SHORT COMMUNICATION

Development and characterization of microsatellite markers in Rosy-faced and other lovebirds (*Agapornis* spp.) using next-generation sequencing

Derek Kong Lam[1] · Simon Yung Wa Sin[1,2]

Abstract

Agapornis are a group of small African parrots that are heavily traded around the world. They are invasive species in many places, but some of them are listed as Vulnerable or Near Threatened. However, the genetic tools for assessing inter-individual relationships, population structure, and genetic diversity of these birds are very limited. Therefore, we developed polymorphic microsatellite markers in *A. roseicollis* and tested the transferability on 5 lovebird species including *A. personatus*, *A. nigrigenis*, *A. fischeri*, *A. pullarius*, and *A. canus*, and two closely related outgroups (i.e. *Bolbopsittacus lunulatus* and *Loriculus galgulus*). We first performed whole-genome re-sequencing on five individuals of *A. roseicollis* to identify potential polymorphic loci. Out of 37 loci tested in 11 *A. roseicollis*, 27 loci were demonstrated to be polymorphic, with the number of the alleles ranging from 2 to 7 (mean = 3.963). The observed heterozygosity ranged from 0 to 0.875 (mean = 0.481) and expected heterozygosity ranged from 0.233 to 0.842 (mean = 0.642). Five loci (*Agro-A13*, $p < 0.01$; *Agro-A15*, $p < 0.05$; *Agro-A43*, $p < 0.05$, *Agro-A65*, $p < 0.05$; *Agro-A67*, $p < 0.05$) were detected to deviate from Hardy-Weinberg equilibrium, with the presence of null alleles suggested in locus *Agro-A13* and *Agro-A77*. The exclusion powers for PE1 and PE2 are 0.997 and 0.999, respectively. The 27 novel polymorphic markers developed here will be useful for parentage and kinship assignment and population genetics study in *Agapornis*, and provide a tool for scientific research, captive breeding industry, and invasion and conservation management of these species.

Keywords *Agapornis roseicollis* · High throughput sequencing · Kinship inference · Parentage analysis · Peach-faced lovebirds · Polymorphic microsatellite loci

Introduction

The genus *Agapornis* (lovebirds) is a group of small African parrots, including *A. personatus*, *A. nigrigenis*, *A. lilianae*, *A. fischeri*, *A. roseicollis*, *A. pullarius*, *A. taranta*, *A. canus*, and *A. swindernianus* (Online Resource 1; [1]). Due

to small body size and colorful plumage, they are very popular as companion pets, and thus being traded heavily around the world each year [2, 3]. Although sources from captive breeding dominates lovebird trades, the wild populations of six lovebird species (i.e. *A. nigrigenis*, *A. lilianae*, *A. fischeri*, *A. roseicollis*, *A. pullarius*, and *A. swindernianus*) were suspected to be declining, likely due to illegal bird trades, hunting, habitat loss, etc. [4]. Besides, out of the nine lovebird species, *A. lilianae* and *A. fischeri* are listed as Near Threatened and *A. nigrigenis* is listed as Vulnerable [4]. Conservation measures such as captive breeding and ex-situ conservation program have been proposed to sustain these species in the wild in order to prevent them from population decline. Knowledge of population structure and genetic diversity is thus critical for designing conservation strategies. However, there is a lack of population genetic study on any of the above species.

✉ Simon Yung Wa Sin
sinyw@hku.hk

1   School of Biological Sciences, The University of Hong Kong, Pok Fu Lam Road, Hong Kong, China

2   Department of Organismic and Evolutionary Biology, Museum of Comparative Zoology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA

Besides, with the large trade volume, some lovebirds became introduced species with wide-ranged distributions worldwide [2]. Occurrences of non-native lovebird populations have been reported in African, Europe, Australia, and Northern America, with *A. fischeri* and *A. roseicollis* having the highest abundance [2]. Since they feed on seeds, leaves, fruits, and some cultivated plants, they could be agricultural pests [5–7], such as the introduced *A. personatus* in Kenya [6]. Also, lovebirds are famous for their destructive behaviors. They could cause potential damages to human infrastructures and native fauna [7, 8]. Breeding site competitions have also been suggested between *A. fischeri* and *Passer domesticus* in France [8] as well as *A. personatus* and *Cypsiurus parvus* in Kenya [9]. Despite all of these possible threats to native fauna, the knowledge on their invasion histories, dispersal patterns and invasion assessments are under-investigated [2]. In addition, in the captive breeding industry, lovebird breeders usually rely on pedigree to identify kinship and parentage to breed for commercially valuable color mutants [10]. However, this approach has many limitations and is not always reliable.

Regarding the above issues, there is a demand to develop an easily accessible genetic tool to investigate inter-individual relationships and population genetics of *Agapornis* spp. Microsatellite markers are commonly employed for these purposes. In this study, we therefore aimed to develop polymorphic microsatellite markers in *A. roseicollis* first as its draft genome is available [11]. We performed whole-genome re-sequencing on five individuals of *A. roseicollis* to screen for potential polymorphic microsatellite loci. We then tested the transferability of the characterized markers on five other *Agapornis* species (i.e. including *A. personatus*, *A. nigrigenis*, *A. fischeri*, *A. pullarius*, and *A. canus*) and two closely related outgroups (i.e. *Bolbopsittacus lunulatus* and *Loriculus galgulus*). The novel polymorphic microsatellite markers developed here are not only useful for the lovebird breeding industry but also for conservation planning, invasion management, and scientific research.

