

Mitochondrial DNA reveals phylogenetic structuring and cryptic diversity in Australian freshwater macroinvertebrate assemblages

Andrew M. Baker^{A,C}, Jane M. Hughes^A, John C. Dean^B and Stuart E. Bunn^A

^ACooperative Research Centre for Freshwater Ecology, Centre for Riverine Landscapes, Griffith University, Nathan Campus, Qld 4111, Australia.

^BEnvironment Protection Authority, Freshwater Sciences Unit, Ernest Jones Drive, Macleod, Vic. 3085, Australia.

^CCorresponding author. Email: am.baker@qut.edu.au

Abstract. Freshwater catchments of south-east Australia possess generally rich and diverse macroinvertebrate faunas, although the genetic structuring of these assemblages is poorly known. In this study, we assessed mitochondrial phylogenetic structure within four genera of macroinvertebrates from the Sydney Water Supply Catchment, south-east Australia: *Euastacus* (parastacid crayfish), *Cheumatopsyche* (hydrpsychid caddisflies), *Atalophlebia* (leptophlebiid mayflies) and *Paratya* (atyid shrimp), with a view to prioritising areas of high diversity for future conservation efforts. We found extremely divergent (~4–19%) cytochrome *c* oxidase subunit I (COI) lineages within all surveyed groups, many of which corresponded to recognised taxa, although there was also evidence of cryptic species within three genera; *Euastacus*, *Atalophlebia* and *Paratya*. Distributions of these three genera were associated with high altitude streams, above dam impoundments. Our results have important implications for management of the Sydney Water Supply Catchment. Future disturbance in this region is likely to be high and priority should be directed towards preserving the diversity of fauna in these upland areas. This comparative phylogenetic approach may have value as a means to focus and direct conservation efforts in other areas supporting high biodiversity.

Extra keywords: *Atalophlebia*, *Cheumatopsyche*, cryptic diversity, *Euastacus*, mitochondrial DNA, *Paratya*, phylogenetic.

Introduction

The significance of genetic diversity as a component of biodiversity evaluation is widely recognised, and its protection has been incorporated into national and international conventions (Moritz and Faith 1998). In the context of biodiversity evaluation, genetic variation can be analysed at various levels, from the individual, through populations to species (Feral 2002). For single species conservation planning, it is paramount to identify component evolutionary lineages to retain maximum genetic diversity, whereas at a higher level, interspecific molecular phylogenies may be used to assess phylogenetic distinctiveness (and conservation values) of species and the areas that they occupy (Moritz and Faith 1998).

In recent years, several algorithms have been proposed to provide a more objective (phylogenetic) basis for making conservation decisions (Bininda-Emonds *et al.* 2000). Such methods have arisen out of a realisation that under a phylogenetic species concept, the number of species in a given area may be inflated greatly compared to morphology-based species designations (Faith 2002). Phylogenetic diversity (PD; Faith 1992), for example, attempts to measure directly the evolutionary history of taxa by estimating the relative

feature diversity of any nominated set of species, by the sum of the lengths of all phylogenetic branches spanned by the set (Faith 2002). However, such quantitative techniques will be of most use where extensive genetic, ecological and taxonomic data have already been compiled for fauna in the area under investigation, and for many regions, these data are not yet available.

Macroinvertebrate communities form an important component of total biodiversity in freshwaters, and as such are widely used in monitoring programmes designed to assess the health of rivers and their catchments (Chessman and Williams 1999). However, available data on macroinvertebrate diversity are generally sparse at both the local and regional scale. Recently, Boyero (2002) compared biodiversity data of two macroinvertebrate taxa (Ephemeroptera and Odonata) at a regional scale in the Americas and found a generally high relative diversity of tropical compared to temperate regions, although there were considerable differences between the two tropical and two temperate regions studied. However, local studies in streams throughout the world have failed to show any clear latitudinal pattern of macroinvertebrate diversity (Boyero 2002).

In Australia, there also appears to be no general pattern attributable to invertebrate species diversity. Due to Australia's unpredictable and arid climate, one may expect generally low invertebrate species diversity. This appears to be true of temperate streams in south-west Australia, which exhibit both low species and family diversity (Bunn and Davies 1990). However, this apparently is not true of temperate streams in south-east Australia, which may exhibit even greater species diversity than many northern hemisphere streams (Lake *et al.* 1985). A case in point is the Hawkesbury-Nepean River catchment in south-east New South Wales, which covers an area of 2.2 million hectares that encloses the Sydney Basin and regions to the north and west. Agriculture, industry and urban expansion have had major effects on the Cumberland Plain, the Southern Tablelands and on riverine floodplains within the catchment. In contrast, the more rugged and nutrient poor sandstones and mountains remain largely undeveloped (Recher *et al.* 1993). A recent broad-scale survey of aquatic macroinvertebrate biodiversity in the Hawkesbury-Nepean and neighbouring Georges River catchment, indicated high regional richness (with 443 recorded species and morphospecies), despite the considerable pressures associated with urban expansion since the 1970s (Chessman and Williams 1999). The true levels of diversity in this region may be even higher, since previous studies have not taken into account the possibility of cryptic diversity.

The Sydney Water Supply Catchment (SWSC) encompasses the Hawkesbury-Nepean, Georges and Shoalhaven River catchments and is governed by the Sydney Catchment Authority (SCA). The SCA was established in 1998 to investigate a series of water supply contamination events, and its primary function is to supply water of sufficient quantity and quality to its customers. In relation to its primary directive, however, the SCA aims to manage genetic and species diversity and maintain habitat, particularly in relation to the management of catchment processes that conserve and protect water quality, public health and safety, and community values. Thus, habitat protection in areas of minimal disturbance and habitat restoration in highly modified areas are essential components of SCA efforts to achieve biodiversity conservation (SCA 2002).

Genetic markers can be used to describe the pattern and extent of phylogenetic structuring within genera of macroinvertebrate fauna occurring in the SWSC. These data can then be used to predict the consequences for biodiversity if certain developments go ahead, and provide information on how to moderate the impact of such developments.

The aim of the present study was to provide a first-order, qualitative comparison of the level and pattern of genetic structuring within targeted aquatic macroinvertebrate genera occurring in the SWSC, with a view to prioritising areas of high diversity for future conservation efforts. Specifically, in this paper we assessed patterns of mitochondrial phylogenetic structuring within four diverse macroinvertebrate genera:

(1) crustaceans; spiny crayfish (*Euastacus*) and freshwater shrimp (*Paratya*); and (2) insects; caddisflies (*Cheumatopsyche*) and mayflies (*Atalophlebia*). Before genetic screening, where possible, we used the most recent taxonomic keys to identify each group to a species. This enabled us to determine the extent of mitochondrial structuring within each genus and the presence of any cryptic diversity.

Materials and methods

The study area

The SWSC is a network delivery system drawing water from three major catchments in south-east New South Wales, Australia: The Hawkesbury-Nepean, Georges and Shoalhaven catchments (Fig. 1). The Hawkesbury-Nepean catchment covers more than ~20 000 km² to the south-west, west and north of Sydney, draining into the sea north of Sydney at Broken Bay. The Georges catchment covers ~100 km² and flows into the sea south of Sydney, through Botany Bay. The Shoalhaven catchment covers an area of ~5700 km² and drains into the Tasman Sea, east of Nowra (SCA 2002).

In this study, we also obtained specimens for some taxa from three other (unregulated) catchments: the Hacking, Macquarie and Minnamurra rivers (Fig. 1). These three catchments are proximate to, and fall within the latitudinal range defined by, the three SWSC water supply catchments. Therefore, all six catchments are henceforth referred to as component parts of the SWSC.

The study taxa

The four genera screened in this study were chosen: (1) to represent a diverse sample of the invertebrate fauna; (2) because they occur across several catchments in the study area; and (3) because they are caught relatively easily.

