Development of microsatellite markers for the American species *Vachellia aroma* (Fabaceae, Caesalpinodeae)

Desarrollo de marcadores microsatélites para la especie americana *Vachellia aroma* (Fabaceae, Caesalpinodeae)

Carolina L. Pometti[1], Cecilia F. Bessega[1], Mauricio Ewens[2], Juan C. Vilardi[1] & Beatriz O. Saidman[1]

1. Laboratorio de Genética, Departamento de Ecología, Genética y Evolución, Instituto IEGEBA (CONICET- UBA), Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina
2. Estación Experimental Fernández-UCSE (Convenio Provincia Santiago del Estero-Universidad Católica Santiago del Estero), Santiago del Estero, Argentina

*caritoege@gmail.com; cpometti@ege.fcen.uba.ar

Citar este artículo
POMETTI, C. L., C. F. BESSEGA, M. EWENS, J. C. VILARDI & B. O. SAIDMAN. 2024. Development of microsatellite markers for the American species *Vachellia aroma* (Fabaceae, Caesalpinodeae). *Bol. Soc. Argent. Bot.* 59: 27-32.

DOI: https://doi.org/10.31055/1851.2372.v59.n1.40763

Summary

Background and aims: There are currently no microsatellite markers available for any American species of *Vachellia*, and particularly, for *V. aroma*. Then the aims of this study were to develop SSR markers specific for *V. aroma*, for the first time, and test its amplification in a close related species.

M&M: For the development of the SSR in *V. aroma,* total genomic DNA was extracted and it was sequenced in a one-fourth run on a Roche 454 GS FLX+ platform. The study area included two Argentinean populations: San José and Robles, Santiago del Estero province.

Results: We detected 422 sequences containing SSR *loci.* A set of 39 primer pairs presented amplified products in *V. aroma* and *V. caven*, but 12 revealed clear, replicable and polymorphic *loci* in *V. aroma*.

Conclusions: The results of this work indicate that a new set of SSR markers was developed for *V. aroma* and their transferability to *V. caven* was assessed. The analysis of variability, showed that these 12 polymorphic markers are highly informative, and a powerful tool to investigate population genetics parameters in *V. aroma* and related species.

Keywords

454 pyrosequencing, polymorphic SSR, *Vachellia aroma, Vachellia caven*.

Resumen

Introducción y objetivos: Actualmente, no existen marcadores microsatélites disponibles para ninguna especie americana del género *Vachellia*, y particularmente, para *V. aroma*. Por lo tanto, los objetivos de este estudio fueron desarrollar marcadores SSR específicos para *Vachellia aroma* por primera vez, y probar su amplificación en una especie emparentada.

M&M: Para el desarrollo de los microsatélites en *V. aroma*, se extrajo ADN genómico total de un individuo y este fue secuenciado en un cuarto de placa en una plataforma Roche 454 GS FLX+. El área de estudio comprendió dos poblaciones argentinas: San José y Robles, provincia de Santiago del Estero.

Resultados: Detectamos 422 secuencias que contenían *loci* microsatélites. Un set de 39 pares de cebadores presentó amplificación de productos en *Vachellia aroma* y *Vachellia caven*, pero doce de ellos mostraron un patrón claro, reproducible y polimórfico en *V. aroma*.

Conclusiones: Los resultados de este trabajo indican que se ha desarrollado un nuevo set de marcadores SSR específicos para *V. aroma* y su transferencia a *V. caven* fue exitosa. El análisis de variabilidad mostró que estos 12 marcadores polimórficos son altamente informativos, constituyendo una poderosa herramienta para investigar parámetros de la genética de poblaciones en *V. aroma* y especies relacionadas.

Palabras clave

454 pirosecuenciación, SSR polimórficos, *Vachellia aroma*, *Vachellia caven*.

