Journal of the Arkansas Academy of Science

Volume 73

Article 11

2019

De novo Development and Characterization of Tetranucleotide Microsatellite Loci Markers from a Southeastern Population of the House Finch (Haemorhous mexicanus)

Edgar E. Sanchez
Arkansas Tech University, esanchez3@atu.edu

J Dylan Maddox
Field Museum of Natural History, jdylanmaddox@gmail.com

Douglas G. Barron
Arkansas Tech University, dbarron@atu.edu

Cover Page Footnote

We thank Kevin Feldheim for facilitating work in the Pritzker Laboratory and the Arkansas Tech University Undergraduate Research Grant for funding. Partial funding was provided by the Pritzker Laboratory for Molecular Systematics and Evolution with support from the Pritzker Foundation and Grainger Bioinformatics Center. We also thank Bailey Coffelt, Stetson Collard, Kagan Davis, and Stephanie Nefas for field assistance.

De novo Development and Characterization of Tetranucleotide Microsatellite Loci Markers from a Southeastern Population of the House Finch (*Haemorhous mexicanus*)

E.E. Sanchez[1,*], J.D. Maddox[2,3], and D.G. Barron[1]

[1]*Department of Biological Sciences, Arkansas Tech University, Russellville, AR 72801*
[2]*Pritzker Laboratory for Molecular Systematics and Evolution, Field Museum of Natural History, 1400 S. Lake Shore Drive, Chicago, IL 60605*
[3]*Environmental Sciences, American Public University System, Charles Town, WV 25414*

*Correspondence: esanchez3@atu.edu

Running Title: House Finch Microsatellite Development

Abstract

Microsatellites are short tandem repeats (e.g. TAGATAGA) of base pairs in a species' genome. High mutation rates in these regions produce variation in the number of repeats across individuals that can be utilized to study patterns of population- and landscape-level genetics and to determine parentage genetically. In this project our objective was to develop microsatellite markers for the House Finch, *Haemorhous mexicanus*. This species has become one of the most well-studied species of songbirds due to its unique geographical, evolutionary, and epidemiological history. Using mist-nets we captured birds on the Arkansas Tech University campus and collected blood samples to obtain genomic DNA. Samples were processed in The Field Museum's Pritzker Laboratory for Molecular Systematics and Evolution, where we fragmented genomic DNA and isolated fragments that contained potential microsatellites using specially designed biotin labelled probes. These DNA fragments were transformed into competent *E. coli* cells which were then PCR-amplified and Sanger sequenced. After sequencing DNA fragments from approximately 500 *E. coli* colonies, we designed and characterized a set of 13 tetranucleotide microsatellite loci. The average number of alleles and heterozygosity found in 12 individuals from Arkansas was 8.69 and 0.80, respectively. This finalized set of microsatellites can be utilized by researchers to determine parentage and characterize genetic differences across House Finch populations.

Introduction

House Finches (*Haemorhous mexicanus*) are one of the most common and well-studied passerine species in North America. Their geographic range was originally restricted to the arid southwest, though in 1939 they were introduced to Long Island, NY (Aldrich and Weske 1978). Since this time the species' native (western) and introduced (eastern) ranges have each expanded dramatically to make them common breeders throughout most of the United States. The initial introduction caused a molecular founder effect (Hawley *et al*. 2006) and produced substantial variation in genetics (Hawley *et al.* 2006), morphology (Bock and Lepthein 1976; Shultz *et al.* 2016), and physiology (Bock and Lepthein 1976) across their range. Low genetic diversity among House Finches may decrease population fitness (Reed and Frankham 2003; Briskie and Mackintosh 2004) and their susceptibility to parasites and pathogens in these populations (Hedrick *et al.* 2001; Acevedo-Whitehouse *et al*. 2003). In fact, Hawley et al. (2005, 2006) suggested that reduced genetic diversity in the eastern population of House Finches may have contributed to their susceptibility to *Mycoplasma gallisepticum*. This poultry pathogen was first reported in Washington D.C. in the mid-1990s (Dhondt *et al*. 1998) and has since spread across eastern and western populations and caused dramatic losses in many House Finch populations (Hawley *et al*. 2006). In recent decades, this species has become a model organism for studies of population genetics, invasion biology, and disease ecology.

