

This document is a self-archived version of published work. For the published version see:

Vaughn, E.E., J. F. Dwyer, M. Culver, and J. Morrison (2015) Development and characterization of polymorphic microsatellite markers for the crested caracara, *Caracara cheriway*. *Conservation Genetics Resources* 7(2):557-559. doi:10.1007/s12686-015-0423-9

Word Count: 797

Development and characterization of polymorphic microsatellite markers for the crested caracara, *Caracara cheriway*

Erin E. Vaughn*§, Joan L. Morrison+, and Melanie Culver§^

*Genetics Graduate Interdisciplinary Program, University of Arizona, Tucson, AZ 85721 USA

§Conservation Genetics Laboratory, School of Natural Resources and Environment, University of Arizona, BSE-325, Tucson, AZ, 85721 USA

+Department of Biology, Trinity College, 300 Summit St., Hartford, CT, 06106 USA

^ U.S. Geological Survey, Arizona Cooperative Fish and Wildlife Research Unit

Corresponding author: Erin Vaughn office: (520)626-1636, fax: (520)621-8801, evaughn@email.arizona.edu

Abstract

We isolated novel microsatellites from the crested caracara (*Caracara cheriway*) with a shotgun pyrosequencing approach. We tested 80 loci for polymorphism among 20 individuals from the threatened Florida population. Fourteen loci were polymorphic. The mean number of alleles was 2.21 (range 2-3) and the mean observed heterozygosity was 0.41 (range 0.15-0.65). None of the 14 polymorphic loci exhibited significant linkage disequilibrium nor did they deviate significantly from Hardy-Weinberg expectations. We also report 16 monomorphic loci.

Key words: microsatellite, Falconidae, Florida, 454 pyrosequencing

The crested caracara (*Caracara cheriway*) is a medium-sized raptor found commonly throughout Mexico, Central America, and northern regions of South America; it also occurs sparsely in the southern United States. There are currently stable populations in Texas while a small population in Florida is listed as threatened (Morrison and Dwyer 2012). The decline of the Florida population is primarily attributed to habitat loss due to rapid urbanization of agricultural and pasture lands (Morrison and Humphrey 2001). To date, no genetic analyses have been conducted on any population of caracara. To supplement current efforts to restore the Florida population and measure genetic diversity within the species, we have developed microsatellite primers from *Caracara cheriway* genomic data.

We used 2 µg of genomic DNA to prepare a library for sequencing at the University of Arizona Genetics Core. We used ¼ plate on the 454 GS FLX platform (Roche Applied Science) to sequence our library. This run generated 371,387 raw individual reads for a total of 137.8 Mbp. The reads were trimmed to eliminate the first 4bp corresponding to the library tag. Also, the 3' end of each sequence was trimmed until the Q20 criterion was met over a 10 bp window. The resulting mean read length was 372 bases (S.D. = 160). We used the program QDD (Meglécz et al. 2010) to search within quality-filtered reads to identify and design primers for microsatellite loci containing a minimum of six perfect repeats for dinucleotides and five perfect repeats for larger motifs. From 3801 dinucleotide, 1424 trinucleotide, 392 tetranucleotide, 195 pentanucleotide, and 49 hexanucleotide repeat loci found, QDD designed primers for 1629 unduplicated loci. From these unduplicated loci, we selected tetranucleotide repeat loci with greater than six repeats (41 loci) and dinucleotide repeat loci with greater than 11 repeats (39 loci) for primer synthesis. We added the M13 primer sequence at the 5' end of all forward primers for economical fluorescent labeling and a "pig-tail" sequence (GTGTCTT) to the 5' end of all reverse primers to reduce ambiguity in fragment size due to incomplete adenylation. We conducted PCRs in 15 µl volumes with 10-20 ng genomic DNA, 1X PCR buffer, 0.2 mM each dNTP, 1.5 mM MgCl₂, 0.5 U *Taq* DNA Polymerase (New England Biolabs), 0.2 µM unlabeled M13-tailed forward primer, 1 µM 6-FAM-labeled M13 primer, and 1 µM reverse primer. We used a touchdown PCR protocol consisting of 94°C for 5 min, 12 cycles of 95°C for 30 s, 67-55°C for 20 sec (1°C decrease each cycle), 72°C for 20 sec, followed by 33 cycles at 95°C for 30 s, 55°C for 20 sec, 72°C for 20 sec, and a final extension of 72°C for 30 min. We used an ABI3730 DNA analyzer (Applied Biosystems) to genotype PCR products and scored alleles in GENOTYPER version 3.7 (Applied Biosystems). To estimate deviations from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD) we used GENEPOP version 4.2 (Rousset 2008) and corrected for multiple comparisons using a sequential Bonferroni test ($\alpha = 0.05$). Finally, we used MICRO-CHECKER version 2.2.3 (Van Oosterhout et al. 2004) to test for null alleles and allele drop-out.

