Full Length Research Paper

Characterization of thirteen microsatellite loci from the Ghanian antimalarial plant *Cryptolepis sanguinolenta*

Naalamle J. Amissah[1], Phillip A. Wadl[2,3], Denita Hadziabdic[2], Sarah L. Boggess[2] and Robert N. Trigiano[2]*

[1]Department of Crop Science, P. O. Box LG 44, University of Ghana, Accra, Ghana.
[2]Department of Entomology and Plant Pathology, the University of Tennessee, 2505 E. J. Chapman Drive, Knoxville, TN, USA.
[3]USDA-ARS-U.S, Vegetable Laboratory, 2700 Savannah Hwy, Charleston, SC, USA.

Cryptolepis sanguinolenta* (Lindl.) Schlechter (Periplocaceae) is an herbaceous plant used in traditional medicine to treat malaria and populations of the species are diminishing due to overharvesting and lack of conservation. Codominant microsatellite markers that can be used to characterize genetic diversity and population structure are currently not available. Therefore, the study isolated 75 microsatellite loci from genomic sequence data, which were then screened for the ability to reveal polymorphisms. From the 75 candidate loci, 13 polymorphic microsatellite loci were optimized for future population genetics studies. Twenty-two *C. sanguinolenta* samples were collected from eight different geographical locations in Ghana. Alleles per locus ranged from 3 to 7 with a mean of 4.4. Expected heterozygosity ranged from 0.24 to 0.77, and all but one locus deviated significantly from Hardy-Weinberg equilibrium. Genetic differentiation mean was 0.06 among all loci, indicating relatively low genetic diversity in these samples. These microsatellite loci should be useful to study genetic diversity, gene flow and population structure as well as in a project involving breeding and conservation of *C. sanguinolenta.

Key words: *Cryptolepis sanguinolenta*, genetic diversity, malaria, medicinal plant, microsatellites.

INTRODUCTION

Cryptolepis sanguinolenta (Lindl.) Schlechter (Periplocaceae) is an important perennial medicinal plant species indigenous to West and Central Africa with a long history of use in the treatment of malaria (Ankrah, 2010; Tempesta, 2010; Osei-Djarbeng et al., 2015). An aqueous extraction of its roots, the portion of the plant with the highest concentration of active antimalarial ingredients, yields indoloquinoline alkaloids, primarily cryptolepine, an N-methyl derivative of the indoloquinoline compound quindoline (Dwuma-Badu, 1978; Tachie et al., 1991). The extract may also have some anticancer properties as well (Ansah and Mensah, 2013). The harvested roots are sold in local markets. Non-sustainable destructive harvesting of the plants for the roots has resulted in a substantial decrease in wild populations (Jansen and Schmelzer, 2010).

*Corresponding author. E-mail: rtrigian@utk.edu. Tel: (865) 974-0221.

Its widespread use as a medicinal plant coupled with over-harvesting, calls for the formulation of effective management plans and conservation through cultivation with the ultimate goal of selecting high active ingredient producing plants for breeding (Amissah et al., 2016).

There is very limited information about the genetic diversity in populations of *C. sanguinolenta*. Amplified Fragment Length Polymorphism (AFLP) revealed low genetic diversity among 116 plants sampled from three regions in Ghana, but did not characterize the population structure or gene flow within the collected sites (Amissah et al., 2016). Microsatellites or single sequence repeats (SSRs) are codominant molecular markers, as compared to dominant AFLPs markers, that typically are polymorphic (Gupta et al., 1996) and well-suited for studying population genetic diversity and dynamics. They are most often employed to evaluate population structure, gene flow and inbreeding (Arnold et al., 2002; Zhang and Hewitt, 2003).

Currently, microsatellite molecular markers are not available to genetically characterize *C. sanguinolenta* individuals and populations. The study generated genome sequence data for *C. sanguinolenta,* and identified 13 polymorphic microsatellites that were used to characterize individuals from eight populations in Ghana, West Africa. The microsatellites from this study will be used to characterize a larger sample of the population and provide information on genetic diversity, population structure and gene flow of the species. The microsatellites should be very useful for a breeding program to create elite genotypes that produce high quantities of antimalarial compounds.