Studies of House Finches have also been instrumental in formulating our understanding of the evolution and maintenance of sexual signals in songbirds. Male finches express carotenoid-based pigmentation that can range from yellow to red (Hill 1993). Pigment concentration and corresponding extent and hue of colorful plumage varies widely across populations (Hill 1993; Inouye *et al.* 2001). While some individual- and population-level variation is explained by diet composition (Hill 1992; Inouye *et al*. 2001; Hill *et al*. 2002), correlative and experimental studies have demonstrated the complex nature of this connection and suggested an additional role for physiological and/or genetic mechanisms of control (Hill 1993, 2002). Multiple studies have demonstrated an association between male plumage coloration and reproductive success (Hill *et al*. 1999, Badyaev *et al*. 2001). Although this species forms socially monogamous pairs, up to 10% of their offspring may be the product of extra-pair matings (Oh and Badyaev 2006). Estimates of male reproductive success must therefore differentiate within-pair versus extra-pair offspring using genetic techniques.

To further advance our understanding of this species' population, landscape, disease, and evolutionary ecology we must utilize molecular markers to characterize genetic variation at the individual and population levels. In this study, we describe the development of microsatellite markers in an Arkansas population of House Finches. Microsatellites are regions of repetitive DNA containing short tandem repeats (e.g. AGATAGATAGAT). While the repeating sequence of base pairs is consistent across individuals, the number of times the sequence repeats can be highly variable across individuals. These non-coding regions are not thought to be under selection and thus can show non-selective patterns of evolutionary divergence.

Although microsatellites have been developed previously for this species (Hawley 2005, Oh and Badyaev 2009), the majority (18 of 25) were dinucleotide repeats which are generally more difficult to score than tetranucleotide due to the presence of shadow or stutter bands (Ginot *et al*. 1996; Daniels *et al*. 1998; Nater *et al*. 2009). Furthermore, few microsatellites have been developed from populations in the southeastern United States despite wide occurrence through this region. It is often preferable to utilize locally-developed microsatellites to avoid problems (e.g. low heterozygosity) resulting from rapid evolution of these loci. Here we describe the characterization of 13 tetranucleotide repeats that, when combined with previously developed microsatellites, will provide a robust microsatellite panel suitable for studies of paternity and population genetics.

Materials and Methods

Field methods

In late 2016 we captured 12 House Finches (3 females, 8 males, 1 unknown sex) using mist nests placed near bird feeders on the Arkansas Tech University campus (35.2945° N, 93.1363° W). We collected blood samples from each bird by puncturing the brachial vein with a hypodermic needle and collecting up to 60 µl upwelling blood in a heparinized capillary tube. Whole blood was applied to non-indicating FTA Elute micro cards (GE WB120410) which lyse cells and denature proteins while preventing DNA degradation. Samples were stored at room temperature (22 °C) until transfer to the Field Museum for genetic processing. Prior to release we aged and sexed the birds and collected basic morphological measurements. Birds were banded with a metal numbered band from the United States Fish and Wildlife Service as well as a passive integrated transponder (PIT) tag and a unique combination of plastic color bands (for related study objectives).

All birds were captured, handled, and released safely and in accordance with procedures approved by the Institutional Animal Care and Use Committee at Arkansas Tech University (approval no. 103116), Arkansas Game and Fish (permit no. 051020161), and United States Fish and Wildlife Service (permit no. 24044).

