Molecular Phylogenetics and Evolution 50 (2009) 317–325

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Resolving genetic diversity in Australasian *Culex* mosquitoes: Incongruence between the mitochondrial cytochrome *c* oxidase I and nuclear acetylcholine esterase 2

Stéphane Hemmerter [a], Jan Šlapeta [b], Nigel W. Beebe [c,d,*]

[a] *Institute for the Biotechnology of Infectious Diseases, University of Technology Sydney, Ultimo, NSW 2007, Australia*
[b] *Faculty of Veterinary Science—B14, University of Sydney, NSW 2006, Australia*
[c] *School of Integrative Biology, University of Queensland, Room 170B, Goddard Building, St. Lucia, Qld. 4072, Australia*
[d] *CSIRO, Long Pocket Laboratories, Indooroopilly, Qld. 4068, Australia*

ARTICLE INFO

Article history:
Received 11 June 2008
Revised 27 October 2008
Accepted 6 November 2008
Available online 27 November 2008

Keywords:
Acetylcholine esterase 2
Cytochrome oxidase I
Barcode
Culex annulirostris
Cryptic species
Japanese encephalitis

ABSTRACT

Insects that vector pathogens are under constant surveillance in Australasia although the repertoire of genetic markers to distinguish what are often cryptic mosquito species remains limited. We present a comparative assessment of the second exon–intron region of the acetylcholine esterase 2 gene (ace-2) and the mitochondrial DNA cytochrome *c* oxidase I (COI) using two closely related Australasia mosquitoes *Culex annulirostris* and *Culex palpalis*. The COI revealed eight divergent lineages of which four were confirmed with the ace-2. We dissect out the nuclear chromosomal haplotypes of the ace-2 as well as the exon–intron regions by assessing the protein's tertiary structure to reveal a hypervariable 5'-exon that forms part of an external protein loop and displays a higher polymorphic rate than the intron. We retrace the evolutionary history of these mosquitoes by phylogenetic inference and by testing different evolutionary hypotheses. We conclude that DNA barcoding using COI may overestimate the diversity of *Culex* mosquitoes in Australasia and should be applied cautiously with support from the nuclear DNA such as the ace-2. Together the COI and ace-2 provide robust evidence for distinct cryptic *Culex* lineages—one of which correlates exactly with the southern limit of Japanese encephalitis virus activity in Australasia.

1. Introduction

Despite the importance of mosquitoes as vectors of pathogens that cause significant human and animal diseases, their genetic diversity still requires accurate detailing while an understanding of their phylogeography and dispersal has the potential to resolve longstanding taxonomic and epidemiological questions. *Culex annulirostris* Skuse (Diptera: Culicidae) is an important vector of the Japanese encephalitis virus (JEV) in Papua New Guinea, Indonesia and surrounding islands and territories (Russell and Dwyer, 2000). This morphospecies often occurs in sympatry with its closely related sister species *C. palpalis* and their similar morphology makes studies on these mosquitoes difficult (Beebe et al., 2002; Chapman et al., 2003, 2000). Allozyme studies and observations from a PCR diagnostic developed for these mosquitoes suggested the existence of a complex of cryptic species (Beebe et al., 2002; Chapman et al., 2003, 2000). Mitochondrial DNA cytochrome oxidase I (COI) studies supported these findings by identifying eight separate lineages within the COI sequences (Hemmerter et al., 2007). Importantly, mitochondrial lineages of *C. annulirostris* showed a southern limit matching that of JEV activity in Australasia, leading to the hypothesis that JEV's establishment may be dependent on a specific mtDNA lineage or cryptic taxa (Hemmerter et al., 2007). However, confirmation is needed to determine if the *C. annulirostris* and *C. palpalis* lineages identified are reproductively isolated.

The acetylcholinesterase 2 gene (ace-2) is currently the only protein coding nuclear marker used for the molecular identification of mosquitoes and has been used for discrimination of the *C. pipiens* complex (Bourguet et al., 1998; Kasai et al., 2008; Sanogo et al., 2007; Smith and Fonseca, 2004). The ace-2 codes for a crucial nerve response enzyme acetylcholinesterase (AChE). The function of synaptic AChE is to stop the neurotransmission in the central nervous system of insects by hydrolyzing the neurotransmitter acetylcholine into acetate and choline (Toutant, 1989). Due to AChE's function, it is the target of two major classes of insecticides used for pest management in agriculture and public health (Hemingway et al., 2004).

* Corresponding author. Address: School of Integrative Biology, University of Queensland, Room 170B, Goddard Building, St. Lucia, Qld. 4072, Australia. Fax: +61 7 3365 1655.
E-mail addresses: stephane.hemmerter@uts.edu.au (S. Hemmerter), jslapeta@usyd.edu.au (J. Šlapeta), n.beebe@uq.edu.au (N.W. Beebe).

1055-7903/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2008.11.016

The aim of this study was to identify a polymorphic region of ace-2 and test its usefulness as a phylogenetic marker for identification of cryptic taxa within *C. annulirostris* and *C. palpalis* in Australasia. We used the three-dimensional structure of *Drosophila melanogaster* AChE as the proxy to map the nucleotide diversity of the second intron to the flanking exons which code for a hypervariable external loop in *Culex* mosquitoes. The genetic diversity of ace-2 and COI is compared and the evolutionary history of these mosquitoes is retraced in a phylogenetic framework. We conclude that the external loop region of ace-2 is a useful marker to investigate diversity of *Culex* mosquitoes and strongly supports the existence of a novel divergent COI *Culex* lineage matching the southern limit of JEV activity in Australasia.

2. Material and methods

2.1. Specimen collection and identification

Mosquitoes were collected from a wide geographic range that included Australia, PNG, Timor Leste and the Solomon Islands (Bougainville to the north and Guadalcanal to the south) (Fig. 1, Supplementary Table S1). Adult mosquitoes were collected using CO_2-baited encephalitis virus surveillance traps both with and without 1-octen-3-ol (octenol). Specimens were morphologically identified using the available keys (Lee et al., 1989; Marks, 1982). Mosquitoes were then stored in either liquid nitrogen or on dry ice, on silica gel, or in 70–100% ethanol prior to DNA extraction.

