

Catchments catch all: long-term population history of a giant springtail from the southeast Australian highlands – a multigene approach

R. C. GARRICK,* C. J. SANDS,*† D. M. ROWELL,‡ D. M. HILLIS§ and P. SUNNUCKS*¶

*Department of Genetics, La Trobe University, Bundoora, Vic. 3086, Australia, †School of Botany and Zoology, Australian National University, Canberra, ACT 0200, Australia, ‡Section of Integrative Biology and Center for Computational Biology and Bioinformatics, University of Texas, Austin, TX 78712, USA

Abstract

Phylogeography can reveal evolutionary processes driving natural genetic-geographical patterns in biota, providing an empirical framework for optimizing conservation strategies. The long-term population history of a rotting-log-adapted giant springtail (*Collembola*) from montane southeast Australia was inferred via joint analysis of mitochondrial and multiple nuclear gene genealogies. Contemporary populations were identified using multilocus nuclear genotype clustering. Very fine-scale sampling combined with nested clade and coalescent-based analyses of sequences from mitochondrial cytochrome oxidase I and three unlinked nuclear loci uncovered marked population structure, deep molecular divergences, and abrupt phylogeographical breaks over distances on the order of tens of kilometres or less. Despite adaptations that confer low mobility, rare long-distance gene flow was implicated: novel computer simulations that jointly modelled stochasticity inherent in coalescent processes and that of DNA sequence evolution showed that incomplete lineage sorting alone was unable to explain the observed spatial-genetic patterns. Impacts of Pleistocene or earlier climatic cycles were detected on multiple timescales, and at least three putative moist forest refuges were identified. Water catchment divisions predict phylogeographical patterning and present-day population structure with high precision, and may serve as an excellent surrogate for biodiversity indication in sedentary arthropods from topographically heterogeneous montane temperate forests.

Keywords: *Collembola*, comparative phylogeography, Great Dividing Range, Pleistocene forest refugia, short-range endemism, *Tallaganda*

Received 9 June 2006; revision accepted 15 September 2006

Introduction

Reconstructing the long-term population history of diverse taxa from the same ecological community can yield generalities about evolutionary processes that generate and maintain biodiversity that extend to unstudied codistributed organisms (Avice 2000). Sedentary invertebrates may be exceptional in

their ability to capture geographical patterning in biota resulting from long-acting processes such as climatic cycles. Viable populations can persist in isolated refugia too small to support vertebrates, and more ancient phylogeographical patterns are less likely to be obscured or completely overwritten by Quaternary glaciation-induced local extinctions (Heatwole 1987; Moritz *et al.* 2001; Hugall *et al.* 2002; Yeates *et al.* 2002). Although highly restricted dispersal is conducive to historical inference (Templeton 1998), phylogeographical studies of animals with low vagility remain rare (Pfenninger & Posada 2002), and comparative phylogeographical studies conducted over very fine spatial scales are rarer still.

Short-range endemics – species with naturally small distributions (< 10 000 km²) – are prevalent in the Australian

Correspondence: Ryan Garrick, Fax: +61 394792480; E-mail: r.garrick@latrobe.edu.au.

†Present address: British Antarctic Survey, Cambridge CB3 0ET, UK

¶Present address: Australian Centre for Biodiversity: Analysis, Policy & Management, School of Biological Sciences, Monash University, Clayton, Vic. 3800, Australia

fauna. Many appear to be wet-adapted ancient Gondwanan relicts with similar ecological and life-history characteristics, most notably, poor powers of dispersal (Harvey 2002). Effective conservation planning may be undermined by the lack of baseline information on microgeographical patterns of endemism in low-mobility taxa. Indeed, current forest management practices are probably not optimal for a large proportion of biodiversity (Garrick *et al.* 2006).

To address this, we are studying phylogeography of rotting-log-dependent (saproxylic) invertebrate assemblages over a 100-km section of Tallaganda National Park and State Forest (collectively 'Tallaganda' herein) on the Gourock Range, an isolated section of the Great Dividing Range, in southeast New South Wales, Australia. We aim to determine whether whole communities of codistributed short-range endemic invertebrates exhibit concordant fine-scale regional structure as a result of shared dependence on moist forest habitats coupled with similar responses to past climatic cycles.

