Gel-free species identification using melt-curve analysis

Oliver Berry[1,3] & Stephen D. Sarre[2]

[1]Invasive Animals Co-operative Research Centre and School of Animal Biology (M092),
The University of Western Australia, Crawley, WA, 6009, Australia.
[2]Applied Ecology Research Group, University of Canberra, ACT, 2615, Australia
[3]Author for correspondence.

E-mail. ofb@cyllene.uwa.edu.au
Tel. +61 8 6488 4509
Fax. +61 8 6488 1029

Keywords: wildlife forensics, real-time PCR, DNA, melt-curve

Running head: Gel-free species identification

Abstract

DNA-based identification of organisms is an important tool in bio-security, ecological monitoring, and wildlife forensics. Current methods usually involve post-PCR manipulations (e.g. restriction digest, gel electrophoresis), which add to the expense and time required for processing samples, and may introduce error. We developed a method of species identification that uses species-specific primers and melt-curve analysis, and avoids post-PCR manipulation of samples. The method was highly accurate when trialled on DNA from six large carnivore species from Tasmania, Australia. Because of its flexibility and cost-effectiveness, this method should find wide use in many areas of applied biological science.

Introduction

Species identification is a focus of biological surveys, human and wildlife forensics, ecological monitoring and biosecurity. Historically, species have been identified from morphological characters. However, this approach is impractical in many situations due to time constraints, the need for specialized taxonomic knowledge, and the difficulty of species identification from partial or trace samples. DNA-based techniques based on the polymerase chain reaction (PCR) have brought tremendous advances in the efficiency and accuracy of species identification and are routinely used in many fields (e.g. Baker & Palumbi, 1994; Liu *et al.*, 1997; Mills *et al.*, 2000).

Most DNA-based techniques for species identification involve multiple post-PCR manipulations of samples, such as sequencing reactions, restriction digests, and gel electrophoresis before results can be obtained. These manipulations add time and cost to the processing of samples, and increase the chances of introducing human error and contamination. Here, we present a method for identifying species from DNA samples that offers advantages over existing methods because it requires no post-PCR manipulation of samples. The method utilizes species-specific primers and melt-curve analysis (MCA) to distinguish species. Briefly, DNA fragments are PCR amplified from an unknown DNA sample using a pool of species-specific primers. Species are then identified by the diagnostic melt temperature (T_m) of their DNA fragment. The method makes use of the dsDNA-specific dye SYBR Green I (Molecular Probes), which shows a rapid loss of fluorescence as the temperature is raised above the sample's melt temperature. Since the T_m of a DNA fragment is determined by its nucleotide content and length, it is possible to design primers that will amplify DNA fragments with species-diagnostic melting temperatures.

We illustrate the method with an assay that distinguishes the six large (>800 gram) mammalian carnivores that occur on the island of Tasmania, Australia. The species targeted include three native marsupials: tiger quoll (*Dasyurus maculatus*), eastern quoll (*Dasyurus viverinus*), Tasmanian devil (*Sarcophilus harrisii*), and three exotic eutherian species: feral cat (*Felis catus*), feral dog (*Canis familiaris*), and the European red fox (*Vulpes vulpes*). Monitoring the distribution of these species is of particular interest to wildlife management agencies in Tasmania for the purposes of conservation of biodiversity and the protection of agricultural industries (Resource Planning and Development Commission, 2003). The methods described here form part of a larger project that we are undertaking, which aims to monitor these species remotely from trace DNA samples.

Materials and Methods

Short (157-176 nt) PCR fragments were designed to be amplified between a universal anchor primer within the tRNA-phenylanaline of the mitochondrial DNA genome and species-specific primers for each target species located in the adjacent 12s rRNA gene (Figure 1). We maximized mismatches at the 3' end of the species-specific primers to ensure their species-specificity. Uncorrected genetic distance between entire fragments (including primers) ranged from 4.0% between the eastern and tiger quolls, to 28.7% between the cat and Tasmanian devil. 5' Poly-C tails were added to some primers to facilitate greater resolution of fragments during MCA (see Figure 1 and Results).