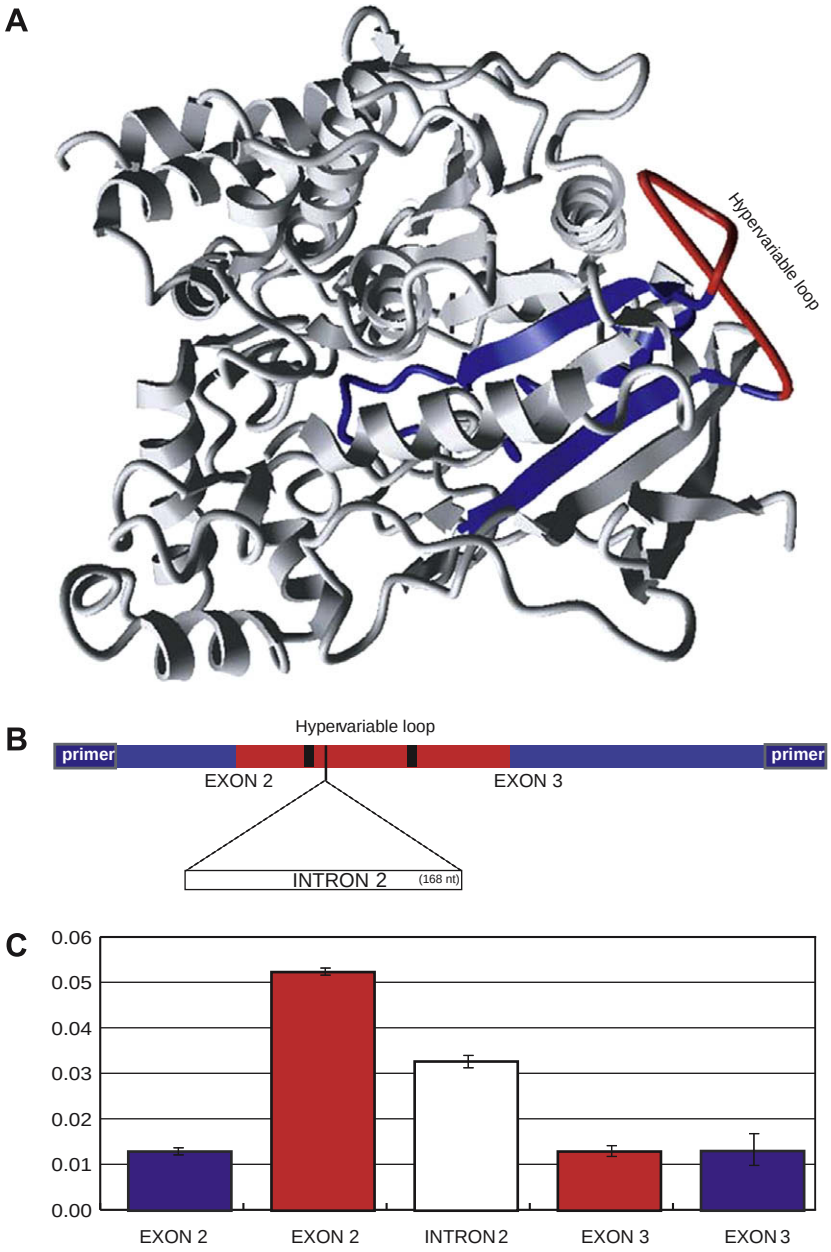

Fig. 1. Three-dimensional structure of the *Drosophila melanogaster* acetylcholinesterase, DmAChE (PDB: 1QO9) with the external hypervariable loop insertion added manually using YASARA (A). This "insertion" is not drawn to scale and is 36 amino acids long for the DmAChE. (B) Graphical representation of the ace-2 intron–exon region amplified in this study which is 392 residues. The vertical bars within the external loop region represent amino acid changes within *C. annulirostris* and *C. palpalis*. (C) Frequency of nucleotide diversity per site of the ace-2 sequence for *C. annulirostris* and *C. palpalis* relative to its position and functionality. The nucleotide diversity per site is showed as a histogram: blue and red = exon; white = intron; red = exon coding for the external loop of the ace-2 protein. Standard deviations are represented for each region.

Total DNA was extracted from mosquitoes using a salt extraction and ethanol precipitation procedure (Beebe et al., 2002). Due to problems associated with separating these mosquitoes using adult morphology, all material was genetically identified to species using an rDNA ITS1 PCR-RFLP that discriminates *C. annulirostris*, *C. palpalis* and *C. sitiens* (Beebe et al., 2002).

2.2. Amplification and sequencing of ace-2

A 382 bp fragment of the ace-2 was PCR amplified, including a partial sequence of the 3′-region of exon 2 and the 5′-region of exon 3 as well as 168 bp for the second intron. Primers were designed using an alignment of the ace-2 gene from the following *Culex* subgenus mosquitoes: *C. pipiens quinquefasciatus* (AY196911), *C. pipiens* (AY196910), *C. nigripalpus* (AY196914), *C. salinarius* (AY196913), *C. restuans* (AY196912), *C. pipiens pallens* (AY497524), *C. torrentium* (AY497525), *C. australicus* (AY497523), *C. pipiens* (AM159193) and *C. tritaeniorhynchus* (AB122151). The same primer pair was also used for both PCR amplification and sequencing; ACE2SHF (5′-CAC GAA CGT ATC CGA AGA CTG-3′) and ACE2SHR (5′-GCC ACG ATT ACG TTT CCA AC-3′). The 25 μl PCR reaction contained 1.75 mM $MgCl_2$, 200 pM for each dNTP, 0.4 μM for each primer, 0.6 U of *Taq* DNA polymerase (Fisher Biotech, WA, Australia) and approximately 1–20 ng of genomic DNA template. Cycling conditions included initial denaturing at 93 °C for 4 min followed by 35 cycles of 93 °C for 1 min, 52 °C for 1 min, 72 °C for 1.5 min with a final elongation for 10 min at 72 °C. The PCR product was size verified on 1.5% agarose gel and the remainder purified using a QIAquick PCR purification kit (Qiagen). Individual PCR products were directly sequenced in both directions on ABI3730*xl* by the Australian Genome Research Facility (University of Queensland, St. Lucia, Australia).

In total, 59 mosquitoes were sequenced for the ace-2 (Supplementary Table S1), of which 49 had previously been sequenced for the COI (Hemmerter et al., 2007). Individual sequences were assembled and edited using Sequencer 4.2.2 (GeneCodes, MI, USA). Base polymorphism was coded using the IUPAC (International Union of Pure Applied Chemistry) ambiguity codes and deposited in GenBank under the Accession Nos. EU710617–EU710675.

2.3. Amplification and sequencing of COI

In addition to the COI sequences used in a previous study and described in Supplementary Table S1 (Hemmerter et al., 2007), an additional 10 mosquitoes were sequenced for the COI. One haplotype was identical to a119 (Hemmerter et al., 2007) and nine were new haplotypes a135–a143 (Supplementary Table S1). Briefly, a 538 bp 5′-fragment of the COI gene was PCR amplified and sequenced using the same primer pair and procedure; F-COI50 (5′-GTA GTT TAG TAG AAA ATG GAG C-3′) and R-COI650 (5′-TAG CAG AAG TAA AAT AAG CTC G-3′) as previously described (Hemmerter et al., 2007). New sequences were deposited in GenBank under the Accession Nos. EU710676–EU710684.

2.4. Assembly and reconstruction of ace-2 haplotypes

The Bayesian statistical method PHASE v2.1 was used to reconstruct distinct nuclear haplotypes for ace-2 of all individuals according to published methods (Stephens and Scheet, 2005; Stephens et al., 2001). PHASE allowed the resolution of the ambiguities within the individual mosquitoes each comprising two haplotypes per genome. For each mosquito PHASE reported two haplotypes (Supplementary Table S1, Supplementary Alignment S1). The nuclear ace-2 sequences of *C. annulirostris* and *C. palpalis* were analyzed separately and—for consistency—five independent runs were executed using different seeds and a thinning interval of one with a burn-in of 100, as recommended in the PHASE documentation.

2.5. Alignment, composition and nucleotide diversity

Nucleotide sequences of ace-2 haplotypes from *C. annulirostris* and *C. palpalis* were aligned in MEGA4 (Kumar et al., 2004; Tamura et al., 2007). The nucleotide diversity of separate regions within the ace-2 sequences was calculated with DNASP v4.10 (Rozas et al., 2003). Composition of the nucleotides and translation to amino acids was performed using MEGA4 (Tamura et al., 2007). Both interspecific (interlineage) and intraspecific (intralineage) pairwise sequence divergences for the ace-2 and the COI were calculated using the Kimura's 2-parameter distance model (Kimura, 1980) in MEGA4 (Tamura et al., 2007). We used the highest intraspecific distance and the smallest interspecific distance rather than means (Meier et al., 2008).

An invertebrate AChE three-dimensional protein structure PDB 1Q09 (Harel et al., 2000) was used to map the ace-2 amino acid sequence. The three-dimensional image was generated using YASARA View (Yet Another Scientific Artificial Reality Application, http://www.yasara.org) with the manual addition of the external unresolved loop.

2.6. Phylogenetic analysis

Unique haplotypes of *C. palaplis* and *C. annulirostris* were used for phylogenetic analysis. A sister species *C. sitiens* previously sequenced for the COI as haplotypes s4 (DQ673853) was used as the outgroup; ace-2 sequence EU710675. For each marker (COI and ace-2), we performed a Maximum likelihood analysis using PhyML v2.4.4 (Guindon and Gascuel, 2003) and a Bayesian analysis using MrBayes v3.1.2 (Ronquist and Huelsenbeck, 2003). The best model of DNA sequenced evolution was determined in Modeltest v3.6 (Posada and Crandall, 1998) in cooperation with PAUP*4b10 (Swofford, 2003) using the Akaike information criterion (AIC).

