capillary electrophoresis microsatellite analysis methods (Table 1). The eight SSRs included 2, 3, 4, 5, and 6 bp repeats and were chosen based on amplification success and the highest number of alleles (band sizes) encounter per locus using the gel method. The Applied Biosystems 3730 Genetic Analyzer (ABI) was used at Iowa State University, and the CEQ 8000 Genetic Analysis system (CEQ) at The Ohio State University. Although each laboratory used their own isolates of *P. sojae*, DNA of 17 isolates was shared between laboratories to allow for comparison of allele sizes using the different methods and dyes.

Primer synthesis for the ABI used universal fluorescent labeling (standard phosphoramidite chemistry). For the forward primers PS01, PS16, PS24, and PS33, 6-carboxy-fluorescine (FAM) dye was used, while hexachloro-6-carboxy-fluorescine (HEX) dye in forward primers PS05, PS10, PS12, and PS29. Amplification was performed in a 96-well Eppendorf Mastercycler thermal cycler (Hamburg, Germany) in 15 µl with 0.2 mM dNTP mixture, 2.5 mM $MgCl_2$, 1X Go Taq Hot Start Colorless Master Mix buffer, 0.08 units Go Taq Hot Start DNA polymerase (Promega Inc., Madison, WI), 0.45 µM of each primer, and 1 µl (100 ng) DNA template. The thermal cycler was programmed for an initial step at 85°C for 2 min, denaturalization step at 94°C for

Table 1 Eight SSRs, primer sequences, and allele size based on the original sequenced isolate P6497

Locus	Super-contig[a]	Repeat motif[b]	Primer sequences (5′–3′)	Actual size isolate P6497 (bp)	GeneBank accession no.
PS01	9	$(GACACT)_{49}$	F: TGATGGGAGATGGCTACAGG	419	EF667485
			R: TCGCAACGACAGATTGATG		
PS05	3	$(TCAG)_{34}$	F: GAAACAATCAACCGAACAACG	263	EF667486
			R: ATAGGAGGGCAAACTGGATG		
PS10	NI[c]	$(CAAAC)_{27}$	F:CGACGAAGAACAACATTACTTG	228	EF667489
			R: ATGAAACCGAACCAAACCTG		
PS12	32	$(GCTGTT)_{23}$	F: GCTGCTTGTTGCTGTTGTTG	306	EF667490
			R: GCGGGTGTTTGGAGAGTATC		
PS16	20	$(ATTAT)_{20}$	F: AATCTGACTTGGACGCTGTG	469	EF667491
			R: GCTTAGTGTTTTGGGTTACGC		
PS24	36	$(CT)_{16}$ +$(CT)_3$	F: GTCATTTCCCTCGCTCACAC	252	EF667495
			R: ACACTGGCAACAAGCAACAG		
PS29	NI	$(TAC)_{15}$	F: CCACTGAAGCGAGGTAGAGG	273	EF667499
			R: GTAGCACAAAATCCGTCTGC		
PS33	2	$(AT)_{15}$	F: CTGCTAGTGCCGTTCGTTG	267	EF667501
			R: TAAAAGGGCTGCTCAAATCG		

[a] Tyler et al. (2006) and Dorrance and Grundwald (2009)

[b] Super-contig or DNA region in the *P. sojae* genome where the microsatellite is located

[c] '+' indicates that two simple repeats were separated by other bases

[d] *NI* not identified

95 s, then 24 cycles at 52°C for 1 min, 72°C for 72 s, 94°C for 30 s, then 52°C for 1 min and 72°C for 30 min. A 96-capillary Applied Biosystem 3730 Genetic Analyzer set up to run samples labeled with these dyes and a GeneScan 500 ROX size standard (Applied Biosystem, Foster City, CA) was used. GeneMapper Software 4.0 (Applied Biosystem, Foster City, CA) was used to size the alleles to the nearest base pair.

