Mammalian Species Identification Using The Cytochrome *b* Gene On The Mitochondria For Forensic Applications

University of Strathclyde

S.S. Tobe; A.M.T. Linacre
Centre for Forensic Science, Department of Pure and Applied Chemistry
University of Strathclyde, Glasgow, UK

University of Strathclyde

INTRODUCTION:

Non-human biological evidence is encountered routinely in forensic casework. This evidence is often left unanalysed due to lack of a standard method. When analysis is undertaken it usually involves sequencing all or part of the mitochondrial genome which is then compared to known sequences on GeneBank. This technique is accepted, but has many problems including: being cost, time and labour intensive due to the extra step of sequencing products; mixtures cannot be separated; and degraded samples may not generate sufficient sequence data.

Therefore, ongoing research into the identification of non-human trace evidence is being undertaken throughout the forensic community to establish a quick, easy and cheap means to identify species. Both nuclear and mitochondrial genomes have been examined and tests have been developed based on both. The nuclear genome is more discriminatory, able to individualise animals, but is close to 400 thousand times larger than the mitochondrial genome (Figure 1a). In most cases individualization of the animal in question is not required and only the identity of the species in question is needed. Therefore, most of the tests recently developed rely primarily on the mitochondrial genome due to its resilience to environmental insult, ubiquity in the body and high copy number per cell (Figure 1b).

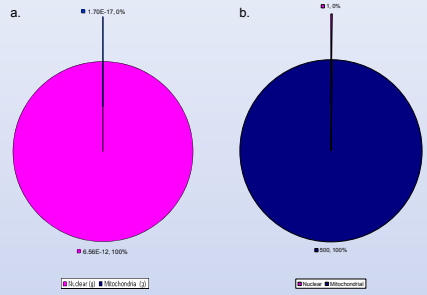

Figure 1:a) The nuclear genome compared to the mitochondrial genome by weight. b) The nuclear genome compared to the mitochondrial genome by copy number per cell. 500 is used as an average value but this number can vary from several hundred copies to several thousand copies per cell.

The use of the cytochrome *b* gene (cyt*b*), located on the mitochondrial genome and well established for taxonomic purposes, is often used in species identification. . The full cyt*b* gene is 1140 base pairs (bp) in size and is well known to show a high degree of interspecies variation, but little intraspecies variation (1). Usually all or part of the cyt*b* gene is amplified, sequenced and compared to an online database, such as NCBI (http://www.ncbi.nlm.nih.gov) for positive identification. Problems with this technique are that it is time and labour intensive, mixtures cannot be separated and degraded samples may not generate sufficient sequence data.

Based on the variation in the sequences it is possible to develop species specific primers that will only react for the species for which they were designed. This would decrease the time and cost of the test by removing the sequencing step, allow separation of mixtures and allow identification of highly degraded samples since large fragments of intact DNA would not be required.

MATERIALS AND METHODS:

Sequence Information and Primer Design

Cytochrome *b* sequence information for the sixteen animals (*F. catus, B. taurus, C. lupus familiaris, V. vulpes, C. hircus, C. porcellus, M. minutes, E. europaeus, E. caballus, M. musculus, H. sapiens, S. scrofa, O. cuniculus, R. norvegicus, C. elaphus* and *O. aries*) were downloaded from the NCBI website and aligned using Clustal W (http://www.ebi.ac.uk/clustalw/). Sequences were then inspected for variations that could be useful for primer design. Once an potential species specific priming site had been found, the sequence for that site was entered into the Basic Local Alignment Search Tool (BLAST) which compared it to every other sequence in the database. In this way, it could be insured that the species specific primers should not react with any other animal but the one for which they were designed. Universal primers from Pääbo *et al.* (2) were altered slightly and a new universal priming site was found to react with the species specific reverse primers.

Table 1: Universal and species specific primers, their location on the cytochrome *b* gene and fragment length

Animal	Location*	5'----------Sequence----------3'	Fragment Length (bp)
Universal 1†	-50	FAM6-GACCAATGATATGAAAAACCATCGTTGT	N/A
Universal 2†	400	FAM6-TGAGGACAAATATCATTYTGAGGRGC	N/A
Universal 3	832	TET-TAYGCHATYYTHCGNTCHATYCCHAAYAAA	N/A
V. vulpes	39	AAGGTCGATGAATGAGTCGTTTACG	106
R. norvegicus	50	GGGCGGGGAGGTCGATAAAGG	114
C. porcellus	85	GAGGGAGCCGAAGTTTCATCACGT	158
E. europaeus	499	GTTAGAGTAGCTTTGTCAACTGAAAATGA	120
C. elaphus	567	AGTAAGTGTACTATAGCGAGTGCTGCG	194
O. cuniculus	571	AAAGAGGAGGTGAATTAAGACTAAAGT	192
S. Scrofa	580	CGTGCAGGAATAGGAGATGTACGGC	199
F. catus‡	615	GGATGTAATTCCTGAGGGGTTGTTA	232
H. sapiens‡	624	ATCGGAATGGGAGGTGATTCCTAGG	241
B. taurus‡	666	TAAGATGTCCTTAATGGTATAGTAG	282
C. lupus familiaris‡	684	GAGTAGGAGTAAGGCTCCTAGGATA	303
C. hircus	693	TTAGAACAAGAATTAGTAGCATGGCG	313
E. caballus	705	AGAATAATACTAGAGTTAGTAGGAGCAAGATC	333
O. aries	906	GCTTTGATGTATGGAGGAGGGGTATAATT	096
M. minutus	960	TACAAGGATTCAGTAAAGTGTTTGGGAG	152
M. musculus	996	CTCCAATTCAGGTTAAGATAAGT	185