2.6.1. Maximum likelihood analysis

The ace-2 alignment was analyzed using Maximum likelihood analysis with the AIC selected HKY + I model (Hasegawa et al., 1985), while the COI alignment was analyzed with the AIC selected GTR + Γ + I model (Lanave et al., 1984). The ML phylogeny tree was generated and its branch robustness assessed by the bootstrapping method with 500 replicates using PhyML v2.4.4 (Guindon and Gascuel, 2003).

2.6.2. Bayesian analysis

In order to better model the nucleotide evolution, we took advantage of MrBayes' v3.1.2 ability to relax the parameters of the nucleotide model over different partitions of data. The COI alignment was partitioned into the three coding position, while the ace-2 alignment was partitioned into the intron region as well as the three coding position of the exon with a specific model nucleotide evolution determined for each partition. For ace-2, the AIC selected the JC model (Jukes and Cantor, 1969) for the first and second codon position and the symmetrical (SYM) model (Zharkikh, 1994) for the third codon position while the intron used the HKY + I model. For the COI data, the AIC favoured the GTR model for the first codon position, the F81 model (Felsenstein, 1981) for the second codon position, and the GTR + I + G model for the third codon position. Metropolis-coupled Markov chain Monte Carlo (MCMCMC) analyses were run with one cold and three heated chains (temperature set to 0.1) for 10,000,000 generations and sampled every 200 generations in MrBayes v3.1.2. We performed four independent runs from a random starting tree per dataset.

We set the burn-in period—when the chains reached stationary phase—using the online program "AWTY: Are We There Yet?", utilizing the cumulative and "sliding window" analysis of the posterior probabilities of the different runs (Nylander et al., 2008).

2.6.3. Alternative tree topology testing using haplotype subsets

To further address the phylogenetic relationships within the obtained haplotypes, we calculated a Maximum likelihood tree for both markers (ace-2 and COI) and applied a variety of constraints to test both our hypotheses and the robustness of the optimal tree using CONSEL v0.1i (Shimodaira and Hasegawa, 2001). The best trees and all constraints were inferred in PAUP*4b10 (Swofford, 2003). We first reconstructed a Neighbour joining tree and used it to estimate the HKY + I model for ace-2 and the GTR + Γ + I likelihood parameter for the COI. We partitioned the dataset into the intron, the first, second and third codon positions. The gamma distribution was calculated separately for each partition. Then the parameters were fixed and used in the inference of a Maximum likelihood tree using a heuristic search with 20 random sequence additions and NNI swapping. The resulting parameters were used for the Maximum likelihood tree reconstruction for the unconstrained tree as well as for all constrained trees. Site likelihoods for individual trees were calculated using PAUP*4b10 (Swofford, 2003) and these were used for the approximately unbiased (AU) test (Shimodaira, 2002) and the Shimodaira–Hasegawa (SH) test (Shimodaira and Hasegawa, 1999), both implemented in CONSEL. A value of $P < 0.05$ was considered statistically significant to reject the hypothesis that the two trees were significantly identical.

3. Results

3.1. Primer selection and sequencing

We compared and analyzed all available sequences of mosquito ace-2 retrieved from public databases using the BLAST capability in MEGA4 (Tamura et al., 2007). The available ace-2 region included the 3′-region of exon 2 and the 5′-region of exon 3 and the entire second intron according to *C. pipiens* AY196910 (Bourguet et al., 1998; Sanogo et al., 2007; Smith and Fonseca, 2004).

A three-dimensional structure of *Drosophila melanogaster* AChE2 was used (DmAChE; ace-2 equates to AChE)—the only available AChE structure published from an invertebrate (Harel et al., 2000). Mapping available mosquito sequences on this structure indicates that these sequences span both a conservative as well as a highly polymorphic region of the unstructured external loop. We designed a pair of primers ACE2SHF (5′-CAC GAA CGT ATC CGA AGA CTG-3′) and ACE2SHR (5′-GCC ACG ATT ACG TTT CCA AC-3′) annealing to the most conservative exon coding regions. The selected region amplified by ACE2SHF/ACE2SHR spans two β-sheets and part of the α-helices of the determined structure of DmAChE (Fig. 1A). It also spans a portion of the DmAChE protein sequence—this structure could not be resolved in the electron density map and has been manually added to the protein complex in Fig. 1A. This external loop (hydrophilic region) represents 31 amino acids of AChE from *C. annulirostris* and *C. palpalis*, and 36 amino acids from *D. melanogaster*.

3.2. Nucleotide diversity within the ace-2 marker region

In order to assess the ace-2 region's sequence diversity for *C. annulirostris* and *C. palpalis*, the 382 nt region of the ace-2 gene was sequenced from 41 *C. annulirostris* and 17 *C. palpalis.* All sequences were of the same length (382 nt) including the 168 nt intron (44%; see Fig. 1B). We also sequenced ace-2 from *C. sitiens*, which was 389 nt consisting of a 175 nt intron (45%), which served as the outgroup (Supplementary Table S1). The alignment was 392 residues long; exon 2 spanned positions 1–76 and exon 3 spanned positions 258–292.

To resolve the individual haplotypes from each of the nuclear chromosomes for *C. annulirostris* and *C. palpalis*, we used PHASE software (Supplementary Alignment A1). PHASE identified 45 distinct haplotypes for *C. annulirostris* and *C. palpalis*; 320 nt positions were constant across all *C. annulirostris* and *C. palpalis* and 62 nt positions were variable (16% of the sequence). The majority of these polymorphic sites (37 positions) were within the intron and 25 polymorphic positions were in the exon regions. The average nucleotide diversity for *C. annulirostris* and *C. palpalis* was 2.5% per site (Fig. 1C). The nucleotide diversity for the intron region was higher (3.3%) than the exon region (1.9%). Surprisingly the 3′-region of the exon 2 containing the external loop showed the highest amount of nucleotide diversity per site (5.3%).

The coding region revealed a relatively homogeneous composition of T, C, A and G nucleotides with 21.9%, 24.5%, 24.4% and 29.1%, respectively. The base composition at the three codon positions differs greatly: at the first codon position, G is represented with the highest frequency of 32.4%; at the second position A is predominant with 33.3%; in the third position C is predominant with 32.6%.

3.3. ace-2 Phylogeny and phylogeography

We first selected an overall model for nucleotide evolution to reconstruct the phylogeny of the ace-2 haplotypes. For the Bayesian analysis we took advantage of MrBayes' ability to relax the parameters of the nucleotide model over subsets of the alignment, in order to better model the nucleotide evolution—particularly at different nucleotide coding positions (Ronquist and Huelsenbeck, 2003).

Using *C. sitiens* as an outgroup for the ace-2 marker, both Maximum likelihood and Bayesian methods showed the same phylogenetic tree topology (Fig. 2A). Both methods discriminate three highly supported lineages: *C. annulirostris*, *C. palpalis* and a third lineage termed *C. annulirostris* "ann-PNG1". The resolution between these three monophyletic lineages is ambiguous and the topology supporting the monophyly of *C. palpalis* and *C. annulirostris* "ann-PNG1" is only weakly supported by 53% bootstrap support using PhyML and 51% posterior probability using MrBayes.

The *C. annulirostris* "ann-PNG1" lineages comprised 12 mosquitoes (haplotypes h15–h18) geographically restricted to PNG and the top of Queensland's Cape York Peninsula (Fig. 3A, Supplementary Table 1). The *C. annulirostris* lineage comprised 29 mosquitoes collected over a large geographic area that included PNG, mainland Australia and the Solomon Islands. Additionally, some individual haplotypes were also dispersed throughout this region such as h9, which was found in Queensland, the Northern Territory, Western Australia, South Australia and PNG (Supplementary Table 1).