MATERIALS AND METHODS

A DNA library (target 400 bp lengthreads) was constructed from a *C. sanguinolenta* individual that was collected in Hweehwee Oboyan, Ghana. One microgram of DNA was used for library preparation using the Ion Xpress™ Plus gDNA Fragment Library Preparation kit (Life Technologies, Carlsbad, CA). Prior to sequencing, the library was quantified with the Ion Library Quantitation Kit (Life Technologies) and diluted to 20 pM. The library was then prepared for sequencing using the Ion PGM™ Template OT2 400 Kit (Life Technologies). The library was loaded on an Ion 318™ Chip v2 (Life Technologies) and sequenced on the Ion PGM™ System (Life Technologies) using the Ion PGM™ Sequencing 400 Kit (Life Technologies). Raw sequencing reads were trimmed and then assembled *de novo* into contigs using the default parameters in CLC Genomics Workbench (Qiagen). Only contigs ≥ 20X coverage were searched for microsatellite motifs (di- to hexa-nucleotide) and primers were designed using BatchPrimer3 (You et al., 2008) using default settings.

Genomic DNA was isolated from leaves of 22 *C. sanguinolenta* plants from 8 locations using a CTAB protocol (Porebski et al., 1997) and one DNA sample of *Pityopsis graminifolia* (Michx.) Small from Tennessee (USA) was used as a negative control for amplification. Ten μl reaction amplifications were completed as described in Hatmaker et al. (2015) using 75 primer pairs. The thermal cycler conditions consisted of 94°C for 3 min and 35 cycles of 94°C for 40 s, 55°C for 40 s and 72°C for 30 s, followed by 72°C for 4 min. Allelic products were separated via electrophoresis using

the QIAxcel Capillary Electrophoresis System (Qiagen, Valencia, California USA) and the data binned into allelic classes according to the protocols described by Hatmaker et al. (2015). Data was analysed for number of alleles/locus, observed heterozygosity, expected heterozygosity, probability of deviation from Hardy-Weinberg equilibrium at $P<0.05$, Shannon's information index, and genetic differentiation (F_{ST}) using the program GenALEx 6.5 (Peakall and Smouse, 2012).

RESULTS AND DISCUSSION

A total of 4,910,467 reads (mode = 405 bp) and 1.3 Gb were generated from sequencing with a single Ion 318™ Chip (v2) on the Ion PGM™ System. After quality trimming, the reads were assembled *de novo* into 38,029 contigs with an average length of 1,721 bp and a N50 of 1,670 bp. Only contigs ≥ 20X coverage (n=1,928) were screened for microsatellites, and 821 were detected. Tri- and tetra-nucleotide motifs were the most common, with 36.1% of the motifs in each of these classes. The next most common class was the dinucleotide motifs at 13%. A total of 574 primer pairs were designed and of these 75 were randomly selected for amplification and detection of polymorphic alleles.

Seventy of the seventy-five primer pairs amplified microsatellite loci. Nineteen of the primers consistently yielded more than two products, twenty-five amplified only one allele/locus (monomorphic) for all plants included in the study and another thirteen primers amplified only two alleles (either as homozygotes or heterozygotes) per locus. Thirteen primers were polymorphic, revealed more than two alleles per locus, and were used for analysis in this study (Table 1). DNA from *P. graminifolia* did not amplify with any of the primer pairs, indicating accuracy of our results. The number of alleles/ locus was relatively low and ranged from 3 to 7 (mean 4.4), which has been observed in other endemic perennial species (Arroyo et al., 2016). The mean observed frequency of heterozygotes (H_o) in the total sample was 0.44, which deviated from the expected heterozygosity (H_e), indicating potential presence of population structure among the eight population collection areas (Hadziabdic et al., 2012). Expected heterozygosity ranged from 0.24 to 0.77, and all but one locus significantly deviated from Hardy-Weinberg equilibrium. However, more samples from each of these areas will be needed to determine population structure and gene flow. Shannon's diversity index averaged 0.85 among all loci. This index measures the number of unique genotypes in a sample population and how evenly the genotypes are distributed (Brown and Weir, 1983). Genetic differentiation (F_{ST}) value of 0.06 was calculated across the 13 loci Three of loci had a calculated F_{ST} ≥0.05, which showed moderate diversity, however 10 of the loci had a F_{ST} of ≤0.05 that indicated relatively low genetic diversity in these samples may portend the same for the general population (Hartl and Clark, 2007).

Annual and herbaceous perennial plants in

Table 1. Primer sequences, repeat motifs, annealing temperatures (T_a), number of alleles (A), allele size range, observed (H_o) and expected heterozygosity (H_e), Shannon's information index (I), and genetic differentiation (F_{ST}) of 13 microsatellite loci isolated from *Cryptolepis sanguinolenta*.