We first confirmed the species-specificity of the primers by trialling each pair on DNA from the other candidate species with conventional PCR and agarose electrophoresis. Using a gradient of annealing temperatures, we determined that all primers were species-specific above 56˚C (see Shivji *et al.*, 2002 for a protocol for testing species-specific primers). We used a Platinum® SYBR® Green qPCR SuperMix UDG (Invitrogen, CA, USA) to conduct real-time PCR and melt curve analysis. Reactions were conducted on a Rotorgene 3000 real-time machine (Corbett Research) and results analysed using Rotor-Gene 6 software. The 25 microlitre reactions consisted of 12.5 µL SuperMix (containing 0.03 units Platinum® *Taq*, 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 3 mM $MgCl_2$, 0.8 mM dNTPs), 0.4 µM of each primer (7 primers in total), 50 ng DNA, and water. Cycling conditions consisted of an initial incubation of 50˚C for 2 minutes, then 95˚C for 2 minutes, followed by 35 cycles of 95˚C 15 seconds, 57˚C 30 seconds, 72˚C 30 seconds. This was immediately followed by the melting profile, which consisted of an initial annealing of 57˚C for 45 seconds followed by temperature ramping at 0.5˚C per step with a 20 second pause at each step.

We conducted melt-curve assays on multiple individuals of each species (N = 11-29 per species, Figure 2). We also repeated the assay for five examples of each species five times to check for within-individual repeatability. In addition, because our method potentially has application to trace DNA samples, we ran assays on a dilution series of template DNA. For each species we ran DNA from two individuals at 1/10 and 1/100, 1/200, and 1/400 dilutions (50, 5, 0.5, 0.25, 0.125ng DNA template). We also trialled the method on DNA from 2 day old faeces (scats) extracted using methods described in Banks et al. (2002). Scats were available only from the fox, tiger quoll (from the Australian mainland), the Tasmanian devil, cat and dog. Finally, we conducted a blind trial using DNA from four individuals of each species. The cat and dog DNA samples were of mixed breed (Melbourne University Veterinary Hospital), except for a single dingo specimen from New South Wales, Australia. The other species were sampled from their total geographic range within Tasmania and mainland Australia.

Results and Discussion

The MCA approach to species identification proved highly effective. Each carnivore species tested produced a highly consistent melting temperature (Figure 2; coefficient of variance 0.06 - 0.12%, maximum intra-specific difference 0.35˚C), which ranged between a mean of 80.06˚C for the eastern quoll and a mean of 82.94˚C for the domestic cat. These results also held for assays on highly dilute samples and highly dilute and/or degraded scat DNA. Repeated assays on DNA from five individuals of each species also produced highly consistent T_ms (average difference between replicates 0.07˚C, range 0-0.2˚C), which is consistent with results from SNP genotyping studies that used the MCA approach (Wittwer *et al.*, 2003).

The mean T_m for each species (determined above) was an accurate species-diagnostic "benchmark", for all samples in the blind trial except the quolls, which could be distinguished from all other species, but not from one another. The two quolls consistently differed in their relative T_m within each MCA run. Failure to distinguish them was due to the small magnitude of the difference in their T_ms (*ca.* 0.15˚C within runs), which was similar in magnitude to that observed within-individuals, between MCA runs (average 0.16˚C ± 0.03SE, n=15), making absolute T_m benchmarks problematic. We found that inclusion of positive controls was an effective way to account for inter-run T_m variation in these species. When positive controls were included and their T_m used as a species-diagnostic benchmark, all species, including the quolls were correctly identified. Therefore, although our results suggest that a mean T_m difference of 0.35˚C is sufficient to accommodate the combined inter-run and within-species variation (here all species except the quolls), correct identification when differences are as low as 0.15˚C can be ensured by including positive controls for the species of interest. These values may vary according to the precision of the hardware used for MCA.

The ability to *a priori* predict whether DNA fragments will have species specific T_ms would greatly improve the efficiency of designing an assay for new species. One approach to predicting whether species will have diagnostic T_ms is to use a T_m predicting algorithm. We found a strong correlation between observed and T_m predicted by a nearest-neighbour algorithm (Breslauer *et al.*, 1986; r = 0.94). However, the difference between observed and predicted T_m was usually larger than the difference in T_m we observed between species (e.g. as low as 0.15˚C between the two quoll species), indicating that predictive algorithms cannot provide the precision necessary to identify whether DNA fragments will have species-diagnostic T_ms where those T_ms are very similar.