The freshwater spiny crayfish, genus *Euastacus*, is endemic to Australia and contains approximately 37 recognised species (Horwitz 1995). *Euastacus* prefer cool, clear, running water and generally inhabit highlands or streams adjacent to elevated areas. Twenty-one currently described *Euastacus* species are recognised as being restricted to, or extending into, New South Wales. Most of these species have extremely restricted distributions and because of their reliance on forested, high altitude streams, are particularly susceptible to habitat disturbance (Merrick 1993). *Euastacus* species are extremely difficult to discern morphologically (even experts have difficulty, using the current keys (Morgan 1997)). Therefore, we also compared our sequences with a comprehensive mitochondrial genetic dataset (K. Crandall *et al.*, unpublished data) containing representative sequences of all known *Euastacus* species (phylogenetic comparisons of the combined dataset are described in the results but cannot be presented as figures in the present study, since the comparative data are unpublished). In this way, we were able to resolve monophyletic clades of haplotypes from the present study that were associated with specimens accurately identified and lodged with Australian state museums.

The caddisfly family Hydropsychidae represents an important component of many Australian running water communities. Eight genera and 27 species are currently recognised from Australia, although the fauna is considerably more diverse and requires further investigation (Dean 1999a). The genus *Cheumatopsyche* contains seven formally named species. However, Dean (1999a) recognised 18 taxa on the basis of larval morphology. Larvae are restricted to moderate- to fast-flowing water where they construct fixed retreats and use a silken capture net to filter food particles from the water column (Dean 1999a). Before genetic analysis we identified larvae using the morphological key in Dean (1999a).

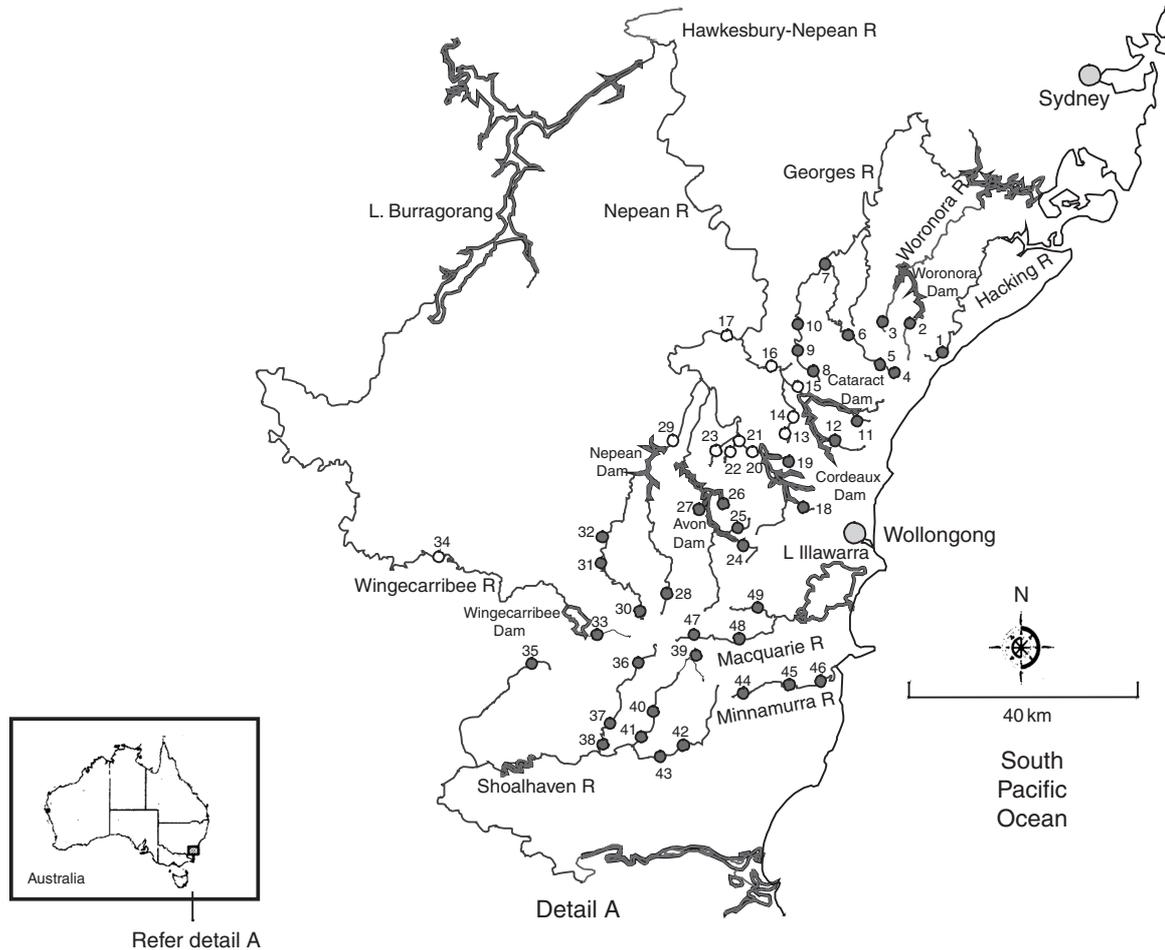


Fig. 1. Sampling sites (1–49) from six catchments (Hacking, Georges, Hawkesbury-Nepean, Shoalhaven, Minnamurra and Macquarie) in the Sydney Water Supply Catchment (SWSC). Hollowed circles indicate sites below dam impoundments.

The family Leptophlebiidae is the most diverse Australian mayfly family, with 16 described genera and in excess of 60 described species. Larvae of the genus *Atalophlebia* are found under rocks and debris of streams and billabongs throughout Australia (Dean 1999b). Dean (1999b) recognised 21 morphospecies (two formally named) of which 12 have been recorded from New South Wales. Before genetic analysis we identified larvae using the morphological key in Dean (1999b).

The diversity data from the crayfish, caddisfly and mayfly genera are presented here for the first time. However, we also analysed the *Paratya* data taken from Baker *et al.* (in press), for the first time in a phylogenetic context, as a point of comparison to data from the other genera.

There is presently only a single described species of the genus *Paratya* in Australia, *P. australiensis*. This species has a wide distribution in upland and lowland streams and non-flowing freshwaters, throughout several thousand kilometres of Australia's east coast (Williams 1977). However, in a previous connectivity study Baker *et al.* (in press) reported populations of these shrimp in the SWSC to be highly structured and to contain several divergent types.

Specimen collection

We sampled macroinvertebrates from six river catchments within the SWSC: Hacking, Georges, upper Nepean, Macquarie, Minnamurra and Shoalhaven. Forty-nine sites were sampled (Fig. 1; Table 1).

Euastacus were collected using a dip net or seine net and one of the first pereopods was removed for genetic analysis. Animals were then returned to the site of capture (pereopods are routinely lost in conspecific agonistic interactions and regrow quickly).

Cheumatopsyche late instar larvae were collected with a dip net by kick-sampling riffles.

Atalophlebia late instar larvae were collected from rocks and stream leaf litter in still pools, using featherlight forceps.

Paratya were collected as described in Baker *et al.* (in press).

All samples were immersed under liquid nitrogen and transferred back to the laboratory where they were held at -80°C , until required for analysis. For each morphologically identified species at each site, several individuals were randomly selected for genetic screening. Representative taxa from several sites were stored in 70% ethanol in the field for subsequent identification under high-powered stereo and compound microscopes.

Genetic typing

Genetic typing protocols for *P. australiensis* are described in Baker *et al.* (in press).

In the present study, we isolated total genomic DNA for all the taxa using a modification of the cetyltrimethylammonium bromide (CTAB)/phenol-chloroform DNA extraction protocol (Doyle and

Doyle 1987). Polymerase chain reaction (PCR) of *Euastacus* spp. cytochrome *c* oxidase subunit I (COI) and 16s mitochondrial regions were undertaken using primer sets LCO1490/HCO2198 (Folmer *et al.* 1994) and 16sar/16sbr (Palumbi *et al.* 1991), respectively. Polymerase chain reaction of *Cheumatopsyche* spp. was undertaken using the aforementioned Folmer *et al.* (1994) COI primers. Since amplification with these COI primers was unsuccessful in *Atalophlebia* spp.,

primers COIf-L and COIa-H were used in this genus (Palumbi *et al.* 1991).

