

Standardization and validation of a panel of cross-species microsatellites to individually identify the Asiatic wild dog (*Cuon alpinus*)

Shrushti Modi¹ Equal first author, 1, Bilal Habib¹ Equal first author, 1, Pallavi Ghaskadbi¹, Parag Nigam¹, Samrat Mondol¹ Corresp. 1

¹ Wildlife Institute of India, Chandrabani, Dehradun, Uttarakhand, India

Corresponding Author: Samrat Mondol

Email address: samrat@wii.gov.in

Background. The Asiatic wild dog or dhole (*Cuon alpinus*) is a highly elusive, monophyletic, forest dwelling, social canid distributed across south and Southeast Asia. Severe pressures from habitat loss, prey depletion, disease, human persecution and interspecific competition resulted in global population decline in dholes. Despite a declining population trend, detailed information on population size, ecology, demography and genetics is lacking. Generating reliable information at landscape level for dholes is challenging due to their secretive behaviour and monomorphic physical features. Recent advances in non-invasive DNA-based tools can be used to monitor populations and individuals across large landscapes. In this paper, we describe standardization and validation of faecal DNA-based methods for individual identification of dholes. We tested this method on 249 field-collected dhole faeces from five protected areas of the central Indian landscape in the state of Maharashtra, India.

Results. We tested a total of 18 cross-species markers and developed a panel of 12 markers for unambiguous individual identification of dholes. This marker panel identified 101 unique individuals from faecal samples collected across our pilot field study area. These loci showed varied level of amplification success (57-88%), polymorphism (3-9 alleles), heterozygosity (0.23-0.63) and produced a cumulative misidentification rate or $PID_{(unbiased)}$ and $PID_{(sibs)}$ value of 4.7×10^{-10} and 1.5×10^{-4} , respectively, indicating a high statistical power in individual discrimination from poor quality samples.

Conclusion. Our results demonstrated that the selected panel of 12 microsatellite loci can conclusively identify dholes from poor quality, non-invasive biological samples and help in exploring various population parameters. This genetic approach would be useful in dhole population estimation across its range and will help in assessing population trends and other genetic parameters for this elusive, social carnivore.

Standardization and validation of a panel of cross-species microsatellites to individually identify the Asiatic wild dog (*Cuon alpinus*)

Shrushti Modi^{*}, Bilal Habib^{*}, Pallavi Ghaskadbi, Parag Nigam, Samrat Mondol[#]

Affiliation: Wildlife Institute of India

**Equal contribution*

[#]Corresponding author: Samrat Mondol, Animal Ecology and Conservation Biology Department, Wildlife Institute of India, Chandrabani, Dehradun, Uttarakhand, India 248001
Email: samrat@wii.gov.in

21

22 **Abstract**

23 ***Background***

24 The Asiatic wild dog or dhole (*Cuon alpinus*) is a highly elusive, monophyletic, forest dwelling,
 25 social canid distributed across south and Southeast Asia. Severe pressures from habitat loss, prey
 26 depletion, disease, human persecution and interspecific competition resulted in global population
 27 decline in dholes. Despite a declining population trend, detailed information on population size,
 28 ecology, demography and genetics is lacking. Generating reliable information at landscape level
 29 for dholes is challenging due to their secretive behaviour and monomorphic physical features.
 30 Recent advances in non-invasive DNA-based tools can be used to monitor populations and
 31 individuals across large landscapes. In this paper, we describe standardization and validation of
 32 faecal DNA-based methods for individual identification of dholes. We tested this method on 249
 33 field-collected dhole faeces from five protected areas of the central Indian landscape in the state
 34 of Maharashtra, India.

35 ***Results***

36 We tested a total of 18 cross-species markers and developed a panel of 12 markers for
 37 unambiguous individual identification of dholes. This marker panel identified 101 unique
 38 individuals from faecal samples collected across our pilot field study area. These loci showed
 39 varied level of amplification success (57-88%), polymorphism (3-9 alleles), heterozygosity
 40 (0.23-0.63) and produced a cumulative misidentification rate or $PID_{(unbiased)}$ and $PID_{(sibs)}$ value of
 41 4.7×10^{-10} and 1.5×10^{-4} , respectively, indicating a high statistical power in individual
 42 discrimination from poor quality samples.

43 ***Conclusion***

44 Our results demonstrated that the selected panel of 12 microsatellite loci can conclusively
 45 identify dholes from poor quality, non-invasive biological samples and help in exploring various
 46 population parameters. This genetic approach would be useful in dhole population estimation
 47 across its range and will help in assessing population trends and other genetic parameters for this
 48 elusive, social carnivore.

49

50

51

52

53

54

55

56

57

58

59

60

61

Introduction

The Asiatic wild dog or dhole (*Cuon alpinus*) is a highly elusive, endangered, social canid distributed in south and southeast Asia (Johnsingh, 1982; Durbin, et al., 2004) occupying a range of habitat types including alpine, temperate, sub-tropical and tropical forests (Durbin et al., 2004). Driven by habitat loss, prey depletion, disease transmission from domestic dog, human persecution and interspecific competition (Hayward, Lyngdoh & Habib, 2014; Kamler et al., 2015), dholes are currently found in about 75% of their historical global range (Durbin et al., 2004; Kamler et al., 2015). Global dhole population is roughly estimated to be about 4500-10500 with only 949-2215 mature individuals, but accurate estimates and population trends are not available from any part of its range (Kamler et al., 2015). They are considered as ‘Endangered’ by IUCN under criteria C2a(i) and included in Appendix II of the Convention on International Trade in Endangered Species (CITES). The Indian subcontinent currently retains majority of the remaining dhole populations (Kamler et al., 2015) in the Western Ghats and central Indian forests (Karanth et al., 2009), along with smaller populations in the Eastern Ghats (Karanth et al., 2009), northeast India (Gopi, Lyngdoh & Selvan, 2010; Lyngdoh et al., 2014) and Himalayan region (Bashir et al., 2014). The species has faced about 60% decline in their historical distribution in the subcontinent (Karanth et al., 2010).

Given the current anthropogenic disturbance scenario across its range, the future survival of this monotypic genus depends on integrated conservation measures involving detailed, accurate

information on ecology, demography and genetics. However, generating reliable information for this elusive, forest-dwelling and pack-living canid at landscape scale is challenging. Traditional ecological techniques such as regular photographic capture approach are ineffective for dholes due to absence of unique coat patterns and their monomorphic forms, and physical tagging methods are impractical at landscape scales due to logistical difficulties, high costs and small numbers of captures possible. In this context, genetic tools have tremendous potential to generate critical information such as population size estimation (Mondol et al., 2009b), phylogeography (Luo et al., 2014; Waits et al., 1998), pack dynamics and reproductive fitness (Sillero-Zubiri, Gottelli & Macdonald, 1996; Girman et al., 1997), dispersal patterns (Epps et al., 2007; Gour et al., 2013) for elusive species conservation across large landscapes (Mondol, Bruford & Ramakrishnan, 2013). The ability to identify individuals from non-invasive samples collected over large space provides a feasible option to generate detailed information on elusive, forest-dwelling dholes as they cannot be identified using other approaches.

