Characterization of six cross-species microsatellite markers suitable for estimating the population parameters of the black-backed jackal (*Canis mesomelas*) using a non-invasive genetic recovery protocol

Robert S. James, Phillip L. James, Dawn M. Scott and Andrew D.J. Overall

*Cogent Biology* (2015), 1: 1108479
EVOLUTIONARY BIOLOGY & MOLECULAR ECOLOGY | RESEARCH ARTICLE

Characterization of six cross-species microsatellite markers suitable for estimating the population parameters of the black-backed jackal (Canis mesomelas) using a non-invasive genetic recovery protocol

Robert S. James1*, Phillip L. James2, Dawn M. Scott1 and Andrew D.J. Overall1

Abstract: The black-backed jackal has proven difficult to study in the wild due to its elusive behaviour. As a rabies vector and predator of livestock, understanding dispersal patterns and population structure is of importance when considering the management of this species in South Africa's rural areas. We have characterised a suite of six cross-species microsatellite markers suitable for this purpose and present a method to isolate suitable quantities of host DNA from faeces in order to undertake endpoint PCR. Allelic dropout and null allele frequencies were shown to be negligible for all loci. No evidence of selection could be detected at any of the loci examined. The genotype frequencies of the total sampled population were tested for concordance with Hardy–Weinberg equilibrium. No significant deficit of heterozygote individuals was detected globally when tested by locus and marker set. An excess of heterozygosity was recorded at three loci within a single study site. The

ABOUT THE AUTHORS

Robert S. James recently completed his PhD at the University of Brighton, researching the population dynamics of the black-backed jackal in game farm ecosystems of South Africa. His research interests include the ecology and population genetics of carnivore species within agricultural landscapes.

Phillip L. James is a research fellow at the Imperial College London and the Royal Brompton Biomedical Research Unit. He is a microbial ecologist studying the micro-organisms residing within and upon mammals and their environments.

Dawn M. Scott is a mammalian biologist with an interest in applied conservation biology. Her principle areas of research are focused on the interactions between humans and wildlife and the consequential effects on mammalian communities.

Andrew D.J. Overall has been involved in various research projects including the molecular ecology of the grey seal, Soay sheep and the hazel dormice. His particular interest is in the roles inbreeding and population substructure play in the genetic health of populations.

PUBLIC INTEREST STATEMENT

The ability to describe a group of individuals in terms of their genetic composition provides a means to measure the degree of isolation between populations in the wild. By comparing the genetic diversity within and between individuals and groups, information about the migration, dispersal and breeding structure of a population can be deduced. While this information is highly relevant to the fields of ecology, evolution and conservation, the success of such studies is often reliant on both the availability of genetic markers and the effort required to sample a wild population. In this study, we developed and tested six genetic markers for the black-backed jackal, a common disease vector and conflict species in South Africa, and presented a sampling method that is capable of isolating DNA from the faecal deposits of this species. This has enabled us to identify the genetic composition of a wild population of black-backed jackals.
total inbreeding coefficient $F_{st}$ was calculated at 3% across the sampled population. The average DNA copy number recovered from faecal deposits was quantified using a qPCR-TaqMan reporter probe method. The results of this study indicate that the quantity of DNA recovered from faeces was sufficient to undertake endpoint PCR.

**Subjects: Animal Ecology; Genetics; Zoology**

**Keywords: microsatellite; black-backed jackal; qPCR; faeces; non-invasive genetic sampling**

1. Introduction

The analysis of populations in the natural environment is a focal point for many ecological and conservation-based studies. The ability to describe a population in terms of genetic isolation and breeding structure provides biologically relevant information that can be applied to both species management and conservation programmes in the wild. Simple sequence repeat polymorphisms, also known as microsatellites, have been extensively used in this regard with previous studies aiming to quantify the amount of inbreeding in endangered species (Cain et al., 2014; Coulson, Albon, Slate, & Pemberton, 1999; Curik et al., 2003); monitor reintroduction programs (Garcia-Moreno, Matoqa, Roy, Geffen, & Wayne, 1996; Storfer, Mech, Reudink, & Lew, 2014; Vernesi et al., 2003), detect genetic bottlenecks and events (Jones, Paetkau, Geffen, & Moritz, 2004; Luikart, Sherwin, Steele, & Allendorf, 1998) and quantify absolute population sizes (Miller, Joyce, & Waits, 2005; Mowat & Strobeck, 2000; Petit & Vialle, 2006). However, robust and informative microsatellite marker sets are often species specific in their primer sequences and must often be developed and characterized for each new and novel species in question. Furthermore, multiple microsatellite loci are required to sample the genome wide effects of drift as the informative power of microsatellite markers predominantly arises from their polymorphic nature and allelic diversity (Kalinowski, 2002; Morin, Martien, & Taylor, 2009). Whole genome genotyping using Single Nucleotide Polymorphisms (SNPs) offers an attractive alternative, a microsatellite-based approach, as they provide genome-wide haplotyping and are very common to both functional and non-coding regions of DNA. However, due to the low allelic diversity of SNPs, the experimental power of this technique is reliant on the number of loci examined. Therefore the total number of SNPs and associated cost required to undertake an analysis of population structure is often many times that of a microsatellite study (Selkoe & Toonen, 2006). Although SNPs have been used extensively to identify population structure in a number of model species (Morin et al., 2009), their use is often limited in novel species due to the cost of development and application.

The focal species of many ecological studies are often free-ranging, exist at low densities and can be challenging to capture and sample due to their cryptic life history strategies. As a result, the difficulty in attaining enough samples to undertake a microsatellite study of sufficient power can be of concern when attempting to infer meaningful demographic factors about populations in the wild (Piggott & Taylor, 2003). Genetic material acquired from indirect sources, such as hair and faeces, can be advantageous over direct genetic sampling as valuable information of a population’s past demography and diversity can be described without the effort required to capture study animals (Ernest, Penedo, May, Syvanen, & Boyce, 2000; Morin, Chambers, Boesch, & Vigilant, 2001; Reddy et al., 2012). The use of faeces as a genetic source material offers additional advantages when studying cryptic mammalian species as deposits are often abundant in the environment (Piggott & Taylor, 2003), spatially accountable and, theoretically, contain a substantial quantity of host genetic material (Davison, Birks, Brookes, Braithwaite, & Messenger, 2006).

Raw faecal samples recovered from the environment are often highly impure, both genetically and chemically, and have undergone a degree of mechanical and microbial degradation prior to collection (Taberlet, Waits, & Luikart, 1999). As a result, the often poor recovery of host DNA from faeces (Piggott, 2005; Taberlet et al., 1996, 1999) combined with the presence of naturally occurring organic PCR inhibitors (Hajkova et al., 2006; Morin et al., 2001) can result in the stochastic amplification of alleles during PCR and lead to a biased representation of a population’s genetic diversity. This
PCR artefact, known as allelic dropout, has been highlighted as a specific complexity when genetically profiling a population using faeces as the frequency of miss-amplification is often not apparent until marker performance can be externally validated using high-quality sources of DNA (Broquet & Petit, 2004).

A number of steps can be taken to minimise the potential for allelic dropout during the collection of samples in the field. The recovery of samples promptly after deposition by the study animal and the fixation of samples in a biologically inert environment prior to analysis are two factors that have previously been used to limit environmental degradation of target DNA. For example, studies investigating the use of faeces as a genetic source material in the pine marten (Martes martes) and red fox (Vulpes vulpes) used a clinical transport buffer designed to reduce the microbial action within faecal samples prior to analysis (O'Reilly, Statham, Mullins, Turner, & O'Mahony, 2008). Subsequent to collection, sample processing is often used to reduce the concentration of organically derived PCR inhibitors in samples prior to DNA extraction (Butler, 2005; Griffiths, Whiteley, O'Donnell, & Bailey, 2000). Previous studies have used standard DNA purification processes involving organic solvents in an attempt to reduce the concentration of PCR inhibitors in a range of adulterated samples from the environment (Griffiths et al., 2000). However, as PCR templates derived from faeces are fundamentally prone to miss-amplification (Pompanon, Bonin, Bellemain, & Taberlet, 2005; Taberlet et al., 1996, 1999), an external validation of both marker performance and DNA yield would be appropriate before reliable estimates of inbreeding and structure can be made.

