Development of microsatellite loci exhibiting reverse ascertainment bias and a sexing marker for use in Emperor Geese (Chen canagica)

Meg C. Gravley*, George K. Sage, Joel A. Schmutz and Sandra L. Talbot

U.S. Geological Survey, Alaska Science Center, 4210 University Drive, Anchorage, Alaska 99508
E-mail: mgravley@usgs.gov

ABSTRACT

The Alaskan population of Emperor Geese (Chen canagica) nests on the Yukon-Kuskokwim Delta in western Alaska. Numbers of Emperor Geese in Alaska declined from the 1960s to the mid-1980s and since then, their numbers have slowly increased. Low statistical power of microsatellite loci developed in other waterfowl species and used in previous studies of Emperor Geese are unable to confidentially assign individual identity. Microsatellite loci for Emperor Goose were therefore developed using shotgun amplification and next-generation sequencing technology. Forty-one microsatellite loci were screened and 14 were found to be polymorphic in Emperor Geese. Only six markers — a combination of four novel loci and two loci developed in other waterfowl species — are needed to identify an individual from among the Alaskan Emperor Goose population. Genetic markers for identifying sex in Emperor Geese were also developed. The 14 novel variable loci and 15 monomorphic loci were screened for polymorphism in four other Arctic-nesting goose species, Black Brant (Branta bernicla nigricans), Greater White-fronted (Anser albifrons), Canada (B. canadensis) and Cackling (B. hutchinsii) Goose. Emperor Goose exhibited the smallest average number of alleles (3.3) and the lowest expected heterozygosity (0.467). Greater White-fronted Geese exhibited the highest average number of alleles (4.7) and Cackling Geese the highest expected heterozygosity (0.599). Six of the monomorphic loci were variable and able to be characterised in the other goose species assayed, a predicted outcome of reverse ascertainment bias. These findings fail to support the hypothesis of ascertainment bias due to selection of microsatellite markers.

Keywords: Emperor Goose, Chen canagica, microsatellite markers, molecular sexing, reverse ascertainment bias

1. INTRODUCTION

The Emperor Goose (Chen canagica) is a stocky, medium-sized goose endemic to the Bering Sea region. The vast majority of Emperor Geese nest in the Yukon-Kuskokwim Delta (YKD) in southwest Alaska, with small numbers nesting in Alaska’s Seward Peninsula and in the Anadyr region of eastern Russia. There are no morphological or plumage differences between male and female Emperor Geese; both sexes have a dark grey body and distinctive white head and neck which, in spring and summer, are stained orange from feeding in wetland habitats where ferrous soils are abundant (Schmutz et al., 2011). When occupying marine and estuarine habitats, Emperor Geese feed on benthic invertebrates (mussels and clams), intertidal and dune vegetation, and algae (Schmutz, 1994). Emperor Geese principally migrate in winter to the Aleutian Island Archipelago (including the Commander Islands of Russia), but also winter in the Kodiak Island Archipelago and the southern coasts of the Alaska Peninsula (Eisenhauer and Kirkpatrick, 1977; Schmutz et al., 2011). The eastern coasts of the Chukotka Peninsula and the Anadyr lowlands in Russia are also home to a small nesting population of Emperor Geese that likely winter in Alaska (Schmutz et al., 2011).

Like other Arctic-nesting geese, the Emperor Goose population declined from 1964 to 1986, dropping in number from 139,000 to 42,000 birds (Schmutz et al., 2011). The IUCN Red List status of the Emperor Goose is ‘near threatened’ with a decreasing population trend (BirdLife International, 2012). However, recent U.S. Fish and Wildlife Service surveys suggest the population has been slowly increasing (U.S. Fish and Wildlife Service, 2016). The species’ slow population recovery, relative to most other Arctic-nesting geese, is likely because Emperor Geese only eat native flora and fauna of the Beringia region, whereas most other goose species have access to agricultural crops on their wintering grounds in lower latitudes of North America (Alisauskas, 2002; Abraham et al., 2005).