In this study, we addressed key methodological issues related to selection and standardization of a set of molecular markers for individual identification of dholes. Subsequently, we tested these markers on field-collected dhole samples from five protected areas of the central Indian landscape in the state of Maharashtra, India for individual identification. In addition to the utilization of these markers in dhole population estimation at landscape level, we believe that this approach has wider relevance in non-invasive, faecal DNA based population assessments of many other low density, elusive, wide-ranging species.

Methods

Research permits and ethical considerations

All required permissions for fieldwork and sampling were provided by the Maharashtra Forest Department (Permit No. 09/2016). The entire study was non-invasive through field-collected faecal samples, and thus did not require any ethical clearance from the institute. Reference dhole blood samples (n=4) were collected as part of another ongoing study in Tadoba-Andhari Tiger Reserve (Permit no. SPP-12/2016), where blood sampling was conducted during radio collaring of dholes.

Study Area

The study was focused in five protected areas Melghat Tiger Reserve (MTR), Pench Tiger Reserve (PTR), Navegaon-Nagzira Tiger Reserve (NNTR), Tadoba-Andhari Tiger Reserve (TATR) and Umred-Karandhla Wildlife Sanctuary (UKWLS) of the central Indian landscape in the state of Maharashtra, India. The entire area is a complex of forested areas (core zone) with different levels of connectivity. NNTR and PTR are geographically closer as compared to MTR-NNTR and MTR-PTR. MTR and PTR are part of the Satpura-Maikal-Pench corridor in the Satpura-Maikal landscape. The forest type is of dry deciduous to moist deciduous nature with major vegetation consisting of *Tectona grandis*, *Anogeissus latifolia*, *Lagerstroemia parviflora*, *Terminalia spp.*, *Heteropogon contortus*, *Themeda quadrivalvis*, *Cynodon dactylon* etc.

Field Sampling

Dholes prefer dense forested habitats (Johnsingh, 1985) where the social groups defecate in communal latrine sites mostly found on the junctions of roads/trails (Johnsingh, 1982). Their elusive nature and highly social behaviour present unique challenges in scat sampling for individual identification. In this study, sampling was conducted through intensive foot and vehicle surveys covering the entire study area. We sampled a total of 49 latrine sites covering five protected areas. Once a latrine site was found only fresh scats were targeted for collection.

One bolus from each fresh scat was collected assuming it to derive from one individual. Separate gloves were used to collect each sample. All samples were collected directly in wax paper and stored in separate ziplock bags. Once brought to the field station, the sample containing ziplock bags were temporarily stored in a large box containing silica gel to minimize fungal growth in humid conditions. Samples were then shipped to the laboratory, where they were stored in a -20°C freezer. GPS co-ordinates and other associated information (track marks, substrata etc.) were collected for each sample. Entire sampling was conducted once per site between January 2015 – June 2017 ensuring maximum coverage of the study area, covering PTR (257.3 km²), MTR (1500.49 km²), NNTR (152.8 km²), TATR (627.5 km²) and UKWLS (189 km²), Maharashtra. A total of 249 samples were collected (PTR- 92, MTR- 76, NNTR- 37, UKWLS- 34, TATR- 10, respectively) for this study. Details of all sample locations are given in Figure 1. For blood sampling, four animals (3 males and 1 female) were remotely administered with reconstituted lyophilized mixture of Telatamine-Zolezepam (Zoletil 100, Virbac, France) at the dose rate of 8.6 mg kg⁻¹ body weight using Dan-inject syringe projector (Model-IM) at a distance of 15-25 meters. After ensuring sedation safe for handling, animal was approached, blindfolded and 1 ml of blood was collected through femoral-saphenous vein puncture. Blood was collected in EDTA vacutainers and preserved at -20°C in the laboratory for genetic work.

DNA extraction

DNA extraction was performed in the laboratory from the frozen faecal samples using QIAamp DNA Tissue Kit (QIAGEN Inc., Hilden, Germany) with a modified approach, depending on sample quality. If the sample had the entire top mucous layer available (i.e. not covered by dust, soil etc.) then it was swabbed with phosphate buffer saline soaked sterile cotton swab and was stored in sterile Eppendorf tube at -20°C (Biswas et al., 2019). However, if the mucous layer was

covered then the top layer was scraped using sterile blades and stored in similar conditions (Biswas et al., 2019). Subsequently, faecal samples collected by both methods were lysed overnight in 300/600 µl of lysis buffer for swabs and scraped samples, respectively and 20 µl proteinase K followed by extraction using the kit's protocol. DNA was eluted twice with 100 µl of 1X TE and stored in -20°C for long-term use. Each set of 22 extractions was accompanied with two negative controls to monitor possible contamination.

DNA from blood samples was extracted using standard protocol given in the QIAamp DNA Tissue Kit (QIAGEN Inc., Hilden, Germany). Negative control was incorporated to monitor any possible contaminations.

Selection of microsatellite markers

There are no dhole specific microsatellites developed so far and the only study focusing on dhole population genetics had used 13 cross-species markers from domestic dogs to study genetic variation (Iyengar et al., 2005). These markers showed low levels of polymorphism and low PID_{sibs} value (3.3×10^{-4}), providing a misidentification rate of 1 in 3000 siblings. Given that India is considered to retain high (about 1500-3000) number of dholes (Kamler et al., 2015), this panel will not provide sufficient statistical power for unambiguous individual identification at landscape levels with large population sizes. For this study, we developed a panel following stringent cross-species marker selection and testing process. The entire process was conducted in two steps: marker selection and rigorous testing before developing a final microsatellite panel for dhole individual identification.

As most of the cross-species markers were found to be from dogs and earlier used markers were less polymorphic for individual identification, we decided to first examine if both species (domestic dog and wild dog) share genetic similarity. Earlier karyotype and chromosomal

banding studies (Graphodatsky et al., 2008) showed almost identical G-banding patterns, indicating high chromosomal level similarity between both species. Subsequently, we identified a total of 37 dog microsatellite loci from earlier published literature (Holmes et al.; Ostrander, Sprague & Rine, 1993; Fredholm & Winterø, 1995; Ostrander et al., 1995; Francisco et al., 1996; Neff et al., 1999). These markers were selected based on their polymorphism (number of alleles, PIC, observed heterozygosity etc.) and amplicon sizes in published literatures. Further, we mapped all the markers on available dog genome canFam 3.1 in UCSC Genome Browser (<http://genome.ucsc.edu/>; Accession ID: GCA_000002285.2) to assess the chromosome number to which each marker is associated with. Finally, a total of 18 microsatellites were selected based on their amplicon size, chromosome number and polymorphism (based on published data) for further testing. The details of the markers are given in Table 1.