For the capillary electrophoresis with the CEQ, all forward primers were designed with a M13(−21) (TGTAAAACGACGGCCAGT) tail at the 5′-end and used a universal WellRed labeled M13(−21) primer as a nested primer (Schuelke 2000). PCR conditions were modified when using universal labeled M13(−21) primer. Two PCR reactions were carried out with the same reverse primer. However, in the first run, M13(−21) tagged forward primer was used while universal M13(−21) labeled primer was used in the second PCR. Amplification was performed in a 96-well DNA Engine Tetrad 2 Peltier Thermal Cycler (Biorad, USA) in 25 µl with 0.2 mM dNTP mixture, 2 mM $MgCl_2$, 1X Go Taq Flexi Colorless buffer, 1 units Go Taq Flexi DNA polymerase (Promega Inc., Madison, WI), 0.2 µM of each primer, and 1 µl (20 ng) DNA template. The thermal cycler was programmed for an initial step at 95°C for 5 min, denaturalization step at 94°C for 30 s, then 24 cycles at annealing temperature for 30 s, 72°C for 30 s and 72°C for 10 min. For annealing temperature, 58 and 56°C were used in the first and second PCRs, subsequently. For comparison, forward primers of two primer sets, PS01 and PS05 were directly labeled with WellRed and used with same PCR protocol and conditions except only one PCR reaction was used with an annealing temperature 58°C. PCR products were electrophoresed on a CEQ 8000 Genetic Analyzer (Beckmann Coulter) with either 400 or 600 bp size standard depending on the size of the fragment. Allele sizes were determined using the software provided by the Genetic Analyzer. Eighteen bases were removed from the allele data from the primer pairs that used universal labeled M13 (−21) primer.

Results

The 190 *P. sojae* isolates analyzed with 25 SSRs using gel method resulted in a total of 75 alleles, ranging from 2 to 6 alleles per locus, with an average of three (Supplementary Table 1). The agarose gel

electrophoresis method detected 86% of the bands which differed by more than 5 bp. In contrast, the capillary methods were more sensitive and able to separate PCR products that differed by only 2 bp. For the eight selected SSRs (Table 1) that were assayed in this study, a total of 70 alleles, with an average of 8.8 alleles per locus, were detected using the capillary electrophoresis methods from 219 *P. sojae* isolates compared to a total of 29 alleles, with an average of 3.6 alleles per locus, using the gel method (Table 2). Thus the capillary electrophoresis methods were able to detect 2.4-fold more alleles on average than the gel method. The number of alleles detected per locus using the three methods however, varied depending on the specific SSR. For PS29 which has 3 bp repeat, the number of alleles detected by the three methods was identical (Table 2). In contrast, 14 alleles were detected for PS05 (5 bp repeat) with the capillary methods, while only five were detected using the gel method (Table 2).

The two capillary methods had similar results in both labs. For the reference isolates shared between labs, the same alleles were detected although allele sizes for all of the primer pairs differed from 1 to 13 bp with an average of 3.2 bp (Table 2). The longer the PCR product size, the bigger the base pair differences between the ABI system and the CEQ. Specifically, PCR products over 400 bp differed by 5–13 bp (PS01 and PS16), however, these base pair differences were consistent across isolates (Table 2). Similarly, PCR products of less than 400 bp differed consistently, but by only 1 or 2 bp between the two capillary methods.

Discussion

When studying genetic variation of an organism, it is important to use a method that gives reliable and reproducible results that can be repeated from laboratory-to-laboratory and across genotyping systems. In this study, the ABI and CEQ system detected more alleles for seven of eight SSRs than an agarose gel electrophoresis method. Although, the number of alleles detected by the capillary systems were similar when the CEQ and ABI systems were compared, an average variation of 3 bp for each amplified fragment was observed, with larger variations in size between methods being observed as the size of the PCR product increased. While this was not unexpected, since migration of the PCR products is affected by the labeling and the capillary system (Jones et al. 1997; Weeks et al. 2002; Widmark et al. 2011), it does mean that representative standard isolates should be included in similar assays to ensure lab-to-lab comparisons.