Many waterfowl species, including the Emperor Goose, engage in a reproductive strategy called egg-dumping or intraspecific nest parasitism (INP), in which females of the same species lay eggs in another female’s nest (McCloskey, 2013). It is important to identify the parentage of eggs in nests because nest parasitism affects fitness of both the parasite and the host (Waldecker et al., 2011). Nest parasitism can be detected using genetic markers, such as microsatellites, that are sufficiently
variable to distinguish relatives from among non-relatives. However, there are currently no microsatellite markers specifically developed for Emperor Geese, and sexing markers routinely used on other bird species (Griffiths et al., 1998) do not distinguish between sexes in Emperor Geese. Using microsatellite loci developed for other species (see Table 1; Mallard [Anas platyrhynchos]: Aph02; Canada Goose [Branta canadensis]: Bcau4, Bcau5, Bcau11; Ruddy Duck [Oxyura jamaicensis]: Oxy13; and Common Eider [Somateria mollissima]: Sm01, Sm010, Sm011, Sm013), Watkins (2006) found that over 80% of Emperor Goose nests on the YKD had apparently been parasitised by other Emperor Geese, and that on average, only half the eggs in those nests had been laid by the host goose. These rates of INP in Emperor Geese were higher than previous estimates (Petersen, 1991) and higher than any other waterfowl species known to date (Watkins, 2006). However, the microsatellite markers in that study (Table 1) had low resolving power for use in individual identification within the Alaskan population of Emperor Geese, and only marginally better resolution for detecting egg dumping. For instance, based on the population size estimated for Emperor Geese, the probability of identity (PID) values reported in Watkins (2006) were not sufficiently small to delineate between some individuals. In addition, some of the eggs in the nests could not be eliminated as parasitic or determined to have originated from the host birds. Thus, Watkins (2006) hypothesised that levels of INP may be higher than her study suggested. To better estimate levels of INP, family relationships, and population genetics characteristics among Emperor Geese for use in future studies, it was necessary to develop markers with more resolution, in addition to sexing markers.

Because male and female Emperor Geese are not sexually dimorphic, field sexing of adults is difficult and reliant upon the behaviour of the geese. Females typically sit on the nest, while the male stands guard nearby (Schmutz et al., 2011). Nestlings are even more difficult to sex and often require handling to inspect their reproductive tracks (Griffiths et al., 1998). Genetic sexing markers can either confirm or correct field sexing of birds and do not necessarily require a biologist to handle nestlings or monitor nesting birds for behavioural cues, since it allows an individual bird to be sexed using non-invasively collected DNA sources, such as feathers and eggshell membranes. Certain types of parentage analyses and programs use exclusion techniques requiring that known sexes be associated with parental genotypes to delineate between potential mothers and fathers (Jones and Ardren, 2003). For example, the kinship program CERVUS uses likelihood methods that can account for potential null alleles, allelic dropout, and other genotyping errors that can cause false exclusions of parents (Kalinowski et al., 2007) which would be problematic in studies involving INP.

In this study, 14 novel polymorphic microsatellite loci and a sexing marker were developed to assess levels of genetic diversity in Alaska’s Emperor Goose population and assist in genetic identification of individuals for future studies of INP. A cross-species amplification of the 14 loci found to be polymorphic in Emperor Geese was also performed on four other co-distributed Arctic-nesting goose species: Greater White-fronted Goose (Anser albifrons); Canada Goose (B. canadensis); Cackling Goose (B. hutchinsii); and Black Brant (B. bernicla nigricans). As well, 15 loci monomorphic in the focal species were also amplified in the cross-species screen to examine findings of reverse ascertainment bias in the loci.

2. METHODS

2.1 Next-generation sequencing

Genomic DNA was extracted from muscle tissue of a male Emperor Goose (CPD1611; collected from Chagvan Bay, Alaska in 1991) and subjected to next generation sequencing.
Microsatellite and sexing markers in Emperor Geese

2.2 Microsatellite development

A modified salt extraction technique (Medrano et al., 1990; Sonsthagen et al., 2004) was used to extract genomic DNA from blood and heart samples collected from 30 Emperor Goose carcasses donated to U.S. Fish and Wildlife from both the YKD nesting grounds and wintering locales in Alaska between 1991 to 2014 (collection and microsatellite data are available; Talbot, 2017). Forty-one loci with dinucleotide repeats were initially screened in 11 Emperor Geese, of which 14 useable loci were found to be polymorphic. These loci were thereafter amplified in each of the 30 individuals, which included CPD1611 (Tables 2 and 3). The polymorphic and monomorphic loci were also screened in four other goose species from Alaskan populations: Greater White-fronted Goose (n = 10) from Selawik; Canada Goose (n = 10) from Copper River Delta; Cackling Goose (n = 9) from the YKD; and Black Brant (n = 11) from Kigigak Island.