Methods

Agapornis blood and tissue samples were collected from local breeders and museums, respectively (Online Resource 2). Blood samples were stored in Queen's lysis buffer (0.01 M Tris, 0.01 M sodium-EDTA, 1.0% *n*-lauroylsarcosine, 0.01 NaCl, pH 8.0) until DNA extraction. We performed whole-genome re-sequencing on 5 individuals of *A. roseicollis*. The DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen, Germany). The extracted DNA was sheared to targeted size using Covaris S220. We used the PrepX Complete ILMN 32i DNA Library Kit (TaKaRa, US) to prepare the 220 bp fragment libraries. The libraries were quantified using the Qubit dsDNA HS Assay Kit (Invitrogen, US), and then multiplexed in equamolar ratio for sequencing using the Illumina HiSeq 2500 instrument. Low coverage sequencing (~5X) was performed.

Primers for potential polymorphic microsatellite loci were designed based on the draft genome of *A. roseicollis* (GenBank: NDXB00000000.1) and the whole-genome re-sequencing data (Sin et al. in prep). MSATCOMMANDER 1.0.8 [12] was first used to search for the nucleotide repeats of dinucleotide, trinucleotide, and tetranucleotide with more than eight repeat units in the draft genome, coupled with primer design by Primer3 [13] implemented in the software. The forward primers were tagged with M13 sequence labeled with FAM (5′-GGAAACAGCTATGACCAT-3′) [14]. Next, to determine if the identified loci were polymorphic, we mapped the re-sequencing data of the 5 individuals to the loci. The raw reads of re-sequencing were first filtered and adapter-trimmed by Trimmomatic 0.35 [15]. The trimmed reads were mapped onto the draft genome by BWA-MEM algorithm in Burrows-Wheeler Aligner (BWA 0.7.17) with default settings [16]. The mapping results were visualized in IGV 2.4.14 [17]. A locus was identified to be polymorphic if a tandem repeat number variation was found between the re-sequencing data and the reference genome for at least one of the 5 individuals. Among the potential polymorphic loci, 34 loci were randomly selected from the database with the following criteria: (1) the repeat number of the short tandem repeat in the reference genome is more than 8 but less than 20; (2) loci were selected from each class of repeats evenly (dinucleotide, trinucleotide, and tetranucleotide); (3) Repeats with more than 70% GC content were avoided (e.g. CG and TGGG); (4) the selected loci are at different scaffolds of the draft genome; (5) no indel or single nucleotide polymorphism (SNP) was found in the primer binding sites based on the whole-genome re-sequencing mapping; (6) the size of the PCR product is between 100 bp to 500 bp; and (7) several primer design criteria have to be fulfilled (e.g. primer GC content around 50%, at least 1-bp GC clamp at 3′ site of the primers, low levels of self- or pair-complementarity, etc.; [12, 18]). In addition, three loci from the Dawson et al. [19] study were also selected for characterization and the primers from that study were modified.

Genotyping PCR was performed in 10 μl reaction volume containing 1X GoTaq Reaction Buffer, 0.1 μg/ μl BSA, 2 mM MgCl$_2$, 0.2 mM dNTP, 0.01 μM forward primer tagged with M13R, 0.15 μM reverse primer, 0.15 μM fluorescently labelled M13R primer, 20 ng gDNA and 0.25 unit GoTaq Polymerase (Promega, US). The cycling conditions were 1 cycle of 2 min at 95 °C; 20 cycles of 30 s at 95 °C, 30 s at 54–62 °C, 30 s at 72 °C; 20 cycles of 30 s at 95 °C, 30 s at 55 °C, 30 s at 72 °C; 1 cycle of 10 min at 72 °C. The PCR products were sent to BGI (Shenzhen, China) or CPOS (Centre for PanorOmic Sciences, The University of

Hong Kong, HKSAR) for sizing using 3730xl DNA Analyzer (Applied Biosystems) with 500 LIZ Dye Size Standard (Applied Biosystems). In total, 11 individuals of *A. roseicollis* were genotyped. The same protocol was used for the cross-species amplification.

Monomorphic loci were excluded from the analysis. The pairwise relatedness using Lynch & Ritland estimator [20] was first calculated by GenAlEx 6.5 [21] implemented in Excel (Microsoft). If the pairwise relatedness (*r*) between two individuals were > 0.2, only one of them was included randomly for downstream analyses. Three pairs of *A. roseicollis* were detected to have *r* > 0.2 value (Online Resource 2). Two pairs (*r* = 0.479 and 0.521) were known first-order kins from the local breeders in Hong Kong. Another case (*r* = 0.3404) was a pair with an unknown relationship, where both individuals were from the same museum. The loci were then tested for Hardy-Weinberg equilibrium and linkage disequilibrium using GENEPOP 4.2 [22]. The observed and expected heterozygosity were calculated. The presence of null allele, allele dropout and scoring error due to stutter were checked using MICRO-CHECKER 2.2.3 [23]. To determine the power of our characterized loci in parentage analysis, the polymorphic information content (PIC), probability of exclusions (both parents are unknown, PE1; one parent is known, PE2), and probability of identity (PI) were calculated using CERVUS 3.0.7 [24]. The accumulated PE1, PE2, and PI were also calculated for all possible combinations of loci, excluding those loci with null alleles (*p* < 0.05). We performed BLASTN against the online Nucleotide collection (nr/nt) database (NCBI) to check whether the scaffolds that contain the loci-of-interest are located on the sex chromosome.