The black-backed jackal is considered a major conflict species within the livestock-producing areas of South Africa. As a significant disease vector (Loveridge & Macdonald, 2001) and accomplished predator of game and livestock (Klare, Kamler, Stenkewitz, & Macdonald, 2010), understanding the dispersal and breeding patterns of the black-backed jackal is fundamental in the successful management of this species. To date, many studies have attempted to glean information on dispersal, disease transmission and depredation rates using both direct observations and population sampling methods (e.g. Ferguson, Nel, & Wet, 1983; Hiscocks & Perrin, 1988; Jenner, Groombridge, & Funk, 2011). However, very few studies have attempted to estimate the population parameters of free ranging black-backed jackals using a microsatellite-based approach. Six microsatellite loci were developed and characterised for the black-backed jackal in this investigation in order to provide a stable platform from which further analysis of population structure can be undertaken. Due to the cryptic and territorial behaviour of the black-backed jackal, faecal deposits were assessed for their suitability in providing non-invasive genetic source material for a microsatellite-based analysis of population structure. Due to the known complexities involved with isolating DNA from faecal samples and the variability of success across differing species, the feasibility of using DNA recovered from faecal samples was first ascertained before any meaningful demographic information was gleaned from the population. This study aimed to define a suitable and robust non-invasive faecal sampling method, capable of recovering adequate concentrations of host DNA in order to amplify a suite of polymorphic microsatellite markers originally developed for the domestic dog (Canis familiaris).

2. Methods and materials

2.1. Collection and storage of DNA source material

2.1.1. Tissue
A total of 18 black-backed jackals were directly sampled for genetic material by taking tissue biopsies of the ear lobe. All tissue samples were collected from recently deceased individuals within the North West Province of South Africa. Tissue samples were collected over an 18-month period and primarily originated from animal–traffic collisions. All tissue samples originate from individuals positively identified as C. mesomelas and spatially referenced with GPS data upon collection. Tissue samples were placed in absolute ethanol (EtOH) upon collection and stored at −20°C prior to transport to the UK for analysis.
2.1.2. Faeces
Seventy-five faecal samples were collected from six separate locations within the North West Province of South Africa. The minimum, maximum and average distance between sampling sites was 27, 110.4 and 66.1 km respectively. All samples were geographically referenced upon collection using a Garmin GPS 62S. Collected samples showed little to no signs of environmental degradation and retained an undesiccated mucosal coat. Stool Transport and Recovery (S.T.A.R.) buffer (Roche diagnostics: 03335208001) was used to minimise microbial degradation post collection and prevent sample desiccation during transport and storage. Upon collection, approximately 1 g of faecal material was removed from the outermost layer of the faecal pellet using a sterile razor blade, and subsequently deposited into 5 ml of S.T.A.R. buffer. The outermost layer of the faecal pellet was targeted to maximise the number of shed host cells recovered from the mucosal coat. Samples were shone until a homogeneous solution was formed. Variation in stool consistency required additional S.T.A.R. buffer to be used to ensure homogenization of some samples. Samples were stored at −20°C prior to transport to the UK for further extraction and analysis.

2.2. DNA recovery and extraction

2.2.1. Tissue
Host DNA was extracted from 25 mg of ear lobe tissue fixed in absolute EtOH using the DNeasy Blood and Tissue Kit (Qiagen cat No: 69504) and the manufacturer’s protocol for tissue extraction. Cartilaginous tissue was removed from epidermal skin cells prior to proteinase K digestion at 56°C. DNA was recovered in 150 μl of manufacturer-supplied PCR-compatible elution buffer.

2.2.2. Faeces
Template DNA was extracted from faecal samples using a Qiagen DNeasy spin column modified with a chloroform DNA extraction procedure. Raw samples stored in S.T.A.R. buffer were defrosted in batches of 10 at 4°C for 2 h prior to extraction. Samples were again homogenised by vortexing, then 10 ml of each sample was pipetted into a clean and sterile 15 ml collection tube. One ml of ≥99.8% chloroform + ethanol (GC) was added to the 10 ml of S.T.A.R. buffer/faecal solution and mixed gently until emulsified. Samples were then centrifuged at 1,000 × g for 3 min. The supernatant was then carefully removed and stored in a clean 15 ml collection tube for subsequent DNA extraction.

In order to extract DNA from the supernatant, a Qiagen tissue extraction protocol was followed. Columns were centrifuged at 1400 × g for 3 min and allowed to dry at room temperature for 5 min prior to DNA elution. DNA was eluted using 75 μl of warmed Qiagen elution buffer at 54°C. The elution buffer was added to the membrane and allowed to stand at room temperature for 5 min. Samples were then centrifuged at 6,100 × g and the DNA eluted into a 1.5 ml micro-centrifuge tube.

2.3. DNA template quantification
The average absolute copy number of host-specific DNA recovered from faecal samples was ascertained to ensure reliable DNA template quantity for accurate genotyping analysis. A real-time quantitative PCR assay was used to quantify the absolute DNA copy number in a subsample of four tissue and ten faecal extracts. The TaqMan probe used in this study had been designed for the quantification of cellular copy number in the domestic dog by targeting a highly conserved 287 bp fragment of the β-actin housekeeping gene (Table 1).

The conservation of the β-actin gene is expected to be high within the Canid genus due to its functional importance in cellular structural maintenance and thus suitable for quantifying the DNA copy number in C. mesomelas. However, to ensure the cross-species compatibility of this probe in C. mesomelas, the fragment of the β-actin gene containing the TaqMan probe binding site was first isolated and sequenced to insure homology in the black-backed jackal.
2.4. **β-actin probe homology**

To validate the homology of the TaqMan probe binding site in *C. mesomelas*, the probe primers were used to amplify the homologous locus in two high-quality jackal tissue extracts using a gradient PCR protocol (*n* = 27). The gradient PCR protocol consisted of an initial denaturation of 94°C for 5 min followed by 35 cycles of 94°C for 1 min, 53, 53.2, 53.7, 54.5, 55.5, 56.7, 57.8, 59.0 and 60°C for 1 min and 72°C for 30 s. A final extension stage was used at a temperature of 72°C for 5 min. PCR conditions consisted of 25 μl reactions comprised of 1 × Promega environmental PCR mastermix, 0.4 pmol/μl β-actin fragment primer mix and 2 μl of 1:10 dilution template DNA. Positive amplicons were visualised under UV light on a 2% agarose gel stained with ethidium bromide.

One optimal PCR product from each of the two individuals examined on a 2% agarose gel was then reclaimed using a Qiagen QIAquick gel purification kit (cat No: 28704 and 28706). Purified PCR products were then cloned into chemically competent *E. coli* using a TOPO TA Cloning kit, PCR2.1-TOPO TA vector and manufacturer-based chemical cloning protocol. *E. coli* was then spread onto ampicillin selective plates to identify positive uptake of the plasmid. Lac-Z gene expression was upregulated using 100 mM IPTG (isopropyl-β-D-1-thiogalactopyranoside) to aid in x-gal blue-white colony screening for successful insert uptake. *E. coli* were then grown for 24 h at 37°C.

Ten positively screened colonies were selected at random and removed from the agar plate using sterile pipette tips. Colony PCR was undertaken on each colony stab using the vector-based M13 primers in order to ensure successful insertion of the β-actin fragment within the recovered plasmid. PCR conditions consisted of an initial denaturation stage of 94°C for 10 min followed by 30 cycles of 94°C for 1 min, 55°C for 1 min and 72°C for 1 min. A final extension stage of 72°C for 10 min was used. The final concentration used in a PCR consisted of 1 × Promega environmental master mix, 0.25 pmol/μl β-actin forward and reverse primer, 0.5 μM BSA and up to 25 μl DI H₂O with 1 colony stab as DNA template. Ten millilitres of amoxicillin selective nutrient broth was subsequently inoculated with each colony stab pipette tip and incubated on a mechanical shaker at 37°C for 24 h. Colony PCR amplicons were visualised under UV light on a 2% agarose gel stained with ethidium bromide.

Ten millilitres of amoxicillin selective nutrient broth was subsequently inoculated with each colony stab pipette tip and incubated on a mechanical shaker at 37°C for 24 h. Two nutrient broth inoculations were selected due to the success of their corresponding colony PCR. Plasmids were recovered from *E. coli* in the nutrient broth solutions using a Qiagen mini prep recovery kit. Plasmid recovery was screened on a 1.2% agarose gel stained with ethidium bromide under UV light. To insure the homology of the β-actin primers and to accurately quantify the length and composition of the insert sequence recovered from the black-backed jackal, 5 μl of a 1:5 dilution of mini prep solution was sent for commercial Sanger sequencing. Both forward and reverse strands of the plasmid were sequenced using the M13 insert flanking primers by the commercial sequencing company GATC bioscience. Sanger sequencing was undertaken on an ABI 3730xl. Forward and reverse consensus sequences were compiled, trimmed and analysed using CLC Genomics Workbench v. 6.8.

