Conservation Genetics Resources (2022) 14:147–150
https://doi.org/10.1007/s12686-022-01255-6

TECHNICAL NOTE

Development of a 95 SNP panel to individually genotype mountain lions (*Puma concolor*) for microfluidic and other genotyping platforms

Michael R. Buchalski[1] · Benjamin N. Sacks[2] · Kristen D. Ahrens[1] · Kyle D. Gustafson[3] · Jaime L. Rudd[4] · Holly B. Ernest[5] · Justin A. Dellinger[4]

Received: 8 September 2021 / Accepted: 11 February 2022 / Published online: 23 February 2022

Abstract

The mountain lion (*Puma concolor*) is one of the few remaining large predators in California, USA with density estimation from fecal genotypes becoming an essential component of conservation and management. In highly urbanized southern California, mountain lions are fragmented into small, inbred populations making proper marker selection critical for individual identification. We developed a panel of single nucleotide polymorphism (SNP) markers that can be used for consistent, routine mountain lion monitoring by different laboratories. We used a subset of existing Illumina HiSeq data for 104 individuals from throughout California to design a single, highly heterozygous multiplex of 95 SNPs for the Fluidigm platform. This panel confidently differentiates individual mountain lions, identifies sex, and discriminates mountain lions from bobcats. The panel performed well on fecal DNA extracts and based on design, had sufficient resolution to differentiate individual genotypes in even the population with lowest genetic diversity in southern California.

Keywords California · Fecal DNA · Fluidigm · Mountain lion · *Puma concolor* · Single nucleotide polymorphism

The mountain lion (*Puma concolor*) is an elusive, common large predator native to California, USA. Monitoring of wild felids often requires non-invasive genetic tools (Ernest et al. 2000). In California, population monitoring has increasingly relied on canine surveys for feces followed by fecal genotyping and mark-recapture analyses to estimate population abundance. Although microsatellites have previously been relied upon for such noninvasive surveys, they have several weaknesses that must be overcome to prevent errors (Taberlet et al. 1999), including low amplification rates (Brinkman et al. 2010; Murphy et al. 2007; Piggott 2004), challenges calibrating between laboratories (von Thaden et al. 2017), allelic dropout, and false alleles (Creel et al. 2003; Pompanon et al. 2005).

There is great need for genetic methods that can be applied to low quality DNA samples while maintaining statistical power to differentiate individuals in low-diversity populations. In southern California, many mountain lion populations are impacted by urban development, and exist in small numbers with low genetic diversity (Ernest et al. 2014; Gustafson et al. 2017; Riley et al. 2014), making it difficult to apply genetic methods without high probability of monomorphic loci and identification errors. A previous SNP panel (PumaPlex100; Erwin et al. 2021) demonstrated the utility of SNPs to differentiate among mountain lion individuals. However, this panel was not optimized for low-diversity populations and requires up to four multiplexes. Therefore, we developed a single multiplex of 95 SNPs for the Fluidigm platform (Fluidigm, San Francisco USA) that can be used across laboratories for surveys of California and potentially other mountain lion populations. These SNPs

✉ Michael R. Buchalski
michael.buchalski@wildlife.ca.gov

1 Wildlife Genetics Research Unit, Wildlife Health Laboratory, California Department of Fish and Wildlife, Sacramento, CA, USA

2 Mammalian Ecology and Conservation Unit, Veterinary Genetics Laboratory, Department of Population Health and Reproduction, School of Veterinary Medicine, University of California, Davis, CA, USA

3 Department of Biological Sciences, Arkansas State University, Jonesboro, AR, USA

4 Wildlife Health Laboratory, California Department of Fish and Wildlife, Sacramento, CA, USA

5 Wildlife Genomics and Disease Ecology Laboratory, Department of Veterinary Sciences, University of Wyoming, Laramie, WY, USA

can identify individuals for abundance monitoring, assign sex, and discriminate mountain lion feces from bobcat (*Lynx rufus*), the most common non-target species detected during canine surveys.

