Int J Legal Med (2011) 125:27–32
DOI 10.1007/s00414-009-0395-y

ORIGINAL ARTICLE

DNA-based identification of forensically important Australian Sarcophagidae (Diptera)

Kelly A. Meiklejohn · James F. Wallman · Mark Dowton

Received: 20 March 2009 / Accepted: 12 November 2009 / Published online: 8 December 2009

Abstract The utility of the forensically important Sarcophagidae (Diptera) for time since death estimates has been severely limited, as morphological identification is difficult and thermobiological histories are inadequately documented. A molecular identification method involving the sequencing of a 658-bp 'barcode' fragment of the mitochondrial cytochrome oxidase subunit I (COI) gene from 85 specimens, representing 16 Australian species from varying populations, was evaluated. Nucleotide sequence divergences were calculated using the Kimura-two-parameter distance model and a neighbour-joining phylogenetic tree generated. All species were resolved as reciprocally monophyletic, except *Sarcophaga dux*. Intraspecific and interspecific variation ranged from 0.000% to 1.499% (SE=0.044%) and 6.658% to 8.983% (SE=0.653%), respectively. The COI 'barcode' sequence was found to be suitable for the molecular identification of the studied Australian Sarcophagidae: 96.5% of the examined specimens were assigned to the correct species. Given that the sarcophagid fauna is poorly described, it is feasible that the few incorrectly assigned specimens represent cryptic species. The results of this research will be instrumental for implementation of the Australian Sarcophagidae in forensic entomology.

Keywords Sarcophagidae · Diptera · Forensic entomology · COI · DNA 'barcoding' · Identification

Introduction

Insect evidence can prove valuable in estimating the minimum time since death or the postmortem interval (PMI) in forensic investigations. If accurate, PMI estimation can narrow the field of suspects and aid in the identification of the deceased. Initial corpse colonisers (such as dipterans, flies) are used preferentially over late colonisers (such as coleopterans, beetles) to correctly estimate the PMI due to their presence during primary corpse decomposition [1–4].

The Sarcophagidae or flesh flies (Diptera) comprise over 2,500 species in over 100 genera globally, with many species being carrion breeders and initial corpse colonisers [5, 6]. Such species have the potential to be used to estimate the PMI [1–4, 7]. Additionally, the reproductive cycle of the sarcophagids makes them prospectively more reliable for PMI estimations compared with other initial dipteran colonisers (Calliphoridae: blow flies) [4, 8, 9]. Sarcophagids are viviparous, depositing live larvae directly onto a meat source, providing instantly developed immatures for corpse decomposition [6, 8, 10]. By contrast, most calliphorids are oviparous, laying eggs onto a meat source, which will only hatch if the correct environmental conditions are met [1–3]. Despite the potential for using

Electronic supplementary material The online version of this article (doi:10.1007/s00414-009-0395-y) contains supplementary material, which is available to authorized users.

K. A. Meiklejohn (✉) · J. F. Wallman
Institute for Conservation Biology and Environment Management, School of Biological Sciences, University of Wollongong, NSW 2522, Australia
e-mail: km988@uow.edu.au

M. Dowton
Centre for Medical Bioscience, School of Biological Sciences, University of Wollongong, NSW 2522, Australia

sarcophagids for PMI estimations, they are currently used much less than the calliphorids [1, 2, 4, 8].

Use of the sarcophagids for PMI estimation is limited as species are often not morphologically distinct, especially as immatures [4, 5, 8, 10]. Family level identification is straightforward in adults as all sarcophagids share common characteristics including grey to black longitudinal stripes on the thorax, a checkerboard or tessellated abdomen, a heavily bristled body and widely spread eyes in both sexes [5, 6, 8]. Species-level identification of adults is only possible by taxonomic experts, where meticulous examination of subtle morphological differences, including regional hair presence and colour, body pigmentation and bristle length, placement and abundance is necessary [5, 6, 8].

Aside from morphology, the use of sarcophagids as PMI estimators is restricted due to inadequate documentation of their thermobiological histories, especially for the Australian species. To implement the use of sarcophagids for PMI estimation, a method for easy yet accurate species-level identification at any life stage is required, followed by thermobiological studies. Molecular identification is an alternative method proposed to eliminate the issues surrounding correct species-level identification based solely on taxonomy. The mitochondrial genome has been broadly used for species-level identifications: cytochrome oxidase subunit I (COI) was found to be descriptive for the identification of Chrysomyinae (Diptera: Calliphoridae) [11] and more specifically the COI 'barcode' region was found to be the most reliable for distinguishing between Australian *Chrysomya* (Diptera: Calliphoridae), after testing a range of gene regions [12].