### Table 1. β-actin forward and reverse primer sequences and custom TaqMan probe design for *Canis familiaris*

| Oligonucleotide | Sequence | Tm |
|-----------------|----------|----|
| β-actin forward | 5’-ccctgcgcctccaggtgaaa | 54 |
| β-actin reverse | 5’-gggtgcgctcctggctctt | 55 |
| TaqMan probe | FAM-aggacctctatgcacaagtc-TAMRA | 61 |

2.5. **Synthesis of DNA copy number standards for qPCR**

To accurately estimate the absolute copy number of host DNA within the extracted tissue and faecal samples, a set of DNA template standards of known copy numbers were created. DNA plasmids containing the β-actin fragment insert were recovered from the selected *E. coli* using a Qiagen DNA
mini prep plasmid recovery kit. The average DNA concentration of the mini prep elution was estimated using a NanoDrop 2000 spectrophotometer and five replicate samples. The total plasmid copy number per μl of mini prep elution was calculated from the combined molecular mass of the recovered plasmid and insert based on an average nucleotide mass of 650 Daltons and a total plasmid length of 3908 bp. The plasmid copy number was then standardised to 1 × 107 copies/μl by diluting 4.13 μl of mini prep elution in 495.87 μl of DI H2O. To reduce a qPCR bias from pipetting error, no less than 5 μl of DNA template was pipetted during the qPCR. A 1:5 dilution of the standard was created by pipetting 100 μl of standard at 1 × 107 copies/ μl into 400 μl of DI H2O. From this, six 1:10 serial dilutions were created from 2 × 106 to 2 × 10⁻¹ by diluting 50 μl of stock into 450 μl of DI H2O.

2.6. Absolute quantification of DNA template concentration via qPCR

Ten faecal sample extracts and four tissue sample extracts were used in the quantification of DNA template copy number by qPCR using a Canid specific reporter probe. Each extract was run in triplicate with a negative and positive control. qPCR was performed on an ABI Fast 7,000 real-time thermal cycler. Copy number quantification and analysis was undertaken using ABI software 7000 v. 2.0.5.

Seven standards of known fragment copy number were used to calculate sample template copy number and cycle threshold (Ct). Each standard was run in triplicate to ensure copy number accuracy. A total of 96 reactions were run with the following PCR conditions: 50°C for 2 min followed by 95°C for 10 min then 40 cycles of 95°C for 15 s and 60°C for 1 min. The detection of TaqMan probe dissociation and subsequent fluorescence took place at the end of each annealing cycle. The following PCR reagents were used in 25 μl reactions with a final reaction concentration of 1 × Promega environmental master mix 900 μM β-actin forward primers, 900 μM β-actin reverse primer, 0.25 μM probe, and 5 μl of DNA template and up to a 25 μl reaction volume with RNA-free DI H2O.

2.7. Microsatellite characterisation, allelic dropout and null allele frequencies in tissue and faecal-derived DNA templates

A total of six microsatellite loci were selected for homology testing within C. mesomelas (Table 2) from those originally described for pedigree analysis in the domestic dog (Ostrander, Mapa, Yee, & Rin, 1995; Ostrander, Sprague, & Rin, 1993; Wayne & Ostrander, 1999; Wictum et al., 2013). All primers used were synthesised by Eurofins MWG Operon and purified by HPLC. The primer pairs for all six markers were selected as they had previously shown the potential for cross-species homology within the Canid genus (Magory Cohen et al., 2013; Zachos et al., 2009). Careful consideration was given to maintaining consistent PCR annealing temperatures across all selected primers in order to facilitate their use in batch processing of large sample sizes. Individual PCR reactions were undertaken in 25 μl volumes containing approximately 40 ng of DNA template, 1 × Invitrogen PCR buffer, 1.5 mM MgCl2, 1 unit of Invitrogen hot start PlatinumTaq DNA polymerase (Invitrogen), 1 unit of Qiagen Q-solution, 0.5 μl/ng BSA, 0.2 mM dNTP mix and 0.2 μM primer mix. Amplification conditions used on a Techne TC-4000 thermal cycler consisted of: initial denaturation at 94°C for 2 min, followed by 35 cycles of 94°C for 1 min, 55°C for 45 s and 72°C for 1 min finishing with a final extension stage of 72°C for 5 min.

| Locus     | F Primer                     | R Primer                     | Repeat type | AN   |
|-----------|------------------------------|------------------------------|-------------|------|
| DogP109   | aaccttgaaccacactttgca       | accttgctctggttttaagc        | (CA)ₙ       | L15666 |
| DogP123   | aacttcaccacatccctctctctctct | ttcattacccgtcctctctctctctct | (CA)ₙ       | L15700 |
| DogP204   | cgagctcctgatggaagagctta     | caaagtctgatggaagagctta      | (CA)ₙ       | L15664 |
| DogP374   | aggaggacagaaagacagaaagaaag | aaggtgtctgatggaagagctta     | (CA)ₙ       | L24264 |
| VGL1541   | gacactctcatgtggaagagctta    | accttctgtgacttctctctctctct | (CTTT)ₙ     | NA    |
| DogPa622  | ctttgtgaacatcctctctctctcta | cccggctgtacttctctctctctctct | (CA)ₙ       | L27184 |

Table 2. Microsatellite loci, forward (F) and reverse (R) primer sequences (5′-3′), repeat type and NCBI accession numbers (AN) where available.
2.8. Microsatellite repeat motif characterisation
To characterise the basic repeat motif at each microsatellite locus, two PCR amplicons from 14 independent homozygote individual samples were sequenced. Direct Sanger sequencing was undertaken on both of the DNA strands for each locus to ensure accurate motif characterisation. Commercial sequencing companies Source Bioscience: Life Science and GATC Bioscience were used to undertake high-quality Sanger sequencing analysis on an ABI 3730xl. Consensus sequences were aligned and compiled using BIOEDIT v. 1.0 (Hall, 1999). Sequence repeat motifs were visually inspected and quantified using BIOEDIT v. 1.0 (Hall, 1999).

2.9. Allelic dropout and null alleles
Allelic diversity at each locus was examined by scoring a sample of 18 tissue samples and 65 faecal samples that showed positive amplification at one or more loci. PCR products were visualised under UV light using 10 μl of PCR product on a 2% agarose gel. Subsequently, 6 μl of PCR product was then sent for direct commercial genotyping (Source Biosciences: Life Sciences, UK). Forward primers were fluorescently labelled at the 5’ end with either FAM or HEX and amplicons were separated using capillary electrophoresis using an ABI Prism 3730. PCR fragment size was determined using the internal size standard Rox 500 (ABI Dye Set-30) and GENEMARKER v. 2.1.2.

All PCR products were single loaded to ensure accurate characterisation of allele polymorphisms and identification of potential PCR artefacts. Allelic dropout estimates were calculated using CERVUS v. 3.0 (Kalinowski, Taper, & Marshall, 2007), null allele estimates were undertaken using the program MICROCHECKER v. 2.2.3 (Van Oosterhout, Hutchinson, Wills, & Shipley, 2004). Fishers combined probability test is performed for all homozygote classes and significant differences in H-W proportions are identified between the observed and randomized genotypes (Van Oosterhout et al., 2004). This approach was used as it encompasses both a measure of overall heterozygosity and repeat length consistency. This permits the differentiation of null alleles and genotyping errors as a causative factor for excess heterozygosity in the observed genotypes.

2.10. Identification of neutrality on individual microsatellite loci
Identifying neutrality at microsatellite loci is based on the assumption that the effect of inbreeding at different loci should be consistent across loci, given the demographic history of the population in question, as inbreeding has a genome wide effect. As inbreeding and differentiation are strongly linked to the heterozygosity at each locus, outlying $F_{ST}$ values can be identified by plotting locus $F_{ST}$ values against the heterozygote frequency of the total population using a null distribution generated from the data. Expected $F_{ST}$ is estimated from the data using the mean $F_{ST}$ between loci, weighted by heterozygosity at individual locus. Heterozygosity is estimated from the average pairwise difference in $F_{ST}$ between all gene combinations. $F_{ST}$ is estimated using the method of Cockerham and Weir (1993). Loci outside the expected $F_{ST}$ distribution are assumed to be subject to selection.

The program Lonisan (Antao, Lopes, Lopes, Beja-Pereira, & Luikart, 2008) is an open source program used to detect selection in a range of molecular markers based on the $F_{ST}$ outlier method (Beaumont & Nichols, 1996). All 65 individual genetic profiles were used in the analysis collected from the six study sites. The program was run with the default settings using the infinite alleles model with 50,000 iterations at a confidence interval of 0.99.