Microsatellite Enrichment

All laboratory methods were carried out in the Pritzker Laboratory for Molecular Systematics and Evolution at the Field Museum in Chicago, IL. Microsatellite markers were developed following the enrichment protocol of Glenn and Schable (2005). Approximately 1 µg of genomic DNA (gDNA) from one individual was digested with RsaI and XmnI, and SuperSNX24 linkers were ligated onto the ends of gDNA fragments, which act as priming sites for polymerase chain reactions (PCR) in subsequent steps. Five biotinylated tetranucleotide probes [$(AAAT)_8$; $(AACT)_8$; $(AAGT)_8$; $(ACAT)_8$; $(AGAT)_8$] were hybridized with gDNA for 45 min. The biotinylated probe-gDNA complex was added to magnetic beads coated with steptavidin (Dynabeads® M-280 Invitrogen, Carlsbad, California). This mixture was washed twice with 2xSSC, 0.1% SDS and four times with 1xSSC, 0.1% SDS at 52 ºC. For the final two washes, the mixture was incubated for 1 min in a 52 ºC water bath. Between washes, a magnetic particle collecting unit was used to capture the magnetic beads which are bound to the biotin-gDNA complex. This allowed us to capture gDNA containing tetranucleotide repeats while other fragments (i.e. those not containing repeats) were washed away. Enriched fragments were removed from the biotinylated probe by denaturing at 95 ºC and precipitated with 95% ethanol and 3M sodium acetate. To increase the proportion of enriched fragments, a "recovery" PCR was performed in a 25 µl reaction containing 1X PCR buffer (10 mM Tris-HCl, 50 mM KCl, pH 8.3), 1.5 mM $MgCl_2$, 0.16 mM of each dNTP, 10X BSA, 0.52 µM of the SuperSNX24 forward primer, 1 U *Taq* DNA polymerase, and approximately 25 ng enriched gDNA fragments. Thermal cycling, performed in an MJ Research DYAD, was performed as follows: 95 ºC for 2 min followed by 25 cycles of 95 ºC for 20 s, 60 ºC for 20 s, and 72 ºC for 90 s, and a final elongation step of 72 ºC for 30 min. Subsequent PCR fragments were cloned using the TOPO-TA Cloning® kit following the manufacturer's protocol (Invitrogen). Bacterial colonies containing a vector with gDNA (i.e. white colonies) were used as a template for subsequent PCR in a 25 µl reaction containing 1X PCR buffer (10 mM Tris-HCl, 50 mM KCl, pH 8.3), 1.5 mM MgCl2, 0.12 mM of each dNTP, 10X BSA, 0.25 µM of the M13 primers, and 1 U *Taq* DNA polymerase. Thermal cycling was as follows: an initial denaturing step of 95 ºC for 7 min will be followed by 35 cycles of 95ºC for 20 s, 50 ºC for 20 s, and 72 ºC for 90 s. These PCR products were cleaned using MultiScreen-PCR Filter Plates following the manufacturer's protocol (Millipore, Billerica, Massachusetts). DNA sequencing was performed using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, California). Sequencing reactions were precipitated with ethanol and 125 mM EDTA and run on an ABI 3730 DNA Analyzer. We then developed primers flanking core microsatellite repeats using Primer3 (http://primer3.ut.ee). Forward primers were designed with M13-tails (5′-TGTAAAACGACGGCCAGT-3′) and reverse primers with a "pigtail" (5′-GTGTCTT-3′), the former to incorporate fluorescently labeled M13 primers via PCR (Schuelke 2000) and the latter to adenylate the 3′ end of the forward product (Brownstein et al. 1996).

Genotyping Individuals

Genomic DNA was extracted using DNeasy Blood & Tissue Kits (Qiagen, Hilden, Germany) following the manufacturer's instructions. Microsatellite loci were amplified separately in 10µl reactions using the following two-step thermal protocol: an initial denaturing step at 94 ºC for 4 min followed by 35 cycles of 94 ºC for 15 s, 58 ºC for 15 s, 72 ºC for 45 s, then 8 cycles of 94 ºC for 15 s, 53 ºC for 15 s, 72 ºC for 45 s and a final extension step at 72 ºC for 10 min. Each reaction had a final concentration of 10 mM Tris-HCL, 50 mM KCL, 0.5 mM dNTPs, 1.5 mM $MgCl_2$, 1 µg BSA, 0.16 µM fluorescently labeled M13 primer (6-FAM), 0.04 µM forward primer, and 0.16 µM reverse primer. PCR products were then combined with the ALEXA-725 size standard (Maddox and Feldheim 2014) and run on an ABI 3730 DNA analyzer. Allele sizes were determined using the Microsatellite Plugin (v1.4.6) in Geneious Prime (v2019.0.4) using the local southern sizing method. Loci were tested for Hardy-Weinberg equilibrium using GenAlEx (v6.5; Peakall and Smouse 2006, 2012) and linkage disequilibrium with Genepop (v4.2; Raymond and Rousset 1995; Rousset 2008).