Results and discussion

Using MSATCOMMANDER, 14626 short tandem repeats (> 8 repeat units) were identified in the draft genome of *A. roseicollis*, including 9532 dinucleotides, 1351 trinucleotides, and 3743 tetra-nucleotides. The mapping results of the whole-genome re-sequencing data allowed us to identify polymorphic loci for PCR amplification and characterization. Out of the 37 selected loci, 29 loci were successfully amplified and optimized in 11 individuals of *A. roseicollis*. Two of the 29 loci were demonstrated to be monomorphic in our tested samples. This result was inconsistent with the mapping results, in which the tandem repeat variants of the mapped individuals were different from the draft genome. Possible explanations include genotyping errors and mapping errors due to low converge re-sequencing data. The number of alleles for the 27 polymorphic loci ranged from 2 to 7 (mean = 3.963; Table 1). The observed heterozygosity ranged from 0 to 0.875 (mean = 0.481) and expected

heterozygosity ranged from 0.233 to 0.842 (mean = 0.642). Loci *Agro-A13* (*p* < 0.01), *Agro-A15* (*p* < 0.05), *Agro-A43* (*p* < 0.05), *Agro-A65* (*p* < 0.05), and *Agro-A67* (*p* < 0.05) were detected to be deviated from Hardy-Weinberg equilibrium significantly. Eight pairs of loci showed significant linkage disequilibrium (*Agro-A45* & *Agro-A59*, *Agro-A07* & *Agro-65*, *Agro-59* & *Agro-65*, *Agro-59* & *Agro-67*, *Agro-23* & *Agro-A73*, *Agro-A27* & *Agro-A73*, *Agro-A57* & *Agro-A73*, and *Agro-A15* & *Agro-A75*; *p* < 0.05). The deviations from Hardy-Weinberg equilibrium and the linkage disequilibrium could probably be an artefact as all tested individuals were likely originated from different populations (Online Resource 2). Null alleles were significantly evidenced for the locus *Agro-A13* (*p* < 0.01) and *Agro-A75* (*p* < 0.05) as was indicated by the general excess of homozygotes for most of the allele classes. The presence of null alleles at the locus *Agro-A13* might also explain its deviation from Hardy-Weinberg equilibrium. There was no evidence of sex-linked loci because the scaffolds containing the loci were not located on the sex chromosome based on the BLAST results, and neither sex demonstrated only homozygotes. The presence of null alleles was probably due to preferential amplification during PCR given that there was also no evidence of allele dropout (short allele dominance) or scoring error due to stuttering [23]. After removing the two loci with null alleles, the remaining 25 polymorphic loci showed reasonable PIC, ranging from 0.195 (*Agro-29*) to 0.759 (*Agro-A47*) with a mean of 0.532. The total exclusion powers of the 25 loci were 0.997 (PE1) and 0.999 (PE2) with the total PI of 2.11×10^{-17} (Fig. 1), which indicates the great potential of these 25 loci as a tool for parentage and kinship assignment as well as population genetics study.

The cross-species utility of the 27 polymorphic microsatellite loci developed in *A. roseicollis* was also tested on another five *Agapornis* species (i.e. *A. personatus*, *A. nigrigenis*, *A. fischeri*, *A. pullarius*, and *A. canus*) and two closely-related outgroups (i.e. *B. lunulatus* and *L. galgulus*; Table 2). In ten individuals of *A. fischeri*, all 27 loci were shown to be polymorphic with the number of alleles ranging from 2 to 8 (mean = 4.52). The observed and expected heterozygosity were from 0 to 1 (mean = 0.496) and from 0.189 to 0.853 (mean = 0.612), respectively. Three loci (*Agro*-A13, *p* < 0.05; *Agro-A51*, *p* < 0.01 and; *Agro-A79*, *p* < 0.05) showed a significant deviation from Hard-Weinberg equilibrium while two pairs of loci (*Agro-39* & *Agro-57* and *Agro-51* & *Agro-79*; *p* < 0.05) showed a significant linkage disequilibrium. Similar to *A. roseicollis*, these deviations from Hard-Weinberg equilibrium and linkage disequilibrium could probably result from mixing individuals from different populations (Online Resource 2). Null alleles were detected in two loci (*Agro-A13* and *Agro-A51*; *p* < 0.01) with no evidence of allele dropout (short allele dominance), scoring error due to stuttering, and sex-linked loci. The presence of

Table 1 Development and characterization of 27 polymorphic microsatellite markers for *Agapornis roseicollis* (n = 11)