Phylogeography of a saproxylic springtail

Here we investigate the phylogeography of a new species of *Acanthanura* (Collembola: Neanuridae: Uchidanurinae), to be described elsewhere (P. Greenslade, in preparation). This unusually large-bodied Collembolon (up to 9 mm) currently appears to be restricted to Tallaganda. Mobility of the Uchidanurinae is thought to be extremely limited because, being soft-bodied and dorsoventrally flattened, they are vulnerable to desiccation (Greenslade *et al.* 2002). The preferred microhabitat of *Acanthanura* sp. n. is large-diameter rotting *Eucalyptus* logs (R. Garrick, personal observation), which harbour plasmodial slime mould on which congeneric species graze (Greenslade *et al.* 2002). Log interiors remain very moist throughout the year (95–100% humidity, J. Woodman, unpublished data) and are buffered from extremes in temperature by up to 16.5 °C (C. Schmuki, unpublished data). Dependence on this stable microenvironment makes it likely that *Acanthanura* sp. n. tracked the changing distribution of closed-canopy moist forests during Pleistocene climatic cycles with high fidelity.

Climatic impacts on temperate Australian montane forest biota

Topographically heterogeneous montane landscapes that retained moist forest refuges during Pleistocene glacial cycles currently harbour many wet-adapted, short-range endemics (Harvey 2002). The Great Dividing Range of southeast Australia has a long history of geological stability [≥ 50 –70 million years, (Myr)] compared to mountain ranges in most other parts of the world, resulting in the preservation of ancient drainage systems (Frakes *et al.* 1987).

Accordingly, five microbiogeographical regions, defined *a priori* by major and minor catchment divisions, have been recognized at Tallaganda (Garrick *et al.* 2004; Fig. 1).

Low to moderate elevation temperate forests on the Great Dividing Range have been inferred as refuges for more vagile amphibians, reptiles, and invertebrates (Keast 1981 and references therein; Moritz *et al.* 2000, 2001), indicating recurrent impacts of climatic cycles on biota. A prolonged cool, dry phase extending from 6 to 2.5 My before present (Mybp) caused widespread contraction of rainforest (Bowler 1982), but not to the extent of preventing re-expansion from sheltered gullies during times of climate amelioration (Kershaw *et al.* 1991). Following an