GenBank Accession No.	Locus	Primer sequences (5'-3')	Repeat Motif	Ta (°C)	A	Size range (bp)	Ho	He	I	F_ST
KU361063	Cry007	F: GACATAGCTTTGGAAGGGTAG R: TCAGGATCAATTTCCACTTG	$(AAAT)_3$	58	4	265-285	0.88	0.70*	1.27	0.03
KU361064	Cry021	R: GCTCCTGACCCCATATCAGTT F: GGTGACCCTCAATCTGATGAA	$(GAT)_4$	60	3	179-199	0.97	0.51*	0.75	0.00
KU361065	Cry025	F: GCCATTTGGAATGGTATCTCC R: TCACACAAGTGACCATGATTGA	$(TATTT)_3$	60	5	266-280	0.00	0.61*	1.12	0.07
KU361066	Cry030	F: GGCATTGGCCATATATACTCCT R: CATGTTCCAGGGCAGTAGAAA	$(AT)_7$	60	5	143-155	0.33	0.61*	1.06	0.01
KU361067	Cry031	F: AGGATACGTGGCTCCAAGATT R: AGGATGCCAATGAAAGGATCT	$(TAA)_4$	60	6	135-147	0.97	0.75*	1.47	0.04
KU361068	Cry032	F: GGAAGCATCCATAAGGAGGAG R: TGGATATTCTAGTTGCCTTGTGG	$(TTTA)_4$	60	5	107-122	0.97	0.64*	1.15	0.05
KU361069	Cry033	F: GAAGCCCCATTCTTTCGATAC R: TTAAACCTCCCAAGTGCTCAG	$(TTA)_5$	59	3	121-125	0.00	0.45*	0.74	0.09
KU361070	Cry035	F: CAACTCATAATCAGGGGTCAAA R: TGGAACTCTAAGGATTTGTGTC	$(ATTT)_3$	57	3	167-171	0.00	0.54*	0.82	0.02
KU361071	Cry036	F: GCGCAATTAAGAAATGCGTA R: CAGCGCATTTCCAACAATAA	$(TAT)_5$	59	4	244-262	0.13	0.56*	0.93	0.03
KU361072	Cry037	F: GCGCAATTAAGAAATGCGTA R: CAGCGCATTTCCAACAATAA	$(ATT)_5$	59	4	244-263	0.17	0.24*	0.45	0.12
KU361073	Cry 041	F: GAACAGTAGCGTAGCCCCAGT R: AAAGCACCAAGAAAAGATGAGG	$(ATAA)_3$	60	4	135-141	0.29	0.65	1.11	0.02
KU361074	Cry044	F: CGGTTCCCCTTCTACAATTTC R: CCCATCAACTAACGAAAAAGG	$(AAC)_5$	59	7	237-251	1.00	0.77*	1.55	0.03

Table 1. Cont'd.

KU361075	Cry047	F: TAGGGCTTCTTTCATGCGTTT R:CAGATTCAAGTGGACGAAACC	(TTCC)$_3$	60	4	134-146	0.03	0.39*	0.68	0.02
Means					4.4		0.44	0.57	0.85	0.06

*Significant deviation from Hardy-Weinberg equilibrium at $P \leq 0.05$.

comparison to woody plant species have lower mean genetic diversity due to lack of polymorphic loci as well as narrow geographical distribution (Hamrick et al., 1992). The study finding agrees with the report by Amissah et al. (2016) that used AFLP analysis to reveal a low (25%) genetic diversity in 116 sampled plants. Additional samples and analyses need to be completed to confirm our preliminary findings using micro-satellite analysis.

Conclusion

The study provides defined codominant molecular markers for *C. sanguinolenta.* A total of 13 primer pairs can be used to amplify polymorphic microsatellite loci that should provide more complete assessment of genetic diversity, gene flow and population structure of this species. Additionally, these markers should be a valuable tool for breeding elite genotypes and conservation of *C. sanguinolenta.*

Conflict of Interests

The authors have not declared any conflict of interests.

ACKNOWLEDGMENTS

REFERENCES

Amissah JN, Spiller M, Oppong A, Osei-Safo D, Owusu-Darko R, Debener, T, Danquah EY, Addae-Mensah I (2016). Genetic diversity and cryptolepine concentration of *Cryptolepis sanguinolenta* (Lindl.) Schlt from selected regions of Ghana. J. Appl. Res. Med. Arom. Plants 3:34-41.