Locus	Repeat motif[a]	Primer sequences (5′–3′)[b]	Ta (°C)[c]	Expected allele size[a]	Allele size range (bp)[d]	No. of alleles	Ho[e]	He[e]	HWE[e]	Null alleles[e]	PIC[e]	PE1[e]	PE2[e]	PI[e]
TG13-017	(TA)7	F: GGAAACAGCTATGACCA TGCTTTGCATCTTGCCTT AAAG R: GGTAACTACAACATT CCAACTCCTG	57	208	205–207	2	0.250	0.500	ns	ns	0.359	0.110	0.179	0.392
TG03–002	(TA)10	F: GGAAACAGCTATGACCA TCTTGCCTTTCTGGTATG AGTATAGC R: AAGTACAAAGCACTG TGGAGCAG	58	129	126–130	3	0.625	0.708	ns	ns	0.590	0.220	0.368	0.187
Agro-A07	(AC)15	F: GGAAACAGCTATGACCA TGAGATTAACCACCATAG CCAC R: GAGTTGGGAAGGCAA GAAGG	58	376	365–371	4	0.750	0.717	ns	ns	0.612	0.245	0.406	0.168
Agro-A09	(TGA)12	F: GGAAACAGCTATGAC CATGTACACCCTGAGAGC TCCTG R: TCTGAGGTGGTGCTG AGTTG	61	310	302–317	5	0.875	0.708	ns	ns	0.618	0.254	0.426	0.159
Agro-A13	(AC)15	F: GGAAACAGCTATGACCA TGGAACGCACCTTAAAGC AG R: TGTGCACCAAGATCA GCAAC	62	109	98–114	5	0.250	0.742	**	**	0.645	0.272	0.444	0.144
Agro-A15	(GATA)11	F: GGAAACAGCTATGAC CATTCTCACCTCATGCTC TTGGG R: GCTCTCGGGATGCATT GTGAC	58	181	176–184	3	0.375	0.685	*	ns	0.544	0.190	0.333	0.220
Agro-A23	(ATAG)8	F: GGAAACAGCTATGACCA TGGAAGTGCAGGTTCAGG AG R: TGTGCACAAATATTCATG GCC	58	353	346–354	3	0.625	0.692	ns	ns	0.575	0.210	0.357	0.197
Agro-A25	(GAT)8	F: GGAAACAGCTATGAC CATAAAGCCTCTTCTCCA TCGCC R: CGCTTTGAAAATCGCG GTTTC	57	343	334–340	3	0.625	0.692	ns	ns	0.575	0.210	0.357	0.197

Table 1 (continued)

Locus	Repeat motif[a]	Primer sequences (5′–3′)[b]	Ta (°C)[c]	Expected allele size[a]	Allele size range (bp)[d]	No. of alleles	Ho[e]	He[e]	HWE[e]	Null alleles[e]	PIC[e]	PE1[e]	PE2[e]	PI[e]
Agro-A27	$(CT)_6(CA)_{12}(GCT)_3$	F: GGAAACAGCTATGAC CATCCTCTCAGCGTCCTC AGTC R: GGAGAAACAAGCAGC AGTGG	58	248	247–253	4	0.250	0.592	ns	ns	0.456	0.154	0.257	0.297
Agro-A29	$(GT)_{13}$	F: GGAAACAGCTATGAC CATCTCCAGGAGTTAATG TTTGCAC R: CAGCCTGCACACGTG TTTC	58	308	301–307	2	0.000	0.233	ns	ns	0.195	0.024	0.097	0.634
Agro-A37	$(GAG)_{13}$	F: GGAAACAGCTATGAC CATCCCTATTCCTCTCGT CCCAG R: GATCGCGCCGTCATC TTTAG	58	359	336–342	3	0.250	0.342	ns	ns	0.294	0.051	0.163	0.488
Agro-A39	$(ATCC)_{13}$	F: GGAAACAGCTATGACCA TGTCCGTCCATCCACTTC TC R: ACATCCAGGGTAAGA GTGCC	57	253	242–254	4	0.500	0.767	ns	ns	0.667	0.293	0.465	0.130
Agro-A43	$(AGAT)_{12}$	F: GGAAACAGCTATGAC CATAAAGACACGCCAGCA CAAC R: ACTGCTTTGAGAATG AGATCCC	60	402	385–401	4	0.250	0.575	*	ns	0.447	0.145	0.252	0.305
Agro-A45	$(GAG)_{10}$	F: GGAAACAGCTATGACCA TGATTGACTGGCCTTCCC TC R: GGCGGAATGTTAAGC AGGAC	62	388	398–404	3	0.250	0.342	ns	ns	0.294	0.051	0.163	0.488
Agro-A47	$(CAAC)_{10}$	F: GGAAACAGCTATGACCA TGCTTGTCTCCTGGATTT GC R: ACTTCTCAGCTTGTC CCTCC	62	188	178–198	6	0.750	0.842	ns	ns	0.759	0.412	0.591	0.075

Table 1 (continued)