2.11. Probability of identity
The ability to resolve population substructuring and individual identities using the novel microsatellite data-set was assessed using API-CALC v. 1.0 (Ayres & Overall, 2004). The probability that two identical copies of a constructed genotype from a population are from distinct individuals is the average probability of identity ($PI_{ave}$), and was calculated using all microsatellite genotype profiles generated in this study. As little is yet known about the potential for population substructuring, and the degree of relatedness between C. mesomelas individuals within a population, $PI_{ave}$ was estimated for a range of $F_{ST}$ values. The probability of identity sibs ($P(ID)_{sib}$) (Evett & Weir, 1998) was also
calculated using the program GENECAP (Wilberg & Dreher, 2004) as it provides a conservative estimate of PI and can remain informative when used on data from inbred populations.

2.12. Weir and Cockerham fixation statistics and pairwise genetic distances between samples

To assess the quality of microsatellite marker performance in estimating the population parameters in the black-backed jackal, the total population fixation index ($F_{ST}$) and allele frequency correlation estimates between all six sampled subpopulations (Weir & Cockerham, 1984) were calculated using the program GENEPOP v. 4.2.1. Evidence for genetic isolation by distance was also assessed by plotting a pairwise genetic distance matrix and a pairwise spatial distance matrix. Tissue samples were used in the latter analysis as they provided the highest quality and quantity of host DNA. Pairwise $F_{ST}$ significance values and Bonferroni corrections were undertaken using the program FSTAT v. 2.9.3.2 (Goudet, 1995). A Mantel test for dissimilarity was performed between the two matrices using the program R v. 3.0.2 (permutation = 999 model = strata). Population differentiation between localities was tested using the exact G test and tested for significance using the Markov chain algorithm (Dememorisation = 1,000, batches = 100, iterations/batch = 1000).

3. Results

3.1. DNA template quantification

3.1.1. β-Actin probe homology

All 27 PCRs from the two tissue samples produced positive amplicons using the β-actin primers. All amplicons were within the expected size range of 500 bp when analysed on an agarose gel. Gradient PCR provided the optimum annealing temperature as 55°C for the use of the β-actin primer pairs on the DNA templates originating from *C. mesomelas*. Following plasmid recovery, the insert fragment in the vector was sequenced using the flanking M13 primer pair to ensure the correct sequence insert and accurate characterisation of the primer and probe annealing sites. Both the primer-binding sites and TaqMan probe target site remained conserved and intact within *C. mesomelas*. In order to ensure the correct amplification of the DNA fragment in *C. mesomelas* using the β-actin primers designed for *C. familiaris*, the NCBI nucleotide database was queried with the NCBI nucleotide BLAST algorithm for a corresponding fragment sequence to reveal a 94% coverage similarity with the β-actin DNA sequence within *C. familiaris*. In addition, a 1 bp deletion and an 8 bp mismatch are apparent within the sequence acquired from amplicons originating in *C. mesomelas* when compared with the original sequence from *C. familiaris* (NCBI Reference Sequence: NM_001195845.1).

3.1.2. Absolute quantification of DNA template concentration via qPCR

During the qPCR, no amplification was seen within the three negative controls. All four DNA extracts from tissue source material showed positive amplification in all three repeats. Of the ten faecal samples run in triplicate, six samples showed clear and positive amplification in all three repeats (Figure 1). After removing the failed reactions from the analysis, the qPCR efficiency was calculated at 103.461% with a $R^2$ value of 0.935. Approximately 3.7 cycles separated the cycle threshold ($C_T$) value for each of the tenfold standard dilutions. The average number of copies of host DNA recovered from tissue source material was significantly greater than the average DNA copy number recovered from faecal source material (Figure 2).

The difference in detection cycle threshold ($C_T$) was examined between the two source material extracts. Amplification of the β-actin target fragment was seen consistently to have lower $C_T$ values than the amplification of target DNA in faecal extracts, by a maximum range of eight cycles.
3.2. Microsatellite characterisation and allelic dropout in tissue and faecal-derived DNA templates

3.2.1. Microsatellite repeat motif characterisation

All DNA extracts from *C. mesomelas* tissue provided good-quality DNA templates of >50 ng/μl; suitable for PCR analysis. All DNA templates derived from tissue produced positive PCR amplicons for all six sets of primers screened for cross-species amplification. Successful amplification of domesticated dog control templates was also seen for all loci. Direct Sanger sequencing of PCR amplicons was used to determine the basic microsatellite repeat motif at each locus. Repeat sequence motifs for each allele sequenced are shown in Table 3. All repeat length variations were consistent with allele size.
variation within each locus examined. The number of repeated nucleotide subunits ranged between 6 and 13 across all loci. Four of the five di-nucleotide microsatellite loci sequenced were comprised of pure and uninterrupted (CA)$_n$ based simple sequence repeats. Locus DogP374 was found to repeatedly produce uninterrupted compound repeat motifs (GC)$_m$(CA)$_n$ for all three alleles sequenced. The single tetra-nucleotide marker VGL1541 comprised an uninterrupted variable (CTTT)$_n$ repeat sequence for three separately genotyped individuals.

3.2.2. Allelic dropout in tissue and faeces
All six loci gave repeatable and scorable allele sizes from a total sample size of 18 tissue samples (Table 4).

Accumulative PCR and fragment analysis success rate for all microsatellite loci amplified from tissue was calculated at 98.15% ($n$ = 108). All microsatellite markers were seen to be polymorphic in *C. mesomelas* with between 6 and 13 alleles detected per locus. No significant allelic dropout could be detected and the estimated frequency of null alleles was low within the sampled population (Table 4).

Sixty-five out of 72 faecal extracts gave positive amplification of one or more microsatellite loci using end-modified fluorescently labelled primers when visualised on a 2% agarose gel. PCR amplicon size was then estimated using capillary electrophoresis for the successful samples. Null allele

Table 3. Locus name; microsatellite composition; sequenced allele size and repeat characterisation for microsatellites amplified from 14 individual *Canis mesomelas* samples

| Locus   | Composition     | Allele size (bp) | Repeat motif       |
|---------|-----------------|------------------|--------------------|
| DogP109 | Pure            | 144              | (CA)$_{14}$        |
| DogP109 | Pure            | 136              | (CA)$_{12}$        |
| DogP123 | Pure            | 138              | (CA)$_{13}$        |
| DogP123 | Pure            | 136              | (CA)$_{12}$        |
| DogP204 | Pure            | 214              | (CA)$_{18}$        |
| DogP204 | Pure            | 204              | (CA)$_{16}$        |
| DogP374 | Uninterrupted compound | 196 (CG)$_{16}$ (CA)$_{19}$ | |
| DogP374 | Uninterrupted compound | 196 (CG)$_{16}$ (CA)$_{19}$ | |
| DogP374 | Uninterrupted compound | 196 (CG)$_{16}$ (CA)$_{19}$ | |
| VGL1541 | Pure            | 193              | (CTTT)$_{16}$      |
| VGL1541 | Pure            | 193              | (CTTT)$_{16}$      |
| VGL1541 | Pure            | 185              | (CTTT)$_{16}$      |
| DogP622 | Pure            | 216              | (CA)$_{16}$        |
| DogP622 | Pure            | 210              | (CA)$_{16}$        |

Table 4. Number of allelic length variations in each microsatellite locus; PCR success rate and null allele estimate calculated by Cervus v. 3.0 and MICROCHECKER v. 2.2.3 ($n$ = 18)

| Locus   | Allele number | Size range (bp) | Positive PCR (%) ($n$ = 18) | Null allele frequency estimate |
|---------|---------------|-----------------|-------------------------------|-------------------------------|
| DogP109 | 13            | 126–192         | 100                           | −0.0539                       |
| DogP123 | 10            | 134–152         | 100                           | −0.0459                       |
| DogP204 | 11            | 198–222         | 100                           | −0.0073                       |
| DogP374 | 7             | 170–206         | 94.5                          | 0.0231                        |
| VGL1541 | 7             | 181–227         | 100                           | −0.0079                       |
| DogP622 | 6             | 210–220         | 100                           | 0.0334                        |
estimates and allelic diversity for each locus are summarised in Table 5. No significant false allele calling, due to stutter, was detected via MICROCHECKER v. 2.2.3 (Van Oosterhout et al., 2004). The expected and observed allele frequencies were highly similar for all loci examined. However, the total observed heterozygosity at locus DOGP123 was much lower than expected under the assumption of random assortment, given the number of alleles present within the sampled population.