Recibido: 24 Mar 2023
Aceptado: 4 Mar 2024
Publicado impreso: 31 Mar 2024
Editora: Viviana Solis Neffa

ISSN versión impresa 0373-580X
ISSN versión on-line 1851-2372

Introduction

Vachellia aroma, synonym *Acacia aroma* (Gillies ex Hook. & Arn.) Seigler & Ebinger (Seigler & Ebinger, 2006), (Fabaceae, Caesalpinoideae) is a tree that inhabits Central and South America, and the Caribbean. In South America, the applications for this tree are several. The wood has many uses, like charcoal and fuel, and more important, construction of hard structures (del Valle Perea *et al.*, 2007). A study of the properties of its wood, yielded to a hard, dense and durable one with the possibility of even more potential uses (Pometti *et al*., 2009). Furthermore, its twigs and leaves are used in traditional medicine as infusion to treat gastritis, liver and stomach disorders, and as a digestive aid (Carrizo *et al.*, 2005); also, the leaves have antiseptic properties, useful in several affections of skin, throat, eyes and canker (Martínez Crovetto, 1981; Cialdella, 1984; Alonso & Desmarchelier, 2005; del Valle Perea *et al*., 2007). As regards its cultivation, it adapts to silvopastoral production systems. Their fruits and leaves serve as forage for goats, sheep and cattle (Demaio *et al*., 2002; del Valle Perea *et al*., 2007), but the size of the thorns turn the access to the higher branches very difficult. In this line, several years ago, we started an improvement program of controlled crossings aiming to reduce the size of the thorns in the Estación Experimental Fernández, Santiago del Estero, Argentina.

The microsatellites or Simple Sequence Repeat (SSR) are short sequences, from two to six nucleotides, tandem repeated that can vary in number and sequence across the different individuals. The main advantage that these markers offer for the genome analysis is their abundance and dispersion in the DNA. Its value consists in their multialellic nature and codominat inheritance, allowing the detection of several variants in the population. The SSR can be found in all the eukaryote organisms, they are highly polymorphic, and, although they are genome-specific, they could be transferred across species, according to the phylogenetic distance between them and the *loci* conservation (Ferreira, 2003). Therefore, these markers have become a very valuable tool for the genetic mapping of species, diversity analyses, studies of gene flow and differentiation between populations, and genetic improvement and breeding (Contreras *et al*., 2020).

There are currently, to our knowledge, no microsatellite markers available for any American species of *Vachellia*, and particularly, for *V. aroma*. Then, this motivated us to the aims of this study that were to develop, for the first time, SSR markers specific for *Vachellia aroma*, an American species, and test its amplification in a close related species.

Materials and Methods

One silica gel-dried leaf sample of *V. aroma* belonging to one individual from Robles, Santiago del Estero, Argentina (S 63°58' 59.76"; W 28° 3' 12.78") was used in a simple sequence repeat (SSR) scan at the whole genome level. Total genomic DNA was extracted with Mini Plant DNAeasy Kit (Quiagen) according to manufacturer manual. Then, it was sequenced in a one-fourth run on a Roche 454 GS FLX+ platform (454 Life Sciences, a Roche Company, Branford, Connecticut, USA) by the INDEAR, Rosario, Argentina, service.

In order to identify microsatellite sequences in the contigs obtained, we used the software MSATCOMMANDER v. 0.8.2 (Faircloth, 2008). We used the option Design primer, in which the software searches for microsatellite repeats and identifies possible primer annealing sites in one step. Primer3 (Rozen & Skaletsky, 2000) is implemented in MSATCOMMANDER for primer design according to the following criteria: amplification products within the size range of 100-500 bp, optimal melting temperature (range 57-62 ºC), optimal GC content of 50%, possession of at least 1-bp GC clamp, low levels of self- or pair-complementary, and maximum end stability (D G) of 8.0 (Faircloth, 2008).

Amplification through PCR was performed on individual *loci* in 50µl reactions containing 10/30 ng DNA, 0.6 mM each primer, 0.2 mM dNTPs, 0.3 U Taq DNA polymerase (Invitrogen, Buenos Aires, Argentina), and 1.5 mM $MgCl_2$. A T21 IVEMA thermal cycler (IVEMA, Buenos Aires, Argentina) was used for amplifications, where the cycling profile was initial denaturation at 94 ºC for 5 min; followed by 40 cycles at 94 ºC for 45s denaturation, primer-specific annealing temperature (53-58 ºC; see Table 1) for 45 s, extension at 72 ºC for 45 s; and a final extension at 72 ºC for 10 min. PCR products were run in a 6% (w/v) polyacrylamide gel containing 5 M urea in 1% TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM EDTA, pH 8) with a 10-bp DNA Ladder (Invitrogen) size marker. Gels were

Table 1. Characteristics of the polymorphic microsatellite markers developed in *Vachellia aroma*. Abbreviations= F: forward; R: reverse; Ta: annealing temperature.