Results and Discussion

A total of 12 House Finch individuals were screened using the 13 microsatellite loci developed here (Table 1). Across all loci, heterozygosity averaged 0.80 ± 0.05 SE (range: 0.58 to 1.00) and the number of alleles 8.69 ± 0.76 SE (range: 4 to 14). All loci were in Hardy-Weinberg equilibrium and no linkage disequilibrium or sex-linkage was detected. To determine the repeatability of our marker set we genotyped the blood sample of a recaptured individual. The sample was blindly processed exactly the same as the other samples and resulted in the same genotype as its original sample.

Even relatively minor scoring errors can affect paternity results. For example, Hoffman and Amos (2005) found that relatively minor error rates of 0.01 per allele could increase incorrect rates of paternity exclusion above 20%. Dinucleotide sequences like those from previously developed House Finch microsatellites (Hawley 2005; Oh and Badyaev 2009) are more difficult to score due to shadow or stutter peaks which can lead to scoring errors. The tetranucleotides used in our research will give more accurate and precise results in terms of paternity analysis. Ultimately, however, scoring error rates of known mother-offspring pairs will be needed to determine 'true' genotyping error rates.

Along with the two tetranucleotides developed by Hawley (2005) and five tetranucleotides developed by Oh and Badyaev (2009), these additional 13 loci will provide a robust market set that should minimize genotyping error rates. The microsatellites developed and characterized herein will enable researchers studying House Finches to more accurately determine genetic paternity and elucidate population- and landscape-level patterns of genetic diversity.

Acknowledgements

We thank Kevin Feldheim for facilitating work in the Pritzker Laboratory and the Arkansas Tech University Undergraduate Research Grant for funding. Partial funding was provided by the Pritzker Laboratory for Molecular Systematics and Evolution with support from the Pritzker Foundation and Grainger Bioinformatics Center. We also thank Bailey Coffelt, Stetson Collard, Kagan Davis, and Stephanie Nefas for field assistance.

Table 1. Characteristics of 13 tetranucleotide microsatellites loci isolated from a southeastern population of the House Finch (*Haemorhous mexicanus*). Sequences have been deposited in GenBank under the accession numbers listed.