For these reasons, we investigated the ability of the COI 'barcode' region to accurately identify adult sarcophagid specimens from the Australian east coast. This study was aimed at assisting the future implementation of Australian Sarcophagidae as a tool in forensic investigations.

Materials and methods

Specimens

Sarcophagidae specimens were collected using decayed meat baits and hand nets at several locations across the Australian east coast (Table 1 of the Electronic supplementary material). Collected adult specimens were identified morphologically by KAM through the use of relevant taxonomic keys [13, 14]. Genitalia dissections were prepared and examined when necessary to validate the identification of a specimen. Specimens were collected directly into absolute ethanol and retained at 4°C in the Diptera collection of the School of Biological Sciences, University of Wollongong.

DNA extraction, amplification and sequencing

From each adult Sarcophagidae specimen, two legs were used for the extraction of total genomic DNA by the 'salting out' protocol [15]. The DNA was resuspended in 50 µl of fresh TE solution and stored at 4°C. The 658-bp 'barcode' region of COI was amplified using the primer combination described by Nelson and colleagues [12], LCO1490-L (5′-GGTCWACWAATCATAAAGATATTGG-3′) and HCO2198-L (5′-RAAACTTCWGGRTGWC CAAARAATCA-3′). Each 20 µl reaction mixture contained 16 mM $(NH_4)_2SO_4$, 67 mM Tris-HCl (pH 8.8 at 25°C), 0.01% Tween-20, 490 µM of deoxynucleotide triphosphates (dNTPs), 2.5 mM $MgCl_2$, 400 nM of each primer (Sigma, Castle Hill, NSW, Australia), 1.2 U of BIOTAQ™ DNA Polymerase (Bioline, Alexandria, NSW, Australia) and 1 µl of extracted total genomic DNA or distilled H_2O (negative control). The PCR temperature cycles, carried out on a Palm Cycler™ II (Corbett Research, Mortlake, NSW, Australia), consisted of an initial 2 min denaturation step at 94°C, followed by 30 cycles of denaturation at 94°C for 30 s, primer annealing across a 45–65°C temperature gradient for 30 s and elongation for 2 min at 72°C. A final 5 min incubation at 72°C allowed for the completion of any partially synthesised strands.

All PCR amplicons were resolved by agarose gel electrophoresis to identify the optimal annealing temperature (generally 50–52°C). These amplicons were treated with ExoSAP-IT® (GE Healthcare, Bucks, HP8 4SP, UK) as per the manufacturer's instructions, to digest unincorporated primers and dNTPs and were stored at 4°C.

Sequencing of the ExoSAP-IT® treated PCR products was performed using the ABIPRISM® BigDye™ Terminator Version 3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA). Reactions were setup according to the manufacturer's instructions, and PCR primers were used to initiate the sequencing reactions. Sequencing reactions were thermocycled using the Palm Cycler™ II, following the protocol described by Platt and colleagues [16].

Individual sequencing products were purified by combining 2 µl of EDTA (125 mM, pH 8), 2 µl of sodium acetate (3 M, pH 5.2) and 50 µl of absolute ethanol (95%) to the 20 µl sequencing reaction. The mixture was vortexed and incubated at room temperature for 15 min and then centrifuged at 16,100×g for 15 min. The resulting supernatant was removed, and the pellet was washed with 180 µl of ice-cold 70% ethanol and centrifuged for a further 5 min at 16,100×g. Again the supernatant was discarded, the pellet allowed to air-dry and stored at −20°C. The sequencing products were separated using an ABI 3130*xl* Genetic Analyzer (Applied Biosystems, Scoresby, VIC, Australia).

DNA sequence analysis

ChromasPro Version 1.33 (Technelysium, Tewantin, QLD, Australia; available online at www.technelysium.com.au/ChromasPro.html) was used to edit and confirm the sequence electropherograms. Each sequenced fragment was submitted to the Basic Local Alignment Search Tool (National Center for Biotechnology Information, Bethesda, MD, USA; available online at www.ncbi.nlm.gov/blast/BLAST) to verify sequence similarity to previously identified COI dipteran sequences. Mitochondrial gene sequences were translated into amino acid sequences, using the programme EMBOSS Transeq (available online at www.ebi.ac.uk/Tools/emboss/transeq/index.html). Mitochondrial gene sequences were aligned using BioEdit Sequence Alignment Editor Version 7.0.5.3 [17]. Sequences were entered into the Barcode of Life Datasystems Management and Analysis System (www.boldsystems.org) [18], where storage and preliminary data analysis was undertaken.