The methods used in this study vary considerably in cost. In addition to the equipment and the requirement of core lab facilities, both the ABI system and CEQ require fluorescent labeled primers which have additional costs. At the time of this study, each sample analyzed on a gel cost approximately $0.62 (using Agarose wide range/standard 3:1 and rapid-run agarose buffer), compared to $1.48 and $1.65 per sample for 96 samples for FAM and HEX dye, respectively, for the ABI method (includes cost of fluorescent label and genotyping). For one sample, the cost of using the CEQ machine for genotyping is $1.75 at Ohio State University. The prices are different in the two genomic core facilities, making sample cost significantly different between them. When the universal primer M13(−21) was used the cost was $2.49/sample for 96 samples, however, when direct labeling method was used, the cost increased to $4.64/sample for 96 sample. The cost of a sample using ABI or CEQ does decrease, as the number of samples increases ($1.27 to $1.29 per sample in ABI system and $2.49 to $1.82–2.08 per sample in CEQ system for 960 samples), and/or by multiplexing more than one sample per well.

A challenge in using the M13(−21) primer labeling method described by Schuelke (2000) is that the annealing temperatures of SSR primers should not be close to that of the universal labeled M13(−21) primer. In this study, the annealing temperatures for SSR primers and M13 primer differed by only 2–5°C, which made it difficult to amplify specific bands when all three primers were used simultaneously. To overcome this, a two-step PCR was necessary, where the first PCR was performed using M13(−21) tailed forward primer and SSR reverse primer, and the second PCR used the universal fluorescent labeled M13(−21) forward and same SSR reverse primer. The PCR products of the first PCR were used as the template for the second PCR. Direct primer labeling, as done for PS01 and PS05, was easier and allowed for multiplex by labeling different colored dyes. Using this approach

Table 2 Loci, primers, labeling method and alleles detected in this study (page 1 of 2)

Locus	Primer	Gel method Approximate band size (bp)[a]	Applied biosystems 3730 genetic analyzer (ABI)[b] Allele size (bp)[c]	CEQ 8000 genetic analysis system (CEQ)[d] Allele size (bp)[e]	Allele name	Difference in bp between ABI and CEQ
PS01	PS01-F	–	n/d[f]	206	PS01-B	–
		257	n/d	260	PS01-C	–
	PS01-R	–	n/d	266	PS01-D	–
		269	*269*	*272*	PS01-E	3
		281	*281*	*284*	PS01-F	3
		–	*287*	*290*	PS01-G	3
		293	293	296	PS01-H	3
		–	n/d	418	PS01-I	–
		419	*419*	*424*	**PS01-A**	5
		–	*425*	*430*	PS01-J	5
		–	431	436	PS01-K	5
		–	437	442	PS01-L	5
		–	443	448	PS01-M	5
PS05	PS05-F	–	n/d	260	PS05-B	–
		263	*263*	*264*	**PS05-A**	1
	PS05-R	–	n/d	268	PS05-C	–
		290	n/d	296	PS05-D	–
		–	n/d	304	PS05-E	–
		307	*307*	*308*	PS05-F	1
		–	n/d	312	PS05-G	–
		–	n/d	316	PS05-H	–
		–	n/d	320	PS05-I	–
		335	*335*	*336*	PS05-J	1
		–	*339*	*340*	PS05-K	1
		–	343	344	PS05-L	1
		–	347	n/d	PS05-M	–
		360	355	n/d	PS05-N	–
PS10	PS10-F	–	n/d	144	PS10-B	–
		153	*146*	*149*	PS10-C	3
	PS10-R	–	*191*	*194*	PS19-D	3
		198	196	199	PS10-E	3
		–	211	214	PS10-F	3
		218	n/d	219	PS10-G	–
		–	*221*	*224*	**PS10-A**	3
		228	n/d	229	PS10-H	–
		233	n/d	234	PS10-I	–
		238	n/d	249	PS10-J	–
		–	251	n/d	PS10-K	–
PS12	PS12-F	258	*251*	*252*	PS12-B	1
		264	*257*	*258*	PS12-C	1