PCR standardization and data validation

All initial standardization of the markers was conducted using dhole blood samples (n=4). PCR reactions were performed for selected 18 microsatellites in 10 µl reactions containing 3.5 µl Qiagen multiplex PCR buffer mix (QIAGEN Inc., Hilden, Germany), 0.2 µM labelled forward primer (Applied Biosystems, California, United States), 0.2 µM unlabelled reverse primer, 4 µM BSA and 2 µl of 1:50 dilution of blood DNA extract. The PCR conditions included an initial denaturation (95 °C for 15 min); 50 cycles of denaturation (94 °C for 30 s), annealing (50-60 °C gradient for 30 s) and extension (72 °C for 35 s); followed by a final extension (72 °C for 30 min). Following post-temperature standardizations markers with same annealing temperatures but with different labels or allele sizes were standardized as multiplex assays (see Table 1 for details). During all amplifications, both extraction and PCR negative controls (one PCR negative every set of 11 reactions) were included to monitor any possible contamination. Post

amplification, 2 µl of PCR product was mixed with HiDi formamide (Applied Biosystems, California, United States) and LIZ 500 size standard (Applied Biosystems, California, United States) and genotyped in an ABI genetic analyser (Applied Biosystems, California, United States). The fragment lengths were scored manually using the program GENEMARKER (Softgenetics Inc., Pennsylvania, United States). Each reaction was repeated three times to ensure good data quality.

Once the initial temperature and multiplexing standardizations were performed using reference blood DNA samples, final standardization was conducted with dhole faecal DNA. Species identification was performed for all field-collected faeces using specific mtDNA primers described in Modi et al., (2018). PCR reactions were performed with 4 µl of hotstart taq mix (QIAGEN Inc., Hilden, Germany), 4 µM BSA, 0.5 µM of primer mix and 3 µl of DNA extract with conditions including initial denaturation (95 °C for 15 min); 50 cycles of denaturation (94 °C for 30 s), annealing (50 °C for 30 s) and extension (72 °C for 35 s); followed by a final extension (72 °C for 10 min). Negative controls were included to monitor contaminations. Samples that produced species-specific bands (n=225) were further processed for microsatellite analyses.

For faecal samples, data validation was performed through a modified multiple-tube approach as described in Mondol et al., (2009b). All faeces that had amplified in 50% of the loci in the panel during first PCR were repeated two more times for all loci. Following allele calling, a consensus genotype was prepared using the ‘Quality index’ protocol (Miquel et al., 2006), during which alleles were called manually and scored as ‘1’ if the repeat is identical with the first call, or ‘0’ if calls do not match due to no amplification, allelic dropout (one allele in heterozygote is erroneously not amplified), false allele (slippage artefact during PCR) etc. To calculate the

quality index for each locus/sample the scores assigned to each repeat are summed and divided by the total number of repeats, and only quality index of 0.75 or more (at least three out of four repeats) for each locus was considered for downstream analyses. We calculated average amplification success as the percent positive PCR for each locus, as described by Broquet & Petit (2004). We quantified allelic dropout and false allele rates manually as the number of dropouts or false alleles over the total number of amplifications, respectively (Broquet & Petit, 2004), as well as using MICROCHECKER v 2.2.3. (Oosterhout et al., 2004). The false allele frequency was calculated for both homozygous and heterozygous genotypes as the ratio of the number of amplifications having one or more false alleles at a particular locus and the total number of amplifications while allele dropout rate (ADO) was calculated as the ratio between the observed number of amplifications having loss of one allele and the number of positive amplifications of the heterozygous individuals. Program FreeNA (Chapuis & Estoup, 2007) was used to determine the frequency of null alleles, which estimates the null allele frequency using EM algorithm (Dempster et al., 1977).

Molecular sexing of the identified individuals was conducted using already developed multiplex sexing approach (Modi et al., 2018), where three sex chromosome specific markers (DBY, AHT-X40 and SRY) were combined to generate a three-band pattern for males and a single band for females. This approach reduces identification of ‘false negatives’ of males due to allelic dropout from the Y chromosome from poor quality samples.

Data analyses

The identity analysis module implemented in program CERVUS (Kalinowski, Taper & Marshall, 2007) was used to identify identical genotypes (or recaptures) by comparing data from all samples for all amplified loci. All genetic recaptures were removed from the data set.

GIMLET (Valiere, 2002) was used to calculate the $PID_{(sibs)}$ for all the individuals. Following this, any allele having less than 10% frequency across all amplified loci were rechecked for allele confirmation. ARLEQUIN (Excoffier, Laval & Schneider, 2005) was used to determine Hardy Weinberg equilibrium and linkage disequilibrium for all the loci.

Results

During initial standardizations we tested all 18 selected markers (see Table 1) with four wild-caught dhole blood DNA samples. Three of these tested markers (WD2159, CPH16, AHT136) did not show any amplification in the blood DNA samples and were removed from subsequent analyses. The remaining 15 markers were then amplified with 225 genetically confirmed dhole faecal samples. Following data validation through multiple repeats, amplification success rates and polymorphism for these loci were calculated. The results show that loci CXX608 and CXX140 were monomorphic in all amplified samples, and locus CPH6 has low amplification success rate (~ 35%) from faecal DNA and thus were removed from the panel. The remaining 12 markers were finally standardized as four multiplex panels (see Table 1) for dhole individual identification.

None of these final 12 loci showed any signatures of large-scale allelic dropouts. The mean allelic dropout rate was found to be 0.1, whereas mean false allele frequency for all the 12 loci was 0.092. Overall frequency of null alleles was calculated as 0.11, indicating this 12 loci panel has low genotyping error rates. Amplification success ranged between 57% to 88% from dhole faecal DNA. The loci showed relatively higher (WD2201-9 alleles, H_o -0.63) to medium (CXX251-3 alleles, H_o -0.23) levels of polymorphism (Table 1). Except locus WD2001, none of the other loci were found to deviate from the Hardy-Weinberg equilibrium and there were no evidences for strong linkage disequilibrium between any pair of loci. Summary statistics for

various measures of polymorphism (observed and expected heterozygosity, number of alleles and allelic size range) for all loci in the final panel are presented in Table 1.

For individual identification, we only considered samples that produced good quality data for at least seven of the 12 panel loci. This cut-off value of average of seven loci was decided based on the statistical support (PID_{sibs} value of 1 in 500 siblings) produced by these loci. Given that any single largest dhole population is about 250-300 individuals (Kamler et al., 2015), this value is sufficient for individual identification at local scales in India. Out of the 225 field-collected dhole faecal samples amplified with the panel of 12 markers, 98 produced seven or more loci data. Overall, we generated genetic data from a total of 102 samples (4 blood and 98 faecal samples) (Figure 1). About 70% of these samples ($n=71$) have successfully amplified for 10-12 loci. Following analyses with CERVUS, we identified 101 unique dhole individuals from the entire data, whereas one individual from NNTR was found to be a 'genetic recapture'. Cumulative PID_{sibs} and $PID_{unbiased}$ values were found to be 1.5×10^{-4} and 4.7×10^{-10} , respectively, indicating a strong statistical support for unambiguous individual identification. The number of unique individuals from each sampled area was found to be: PTR= 33, MTR= 35, NNTR= 16, UKLWS= 9 and TATR= 4 (Figure 2). Molecular sexing showed a success rate of 67%, with a male:female sex ratio of 4:1 in all identified dhole individuals ($n=101$).