Locus	Primer sequence (5´-3´)	Ta (ºC)	Motif	Product size (bp)	Accession
TU03	F: TGAGAGTTAGCGGTACAGGC	56	$(AATC)_4$	300- 306	OR231272
	R: GCCAACTCTAAGAAGCAGCG				
TU04	F: ATCCCATTCGATCCCTGAGC	56	$(AATG)_4$	382- 390	OR231273
	R: GTGTTGTAAGTCGCGACCTG				
TU08	F: TCCCTTTACTTCCCTTGGCC	55	$(AACTG)_4$	385- 390	OR231274
	R: AGAGAGCAAGACAGTCCAGC				
TU13	F: CGACCCACACTTGAGAAAGC	54	$(ATC)_4$	213-222	OR231275
	R: CGCCATGATTCTTCTCCTGC				
TU39	F: TCCTCCTTCAGCACTCCTTC	57	$(AACTG)_4$	400- 410	OR231276
	R: CTCAGAAGTGCAGGGAGGAG				
TU61	F: CATGTCCGTTGTCCCTTCTC	53	$(AAAAT)_4$	284- 294	OR231277
	R: CTCTCTTTGCAGCACCTTGG				
TU80	F: CAGAAAGCCGAGCAAGAGTG	58	$(ACTG)_4$	326- 334	OR231278
	R: TCAGGGTCAAGAAACGTTGTG				
TU114	F: TTTCTTCCCATCCAAGCGG	55	$(AAAAG)_4$	357- 362	OR231279
	R: CCTTTACACCATCACCGCAC				
TU123	F: CATCTAGAGGCTTTGCTTGGTG	57	$(ACGC)_5$	324- 328	OR231280
	R: TGACATACTTCTGCTGGGCG				
TU131	F: GGGAGTGCGCATTAAGATCAG	57	$(AAAG)_5$	317- 325	OR231281
	R: GATTCATGCATGTCCCACAAG				
TU134	F: GTGGCAGACTCACCTCCTAC	56	$(AAAT)_5$	354- 358	OR231282
	R: GACGAGGAGTTCAATGTTTGTC				
TU154	F: TCGGTTGCATAATTGGAGGG	54	$(AAAGG)_4$	170- 180	OR231283
	R: TCGAAGGCCAACTGTAATATCC				

stained with silver nitrate. For all the polymorphic *loci* that generated clear and reproducible patterns, we calculated observed, expected, and unbiased expected heterozygosity, the effective number of alleles, Nei's genetic distance (Nei, 1978), Wright (1978) fixation indices (F_{IS} or inbreeding coefficient and F_{ST} that estimates the genetic differentiation between populations) and Chi-Square Hardy-Weinberg equilibrium test using GenAlEx 6.5 (Peakall & Smouse, 2012) and the polymorphism information content (PIC) using MolMarker (Jahnke *et al*., 2022). The PIC index describes diversity within populations and characterizes the degree of polymorphism at each locus (Botstein *et al*., 1980). In order to calculate all these indices, we used 20 individuals in total, belonging to two Argentinean populations (10 from San José (SJ): S 63°49'54.90"; W 27°52'52.92" and 10 from Robles (RO): S 63°58'59.76"; W 28° 3'12.78", Santiago del Estero province). The primers were also assayed in two individuals of *Vachellia caven* (Molina) Seigler & Ebinger, a related species, to test cross amplification.

Results and Discussion

The 454 sequencing of *V. aroma* was assembled into 9449 contigs where we detected 422 sequences containing SSR *loci*; 239 (56.7%) di-, 113 (26.8%) tri-, 15 (3.5%) tetra-, 27 (6.4%) penta-, 12 (2.8%) hexanucletide and 16 (3.8%) with complex motifs.