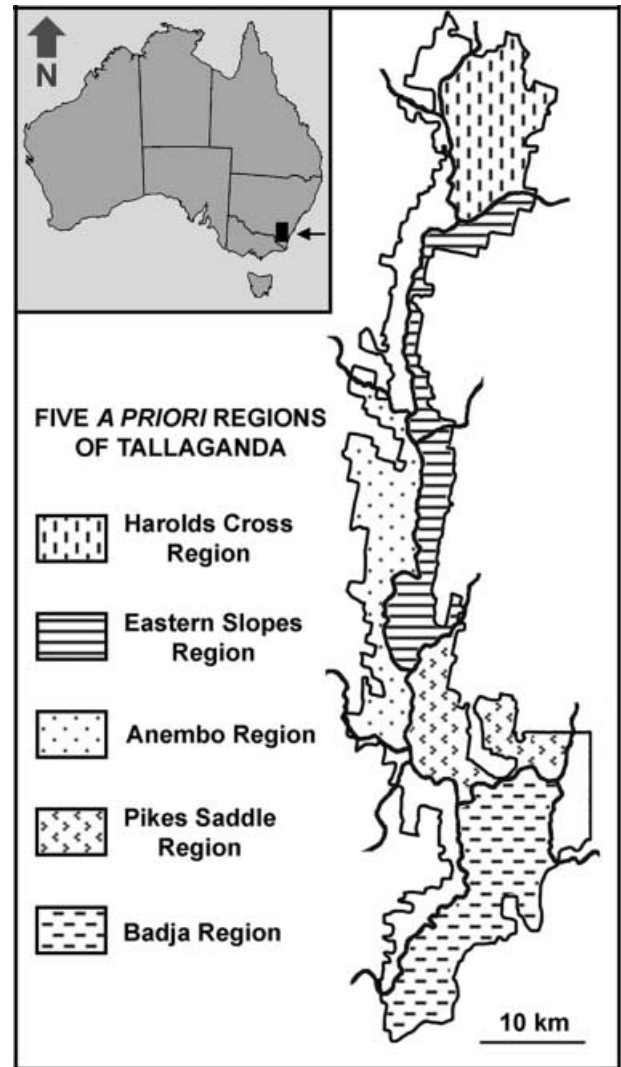


Fig. 1 Five *a priori* microbiogeographical regions of Tallaganda. Major and minor catchment divisions (black lines extending outside the perimeter of the study area) define regions described by Garrick *et al.* (2004). Affiliations of areas in white were not delineated in the original landscape model. Inset: map of Australia showing location of Tallaganda.

extended Pliocene weathering phase, rapid changes in composition of floristic assemblages occurred with the onset of climatic oscillations. However, in contrast to Europe and North America (Hewitt 2000), mainland Australia remained almost entirely free of ice sheet advances during Pleistocene glacials. Periglacial conditions and open herbaceous vegetation dominated higher elevations of Great Dividing Range (Bowler 1982; Barrows *et al.* 2001) permitting persistence of specific microhabitats, niches and local taxa. Ecologically, the past 400 000 years BP are considered particularly significant in shaping Australia's zoogeographical subregions (Bowler 1982). However, the two major periods of periglacial activity were restricted to relatively recent timescales (Early Kosciuszko glaciation, ~59 000 years BP, and the last glacial maximum, 17 000 to 20 000 years BP; Barrows *et al.* 2001, 2004). Based on the palaeoclimatic history of southeast Australia, short-range endemics might be expected to exhibit very fine-scale phylogeographical patterning of considerable antiquity.

This understanding of Pleistocene climatic impacts at Tallaganda allows us to generate spatially explicit, testable phylogeographical hypotheses for sedentary taxa (Materials and methods). Using a variety of genetic data and analytical techniques that exploit phylogeographical signal across a broad temporal spectrum, we assessed the degree to which catchments reflect long-term population history of *Acanthanura* sp. n.

Materials and methods

Prior expectations

- 1 Owing to repeated impacts of historical climatic cycles on the distribution of moist forest habitats across southeast Australia, we predict that impacts of historical climatic cycles will be detectable on multiple timescales.
- 2 Based on topographical and hydrological characteristics of Tallaganda, Garrick *et al.* (2004) proposed that the

Eastern Slopes Region (Fig. 1) contains the most potential good refuges for wet-adapted saproxylic invertebrates. Thus we expect high genetic diversity, deep molecular divergences, and ancestral/basal DNA sequences and lineages to be concentrated in this region.

- 3 Evidence for postglacial range expansion is expected within catchments likely to have experienced the greatest retraction (or total loss) of moist forests during Pleistocene periglaciation. Specifically, the Harolds Cross, Anembo, Pikes Saddle and Badja Regions (Fig. 1) are predicted to have been dominated by dry woodland or treeless steppe during cool dry periods (Garrick *et al.* 2004).
- 4 Given the long-term stability of Tallaganda's drainage divisions, we predict the five *a priori* microbiogeographical regions (Fig. 1) harbour distinct gene pools for low-mobility saproxylic invertebrates.

Table 1 summarizes how analytical approaches described below were used to address each of these prior expectations.

Taxon sampling and DNA isolation

Between 1997 and 2004, 206 individuals were collected from 72 rotting logs across Tallaganda. Genomic DNA was extracted using a 2× cetyltrimethyl ammonium bromide procedure, modified after Murray & Thompson (1980). Ethanol-preserved and slide-mounted voucher specimens will be lodged at the South Australian Museum following description of the species (P. Greenslade, in preparation).