Discussion

In this paper, we standardized protocols for individual identification of Asiatic wild dogs from poor quality DNA samples, and the final marker panel could unambiguously identify individual dholes in our field-based pilot study from five protected areas of Maharashtra, India. The systematic protocols followed here offer some advantages over earlier efforts on dhole individual identification from faecal samples by Iyengar et al. (2005). Firstly, use of a large panel of 37

microsatellite loci for preliminary assessment of marker suitability along with genomic mapping-based selection of final markers ($n=18$) allowed us to ascertain a combination of loci for unambiguous individual identification with high statistical power. The rigorous testing of the loci with large number of DNA samples from different sources also allowed us to exclude loci that might be problematic due to low amplification success from non-invasive samples. The final panel consisting 12 markers were further standardized into four multiplex reactions to provide time and cost-effective options during data generation. We were very careful to initially select a large number of tetranucleotide markers as they are known to have low stutter peak problems and better allele characteristics from poor quality samples (Walsh, Fildes & Reynolds, 1996), while dinucleotide markers generally have higher amplification success (Broquet, Menard & Petit, 2007). Thus, our final panel with a ratio of 2:1 tetra vs. dinucleotide microsatellites would provide the ideal combination in terms of high success rate and less technical issues in allele calling during dhole individual identification. The amplification success rate for all loci was > 70% except locus PEZ5 (~ 60%), but it was found to be polymorphic and was included in the panel. The overall genotyping error frequency was found to be < 0.2 from dhole faeces, which is within the recommended limits for non-invasive population genetic research (Smith & Wang, 2014).

Our motivation in this study was to develop effective protocols that could be applied for individual identification of Asiatic wild dogs as they are difficult to identify from physical characteristics (spots, marks, stripes etc.). Their elusive nature also makes it challenging to estimate population size using traditional techniques (photographic capture, field-based observations etc.) at landscape levels. For genetic estimation of population size Waits et al., (2001) recommended a threshold PID_{sibs} value that is at least double than the approximate

number of animals in any given area. The cumulative PID_{sibs} value of 1.5×10^{-4} achieved in this study is better than Iyengar et al., (2005) (PID_{sibs} of 3.3×10^{-4}) and should be sufficient to study dhole genetics and specifically population estimation across its range. Among all the dhole range countries India is considered to retain the highest (about 1500-3000) number of individuals (Kamler et al., 2015) and our misidentification rate achieved in this study (1 in 6700 siblings) would provide strong statistical power in individual identification. The most recent assessment suggests that the largest dhole population in Western Ghats, India holds about 207-304 individuals (Kamler et al., 2015), thereby assuring that our seven loci cut-off (misidentification rate 1 in 500 siblings) to select samples and 12 loci panel would be useful in population estimation at local scale. However, it is important to point out that we generated individual level information from about 43.5% (98 out of 225 faeces) of the field-collected samples in this study. Similar patterns of low amplification success rate from field-collected faecal samples have been observed in earlier genetic study of dhole (Iyengar et al., 2005), leopard (Mondol et al., 2009) and other species (Smith & Wang, 2014). Considering dhole cryptic nature, social behaviour and ecology in corroboration with low amplification success rate, we suggest an intensive faecal sampling effort for estimation of population size for this species. It is also noteworthy to point out that literature survey for dhole marker selection in this study was mostly based on available information on non-invasive canid population genetic research with specific information available on markers such as marker polymorphism (number of alleles, observed heterozygosity) and amplicon size etc. (Holmes et al. 1995; Ostrander, Sprague & Rine, 1993; Fredholm & Winterø, 1995; Ostrander et al., 1995; Francisco et al., 1996; Neff et al., 1999). However, future studies should also consider additional markers those are tested as part canid forensic studies (for example see van Asch et al., 2009; Berger et al., 2004; Hellmann et al., 2006; Eichmann, Berger

& Parson, 2004) on dholes. In addition, already available dhole genome information (Campana et al., 2016; Habib et al., 2018) also can be used to develop a suitable SNP panel for in depth analyses of dhole population and demographic parameters.

During individual identification we had identified only one genetic recapture from the field-collected faecal samples. This pattern of low dhole recapture could be attributed to our sampling strategy to cover large geographical area and maximize collection of faeces from potentially different individuals, as well as relatively low amplification success rates from faecal samples. We have sampled the entire study area only once and focused on collecting fresh samples, thereby probably missed recapturing the same individuals from same latrine sites. Further, low amplification success rate (101 genotypes from 225 fresh samples) might have resulted in getting lesser number of recaptures. We got a male biased sex-ratio (4:1) in this pilot study. While there is no conclusive information on dhole sex ratio across its range, our earlier study (Modi et al. 2018) in the same landscape has shown a male biased (M: F ratio of 3:1) sex ratio, and ecological study in southern India by Venkataraman (1998) suggested male biased packs. Future studies with extensive genetic sampling across this landscape would potentially provide more accurate sex ratio for dholes.

Conclusion

In the broader context of understanding dhole population dynamics at local or landscape scales, genetic sampling is possibly the only way to generate information with spatial and temporal coverage for this elusive, social carnivore as photographic sampling or conventional tagging cannot be employed due to lack of distinguishing natural marks and logistical difficulties of physical captures of large number of animals. Results from this study provide a robust tool to generate individual level information from field-collected faecal samples. In combination with a

good sampling strategy, our methods can be used in a cost-effective way to investigate species biology (including patterns of genetic diversity, relatedness and population connectivity) as well as to estimate population abundance of dholes in the wild.

Acknowledgement

We acknowledge the Maharashtra State Forest Department for providing research permission, and the forest officials and staff for their assistance in fieldwork. We thank the Dean, Faculty of Wildlife Sciences and Director, WII for their continuous support in this research. We thank A. Madhanraj, Tanima and Aditi for their help in the laboratory and Shaheer for his help with GIS. We sincerely thank the editor and the two reviewers for their comments to improve the earlier version of the manuscript.