Locus	Repeat motif[a]	Primer sequences (5′–3′)[b]	Ta (°C)[c]	Expected allele size[a]	Allele size range (bp)[d]	No. of alleles	Ho[e]	He[e]	HWE[e]	Null alleles[e]	PIC[e]	PE1[e]	PE2[e]	PI[e]
Agro-A51	$(TA)_{13}$	F: GGAAACAGCTATGAC CATAACCACCAGCAGAGA TTCAG R: TGGGACAGACTTAAG CAGGC	62	427	441–451	4	0.375	0.592	ns	ns	0.510	0.165	0.325	0.243
Agro-A53	$(ATG)_{14}(AT)_3(TG)_3$	F: GGAAACAGCTATGAC CATCTTGCCATACCAACC TTTGC R: GGAAATCACCTACCC ATGCAC	58	319	327–354	4	0.500	0.642	ns	ns	0.525	0.190	0.325	0.235
Agro-A57	$(TG)_{14}$	F: GGAAACAGCTATGAC CATACCTCAGCCACACCC TC R: CCTCTCTGCCAGTCTG ACTTG	58	416	423–429	3	0.500	0.567	ns	ns	0.468	0.141	0.277	0.283
Agro-A59	$(ATA)_{10}$	F: GGAAACAGCTATGAC CATCCATCCTTCTCTCCA GTGGG R: ATTGGTTGGCAGGTG GTAAG	54	234	240–249	4	0.750	0.608	ns	ns	0.496	0.163	0.295	0.259
Agro-A63	$(AAT)_{11}$	F: GGAAACAGCTATGAC CATCCTGCCAAATTCATA AAGCAGC R: TACTGTTGGCTTCTTGGG AC	58	323	331–337	3	0.250	0.575	ns	ns	0.482	0.145	0.289	0.270
Agro-A65	$(AAT)_{14}$	F: GGAAACAGCTATGAC CATAACACTCTCACTGGC CTTCC R: CAATGAACAGAGCAC GGTGG	62	185	196–208	5	0.500	0.808	*	ns	0.719	0.357	0.535	0.098
Agro-A67	$(CCT)_8$	F: GGAAACAGCTATGAC CATCCAAGCACAAAACACC TCCAG R: GGAGGATTGCTGGTG GAGAG	54	355	360–372	4	0.375	0.650	*	ns	0.559	0.201	0.364	0.203

Table 1 (continued)

Locus	Repeat motif[a]	Primer sequences (5′–3′)[b]	Ta (°C)[c]	Expected allele size[a]	Allele size range (bp)[d]	No. of alleles	Ho[e]	He[e]	HWE[e]	Null alleles[e]	PIC[e]	PE1[e]	PE2[e]	PI[e]
Agro-A73	(GATA)$_{11}$	F: GGAAACAGCTATGACCATGTTAGAGGAGGGGCACACTC R: TGAGTGGAGCAGTACTGAGC	62	432	439–451	4	0.625	0.708	ns	ns	0.616	0.247	0.419	0.160
Agro-A75	(TCTA)$_{12}$	F: GGAAACAGCTATGACCATGAAAGACAGCAGCCCTAAG R: AGATGTTCTGGTGCTAAGACAG	62	126	122–142	6	0.750	0.792	ns	ns	0.701	0.335	0.513	0.108
Agro-A77	(GATG)$_{11}$ GATA (GATG)$_4$	F: GGAAACAGCTATGACCATCTCCCTCTGCAGCTGTAG R: GAGTGGTTTCAAGCCAGCTC	62	302	274–326	7	0.500	0.833	ns	*	0.755	0.412	0.593	0.074
Agro-A79	(TATC)$_{13}$	F: GGAAACAGCTATGACCATGGTTCTGCTTCTCACAGAC R: GCACCTTCAGAAGCAACCG	62	365	360–380	5	0.625	0.775	ns	ns	0.682	0.316	0.491	0.120
Agro-A81	(CA)$_{13}$	F: GGAAACAGCTATGACCATCCCACATGTGATCCAAACCC R: CAATGTCTTGTAGCAGGCAGG	62	343	357–367	4	0.625	0.650	ns	ns	0.559	0.201	0.364	0.203

Five individuals were from museums (see Online Resource 2) and six were from local breeders in Hong Kong

[a]The repeat motif and expected allele size were estimated from the draft genome of *A. roseicollis* (GeneBank: NDXB 00000000.1)

[b]The forward primers (F) are tagged with the universal M13 sequence (underlined). Reverse primers (R)

[c]The annealing temperatures (Ta) were optimized in *A. roseicollis*

[d]Observed allele size range

[e]Two known full-sibs and one speculated half-sib were excluded from the calculation of observed heterozygosity (Ho), expected heterozygosity (He), probability of deviation from Hardy-Weinberg equilibrium (HWE), presence of null alleles (null alleles), polymorphic information content (PIC), probability of exclusions (bath parents unknown, PE1; one parent known, PE2), and probability of identity (PI)