Locus	Primer sequence (5'-3')[a]	Repeat motif	N_A	Size range (bp)	H_O	H_E	Accession no.
Haem001	F: TGGACATACCACAACATCTTAGGA R: TGCTCTAGCTTCCAGCCCTA	$(AACT)_{14}$	7	190-223	0.58	0.75	MN333897
Haem036	F: TAGCTGCTGTCAGGAAACCC R: CACAGCACAGCAGAGAGGAA	$(TAGA)_{12}$	8	179-199	1.00	0.81	MN333898
Haem086	F: ACAACATCAATGTCAGGTGATTCA R: ACCTCAAGGACTGGGACACT	$(GGAT)_{14}$	4	351-363	0.58	0.64	MN333899
Haem089	F: ACAGCAAAGAAGATTGTCATGCA R: AGAGAAGCTGAGGGGTCACA	$(AGAT)_{15}$	9	220-264	0.92	0.82	MN333900
Haem092	F: CCCAGAAGAGGGTCAGGAAA R: AGCCTACCCTCTTTAAATTTGAAACC	$(AGAT)_{16}$	9	286-326	0.92	0.79	MN333901
Haem110	F: CAGGAGTGCAGAAGTTGGCA R: ACTTCTGTTGCCATGTTTATCAAT	$(GATA)_{13}$	7	226-250	0.75	0.77	MN333902
Haem137	F: TGCAGAAGTTGGCACGTTTTT R: TACTTGATCCAATTGTGTGGTCT	$(AGAT)_{13}$	7	195-219	0.75	0.77	MN333903
Haem298	F: CGTACAAATGGAAGCTGTGCC R: TGGGTAGTAGCTTTGCTGCC	$(TAGA)_{14}$	10	246-294	1.00	0.85	MN333904
Haem309	F: TCCTGGTCTTTGCTGTTGTGT R: GTCTATGTCTCAGATGCAATGTGC	$(TAGA)_{14}$	8	266-280	0.75	0.84	MN333905
Haem325	F: TCAGTTGGAAGGGACCTAGTC R: TGAGCATCTGGAACATACTCCA	$(TCTA)_{12}$	14	250-350	0.92	0.85	MN333906
Haem326	F: TGATCTCATCTGCATTTATCTTCATTG R: GCTTAGCTACCATGAACCTTGC	$(TCTA)_{13}$	8	165-200	0.67	0.78	MN333907
Haem329	F: CTTCATGCCATGTCCTGCCT R: TGCTCCTCTGATTGACTCCAG	$(TCTA)_{16}$	8	213-245	0.58	0.84	MN333908
Haem330	F: CAGGAATCCCTCTTTTCAGCTG R: GCCTATGCTGTGATAATTGCAC	$(TCCA)_{14}$	14	204-328	1.00	0.90	MN333909

[a]TGTAAAACGACGGCCAGT and GTGTCTT were added to the 5' end for forward and reverse primers, respectively
F: forward primer, R: reverse primer; N_A: number of alleles; H_O: observed heterozygosity; H_E: expected heterozygosity

Literature Cited

Acevedo-Whitehouse K, F Gulland, D Grei, and **W Amos.** 2003. Disease susceptibility in california sea lions. Nature 422:35-35.

Aldrich JW and **JS Weske.** 1978. Origin and evolution of the eastern house finch population. The Auk 95:528-536.

Badyaev AV, GE Hill, PO Dunn, and **JC Glen.** 2001. Plumage color as a composite trait: Developmental and functional integration of sexual ornamentation. The American Naturalist 158:221-235.

Bock CE and **LW Lepthein.** 1976. Growth in the eastern house finch population, 1962-1971. American Birds 30:791-792.

Briskie JV and **M Mackintosh.** 2004. Hatching failure increases with severity of population bottlenecks in birds. Proceedings of the National Academy of Sciences of the United States of America 101:558-561.

Brownstein MJ, JD Carpten, and **JR Smith.** 1996. Modulation of non-templated nucleotide addition by *taq* DNA polymerase: Primer modifications that facilitate genotyping. Biotechniques 20:1004-1010.

Daniels J, P Holmans, N Williams, D Turic, P Mcguffin, R Plomin, and **MJ Owen.** 1998. A simple method for analyzing microsatellite allele image patterns generated from DNA pools and its application to allelic association studies. The American Journal of Human Genetics 62:1189-1197.

Dhondt AA, DL Tessaglia, and **RL Slothower.** 1998. Epidemic mycoplasmal conjunctivitis in house finches from eastern north america. Journal of Wildlife Diseases 34:265-280.

Ginot F, I Bordelais, S Nguyen, and **G Gyapay.** 1996. Correction of some genotyping errors in automated fluorescent microsatellite analysis by enzymatic removal of one base overhangs. Nucleic Acids Research 24:540-541.

Glenn TC and **NA Schable.** 2005. Isolating microsatellite DNA loci. *In* Zimmer EA and EH Roalson, editors. Molecular evolution: Producing the biochemical data, part b. Elsevier Academic Press Inc (San Diego). p 202-222.