Table 2 continued

Locus	Primer	Gel method Approximate band size (bp)[a]	Applied biosystems 3730 genetic analyzer (ABI)[b] Allele size (bp)[c]	CEQ 8000 genetic analysis system (CEQ)[d] Allele size (bp)[e]	Allele name	Difference in bp between ABI and CEQ
	PS12-R	–	292	293	PS12-D	1
		300	*298*	*299*	PS12-E	1
		–	*304*	*305*	PS12-F	1
		306	*310*	*311*	**PS12-A**	1
		–	316	317	PS12-G	1
PS16	PS16-F	–	395	403	PS16-B	8
		405	*405*	*413*	PS16-C	8
	PS16-R	–	n/d	418	PS16-D	–
		–	*455*	*463*	PS16-E	8
		–	460	n/d	PS16-F	–
		–	n/d	478	PS16-G	–
		470	*470*	*483*	**PS16-A**	13
		–	*475*	*488*	PS16-H	13
PS24	PS24-R	236	*236*	*238*	PS24-B	2
		252	*252*	*254*	**PS24-A**	2
	PS24-F	262	*262*	*264*	PS24-C	2
		–	n/d	310	PS24-D	–
		–	n/d	338	PS24-E	–
PS29	PS29-R	252	*249*	*251*	PS29-B	2
	PS29-F	273	*270*	*272*	**PS29-A**	
PS33	PS33-R	–	*250*	*252*	PS33-B	2
		257	256	258	PS33-C	2
	PS33-F	–	258	260	PS33-D	2
		–	262	264	PS33-E	2
		–	264	266	PS33-F	2
		267	*266*	*268*	**PS33-A**	2
		–	*268*	*270*	PS33-G	2
		–	270	272	PS33-H	2
		–	274	276	PS33-I	2
		–	278	280	PS33-J	2

[a] Total number of isolates analyzed $n = 190$; 126 from Ohio and 64 from Iowa

[b] Forward primer PS01, PS16, PS24, and PS33 were labeled with FAM while PS05, PS10, PS12, and PS29 were labeled with HEX dye

[c] Total number of isolates analyzed $n = 106$; 93 isolates from Iowa (includes 64 isolates used for the gel method) and 13 isolates received from Ohio's lab. Size is based on approximate band sizes as determined by GeneMapper® Software 4.0

[d] All forward primer were designed with a M13(−21) tail at the 5′-end and used a universal WellRED D4 labeled M13(−21) primer as a nested primer, except for primers PS01 and PS05 that were directly labeled using WellRED D4

[e] Total number of isolates analyzed $n = 130$; 126 isolates from Ohio (same isolates used for the gel method) and 4 isolates received from Iowa's lab. Size is based on approximate band sizes as determined by CEQ 8000 genetic analysis software

[f] Not detected in the isolates from that laboratory

Alleles found on the isolates shared by laboratories ($n = 17$) are *italicized*

Allele corresponding to sequenced isolate P6497 is *bolded*

can reduce costs when large numbers of samples need to be analyzed.

The capillary methods used in this study resulted in higher number of alleles detected, and although more expensive, the results allow for greater detection of genetic variation in *P. sojae*. The agarose method has the advantage that it can be accomplished by any laboratory with minimum infrastructure and is more economical, although it is best suited for those SSRs with longer repeats as their differences in amplicons can be easily distinguished on a gel. In addition, interpretations of data from studies with gel systems should be made with caution as some alleles will be missed.