No significant evidence for large allelic dropout or null alleles was detected within the sampled population by analysis either via MICROCHECKER or CERVUS.

Sampling effort was plotted as a function of accumulative allelic diversity to ensure an adequate detection of allelic variation was achieved within the sampled population (Figure 3).

3.2.3. Identification of selection on individual microsatellite loci

No significant evidence for selection could be identified for any microsatellite marker when examined using the program Lositan (Figure 4). No significant outlier values (Simulated $\hat{F}_{ST} - \text{Sample } F_{ST} > 0.95$) could be ascertained after 50,000 simulations. All markers were identified as being void of significant selection relative to the marker set. $F_{ST}/H$ simulations placed candidate markers within the bounds of neutral selection (grey) at the 99% significance interval. Marker DogP204 showed a level of heterozygosity and $F_{ST}$ value that places it on the boundary of acceptability. However, an extension of the simulation run could find no additional signal placing this marker as an outlier.

3.2.4. Average probability of identity ($PI_{ave}$ and $P(ID)_{sib}$)

Individual locus heterozygosity, expected average probability of identity for each locus, and total accumulative observed and expected $PI_{ave}$ across all loci were estimated using API-CALC v. 1.0 across a range of $F_{ST}$ values. The microsatellite markers examined in this study are seen to be highly informative of identity in non-inbred subpopulations under simulated $F_{ST}$ values of $<0.75$. As inbreeding is known to reduce the probability of identity through the duplication of alleles resulting from mating between related individuals with common ancestors, a range of $F_{IS}$ values were plotted against $PI_{ave}$ at an $F_{ST}$ value of 3%. The $F_{ST}$ value of 3% was chosen as this was the average $F_{ST}$ value across all loci. Cumulative and individual observed and expected $PI_{ave}$ for a population with assumed $F_{IS}$ and $F_{ST}$ values of 0 are shown in Table 6.

Probability of identity (sib) (Evett & Weir, 1998) was calculated across all six loci using GENECAP. This method was chosen as it provides a conservative estimate of $PI$ as the effects of linkage and

| Locus  | Allele number | Null allele frequency estimate | $H_{obs}$ | $H_{exp}$ | Total $H_{o}$ | Total $H_{E}$ |
|--------|---------------|-------------------------------|-----------|-----------|---------------|---------------|
| DogP109 | 12            | -0.0160                       | 0.824     | 0.818     | 11            | 7.854         |
| DogP123 | 14            | -0.1038                       | 0.971     | 0.824     | 2             | 10.254        |
| DogP204 | 11            | 0.0409                        | 0.727     | 0.800     | 10            | 7.302         |
| DogP374 | 9             | -0.0221                       | 0.667     | 0.675     | 21            | 19.877        |
| VGL1541| 10            | -0.0509                       | 0.818     | 0.759     | 12            | 10.489        |
| DogP622 | 6             | -0.0653                       | 0.727     | 0.663     | 17            | 21.211        |

Notes: Null allele frequency, $H_{obs}$, and $H_{exp}$ was estimated using CERVUS v. 3.1.2. No significant evidence for allelic dropout or null alleles was detected by MICROCHECKER v. 2.2.3 and repeat type specific analysis at 95% ($n = 65$). $H_{obs}$ is the observed heterozygote frequency; $H_{exp}$ is the expected heterozygote frequency; $H_{o}$ is the total number of observed heterozygotes; $H_{E}$ is the total number of expected heterozygotes.
structure are considered in this analysis. $P(ID)_{sib}$ for the total sampled population was $1.03 \times 10^{-3}$. Locus-specific $P(ID)_{sib}$ and accumulative $P(ID)_{sib}$ are shown in Table 7.

### 3.3. Weir and Cockerham fixation statistics and pairwise genetic distances between samples

A high degree of allelic variation was seen at the six microsatellite loci examined. Between 6 and 15 alleles were recorded for each locus within the entire sampled population. When the sampled individuals from all six sites were clustered into one data population, a significant deviation from the Hardy–Weinberg equilibrium was apparent ($X^2 = \text{Infinity}, df = 12, p < 0.01$). The Hardy–Weinberg probability test for each locus within the total sampled population showed a high degree of
variability, with both positive and negative $F_{IS}$ values calculated as being significantly different from zero (Table 8).

The Hardy–Weinberg exact test and a multi-locus test for heterozygote frequency revealed five of the six sampled populations did not deviate from Hardy–Weinberg expectations at the 5% significance level. A single locality showed a significant deviation from Hardy–Weinberg expectations ($X^2 = 33.4919, df = 12, p = 0.0008$) and multi-locus excess of heterozygote individuals ($p = 0.016, S.E. = 0.0022$).
The total proportion of variation in allele frequency explained by sample site isolation ($F_{ST}$) is approximately 3% (Table 9); however, a large degree of variation is apparent when each locus is examined independently.

**Table 8. Weir and Cockerham’s $F_{IS}$ estimates calculated for each locus for the total population**

| Locus   | $p$-Value | S.E.  | $F_{IS}$  |
|---------|-----------|-------|-----------|
| DogP109 | 0.0069    | 0.0033| 0.0215*   |
| DogP123 | 0.0001    | 0.0001| −0.0791*  |
| DogP204 | 0.0814    | 0.0169| 0.0740    |
| DogP374 | 0.8060    | 0.2250| 0.0448    |
| VGL1541 | 0.5325    | 0.0208| 0.0189    |
| DogP622 | 0.2391    | 0.0091| −0.0156   |

*Significant at 5% level.

**Table 9. Weir and Cockerham fixation statistics for each locus and combined across six loci between all localities**

| Locus   | $F_{IS}$  | $F_{ST}$  | $F_{IT}$  |
|---------|-----------|-----------|-----------|
| DogP109 | 0.0835    | −0.0070   | 0.0771    |
| DogP123 | −0.1788   | 0.0074    | −0.1701   |
| DogP204 | −0.0024   | 0.0834    | 0.0812    |
| DogP374 | 0.1005    | −0.0015   | 0.0991    |
| VGL1541 | 0.0223    | 0.0062    | 0.0284    |
| DogP622 | −0.1823   | 0.1146    | −0.0468   |
| All     | −0.0253   | 0.0307    | 0.0062    |

Figure 5. Pairwise $F_{ST}$ distances as a function of spatial distance.

Note: Mantel statistic = 0.1017, $p = 0.336$. 
Significant genetic differentiation was apparent between sample sites when examining the variation in allele frequencies between sites using the exact $G$ test ($X = 53.768, \text{df} = 12, p = 3 \times 10^{-7}$).

Pairwise ($F_{ST}$) estimates were calculated for each possible population pair and summarised across all loci to examine isolation by distance for evidence of a continuous population. Pairwise ($F_{ST}$) estimates were plotted against site distance matrix and tested for correlation by estimating a mantel statistic (Figure 5).

Pairwise ($F_{ST}$) gave no indication of population structuring due to isolation by distance alone as no significant correlation between pairwise $F_{ST}$ distance estimates and site distance was apparent (Mantel statistic $r = -0.1017, p = 0.336$, based on 999 permutations).

4. Discussion

Both the β-actin primer annealing sites and Taqman probe annealing site within *C. mesomelas* were shown to be homologous to that of *C. familiaris* by plasmid sequence analysis. A single base pair deletion and 8 sequence mismatches were seen within the amplified β-actin fragment from *C. mesomelas* when compared with the amplified fragment from *C. familiaris*. The 1 bp deletion and 8 independent bp substitutions detected in the jackal template seems to have little effect on the probe and primer efficiency as no mutations were detected within the probe and primer annealing sites. The presence of such genotypic inconsistencies between *C. mesomelas* and *C. familiaris* is not unexpected, given the 3–7 million years of evolutionary divergence separating these species (Lindblad-Toh et al., 2005; Vila, Maldonado, & Wayne, 1999; Wayne & Ostrander, 1999).

The cumulative efficiency of the qPCR assay across all reactions was 103.5% which lies within the acceptable reaction efficiency of 95–105% for detection of fragment amplification. The predictive power of the standard dilutions for estimating the DNA copy number in unknown samples was sufficient with an $R^2$ value of 0.935. The recovery, quantification and serial dilution of the purified plasmid standard was efficient and thus accurate estimates of unknown DNA template copy number can be ascertained with sufficient statistical power.