DNA 'barcode' analysis

Nucleotide sequence divergences for the sarcophagids were calculated using the Kimura-two-parameter (K2P) [19] distance model, available within *Phylogenetic Analysis Using Parsimony* (PAUP** and other methods) Version 4.0b10 [20]. A bootstrap (1,000 replicates) neighbour-joining (NJ) [21] analysis was also performed in *PAUP**, which provided a visual representation of the divergence between specimens. A semistrict consensus of the 1,000 bootstrap NJ trees was generated, only retaining groups that had occurred in >50% of the trees. Three species of Calliphoridae (Diptera; *Calliphora augur*, *Chrysomya rufifacies* and *Lucilia cuprina*) were included as the out-groups.

Results and discussion

Amplification and sequencing of the COI 'barcode' region

The 'barcode' region of COI was straightforward to both amplify and sequence. A 658-bp fragment of the COI gene was sequenced from 85 specimens, representing 16 species of Australian Sarcophagidae. Electropherograms did not contain sequence ambiguities, and the translated sequences did not contain premature stop codons, features consistent with *bona fide* mitochondrial sequences and not those of nuclear pseudogenes. The absence of insertions and deletions allowed the sequences to be aligned with ease. From this alignment, an entropy plot, which highlights conserved and variable nucleotide regions within an alignment, was generated in BioEdit. This plot indicated that variation was relatively evenly distributed along the fragment, such that any region could be targeted if degraded forensic material was investigated. Similarly to other dipteran families [12], the nucleotide composition showed an AT bias within the Australian Sarcophagidae (mean A=29.66%, T=37.02%, C=17.43, G=15.89%).

Neighbour-joining analysis of the Australian Sarcophagidae COI 'barcode' sequences

The purpose of this study was to investigate whether the 'barcode' region of COI provided sufficient resolution to identify morphologically indistinct Australian Sarcophagidae specimens. Effective species-level resolution for the Australian Sarcophagidae studied was obtained, as assessed by a NJ analysis based on K2P distances.

Individual sarcophagid specimens (from different populations) that had been identified as the same species morphologically were nearly always resolved within the same monophyletic group (96.5% of cases). This resolution is comparable to other 'barcode' studies [12, 22–27]. However, three exceptions/outliers were evident (highlighted by an asterisk; Fig. 1). Further examination of the morphology of the outlier specimens, especially the genitalia, was undertaken to verify the initial identifications. These initial identifications were confirmed, since no additional morphological features were noted that might explain the sequence divergence of the *Sarcophaga dux* and *Sarcophaga crassipalpis* outliers. Taxonomic descriptions of the Australian Sarcophagidae have not been updated since the 1950s, so it is plausible that the outlier specimens represent recently diverged cryptic species. Increased taxon sampling of *S. dux* and *S. crassipalpis* across Australia may help to resolve this issue. Interior branches for each of the monophyletic species have bootstrap values of 100, providing a preliminary indication that species-level resolution of the Australian Sarcophagidae is well supported.

In order to test the validity of the barcode approach, international sarcophagid sequences from species represented in the Australian taxon set and covering the entire barcode region were obtained from Genbank. Only two comparable sequences could be confirmed as international specimens, those being *S. crassipalpis* both from California, USA (GQ223337; AF259510, the latter is from Wells et al. [4]), and were included in a subsequent NJ 'barcode' analysis. These non-Australian representatives clustered with the other Australian *S. crassipalpis* specimens, consistent with barcode analysis of the Australian taxa.