All PCR amplifications were undertaken in a 25 μ L total volume containing: 16.9 μ L ddH₂O, 1 μ L each of 10 μ M primer, 0.5 μ L 10 mM dNTPs, 1 μ L 50 mM MgCl₂, 2.5 μ L 10 \times Reaction Buffer, 0.1 μ L DNA *Taq* polymerase and 1 μ L template DNA. Polymerase chain reaction was performed on the Geneamp PCR System 9700 (PE Applied Biosystems,

Table 1. Numbers of individuals sequenced corresponding to each lineage, for each genus, at each site

Blank cells indicate none of the specified genus was sequenced at that site. Lineages, denoted by upper case letters (in brackets) correspond to those shown in figures. All *Cheumatopsyche* sp. AV1 sequences were obtained from Baker *et al.* (2003) and all *Paratya* sequences were obtained from Baker *et al.* (in press)

Site number	<i>Euastacus</i>	<i>Cheumatopsyche</i>	<i>Atalophlebia</i>	<i>Paratya</i>
1		4(AV1)	2(B)	
2	1(E)			6(A), 15(C)
3	5(A)			
4				6(A), 17(B)
5	5(A)			4(A), 15(B)
6				22(A)
7				20(B)
8	10(A)		3(D)	18(A)
9	1(E)			6(A), 16(B)
10		11(AV1), 1(AV2)		1(A), 19(B)
11			2(B)	16(A)
12	1(E)			24(A)
13				16(A)
14				21(A)
15	1(E)	14(AV1)		
16		10(AV1)		1(A), 19(B)
17		1(AV1), 2(AV4), 8(AV7)		17(A)
18				13(A)
19				19(A)
20		12(AV1)		22(A)
21		16(AV1)		
22				23(A)
23				18(A)
24	1(B), 1(E)			
25	2(B)		1(D)	
26				15(A), 3(B), 5(C)
27				14(A), 2(B), 2(C)
28				17(A)
29		12(AV1)		3(A), 4(B), 15(C)
30	5(D)	14(AV1)		
31	2(E)	14(AV1)		2(A), 10(B), 10(C)
32	1(E)	13(AV1)		21(A)
33	15(D)			
34		1(AV1), 8(AV2), 1(AV4)		13(B), 8(C)
35				2(B), 3(D)
36	4(C)	3(AV1)		1(A), 2(D)
37		4(AV1), 6(AV4), 1(AV7)		
38		8(AV1)		2(D)
39	1(C)			
40		11(AV1)	1(C)	20(D)
41		12(AV1)		
42				1(B), 1(D)
43		7(AV1)		1(D)
44		8(AV1)		
45		10(AV1)		
46		6(AV1)		
47		12(AV1)	1(A), 1(B)	
48		5(AV1)		
49		8(AV1), 1(AV4)	1(A)	

Foster City, CA, USA) with an initial denaturation step of 94°C for 3 min; 30 cycles of 94°C for 30 s, 50°C for 30 s, 72°C for 30 s; and a final extension step of 72°C for 5 min.

Polymerase chain reaction products were purified using the QIAquick PCR purification kit (Qiagen, Victoria, Australia) and directly sequenced using the PCR light strand primers on an Applied Biosystems 377 automated sequencer. Representative haplotypes for each group of taxa were sequenced with the relevant heavy strand PCR primer, with complete overlap, to ensure DNA-strand homology.

Statistical analyses

Sequences were aligned using BioEdit Version 5.0.9 (Hall 1999). Exploratory data analysis of sequences was performed using MEGA Version 2.1 (Kumar *et al.* 2001). Estimates of mean pairwise sequence diversity \pm s.e.m. were made using 5000 bootstraps of the data.

Modeltest Version 3.06 was used to determine the DNA substitution model and parameters of best fit for the sequence data (Posada and Crandall 1998).

In the present study, we adhered to the view of Avise (1994) who asserted that it is advisable to compare and contrast multiple methods in phylogeny reconstruction, particularly if these entail philosophically different approaches (e.g. distance *v.* character-state). In the present study therefore, we used four different phylogenetic construction methods to ensure concordance among resolved mitochondrial lineages: Bayesian (Rannala and Yang 1996); maximum likelihood (Felsenstein and Churchill 1996); maximum parsimony (Eck and Dayhoff 1966); and neighbour joining (Saitou and Nei 1987). We present here only the Bayesian trees in the form of figures; results from the other tree construction methods are compared and contrasted descriptively with the Bayesian tree, for each genus. In the case of *Euastacus*, trees were inferred for 16S and COI data separately and for a combined COI/16S dataset. Before phylogenetic analysis of the combined dataset, a partition homogeneity test was carried out in PAUP* version 4.0b10 (Swofford 2001) to determine if significantly different signals were being generated by the COI and 16S fragments.

For all groups of taxa, a Bayesian tree was inferred from the sequences using MrBayes Version 2.01 (Huelsenbeck and Ronquist 2001). MrBayes was invoked using codon partitioning, empirical base frequencies, unconstrained branch length, eight rate categories on the gamma distribution, one million cycles with trees sampled every 100 cycles, and eight chains (seven heated). The programme was repeated at three different chain temperature settings (0.1, 0.2 and 0.5), to ensure topology and posterior probability concordance. Neighbour joining,

maximum parsimony and maximum likelihood trees were also inferred, using PHYLIP Version 3.6a2 (Felsenstein 1993). The neighbour joining method was invoked using: the F84 model of DNA substitution, gamma distributed rates, 1000 bootstrap pseudoreplicates of the data (block size 3) and 10 jumbles. Block size was used to partition the sequence data into codons and the jumble option randomises the input order of taxa during tree construction. The maximum parsimony method was invoked using: search for best tree, most thorough search, 1000 bootstrap pseudoreplicates of the data (block size 3) and one jumble. The maximum likelihood method was invoked using: empirical base frequencies, search for best tree, 100 bootstrap pseudoreplicates of the data (block size 3), one jumble of the data, and three categories in the hidden Markov model.

In the consensus trees, we designated groups of haplotypes as 'lineages'. We used an objective criterion to define a 'lineage', as follows: working towards the root, from the tips of the tree, haplotypes cluster into progressively larger clades. A 'clade' of sampled haplotypes was designated a 'lineage' when the posterior probability (expressed as a percentage) at the coalescent node was greater than 70 and the pairwise per cent divergence, between the most basal haplotype member of the clade and the nearest neighbour haplotype outside the clade, was 2% or greater. This procedure was followed until all sampled haplotypes at the tips of the tree became grouped as members of a designated lineage.

Results

For clarity, descriptive statistics of sequence sets from each genus are shown in Table 2. GenBank accession numbers for comparative ingroup and outgroup sequences in phylogenetic reconstructions are shown in Table 3.

Euastacus

Divergence

Mean COI sequence divergence between pairs of lineages ranged from 4.12% \pm 0.81 (A–B) to 11.24% \pm 1.32 (C–D). Mean 16S sequence divergence between pairs of lineages ranged from 1.85% \pm 0.56 (A–B) to 6.19% \pm 1.05 (A–D).