The *C. palpalis* lineage comprised 13 mosquitoes and 15 different haplotypes collected from northern Australia and PNG with haplotypes h26 and h27 appearing in both northern Australia and PNG. A second well-supported monophyletic lineage of *C. palpalis* was also identified in southern Australia (Albury, NSW), comprising four mosquitoes and six different haplotypes (h38–h43), and was tentatively termed *C. palpalis* "pal S-AUS" (Supplementary Table 1).

3.4. COI phylogeny and phylogeography

In this study, the same 59 field collected mosquitoes that were sequenced for the ace-2 were also assessed for the COI. This data comprised both new COI material and sequences from our previous study (Hemmerter et al., 2007). The new COI sequences included

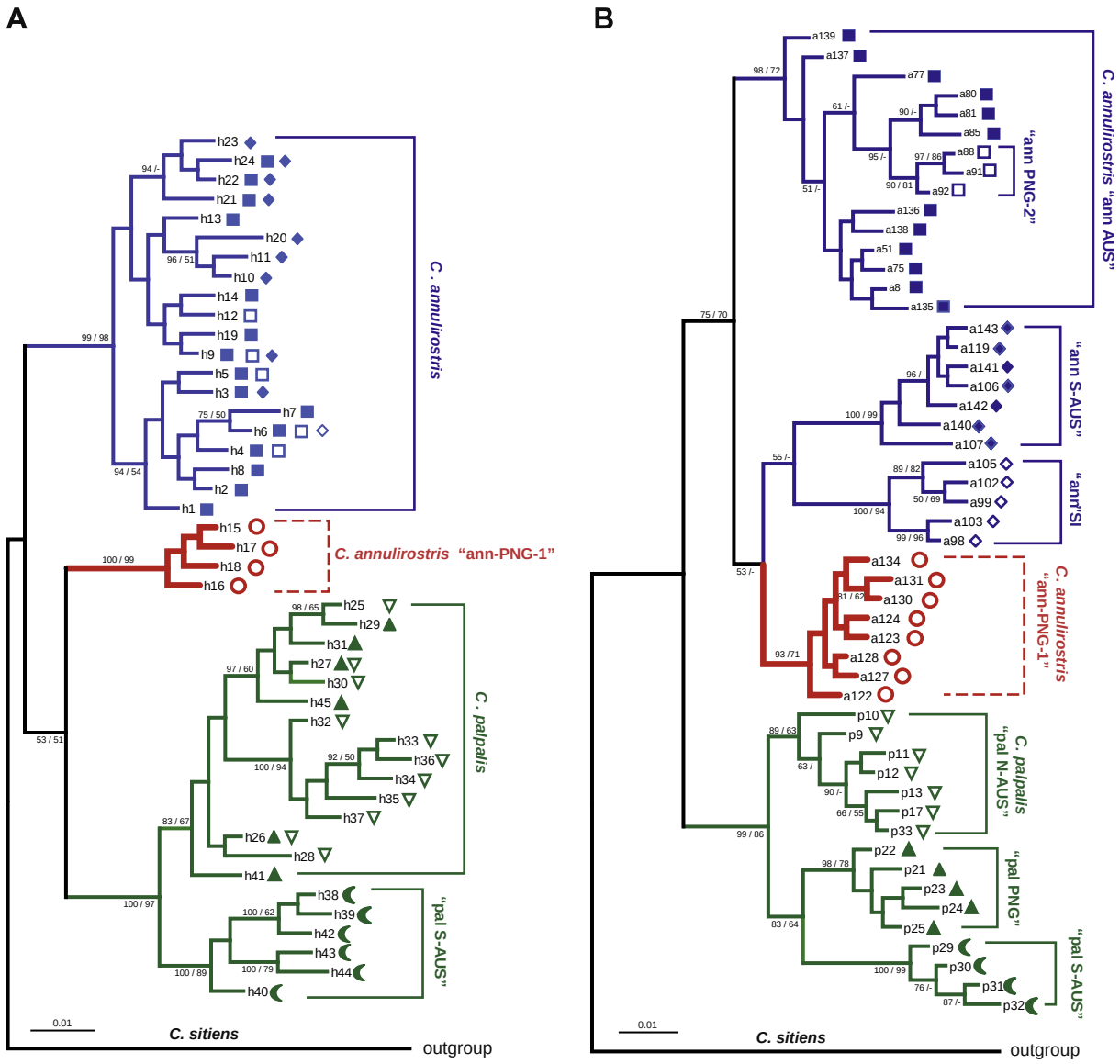

Fig. 2. Phylogenetic tree of *Culex* spp. based on ace-2 and COI markers. (A) The ace-2 Bayesian tree was reconstructed using the nucleotide sequence alignment of 382 nucleotides from 59 field collected mosquitoes (46 different haplotypes). (B) The COI Bayesian tree was reconstructed based on nucleotide sequence alignment of 538 nucleotides from the same 59 field collected mosquitoes (51 distinct haplotypes). Bayesian posterior probabilities calculated using MrBayes v3.1.2 and ML bootstrap support values calculated with PhyML v2.4.4 (500 replicates) above 50% are shown. Lineage designation is indicated on the right. The geometric shapes correspond to the different COI lineages and were mapped onto ace-2.

10 mosquitoes collected from Alice Springs in central Australia (Northern Territory) and Forestdale in Western Australia (Supplementary Table S1). When these new COI sequences were compared to the existing dataset of COI haplotypes (Hemmerter et al., 2007), nine out of the 10 haplotypes were unique and one was identical to a119 (DQ673699).

Again, using *C. sitiens* as the outgroup, both Maximum likelihood and Bayesian methods showed the same phylogenetic tree topology for the major lineages (Fig. 2B). In our previous study using only the COI, *C. annulirostris* revealed four major lineages and one sublineage while and *C. palpalis* revealed three lineages (Hemmerter et al., 2007).

We again found well-supported lineages within *C. annulirostris*—termed "ann S-AUS" and "ann SI"—that were geographically restricted to southern Australia (South Australia and Western Australia) and the Solomon Islands, respectively. The *C. annulirostris* "ann AUS" lineage was distributed throughout Australia and PNG and included a sublineage ("ann PNG-2") which was present in northernmost Cape York and PNG. The *C. annulirostris* "ann PNG-1" lineage is well supported and is also restricted to northernmost Cape York and PNG (Fig. 2B).

Culex palpalis was well resolved and showed the well-supported monophyletic lineages "pal S-AUS" found in New South Wales (Albury), a distinct "pal PNG" lineage from PNG, and a third lineage with reduced branch support, "pal N-AUS" (Fig. 3B).

3.5. Phylogenetic status of the *C. annulirostris* "ann PNG-1" lineage

The seven COI lineages identified within *C. annulirostris* and *C. palpalis* collapsed to four ace-2 lineages. In terms of *C. annulirostris*, the "ann PNG-1" lineage is supported by both markers while the additional COI lineages ("ann SI", "ann S-AUS", "ann AUS") are scattered within a single ace-2 monophyletic lineage of *C. annulirostris* (Fig. 3A and B). *Culex palpalis* remains monophyletic but only two