Given the above result, it is likely that some empirical testing will be required to determine the species-specificity of T_ms. Our initial experiments showed that the dog DNA fragment without a poly-C tail had an identical T_m to fox. We found that adding a C_{10} tail to the dog-specific primer was an effective way to increase the T_m of the product without requiring redesign of primer sites. Experiments adding C5, C10 and C15 tails to the cat and the dog primers showed that T_m increased by approximately 1 degree per five C nucleotides added (data not shown). In our assay, DNA fragments for each species were of similar size (excluding poly nucleotide tails), so that their T_m was predominantly determined by their nucleotide content. A more effective approach to ensuring that species' T_ms don't overlap would be to place primers so that fragments vary in length as well as nucleotide content. Ribosomal genes such as the 12s ribosomal RNA sequence used here make this relatively straightforward because their stem and loop structure, which includes indels, offers interspersed areas of conservatism for anchor primer placement, and variable regions, which make product's T_ms differ and can place species specific primers. Some nuclear genes may also be useful targets, particularly those with multiple copies such as ITS-1, as these would facilitate use with trace DNA samples.

When empirically determining the species-diagnostic "benchmark" T_ms, the choice of appropriate standards for each species is of major importance. These need to adequately sample the intra-specific genetic variation for each species, so should be sourced from throughout the relevant geographic range. The significance of this is illustrated by trials of our assay on tiger quoll specimens from mainland Australia. These assays showed extensive overlap in T_m with both the Tasmanian eastern and tiger quolls (Figure 2D). This result likely reflects the previously described large phylogeographic split between mainland and Tasmanian tiger quolls and the large amount of intraspecific genetic variation present in tiger quolls throughout the entire eastern seaboard of Australia (Firestone *et al.*, 1999). Despite this, our assay would still have utility on mainland Australia because eastern quolls are extinct there.

In principle, the MCA approach to species identification allows efficient use of both time and consumables. For example, because MCA requires no post-PCR processing, the total time taken for PCR plus MCA should be less than for methods that involve agarose electrophoresis, and significantly less than for direct sequencing. Processing is also simplified by associated software that enables genotypes to be called automatically. Savings will also come from the use of fewer consumables. We estimate that the cost of consumables per 25 µL reaction was approximately AUD$1.30 (€0.77 Euros, US$0.94) and this may be reduced by dispensing with the pre-fabricated PCR kit and obtaining the reagents separately. In our experience, this cost is comparable to or less than that required for the more time-consuming methods of species identification. The method will also be less expensive than assays based on sequence-specific molecular beacons such as Taqman probes (Applied Biosystems) because it does not require a customised probe to be designed and manufactured for each species of interest (e.g. Guiver *et al.*, 2001). Rather, so long as complementary species-specific primers with unique melting temperatures can be designed, new species can be added to the multiplex when required.

In summary, we have demonstrated that MCA can provide highly accurate and efficient species identification. Therefore, it should find a range of applications in applied biological research, such as fisheries fraud and wildlife forensics (Baker & Palumbi, 1994; Paxinos *et al.*, 1997; Shivji *et al.*, 2002).

Acknowledgements.

We thank the University of Canberra, The Fox-Free Tasmania Taskforce, Parks ACT, and Pestat Pty Ltd for financial support. We also thank Niccy Aitken for discussions, and Karen Firestone, Monica Ruibal, Menna Jones, David Pemberton, Paul de Torres, staff at the Melbourne University Veterinary School for tissue samples, and staff at the Caversham Wildlife Park for faecal samples. Comments from Gordon Luikart greatly improved the manuscript.

Literature cited

Baker CS, Palumbi SR (1994) Which whales are hunted? A molecular genetic approach to monitoring whaling. *Science* **265**, 1538-1539.

Banks SC, Piggott MP, Hansen BD, Robinson NA, Taylor AC (2002) Wombat coprogenetics: enumerating a common wombat population by analysis of faecal DNA. *Australian Journal of Zoology* **50**, 193-204.