[f]Locus TG13–017 and TG03–002 were modified from Dawson et al. [19]
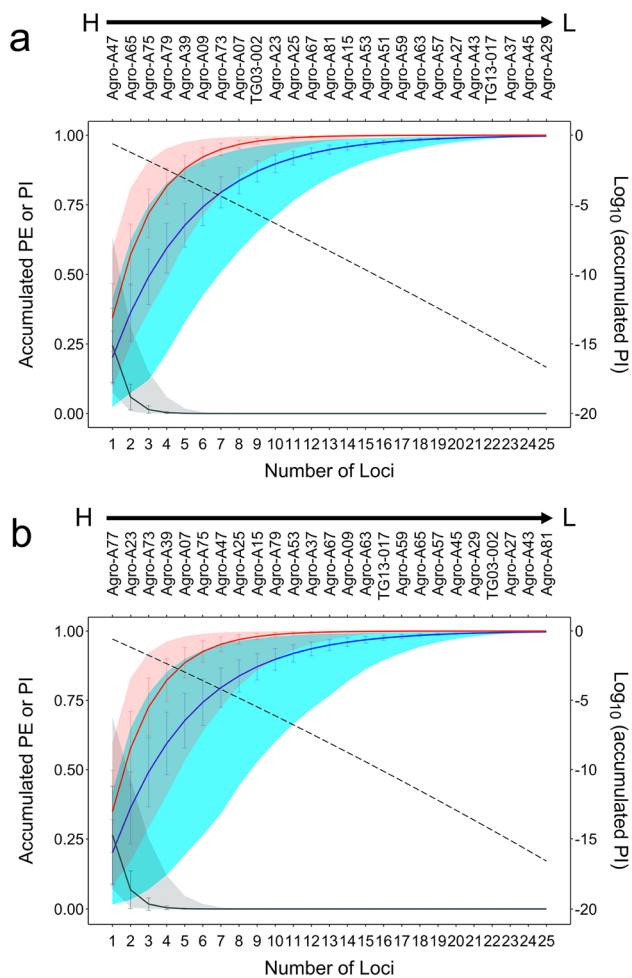
*$p < 0.05$; **$p < 0.01$

Fig. 1 The probability of exclusions (PE) and probability of identity (PI) of the 25 microsatellite markers in **a** *Agapornis roseicollis* and **b** *Agapornis fischeri*. Loci with the presence of null alleles (Agro-A13 and Agro-A77 for *A. roseicollis*; Agro-A13 and Agro-A51 for *A. fischeri*) were excluded from the analysis. The accumulated PE are represented by the blue (both parents are unknown; PE1) and red (one parent is unknown; PE2) colours. The accumulated PI is represented by the grey colour. The shadow areas indicate all possible values of accumulated PE1, PE2, and PI calculated from all combinations of loci with the mean (solid lines) and standard errors (error bars) shown. The dashed line indicates the $\log_{10}$ values of the mean of the accumulated PI. The loci at the top is the combination that corresponds to the upper limits of the accumulated PE1 and PE2 as well as the lower limit of the accumulated PI. The two combinations were arranged from the highest (H) PE1/PE2 to the lowest (L) PE1/PE2 as indicated by the arrows

null alleles could be due to preferential amplification during PCR. Excluding the two loci with the presence of null alleles

in *A. fischeri*, the remaining 25 loci showed comparable PIC (mean = 0.521; from 0.164 to 0.765), total exclusion powers (PE1 = 0.997; PE2 = 0.999), and total PI (2.61×10^{-17}; Fig. 1) to the values observed in *A. roseicollis*. These 25 loci originally developed for *A. roseicollis* could also facilitate parentage and kinship assignment as well as population genetic studies in *A. fischeri*. Besides, the rate of successful amplification of all polymorphic 27 loci in the rest of species varied (*A. personatus*, 100%; *A. nigrigenis*, 88.9%; *A. pullarius*, 59.3%; *A. canus*, 51.9%; *B. lunulatus*, 37.0%; and *L. galgulus*, 29.6%; Table 2). This variation to some extent is correlated with the phylogenetic relationship between those species and *A. roseicollis* (Online Resource 1), with more closely related species such as those eye-ring species (e.g. *A. fischeri*, *A. personatus*, and *A. nigrigenis*) had higher successful rate.

In conclusion, the 27 novel microsatellite markers developed here displayed reasonable polymorphism for both *A. roseicollis* and probably all eye-ring species such as *A. fischeri*, *A. personatus*, and *A. nigrigenis*. The high exclusion powers observed in both *A. roseicollis* and *A. fischeri* (two different panels) further indicated that the isolated loci are useful tools for inferring kinship and parentage in captive breeding program. Since *A. roseicollis*, *A. fischeri* and *A. personatus* have the most records of established non-native populations worldwide and may pose threat to local fauna [2, 7], the newly developed markers will allow more detailed investigations into their invasion histories and dispersal patterns, which could further help biosecurity policy development. This study will also facilitate the conservation management of the threatened or near threatened lovebirds in their native habitats.

Table 2 Cross-species utility of 27 *Agapornis roseicollis* microsatellite markers in another five *Agapornis* species and two closely-related outgroups