The most frequent number of tandem repeats within dinucleotide motifs was 8, in the case of trinucleotides and tetranucleotides the most frequent configuration included 5 repeats, and finally, 4 times was most frequent for penta- and hexanucleotides.

A set of 39 primer pairs was selected according to ascending order on primer pair penalty value to test the amplification success. All these primers presented amplified products in *V. aroma* and *V. caven* with the same intensity, but only 12 revealed clear and polymorphic *loci* in *V. aroma* (Table 1) (the polymorphism was not tested in *V. caven* because we only tested the cross amplification).

Across the sample size, we found between two to four alleles (A), with an average of 2.417 and 2.5 alleles in San José and Robles, respectively. In all cases, Ne is lower than A, the mean number of effective alleles (Ne) was 1.787 and 1.911 in San José and Robles respectively, ranging from 1.308 to 2.916 for the two populations. Mean heterozygosity was 0.336/0.425, 0.412/0.463 and 0.433/0.488, in San José/Robles (Observed Heterozygosity [H_O], Expected Heterozygosity [He] and Expected and Unbiased Heterozygosity [uHe], respectively) (Table 2). All these estimates are similar to those calculated for the SSRs of *Acacia nilotica* ssp. *indica* (synonym *Vachellia*

Table 2. Descriptive statistics for polymorphic microsatellites in *Vachellia aroma*. Abbreviations= N: Sample size; A: Number of different alleles; Ne: Number of effective alleles, 1 / (Sum pi^2); Ho: Observed heterozygosity, No. of Hets / N; He: Expected heterozygosity, 1 - Sum pi^2; uHe: Unbiased expected heterozygosity, (2N / (2N-1)) * He; F_{is}: Fixation index; P H-W test: P value for the Hardy-Weinberg equilibrium test (the asterisks show the level of significance for Hardy-Weinberg equilibrium test: *P< 0.05, **P<0.01); PIC: Polymorphic information content; SE: Standard error; 95% Confidence interval of F_{is} between square brackets.

Population	Locus	N	A	Ne	Ho	He	UHe	Fis	P H-W test	PIC
San José	TU13	10	3	1,847	0,545	0,459	0,481	-0,189	0,300	0,387
	TU61	10	3	2,180	0,091	0,541	0,567	0,832	0,011*	0,539
	TU123	10	2	1,862	0,182	0,463	0,485	0,607	0,044*	0,356
	TU131	10	3	2,916	0,455	0,657	0,688	0,308	0,427	0,598
	TU80	10	3	1,322	0,273	0,244	0,255	-0,119	0,965	0,228
	TU39	10	2	1,308	0,273	0,236	0,247	-0,158	0,601	0,208
	TU154	10	2	1,342	0,100	0,255	0,268	0,608	0,055	0,222
	TU03	10	2	1,541	0,455	0,351	0,368	-0,294	0,329	0,290
	TU04	10	3	1,942	0,400	0,485	0,511	0,175	0,445	0,409
	TU08	10	2	1,724	0,600	0,420	0,442	-0,429	0,175	0,332
	TU134	10	2	1,862	0,364	0,463	0,485	0,214	0,477	0,356
	TU114	10	2	1,600	0,300	0,375	0,395	0,200	0,527	0,305
	Mean	10,000	2,417	1,787	0,336	0,412	0,433	0,146 [-0,078/0,371]		0,445
	SE	0,000	0,149	0,129	0,047	0,037	0,039	0,114		0,260
Robles	**TU13**	10	4	2,564	0,800	0,610	0,642	-0,311	0,844	0,556
	TU61	10	3	1,681	0,100	0,405	0,426	0,753	0,004**	0,379
	TU123	10	2	1,835	0,300	0,455	0,479	0,341	0,281	0,351
	TU131	10	3	1,361	0,100	0,265	0,279	0,623	0,018*	0,247
	TU80	10	3	2,062	0,700	0,515	0,542	-0,359	0,407	0,462
	TU39	10	2	1,923	0,400	0,480	0,505	0,167	0,598	0,365
	TU154	10	2	1,724	0,200	0,420	0,442	0,524	0,098	0,332
	TU03	10	2	1,600	0,500	0,375	0,395	-0,333	0,292	0,305
	TU04	10	3	2,299	0,600	0,565	0,595	-0,062	0,963	0,485
	TU08	10	2	1,923	0,600	0,480	0,505	-0,250	0,429	0,365
	TU134	10	2	1,980	0,700	0,495	0,521	-0,414	0,190	0,372
	TU114	10	2	1,980	0,100	0,495	0,521	0,798*	0,012*	0,372
	Mean	10,000	2,500	1,911	0,425	0,463	0,488	0,123 [-0,142/0,387]		0,383
	SE	0,000	0,195	0,091	0,075	0,026	0,027	0,135		0,083