Breslauer KJ, Frank R, Blocke H, Marky LA (1986) Predicting DNA duplex stability from the base sequence. *Proc. Natl. Acad. Sci. USA* **83**, 3746-3750.

Firestone KB, Elphinstone MS, Sherwin WB, Houlden BA (1999) Phylogeographical population structure of tiger quolls *Dasyurus maculatus* (Dasyuridae: Marsupialia), an endangered canivorous marsupial. *Molecular Ecology* **8**, 1613-1625.

Guiver M, Levi K, Oppenheim BA (2001) Rapid identification of candida species by TaqMan PCR. *Journal of Clinical Pathology* **54**, 362-366.

Liu W, Marsh TL, Cheng H, Forney LJ (1997) Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16s rRNA. *Applied and Environmental Microbiology* **63**, 4516-4522.

Mills LS, Pilgrim KL, Schwartz MK, McKelvey K (2000) Identifying Lynx and other North American Felids based on mtDNA analysis. *Conservation Genetics* **1**, 285-288.

Paxinos E, Mcintosh C, Ralls K, Fleisher R (1997) A noninvasive method for distinguishing among canid species: amplification and enzyme restriction of DNA from dung. *Molecular Ecology* **6**, 483-486.

Resource Planning and Development Commission (2003) State of the Environment Tasmania 2003.

Shivji M, Clarke S, Pank M*, et al.* (2002) Genetic identification of pelagic shark body parts for conservation and trade monitoring. *Conservation Biology* **16**, 1036-1047.

Wittwer CT, Reed GH, Gundry CN, Vandersteen JG, Pryor RJ (2003) High-resolution genotyping by amplicon melting analysis using LCGreen. *Clinical Chemistry* **49**, 853-860.

Figure Captions

Figure 1. Alignment showing positions of primers (boxed) and partial sequences amplified using the MC-species-ID assay described. Triangles mark the part of total fragment that is not shown (approximate positions 57 and 107), and * indicate positions of conserved sequence. 5’ primer is universal and 3’ primers are species-specific (complements to the species-specific primers used for the assay are shown).

Figure 2. Melt curve results for Tasmanian carnivorous mammals. X axis shows temperature, Y axis shows results from various experimental treatments (A-E). (A) Dilution series, from bottom to top 50, 5, 0.5, 0.25, 0.125ng DNA template used in PCR (n=2 for each species). Star symbol indicates result obtained from scat DNA. (B) Five individuals of each species, each tested five times. (C) Multiple individuals of each species, one test each (left to right, n = 29, 23, 25,22, 12, 11). (D) Mean ± 95% confidence intervals for each species, triangle symbol represents data from mainland *D. maculatus*. (E) Examples of melt curves for each species. For melt curves the Y axis scale is the smoothed negative first derivative of fluorescence with respect to temperature. It can be interpreted approximately as the rate of change in fluorescence as temperature increases.

```
                                                                 ▼     ▼
5'                                                                                                                                           3'
Dviv                          TTAGATGGATTCTAATAAATCCCATAAACACAAA-----TTTCCGCATCCCAGTGAGTATGCCCTTTTAGCTTTCATGAAGTATAAAGGAGCTG
Dmac                          TTAGATGGATTCTAATAAATCCCATAAACACAAA-----TTTCCGCACCCCAGTGAGTATGCCCTTTTAACTGTTACTAAGTATAAAGGAGCTGG
Shar  TAAAGCAARGCACTGAAAATGC  TTAGATGGATTCTAATAAATCCCATAAACACAGA-----TTTCCGCACCCCAGTGAGTATGCCCTTTTAACTCCATTAAAGTATAAAGGAGCT
Cfam                          CAAGATGAGTCGCAC--GACTCCATAAACATAAA-----CCTCCACGCCCCAGTGAGAATGCCCTTAAAATCACC-AGTGATCTAAAGGAGCAGGGGGGGGGG
Fcat                          CTAGATGAGTCGCCA--GACTCCATAAACACAAA-----CCTCCGCATCCCGGTGAAAATGCCCTCTAAGTCACCCAGTGACCTAAAGGGGGGGGGG
Vvul                          CAAGATGAGTCGCAC--GACTCCATAAACACAAA-----CCTCCACGCCCCAGTGAGAATGCCCTTAAAATCGCT-AACGATCTAAAGGAGCA
                                *****  *        *  ********* * *       *** *  *** ****  *******   *              ***
```

Figure 1.