Hawley DM. 2005. Isolation and characterization of eight microsatellite loci from the house finch (carpodacus mexicanus). Molecular Ecology Notes 5:443-445.

Hawley DM, D Hanley, AA Dhondt, and **IJ Lovette.** 2006. Molecular evidence for a founder effect in invasive house finch (carpodacus mexicanus) populations experiencing an emergent disease epidemic. Molecular Ecology 15:263-275.

Hawley DM, KV Sydenstricker, GV Kollias, and **AA Dhondt.** 2005. Genetic diversity predicts pathogen resistance and cell-mediated immunocompetence in house finches. Biology Letters 1:326-329.

Hedrick PW, TJ Kim, and **KM Parker.** 2001. Parasite resistance and genetic variation in the endangered gila topminnow. Animal Conservation 4:103-109.

Hill GE. 1992. Proximate basis of variation in carotenoid pigmentation of male house finches. The Auk 109:1-12.

Hill GE. 1993. Geographic variation in the carotenoid plumage pigmentation of male house finches (*carpopacus mexicanus*). Biological Journal of the Linnean Society 49:63-86.

Hill GE. 2002. A red bird in a brown bag: The function and evolution of colorful plumage in the house finch. Oxford University Press (NY). 336 p.

Hill GE, CY Inouye, and **R Montgomerie.** 2002. Dietary carotenoids predict plumage coloration in wild house finches. Proceedings of the Royal Society B-Biological Sciences 269:1119-1124.

Hill GE, PM Nolan, and **AM Stoehr.** 1999. Pairing success relative to male plumage redness and pigment symmetry in the house finch: Temporal and geographic constancy. Behavioral Ecology 10:48-53.

Hoffman JI and **W Amos.** 2005. Microsatellite genotyping errors: Detection approaches, common sources and consequences for paternal exclusion. Molecular Ecology 14:599-612.

Inouye CY, GE Hill, RD Stradi, and **R Montgomerie.** 2001. Carotenoid pigments in male house finch plumage in relation to age, subspecies, and ornamental coloration. The Auk 118:900-915.

Maddox JD and **KA Feldheim.** 2014. A cost-effective size standard for fragment analysis that maximizes throughput on five dye set platforms. Conservation Genetics Resources 6:5-7.

Nater A, AM Kopps, and **M Krutzen.** 2009. New polymorphic tetranucleotide microsatellites improve scoring accuracy in the bottlenose dolphin tursiops aduncus. Molecular Ecology Resources 9:531-534.

Oh KP and **AV Badyaev.** 2006. Adaptive genetic complementarity in mate choice coexists with selection for elaborate sexual traits. Proceedings of the Royal Society B-Biological Sciences 273:1913-1919.

Oh KP and **AV Badyaev.** 2009. Isolation and characterization of 17 microsatellite loci for the house finch (carpodacus mexicanus). Molecular Ecology Resources 9:1029-1031.

Peakall R and **PE Smouse.** 2006. Genalex 6: Genetic analysis in excel. Population genetic software for teaching and research. Molecular Ecology Notes 6:288-295.

Peakall R and **PE Smouse.** 2012. Genalex 6.5: Genetic analysis in excel. Population genetic software for teaching and research-an update. Bioinformatics 28:2537-2539.

Raymond M and **F Rousset.** 1995. Genepop (v1.2): Population genetics software for exact tests and ecumenicism. Journal of Heredity 86:248-249.

Reed DH and **R Frankham.** 2003. Correlation between fitness and genetic diversity. Conservation Biology 17:230-237.

Rousset F. 2008. Genepop'007: A complete reimplementation of the genepop software for windows and linux. Molecular Ecology Resources 8:103-106.

Schuelke M. 2000. An economic method for the fluorescent labeling of pcr fragments. Nature Biotechnology 18:233-234.

Shultz AJ, AJ Baker, GE Hill, PM Nolan, and **SV Edwards.** 2016. Snps across time and space: Population genomic signatures of founder events and epizootics in the house finch (haemorhous mexicanus). Ecology and Evolution 6:7475-7489.