We identified SNPs using existing RADseq data (Illumina HiSeq, paired-end 150 bp reads) from mountain lion tissues sampled as part of a California statewide population genetic study (Gustafson et al. 2021; raw data available at https://doi.org/10.17605/OSF.IO/HUF4K) and mapped reads to the *Puma concolor* scaffold-level genome assembly (PumCon1.0), using 104 individuals, including up to 15 individuals from each of 10 populations identified by Gustafson et al. (2019). We used VCFTOOLS to filter biallelic autosomal SNPs to a PHRED score ≥ 30 (99.9% accuracy) with ≥ 80% of individuals genotyped at each SNP. To maximize heterozygosity and resolution to differentiate individuals, we used PLINK v1.90b6.10 to filter the minor allele frequency to 0.40–0.45. We retained 134 SNPs distributed across 28 scaffolds (mean within-scaffold distance of 11.3 Mbp) for further validation (Supporting File 1). We identified species-diagnostic SNPs using published mitogenome sequences for mountain lion and bobcat (Supporting information Table S1). We also identified sex-linked SNPs from both the SRY locus on the non-recombining region of the Y chromosome and the zinc finger orthologs of the X and Y chromosomes (Supporting information, Table S2).

We used PYSAM v0.15.0 to retrieve 100 base pairs of up- and down-stream flanking sequence for the 134 SNPs. We used Fluidigm's D3 Assay Design Tool (https://d3.fluidigm.com) in conjunction with the flanking sequences (Supporting Information, Table S3) to design primers, which we ordered from Fluidigm (Supporting information, Table S4).

To test the performance of the SNP panel on fecal DNA, we used 19 pairs of tissue and fecal samples collected during mountain lion necropsies conducted by the California Department of Fish and Wildlife from 14 different counties. Feces were experimentally weathered outdoors in direct sunlight for seven days. Tissue and fecal DNA were extracted using the Genomic Tip 20G kit and the QIAamp Fast DNA Stool Mini Kit (Qiagen, Valencia, CA), respectively. To assess the performance of the SNP panel on low-diversity populations, we extracted tissues for 26 mountain lions from the highly inbred Santa Ana Mountains population in southern California. We also extracted 16 bobcat tissues to test the species-diagnostic SNPs.

We used Fluidigm's protocol for 96.96 Dynamic Arrays with integrated fluidic circuits run on the Juno thermocycler paired with the Biomark HD system. We modified the pre-amplification PCR reaction to align with optimizations for degraded samples (see Table S2 in von Thaden et al. 2020). We also increased the manufacturer's PCR cycling profile from 34 to 45 in the final allele specific reaction. We analyzed genotyping results using Fluidigm SNP Genotyping Analysis Software version 4.5.1. We selected a final panel of 91 autosomal SNPs based on consistent cluster separation and high call rates. We also selected two species-diagnostic and two sex-linked SNPs. For the Santa Ana population, we tested for linkage disequilibrium LD using the R package POPPR (Kamvar et al. 2014) to calculate the index of association among pairs of loci and corrected for a false detection rate FDR of 5%. We estimated expected heterozygosity H_E using the ADEGENET package (Jombart and Ahmed 2011) and calculated the probability of identity P_{ID} and probability of identity in siblings P_{IDsibs} using the POPGENUTILS package (Tourvas 2021). All packages were implemented in R 4.1.1 (R Core Team 2021).