Genes analysed, genotyping assays, and sequence alignment

We generated a dataset comprised of six nuclear DNA (nDNA) loci and a 570-bp fragment of the mitochondrial cytochrome oxidase I gene (mtCOI). Genotypic, genic, and genealogical information (*sensu* Sunnucks 2000) was obtained for three noncoding nDNA genes: Elongation factor-1α intron (UcEF-1α, 266 bp), and anonymous loci

Table 1 Summary of analytical methods used, the temporal scale under consideration, prediction(s) addressed, and associated *a priori* expectations. Population abbreviations are: HCR, Harolds Cross Region; AR, Anembo Region; PSR, Pikes Saddle Region; BR, Badja Region.

Analytical method	Temporal period	Prediction(s) addressed	Expected outcome
ML phylogenetic analysis	Ancient	Prior 2 Prior 4	ESR lineages basal and highly divergent Major clades localized within regions
Divergence time estimation	Ancient	Prior 1	Multiple non-overlapping date ranges
Nested clade analysis	Short/intermediate	Prior 1 Prior 3 Prior 4	Inferences indicate multiple temporal 'phases' Range expansion in HCR, AR, PSR and BR Inferences localized within regions
Population identification	Present	Prior 4	Populations localized within regions
Demographic changes	Short/intermediate	Prior 3	Growth in HCR, AR, PSR and BR

Uc180 (96 bp) and Uc3 (150 bp). Alleles were isolated from diploid polymerase chain reaction (PCR) products via single-stranded conformation polymorphism (SSCP) prior to sequencing, allowing phase of segregating sites to be observed directly. Restriction-fragment length polymorphism assays were employed for two nDNA loci: adenine nucleotide transporter intron (UcANT), and a member of wingless gene family (UcWnt). Insertion/deletion mutations (indels) were assayed for the anonymous nDNA locus Uc44. Marker development, primer sequences and genotyping protocols are given in Garrick & Sunnucks (2006). DNA sequences of all alleles observed at each nuclear locus have been deposited in GenBank (accessions DQ322529–DQ322590). Amplification and SSCP screening of mtCOI followed Garrick *et al.* (2004) and Sunnucks *et al.* (2000a), respectively. All sequencing was performed by Macrogen. Sequences were edited with reference to chromatographs and aligned using ALIGNR version 2.0 (LI-COR). Areas of ambiguous alignment were omitted. To confirm recovery of true mtCOI, haplotype sequences were translated, inspected for open reading frames, and compared with Lunt *et al.*'s (1996) model of insect mtCOI.

Phylogenetic analysis of mtCOI and estimating divergence times

To identify geographical locations of the deepest genetic subdivisions and assess the chronology of lineage branching order, we used traditional phylogenetic analysis of the single most information-rich gene available – mtCOI. This addressed prior expectations 4 and 2, respectively. To consider prior expectation 1, we used molecular dating to examine evidence for lineage divergences on multiple timescales.

The mtCOI phylogeny was estimated using maximum likelihood (ML) in PAUP* version 4.0b10 (Swofford 2002). *Acanthanura* sp. n. 2 from Wadbilliga National Park, New South Wales (P. Greenslade, in preparation) was used as an outgroup (GenBank accession DQ518749). The best-fit model of molecular evolution was determined via the Akaike Information Criterion using MODELTEST version 3.06 (Posada & Crandall 1998). To define the substitution model fully, parameters were initially estimated following a heuristic tree search using maximum parsimony (MP) with starting trees obtained via stepwise addition, 100 random addition sequence replicates, and tree-bisection–reconnection (TBR) branch-swapping in PAUP*. Likelihood analysis was then conducted using parameters estimated from the MP tree (heuristic search, one random addition sequence, TBR). The iterative procedure, with starting parameters obtained from the previous search, was repeated until successive ML searches converged on the same tree score, tree topology, and model parameter estimates (Swofford *et al.* 1996). Node support was assessed via 1000

bootstrap replicates (ML heuristic search with parameters fixed from the final tree, one random addition sequence, nearest-neighbour interchange branch-swapping).

The molecular clock hypothesis was assessed using Tajima's (1993) relative rate test in MEGA version 3.1 (Kumar *et al.* 2004). Absolute ages of divergences among major mtCOI clades with at least moderate (> 70%) bootstrap support on the ML phylogeny were estimated following Nei & Li (1979) (see also Takahata & Nei 1985) using uncorrected *P* distances. To calibrate the molecular clock we employed the commonly used rate for arthropod cytochrome oxidase genes 2.3% per Myr, corresponding to a mutation rate of 0.012 substitutions per site per lineage per Myr (Brower 1994). This rate has provided divergence date estimates that fit prior expectations in population- and species-level studies of Collembola (Garrick *et al.* 2004; Stevens *et al.* 2006, respectively).