Phylogenetic reconstruction

Cytochrome *c* oxidase subunit I trees were constructed under the HKY + G model of sequence evolution with

Table 2. Descriptive statistics for macroinvertebrate sequences screened from the study area

Genus	Gene region	Sequence length	No. haplotypes	GenBank accession	No. individuals	Indels	Variable sites	Phylogenetically informative sites			Ti:Tv ratio
								1st codon	2nd codon	3rd codon	
<i>Euastacus</i>	COI	531 bp	19	AY380460-78	56	No	99	9	0	83	2.8 : 1
	16S	485 bp	10	AY380519-28	15	No	45		37 non-coding		7.7 : 1
<i>Cheumatopsyche</i>	COI	549 bp	14	AY380481-7 AY380488-94 ^A	256 (216 ^A)	No	137	28	1	105	1.3 : 1
<i>Atalophlebia</i>	COI	564 bp	12	AY380506-17	12	No	166	10	0	120	2.5 : 1
<i>Paratya</i>	COI	456 bp	34	AY308075-105 ^B AY308109 ^B AY308120 ^B AY308172 ^B AY308174 ^B	569 ^B	No	69	8	0	45	2.3 : 1

COI, cytochrome *c* oxidase subunit I; Ti/Tv, transition/transversion.

^AThose sequences taken from Baker *et al.* (2003), ^BSequences taken from Baker *et al.* (in press).

Table 3. GenBank accession numbers of sequences used as comparative ingroup taxa in phylogeny reconstructions

Taxa group (genus)	Species	Gene region	GenBank accession	Remarks
Crayfish (<i>Euastacus</i>)	<i>E. claytoni</i>	COI	AY380479	Queensland Museum RegNo. QMW26600
	<i>E. dharawalus</i>	COI	AY380480	Queensland Museum RegNo. QMW26607
	<i>E. bispinosus</i> ^A	COI	AF493634	–
Crayfish (<i>Cherax</i>)	<i>C. destructor</i> ^A	COI	AY153887	–
Crayfish (<i>Euastacus</i>)	<i>E. claytoni</i>	16S	AY383598	Queensland Museum RegNo. QMW26600
	<i>E. dharawalus</i>	16S	AY380529	Queensland Museum RegNo. QMW26607
	<i>E. bispinosus</i> ^A	16S	AF235991	–
	<i>E. bispinosus</i> ^A	16S	AF492813	–
	<i>E. yarraensis</i> ^A	16S	AF044245	–
	<i>E. armatus</i> ^A	16S	AF044242	–
	<i>E. armatus</i> ^A	16S	AF135986	–
	<i>E. rieki</i> ^A	16S	AF135984	–
	<i>E. bidwalus</i> ^A	16S	AF135987	–
	<i>E. suttoni</i> ^A	16S	AF135989	–
	<i>E. hystricosus</i> ^A	16S	AF135988	–
	Crayfish (<i>Astacopsis</i>)	<i>A. gouldi</i> ^A	16S	AF135969
Caddisfly (<i>Cheumatopsyche</i>)	sp. AV3	COI	AY380495-97	–
	sp. AV5	COI	AY380498-500	–
Caddisfly (<i>Asmicridea</i>)	spp.	COI	AY380501-2	–
Caddisfly (<i>Taschorema</i>)	spp.	COI	AY380503-4	–
Caddisfly (<i>Economus</i>)	sp.	COI	AY380505	–
Mayfly (<i>Ulmerophlebia</i>)	sp.	COI	AY380518	–

COI, cytochrome *c* oxidase subunit I. ^ADownloaded from GenBank, the remainder were sequenced in this study.

gamma shape parameter 0.1824. 16S trees were constructed under the HKY + I + G model of sequence evolution with a proportion of 0.4415 invariable sites and gamma shape parameter 0.2938. The partition homogeneity test indicated there was no significant difference in signals generated by the COI and 16S datasets ($P = 1.00$). The combined dataset phylogeny was constructed under the HKY + G model of sequence evolution with gamma shape parameter 0.0136.

The Bayesian trees (Fig. 2a–c) were consistent and recovered five well-supported divergent lineages (A–E). ML, NJ and MP trees (not shown) recovered the same five lineages with bootstrap support values of at least 75% (note that several GenBank sequences from *Euastacus* species are included in these tree constructions).

The representative *Euastacus dharawalus* sequence formed a well-supported monophyletic group with haplotypes 3 and 4 in the 16S tree and haplotypes 14 and 15 in the COI tree, indicating that lineage D corresponded to *E. dharawalus*.

Separate COI and 16S NJ trees (not shown) with all haplotype sequences from this study combined with sequences representing most of the 37 recognised *Euastacus* species, indicated strong (>70%) monophyletic bootstrap support for lineage A and E haplotypes with *E. australasiensis* and *E. spinifer*, respectively (K. Crandall, unpublished data). Our lineages B and C, however, formed separate well-supported monophyletic groups, with respect to all included known species sequences. It is conceivable that one of these lineages represents *E. hirsutus*, which is known from an extremely

limited range in the upper Shoalhaven (Merrick 1993), but types of this species were unfortunately not available for inclusion in the phylogeny.

Distribution

Euastacus were found at 15 sites and with one exception, all of these sites were in headwater streams above dam impoundments (Fig. 2d).

Cheumatopsyche

Divergence

Mean COI sequence divergence between pairs of lineages ranged from 7.53% ± 1.08 (AV3–AV5) to 12.93% ± 1.39 (AV3–AV7).

Phylogenetic reconstruction

Cytochrome *c* oxidase subunit I trees were constructed under the GTR + G model of sequence evolution with gamma shape parameter 0.3529. The Bayesian tree recovered six (AV1, AV2, AV3, AV4, AV5, AV7) well-supported divergent monophyletic *Cheumatopsyche* lineages, which matched species designations based on larval morphology (Fig. 3a). ML, NJ and MP trees (not shown) recovered the same six lineages with bootstrap support values of at least 75%. Note that 31 *Cheumatopsyche* sp. AV1 haplotypes had previously been described by Baker *et al.* (2003) in a population genetic study carried out in the region, and we included a randomly chosen sample of these in the present phylogenetic reconstruction, for comparative purposes.