Thirty loci successfully amplified, of which, 14 (four tetranucleotide and ten dinucleotide motifs) were polymorphic within a set of 20 individuals from the Florida caracara population (Table 1). The mean number of alleles per locus was 2.21 (range 2-3) and the mean observed heterozygosity was 0.41 (range 0.15-0.65) (Table 1). We detected no significant LD among loci pairs, deviation from HWE, or evidence for null alleles and allele drop out (adjusted $p > 0.05$). Sixteen loci (seven tetranucleotide and nine dinucleotide motifs) amplified well but were monomorphic within the same set of 20 individuals. We have reported the monomorphic loci here (Table 1) as they may be found to be polymorphic in other populations of crested caracara. Sequences for all 30 loci are provided (Online Resource 1).

Acknowledgements and Disclaimers

We would like to thank Debbie Chen and Mila Crevar for their assistance in primer testing. Any use of trade, firm, or product names is for descriptive purpose only and does not imply endorsement by the U.S. Government.

References

- Meglécz E, Costedoat C, Dubut V, Gilles A, Malausa T, Pech N, and Martin J (2010) QDD: a user friendly program to select microsatellite markers and design primers from large sequencing projects. *Bioinformatics* 26(3):403-404
- Morrison JL and Dwyer JF (2012) Crested caracara (*Caracara cheriway*), *The birds of North America online* (A. Poole, Ed.). Ithaca: Cornell Lab of Ornithology; retrieved from the birds of North America online: <http://bna.birds.cornell.edu/bna/species/249>

Morrison JL and Humphrey SR (2001) Conservation value of private lands for crested caracaras in Florida. *Conserv Biol* 15:675-684

Rousset F (2008) GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Mol Ecol Resour* 8(1):103-106

Van Oosterhout C, Hutchinson WF, Wills DPM, and Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* 4:535-538

Table 1. Characterization of 30 microsatellite loci isolated from *Caracara cheriway* and tested in 20 individuals from Florida.