Nested clade analysis

To distinguish recurrent population processes from historical events on short to intermediate temporal scales, we used nested clade analysis (NCA, Templeton *et al.* 1995). NCA is spatially explicit, enabling prior expectations 3 and 4 to be addressed. Because temporal chronology of phylogeographical inferences can be approximated without assuming a molecular clock (Templeton 2004), NCA is also appropriate for evaluating prior expectation 1.

Genealogical relationships among haplotypes/alleles at loci mtCOI, UcEF-1 α , Uc180 and Uc3 were estimated using Templeton *et al.*'s (1992) statistical parsimony (SP), implemented in rcs version 1.21 (Clement *et al.* 2000). There is no evidence for recombination and little or no evidence for selection affecting the three nuclear genes (Garrick & Sunnucks 2006). Contiguous alignment gaps (nuclear genes only) were treated as a fifth character, coded so that each identical indel represented a single event. Cladograms were constructed using the 95% confidence criterion (Templeton *et al.* 1992). Hierarchical nesting categories were assigned following Templeton *et al.* (1987), Templeton & Sing (1993) and Crandall (1996). The null hypothesis of no association between geography and genealogy was tested using a contingency χ^2 test in GEODIS version 2.2 (Posada *et al.* 2000). Significance was evaluated by permuting 10 000 times. When the null hypothesis was rejected ($P < 0.05$), Templeton's (2004) inference key (updated November 2005, available at <http://darwin.uvigo.es/software/geodis.html>) was used to make biological interpretations of relationships between D_c (clade distance) and D_n (nested clade distance), and I-T (interior-tip) contrasts. Each locus was analysed separately, and integration of inferences across loci was achieved using four criteria: (i) concordance of primary inferences; (ii) geographical concordance; (iii) no contradiction of the temporal sequence

of events/population processes inferred for each locus via I-T contrasts; and (iv) correspondence of inferences from lower-level nDNA nested clades with those from higher-level mtCOI nested clades (owing to presumed fourfold difference in effective population size, N_e). A caveat of this conservative approach for integrating inferences across loci is that the minimum number of events/processes and temporal 'phases' are inferred; this does not necessarily reflect high precision with respect to comparison of nuclear and mitochondrial genes.

We performed tests for exponential population growth using the g statistic (see Demographic changes, below) on datasets comprised of haplotypes/alleles from nested clades for which range expansion events were inferred (Results). This test is inappropriate when the assumption of panmixia is violated (our assessment was based on results from nuclear genotype clustering).

Contemporary population structure and partitioning of molecular variation

The present-day distribution of genetically distinct populations, and the degree of differentiation among them, was estimated using genotypic and genic data, respectively. These analyses address prior expectation 4, focusing on the most recent temporal period in the history of *Acanthanura* sp. n. at Tallaganda.

To identify the number of populations (K), and their members, we used Bayesian clustering of genotype data from six nuclear loci, implemented in STRUCTURE version 2.1 (Pritchard *et al.* 2000). All runs employed 'correlated allele frequencies' and 'admixture ancestry'. Estimated $-\ln$ likelihoods were obtained for $K = 1-8$, with five replicates of each K . A burn-in of 10^5 Markov chain Monte Carlo (MCMC) generations and run length of 10^6 MCMC generations were used. Remaining parameters were default. The smallest K that captured the major structure in the data was accepted. Where phylogenetic analysis of mtCOI identified a spatially localized monophyletic clade nested within a genetic population, we separated those samples and erected an additional population to approximate within-group panmixia better. Geographical populations were delineated via individuals with membership coefficient $Q \geq 0.90$. Other individuals with intermediate values (between 0.90 and 0.80) were grouped in the geographical population where they were sampled. Admixed individuals ($Q < 0.80$) from a single sampling site (rotting log) within a genetic contact zone were randomly assigned as a group to a proximate geographical population.