When genotyping the 19 paired tissue and fecal samples at the 91 autosomal SNPs, we obtained > 95% call rates on 18 of the 19 fecal samples. The overall agreement between the paired samples was 99.9% (SD = 0.6%), indicating a genotyping error rate < 0.1%. For the low-diversity Santa Ana population, no loci were in LD following FDR correction. The average H_E for the 91 autosomal SNPs was 0.404, an estimate higher than previously measured using 42 microsatellites ($H_E = 0.33$; Gustafson et al. 2019), suggesting our SNP filtering protocol successfully inflated heterozygosity and increased power for individual identification. In fact, values for P_{ID} and P_{IDsibs} across all SNPs were $< 1 \times 10^{-10}$ (Supporting information, Table S5) and individual identifications could be made with high confidence from as few as 15 SNPs (Fig. 1). For 33 samples where both sex-linked SNPs (2FelidSRYSNP-GT, mlbcZfy-680) yielded genotypes, 30 pairs (91%) agreed with known sex, suggesting < 5% sex-typing error rate. The species-diagnostic SNPs (mtdna_658, mtdna_2089) were 100% successful at differentiating mountain lions from bobcats. Further, bobcat reference genotypes were monomorphic at all autosomal SNPs.

This high-resolution panel of 95 SNPs was optimized for the Fluidigm system but can be genotyped in any SNP-typing platform. The panel was designed from mountain lions throughout California and may be subject to ascertainment bias when applied to other geographic regions. As a result, the utility of this panel for individual identification throughout the species range requires further verification. The high minor allele frequency screening step for SNP selection proved useful for differentiating individuals but also inflates estimates of heterozygosity. Therefore, certain analyses, such as estimating and comparing heterozygosity or identifying population structure and gene flow would not be appropriate uses of these markers.

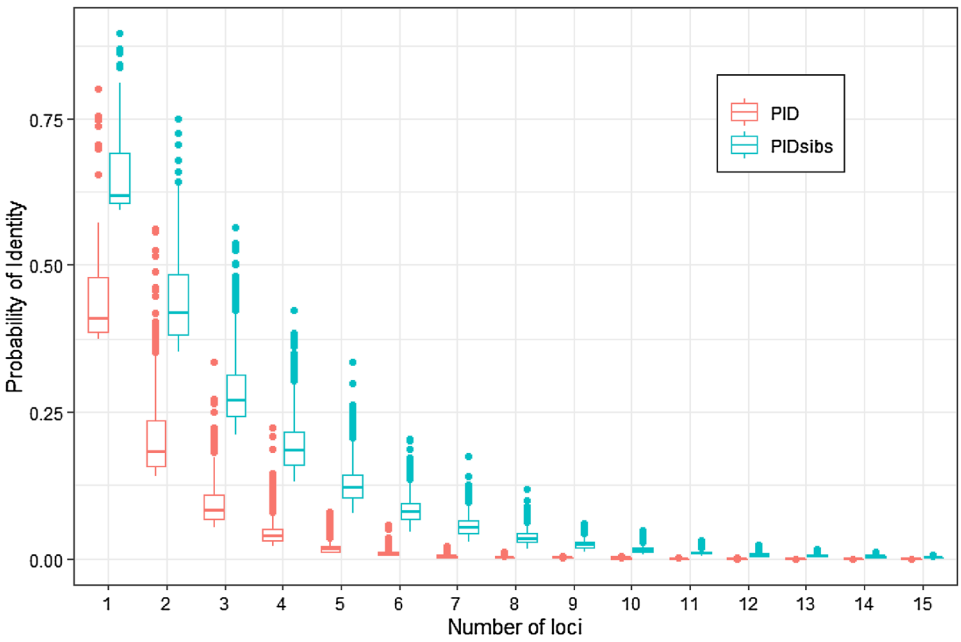

Fig. 1 Decrease in P_{ID} and P_{IDsibs} with increasing number of loci for 26 mountain lions from the low-diversity Santa Ana Mountains population in California, USA. Figure generated using the *pid_permute* function for the R package POPGENUTILS (Tourvas 2021) and based on 1000 permutations across different combinations of loci. Only 1 to 15 of 91 total autosomal loci are shown for clarity

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12686-022-01255-6.

Acknowledgements S. Vanderzwan and A. Kubicki were instrumental to completing the laboratory work.

Author contributions JAD and JLR collected the samples; MRB designed the study; MRB, KDG, HBE, and BNS discovered the SNPs; BNS and KDA tested and selected the final panel; all authors contributed to writing.