Locus	GenBank accession no.	Repeat motif	Primer sequences	Size range (bp)	No. alleles	H _o	H _e
CRCA1	KJ169554	AAAT	F: GCAATACATCGGAGACAGGC R: TGTTGGTCTCAGGGCAGAGT	289-297	3	0.60	0.66
CRCA2	KJ169555	AAAT	F: GCGTGATAGGTGAAATGCAA R: TCAGGCTTGGAAATCCTGTTT	146-150	2	0.15	0.14
CRCA3	KJ169556	AGGT	F: TCTGGAAATCCGTATCTACCCA R: AACTGAAATTTGTCTTGCAGAGG	136-144	3	0.50	0.49
CRCA4	KJ169557	ACGG	F: AGTCTGAACGCTGAAGTGTAGG R: AGCCTTGCAATGAAGCAAT	117-121	2	0.45	0.47
CRCA5	KJ169558	AC	F: ACCTCTGGCAAGAGTGTCA R: GGTGGGACTGTGCAATGTAA	196-202	2	0.20	0.18
CRCA6	KJ169559	AC	F: GCAGCAATGTACTCTTTCAG R: TGCAGAGAGATCGGGATGTT	164-174	2	0.25	0.35
CRCA7	KJ169560	AC	F: GTCGCTGTTAGTACCGTGGC R: TGAACGGTCTATGTGCGCC	130-140	3	0.45	0.53
CRCA8	KJ169561	AC	F: CAGAGGTCTGGGTTTCTGTGA R: CCACTCTGGGAACAGTTTG	301-303	2	0.60	0.5
CRCA9	KJ169562	AC	F: TCCATAAGCCTCACCACCAT R: GCTTTCAGCTGCCAGTCAGT	123-129	2	0.45	0.49
CRCA10	KJ169563	AC	F: AATCTAGCTCCCAGCCAAGC R: CATGCTTCCTAGAGTTCGCTTT	163	1	-	-
CRCA11	KJ169564	AC	F: TTGGTTCCTCGTTTACTACACC R: TCAAGACTTGCCACCCTCAT	147-151	2	0.45	0.35
CRCA12	KJ169565	AC	F: TGATCTCTGGTGTGCTTGTAGG R: AAGAATTAGAAGTGGTCTTTCG	108-114	2	0.45	0.49
CRCA13	KJ169566	AT	F: AGGGTCTGGATAGATAAGGCTC R: AGAACAGCATTTCCTCTGCG	263-265	2	0.65	0.50
CRCA14	KJ169567	AC	F: GGAATGTGTCAGAACAGTTTGC R: GCAGCACCTGTAAGAATGG	312-316	2	0.20	0.32
CRCA15	KJ728767	AC	F: TGCAAGTGGTGAAGGAAGT R: GGGCAGCAGAACAGCAGTAT	207-209	2	0.30	0.26
CRCA16	KJ728768	AAAC	F: TGAAAGAACAGTAGCATCCCAG R: TGAACAGTTTAACATCAAATCCTGA	285	1	-	-
CRCA17	KJ728769	AAAC	F: TGATGTTGACCTTCCTGCAC R: CATCAGAAATCAGGGATGGG	176	1	-	-
CRCA18	KJ728782	AAAT	F: TCCGTGATATTGCACGCTTA R: CCGTCAGCTTCCGTAGTTTG	215	1	-	-
CRCA19	KJ728770	ACAG	F: CTTCCACAAGATGGAGCCTG R: GCCATTCTGCATCGCTCTC	182	1	-	-
CRCA20	KJ728771	AAAT	F: GAGGCCTTTCGAAGTGAAT R: TTTCTCCTTGCTGGCAACTT	190	1	-	-
CRCA21	KJ728772	AAAT	F: AGGATGGACAGCCTTCATCA R: CCATTCAGGGAGTGACAGG	175	1	-	-
CRCA22	KJ728773	ACCT	F: ATAGCTCAGCAGGTCCACA R: CGCTACCATTTCAGGCTTA	306	1	-	-
CRCA23	KJ728774	AC	F: GCCATGTGTAGCAAAGGACC R: GCGTGTGTGCTTGCATGTAT	144	1	-	-
CRCA24	KJ728775	AC	F: TATGCAATGTGCAGGCATGT R: ACCCGTACTCAGCAATCTG	115	1	-	-
CRCA25	KJ728776	AC	F: TTACTTCAGTCATGTCTTCTTT R: CGGGAATCCTACAGGTAGCC	169	1	-	-
CRCA26	KJ728777	AG	F: ACAACTGCCTGAATTCACA R: TAGATGCTCCAGCCTGCATT	185	1	-	-
CRCA27	KJ728778	AC	F: TTTCTACATGCACGACGG R: TGGCTACTAGCACTGCTGCTT	119	1	-	-
CRCA28	KJ728779	AG	F: TGCTCACTTGCCCCAAATGTT R: CTGGAGAATCTGCCGTATGC	205	1	-	-

CRCA29	KJ728780	AG	F: TTTGTCACAACACTGGTGC R: ATGGCTGGCTCATCTTCTT	321	1	-	-
CRCA30	KJ728781	AC	F: GCTGGACCACCAATAAGC R: TCGGTTAATTATTGCCCTG	298	1	-	-