To test the significance of genetic differentiation, partitioning of variation within and among geographical populations was determined via analysis of molecular variance (AMOVA) in ARLEQUIN version 3.0 (Excoffier *et al.* 2005). AMOVA was calculated for all nuclear loci combined

using F_{ST} (Weir & Cockerham 1984), and for mtCOI and each of the three nDNA sequence markers independently, using Φ_{ST} (Excoffier *et al.* 1992). In the latter cases, the best-fit model of molecular evolution was determined in MODELTEST, and ML-corrected distance matrices generated using PAUP*. Significance of F_{ST} and Φ_{ST} was evaluated by permuting 16 000 times.

Demographic changes

To test the predictions of prior expectation 3, coalescent-based (and other) analyses were used to assess evidence for recent postglacial demographic expansion. Maximum-likelihood estimates of Θ ($N_e\mu$ for mtDNA or $4N_e\mu$ for diploid autosomal genes, where μ is the per-site per-generation mutation rate) and g (the exponential growth parameter) were generated for mtCOI, UceEF-1 α , Uce180 and Uce3 separately in each geographical population using FLUCTUATE version 1.4 (Kuhner *et al.* 1998). All runs employed the following search strategy: 10 short Monte Carlo chains (mtCOI, 4000 steps; nDNA, 20 000) and five long chains (mtCOI, 40 000 steps; nDNA, 200 000) sampling every 20th step, random starting trees, transition/transversion (ti/tv) ratio 2.0, empirical base frequencies, and starting Θ -value from Watterson's (1975) estimate. The parameter Θ was estimated jointly with g (initial value $g = 0.1$), and with g held at zero (constant population size). Runs were repeated five times, and then mean and standard deviation (SD) of ML parameter estimates were calculated. We interpreted g to indicate population growth only when $g - 3 \times SD(g) > 0$ [population decline, $g + 3 \times SD(g) < 0$] (Lessa *et al.* 2003).

Demographic change was also assessed using Ramos-Onsins & Rozas's (2002) R_2 and Fu's (1997) F_S , calculated in DNASP version 4.10.3 (Rozas *et al.* 2003) and ARLEQUIN, respectively. Significance was assessed by comparing the observed value with a null distribution simulated via the coalescent process under constant population size (10 000 replicates). We considered there to be strong evidence of population growth/decline only when the inference was corroborated by two or more genes.

Long- vs. short-distance movements, and distinguishing incomplete lineage sorting from gene flow

To gain new insights into the population biology and ecology of a previously unstudied taxon, we assessed the spatial scale over which nuclear gene flow occurs. Because inferences of contemporary interbreeding can be confounded by common ancestry, we developed a novel computer simulation-based approach to aid interpretation of empirical data.

We calculated Slatkin & Maddison's (1989) s , the minimum number of dispersal events needed to reconcile the geographical distribution of alleles with their phylogeny

(assuming complete lineage sorting), for all pairs of geographical populations, in MESQUITE version 1.06 (Maddison & Maddison 2005). MP phylogenies were estimated for UcEF-1 α , Uc180 and Uc3 in PAUP*. Owing to unavailability of outgroup alleles, the likely root of nuclear gene trees was determined following Crandall & Templeton (1993), implemented in TCS. Redundant alleles were represented as a polytomy on the best gene tree(s) and identical alleles found in two or more populations were treated as being different but forming a monophyletic group. To consider short- vs. long-distance movements, we divided population pairs into two categories (i) neighbours; and (ii) non-neighbours, then compared *s* values (locus-by-locus, and averaged across loci) within and between partitions.

Currently, no single method can discriminate incomplete lineage sorting (ILS) from other population-genetic processes (Masta *et al.* 2003). Qualitative approaches are often used to distinguish ILS from dispersal: when a haplotype occupying a terminal (tip) position on a phylogeny is apparently 'misplaced' with respect to geographical locality, dispersal is usually inferred (e.g. Stevens & Hogg 2003). However, the frequency with which ILS produces the same pattern is generally unknown. To address this, we used simulations in MESQUITE to quantify the probability of incompletely sorted alleles occupying tip positions on an intraspecific phylogeny by jointly modelling stochasticity inherent in the coalescent process and nucleotide substitution.