Four of the six faecal samples examined via qPCR failed to indicate significant amplification of the β-actin target fragment in any of the three repeats examined. The failure of amplification in these samples may be due to inadequate initial DNA template concentration or a failure of the assay itself. The sensitivity of qPCR to reaction inhibition by chemical adulterants is significantly greater than that of standard endpoint PCR due to the additional requirement of successful probe annealing and dissociation. In addition, the total length of the β-actin amplicon is approximately 80 bp longer than the largest microsatellite examined, thus the success rate of qPCR should be considered a conservative estimate when compared with amplification of smaller microsatellite loci from faecal templates using standard endpoint PCR protocols.

The amplification of the six remaining faecal samples showed positive amplification with $C_t$ detection values of <34 cycles. The average DNA copy number recovered from the faecal template was 81.1 copies/μl with a standard deviation of 96.6 copies/μl. This indicates that recovery of significant quantities of host-specific DNA from faecal source material is possible; however, it is clear that the absolute quantity of DNA recovered was highly variable. The large difference in initial template copy number between faeces and tissue is undoubtedly due to the indirect nature of the faecal sample collection and the associated low quantity and quality of host DNA. Microbial and mechanical degradation of host DNA will also play a factor in low initial DNA template concentration.

This analysis of absolute template quantification using a highly specific reporter probe method, combined with the statistical analysis for allelic anomalies arising from faecal-derived DNA templates, strongly supports the use of jackal faecal samples as a source of genetic material in population studies. While the average concentration of host DNA recovered from faeces is theoretically suitable for undertaking genome-wide SNP analysis and is arguably more suitable for degraded
templates when using plate technology. However, the initial cost and labour required to identify thousands of SNPs for a genome-wide coverage would be many times higher than a microsatellite-based study of similar resolution. Therefore, while extensive SNP genotyping remains unavailable for this species, microsatellites offer a viable and effective method of estimating the population parameters of the black-backed jackal. The variability in template quantity recovered from faecal samples may be a hindrance in a subset of samples collected in the field. Therefore, it is advised that a crude visual screening process as proposed by Hogan, Cooke, Burridge, and Norman (2008) is undertaken prior to fragment analysis to avoid the use of poor-quality DNA templates (Hogan et al., 2008).

The complete consensus genotype data from all DNA source material was required to maximise the sample size needed to accurately characterise the microsatellite set developed for *C. mesomelas* in this investigation. Again, the full consensus data-set of 65 individuals indicated no significant presence of null alleles or allelic dropout in the sampled population or misidentification of allele sizes due to inconsistent stutter patterns in the electrophoretogram. The cross-species amplification of these microsatellite loci, which have been specifically refined for pedigree analysis in the domestic dog, has proven consistent in both faecal and tissue-derived DNA templates from *C. mesomelas*. The genetic homology of primer-binding sites between *C. familiaris* and *C. mesomelas* must therefore be highly conserved over an estimated 3–5 my evolutionary divergence separating these two species from their common ancestor (Lindblad-Toh et al., 2005; Vila et al., 1999; Wayne, Benveniste, Janczewski, & O’Brien, 1989). The effort invested in the development and characterisation of microsatellite loci required to undertake accurate pedigree analysis in the domestic dog is substantial and has driven the creation of highly informative, polymorphic and unlinked markers (Ostrander et al., 1993). The consistency in amplicon size and lack of erroneous stutter patterns in amplicons observed from samples taken from *C. mesomelas* may again be in part due to the stringent screening process originally undertaken to characterise these markers in the domestic dog (Ostrander et al., 1993).

No significant effect of selection could be found acting on the marker set chosen in this investigation. However, locus DogP204 did show a high level of genetic differentiation in the populations studied compared with the remaining five markers (Figure 4). The *F*<sub>ST</sub> outlier method was chosen as the comparison of pairwise variation in locus. *F*<sub>ST</sub> across a population captures the effect of independent genealogy and thus is less susceptible to sampling artefacts arising from small sample sizes. While this method remains a reliable estimate of selection at multiple loci in sample sizes >25 (Beaumont & Nichols, 1996) and across a range of inbreeding and structured populations (Beaumont & Nichols, 1996), calculation of the null distribution of *F*<sub>ST</sub> is reliant on variation between loci. Previous studies have used Lositan to successfully identify selection within marker sets with between 9 and 12 loci (Thaler et al., 2010; Zelnio et al., 2010), however relatively few studies have attempted to estimate a null distribution of *F*<sub>ST</sub> using only six loci, (e.g. An, Lee, & Dong, 2012). As such in order to get an accurate estimate of *F*<sub>ST</sub>, additional loci may be required to produce a more informative estimate of expected heterozygosity (Beaumont & Nichols, 1996), and further investigate the potential for selective neutrality in the marker set examined in this study. Despite this shortcoming, it is proposed that the cross-species marker set, characterised to investigate the population parameters of the black-backed jackal in this study, meet the assumption of neutral Mendelian inheritance and are suited to investigating the population parameters of this species.

The total probability of identity, *PI*<sub>ave</sub>, across all loci was used to quantify the ability of the microsatellites to uniquely identify individuals using both non-invasive faecal and tissue samples. The total *PI*<sub>ave</sub> indicates a low probability of shared genotype profiles occurring within the sample population based on the assumption of a random sampling of a non-structured population (Table 6). The least informative single locus marker was identified as locus DogP622 with an observed *PI*<sub>ave</sub> of 0.1578. This result is concurrent with the low allelic diversity associated at this marker when compared with the rest of the marker set. The most informative marker was DogP204 with a *PI*<sub>ave</sub> of 0.0216, closely followed by DogP109 with an observed *PI*<sub>ave</sub> value of 0.0306. These two microsatellite markers were the most polymorphic of the analysed marker set with 14 and 15 alleles respectively. However,
although Locus DogP123 was also seen to be highly polymorphic with an allelic number of 13, the
large proportion of heterozygote individuals recorded at this locus was thought to increase the
expected $\text{PI}_{\text{ave}}$ to 0.13828 when $F_{ST}$ and $F_{IS}$ were set at 0. The comparatively high $\text{PI}_{\text{ave}}$ value, com-
bined with the unexplained deficit in expected number of homozygote individuals at this marker, is
of potential concern with respect to the validity and reliability of this marker when used with
$C. \text{mesomelas}$. However, the full six-locus microsatellite profile is estimated to have a cumulative
$\text{PI}_{\text{ave}}$ of $1.96 \times 10^{-8}$ (Table 6). This means that the probability that two profiles matching at all six
markers originating from two distinct individuals is less than 1 in 51 million.

As population substructuring can have a significant effect on the probability of shared genotypes
within a sample, cumulative probability of identity was calculated for a range of $F_{ST}$ values. A sharp
increase in the cumulative probability of identity for the six microsatellite loci examined was seen
when simulated $F_{ST}$ values were greater than 0.75. As values of $F_{ST}$ are rarely greater than 0.2 in wild
populations (Allendorf, Luikart, & Aitken, 2012; Pritchard, Stephens, & Donnelly, 2000), it is therefore
apparent that the microsatellite set developed in this investigation consists of informative molecular
markers capable of accurately differentiating individuals in populations subject to partially reduced
gene flow and migration. However, in populations that are highly inbred, the ability for these mark-
ers to resolve individuals can be substantially reduced in inbred populations with $F_{IS} > 0.6$.

When using these microsatellite markers to investigate the population parameters of the black-
backed jackal, this study found that the observed genotype frequencies deviated significantly from
those expected under the Hardy–Weinberg equilibrium, given the observed allele frequencies within
the total population ($X^2 = \text{Infinity}, df = 12, p < 0.01$). This indicated that the total sampled population
cannot be described as one large randomly mating population. An excess of heterozygosity at the
individual level ($F_{IS} = -0.0253$) indicated potential outbreeding. Black-backed jackal subpopulations
at five of the six sample localities were shown to be in Hardy–Weinberg equilibrium. The variation in
allele frequencies between subpopulations gives rise to a deviation from the expected Hardy–
Weinberg proportions when examined at the total population level, indicating barriers to gene flow
and the non-random association of gametes within the population. The analysis of allele frequen-
cies within the sampled subpopulations relative to the total sampled population indicates popula-
tion structuring with an $F_{ST}$ value of 0.0307. Significant subpopulation differentiation was also
calculated within the total sampled population ($X^2 = 53.768, df = 12, p = 3.0 \times 10^{-7}$). No evidence of
spatial correlation with genetic isolation by distance could be ascertained using the mantel test
within the total population when examining the variation in allele frequencies within and between
subpopulations. The fixation statistic for the total population ($F_{ST}$), the significant level of subpopula-
tion differentiation expected through genetic isolation by distance (Mantel test, Figure 5), are indica-
tive of a population that is not continuous across its range and thus is subject to barriers to gene flow
across the study area.