References

- Bashir T, Bhattacharya T, Poudyal K, Roy M, Sathyakumar S. 2014. Precarious status of the endangered dhole *cuon alpinus* in the high elevation eastern himalayan habitats of khangchendzonga biosphere reserve, Sikkim, India. *Oryx* 48:125–132. DOI: 10.1017/S003060531200049X.
- Berger B, Berger C, Hecht W, Hellmann A, Rohleder U, Schleenbecker U, Parson W. 2014. Validation of two canine STR multiplex-assays following the ISFG recommendations for non-human DNA analysis. *Forensic Science International: Genetics* 8(1):90-100.
- Biswas S, Bhatt S, Paul S, Modi S, Ghosh T, Habib B, Nigam P, Talukdar G, Pandav B & Mondol S. 2019. A practice faeces collection protocol for multidisciplinary research in wildlife science. *Current Science (in press)*. Available at <https://www.currentscience.ac.in/php/forthcoming/2019/30500.pdf>

- Broquet T, Ménard N, Petit E. 2007. Noninvasive population genetics: A review of sample source, diet, fragment length and microsatellite motif effects on amplification success and genotyping error rates. *Conservation Genetics* 8:249–260. DOI: 10.1007/s10592-006-9146-5.
- Broquet T, Petit E. 2004. Quantifying genotyping errors in noninvasive population genetics. *Molecular Ecology* 13:3601–3608. DOI: 10.1111/j.1365-294X.2004.02352.x.
- Campana MG, Parker LD, Hawkins MT, Young HS, Helgen KM, Gunther MS, Woodroffe R, Maldonado JE, Fleischer RC. Genome sequence, population history, and pelage genetics of the endangered African wild dog (*Lycaon pictus*). 2016. *BMC genomics* 17(1):1013.
- Chapuis MP, Estoup A. 2007. Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* 24:621–631. DOI: 10.1093/molbev/msl191.
- Durbin L, Venkataraman A, Hedges S, & Duckworth JW. Dhole (*Cuon alpinus*) Status survey and conservation action plan. In *Canids: Foxes, Wolves, Jackals and Dogs*. 2004. p 210–219.
- Eichmann C, Berger B & Parson W. A proposed nomenclature for 15 canine-specific polymorphic STR loci for forensic purposes. 2004. *International journal of legal medicine* 118(5):249-66.
- Epps CW, Wehausen JD, Bleich VC, Torres SG, Brashares JS. 2007. Optimizing dispersal and corridor models using landscape genetics. *Journal of Applied Ecology* 44:714–724. DOI:

10.1111/j.1365-2664.2007.01325.x.

Excoffier L, Laval G, Schneider S. 2005. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics* 1:117693430500100. DOI: 10.1177/117693430500100003.

Francisco LV, Langston AA, Mellersh CS, Neal CL, Ostrander EA. 1996. A class of highly polymorphic tetranucleotide repeats for canine genetic mapping. *Mammalian Genome* 7:359–362. DOI: 10.1007/s003359900104.

François O, Ancelet S, Guillot G. 2006. Bayesian clustering using hidden Markov random fields in spatial population genetics. *Genetics* 174:805–16. DOI: 10.1534/genetics.106.059923.

Fredholm M, Winterø a K. 1995. Variation of short tandem repeats within and between species belonging to the Canidae family. *Mammalian Genome* 6:11–18. DOI: 10.1007/BF00350887.

Girman DJ, Mills MGL, Geffen E, Wayne RK, Ecology SB, Wayne RK. 1997. A Molecular Genetic Analysis of Social Structure , Dispersal , and Interpack Relationships of the African Wild Dog (*Lycaon pictus*). *Behavioral Ecology and Sociobiology* 40(3):187–198.

Gopi GV, Lyngdoh S, Selvan KM. 2010. Conserving the endangered Asiatic wild dog *Cuon alpinus* in Western Arunachal Pradesh : Fostering Better Coexistence For Conservation. *Final Technical Report Submitted to Rufford Small Grant Program, UK. 54 pp.*

Gour DS, Bhagavatula J, Bhavanishankar M, Reddy PA, Gupta JA, Sarkar MS, Hussain SM, Harika S, Gulia R, Shivaji S. 2013. Philopatry and Dispersal Patterns in Tiger (*Panthera tigris*). *PLoS ONE* 8:14–17. DOI: 10.1371/journal.pone.0066956.

- Graphodatsky AS, Perelman PL, Sokolovskaya NV, Beklemisheva VR, Serdukova NA, Dobigny G, O'Brien SJ, Ferguson-Smith MA, Yang F. 2008. Phylogenomics of the dog and fox family (Canidae, Carnivora) revealed by chromosome painting. *Chromosome Research* 16:129–143. DOI: 10.1007/s10577-007-1203-5.
- Habib B, Ghaskadbi P, Nigam P, Modi S, Kumar PS, Sharma K, Singh V, Kumar B, Tripathi A, Kothandaraman H, Yellaboina S, Baghel DS, Mondol S. 2018. The Cuon Enigma: Genome survey and comparative genomics of the endangered Dhole (*Cuon alpinus*). *bioRxiv*. 2018 Jan 1:443119.
- Hayward MW, Lyngdoh S, Habib B. 2014. Diet and prey preferences of dholes (*Cuon alpinus*): Dietary competition within Asia's apex predator guild. *Journal of Zoology* 294:255–266. DOI: 10.1111/jzo.12171.
- Hellmann AP, Rohleder U, Eichmann C, Pfeiffer I, Parson W, Schleenbecker U. 2006. A proposal for standardization in forensic canine DNA typing: allele nomenclature of six canine-specific STR loci. *Journal of forensic sciences* 51(2):274-81
- Holmes NG, Dickens HF, Parker HL, Binns MM, Sampson J. 1995. Eighteen canine microsatellites. *Animal Genetics* 126(2):132–3.
- Iyengar A, Babu VN, Hedges S, Venkataraman AB, Maclean N, Morin PA. 2005. Phylogeography, genetic structure, and diversity in the dhole (*Cuon alpinus*). *Molecular Ecology* 14:2281–2297. DOI: 10.1111/j.1365-294X.2005.02582.x.
- Johnsingh AJT. 1982. Reproduction and social behaviour of the dhole, *Cuon alpinus* (Canidae). *Journal of Zoology* 198:443–463. DOI: 10.1111/jzo.1982.198.4.443.

Johnsingh AJT. 1985. Distribution and status of dhole *Cuon alpinus* Pallas, 1811. *Mammalia* 49:203–208. DOI: 10.1515/mamm.1985.49.2.203.

Kalinowski ST, Taper ML, Marshall TC. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* 16:1099–1106. DOI: 10.1111/j.1365-294X.2007.03089.x.

Kamler JF, Songsasen N, Jenks K, Srivathsa A, Sheng L, Kunkel K. *Cuon alpinus*. 2015. The IUCN Red List of Threatened Species. e.T5953A72477893. <http://dx.doi.org/10.2305/IUCN.UK.2015-4.RLTS.T5953A72477893.en>.

Karanth KK, Nichols JD, Hines JE, Karanth KU, Christensen NL. 2009. Patterns and determinants of mammal species occurrence in India. *Journal of Applied Ecology* 46:1189–1200. DOI: 10.1111/j.1365-2664.2009.01710.x.

Karanth KK, Nichols JD, Karanth KU, Hines JE, Christensen NL. 2010. The shrinking ark: Patterns of large mammal extinctions in India. *Proceedings of the Royal Society B: Biological Sciences* 277:1971–1979. DOI: 10.1098/rspb.2010.0171.