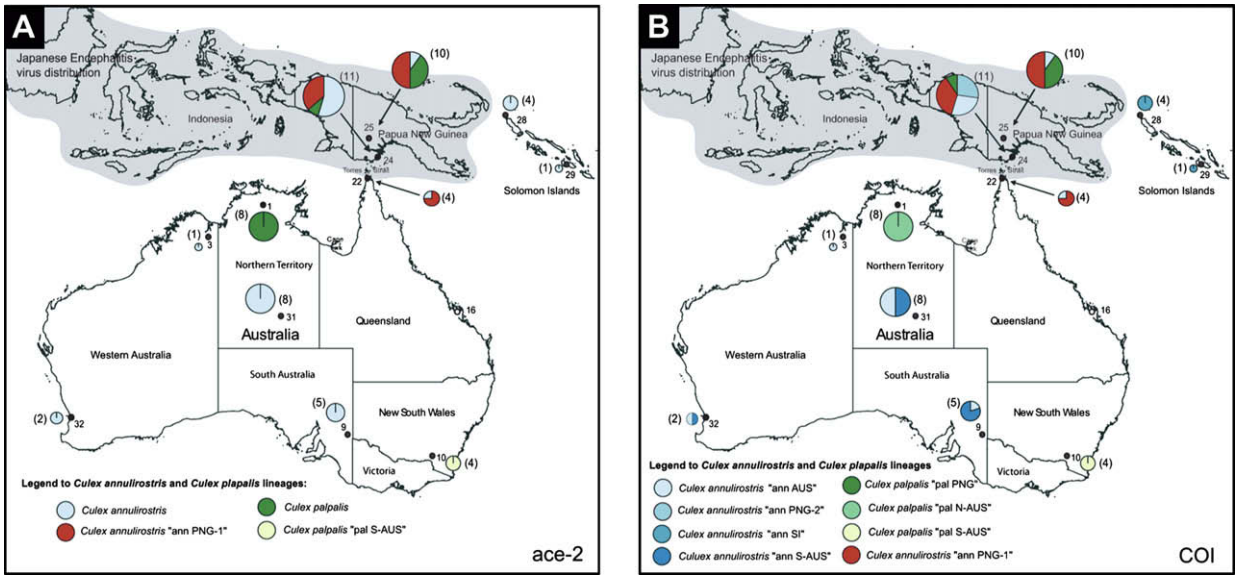

Fig. 3. Phylogeography of *C. annulirostris* and *C. palpalis*. Map of collection sites in Australasia with the proportional distribution of ace-2 lineages (A) and COI lineages (B) and with collection site numbering the same as that used in Hemmerter et al. (2007). Pie chart graphs indicate the distributional frequency of identified lineages; four ace-2 lineages and eight COI lineages. The size of the pie charts' segments is proportional to the number of mosquitoes identified as *C. annulirostris* and *C. palpalis*, which is also indicated in brackets. Fifty-eight *C. annulirostris* and *C. palpalis* mosquitoes were genotyped and the outgroup taxa *C. sitiens* is from locality 16 (Queensland, empty circle).

lineages are recognized for the ace-2, and three lineages for the COI—the "pal PNG" and "pal N-AUS" were not supported by the ace-2.

To test the effectiveness of ace-2 and COI as molecular markers for *C. annulisorstris* and *C. palpalis* we calculated the maximum intraspecific (intralineage) distance and the minimum interspecific (interlineage) distance within the respective sequences. The more discriminating a genetic marker, the lower its intraspecific diversity and the higher its interspecific diversity; if plotted against each other this value will move towards the top left corner of the graph (Fig. 4). First, we plotted two formerly described species *C. annulirostris* and *C. palpalis*, which gave values near or below the diagonal threshold line of resolution for ace-2 and COI (Fig. 4A). Second, we assumed the lineage *C. annulirostris* "ann PNG-1" to be distinct (Fig. 4B), and in the third alternative we included the fourth lineage *C. palpalis* "pal S-AUS" (Fig. 4C). This analysis indicates that the second scenario (Fig. 4B), using the ace-2 with three recognized lineages, has the lowest intralineage diversity and the highest interlineage diversity (i.e. ace-2 ▲ ■ ● are the nearest, towards the top left corner of the graph, Fig. 4B). Thus the preferred assumption about diversity is to recognize three distinct lineages: *C. annulirostris*, *C. annulirostris* "ann PNG-1" and *C. palpalis*.

Because of the lack of resolved relationships between *C. palpalis*, *C. annulirostris* and *C. annulirostris* "ann PNG-1", we attempted to test three different evolutionary hypotheses at both loci (Table 1). Either (i) that *C. annulirostris* "ann PNG-1" is a sister to *C. annulirostris* or (ii) that *C. annulirostris* "ann PNG-1" is sister to *C. palpalis*, or, lastly, (iii) that *C. palpalis* is sister to *C. annulirostris*. The SH test could not reject any alternative topology for ace-2 and COI. The AU test could not reject any topology for the COI tree while the AU test

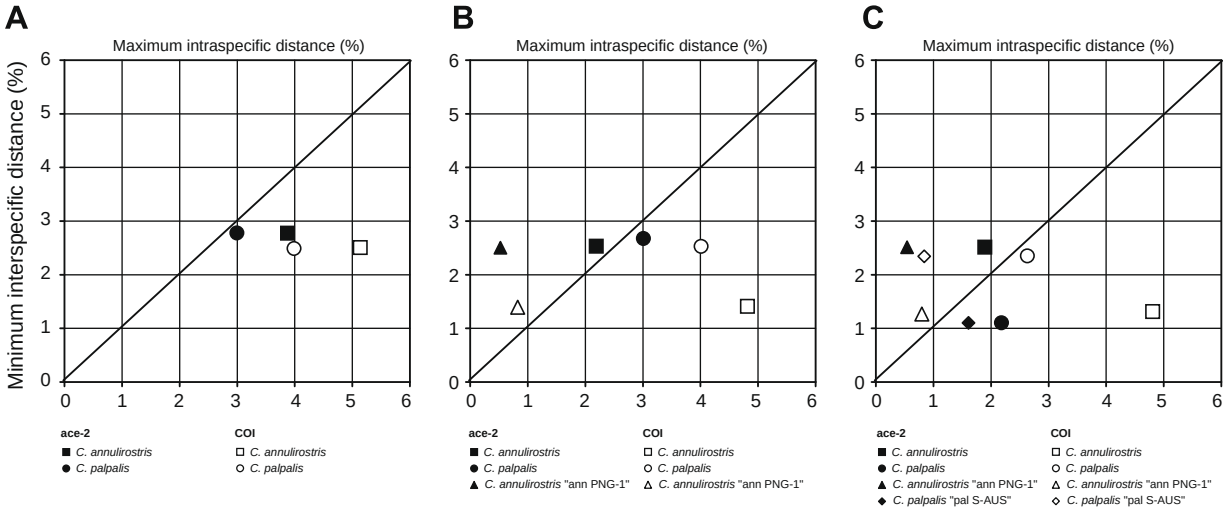

Fig. 4. The ace-2 and COI sequence divergence for the *C. annulirostris* and *C. palpalis* lineages. Minimum interspecific (interlineage) divergence is plotted against maximum intraspecific (intralineage) divergence. Points above the diagonal represent cases where the divergence between lineages is higher than the divergence within lineages. The best molecular markers should occur above the diagonal line towards the top left corner and display low intraspecific divergence and high interspecific divergence. (A) The two formerly described species *C. annulirostris* and *C. palapalis* are recognized; (B) three distinct lineages including *C. annulirostris* "ann PNG-1" are assumed; and (C) the fourth lineage *C. palpalis* "pal S-AUS" is included. Note that ace-2 (B) with three recognized lineages has the lowest intralineage diversity and a high interlineage diversity and is the preferred assumption, revealing *C. annulirostris*, *C. annulirostris* "ann PNG-1" and *C. palpalis*.

Table 1
Confidence values from the approximately unbiased (AU) and Shimodaira–Hasegawa (SH) tests for the unconstrained and constrained tree topologies.