Locus	*Agapornis fischeri* (n = 10)										*Agapornis personatus* (n = 4)	*Agapornis nigrigenis* (n = 1)	*Agapornis canus* (n = 3)	*Agapornis pullarius* (n = 3)	*Bolbopsittacus lunulatus* (n = 1)	*Loriculus galgulus* (n = 1)
	Obs. Allele size range (bp)	No. of alleles	Ho	He	HWE	Null alleles	PIC	PE1	PE2	PI						
TG13–017[a]	205–209	3	0.400	0.584	ns	ns	0.491	0.154	0.294	0.262	207, 209	209	205	204	218	217, 221
TG03–002[a]	126–130	3	0.200	0.353	ns	ns	0.303	0.056	0.166	0.474	126, 128	126, 128	X	128	X	128
Agro-A07	363–371	5	0.600	0.768	ns	ns	0.688	0.320	0.498	0.115	369	365	X	X	X	X
Agro-A09	293–317	5	0.500	0.568	ns	ns	0.508	0.163	0.331	0.244	292, 310	302	303, 305, 015	289, 295, 298	X	X
Agro-A13	88–114	8	0.500	0.853	*	**	0.790	0.466	0.643	0.056	88, 90, 94, 98, 110	90, 108	X	X	X	X
Agro-A15	172–184	4	0.600	0.711	ns	ns	0.612	0.245	0.403	0.169	176, 180, 184	184	X	X	X	X
Agro-A23	341–361	6	0.600	0.821	ns	ns	0.749	0.399	0.579	0.079	333, 341	357	345, 353, 357, 361	353	X	365
Agro-A25	327–339	5	0.600	0.742	ns	ns	0.653	0.285	0.455	0.139	326, 332, 335, 338	338	330, 336	325, 331	338	341
Agro-A27	245–261	4	0.300	0.284	ns	ns	0.259	0.038	0.147	0.544	241, 243, 249, 251	239	241, 245	231	229	X
Agro-A29	301–307	3	0.300	0.426	ns	ns	0.368	0.082	0.211	0.391	301	301		308, 310, 312	X	X
Agro-A37	339–348	4	0.700	0.684	ns	ns	0.584	0.225	0.379	0.188	342, 345, 348	345	X	X	X	X
Agro-A39	242–258	5	0.700	0.789	ns	ns	0.709	0.342	0.520	0.104	246, 250	239	X	X	X	X
Agro-A43	368–396	3	0.100	0.195	ns	ns	0.177	0.017	0.095	0.672	367	367	360	361	349	X
Agro-A45	391–424	4	0.400	0.432	ns	ns	0.379	0.087	0.224	0.379	391, 394	X	X	X	X	X
Agro-A47	183–199	5	0.700	0.737	ns	ns	0.656	0.288	0.464	0.134	186, 190, 194	177, 181	158, 162	183, 187, 191	X	X
Agro-A51	436–450	7	0.300	0.826	**	*	0.755	0.412	0.590	0.076	444, 446, 448, 450	452	X	X	X	X
Agro-A53	321–333	5	0.700	0.674	ns	ns	0.603	0.238	0.417	0.167	321, 330	330	333, 339, 342, 345, 357	327, 330, 333	334, 337	X
Agro-A57	423–429	3	0.400	0.484	ns	ns	0.41	0.106	0.237	0.341	426, 428	427	424	419	440	449
Agro-A59	240–249	4	0.500	0.574	ns	ns	0.476	0.154	0.286	0.276	240, 243	X	234	X	X	X
Agro-A63	329–335	3	0.300	0.595	ns	ns	0.482	0.160	0.283	0.272	332, 335, 341	323	317	319	315	314
Agro-A65	184–205	5	0.400	0.558	ns	ns	0.489	0.153	0.311	0.262	187, 190, 196, 208	199	X	187, 196	X	X
Agro-A67	357–369	4	0.600	0.595	ns	ns	0.509	0.169	0.319	0.245	357, 363	X	367, 370	374, 377	X	X

Table 2 (continued)

Locus	Agapornis fischeri (n=10)										Agapornis personatus (n=4)	Agapornis nigrigenis (n=1)	Agapornis canus (n=3)	Agapornis pullarius (n=3)	Bolbopsittacus lunulatus (n=1)	Loriculus galgulus (n=1)
	Obs. Allele size range (bp)	No. of alleles	Ho	He	HWE	Null alleles	PIC	PE1	PE2	PI						
Agro-A73	428–448	5	1.000	0.800	ns	ns	0.72	0.357	0.535	0.097	422, 430, 438, 442, 450, 454	441	497, 513, 517, 525, 537, 545	448, 452, 456	X	435
Agro-A75	122–142	6	0.800	0.747	ns	ns	0.671	0.309	0.488	0.123	122, 126, 130, 134	130	118, 122, 126, 130	X	207, 215	136
Agro-A77	293–317	6	0.600	0.837	ns	ns	0.765	0.418	0.597	0.072	309, 313, 317, 325, 329	330	X	309, 313	290	X
Agro-A79	366–382	5	0.600	0.695	*	ns	0.601	0.244	0.403	0.175	366, 370, 374, 378, 382	378	X	X	356, 360	X
Agro-A81	351–357	2	0.000	0.189	ns	ns	0.164	0.016	0.082	0.689	351	353, 355	X	X	X	X
								Number of loci			27	24	14	16	10	8
								Number of polymorphic loci			23	4[b]	9	9	3[b]	1[b]

Markers were characterized in ten individuals of *Agapornis fischeri* with observed allele size range (Obs. allele size range), number of alleles (No. of alleles), observed heterozygosity (Ho), expected heterozygosity (He), probability of deviation from Hardy-Weinberg equilibrium (HWE), presence of null alleles (null alleles), polymorphic information content (PIC), probability of exclusion (bath parents unknown, PE1; one parent known, PE2), and probability of identity (PI) shown. Observed allele sizes and number of polymorphic loci were also determined for *Agapornis personatus* (n = 4), *Agapornis nigrigenis* (n = 1), *Agapornis canus* (n = 3), *Agapornis pullarius* (n = 3), *Bolbopsittacus lunulatus* (n = 3), and *Loriculus galgulus* (n = 1; see Online Resource 2 for sample information). ns = non-significant. X = unsuccessful amplification

[a]Locus TG13–017 and TG03–002 were modified from Dawson et al. [19]

[b]Only one individual was tested

*p < 0.05; **p < 0.01

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

Ethical approval All animals used in this study were approved and handled in accordance with the guidelines provided by the Committee on the Use of Live Animals in Teaching and Research (CULATR) in the Laboratory Animal Unit, HKU (CULATR Approval Number: 4749-18).