Luo SJ, Zhang Y, Johnson WE, Miao L, Martelli P, Antunes A, Smith JLD, O’Brien SJ. 2014. Sympatric Asian felid phylogeography reveals a major Indochinese-Sundaic divergence. *Molecular Ecology* 23:2072–2092. DOI: 10.1111/mec.12716.

Lyngdoh S, Gopi GV, Selvan KM, Habib B. 2014. Effect of interactions among ethnic communities, livestock and wild dogs (*Cuon alpinus*) in Arunachal Pradesh, India. *European Journal of Wildlife Research* 60:771–780. DOI: 10.1007/s10344-014-0846-8.

Miquel C, Bellemain E, Poillot C, Bessi re J, Durand A, Taberlet P. 2006. Quality indexes to

assess the reliability of genotypes in studies using noninvasive sampling and multiple-tube approach. *Molecular Ecology Notes* 6:985–988. DOI: 10.1111/j.1471-8286.2006.01413.x.

Modi S, Mondol S, Ghaskadbi P, Hussain Z, Nigam P, Habib B. 2018. Non-invasive DNA-based species and sex identification of Asiatic wild dog (*Cuon alpinus*). *Journal of Genetics* (In press).

Mondol S, Bruford MW, Ramakrishnan U. 2013. Demographic loss, genetic structure and the conservation implications for Indian tigers. *Proceedings of the Royal Society B: Biological Sciences* 280:20130496–20130496. DOI: 10.1098/rspb.2013.0496.

Mondol S, Karanth KU, Ramakrishnan U. 2009. Why the Indian subcontinent holds the key to global tiger recovery. *PLoS Genetics* 5. DOI: 10.1371/journal.pgen.1000585.

Mondol S, Navya R, Athreya V, Sunagar K, Selvaraj VM, Ramakrishnan U. 2009a. A panel of microsatellites to individually identify leopards and its application to leopard monitoring in human dominated landscapes. *BMC Genetics* 10. DOI: 10.1186/1471-2156-10-79.

Mondol S, Karanth KU, Kumar NS, Gopalaswamy AM, Andheria A, Ramakrishnan U. 2009b. Evaluation of non-invasive genetic sampling methods for estimating tiger population size. *Biological Conservation* 142:2350–2360. DOI: 10.1016/j.biocon.2009.05.014.

Neff MW, Broman KW, Mellersh CS, Ray K, Acland GM, Aguirre GD, Ziegler JS, Ostrander EA, Rine J. 1999. A second-generation genetic linkage map of the domestic dog, *Canis familiaris*. *Genetics* 151:803–820. DOI: 10.1073/pnas.95.6.3048.

Oosterhout CV, Hutchinson WF, Wills DPM, Shipley P. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology*

- Notes 4:535–538. DOI: 10.1111/j.1471-8286.2004.00684.x.
- Ostrander EA, Mapa FA, Yee M, Rine J. 1995. One hundred and one new simple sequence repeat-based markers for the canine genome. *Mammalian Genome* 6:192–195. DOI: 10.1007/BF00293011.
- Ostrander EA, Sprague GF, Rine J. 1993. Identification and characterization of dinucleotide repeat (CA)(n) markers for genetic mapping in dog. *Genomics* 16:207–213. DOI: 10.1006/geno.1993.1160.
- Sillero-Zubiri C, Gottelli D, Macdonald DW. 1996. Male philopatry, extra-pack copulations and inbreeding avoidance in Ethiopian wolves (*Canis simensis*). *Behavioral Ecology and Sociobiology* 38 : 331-340.
- Smith O, Wang J. 2014. When can noninvasive samples provide sufficient information in conservation genetics studies? *Molecular Ecology Resources* 14:1011–1023. DOI: 10.1111/1755-0998.12250.
- Valiere N. 2002 GIMLET: a computer program for analysing genetic individual identification data. *Molecular Ecology Notes* 2(3):377-379. doi.org/10.1046/j.1471-8286.2002.00228.x-i2
- van Asch B, Alves C, Gusmão L, Pereira V, Pereira F, Amorim A. (2009). A new autosomal STR nineplex for canine identification and parentage testing. *Electrophoresis* 30(2):417-423.
- Venkataraman AB. 1998. Male-biased adult sex ratios and their significance for cooperative breeding in dhole, *Cuon alpinus*, packs. *Ethology* 104(8):671-84.
- Waits LP, Talbot SL, Ward RH, Shields GF. 1998. Mitochondrial DNA Phylogeography of the North American Brown Bear and Implications for Conservation. *Conservation Biology*

12:408–17.

Waits LP, Luikart G, Taberlet P. 2001. Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Molecular Ecology* 10(1):249-256.

Walsh PS., Fildes NJ., Reynolds R. 1996. Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. *Nucleic Acids Research* 24:2807–2812.

DOI: 10.1093/nar/24.14.2807.

Figure legend

Figure 1: Protected area map with locations of unique genotypes identified in the study area of Maharashtra, India. The study area names are as following: MTR- Melghat Tiger Reserve, PTR- Pench Tiger Reserve, TATR- Tadoba-Andhari Tiger Reserve, NNTR- Navegaon-Nagzira Tiger Reserve, UKWLS- Umred-Karandhla Wildlife Sanctuary.

Figure 2: Graph showing the number of confirmed dhole scats collected from the field and area-wise unique genotypes identified from them. In addition to faecal-based individual genotypes (n=97), four blood samples were collected from TATR, making the total unique genotype of 101 in this study.

Figure 1

Protected area map with locations of unique genotypes identified in the study area of Maharashtra, India.

The study area names are as folloing: MTR- Melghat Tiger Reserve, PTR- Pench Tiger Reserve, TATR- Tadoba-Andhari Tiger Reserve, NNTR- Navegaon-Nagzira Tiger Reserve, UKWLS- Umred-Karandhla Wildlife Sanctuary.