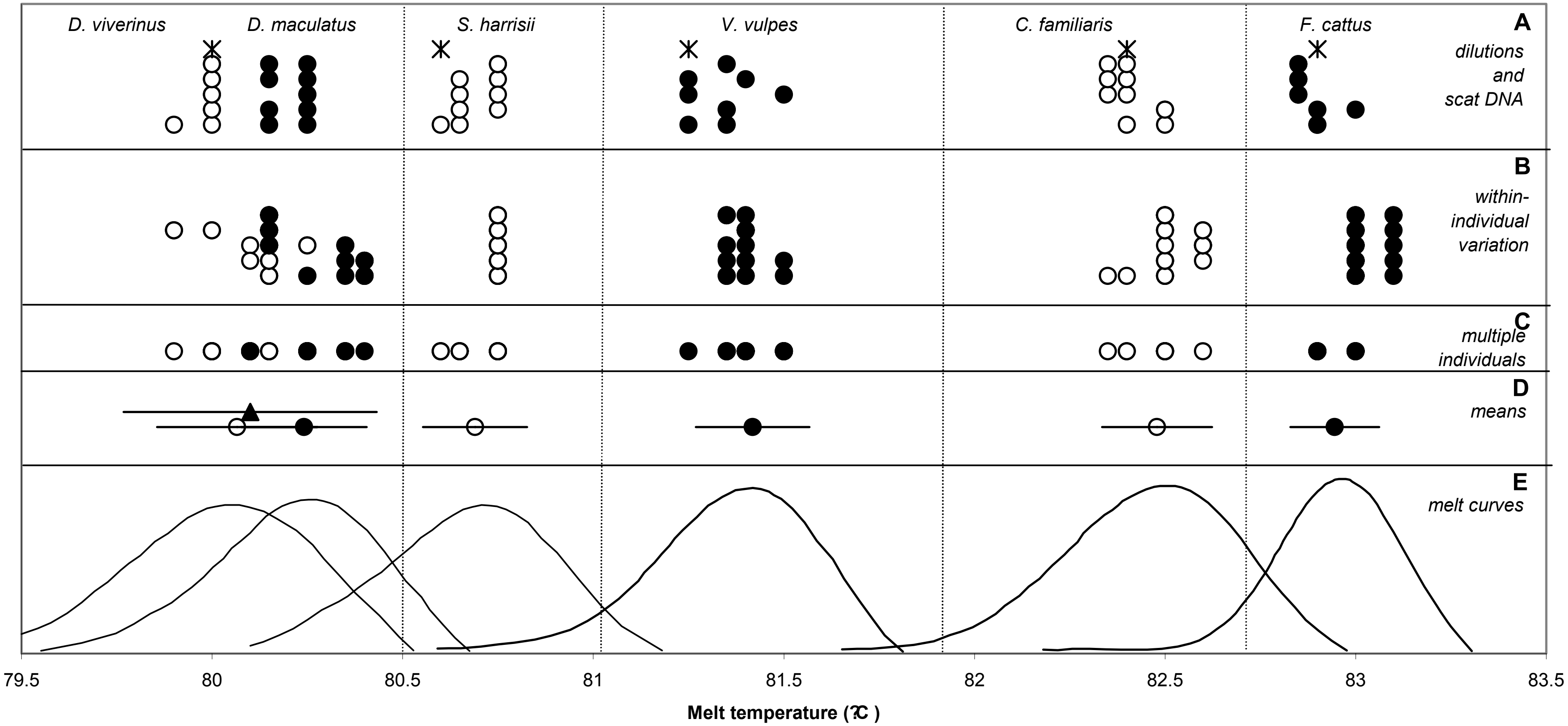
D. viverinus
D. maculatus
S. harrisii
V. vulpes
C. familiaris
F. cattus
A
dilutions and scat DNA
B
within-individual variation
C
multiple individuals
D
means
E
melt curves
79.5
80
80.5
81
81.5
82
82.5
83
83.5
Melt temperature (°C)

Figure 2