Universal-tailed polymerase chain reaction (PCR) chemistry (Oetting et al., 1995) and thermocycler parameters followed Handel et al. (2006), except the PCR reaction contained 40 cycles at 94 °C for 15 s, 50–54 °C for 15 s and 72 °C for 30 s. To obtain size standards for each locus, two to six heterozygous individuals were selected for each locus and scored against a fluorescently labelled M13 sequence ladder of known size. These were used in each subsequent gel as size standards, occupying four to six lanes across each gel. Based on these standards, genotypes for each individual were determined using Gene Profiler 4.05 (Scanalytics, Inc.). Quality control (QC) procedures for microsatellite loci included: (1) performing eight separate extractions of DNA from paired heart and blood tissues from each individual to verify DNA extracted from both tissue types amplified the same genotypes; and (2) duplicating PCR reactions for 10 individuals across loci. This provided an overall QC replication of 60%. All lab work (including for molecular sexing, section 2.3) was performed using sterile laboratory techniques and procedures.

Observed (H_s) and expected (H_E) heterozygosity and polymorphism information content (PIC) were calculated using the Microsatellite Toolkit (Park, 2000). Linkage disequilibrium (LD) and deviations from Hardy–Weinberg equilibrium (HWE) were tested for using GenePop’007 (Raymond and Rousset, 1995; Rousset, 2008). PID assuming a randomly breeding population, and PID among first order relatives (PID_1) were identified using Gimlet, v. 1.3.2 (Valière, 2002). Parental exclusion values (Q) were calculated using PowerStats, v. 1.2 (Tereba, 1999). MICRO-CHECKER, v. 2.2.3 (Van Oosterhout et al., 2004) was used to identify potential scoring errors and null alleles. Given the recent population decline of Emperor Geese in Alaska up until the 1980s, population demography was evaluated for the microsatellite loci using BOTTLENECK 1.2.02 (Cornuet and Luikart, 1996; Maruyama and Fuerst, 1985), stepwise mutation model (SMM [Ohta and Kimura, 1973]), and two-phase model of mutation (TPM [Di Rienzo et al., 1994]). The TPM parameters included a 9% variance and 80% proportion of SMM (Piry et al., 1999; Garza and Williamson, 2001) with 1,000 iterations performed. A Wilcoxon sign-rank test, which determines if the average of standardised differences between H_s and H_E is significantly different from zero, was used to assess significance (Cornuet and Luikart, 1996). Recent population growth is indicated by significant heterozygote deficiency values relative to the number of alleles, whereas a recent population bottleneck is indicated by heterozygote excess relative to the number of alleles (Cornuet and Luikart, 1996). Since BOTTLENECK only compares heterozygote deficiency and excess relative to genetic diversity, and not to HWE expectation (Cornuet and Luikart, 1996) only loci in HWE were included in the analysis.

2.3 Molecular sexing

Emperor Geese of known sex (via necropsy) were screened using the P2 and P8 primers that access the chromo-helicase dehydrogenase (CHD) gene, the molecular marker most commonly employed to determine sex in birds (Griffiths et al., 1998). The CHDZ and CHDW bands did not appear to differ by size between male and female Emperor Geese (data not shown), precluding their use as molecular sexing markers in the species. Prior research found differences in Black Oystercatcher (Haematopus bachmani; Guzzetti et al., 2008) between gametologs of the histidine triad nucleotide binding protein gene family (HINTZW, formerly Wpcki) [Hori et al., 2000] and ASW [O’Neill et al., 2000]; see Ceplitis and Ellegren [2004] for discussion of nomenclature). The HINTW gene is located on the W-chromosome and its gametolog, the highly conserved HINTZ gene, is located on the Z-chromosome (Hori et al., 2000; Ceplitis and Ellegren, 2004). An Emperor Goose sexing marker was therefore developed to target the HINTZ gametologs. Procedures were followed as reported in Guzzetti et al. (2008) to take advantage of differences between the two HINT gametologs in Emperor Geese. Primers ASW12-D3 and ASW12-R2, tailed with 19 base pairs (bp) M13F and 20 bp M13R fluorescent labels, respectively
Meg C. Gravley, George K. Sage, Joel A. Schmutz and Sandra L. Talbot