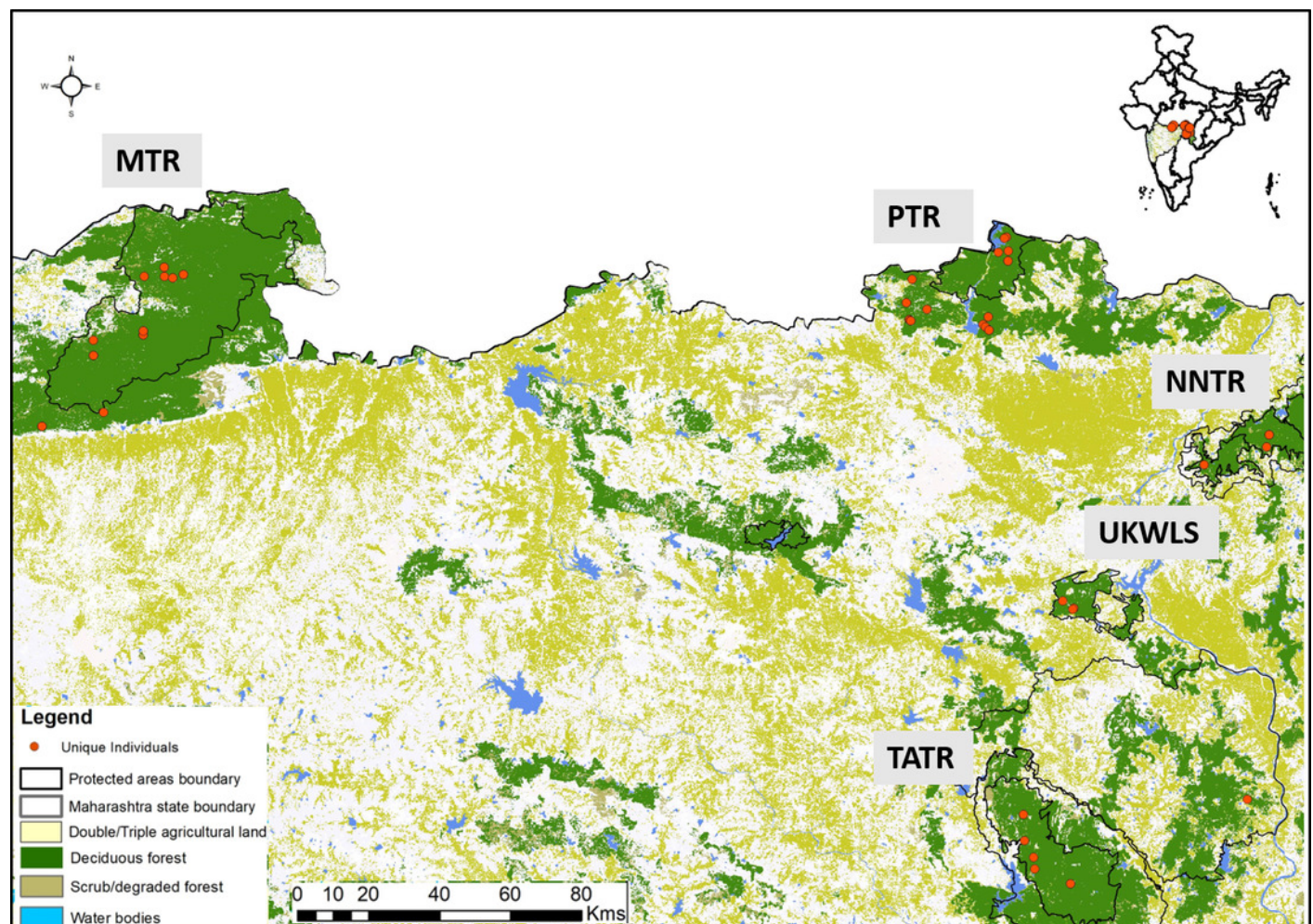


Figure 2

Graph showing the number of confirmed dhole scats collected from the field and area-wise unique genotypes identified from them.

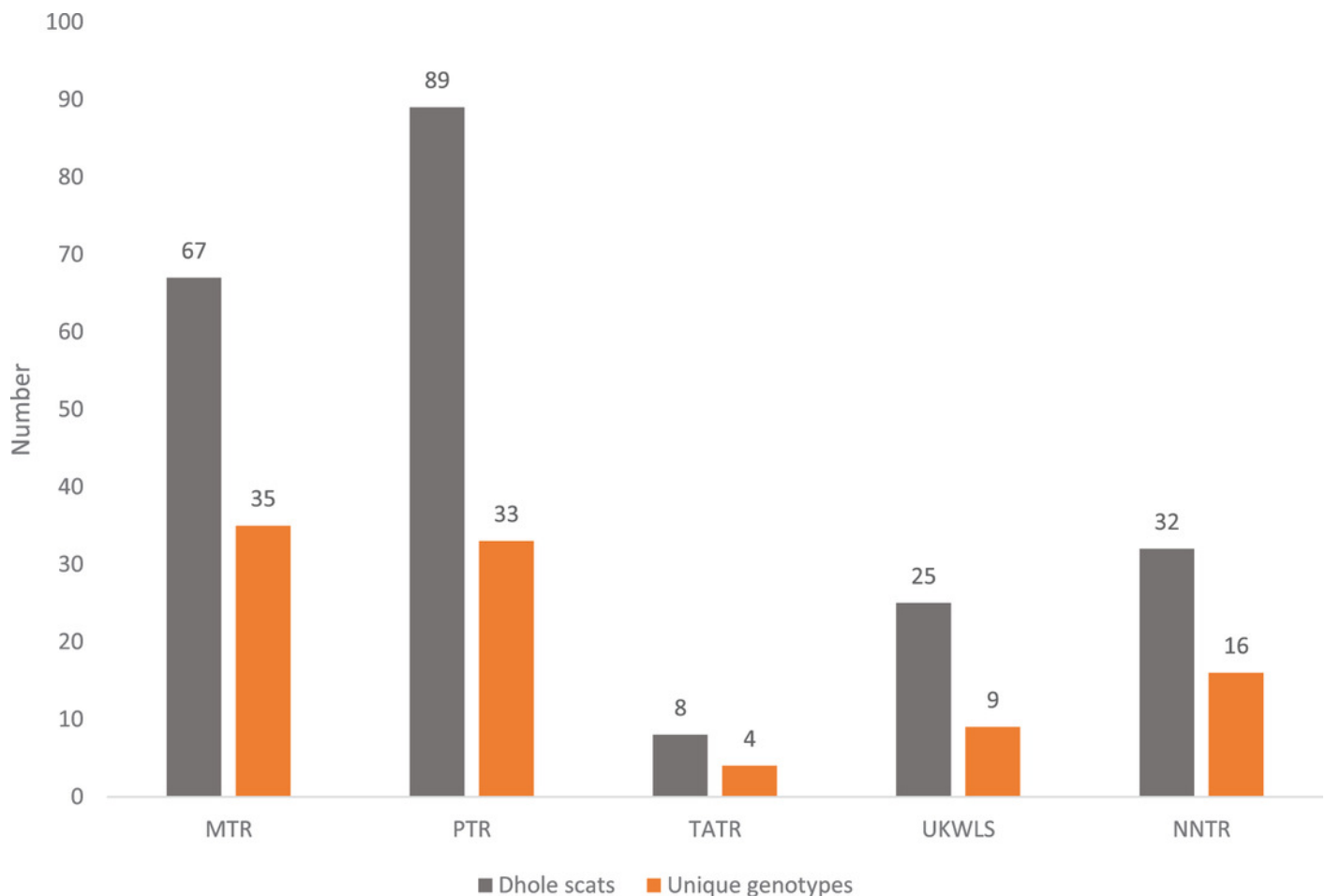


Table 1 (on next page)

Details of microsatellite loci used for Asiatic wild dog.