It is assumed that in a genetically homogeneous population where dispersal is not spatially lim-
ited, gene flow reduces as the distance between individuals increases. The isolation by distance
model (Wright, 1946) is a modification of the island model which considers short distance dispersal
as opposed to assuming random and equal migration between any given island. The IBD model is
therefore used to infer population parameters by comparing genetic distance with linear geographic
distance under the assumptions of equal and unimpeded dispersal in all directions from a point of
origin. This approach provides the null model to which departures arising from barriers to gene flow
can be tested. It is not possible to infer what form of barrier to gene flow is in action from this analy-
sis, nor identify which sampling sites are isolated from the rest, yet it does provide evidence that a
barrier to gene flow is in effect. While this study makes no attempt to identify the ecological factors
that are responsible for the apparent population identified in this investigation, future studies are
required to clarify the causative factors associated with the significant $F_{ST}$ values in this population.

This study presents a robust and reliable suite of six polymorphic microsatellite loci suitable for
use in population studies of the black-backed jackal. Furthermore, this study includes a verified
method for the collection, purification and extraction of host-specific DNA templates from non-invasively collected faecal source material. Although the quantity of DNA recovered from non-invasive sampled faecal source material is lower than that of directly sampled tissue-derived source material, the mean absolute quantity of genetic material was adequate to undertake efficient PCR. However, the organically derived PCR inhibitor compounds inherent to faecal samples were seen to be of potential concern to the progression of the qPCR reaction.

Due to the stringent examination of microsatellite behaviour, the quantification of template quantity and the successful application of this marker suite to a real-world population, it can be said with reasonable confidence that non-invasive sampling of faecal-derived DNA templates presents a feasible and accurate method to facilitate the estimation of population parameters of the black-backed jackal using the novel marker set developed in this investigation.

Acknowledgements
We are particularly indebted to the Earth Watch Institution for providing funds to undertake both field and laboratory work for this study. We are grateful to African Explosives Ltd, Standard Bank and Investec Bank for permitting land access in order to collect fresh faecal samples of the black-backed jackal in game farm environments. Furthermore, we would like to thank Mr Dougal Mactavish and Mrs Lynne Mactavish for the coordination of earth watch project “the scavengers of South Africa”, from which a number of genetic samples were recovered. We would also like to thank the numerous Earth Watch volunteers that helped identify and locate faecal samples in the field.

Funding
The authors received no direct funding for this research.

Competing interests
The authors declare no competing interest.

Author details
Robert S. James1
E-mail: r.s.james@brighton.ac.uk
Phillip L. James2
E-mail: p.james@imperial.ac.uk
Dawn M. Scott1
E-mail: dawn.scott@brighton.ac.uk
Andrew D.J. Overall1
E-mail: a.d.j.overall@brighton.ac.uk
1 School of Pharmacy and Bioscience, University of Brighton, Brighton BN2 4GJ, UK.
2 School of Life Sciences, The University of Warwick, Coventry CV4 7AL, UK.

Citation information
Cite this article as: Characterization of six cross-species microsatellite markers suitable for estimating the population parameters of the black-backed jackal (Canis mesomelas) using a non-invasive genetic recovery protocol, Robert S. James, Phillip L. James, Dawn M. Scott & Andrew D.J. Overall, Cogent Biology (2015), 1: 1108479.

Cover image
Source: Author (Robert S. James).

References
Allendorf, F. W., Luikart, G. H., & Altkin, S. N. (2012). Conservation and the genetics of populations (2nd ed.). Oxford: Wiley-Blackwell.
Ar, H. S., Lee, J. W., & Dong, C. M. (2012). Population genetic structure of Korean pen shell (Atrina pectinata) in Korea inferred from microsatellite marker analysis. Genes Genomics, 34, 681-688. http://dx.doi.org/10.1007/s13258-012-0091-z
Antao, T., Lopes, A., Lopes, R. J., Beja-Pereira, A., & Luikart, G. (2008). LOSITAN: A workbench to detect molecular adaptation based on a FST-outlier method. Bioinformatics, 9, 323. http://dx.doi.org/10.1186/1471-2105-9-323.
Ayres, K. L., & Overall, A. D. (2006). Api-Calc 1.0: A computer program for calculating the average probability of outcrossing allowing for substructure, inbreeding and the presence of close relatives. Molecular Ecology Notes, 4, 315–318. http://dx.doi.org/10.1111/j.1471-8214.2004.00652.x
Beaumont, M. A., & Nichols, R. A. (1996). Evaluating loci for use in the genetic analysis of population structure. Proceedings of the Royal Society B: Biological Sciences, 263, 1619–1626. http://dx.doi.org/10.1098/rspb.1996.0237
Broquet, T., & Petit, E. (2004). Quantifying genotyping errors in noninvasive population genetics. Molecular Ecology, 13, 3601–3608. http://dx.doi.org/10.1111/j.1365-294x.2004.02352.x
Butler, J. M. (2003). Forensic DNA typing. Biology, technology and genetics of STR markers (2nd ed., 660 pp.). Burlington, MA: Elsevier Academic Press. ISBN:0-12-147952-8, 457–462.
Cain, B., Wadero, A. B., Showcraw, S. G., Edwin Harris, W., Stevens-Wood, B., Kemp, S. J., … Watts, P. C. (2014). Sex-biased inbreeding effects on reproductive success and home range size of the critically endangered black rhinoceros. Conservation Biology, 28, 594–603. http://dx.doi.org/10.1111/cobi.12175
Cockerham, C. C., & Weir, B. S. (1993). Estimation of gene flow from F-Statistics. Evolution, 47, 855–863. http://dx.doi.org/10.2307/2410189
Coulson, T., Albion, S., Slate, J., & Pemberton, J. (1999). Microsatellite loci reveal sex-dependent responses to inbreeding and outbreeding in red deer calves. Evolution, 53, 1951–1960. http://dx.doi.org/10.2307/2640453
Curik, I., Zechnier, P., Slikker, J., Achmann, R., Bodo, I., Dojc, P. … Brem, G. (2003). Inbreeding, microsatellite heterozygosity, and morphological traits in Lipizzan horses. Journal of Heredity, 94, 125–132. http://dx.doi.org/10.1093/jhered/esg029
Davison, A., Birks, J. D. S., Brooke, R. C., Brithwaite, T. C., & Messenger, J. E. (2006). On the origin of foeces: Morphological versus molecular methods for identifying rare carnivores from their scats. Journal of Zoology, 267, 141–143.
Ernest, H., Penedo, M., May, B., Syvanen, M., & Boyce, W. (2000). Molecular tracking of mountain lions in the Yosemite Valley region in California: genetic analysis using microsatellites and faecal DNA. Molecular Ecology, 9, 433–441. http://dx.doi.org/10.1046/j.1365-294x.2000.00890.x
Evett, I. W., & Weir, B. S. (1998). Interpreting DNA evidence: Statistical genetics for forensic scientists. Sunderland: Sinauer.
Ferguson, J., Nel, J., & Wet, M. J. (1983). Social organization and movement patterns of black-backed jackals Canis mesomelas of South Africa“, from which a number of genetic samples were recovered. We would also like to thank the numerous Earth Watch volunteers that helped identify and locate faecal samples in the field.
mesomelas in South Africa. Journal of Zoology, 199, 487–502.

Garcia-Moreno, J., Matocq, M. D., Roy, M. S., Geffen, E., & Wayne, R. K. (1999). Relationships and genetic purity of the endangered Mexican Wolf based on analysis of microsatellite loci. Conservation Biology, 10, 376–389. http://dx.doi.org/10.1046/j.1523-1739.1999.10003736.x

Goudet, J. (1995). FSTAT (version 1.2): A computer program to calculate F-statistics. Journal of Heredity, 86, 485–486.

Griffiths, R. I., Whiteley, A. S., O’Donnell, A. G., & Bailey, M. J. (2000). Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA- and RNA-based microbial community composition. Applied and Environmental Microbiology, 66, 5488–5491. http://dx.doi.org/10.1128/AEM.66.12.5488-5491.2000

Hajkova, P., Zemanova, B., Bryja, J., Hajek, B., Roche, K., Tkadlec, E., & Zima, J. (2008). Factors affecting success of PCR amplification of microsatellite loci from otter faeces. Molecular Ecology Notes, 6, 559–562. http://dx.doi.org/10.1111/j.1471-8286.2007.01945.x

Hall, T. A. (1999). BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series, 41, 95–98.