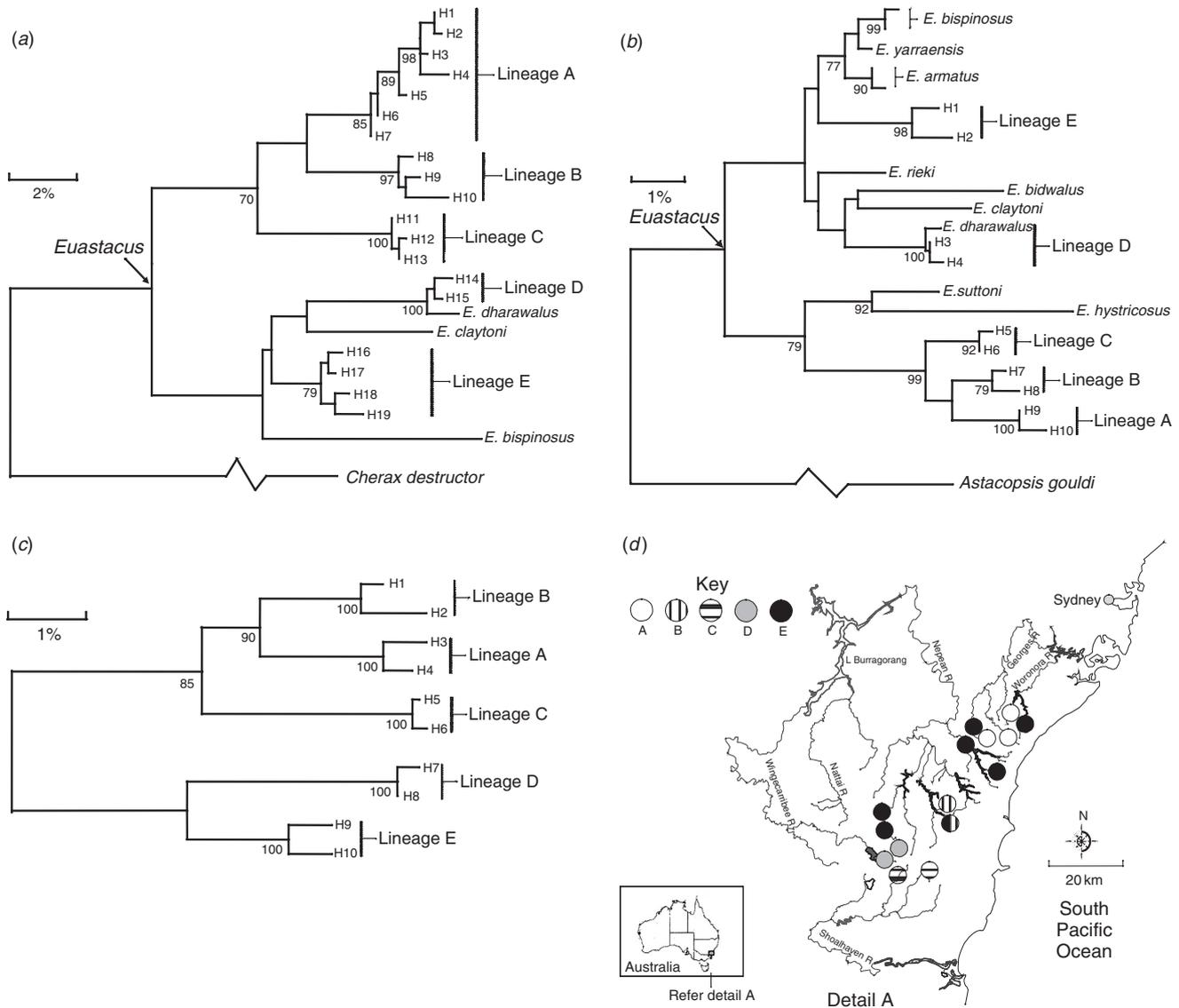


Fig. 2. (a–c) Bayesian phylogenies for the genus *Euastacus* using cytochrome *c* oxidase subunit I (COI), 16S and combined COI/16S, respectively. Posterior probabilities (converted to percentages) of 70 or above are printed below nodes. Lineages (A–E) and component haplotypes (H1–H19) delineate sequences from the present study. (d) Pie charts showing distributions of each lineage in the study area (note that proportion of each component lineage is not indicative of relative frequency).

Distribution

Cheumatopsyche were collected from 24 sites and four recognised *Cheumatopsyche* species were recovered (AV1, AV2, AV4 and AV7). Our sample contained predominantly *Cheumatopsyche* sp. AV1, although 2–3 taxa were sympatric at five sites (Fig. 3b).

Atalophlebia

Divergence

Mean COI sequence divergence between pairs of lineages ranged from 9.87% ± 1.17 (A–C) to 18.74% ± 1.60 (D–E).

Phylogenetic reconstruction

Cytochrome *c* oxidase subunit I trees were constructed under the HKY + G model of sequence evolution with gamma shape parameter 0.1603. The Bayesian tree recovered four well-supported divergent monophyletic lineages (A–D; Fig. 4a). ML, NJ and MP trees (not shown) recovered the same four lineages with bootstrap support values of at least 75%. Based on morphology, lineages A–C keyed out to *Atalophlebia* sp. AV13, but in the mtDNA tree formed three well-supported extremely divergent cryptic lineages nested within a well-supported AV13 clade. Lineage D exactly matched *Atalophlebia* sp. AV2 morphologically and

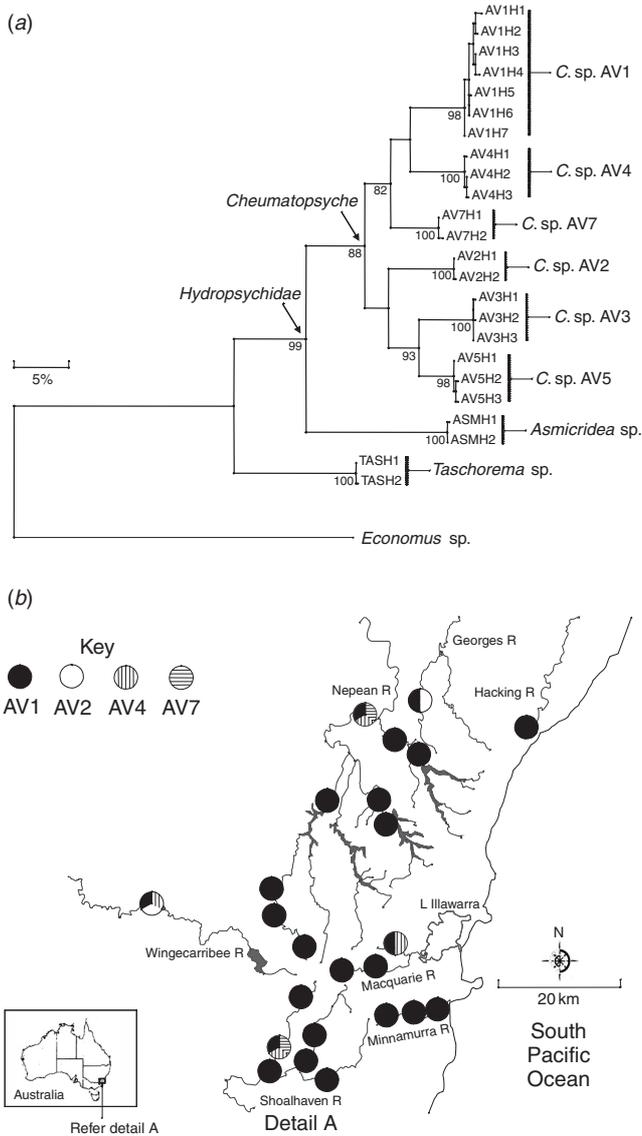


Fig. 3. (a) Bayesian cytochrome *c* oxidase subunit I (COI) phylogeny for the genus *Cheumatopsyche*. Posterior probabilities (converted to percentages) of 70 or above are printed below nodes. Species (AV1, AV2, AV4 and AV7) and component haplotypes (H) delineate *Cheumatopsyche* sequences from the Sydney Water Supply Catchment (SWSC). (b) Pie charts showing distributions of each species in the study area (note that proportion of each component lineage is not indicative of relative frequency).

in the tree formed a well-supported group containing four haplotypes.

Distribution

Although *Atalophlebia* were abundant at many sites above dam impoundments, we were only able to sequence representative samples from seven sites. Lineages A and B were sympatric at one site in the Macquarie River (Fig. 4b).

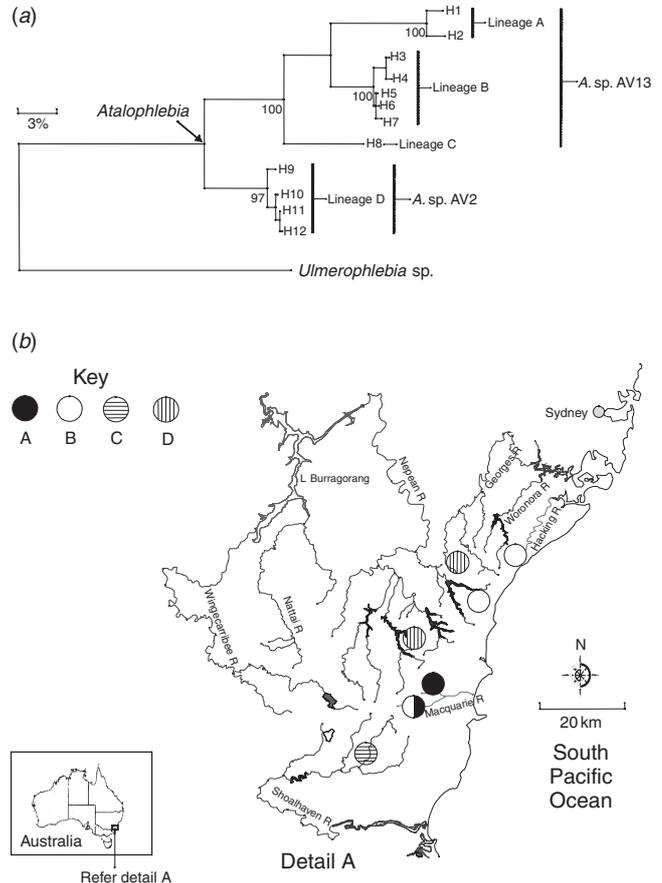


Fig. 4. (a) Bayesian cytochrome *c* oxidase subunit I (COI) phylogeny for the genus *Atalophlebia*. Posterior probabilities (converted to percentages) of 70 or above are printed below nodes. Lineages (A–D) and component haplotypes (H1–H12) delineate sequences. (b) Pie charts showing distributions of each lineage in the study area (note that proportion of each component lineage is not indicative of relative frequency).