Table 1: Details of microsatellite loci used for Asiatic wild dog.

| Locus | Primer sequence 5'-3' | Repeat motif | Dye | T _a | Chromosome number (in dog) | Na | Ho | He | Allelic Range | PID _(unbiased) (cumulative) | PID _(sibs) (cumulative) | Success rate (%) | Genotyping error (%) | | | Reference |
|-------------------------------|---|--------------|-----|----------------|----------------------------|----|------|------|---------------|--|------------------------------------|------------------|----------------------|------|------|------------------------|
| | | | | | | | | | | | | | ADO | FA | NA | |
| WD2201 (Panel 4) | ATCAACAATGCATGCCACAT GAGAACAAATAAATGCAAGCCC | Tetra | FAM | 59°C | 7 | 9 | 0.63 | 0.78 | 170-202 | 7.40E-02 | 3.81E-01 | 73.97 | 10.4 | 6.9 | 3.5 | Francisco et al (1996) |
| PEZ6 (Panel 3) | ATGAGCACTGGGTGTTATAC ACACAATTGCATTGTCAAAAC | Tetra | NED | 53°C | 27 | 7 | 0.38 | 0.78 | 206-230 | 5.88E-03 | 1.46E-01 | 72.05 | 9.1 | 8.2 | 19.2 | Neff et al (1998) |
| WD2140 [#] (Panel 2) | GGGGAAGCCATTTTTAAAGC TGACCCTCTGGCATCTAGGA | Tetra | HEX | 56°C | 5 | 9 | 0.4 | 0.76 | 122-178 | 5.00E-04 | 5.69E-02 | 69.65 | 10.2 | 7.9 | 4.8 | Francisco et al (1996) |
| AHT130 [#] (Panel 2) | CCTCTCCTGGTAAGTGCTGC TGGAACACTGGTCCCCAG | Di | FAM | 56°C | 18 | 7 | 0.44 | 0.74 | 98-112 | 4.69E-05 | 2.29E-02 | 71.47 | 13.9 | 2.4 | 0 | Holmes et al (1994) |
| PEZ3 (Panel 3) | CACTTCTCATACCCAGACTC CAATATGTCAACTATACTTC | Tetra | PET | 53°C | 19 | 8 | 0.52 | 0.73 | 110-146 | 5.26E-06 | 9.53E-03 | 88.45 | 11 | 16.5 | 16.4 | Neff et al (1998) |
| WD2137 (Panel 4) | GCAGTCCCTTATTCCAACATG CCCCAAGTTTTGCATCTGTT | Tetra | FAM | 56°C | 3 | 7 | 0.46 | 0.66 | 156-180 | 7.29E-07 | 4.33E-03 | 68.94 | 9 | 1.6 | 22.1 | Francisco et al (1996) |
| PEZ5 (Panel 3) | GCTATCTTGTTTCCACAGC TCACTGTATACAACATTGTC | Tetra | FAM | 53°C | 12 | 8 | 0.15 | 0.61 | 150-254 | 1.39E-07 | 2.14E-03 | 57.37 | 12.9 | 6.9 | 16 | Neff et al (1998) |
| CXX251 (Panel 2) | TACCACTGTCATTTTCCATGC AAGAGGATACCGGTGGCAG | Di | NED | 56°C | 1 | 3 | 0.23 | 0.58 | 128-136 | 3.32E-08 | 1.11E-03 | 78.99 | 8.8 | 5.6 | 11 | Ostrander (1995) |
| WD2096 [#] (Panel 1) | CCGTCTAAGAGCCTCCAG GACAAGGTTTCCTGGTTCCA | Tetra | FAM | 59°C | 11 | 3 | 0.66 | 0.51 | 93-101 | 1.05E-08 | 6.38E-04 | 81.98 | 8.6 | 8.5 | 24 | Francisco et al (1996) |
| CXX279 [#] (Panel 1) | TGCTCAATGAAATAAGCCAGG GGCGACCTTCATTCTCTGAC | Di | PET | 59°C | 22 | 5 | 0.2 | 0.45 | 125-135 | 3.29E-09 | 2.33E-04 | 84.05 | 7.0 | 6.9 | 13 | Ostrander (1995) |
| WD2001 [#] (Panel 1) | TCCTCCTCTTCTTTCCATTG TGAACAGAGTTAAGGATAGACAC | Tetra | HEX | 59°C | 23 | 3 | 0.41 | 0.48 | 134-142 | 1.24E-09 | 3.86E-04 | 78.92 | 10.4 | 8.2 | 6.2 | Francisco et al (1996) |
| CXX30 (Panel 1) | GCCTTTTAGGGAGCTTTCTTT GAGTCTGCTTTTCTCCTTCCC | Di | NED | 59°C | 2 | 3 | 0.38 | 0.41 | 122-142 | 4.68E-10 | 1.50E-04 | 77.28 | 9.4 | 9.9 | 3.3 | Ostrander (1995) |
| Mean | | | | | | 6 | 0.40 | 0.62 | | | | 75.26 | 10.1 | 9.2 | 11.2 | |
| | | | | | | | | | | | | | | | | |

| | | | | | | | | | | | | | | |
|----------|--|-------|-----|------|----|---|-------|-------|---------|----------|----------|---|---|---------------------------------------|
| AHT136* | GAGAGGGCTGGTGGTAGGGG CGTGGCTATCTTTGGAGGGA | Di | HEX | NA | 11 | - | - | - | - | - | - | - | - | Neff et al.1998, Holmes et al.1994 |
| WD2159* | GAATCCCACATCGGGCTC ATTAAGTTTGGAAAGCCAGGTAAG | Tetra | HEX | NA | 24 | - | - | - | - | - | - | - | - | Francisco et al (1996) |
| CPH6* | CATTGGCTGTTTGA CTCTAGG ACTGATGTGGGTGTCTCTGC | Di | FAM | 56°C | 23 | 4 | 0.189 | 0.577 | 107-136 | - | - | - | - | Fredholm et al (1995) |
| CPH16* | CTACACCAGTTAGGGAATCTAGC CAGATTCAAATCCACTCTCAGAC | Di | HEX | NA | 20 | - | - | - | - | - | - | - | - | Fredholm et al (1995) |
| CXX140* | CAGAGGTGGCATAGGGTGAT TCGAAGCCCAGAGAATGACT | Di | PET | 56°C | 4 | 2 | 0.012 | 0.059 | 149-151 | 7.42E-08 | 9.87E-04 | - | - | Ostrander (1995) |
| CXX608*# | TATTGTAAGTCTTCCTTGAC TCTACCGTCTACAACAAAAGGG | Di | HEX | 53°C | 15 | 2 | 0 | 0.04 | 134 | 7.42E-08 | 9.87E-04 | - | - | Ostrander (1995) |

* Loci omitted from final analyses

Loci used in the study by Iyengar et al.

Na=number of alleles, H_o= Observed heterozygosity, H_e= Expected heterozygosity, PID= Probability of identity, T_a- Annealing temperature, ADO- Allele Drop-out, FA- False Alleles, NA- Null alleles