(see Guzzetti et al., 2008) were used to sequence HINTZ in four male (781 bp) and HINTW in four female (778 bp) Emperor Geese. A new reverse primer (HZW278: 5’-ACACAMGTGYWTTTTGCACA-T-3’), synthesised with the 20 bp M13R primer on the 5’ end, was subsequently designed to amplify and sequence, along with ASW13-D3, a 253 bp product for HINTZ and a 243 bp product for HINTW (primer sequences not included), which revealed conserved regions of nucleotides in both gametologs that flanked a region containing insertions/deletions between them (Figure 1A, sequences archived in GenBank, accession numbers: KY274281-5 [HINTZ], KY274277-80 [HINTW]). Products from subsequent universal-tailed PCR reactions employing these two primers and incorporating labelled M13F primer (similar to the microsatellite PCR reactions, section 2.2) run on a 6% acrylamide gel, yielded a 297 bp product from the Z-chromosome (males and females) and a 287 bp product from the W-chromosome (females only). This size includes genomic DNA and the primer sequences, but not the universal primer sequences (M13F and M13R). Sex assignment is based on the absence (male: ZZ) or presence (female: ZW) of the band for HINTW (Figure 1B). Sexing was performed on DNA from blood samples for all Emperor Geese. For QC replication, PCR amplifications were duplicated using DNA from six individuals, for an overall QC replication of 23% (n = 7).

3. RESULTS

Microsatellite loci developed in other waterfowl species and used in prior analyses of Emperor Geese (Watkins, 2006) were characterised (Table 1). No genetic diversity metrics were shown for the Z-specific loci (Bca4, Smo01). None of the autosomal loci deviated from HWE in Emperor Goose, but LD was found (α < 0.05) between Bca11 and Smo11 (P = 0.013) and Smo11 and Smo13 (P = 0.048). All genetic sexes were found to be identical to the sex determined during necropsy. In Emperor Geese, the W band amplified preferentially in females and only the Z band amplified in males (Figure 1B).

The Emperor Goose NG library generated 58,275 reads. Microsatellite motifs were found in 643 reads (538 dinucleotides) and 41 loci with dinucleotide microsatellite repeat motifs were screened. Of these 41 loci, 15 were polymorphic among Emperor Geese, 15 were monomorphic, five amplified extraneous products, and six failed to amplify. Of the polymorphic loci, one locus was found to have two alleles, but co-amplified with other products and was unable to be optimised by redesigning the primer set, leaving a total of 14 polymorphic loci (Table 2). One locus (Cca02) contained a juxtaposed microsatellite repeat and primers were developed to amplify both repeats (Cca02a and Cca02b) independently.

For Emperor Geese used in this study, two to six alleles, ranging in size from 93 to 236 bp, were amplified per

---

**Figure 1** Size differences between the HINTW and HINTZ gametologs in Emperor Goose (*Chen canagica*). (A) Sequence data with forward primer from Guzzetti et al. (2008) and redesigned reverse (HZW278), underlined. Grey numbers indicate sequence position. (B) Gel image of sexing PCR reaction. Position 297 is the Z band and position 287 is the W band.
Table 2 Characterisation of microsatellite markers developed for Emperor Goose (Chen canagica) with locus name, repeat motif, primer sequences (universal tail in parentheses), and GenBank accession number