Paratyta

Divergence

Mean COI sequence divergence between pairs of lineages ranged from 3.78% ± 0.85 (A–D) to 5.82% ± 1.04 (B–D).

Phylogenetic reconstruction

Cytochrome *c* oxidase subunit I trees were constructed under the TRN + G model of sequence evolution, with gamma shape parameter 0.1642.

The Bayesian tree recovered four well-supported divergent cryptic lineages of *P. australiensis* (A–D; Fig. 5a). ML, NJ and MP trees (not shown) recovered the same four lineages with bootstrap support values of at least 75%.

Distribution

Paratyta were ubiquitous throughout the study area and individuals were obtained from 32 sites. Two to three lineages were sympatric at 14 sites (Fig. 5b). Lineage A *P. australiensis*

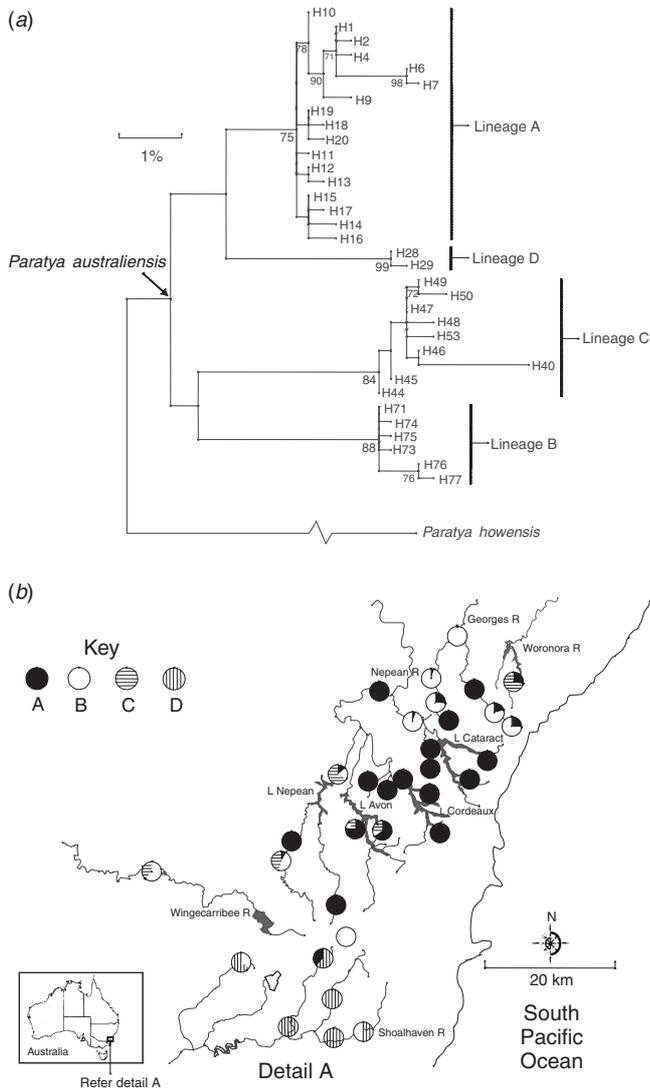


Fig. 5. (a) Bayesian cytochrome *c* oxidase subunit I (COI) phylogeny for the genus *Paratyta*. Posterior probabilities (converted to percentages) of 70 or above are printed below nodes. Lineages (A–D) and component haplotypes (H) delineate sequences. (b) Pie charts showing distributions of each species in the study area (note that proportion of each component lineage is representative of relative frequency).

was the most commonly encountered of the four divergent lineages and appears to be most prevalent in streams above dam impoundments.

Discussion

Our results indicate substantial mitochondrial phylogenetic diversity within the genera of freshwater macroinvertebrate taxa surveyed, considering the relatively small geographic area from which samples were taken.

Previous ecologically based studies of river macroinvertebrates by Recher *et al.* (1993) and Chessman and Williams (1999) also reported a rich and diverse fauna in the

Hawkesbury-Nepean River basin. Chessman and Williams (1999) found that although areas around Sydney that have suffered from urbanisation do indeed support impoverished macroinvertebrate faunas, the Hawkesbury-Nepean basin as a whole supports 443 recorded species and morphospecies. Their study focussed on the area to the west and north of Sydney and did not include streams from the upper reaches of the Georges or Nepean catchments. However, Grown and Grown (2001) carried out a similar survey of the macroinvertebrate taxa extending upland into areas both below and above the dams, in these two catchments. They recorded 297 macroinvertebrate taxa, although the fauna were recognised to be considerably more diverse than this, since taxa were identified to genus level only. Moreover, they found that sites sampled below the dams supported a significantly lower number of taxa, and a different suite of taxa was associated with unregulated sites in upland streams.

Several sites in this study were in the same stream reaches in the Nepean and Georges catchments as those examined by Grown and Grown (2001). In accordance with their study, we found evidence that the unregulated headwater streams appear to possess unique fauna, suggesting that this area is of high conservation significance. For example, although freshwater shrimp (*Paratyta*) were ubiquitous throughout the study area, lineage A *Paratyta australiensis* was the most commonly encountered of the four divergent lineages and appears to be most prevalent in streams above dam impoundments. Similarly, lineage D *P. australiensis* was only found in the upper reaches of the Shoalhaven River catchment, which is unregulated. In the present study, spiny crayfish (*Euastacus*) were collected from 15 sites spanning the Shoalhaven, Nepean and Georges river catchments. With the exception of one site, all localities where crayfish were encountered were in headwater stream reaches above dam impoundments. Two of these crayfish taxa (Lineages B and C), were each only found at two sites. Lineage B was found at two sites in the Shoalhaven catchment and lineage C was found at two sites in the Nepean catchment. The mayfly genus *Atalophlebia* was also primarily encountered in unregulated stream reaches, a pattern observed by Grown and Grown (2001). In contrast, we found *Cheumatopsyche* at most sites throughout the study area, with the exception of the upper Georges catchment, which is predominantly bedrock substrate and generally unsuitable for the larvae. *Cheumatopsyche* were particularly abundant at outflow sites below dams. In fact, we found a generally high diversity of trichopteran genera throughout the study area at water release sites immediately below the Cataract, Cordeaux and Nepean dams (data not shown), probably in association with high nutrient loads in the riffle at these collection points. These results are also in keeping with the study by Grown and Grown (2001), which found the genus *Cheumatopsyche* to be indicators of water supply release sites and associated with high water flow-through areas.

The observations regarding stream position of macroinvertebrate lineages in this study, were formally tested within the Nepean catchment for association with locality above or below dam impoundments. We compared the total number of mitochondrial lineages present within all four macroinvertebrate genera at each sampled site in the Nepean catchment, and found no significant association with stream position (above/below dam impoundments; Mann–Whitney *U*-test, $Z = -0.689$, $P > 0.05$). However, when the caddisflies (*Cheumatopsyche*) were excluded from the dataset, we found significantly more lineages of the other three genera in upstream sites (Mann–Whitney *U*-test, $Z = -2.105$, $P < 0.05$).