nilotica) (Number of alleles between 2-3, Ho ranging from 0- 0.493, He ranging from 0.139-0. 493), an African species that also belongs to the genus *Vachellia* (Wardill *et al*., 2004). The values of heterozygosity, particularly He, were larger than that estimated from 401 AFLP *loci* in six Argentinean populations of *V. aroma* (He= 0.2) (Pometti *et al*., 2018). The Hardy-Weinberg equilibrium was tested *locus* by *locus*, and showed that two out of twelve *loci* were not at the equilibrium for San José and three out of twelve were not for Robles (Table 2). In those cases, where the loci are not in equilibrium, a possible cause could be the excess of homozygotes for each locus. This could be concluded because of the F_{is} values (Table 2). However, the fixation index F_{is} (or inbreeding coefficient) was also estimated in the two populations and the mean value was 0.146 and 0.123 for SJ and RO, respectively, indicating that although the estimates are positive suggesting some amount of inbreeding, both populations are in Hardy-Weinberg equilibrium because its 95% confidence interval contains the 0, P>0.05 in both cases (Table 2). Estimates of Nei's genetic distance (0.091) and Wright's F_{ST} (0.051) between populations showed an appreciable differentiation between them with the SSR markers, and were also consistent with previous estimates for the same populations with AFLP markers (Pometti *et al.*, 2018). In *V. aroma*, the PIC ranged from 0.208 to 0.598 in San José and from 0.247 to 0.556 in Robles (Table 2). PIC values lower than 0.25 indicate low polymorphism, values between 0.25 and 0.5 indicate average polymorphism and values higher than 0.5 indicate high polymorphism. Two *loci* were highly polymorphic for SJ (TU61 and TU131) and three for RO (TU13, TU80 and TU04), the remaining ones were average polymorphic for both populations. Although, these results could be biased by the use of polyacrylamide high resolution gels instead of automatic sequencer, which could underestimate the parameters of genetic diversity, here the data set can be considered useful and can in the future replace other more laborious markers. In future works, a sub-sample of these 12 loci could be used for population-genetics studies, based in the higher A and He, for example TU03, TU04, TU08, TU13, TU80 and TU 131.

Conclusions

The results of this work indicate that a set of SSR markers was developed, for the first time, for *V. aroma*, an American species of the genus *Vachellia* and their transferability to *V. caven* was assessed. The analysis of variability, shows that these 12 polymorphic markers are highly informative, and a powerful tool to investigate population genetics parameters in *V. aroma* and related species.

In conclusion, the present SSR markers could be used for estimating genetic variability and structure, include in developing breeding programs, and in population studies where informative markers are needed, with the potential transferability across the genus *Vachellia*.

Authors' contributions

Conceptualization: CLP; Formal analysis: CLP, CFB; Funding acquisition: CLP, CFB; Investigation: CLP, CFB, ME; Methodology: CLP; Project administration: CLP, CFB; Supervision: CLP; Visualization: CLP, CFB, ME, JCV, BOS; Writing-original draft: CLP, CFB, JCV, BOS

Acknowledgments

The present work was conducted thanks to the following funding: Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) PIP 11220200100477CO to CLP; ANPCyT PICT-2021-00307 to CLP and PICT 2020-1402 to CFB; Universidad de Buenos Aires, UBACYT 20020190200106BA to CFB.

Competing interests

The authors have declared that no competing interests exist.