Acknowledgments Funding for this study was received in part from the Iowa Soybean Association and Ohio Soybean Council. Salaries and research support were provided by State and Federal funds appropriated to the Iowa State University and Ohio Agricultural Research and Development Center, The Ohio State University. Sequencing was done at the DNA Facility at Iowa State University Office of Biotechnology and Molecular Cellular Imaging Center at the OARDC. We wish to thank Vijitha Silva and Lily Zelaya-Molina for assistance with these studies.

References

Chen R, McDonald BA (1995) Sexual reproduction plays a major role in the genetic structure of populations of the fungus *Mycosphaerella graminicola*. Genetics 142:1119–1127

Dorrance AE, Grunwald N (2009) *Phytophthora sojae*: diversity among and within populations. In: Lamour K, Kamoun S (eds) Oomycetes genetics and genomics: diversity, interactions, and research tools. Wiley, Hoboken

Dorrance AE, Inglis AD, Derie ML, Brown CR, Goodwin SB, Fry WE, Deah KL (1999) Characterization of *Phytophthora infestans* population in western Washington. Plant Dis 83:423–428

Drenth A, Whisson SC, Maclean DJ, Irwin JAG, Obs NR, Ryley MJ (1996) The evolution of races of *Phytophthora sojae* in Australia. Phytopathology 86:163–169

Förster H, Tyler BM, Coffey MD (1994) *Phytophthora sojae* races have arisen by clonal evolution and by rare outcrosses. MPMI 7:780–791

Gally M, Ramos AM, Dokmetzian D, Lopez SE (2007) Genetic variability of *Phytophthora sojae* isolates from Argentina. Mycologia 99:877–883

Garnica D, Pinzon A, Quesada-Ocampo L, Bernal A, Barreto E, Grunwald N, Restrepo S (2006) Survey and analysis of microsatellites from transcript sequences in *Phytophthora* species: frequency, distribution, and potential as markers for the genus. BMC Genomics 7:245

Gijzen M, Qutob D (2009) *Phytophthora sojae* and soybean. In: Lamour K, Kamoun S (eds) Oomycetes genetics and genomics: diversity, interactions, and research tools. Wiley, Hoboken

Hartman GL, Sinclair JB, Rupe JC (eds) (1999) Compedium of soybean disease, 4th edn edn. American Phytopathological Society, St. Paul

Hobe M (1981) Pathogenic variability of *Phytophthora megasperma* f.sp. *glycinea* isolated from northwest Ohio soils. M.S. Thesis, Ohio State University, Columbus

Jones CJ, Edwards KJ, Castiglione S, Winfield MO et al (1997) Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. Mol Breed 3:381–390

McDonald BA, Linde C (2002) Pathogen population genetics, evolutionary potential, and durable resistance. Ann Rev Phytopath 40:349–379

Meng XQ, Shoemaker RC, Yang XB (1999) Analysis of pathogenicity and genetic variation among *Phytophthora sojae* isolates using RADP. Mycol Res 103:173–178

Schena L, Cardle L, Cooke D (2008) Use of genome sequence data in the design and testing of SSR markers for *Phytophthora* species. BMC Genomics 9:620

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nat Biotech 18:233–234

Tyler BM, Tripathy S, Zhang X, Dehal P et al (2006) *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. Science 313:1261–1266

Weeks DE, Conley YP, Ferrell RE, Mah TS, Gorin MB (2002) A tale of two genotypes: consistency between two high-throughput genotyping centers. Genome Res 12:430–435

Widmark AK, Andersson B, Sandström M, Yuen JE (2011) Tracking *Phytophthora infestans* with SSR markers within and between seasons—a field study in Sweden. Plant Pathol 60. doi:10.1111/j.1365-3059.2011.02446.x

Zelaya-Molina LX, Ortega MA, Dorrance AE (2011) Easy and efficient protocol for oomycete DNA extraction suitable for population genetic analysis. Biotechnol Lett 33:715–720