† From Pääbo *et al.* (1988) (2) ‡ From Panvisavas (2004) (3)
*in relation to the start of the cytochrome *b* gene which was given the designation of 1
bp = base pairs, A = adenine, T = thymine, C = cytosine, G = guanine, R=A/G, Y=C/T, H=A/T/C, N=A/T/C/G

DNA extraction and PCR amplifications

DNA was extracted from blood, tissue, hair, or buccal cells using the QIAamp® Micro Kit (Qiagen). Hair samples were allowed to digest in Proteinase K and DTT for up to 48 hours. PCR amplification was performed in a 20 µL reaction mixture, which contained 2-5 ng genomic DNA, reaction buffer (200 mM Tris-HCl (pH 8.4), 500 mM KC), 50 mM MgCl2, 1.5 unit of Platinum *Taq* DNA polymerase (Invitrogen) and primers. Amplification was conducted on an ABI 2720 Thermal Cycler with the following cycling conditions: 32 cycles of 95 ºC for 45 sec, 60 ºC for 45 sec and 72 ºC for 60 sec, then 72 ºC for 20 min for further extension. Analysis was conducted on an ABI PRISM 310 Genetic Analyser.

RESULTS:

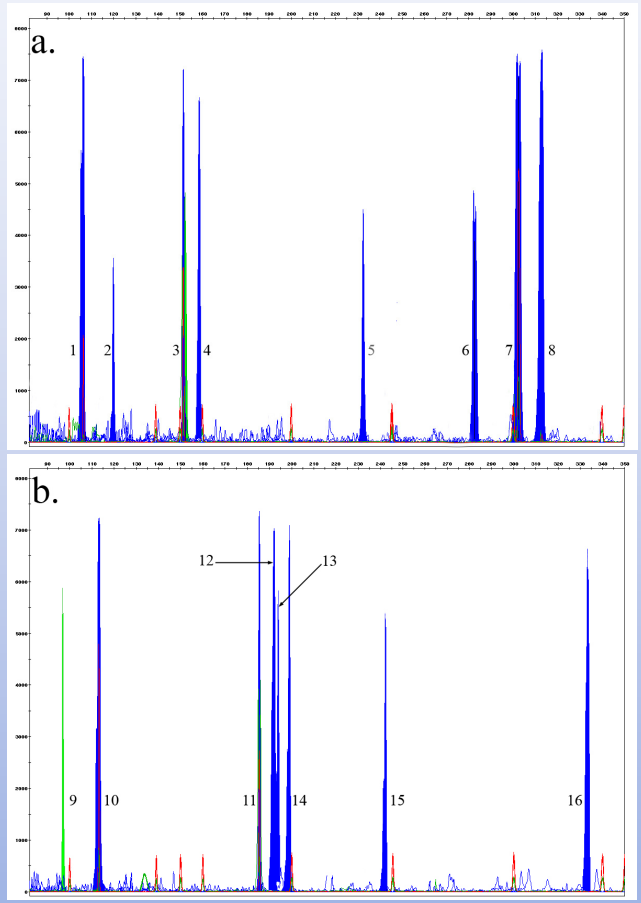

Figure 2: Peaks for the 16 different animals. a) 1. Fox (106 bp); 2. Hedgehog (120 bp); 3. Harvest Mouse (152 bp); 4. Guinea Pig (158 bp); 5. Cat (232 bp); 6. Cow (282 bp); 7. Dog (303 bp); and 8. Goat (313 bp). b) 9. Sheep (96 bp); 10. Rat (114 bp); 11. House Mouse (185 bp); 12. Rabbit (192 bp); 13. Red Deer (194 bp); 14. Pig (199 bp); 15. Human (241 bp); and 16. Horse (333 bp).

CONCLUSIONS:

Primers reacted only with the animal for which they were designed resulting in peaks of expected size (in bp) that were easily distinguishable from the other species tested. Similarly sized peaks such as rabbit (192 bp) and red deer (194 bp) can be easily distinguished even in the case of a mixture. The test is simple, fast, cheap and the results are easily interpreted making the test ideal for use in forensic casework.

Potentials for this test include identification of: components in food, which may have religious implications; dog and fox mixtures (contrary to fox hunting laws in the UK); dog or cat hair used in fur coats; and any other mammalian components in mixtures. The test also has the potential for more species to be added at a later time, including CITIES listed species, and expansion to include other genes on the mitochondria.

REFERENCES:

1. Linacre A. Application of Mitochondrial DNA Technologies in Wildlife Investigations – Species Identification. Forensic Science Review. **17(2)**, 2005.

2. Pääbo, S., J.A. Gifford and A.C. Wilson, Mitochondrial DNA sequences from a 7000-year old brain. Nucleic Acids Research. **16(20)**, 1988.

3. Panvisavas N. Species Testing by DNA Analysis. Unpublished Thesis. Centre for Forensic Science, Strathclyde University. 2004.