The shrimp (*Paratya*) and mayfly (*Atalophlebia*) genera we examined harboured highly divergent cryptic lineages, and two of the spiny crayfish (*Euastacus*) lineages (B and C) could not be matched to any recognised taxa, in a phylogeny containing representative type sequences of Australia's 37 *Euastacus* species (K. Crandall, unpublished data). Together, these data further emphasise the rich diversity of fauna supported within headwater streams in the catchment. Interestingly, populations of the psephenid beetle genus *Sclerocyphon* occurring in the SWSC also harbour a large number of cryptic taxa. Mitochondrial (COI) data indicate the presence of at least nine highly (~8–12%) divergent taxa in our study area alone. These lineages correspond to only three morphologically recognised species and allozyme evidence suggests that several the cryptic taxa are indeed reproductively isolated (A. Wheatley, unpublished data). Of the four groups of taxa studied here, caddisflies (*Cheumatopsyche*) were the only taxa for which current morphologically based designations closely matched the mitochondrial genetic data.

Although there is ongoing debate about the validity of various species concepts, detailed comparative studies incorporating morphological, ecological and genetic data are perhaps most likely to provide a precise diagnosis of species boundaries. For example, over 100 new old-world tree frog species from Sri Lanka were recognised after extensive analyses of mitochondrial DNA, morphology, ecology and bioacoustics, where previously only 18 species were known (Meegaskumbura *et al.* 2002).

In this regard, it is important to recognise that in the present study we have compared data from a single genetic locus, and that the observed mitochondrial gene trees may be incongruent with the organismal species trees for the identified crayfish, shrimp and mayfly lineages. The deep divergence and sympatry of several mitochondrial lineages detected in the present study is certainly consistent with a scenario of historical genetic isolation, evolution along separate trajectories and (in some cases) subsequent recontact, but does not necessarily rule out post-separation (secondary) gene flow.

Although the mitochondrial lineages described here could be designated species under the criterion of a monophyletic species concept (Mishler and Brandon 1987), this approach

is problematic for several reasons (see, de Queiroz and Donoghue 1988; Mishler and Theriot 2000). We view the present mitochondrial dataset as providing important baseline information about macroinvertebrate genetic diversity and a necessary first step towards determining species diversity within these genera. In regard to diagnosing species, we adhere to the opinion of Lee (2003) who asserted that detection of reproductive isolation between sympatric lineages would provide a sharper criterion to objectively define species status. Further, in terms of future management of the SWSC, morphological discrimination between these taxa is desirable.

One way to detect reproductive isolation in future studies would be through formal comparative analysis of nuclear (allozyme or other) markers, to reveal whether diagnostic loci can be used as evidence for an absence of present-day interbreeding between sympatric lineages of the studied macroinvertebrate taxa. Recently, Chenoweth and Hughes (2003) reported divergent cryptic mitochondrial lineages within the Australian atyid freshwater shrimp *Caridina indistincta* sp. B. However, they found no evidence of reproductive isolation between these sympatric lineages from allozyme assays. In contrast, for the shrimp *Paratya australiensis*, preliminary screening of 12 allozyme loci for individuals representing three divergent cryptic mitochondrial lineages (A, B and C), which are sympatric in the present study area, has provided evidence of fixed alternate alleles between lineages A and B/C at the *Aat-2* locus and allele frequency differences among all three lineages at four other loci (*Mpi-1*, *Pgi-1*, *Pgm-1* and *Pgd-1*; A. Baker, unpublished data). Formal allozyme analysis is required to corroborate these preliminary data and studies need to be conducted to assess morphological differences among *Paratya* exhibiting different mitochondrial lineages, to facilitate identification in the field.

In this study, the genus *Atalophlebia* was commonly encountered in headwater streams. Although we were only able to sequence a relatively small number of larvae, mitochondrial DNA diversity was extremely high. Preliminary examination by one of the authors (JCD) of larval morphology of taxa representing the cryptic lineages (A, B and C) of *Atalophlebia* identified in the present study, has indicated the presence of several morphological characters that could potentially be used to separate them. These differences remain to be formalised and incorporated into the current taxonomic key (Dean 1999b) for the genus *Atalophlebia*.

For the six species of *Cheumatopsyche*, our genetic data corroborate species designations based on larval morphology. However, most of these taxa are undescribed (see Dean 1999a); larvae need to be reared to study adult morphology and studies need to be conducted to determine the ecological requirements of these species.

Although we have yet to formalise morphological characters that could be used to separate the identified lineages

of *Euastacus*, pairwise levels of genetic divergence at both COI and 16S mitochondrial regions are the same order of magnitude as for pairs of recognised *Euastacus* species, suggesting that lineages have been genetically diverged for similar evolutionary timeframes.

Our findings have implications for future management of the SWSC. We have demonstrated high levels of mitochondrial DNA diversity within two groups of Crustacea as well as groups of Ephemeroptera and Trichoptera. Further genetic and taxonomic research needs to be conducted on these macroinvertebrate assemblages within and surrounding the SWSC, to determine the degree of endemism of these taxa. Several taxa are associated with unregulated headwater streams, suggesting that conservation efforts should focus on preserving the rich diversity of fauna in these areas. At present, public access to these areas is restricted. However, there are plans to augment sources of water supply to Sydney by constructing new dams or enlarging/upgrading existing ones. More importantly, a commercial coal seam lies beneath the headwaters of the Nepean catchment. Mining operations currently in progress produce subsidence and will potentially irreversibly alter drainage patterns and flow regimes within the SWSC.

A significant finding of our study was the high genetic diversity of spiny crayfish (*Euastacus*) in the region. Mitochondrial data suggest limited gene flow within and among sub-catchments for all species of *Euastacus* occurring in the SWSC (A. Baker, unpublished data), making this genus particularly vulnerable to catchment disturbance. During the preparation of this manuscript, several streams surrounding one of only two sites where *Euastacus* lineage B was found (Site 25) lost all surface water through cracked streambeds, resultant from mining activities. Maintaining the SWSC's rich biodiversity in the face of such disturbances will be the greatest challenge faced by management authorities in the future.

The comparative phylogenetic approach has utility for the strategic management of specific biodiversity 'hot spots' around the world. The phylogenetic diversity (PD) measure championed by Faith (1992), which uses phylogenetic patterns to predict feature diversity of sets of species, has been used recently to highlight high diversity and endemism within 10 genera of five beetle families in Australian forests (Faith *et al.* 2004). Similarly, Barker (2002) used this approach to set priorities for threatened New Zealand forest bird fauna and Sechrest *et al.* (2002) illustrated that almost 70% of the total amount of carnivore and primate evolutionary history was held in 25 biodiversity hotspots around the world.

In the future, a similar formalised approach could be adopted to assess phylogenetic diversity in the SWSC, after obtaining genetic information for a more extensive sample and wider range of the macroinvertebrate fauna occurring in the region. This would also permit a formal comparison

between species diversity (for which considerable data are already available in this region e.g. Growns and Growns 2001) and phylogenetic diversity, to address the central question of whether different estimates of 'biodiversity' relate to one another.

Acknowledgments

We thank Keith Crandall for use of his unpublished *Euastacus* sequences, Arlene Wheatley for unpublished *Sclerocyphon* data and Jemma Somerville for laboratory work. Peter Davie kindly provided access to Queensland Museum *Euastacus* specimens. James Fawcett, Mia Hillyer, Amanda Kotlash, Martin Krogh, Tim Page and Simon Williams provided invaluable field assistance. This research forms a component of a large project on connectivity of macroinvertebrates in the SWSC, funded by the Cooperative Research Centre for Freshwater Ecology.