Funding Funding was from the U.S. Fish and Wildlife Service, Wildlife Restoration Act (Grant No. P1580009) and the California Department of Fish and Wildlife, Big Game Management Program.

Data availability Original SNP data from Gustafson et al. (2021) can be found at https://doi.org/10.17605/OSF.IO/HUF4K. All other data have been provided as online supplements.

Declarations

Conflict of interest Not applicable.

References

Brinkman TJ, Schwartz MK, Person DK, Pilgrim KL, Hundertmark KJ (2010) Effects of time and rainfall on PCR success using DNA extracted from deer fecal pellets. Conserv Genet 11(4):1547–1552

Creel S, Spong G, Sands JL, Rotella J, Zeigle J, Joe L et al (2003) Population size estimation in Yellowstone wolves with error-prone noninvasive microsatellite genotypes. Mol Ecol 12(7):2003–2009

Ernest HB, Penedo MC, May BP, Syvanen M, Boyce WM (2000) Molecular tracking of mountain lions in the Yosemite valley region in California: genetic analysis using microsatellites and faecal DNA. Mol Ecol 9(4):433–441

Ernest HB, Vickers TW, Morrison SA, Buchalski MR, Boyce WM (2014) Fractured genetic connectivity threatens a southern California puma (*Puma concolor*) population. PLoS ONE 9:e107985

Erwin JA, Fitak RR, Culver M (2021) PumaPlex100: an expanded tool for puma SNP genotyping with low-yield DNA. Conserv Genet Resour 13(3):341–343

Gustafson KD, Vickers TW, Boyce WM, Ernest HB (2017) A single migrant enhances the genetic diversity of an inbred puma population. R Soc Open Sci 4(5):170115

Gustafson KD, Gagne RB, Vickers TW, Riley SPD, Wilmers CC, Bleich VC et al (2019) Genetic source–sink dynamics among naturally structured and anthropogenically fragmented puma populations. Conserv Genet 20(2):215–227

Gustafson KD, Gagne RB, Buchalski MR, Vickers TW, Riley SPD, Sikich JA et al (2021) Multi-population puma connectivity could restore genomic diversity to at-risk coastal populations in California. Evol Appl. https://doi.org/10.1111/eva.13341

Jombart T, Ahmed I (2011) adegenet 1.3–1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27(21):3070–3071

Kamvar ZN, Tabima JF, Grünwald NJ (2014) Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2:e281

Murphy MA, Kendall KC, Robinson A, Waits LP (2007) The impact of time and field conditions on brown bear (*Ursus arctos*) faecal DNA amplification. Conserv Genet 8(5):1219–1224

Piggott M (2004) Effect of sample age and season of collection on the reliability of microsatellite genotyping of faecal DNA. Wildl Res 31(5):485–493

Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. Nat Rev Genet 6(11):847–859

R Core Team (2021) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Riley SP, Serieys LE, Pollinger JP, Sikich JA, Dalbeck L, Wayne RK et al (2014) Individual behaviors dominate the dynamics of an urban mountain lion population isolated by roads. Curr Biol 24(17):1989–1994

Taberlet P, Waits LP, Luikart G (1999) Noninvasive genetic sampling: look before you leap. Trends Ecol Evol 14(8):323–327

Tourvas N (2021) PopGenUtils: a collection of useful functions to deal with genetic data in R. R package version 0.1.6

von Thaden A, Cocchiararo B, Jarausch A, Jüngling H, Karamanlidis A, Tiesmeyer A et al (2017) Assessing SNP genotyping of noninvasively collected wildlife samples using microfluidic arrays. Sci Rep 7:10768

von Thaden A, Nowak C, Tiesmeyer A, Reiners TE, Alves PC, Lyons LA et al (2020) Applying genomic data in wildlife monitoring: development guidelines for genotyping degraded samples with reduced single nucleotide polymorphism panels. Mol Ecol Resour 20(3):662–680