<table>
<thead>
<tr>
<th>Locus</th>
<th>Repeat Motif</th>
<th>Primer Sequence (5’→3’)a</th>
<th>GenBank No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cca02ab</td>
<td>(AC)10</td>
<td>F: (M13R)TGTCATGCAGGAGACATGC R: CATGGCCCTGTGGTGGAGC</td>
<td>KY274288</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F: (SP6)TGTCGCGCACATGCACCTTGC R: CATGGCCCTGTGGTGGAGC</td>
<td>KY274288</td>
</tr>
<tr>
<td>Cca04</td>
<td>(CA)n(Ni)(CA)n11</td>
<td>F: (M13R)TGTCATTTAATTCGAGGGG R: TTTCAGAAGACACCTTGGAGC</td>
<td>KY274289</td>
</tr>
<tr>
<td>Cca05</td>
<td>(AC)n0</td>
<td>F: (SP6)CCTCCTTTTCTGTGAAGGC R: CTCAAGATTTCTTTAACAGGC</td>
<td>KY385249</td>
</tr>
<tr>
<td>Cca06b</td>
<td>(AC)n11</td>
<td>F: (M13F)CTTCCAAACACCTGACGCG GC: AACAAGGCCGTTTACATAG</td>
<td>KY385250</td>
</tr>
<tr>
<td>Cca10</td>
<td>(AC)n13</td>
<td>F: (M13F)TTAATGACACACCTGACG GC: AGTGGACGCTGTCTTGAGC</td>
<td>KY274290</td>
</tr>
<tr>
<td>Cca18b</td>
<td>(AC)n0</td>
<td>F: (M13F)CGTCCATTTAATCTAACAGGC GC: GCAGAAGACCTGACG CG: AAGATGTGACGTGTGGAGC</td>
<td>KY385258</td>
</tr>
<tr>
<td>Cca20</td>
<td>(CA)n10</td>
<td>F: (M13F)CTTCCAACCTCTCTAGTCGC GC: GCAGAAGACCTGACG CG: AAGATGTGACGTGTGGAGC</td>
<td>KY274291</td>
</tr>
<tr>
<td>Cca21</td>
<td>(CA)n0</td>
<td>F: (M13F)CTTCCACCTCTCTAGTCGC GC: GCAGAAGACCTGACG CG: AAGATGTGACGTGTGGAGC</td>
<td>KY274292</td>
</tr>
<tr>
<td>Cca22</td>
<td>(CA)n0</td>
<td>F: (M13F)CTTCCAACCTCTCTAGTCGC GC: GCAGAAGACCTGACG CG: AAGATGTGACGTGTGGAGC</td>
<td>KY274293</td>
</tr>
<tr>
<td>Cca23</td>
<td>(CA)n0</td>
<td>F: (M13F)CTTCCAACCTCTCTAGTCGC GC: GCAGAAGACCTGACG CG: AAGATGTGACGTGTGGAGC</td>
<td>KY274294</td>
</tr>
<tr>
<td>Cca24</td>
<td>(GA)n11</td>
<td>F: (M13F)CTTCCAACCTCTCTAGTCGC GC: GCAGAAGACCTGACG CG: AAGATGTGACGTGTGGAGC</td>
<td>KY274295</td>
</tr>
<tr>
<td>Cca25c</td>
<td>(AC)n9</td>
<td>F: (SP6)GACGGACCACTCCTCAGGC GC: GCTGAAGGAGGTGATAAGGC</td>
<td>KY385259</td>
</tr>
<tr>
<td>Cca26</td>
<td>(GA)n11</td>
<td>F: (M13F)CATCTTCACTTCTAGTCGC GC: GCAGAAGACCTGACG CG: AAGATGTGACGTGTGGAGC</td>
<td>KY274296</td>
</tr>
<tr>
<td>Cca29</td>
<td>(GA)n11</td>
<td>F: (M13F)CTTCCAACCTCTCTAGTCGC GC: GCAGAAGACCTGACG CG: AAGATGTGACGTGTGGAGC</td>
<td>KY274297</td>
</tr>
<tr>
<td>Cca31c</td>
<td>(GA)n9</td>
<td>F: (SP6)GACGGACCACTCCTCAGGC GC: GCTGAAGGAGGTGATAAGGC</td>
<td>MF093617</td>
</tr>
<tr>
<td>Cca33c</td>
<td>(TA)n10</td>
<td>F: (M13F)ATGTCATGGAAATCTCTAAAG GC: GGCTCAACAGGTTACATAG</td>
<td>KY385262</td>
</tr>
<tr>
<td>Cca35</td>
<td>(TA)n10</td>
<td>F: (SP6)CTTTTTACTAAGGAAAGGC GC: GCTGAAGGAGGTGATAAGGC</td>
<td>KY274298</td>
</tr>
<tr>
<td>Cca39</td>
<td>(AT)n13</td>
<td>F: (M13F)TATCCACCTGCAACAGGC GC: GGCTCAACAGGTTACATAG</td>
<td>KY274299</td>
</tr>
<tr>
<td>Cca40</td>
<td>(AT)n13</td>
<td>F: (M13F)TGTTCCTTCTCCTTCAGG GC: GCCATTACATAGCTCAGGC</td>
<td>KY274300</td>
</tr>
<tr>
<td>Cca42</td>
<td>(CT)n12</td>
<td>F: (SP6)GACGGACCACTCCTCAGGC GC: GCTGAAGGAGGTGATAAGGC</td>
<td>KY274301</td>
</tr>
</tbody>
</table>

aUniversal tail primer sequences: M13R (GGATAACAATTTCACACAGG); M13F (CAGCAGCGTGAAACGAC); and SP6 (GATTTAGGTGTGACACTATAG).
bJuxtaposed microsatellite sequence.
cLoci found to be monomorphic in Emperor Goose, but polymorphic in other goose species assayed during the cross-species screen.