References

- Avise, J. C. (1994). 'Molecular Markers, Natural History and Evolution.' (Chapman and Hall: New York, USA.)
- Baker, A. M., Hurwood, D. A., Krogh, M., and Hughes, J. M. (in press). Mitochondrial DNA signatures of restricted gene flow within divergent lineages of an atyid shrimp (*Paratya australiensis*). *Heredity*.
- Baker, A. M., Williams, S. A., and Hughes, J. M. (2003). Patterns of spatial genetic structuring in a hydropsychid caddisfly (*Cheumatopsyche* sp. AV1) from southeastern Australia. *Molecular Ecology* **12**, 3313–3324. doi:10.1046/J.1365-294X.2003.02011.X
- Barker, G. M. (2002). Phylogenetic diversity: a quantitative framework for measurement of priority and achievement in biodiversity conservation. *Biological Journal of the Linnean Society* **76**, 165–194. doi:10.1046/J.1095-8312.2002.00055.X
- Bininda-Emonds, O. R. P., Vazquez, D., and Manne, L. L. (2000). The calculus of biodiversity: integrating phylogeny and conservation. *Trends in Ecology & Evolution* **15**, 92–94. doi:10.1016/S0169-5347(99)01781-4
- Boyero, L. (2002). Insect biodiversity in freshwater ecosystems: is there any latitudinal gradient? *Marine and Freshwater Research* **53**, 753–755. doi:10.1071/MF01232
- Bunn, S. E., and Davies, P. M. (1990). Why is the stream fauna of southwestern Australia so impoverished. *Hydrobiologia* **194**, 169–176.
- Chenoweth, S. F., and Hughes, J. M. (2003). Speciation and phylogeography in *Caridina indistincta*, a complex of freshwater shrimps from Australian heathland streams. *Marine and Freshwater Research* **54**, 807–812. doi:10.1071/MF02168
- Chessman, B. C., and Williams, S. A. (1999). Biodiversity and conservation of river macroinvertebrates on an expanding urban fringe: western Sydney, New South Wales, Australia. *Pacific Conservation Biology* **5**, 36–55.
- Dean, J. C. (1999a). 'Preliminary Keys for the Identification of Australian Trichoptera Larvae of the Family Hydropsychidae.' (Cooperative Research Centre for Freshwater Ecology: Albury, Australia.)
- Dean, J. C. (1999b). 'Preliminary Keys for the Identification of Australian Mayfly Nymphs of the Family Leptophlebiidae.' (Cooperative Research Centre for Freshwater Ecology: Albury, Australia.)
- Doyle, J. J., and Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of leaf tissue. *Phytochemistry Bulletin* **19**, 11–15.

- Eck, R. V., and Dayhoff, M. O. (1966). 'Atlas of Protein Sequence and Structure 1966.' (National Biomedical Research Foundation: Silver Spring, MD, USA.)
- Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation* **61**, 1–10. doi:10.1016/0006-3207(92)91201-3
- Faith, D. P. (2002). Quantifying biodiversity: a phylogenetic perspective. *Conservation Biology* **16**, 248–252. doi:10.1046/J.1523-1739.2002.00503.X
- Faith, D. P., Reid, C. A. M., and Hunter, J. (2004). Integrating phylogenetic diversity, complementarity, and endemism for conservation assessment. *Conservation Biology* **18**, 255–261.
- Felsenstein, J. (1993). PHYLIP – Phylogeny Inference Package (Version 3.2). *Cladistics* **5**, 164–166.
- Felsenstein, J., and Churchill, G. A. (1996). A hidden Markov model approach to variation among sites in rate of evolution. *Molecular Biology and Evolution* **13**, 93–104.
- Feral, J. P. (2002). How useful are the genetic markers in attempts to understand and manage marine biodiversity. *Journal of Experimental Marine Biology and Ecology* **268**, 121–145. doi:10.1016/S0022-0981(01)00382-3
- Folmer, O., Black, M., Hoeh, W., Lutz, R., and Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* **3**, 294–299.
- Growns, I. O., and Growns, J. E. (2001). Ecological effects of flow regulation on macroinvertebrate and periphytic diatom assemblages in the Hawkesbury-Nepean River, Australia. *Regulated Rivers Research Management* **17**, 275–293. doi:10.1002/RRR.622
- Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**, 95–98.
- Horwitz, P. (1995). 'A Preliminary Key to the Species of Decapoda (Crustacea: Malacostraca) Found in Australian Inland Waters.' (Cooperative Research Centre for Freshwater Ecology: Albury, Australia.)
- Huelsenbeck, J. P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Biometrics* **17**, 754–755. doi:10.1093/BIOINFORMATICS/17.8.754
- Kumar, S., Tamura, T., Jakobsen, I. B., and Nei, M. (2001). MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics (Oxford, England)* **17**, 1244–1245. doi:10.1093/BIOINFORMATICS/17.12.1244
- Lake, P. S., Barmuta, L. A., Boulton, A. J., Campbell, I. C., and StClair, R. M. (1985). Australian streams and Northern Hemisphere: comparisons and problems. *Proceedings of the Ecological Society of Australia* **14**, 61–82.
- Lee, M. S. Y. (2003). Species concepts and species reality: salvaging a Linnaean rank. *Journal of Evolutionary Biology* **16**, 179–188. doi:10.1046/J.1420-9101.2003.00520.X
- Meegaskumbura, M., Bossuyt, F., Pethiyagoda, R., Manamendra-Arachchi, K., Bahir, M., Milinkovitch, M. C., and Schneider, C. J. (2002). Sri Lanka: an amphibian hotspot. *Science* **298**, 379. doi:10.1126/SCIENCE.298.5592.379
- Merrick, J. R. (1993). 'Freshwater Crayfishes of New South Wales.' (Linnean Society of New South Wales: Milson's Point, New South Wales, USA.)
- Mishler, B., and Brandon, R. N. (1987). Individuality, pluralism and the phylogenetic species concept. *Biology and Philosophy* **2**, 397–414.
- Mishler, B., and Theriot, E. C. (2000). The phylogenetic species concept: monophyly, apomorphy and phylogenetic species concepts. In 'Species Concepts and Phylogenetic Theory: a Debate'. (Eds Q. D. Wheeler and R. Meier.) pp. 44–54. (Columbia University Press: New York, USA.)
- Morgan, G. J. (1997). Freshwater crayfish of the genus *Euastacus* Clark (Decapoda: Parastacidae) from New South Wales, with a key to all species of the genus. *Records of the Australian Museum Supplement* **23**, 1–110.
- Moritz, C., and Faith, D. P. (1998). Comparative phylogeography and the identification of genetically divergent areas for conservation. *Molecular Ecology* **7**, 419–429. doi:10.1046/J.1365-294X.1998.00317.X
- Palumbi, S. R., Martin, A., Romano, S., McMillan, W. O., Stice, L., and Grabowski, G. (1991). 'The Simple Fools Guide PCR.' (Kewalo Marine Laboratory: University of Hawaii.)
- Posada, D., and Crandall, K. A. (1998). Modeltest: testing the model of DNA substitution. *Bioinformatics (Oxford, England)* **14**, 817–818. doi:10.1093/BIOINFORMATICS/14.9.817
- de Queiroz, K., and Donoghue, M. J. (1988). Phylogenetic systematics and the species problem. *Cladistics* **4**, 317–338.
- Rannala, B., and Yang, Z. H. (1996). Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference. *Journal of Molecular Evolution* **43**, 304–311.
- Recher, H. F., Hutchings, P. A., and Rosen, S. (1993). The biota of the Hawkesbury-Nepean catchment: reconstruction and restoration. *Australian Journal of Zoology* **29**, 3–41.
- Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**, 406–425.
- Sechrest, W., Brooks, T. M., da Fonseca, G. A. B., Konstant, W. R., Mittermeier, R. A., Purvis, A., Rylands, A. B., and Gittleman, J. L. (2002). Hotspots and the conservation of evolutionary diversity. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 2067–2071. doi:10.1073/PNAS.251680798
- Swofford, D. L. (2001). 'PAUP* Phylogenetic Analysis Using Parsimony (*and Other Methods) Version 4.' (Sinauer Associates: Sunderland, MA, USA.)
- Sydney Catchment Authority (2002). 'Draft Bulk Water Network Delivery Plan.' (Bulk Water Division: Sydney, Australia.)
- Williams, W. D. (1977). Some aspects of the ecology of *Paratya australiensis* (Crustacea: Decapoda: Atyidae). *Australian Journal of Marine and Freshwater Research* **28**, 403–415.

Manuscript received 12 March 2004; revised 25 May 2004; and accepted 7 June 2004.