Hiscock, K., & Perrin, M. (1988). Home range and movements. Conservation Biology, 2, 1099–1106. http://dx.doi.org/10.1007/BF00293011

Hoganson, G. W., Zdarek, J. A., & Bronik, A. G. (2006). Estimating population size of grizzly bears using hair capture, DNA profiling, and mark-recapture analysis. The Journal of Wildlife Management, 68, 183–193. http://dx.doi.org/10.2193/3802989

Iwaniuk, N., Groombridge, J., & Funk, S. (2011). Commuting, territory and variation in group and territory size in a black-backed jackal population reliant on a clumped, abundant food resource in Namibia. South African Journal of Wildlife Research, 18, 97–100.

Jones, M. E., Paetkau, D., Geffen, E. L. I., & Moritz, C. (2004). Genetic diversity and population structure of Tasmanian devils, the largest marsupial carnivore. Molecular Ecology, 13, 2197–2209. http://dx.doi.org/10.1111/j.1365-3621.2004.01239.x

Kalinowski, S. T. (2002). How many alleles per locus should be used to estimate genetic distances? Heredity, 88, 62–65. http://dx.doi.org/10.1038/sj.hdy.6800009

Kalinowski, S. T., Taper, M. L., & Marshall, T. C. (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

Klare, U., Kamler, J. F., Stenkewitz, U., & Macdonald, D. W. (2010). Diet, prey selection, and predation impact of black-backed jackals in South Africa. Journal of Wildlife Management, 74, 1030–1042. http://dx.doi.org/10.2193/2009-211

Lindblad-Toh, K., Wade, C. M., Mikkelsen, T. S., Klare, U., Kamol, M., … Zody, M. C. (2005). Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature, 438, 803–819. http://dx.doi.org/10.1038/nature04338

Loveridge, A., & Macdonald, D. (2001). Seasonality in spatial organization and dispersal of sympatric jackals (Canis mesomelas and C. adustus): Implications for rabies management. Journal of Zoology, 252, 101–111. http://dx.doi.org/10.1017/S0952836099009977

Luikart, G., Sherwin, W., Steele, B., & Allendorf, F. (1998). Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. Molecular Ecology, 7, 963–974. http://dx.doi.org/10.1046/j.1365-294X.1998.00414.x

Magory Cohen, T., King, R., Dolev, A., Boldo, A., Lichter-Peled, A., & Kahila Bar-Gal, G. (2013). Genetic characterization of populations of the golden jackal and the red fox in Israel. Conservation Genetics, 14, 55–63. http://dx.doi.org/10.1007/s10592-012-0423-1

Miller, C. R., Joyce, P., & Wails, L. P. (2009). A new method for estimating the size of small populations from genetic mark-recapture data. Molecular Ecology, 14, 1991–2005. http://dx.doi.org/10.1111/j.1365-294X.2005.02577.x

Morin, P. A., Martien, K. K., & Taylor, B. L. (2009). Assessing statistical power of SNPs for population structure and conservation studies. Molecular Ecology Resources, 9, 66–73. http://dx.doi.org/10.1111/j.1755-0998.2008.02131.x

Mowat, G., & Strobeck, C. (2000). Estimating population size of wolverines (Gulo gulo) from noninvasive samples for accurate microsatellite genotyping of wild carnivores (Pronghornus verus). Molecular Ecology, 10, 1835–1844. http://dx.doi.org/10.1046/j.1091-6466.2000.00130.x

Morin, P. A., Martien, K. K., & Taylor, B. L. (2009). Assessing statistical power of SNPs for population structure and conservation studies. Molecular Ecology Resources, 9, 66–73. http://dx.doi.org/10.1111/j.1755-0998.2008.02131.x

Paetkau, D., Geffen, E., & Moritz, C. (2004). Genetic diversity and population structure of Tasmanian devils, the largest marsupial carnivore. Molecular Ecology, 13, 2197–2209. http://dx.doi.org/10.1111/j.1365-3621.2004.01239.x

Pigott, M. P. (2005). Effect of sample age and season of collection on the reliability of microsatellite genotyping of faecal DNA. Wildlife Research, 31, 485–493. http://dx.doi.org/10.1071/WR03107

Pigott, M. P., & Taylor, A. C. (2003). Remote collection of animal DNA and its applications in conservation management and understanding the population biology of rare and cryptic species. Wildlife Research, 30, 1–13. http://dx.doi.org/10.1071/WR02077

Pompanon, F., Bonin, A., Bellemain, E., & Taberlet, P. (2005). Genotyping errors: Causes, consequences and solutions. Nature Reviews Genetics, 6, 847–859. http://dx.doi.org/10.1038/nrg1707

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics, 155, 945–959.

Reddy, P. A., Bhavanishankar, M., Bhagavatula, J., Hanika, K., Mahal, R. S., & Shivaji, S. (2012). Improved methods of carnivore fecal sample preservation, DNA extraction and quantification for accurate genotyping of wild tigers. PLoS ONE, 7, e46732. http://dx.doi.org/10.1371/journal.pone.0046732

Selkoe, K. A., & Toonen, R. J. (2006). Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. Ecology Letters, 9, 615–629. http://dx.doi.org/10.1111/ele.2006.9.issue-5

Storfer, A., Mech, S. G., Reudink, M. W., & Lew, K. (2014). Inbreeding and strong population subdivision in an endangered salamander. Conservation Genetics, 15, 137–151. http://dx.doi.org/10.1007/s10592-013-0526-3

Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, V., Escaravage, N., … Bouvet, J. (1996). Reliable genotyping of
samples with very low DNA quantities using PCR. Nucleic Acids Research, 24, 3189–3194.

Taberlet, P., Waits, L. P., & Luikart, G. (1999). Noninvasive genetic sampling: Look before you leap. Trends in Ecology Evolution, 14, 323–327. http://dx.doi.org/10.1016/S0169-5347(99)01637-7

Thaler, A. D., Zel'no, K., Jones, R., Carlson, J., Van Dover, C. L., & Schultz, T. F. (2010). Characterization of 12 polymorphic microsatellite loci in Ifremeria nautilei, a chemoautotrophic gastropod from deep-sea hydrothermal vents. Conservation Genetics Resources, 2, 101–103. http://dx.doi.org/10.1007/s12686-010-9174-9

Van Oosterhout, C., Hutchinson, W. F., Willis, D. P. M., & Shipley, P. (2004). Micro-checker: Software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes, 4, 535–538. http://dx.doi.org/10.1111/j.1471-8286.2004.00797.x

Vernesi, C., Crestanello, B., Pecchioli, E., Tortari, D., Caramelli, D., Hauffe, H., & Bertorelle, G. (2003). The genetic impact of demographic decline and reintroduction in the wild boar (Sus scrofa): A microsatellite analysis. Molecular Ecology, 12, 585–595. http://dx.doi.org/10.1046/j.1365-294X.2003.01763.x

Vila, C., Maldonado, J. E., & Wayne, R. K. (1999). Phylogenetic relationships, evolution, and genetic diversity of the domestic dog. Journal of Heredity, 90, 71–77. http://dx.doi.org/10.1093/1471-8286.2004.00797.x

Wayne, R. K., Benveniste, R. E., Janczewski, D. N., O’Brien, S. J. (1989). Molecular and biochemical evolution of the Carnivora. In Carnivore behavior, ecology, and evolution (pp. 465–492). Springer US. doi:10.1007/978-1-4757-4716-4_18

Wayne, R. K., & Ostrander, E. A. (1999). Origin, genetic diversity, and genome structure of the domestic dog. BioEssays, 21, 247–257. http://dx.doi.org/10.1002/(SICI)1521-8788(199903)21:3<247::AID-BIEJ2-0000000000>3.0.CO;2-X

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. Evolution, 38, 1358–1370. doi:10.2307/2408641

Wilberg, M. J., & Dreher, B. P. (2004). Genecap: A program for analysis of multilocus genotype data for non-invasive sampling and capture-recapture population estimation. Molecular Ecology Notes, 4, 783–785. doi:10.1111/j.1471-8286.2004.00797.x

Zeltino, K., Thaler, A., Jones, R., Saleu, W., Schultz, T., Van Dover, C., & Carlsson, J. (2010). Characterization of nine polymorphic microsatellite loci in Chorocaris sp. (Crustacea, Caridea, Alvinocarididae) from deep-sea hydrothermal vents. Conservation Genetics Resources, 2, 223–226. http://dx.doi.org/10.1007/s12686-010-9243-0

© 2015 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

You are free to:

Share — copy and redistribute the material in any medium or format
Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.
No additional restrictions

You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.