Topology	Nuclear data (ace-2)			Mitochondrial data (COI)		
	Obs.	AU	SH	Obs.	AU	SH
(pal (ann, ann PNG-1)	best	0.765	0.632	0	0.727	0.690
(ann (pal, ann PNG-1)	0.2	0.659	0.585	2.4	0.197	0.365
(ann PNG-1 (ann, pal)	2	0.01*	0.454	1.4	0.435	0.507

Approximately unbiased (AU) and Shimodaira–Hasegawa (SH) tests performed using CONSEL v0.1i. Obs., test statistics; *P-value < 0.05, suggests that the constraints are significantly different—rejected; ace-1, pal—*C. palpalis* + pal S-AUS, ann—*C. annulirostris*; COI, pal—pal PNG + pal N-AUS + pal S-AUS, ann—ann AUS + ann S-AUS + ann SI. *Abbreviations:* ann, *C annulirostris*; pal, *C. palpalis*; for lineage description refer Fig. 3.
* P-value < 0.05 suggested that the constrained tree is significantly different from the unconstrained tree, the rejected values are mark by an triangle.

at ace-2 rejected the monophyly of *C. palpalis* and *C. annulirostris*. These results indicated a lack of robust phylogenetic signal in both ace-2 and COI for these mosquitoes.

3.6. Amino acid changes in ace-2

On assessment of the amino acids sequence of the ace-2 exons, 69 of 71 amino acids were identical. Inclusion of available sequences from *Culex* spp. in GenBank revealed a total of four variable residues across these 71 amino acids. Within *C. palpalis*, *C. annulirostris* and *C. sitiens*, the two amino acid changes occur in the hydrophilic external loop (Fig. 1C). Both residue changes resulted from nucleotide substitution at the second coding position—the first at 23 (68–70 nucleotides) and the second at position 33 (98–100 nucleotides). The first was found as Asparagine in *C. annulirostris* "ann-PNG1", and Serine in *C. annulirostris*, *C. palpalis* and *C. sitiens*. This Asparagine also appears in *C. australicus*, *C. pipines palens* and *C. pipiens quinquefasciatus*. Both Asparagine and Serine are similar in their chemical properties, being small and polar. The second amino acid polymorphism was identified as Tyrosine in *C. annulirostris* "ann-PNG1" and *C. palpalis* while Phenylalanine was found in *C. annulirostris* "ann-AUS" and *C. sitiens*. This change provides evidence to support their monophyly according to the hypothetical (ii) scenario and ace-2 phylogeny (Fig. 2A). No other member of the *Culex* subgenus available in GenBank shows this amino acid change. Both phenylalanine and tyrosine are similar in their chemical properties, being aromatic and hydrophobic.

4. Discussion

Australia's northern Cape York Peninsula has the appropriate conditions for the establishment of the exotic arbovirus JEV, possessing a large quantity of the *C. annulirostris* mosquitoes identified as transmitting JEV (Johansen et al., 2000), large populations of wild pigs which are an excellent amplifying host for the virus, and a close geographic proximity to regular JEV activity (Mackenzie et al., 2002). Consequently, the question of why JEV can cycle for over a decade on islands 70 km north of mainland Australia and not establish on the mainland itself has puzzled both research and public health scientists (Hanna et al., 1999; Hemmerter et al., 2007). We previously identified two potential mtDNA lineages of the JEV vector whose southern distributional limits equate exactly with the current southern limits of JEV—and we proposed that these mtDNA lineages may represent cryptic taxa with the appropriate biology and behavior to cycle JEV (Hemmerter et al., 2007). This subsequent study confirms that a true cryptic species of *C. annulitostris* is present and shares a southern limit with JEV.

4.1. Phylogeny and phylogeography of the ace-2 marker and COI

Maximum likelihood and Bayesian analyses produced a similar phylogeny for each marker considered in this subsequent study (Fig. 3). The major difference between the COI and ace-2 trees was the position of *C. annulirostris* "ann PNG-1", which was placed as the sister taxa to *C. annulirostris* or *C. palpalis* depending on the marker (Fig. 3). In addition the SH test and AU test do not reject these hypotheses for each marker (Table 1). It is not surprising to see tree resolution differences between the ace-2 and COI as the mitochondrial gene generally has a higher rate of substitution than nuclear genes and thus may provide more resolution (Hall and Khromykh, 2004). However, while the nuclear marker did display better branch support than the COI for the main lineages, its evolution was either too slow to reveal the diversity evident in the COI or these branches have recently collapsed due to interbreeding and gene flow between the COI divergent lineages. The potential exists when using the COI barcode to overestimate species diversity due to problems with the amplification of nuclear mitochondrial pseudogenes (numts; Song et al., 2008), and in insects, through the selective effects of maternally inherited symboints (Frezal and Leblois, 2008). We found no evidence of mtDNA pseudogenes, and it seems unlikely that the incongruence we observe is driven by maternally inherited symbionts like *Wolbachia*—which cause selective mtDNA haplotype sweeps through populations that severely reduce the mtDNA haplotype diversity (Hurst and Jiggins, 2005)—as all mtDNA COI lineages showed high haplotype diversity and no evidence was found of *Wolbachia* DNA in mosquitoes sampled from each lineage (Hemmerter et al., 2007; S. Hemmerter, unpublished data).

Mitochondrial markers are frequently used for phylogenetic studies as the genome is represented in high copy numbers per cell and is maternally inherited; because recombination is rare, there is a relatively linear evolution. Also, its generally higher mutation rate in insects when compared to nuclear DNA (Lin and Danforth, 2004)—in combination with its smaller effective population size (compared to nuclear DNA)—means lineages will resolve more quickly (Ballard and Whitlock, 2004). However, as useful as the mtDNA is in revealing historical biodiversity through the identification of divergent lineages such as those revealed by the COI barcode, these lineages may not equate to extant species as this requires some evidence for reproductive isolation which mtDNA cannot provide—i.e. the ace-2 lineages of *C. annulirostris* "ann PNG-1" show strong phylogenetic signal and no evidence of shared alleles in sympatric populations, which can be seen as evidence of reproductive isolation. Thus the advantage of nuclear DNA over the mtDNA is that the haplotypes identified from each chromosome (alleles) can be dissected out to identify the haplotypes shared within and between individuals and COI lineages—in this case we used the PHASE software rather than cloning and sequencing. Evidence of shared haplotypes between divergent lineages may be indicative of recent gene flow or incomplete lineage sorting because recent speciation events may mean ancestral haplotypes are still contained in the population. Examples of shared ace-2 haplotypes from individuals that exist across well-supported COI lineages include the ace-2 haplotype h6, the only haplotype in the Solomon Islands, that is also present in PNG, and h9, which was

found in southern Australia, northern Australia and in PNG-collected individuals (see Fig. 2 and Supplementary Table S1).

Applying a phylogenetic framework to the mtDNA COI and nuclear ace-2, our results discriminate *C. palpalis*, *C. annulirostris* and *C. annulirostris* "ann-PNG1" with a strong branch support (3). The later cryptic taxon will now be tentatively called *Culex* sp. "PNG".

4.2. The ace-2 nuclear marker

In this study, we assessed the intron–exon region of the ace-2 gene as a molecular genetic marker to confirm the presence of mtDNA lineages found using the COI (Hemmerter et al., 2007). We found, through three-dimensional mapping studies, that the DNA sequence flanking this intron codes for an external protein loop that appeared to be highly polymorphic, contributing notably to the resolution of the whole DNA sequence (Fig. 1). Surprisingly, the 5′-external protein loop shows an elevated level of polymorphisms compared to the intron which suggests that the loop's position and orientation on the protein is under a decreased evolutionary selection pressure and that the amino acid substitutions on the 5′- and 3′-exons of this loop also contribute to the identification of the three *Culex* lineages (Fig. 1).

The presence of the external loop was found to be an important distinction between the AChE1 and AChE2 proteins—this region is also called the "hydrophilic insertion" and it is absent from vertebrate and nematode AChE (Weill et al., 2002). The synaptic acetylcholinesterase AChE 2 gene function has been identified in true flies (Cyclorrhapha) as the termination of the transmission of the cholinergic synapse, where in the other diptera this function is performed by AChE1 (Huchard et al., 2006). The functionality of the AChE2 in mosquitoes still remains unclear although results from enzymatic assays suggest that AChE 2 has a small or very localized synaptic function (Huchard et al., 2006).