Microsatellite and sexing markers in Emperor Geese

5

novel locus; \( H_o \) ranged between 0.067 and 0.700, and \( H_i \) ranged from 0.066 to 0.692. Two loci, Cca02a \((P = 0.032)\) and Cca21 \((P = 0.043)\), deviated significantly from HWE (Table 3) and Cca02a was found to have heterozygote deficiency \((P = 0.026)\). Significant LD was observed in six cases: Cca02a and Cca04 \((P = 0.003)\); Cca35 and Cca42 \((P = 0.008)\); Cca02b and Cca23 \((P = 0.023)\); Cca02b and Cca22 \((P = 0.030)\); Cca10 and Cca20 \((P = 0.041)\); and Cca35 and Cca40 \((P = 0.042)\). Significant LD was also observed between Cca10 and Oxy13 \((P = 0.020)\) and Cca40 and Smo11 \((P = 0.043)\). The separately amplified repeats (Cca02a and Cca02b) comprising the juxtaposed locus were not found to be in LD \((P = 0.125)\), despite their close physical proximity of 114 bp. MICRO-CHECKER determined that Cca02a and Cca26 had potential null alleles at a 95% confidence interval. The likelihood
of allelic dropout for Cca02a provides a hypothesis for the lack of LD with Cca02b. Of the juxtaposed microsatellites, Cca02b appears less likely to have allelic dropout based on the MICRO-CHECKER results. The primers for Cca02b produce a smaller fragment (99–109 bp), increasing its relative value for use with non-invasive samples, such as feathers, since DNA extracted from feathers is often degraded and allelic dropout occurs more often in relatively larger loci (Sefc et al., 2003). No signature of recent bottleneck was uncovered in Emperor Geese sampled from Alaska, following the application of Bonferroni correction ($P_{\text{Bonferroni}} = 0.005$, $P_{\text{SAMH}} = 0.996$; $P_{\text{TPMocc}} = 0.226$, $P_{\text{TPMdef}} = 0.785$; $P_{\text{SIMocc}} = 0.464$, $P_{\text{SIMdef}} = 0.551$).

Q ranged from 0.004 to 0.428 across individual loci (Table 3). A multilocus marker suite with an overall combined probability of $Q \geq 0.99$ is theoretically able to exclude erroneous parent pairs from offspring (Villanueva et al., 2002). Thus, the higher the $Q$ value, the more useful a marker is for excluding incorrect parent-pairs from an offspring. Across novel individual loci, PIC values ranged from 0.062 to 0.627 (Table 3). Higher PIC values are associated with higher levels of polymorphism (Shete et al., 2000). For the Alaska population of Emperor Geese with around 98,100 birds – the largest number of geese counted in the past 10 years during a USFWS spring aerial survey (U.S. Fish and Wildlife Service, 2015) – a minimum of six loci (Aph02, Oxy13, Cca40, Cca39, Cca24, Cca04 [PID = 5.09e-6 and $P_{\text{Bonferroni}} = 0.048$] and Cca29 [PID = 9.24e-3 and $P_{\text{Bonferroni}} = 0.005$]) are recommended to perform individual identification using PID (Waits et al., 2001).

All the novel polymorphic loci, except Cca24, amplified in the cross-species screen (Table 4). In addition, among the 15 loci monomorphic in Emperor Geese, eight loci exhibited polymorphism in the cross-species species. However, because two of these loci require optimisation, only six were able to be characterised (Table 4) and included in analyses. Among these, locus Cca31 also needs optimisation, but was scorable for all samples, except one Cackling Goose. In total, 14 loci were polymorphic and amplified two to eight alleles in both Black Brant and Canada Geese. In Cackling Geese, 16 were polymorphic and amplified two to seven alleles. Greater White-fronted Goose samples were the most polymorphic, amplifying across 18 loci with 2 to 12 alleles per locus (see Table 4). Six loci deviated significantly from HWE: Cca05 ($P = 0.048$) in Black Brant; Cca06 ($P = 0.014$) and Cca39 ($P = 0.043$) in Canada Goose; Cca05 ($P = 0.004$) in Cackling Geese; Cca05 ($P = 0.046$) and Cca29 ($P = 0.027$) in Greater White-fronted Geese. Significant LD was observed between Cca21 and Cca26 ($P = 0.004$) in Black Brant; Cca04 and Cca40 ($P = 0.021$), Cca22 and Cca26 ($P = 0.040$) in Canada Goose; and Cca20 and Cca39 ($P = 0.022$) in Greater White-fronted Geese. MICRO-CHECKER uncovered potential null alleles at locus Cca05 in Cackling Geese, potentially due to homozygote excess.