K2P percentage divergences

Examination of the K2P percentage divergence between specimen sequences provides a quantitative approach for DNA 'barcoding' evaluation. Successful species-level

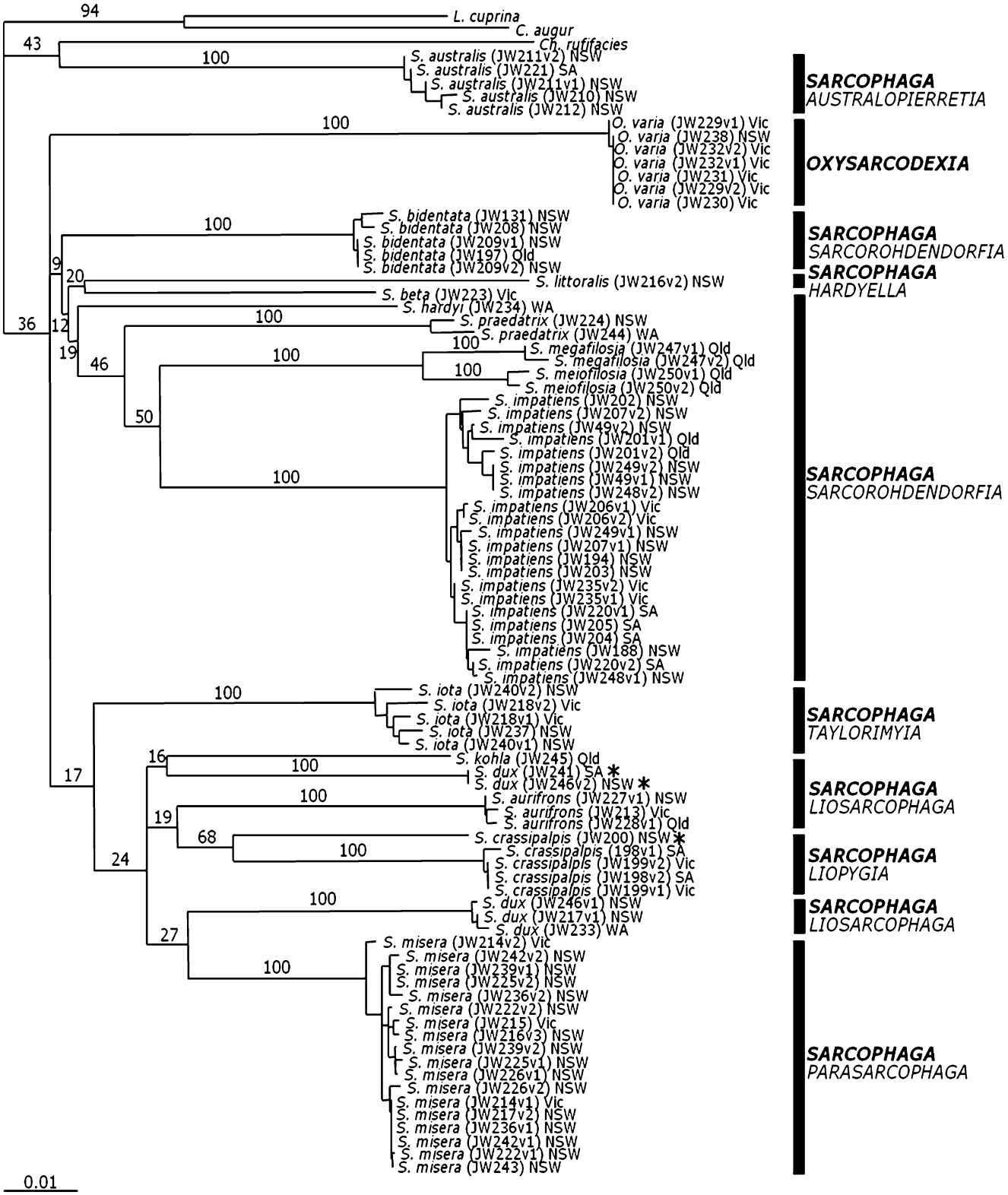

Fig. 1 Neighbour-joining (NJ) tree of Kimura-two-parameter (K2P) distances for 85 cytochrome oxidase subunit I (COI) gene sequences from 16 species of Australian Sarcophagidae. ***GENERA*** and *SUBGENERA* are given on the *right-hand side*. *Numbers* given at *main branches* refer to bootstrap proportions among 1,000 bootstrap replicates (internal monophyletic bootstrap values not shown). *Outliers* highlighted by an *asterisk*. Morphological species identification, voucher ID and state are given in specimen label. *Outgroup* consists of three species of Calliphoridae (*Calliphora augur*, *Chrysomya rufifacies* and *Lucilia cuprina*). Evolutionary distance divergence *scale bar* is 0.01

resolution is obtained by the 'barcoding' approach when intraspecific (within-species) percentage divergences are <3%, in comparison to interspecific (between-species) variation being >3% [25]. The intraspecific and interspecific percentage divergences obtained for the Australian Sarcophagidae species studied lie within these thresholds, as do those reported by Nelson and colleagues [12] for the *Chrysomya* blowflies, Hajibabaei and colleagues [24] for several Lepidoptera families and Smith and colleagues [28] for parasitoid flies.