Ankrah NA (2010). Treatment of falciparum malaria with a tea-bag formulation of *Cryptolepis sanguinolenta* root. Ghana Med. J. 44(1):2.

Ansah C, Mensah KB (2013). A review of the anticancer potential of the antimalarial herbal *Cryptolepis sanguinolenta* and its major alkaloid cryptolepine. Ghana Med. J. 47:137-147.

Arnold C, Rossetto M, McNally J, Henry RJ (2002). The application of SSRs characterized for grape (*Vitis vinifera*) to conservation studies in Vitaceae. Am. J. Bot. 89:22-28.

Arroyo JM, Escudero M, Jordano P (2016). Isolation of 91 polymorphic microsatellite loci in the Western Mediterranean endemic *Carex helodes* (Cyperaceae). Appl. Plant Sci. 4:1500085.

Brown A, Weir B (1983). Measuring genetic variability in plant populations. In: Isozymes in Plant Genetics and Breeding, Part A, Tanksley SD, Orton TJ (eds); Elsevier, Amsterdam. pp. 219-239.

Dwuma-Badu D, Ayim JS, Fiagbe NI, Knapp JE, Schiff PI, Slatkin DJ (1978). Constituents of West African medicinal plants: XX Quindoline from *Cryptolepis sanguinolenta*. J. Pharm. Sci. 67:433-434.

Gupta PK, Balyan HS, Sharma PC, Ramesh B (1996). Microsatellites in plants: A new class of molecular markers. Curr. Sci. (India) 70:45-54.

Hamrick JL, Godt MJW, Sherman-Broyles SL (1992). Factors influencing levels of genetic diversity in woody plant species. In: Population Genetics of Forest Trees, Adams WT, Godt MJ, Sherman-Broyles SL(eds); Kluwer Academic, Dordecht, The Netherlands. pp. 95-124.

Hadziabdic D, Wang XW, Wadl PA, Rinehart TA, Ownley BH, Trigiano RN (2012). Genetic diversity of flowering dogwood in the Great Smoky Mountains National Park. Tree Genet. Genomes 8:855-871.

Hartl DL, Clark AG (2007). Principles of Population Genetics,fourth ed.; Sinauer Associates, Sunderland, Massachusetts, USA.

Hatmaker AE, Wadl PA, Mantooth K, Scheffler BE, Ownley BH, Trigiano RN (2015). Development of microsatellites from *Fothergilla ×intermedia* (Hamamelidaceae) and cross transfer to four genera within the Hamamelidaceae. Appl. Plant Sci. 3(4):1400123.

Jansen PCM, Schmelzer GH (2010). *Cryptolepis sanguinolenta* (Lindl.) Schltr. In: Schmelzer GH, Gurib-Fakim A (eds); Prota 11(2): Medicinal plants/Plantes médicinales 2. PROTA, Wageningen, Netherlands.

Osei-Djarbeng SN, Agyekum-Attobra E, Nkansah R, Solaga D, Osei-Asante S, Owusu-Dapaah G (2015). Medicinal plants constituting antimalarial herbal preparations in the Ghanaian market. Br. J. Pharm. Res. 5:153-162.

Peakall R, Smouse PE (2012). GenAlEx 6.5: genetic

Analysis in Excel.Population genetic software for teaching and research-an update. Bioinformatics 28:2537-2539.

Porebski S, Bailey LG, Baum BR (1997). Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. Plant Mol. Biol. Report. 15:8-15.

Tachie AN, Sharaf MHN, Schiff PL, Boye GL, Crouch RC, Martin GE (1991). Assignment of proton and carbon NMR spectra of the indoquinoline alkaloid cryptolepine. J. Heterocycl. Chem. 28:1429-1435.

Tempesta MS (2010). The clinical efficacy of *Cryptolepis sanguinolenta* in the treatment of malaria. Ghana Med. J. 44(1):1-2.

You FM, Huo N, Gu YQ, Luo MC, Ma Y, Hane D, Lazo GR, Dvorak J, Anderson OD (2008). BatchPrimer3: a high throughput web application for PCR and sequencing primer design. BMC Bioinformatics 29:253-266.

Zhang DX, Hewitt GM (2003). Nuclear DNA analyses in genetic studies of populations: Practice, problems and prospects. Mol. Ecol. 12:563-565.