The three-dimensional structure of crystallized AChE2 from *Drosophila melanogaster* is available for comparison, however, the external loop, located on the surface of the protein, could not be resolved (Harel et al., 2000). The sequence conservation of the remaining ace-2 gene among all insects suggests that most of the protein is functionally imperative and under purifying selection—i.e. most amino acid substitutions are deleterious (Huchard et al., 2006). Nonetheless, it appears that this external "hydrophilic" loop is not under as stringent selection and thus was useful for this molecular genetic study in combination with the intron.

4.3. Just another molecular marker for the identification of insects?

The efforts to understand and control the transmission of the pathogens responsible for human diseases require the use of molecular diagnostic tools to identify and then distinguish morphologically cryptic species to accurately identify those that vector pathogens from those that do not. The nuclear ribosomal DNA (often the internal transcribed spacers) and mitochondrial DNA (in particular the cytochrome oxidase I) have been utilized extensively to identify cryptic mosquito taxa within morphospecies (Collins and Paskewitz, 1996; Norris, 2002; Wells and Sperling, 2000). Ribosomal DNA is a highly polymorphic multicopy gene family that frequently requires a cloning step for sequence analyses. The COI is being utilized in "DNA barcoding" and is now more widely used to visualize recent historical events and species groups (Hebert et al., 2004, 2003; Ward et al., 2005). However, there are limits to the ways in which DNA barcoding using the COI can be serviceable for insect systematics when they fail to identify species due to overlaps between intra- and interspecific genetic variability, recent speciation events causing incomplete lineage sorting, and maternally inherited bacteria that cause mtDNA selective sweeps (Elias et al., 2007; Meier et al., 2006; Whitworth et al., 2007).

When searching for a useful molecular marker for species identification it is important to show that the nucleotide diversity within species is no greater than the nucleotide divergence between species (Meyer and Paulay, 2005). In this, the ace-2 displays a lower intraspecific distance and higher interspecific distance than the COI thus revealing its capacity as a species diagnostic marker as well as its potential as a diagnostic tool. It is unlikely that these results are due to the fact that the COI sequence analyzed in this study does not match exactly the 5′-end COI barcode region, as it has been shown that this barcoding region is no more optimal than other regions downstream in COI (Roe and Sperling, 2007). The second intron–exon region of the ace-2 is a single copy autosomal gene that shows a high diversity between populations and low diversity within a population, making the ace-2 a useful marker for molecular identification.

5. Conclusions

Molecular diagnostic tools resolving complexes of cryptic mosquito species have been utilized successfully for a number of medically and veterinary important vector-borne diseases (Beebe and Cooper, 2000; Krzywinski and Besansky, 2003; Walton et al., 1999). The usefulness of a specific region of ace-2 has previously been assessed as a marker to discriminate between cryptic *Culex* species (Bourguet et al., 1998; Kasai et al., 2008; Sanogo et al., 2007; Smith and Fonseca, 2004). Our work supports the ability of the ace-2 sequence to reveal closely related *Culex* species and to reconstruct the evolutionary history of the mosquitoes implicated in the distribution of JEV in Australasia. Thus we could confirm the presence of a cryptic *Culex* species that shares the same southern limit as JEV—an important mosquito-borne flavivirus already responsible for thousands of human deaths annually in southeastern and southern Asia. A molecular diagnostic can now be developed from this ace-2 marker, delivering another useful mosquito diagnostic tool to assist epidemiological studies and surveillance programs by facilitating accurate identification of pathogen-transmitting vectors.

Additionally, we looked into the functional components of this ace-2 marker to find fast evolving coding DNA flanking the intron that forms an external loop on the protein's tertiary structure. With the growth of protein crystal structure analysis, we would advocate mapping coding sequences onto these structures—where they are available—as this produces useful insights into their functionality. Alternatively, designing molecular markers based on protein structure features such as external protein loops that may be highly polymorphic and can also be better modeled than noncoding introns. Finally the COI marker may be overestimating the extant species diversity of these mosquitoes, although other nuclear markers would be required to confirm this incongruence.

Acknowledgments

S.H. acknowledges insights gained from the 2007 Workshop on Molecular Evolution in Woods Hole US, and to thank Christopher Balakrishnan and Scott Edwards for their help with building the input file for PHASE. This study was supported through a PhD scholarship (S.H.) from the Institute for the Biotechnology of Infectious Diseases, University of Technology, Sydney and through funding from the University of Queensland. We thank Andrew F. van den Hurk, Robert D. Cooper, Peter I. Whelan, Richard C. Russell and Cheryl A. Johansen for providing some of the mosquitoes for this study. A number of these mosquito collections were funded by the Australian Health Minister's Advisory Council (AHMAC) Priority Driven Research Program.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2008.11.016.

References

Ballard, J.W., Whitlock, M.C., 2004. The incomplete natural history of mitochondria. Mol. Ecol. 13, 729–744.

Beebe, N.W., Cooper, R.D., 2000. Systematics of malaria vectors with particular reference to the *Anopheles punctulatus* group. Int. J. Parasitol. 30, 1–17.

Beebe, N.W., van den Hurk, A.F., Chapman, H.F., Frances, S.P., Williams, C.R., Cooper, R.D., 2002. Development and evaluation of a species diagnostic polymerase chain reaction-restriction fragment-length polymorphism procedure for cryptic members of the *Culex sitiens* (Diptera: Culicidae) subgroup in Australia and the southwest Pacific. J. Med. Entomol. 39, 362–369.

Bourguet, D., Fonseca, D., Vourch, G., Dubois, M.P., Chandre, F., Severini, C., Raymond, M., 1998. The acetylcholinesterase gene Ace: a diagnostic marker for the Pipiens and Quinquefasciatus forms of the *Culex pipiens* complex. J. Am. Mosq. Cont. Assoc. 14, 390–396.

Chapman, H.F., Hughes, J.M., Ritchie, S.A., Kay, B.H., 2003. Population structure and dispersal of the freshwater mosquitoes *Culex annulirostris* and *Culex palpalis* (Diptera: Culicidae) in Papua New Guinea and northern Australia. J. Med. Entomol. 40, 165–169.

Chapman, H.F., Kay, B.H., Ritchie, S.A., van den Hurk, A.F., Hughes, J.M., 2000. Definition of species in the *Culex sitiens* subgroup (Diptera: Culicidae) from Papua New Guinea and Australia. J. Med. Entomol. 37, 736–742.

Collins, F.H., Paskewitz, S.M., 1996. A review of the use of ribosomal DNA (rDNA) to differentiate among cryptic *Anopheles* species. Insect Mol. Biol. 5, 1–9.

Elias, M., Hill, R.I., Willmott, K.R., Dasmahapatra, K.K., Brower, A.V., Mallet, J., Jiggins, C.D., 2007. Limited performance of DNA barcoding in a diverse community of tropical butterflies. Proc. Biol. Sci. 274, 2881–2889.

Felsenstein, J., 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. 17, 368–376.

Frezal, L., Leblois, R., 2008. Four years of DNA barcoding: current advances and prospects. Infect. Genet. Evol. 8, 727–736.

Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704.

Hall, R.A., Khromykh, A.A., 2004. West Nile virus vaccines. Expert Opin. Biol. Ther. 4, 1295–1305.

Hanna, J.N., Ritchie, S.A., Phillips, D.A., Lee, J.M., Hills, S.L., van den Hurk, A.F., Pyke, A.T., Johansen, C.A., Mackenzie, J.S., 1999. Japanese encephalitis in north Queensland, Australia, 1998. Med. J. Aust. 170, 533–536.

Harel, M., Kryger, G., Rosenberry, T.L., Mallender, W.D., Lewis, T., Fletcher, R.J., Guss, J.M., Silman, I., Sussman, J.L., 2000. Three-dimensional structures of *Drosophila melanogaster* acetylcholinesterase and of its complexes with two potent inhibitors. Protein Sci. 9, 1063–1072.

Hasegawa, M., Kishino, H., Yano, T., 1985. Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22, 160–174.