Intraspecific variation

The intraspecific variation for the Australian Sarcophagidae studied ranged between 0.000% and 1.499% (SE=0.044%; Table 2 of the Electronic supplementary material). Interestingly, two species had divergences that fell outside the ideal range for effective resolution by the 'barcoding' approach. The intraspecific variation within the five specimens morphologically identified as *S. crassipalpis* was 2.752% (Table 2). However, on removal of the outlier specimen (JW200; highlighted by an asterisk), the percentage divergence dropped to 0.150% within the intraspecific 'barcode' threshold values. Additionally, the intraspecific variation within the nonmonophyletic species of *S. dux* as depicted from the NJ tree (Fig. 1) was 4.658% (Table 2). Based on 'barcode' thresholds, this value is indicative of two distinct species which is portrayed graphically by separate clusters in the NJ tree (Fig. 1).

Importantly, low intraspecific genetic divergence for each species was maintained despite the inclusion of individuals from distinct geographic populations. This is highlighted by *Sarcophaga impatiens* and *Sarcophaga misera*, where broader sampling (across a range of states) was possible. Such preliminary findings indicate that low intraspecific variation is a characteristic of Australian Sarcophagidae, with further taxon sampling likely to confirm this.

Interspecific variation

The range of interspecific variation for directly comparable Australian Sarcophagidae, as determined from examination of the NJ tree, was 6.658–8.983% (SE=0.653%), except between *Sarcophaga megafilosia* and *Sarcophaga meiofilosia* (Table 3 of the Electronic supplementary material). The 2.891% interspecific variation noted between *S. megafilosia* and *S. meiofilosia* suggests that these morphologically distinct specimens could have diverged very recently or possibly belong to the same species (based on the thresholds of the 'barcoding' approach). Both of these species were resolved as monophyletic on the NJ tree (Fig. 1), however, to confirm this positioning, further taxon sampling should be considered.

Additionally, the outliers were compared to the other specimens identified morphologically as the same species with each having divergences well outside the 'barcoding' thresholds (6.658% for the *S. crassipalpis* outlier and 7.653% for the *S. dux* outliers; Table 3 of the Electronic supplementary material). Such divergences suggest that these specimens represent distinct species.

The interspecific divergences for the Australian Sarcophagidae were effective in resolving species based on the DNA 'barcoding' approach. In other studies, different fragments of the COI gene were effective for species resolution of Sarcophagidae in both Europe [11] and the USA [4], with the interspecific variation between species being 2.7–4.7% and 3%, respectively.

Phylogenetic inference

Little is known of the phylogeny among the various genera, subgenera and species of the Australian Sarcophagidae. It should be noted that the evolutionary relationships depicted in this study (Fig. 1) should not be viewed as resolved. Bootstrap support is poor at higher level nodes and appropriate outgroup selection for the Sarcophagidae remains questionable. Most importantly, the NJ method used is effective for species-level resolution in 'barcoding' studies when sequence divergences are low, but performs more poorly than parsimony and likelihood methods for accurately resolving phylogenies. Future studies will investigate evolutionary relationships using a broader taxon sample and a more comprehensive set of phylogenetic approaches.

Conclusions

The Sarcophagidae fly fauna are neglected not only within Australia but globally and therefore, remain unexploited as PMI indicators in forensic entomology. The initial step needed to secure the use of sarcophagids in PMI estimations is the accurate species-level identification of a specimen from any life stage. The COI 'barcode' region was evaluated as a molecular marker for species-level resolution for 16 adult species of Australian Sarcophagidae. This approach proved successful for the accurate identification of the Australian Sarcophagidae examined and is likely to be a sensitive typing strategy for sarcophagids associated with forensic casework.

Once accurate species-level identification of a sarcophagid specimen had been done, relevant thermobiological studies would also be necessary to estimate the age for PMI estimates. Records of thermobiological profiles for sarcophagids are currently inadequate, especially in the Australian context. Further studies should, therefore,

examine the thermobiology of forensically important Australian species. Due to this, we are currently documenting the thermobiology of an endemic and forensically important Australian sarcophagid, *S. impatiens*. With continued research, the use of sarcophagids in forensic entomology should increase and their value as tools in criminal investigations realised.