4. DISCUSSION

This study provides a suite of novel microsatellite markers that can be leveraged with markers developed in other waterfowl species for a range of population genetic studies in the Emperor Goose, including studies involving parentage, INP, and individual identity. A sexing marker specific for Emperor Geese provides a molecular means to determine the sex of individuals, which is not possible using the molecular sexing marker most commonly employed to sex most birds (Griffiths et al., 1998). Cross-

<table>
<thead>
<tr>
<th>Locus</th>
<th>Sizea</th>
<th>$H_0$</th>
<th>$H_1$</th>
<th>PIC</th>
<th>PID</th>
<th>PIDab</th>
<th>Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cca02a</td>
<td>222–236</td>
<td>0.500*</td>
<td>0.692*</td>
<td>0.623</td>
<td>0.1404</td>
<td>0.4495</td>
<td>0.188</td>
</tr>
<tr>
<td>Cca02b</td>
<td>99–109</td>
<td>0.433</td>
<td>0.408</td>
<td>0.360</td>
<td>0.3687</td>
<td>0.6494</td>
<td>0.136</td>
</tr>
<tr>
<td>Cca04</td>
<td>122–132</td>
<td>0.667</td>
<td>0.613</td>
<td>0.536</td>
<td>0.2035</td>
<td>0.5947</td>
<td>0.379</td>
</tr>
<tr>
<td>Cca10</td>
<td>131–137</td>
<td>0.433</td>
<td>0.488</td>
<td>0.412</td>
<td>0.3141</td>
<td>0.5945</td>
<td>0.136</td>
</tr>
<tr>
<td>Cca20</td>
<td>192–194</td>
<td>0.167</td>
<td>0.155</td>
<td>0.141</td>
<td>0.7088</td>
<td>0.8560</td>
<td>0.021</td>
</tr>
<tr>
<td>Cca21</td>
<td>186–194</td>
<td>0.667*</td>
<td>0.608*</td>
<td>0.517</td>
<td>0.2257</td>
<td>0.5118</td>
<td>0.379</td>
</tr>
<tr>
<td>Cca22</td>
<td>144–146</td>
<td>0.300</td>
<td>0.259</td>
<td>0.222</td>
<td>0.5622</td>
<td>0.7694</td>
<td>0.064</td>
</tr>
<tr>
<td>Cca23</td>
<td>143–145</td>
<td>0.067</td>
<td>0.066</td>
<td>0.062</td>
<td>0.8667</td>
<td>0.9371</td>
<td>0.004</td>
</tr>
<tr>
<td>Cca24</td>
<td>137–143</td>
<td>0.567</td>
<td>0.595</td>
<td>0.544</td>
<td>0.1834</td>
<td>0.5108</td>
<td>0.253</td>
</tr>
<tr>
<td>Cca26</td>
<td>128–142</td>
<td>0.400</td>
<td>0.579</td>
<td>0.477</td>
<td>0.2633</td>
<td>0.5348</td>
<td>0.114</td>
</tr>
<tr>
<td>Cca29</td>
<td>110–112</td>
<td>0.567</td>
<td>0.533</td>
<td>0.440</td>
<td>0.2916</td>
<td>0.5654</td>
<td>0.253</td>
</tr>
<tr>
<td>Cca35</td>
<td>136–138</td>
<td>0.233</td>
<td>0.210</td>
<td>0.185</td>
<td>0.6275</td>
<td>0.8098</td>
<td>0.040</td>
</tr>
<tr>
<td>Cca39</td>
<td>103–109</td>
<td>0.567</td>
<td>0.676</td>
<td>0.612</td>
<td>0.1441</td>
<td>0.4588</td>
<td>0.253</td>
</tr>
<tr>
<td>Cca40</td>
<td>165–177</td>
<td>0.700</td>
<td>0.690</td>
<td>0.627</td>
<td>0.1344</td>
<td>0.4493</td>
<td>0.428</td>
</tr>
<tr>
<td>Cca42</td>
<td>93–95</td>
<td>0.433</td>
<td>0.440</td>
<td>0.339</td>
<td>0.3991</td>
<td>0.6375</td>
<td>0.136</td>
</tr>
</tbody>
</table>

*Size determined using an M13 universal ladder.