Informed consent All authors consent to participate. All authors consent to publication.

References

1. Dilger WC (1960) The comparative ethology of the African parrot genus *Agapornis*. Ethology 17:649–685. https://doi.org/10.1111/j.1439-0310.1960.tb00412.x
2. Mori E, Cardador L, Reino L, White RL, Hernández-Brito D, Le Louarn M, Mentil L, Edelaar P, Pârâu LG, Nikolov BP, Menchetti M (2019) Lovebirds in the air: trade patterns, establishment success and niche shifts of *Agapornis* parrots within their non-native range. Biol Invasions. https://doi.org/10.1007/s10530-019-02100-y
3. Van den Abeele D (2016) Lovebirds compendium. About Pets Publishers, The Netherlands
4. IUCN (2019) The IUCN Red List of Threatened Species. Version 2019–3. http://www.iucnredlist.org. Accessed 10 Dec 2019
5. Ndithia H, Perrin MR (2006) Diet and foraging behaviour of the rosy-faced lovebird *Agapornis roseicollis* in Namibia. Ostrich 77(1–2):45–51. https://doi.org/10.2989/00306520609485507
6. Lever C (2005) Naturalized animals of the British Isles. T and AD Poyser Editions, London
7. Menchetti M, Mori E (2014) Worldwide impact of alien parrots (Aves Psittaciformes) on native biodiversity and environment: a review. Ethol Ecol Evol 26(2–3):172–194. https://doi.org/10.1080/03949370.2014.905981
8. Dubois PJ, Maillard JF, Cugnasse JM (2016) Les populations d'oiseaux allochtones en France en 2015 (4e enquête nationale). Ornithos 23:129–141
9. Lever C (1994) Naturalized animals: the ecology of successfully introduced species. Poyser Natural History, London
10. Zwan H, Visser C, Schoonen M, Van Der Sluis R (2019) Development of an SNP-based parentage verification panel for lovebirds. Anim Genet. https://doi.org/10.1111/age.12859
11. Van Der Zwan H, Van Der Westhuizen F, Visser C, Van Der Sluis R (2018) Draft De novo genome sequence of *Agapornis roseicollis* for application in avian breeding. Anim Biotechnol 29(4):241–246. https://doi.org/10.1080/10495398.2017.1367692
12. Faircloth BC (2008) Msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. Mol Ecol Resour 8(1):92–94. https://doi.org/10.1111/j.1471-8286.2007.01884.x
13. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3--new capabilities and interfaces. Nucleic Acids Res 40(15):e115–e115. https://doi.org/10.1093/nar/gks596
14. Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nat Biotechnol 18(2):233–234. https://doi.org/10.1038/72708
15. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30(15):2114–2120. https://doi.org/10.1093/bioinformatics/btu170
16. Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. ArXiv 1303
17. Thorvaldsdottir H, Robinson JT, Mesirov JP (2013) Integrative genomics viewer (IGV): high-performance genomics data visualization and exploration. Brief in Bioinform 14(2):178–192. https://doi.org/10.1093/bib/bbs017
18. Gardner MG, Fitch AJ, Bertozzi T, Lowe AJ (2011) Rise of the machines--recommendations for ecologists when using next generation sequencing for microsatellite development. Mol Ecol Resour 11(6):1093–1101. https://doi.org/10.1111/j.1755-0998.2011.03037.x
19. Dawson DA, Horsburgh GJ, Küpper C, Stewart IRK, Ball AD, Durrant KL, Hansson B, Bacon I, Bird S, Klein Á, Krupa AP, Lee J-W, Martín-Gálvez D, Simeoni M, Smith G, Spurgin LG, Burke T (2010) New methods to identify conserved microsatellite loci and develop primer sets of high cross-species utility - as demonstrated for birds. Mol Ecol Resour 10(3):475–494. https://doi.org/10.1111/j.1755-0998.2009.02775.x
20. Wang J (2004) Estimating pairwise relatedness from dominant genetic markers. Mol Ecol 13(10):3169–3178. https://doi.org/10.1111/j.1365-294x.2004.02298.x
21. Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in excel. Population genetic software for teaching and research-an update. Bioinformatics 28(19):2537–2539. https://doi.org/10.1093/bioinformatics/bts460
22. Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and Ecumenicism. J Hered 86(3):248–249. https://doi.org/10.1093/oxfordjournals.jhered.a111573
23. Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4(3):535–538. https://doi.org/10.1111/j.1471-8286.2004.00684.x
24. Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program cervus accommodates genotyping error increases success in paternity assignment. Mol Ecol 16(5):1099–1106. https://doi.org/10.1111/j.1365-294x.2007.03089.x