Acknowledgments We would like to acknowledge Melanie Archer, Kelly George, Steve and Ruth McKillup and Lisa Mingari for providing specimens. We are grateful to the Australian Biological Resources Study (ABRS) for their financial support.

References

1. Amendt J, Krettek R, Zehner R (2004) Forensic entomology. Naturwissenschaften 91:51–65
2. Catts EP (1992) Problems in estimating the postmortem interval in death investigations. J Agric Entomol 9:245–255
3. Catts EP, Goff ML (1992) Forensic entomology in criminal investigations. Annu Rev Entomol 37:253–272
4. Wells JD, Pape T, Sperling FAH (2001) DNA-based identification and molecular systematics of forensically important Sarcophagidae (Diptera). J Forensic Sci 46:1098–1102
5. Pape T (1996) Catalogue of the Sarcophagidae of the world (Insecta: Diptera). Associated Publishers, Gainesville
6. Shewell GE (1987) Sarcophagidae. Research Branch, Agriculture Canada, Ottawa
7. Greenberg B (1991) Flies as forensic indicators. J Med Entomol 28:565–577
8. Byrd JH, Castner JL (2001) Insects of forensic importance. CRC, Boca Raton
9. Kamal AS (1958) Comparative study of thirteen species of sarcosaprophagous Calliphoridae and Sarcophagidae (Diptera). Ann Entomol Soc Am 51:261–271
10. Zehner R, Amendt J, Schutt S, Sauer J, Krettek R, Povolny D (2004) Genetic identification of forensically important flesh flies (Diptera: Sarcophagidae). Int J Legal Med 118:245–247
11. Wells JD, Williams DW (2007) Validation of a DNA-based method for identifying Chrysomyinae (Diptera: Calliphoridae) used in a death investigation. Int J Legal Med 121:1–8
12. Nelson LA, Wallman JF, Dowton M (2007) Using COI barcodes to identify forensically and medically important blowflies. Med Vet Entomol 21:44–52
13. Lopes HdS (1959) A revision of Australian Sarcophagidae (Diptera). Studia Ent 2:33–67
14. Lopes HdS (1954) Contribution to the knowledge of the Australian sarcophagid flies belonging to the genus "*Tricholioproctia*" Baranov, 1938 (Diptera). An Acad Bras Cienc 26:234–276
15. Sunnucks P, Hales DF (1996) Numerous transposed sequences of mitochondrial cytochrome oxidase I-II in aphids of the genus *Sitobion* (Hemiptera: Aphididae). Mol Biol Evol 13:510–524
16. Platt AR, Woodhall RW, George AL Jr (2007) Improved DNA sequencing quality and efficiency using an optimised fast cycle sequencing protocol. Biotechniques 43:58–62
17. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98
18. Ratnasingham S, Hebert PDN (2007) BOLD: The Barcode of Life Data System (www.barcodinglife.org). Mol Ecol Notes 7:355–364
19. Kimura M (1980) A simple model for estimating the evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120
20. Swofford DL (2001) *PAUP*-Phylogenetic Analysis Using Parsimony* (* and Other Methods). Sinauer Associates, Sunderland
21. Saitou N, Nei M (1987) The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425
22. Ball SL, Hebert PDN, Burian SK, Webb JM (2005) Biological identifications of mayflies (Ephemeroptera) using DNA barcodes. J North Am Benthol Soc 24:508–524
23. Foottit RG, Maw HEL, Von Dohlen CD, Hebert PDN (2008) Species identification of aphids (Insecta: Hemiptera: Aphididae) through DNA barcodes. Mol Ecol Resour 8:1189–1201
24. Hajibabaei M, Janzen D, Burns J, Hallwachs W, Hebert PDN (2006) DNA barcodes distinguish species of tropical Lepidoptera. Proc Natl Acad Sci USA 103:968–971
25. Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. Proc R Soc Lond B 270:313–321
26. Hebert PDN, Ratnasingham S, deWaard JR (2003) Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species. Proc R Soc Lond B 270:s96–s99
27. Ward R, Zemalk T, Innes B, Last P, Hebert PDN (2005) DNA barcoding Australia's fish species. Philos Trans R Soc Lond B Biol Sci 360:1847–1857
28. Smith MA, Woodley NE, Janzen DH, Hallwachs W, Hebert PDN (2006) DNA barcodes reveal cryptic host-specificity within presumed polyphagous members of a genus of parasitoid flies (Diptera: Tachinidae). Proc Natl Acad Sci USA 103:3657–3662