Juxtaposed microsatellite sequence.
Cross-species amplification of microsatellite loci can be subject to ascertainment bias, a characteristic in which polymorphic loci developed in a target species are less polymorphic when utilised in closely related species (Ellegren et al., 1997; Wright et al., 2004; Delport et al., 2006). For example, in a study that used a classical cloning procedure to isolate microsatellite loci in a population of stalk-eyed flies, successful amplification of loci, expected heterozygosity, and polymorphism all decreased as phylogenetic genetic distance increased (Wright et al., 2004). However, Ellegren et al. (2007) noted that ascertainment bias could also be related to the process of selecting loci via the cloning process, which favours longer repeat motifs, and also by a potential lack of published monomorphic loci (i.e. if most monomorphic loci are not published for the focal species, they are not available for cross-species screens). Ascertainment bias is often cited as a limitation of microsatellite loci in population genetics analyses, favouring the use of species-specific markers over cross-species amplification (Hogan, et al., 2009).

However, reverse ascertainment bias – the finding of higher levels of genetic diversity in non-focal species – has been observed in microsatellite markers developed for several avian species, including owls (Hogan et al., 2009), Black Oystercatchers (Williams et al., 2012), and now, Emperor Geese. Thus, the findings of this study fail to support one of the posited hypotheses (selection of microsatellite loci) accounting for ascertainment bias.

Expected heterozygosity was highest in Cackling, Canada, and Greater White-fronted geese, but was relatively similar, ranging from 0.599 to 0.584, lower in Black Brant (0.501) and the lowest in Emperor Geese (0.467). More markers were more highly polymorphic in Greater White-fronted Geese (18) than any other species assayed, with the highest number of alleles across loci (2–12);
this was twice that of Emperor Geese (2–6). However, expected heterozygosity was still slightly higher among some Branta geese (Cackling and Canada goose) than in the Greater White-fronted Goose, which nevertheless exhibited the highest $H_o$ (0.578). Among the geese selected for the cross-species screen, the Greater White-fronted Goose is phylogenetically the closest relative of the Emperor Goose (Donne-Goussé et al., 2002). Thus, the finding of the highest levels of genetic diversity, based on number of alleles and number of polymorphic loci, in Greater White-fronted Goose relative to Branta geese for markers developed in the Emperor Goose is consistent with a hypothesis of ascertainment bias due to increasing phylogenetic distance. However, it begs the question as to why levels of genetic diversity are so low in Emperor Geese such that reverse ascertainment bias was observed.

Reverse ascertainment bias in other species may be attributed to low genetic diversity due to relatively low effective population size (Li and Kimmel, 2013), restricted distribution, unusual mating strategy, and historical demography (Hogan et al., 2009). Certainly, Emperor Geese, while socially monogamous (Schmutz et al., 2011) like other Arctic-nesting geese, have a restricted distribution relative to the other goose species assayed, and presumably a relatively low effective population size. Despite the reduction of the Emperor Goose population in the 1960s and slow recovery, no signature of recent bottleneck was uncovered in Alaska by BOTTLENECK. Luikart et al. (1998) suggests BOTTLENECK can best detect demographic decreases in situations where the population size drops to below 25 effective breeders, a value much lower than the number of breeding individuals in Alaska, even during the years of lowest population size. Hogan et al. (2009) suggest that ascertainment bias in loci may be more common among rare species, and that loci discarded as monomorphic in the target species might be an under-utilised source of polymorphic markers in both phylogenetically closely related and distant species; this study supports that contention. As such, additional loci, which produced either monomorphic or co-amplified products in the Emperor Goose, have been made available on GenBank (accession numbers: KY385249-KY385268, MF093617) for potential use in other goose species.

5. ACKNOWLEDGEMENTS

Funding was provided by the U.S. Geological Survey (USGS), Alaska Science Center. Emperor Goose samples were provided by Christian Dau (formally of USFWS). Other samples were provided by members of the USGS: Greater White-fronted Goose by Craig Ely (collected by Rachel Brubaker and Gene Peltola, USFWS); Canada Goose and Black Brant by Tom Fondell; and Cackling Goose by Paul Flint and Craig Ely. Loci developed for Anser and Brant species were screened in Emperor Geese by Valerie Watkins. This manuscript was improved by comments from J. Pearce and an anonymous reviewer. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Published online: #201#

6. REFERENCES


McCloskey, S.E. (2013) *Variation in metabolic rates of eggs of Emperor Goose: a possible adaptation for intraspecific nest parasitism*. Alaska Pacific University, Anchorage, AK.