Hebert, P.D., Penton, E.H., Burns, J.M., Janzen, D.H., Hallwachs, W., 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proc. Natl. Acad. Sci. USA 101, 14812–14817.

Hebert, P.D., Ratnasingham, S., deWaard, J.R., 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc. Biol. Sci. 270 (Suppl. 1), S96–S99.

Hemingway, J., Hawkes, N.J., McCarroll, L., Ranson, H., 2004. The molecular basis of insecticide resistance in mosquitoes. Insect Biochem. Mol. Biol. 34, 653–665.

Hemmerter, S., Slapeta, J., van den Hurk, A., Cooper, R., Whelan, P., Russell, R., Johansen, C., Beebe, N., 2007. A curious coincidence: mosquito biodiversity and the limits of the Japanese encephalitis virus in Australasia. BMC Evol. Biol. 7, 100.

Huchard, E., Martinez, M., Alout, H., Douzery, E.J., Lutfalla, G., Berthomieu, A., Berticat, C., Raymond, M., Weill, M., 2006. Acetylcholinesterase genes within the Diptera: takeover and loss in true flies. Proc. Biol. Sci. 273, 2595–2604.

Hurst, G.D., Jiggins, F.M., 2005. Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. Proc. Biol. Sci. 272, 1525–1534.

Johansen, C.A., van den Hurk, A.F., Ritchie, S.A., Zborowski, P., Nisbet, D.J., Paru, R., Bockarie, M.J., Macdonald, J., Drew, A.C., Khromykh, T.I., Mackenzie, J.S., 2000. Isolation of Japanese encephalitis virus from mosquitoes (Diptera: Culicidae) collected in the Western Province of Papua New Guinea, 1997–1998. Am. J. Trop. Med. Hyg. 62, 631–638.

Jukes, T.H., Cantor, C.R., 1969. Evolution of protein molecules. Mamm. Protein Metab. 3, 21–132.

Kasai, S., Komagata, O., Tomita, T., Sawabe, K., Tsuda, Y., Kurahashi, H., Ishikawa, T., Motoki, M., Takahashi, T., Tanikawa, T., Yoshida, M., Shinjo, G., Hashimoto, T., Higa, Y., Kobayashi, M., 2008. PCR-based identification of *Culex pipiens* complex collected in Japan. Jap. J. Infect. Dis. 61, 184–191.

Kimura, M., 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111–120.

Krzywinski, J., Besansky, N.J., 2003. Molecular systematics of *Anopheles*: from subgenera to subpopulations. Annu. Rev. Entomol. 48, 111–139.

Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief. Bioinform. 5, 150–163.

Lanave, C., Preparata, G., Saccone, C., Serio, G., 1984. A new method for calculating evolutionary substitution rates. J. Mol. Evol. 20, 86–93.

Lee, D.J., Hicks, M.M., Griffiths, M., Russell, R.C., Bryan, J.H., Marks, E.N., 1989. The Culicidae of the Australian Region. Entomology. Monograph 2, vol. 7. Australian Government Printing Service Press, Canberra, Australia.

Lin, C.P., Danforth, B.N., 2004. How do insect nuclear and mitochondrial gene substitution patterns differ? Insights from Bayesian analyses of combined datasets. Mol. Phylogenet. Evol. 30, 686–702.

Mackenzie, J.S., Johansen, C.A., Ritchie, S.A., van den Hurk, A.F., Hall, R.A., 2002. Japanese encephalitis as an emerging virus: the emergence and spread of Japanese encephalitis virus in Australasia. Curr. Top. Microbiol. Immunol. 267, 49–73.

Marks, E.N., 1982. An Atlas of Common Queensland Mosquitoes. The Queensland Institute of Medical Research, Brisbane.

Meier, R., Shiyang, K., Vaidya, G., Ng, P.K., 2006. DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. Syst. Biol. 55, 715–728.

Meier, R., Zhang, G., Ali, F., 2008. The use of mean instead of smallest interspecific distances exaggerates the size of the "barcoding gap" and leads to misidentification. Syst. Biol. 57, 809–813.

Meyer, C.P., Paulay, G., 2005. DNA barcoding: error rates based on comprehensive sampling. PLoS Biol. 3, e422.

Norris, D.E., 2002. Genetic markers for study of the anopheline vectors of human malaria. Int. J. Parasitol. 32, 1607–1615.

Nylander, J.A., Wilgenbusch, J.C., Warren, D.L., Swofford, D.L., 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. Bioinformatics 24, 581–583.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Roe, A.D., Sperling, F.A., 2007. Patterns of evolution of mitochondrial cytochrome c oxidase I and II DNA and implications for DNA barcoding. Mol. Phylogenet. Evol. 44, 325–345.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Rozas, J., Sanchez-DelBarrio, J.C., Messeguer, X., Rozas, R., 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19, 2496–2497.

Russell, R.C., Dwyer, D.E., 2000. Arboviruses associated with human disease in Australia. Microbes Infect. 2, 1693–1704.

Sanogo, Y.O., Kim, C.H., Lampman, R., Novak, R.J., 2007. A real-time TaqMan polymerase chain reaction for the identification of Culex vectors of West Nile and Saint Louis encephalitis viruses in North America. Am. J. Trop. Med. Hyg. 77, 58–66.

Shimodaira, H., 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. 51, 492–508.

Shimodaira, H., Hasegawa, M., 1999. Multiple Comparisons of Log-Likelihoods with Applications to Phylogenetic Inference. Mol. Biol. Evol. 16, 1114–1116.

Shimodaira, H., Hasegawa, M., 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics 17, 1246–1247.

Smith, J.L., Fonseca, D.M., 2004. Rapid assays for identification of members of the *Culex* (*Culex*) *pipiens* complex, their hybrids, and other sibling species (Diptera: culicidae). Am. J. Trop. Med. Hyg. 70, 339–345.

Song, H., Buhay, J.E., Whiting, M.F., Crandall, K.A., 2008. Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. Proc. Natl. Acad. Sci. USA 105, 13486–13491.

Stephens, M., Scheet, P., 2005. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. Am. J. Hum. Genet. 76, 449–462.

Stephens, M., Smith, N.J., Donnelly, P., 2001. A new statistical method for haplotype reconstruction from population data. Am. J. Hum. Genet. 68, 978–989.

Swofford, D.L., 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol 24, 1596–1599.

Toutant, J.P., 1989. Insect acetylcholinesterase: catalytic properties, tissue distribution and molecular forms. Prog. Neurobiol. 32, 423–446.

Walton, C., Sharpe, R., Pritchard, S., Thelwell, N., Butlin, R., 1999. Molecular identification of mosquito species. Biol. J. Lin. Soc. 68, 241–256.

Ward, R.D., Zemlak, T.S., Innes, B.H., Last, P.R., Hebert, P.D., 2005. DNA barcoding Australia's fish species. Philos. Trans. R. Soc. Lond. B Biol. Sci. 360, 1847–1857.

Weill, M., Fort, P., Berthomieu, A., Dubois, M.P., Pasteur, N., Raymond, M., 2002. A novel acetylcholinesterase gene in mosquitoes codes for the insecticide target and is non-homologous to the ace gene in *Drosophila*. Proc. Biol. Sci. 269, 2007–2016.

Wells, J.D., Sperling, F.A., 2000. A DNA-based approach to the identification of insect species used for postmortem interval estimation and partial sequencing of the cytochrome oxydase b subunit gene I: a tool for the identification of European species of blow flies for postmortem interval estimation. J. Forensic Sci. 45, 1358–1359.

Whitworth, T.L., Dawson, R.D., Magalon, H., Baudry, E., 2007. DNA barcoding cannot reliably identify species of the blowfly genus *Protocalliphora* (Diptera: Calliphoridae). Proc. Biol. Sci. 274, 1731–1739.

Zharkikh, A., 1994. Estimation of evolutionary distances between nucleotide sequences. J. Mol. Evol. 39, 315–329.