Bibliography

ALONSO, J., & C. DESMARCHELIER. 2005. Plantas medicinales autóctonas de la Argentina. Bases científicas para su aplicación en atención primaria de la salud, pp. 136-139. Ed. L.O.L.A., Buenos Aires.

BOTSTEIN, D., R. D. WHITE, M. SKOLNICK, & R. W. DAVIS. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32: 314-331.

CARRIZO, E. DEL V., M. O. PALACIO, & L. D. ROIC (ex aequo). 2005. Uso medicinal de algunas especies nativas en Santiago del Estero (República Argentina). *Dominguezia* 21: 25-32.

CIALDELLA, A. M. 1984. El género *Acacia* (Leguminosae) en la Argentina. *Darwiniana* 25: 59-111.

CONTRERAS, R., F. S. CAREVIC, V. PORCILE & M. ARIAS. 2020. Development of SSR *loci* in *Prosopis tamarugo* Phillipi and assessment of their transferability to species of the *Strombocarpa* section. *Forest Systems* 29: 10.

DEL VALLE PEREA, M., G. PEDRAZA, & J. DEL VALLE LUCEROS. 2007. Relevamiento de la flora arbórea autóctona en la provincia de Catamarca, pp. 38-40. Consejo Federal de Inversiones, Buenos Aires.

DEMAIO, P., U.O. KARLIN, & M. MEDINA. 2002. Árboles nativos del centro de Argentina, Ed. L.O.L.A., Buenos Aires.

FAIRCLOTH, B. C. 2008. MSATCOMMANDER: Detection of microsatellite repeat arrays and automated, locus-specific primer design. *Mol. Ecol. Resour.* 8: 92-94. https://doi.org/10.1111/j.1471-8286.2007.01884.x

FERREIRA, M. A. J. da F. 2003. *Utilizaçao das técnicas de marcadores moleculares na genética de populaçoes, na genética quantitativa e no melhoramento de plantas*. Embrapa Roraima, Boa Vista.

JAHNKE, G., J. SMIDLA, & P. POCZAI. 2022. MolMarker: A Simple Tool for DNA Fingerprinting Studies and Polymorphic Information Content Calculation. *Diversity* (Basel) 14: 497. https://doi.org/10.3390/d14060497

MARTÍNEZ CROVETTO, R. 1981. Las plantas utilizadas en medicina popular en el noroeste de Corrientes. *Miscelánea Fundación Miguel Lillo* 69: 7-139.

NEI, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583-90.

PEAKALL, R., & P. E. SMOUSE. 2012. GenAlEx 6.5. *Bioinformatics* 28: 2537-2539. https://doi.org/10.1093/bioinformatics/bts460

POMETTI, C. L., B. PIZZO, M. BRUNETTI, N. MACCHIONI, … & B. O. SAIDMAN. 2009. Argentinean native woods species: Physical and mechanical characterization of some *Prosopis* species and *Acacia aroma* (Leguminosae; Mimosoideae). *Bioresour. Technol.* 100: 1999-2004. https://doi.org/10.1016/j.biortech.2008.09.061

POMETTI, C., C. BESSEGA, A. CIALDELLA, M. EWENS, … & J. VILARDI. 2018. Spatial genetic structure within populations and management implications of the South American species *Acacia aroma* (Fabaceae). *Plos One* 13: e0192107. https://doi.org/10.1371/journal.pone.0192107

ROZEN, S. & H. J. SKALETSKY. 2000. Primer3 on the WWW for general users and for biologist programmers. In: MISENER, S. & S. A. KRAWETZ (eds.), Methods in molecular biology, vol. 132: Bioinformatics methods and protocols, 365-386. Humana Press, Totowa, New Jersey. https://doi.org/10.1385/1-59259-192-2:365

WARDILL, T. J., K. D. SCOTT, G. C. GRAHAM, & M. P. ZALUCKI. 2004. Isolation and characterization of microsatellite *loci* from *Acacia nilotica* ssp. *indica* (Mimosaceae). *Mol. Ecol. Notes* 4: 361-363. https://doi.org/10.1111/j.1471-8286.2004.00697.x

WRIGHT, S. 1978. *Evolution and the genetics of populations variability within and among natural populations*, vol. 4. University of Chicago Press, Chicago.