We simulated gene coalescence within a two-population constraint tree, with (haploid) $N_e = 100\,000$, divergence time = 25 000 generations, and 40 alleles sampled per population. We assumed zero post-isolation gene flow/dispersal, panmixia within populations, and constant N_e . Because we cannot know the true degree of ILS in our empirical datasets, we used Slatkin & Maddison's (1989) *s* to select 100 coalescent trees representing a range of scenarios (*s* = 3, 4, 5 and 6; 25 trees each). To maximize comparability with *Acanthanura* sp. n. nuclear sequence loci, DNA characters (300 bp) were simulated over coalescent trees using a HKY85 model of evolution (root states and equilibrium nucleotide frequencies A 0.30, C 0.20, G 0.15, T 0.35, *ti/tv* 2.5) and a scaling factor of 8.0×10^{-8} . This yielded sequence divergences within the range observed from our empirical data. SP cladograms were constructed for simulated datasets using TCS, which estimates the root using predictions derived from coalescent theory (Clement *et al.* 2000). Cladogram ambiguities were resolved following Pfenninger & Posada (2002). We quantified how 'interior' incompletely sorted alleles were by counting the number of independent mutational connections they had, and scaled this relative to all observed alleles in the cladogram. Incompletely sorted alleles were divided into three categories: tips, intermediates, and interiors (Appendix I). We then calculated the partitioning of incompletely sorted alleles observed in the 100 simulated datasets across categories.

Acanthanura sp. n. individuals with at least one nuclear sequence allele atypical of the population where it was sampled (excluding animals collected at contact zones possessing allele characteristic of flanking populations) were identified from population allele frequencies, and the relative measure of how 'interior' those alleles were was calculated (as above). For anomalous tip alleles, we inferred gene flow if simulations showed sharing of tip alleles between populations was unlikely under a scenario of ILS ($P < 0.05$ was tentatively chosen as a threshold). If gene flow was inferred, a qualitative assessment of multi-locus genotype composition was used to distinguish between recent and past interbreeding (i.e. several vs. few anomalous alleles in a single animal).

Results

Phylogenetic analysis of mtCOI and estimating divergence times

Sixty-nine mtCOI haplotypes were identified from 203 individuals (GenBank accessions DQ518680–DQ518748). Sequences displayed open reading frames, and amino acid sequences corresponded with Lunt *et al.*'s (1996) model. Consistent with expectations for invertebrate mtDNA, nucleotide composition was A-T biased (A 0.27, C 0.24, G 0.14, T 0.35, ingroup only). On the best ML tree (Fig. 2a), bootstrap support for tip clades was generally low (< 70%), but deeper partitions in the tree had moderate/strong support (70–100%).

Fine-scale geographical structuring of mtCOI is evident, with six major clades identified on the ML phylogeny (Fig. 2a). Most clades are spatially localized, and their distributions are in strong agreement with *a priori* regions of Tallaganda (Fig. 2b). Well-supported monophyletic clades distinguish Harolds Cross Region (HCR, Clade II) and Anembo Region (AR, Clade V). Pikes Saddle Region (PSR) plus Badja Region (BR) haplotypes form a reasonably well-supported monophyletic clade (Clade I). Three major clades within Eastern Slopes Region (ESR, Clades III, IV and VI) are parapatric and as a group are paraphyletic with respect to other major clades (Fig. 2a–c). Together with Clade V from AR, ESR haplotypes occupy the most basal positions in the tree, consistent with prior expectations. We tentatively consider the phylogenetic placement of two singleton haplotypes found in ESR (H50 and H56, marked '*' in Fig. 2) to indicate past gene flow rather than ILS (rationale given below).

Significant departure from rate constancy ($P < 0.05$) was detected for haplotype H31, and for reasons given above, H50 and H56 were also excluded from this analysis. Assuming Brower's (1994) phylogenetic rate, three mostly nonoverlapping timescales of divergence among major mtCOI clades were evident: (i) Early–Mid Pliocene separation

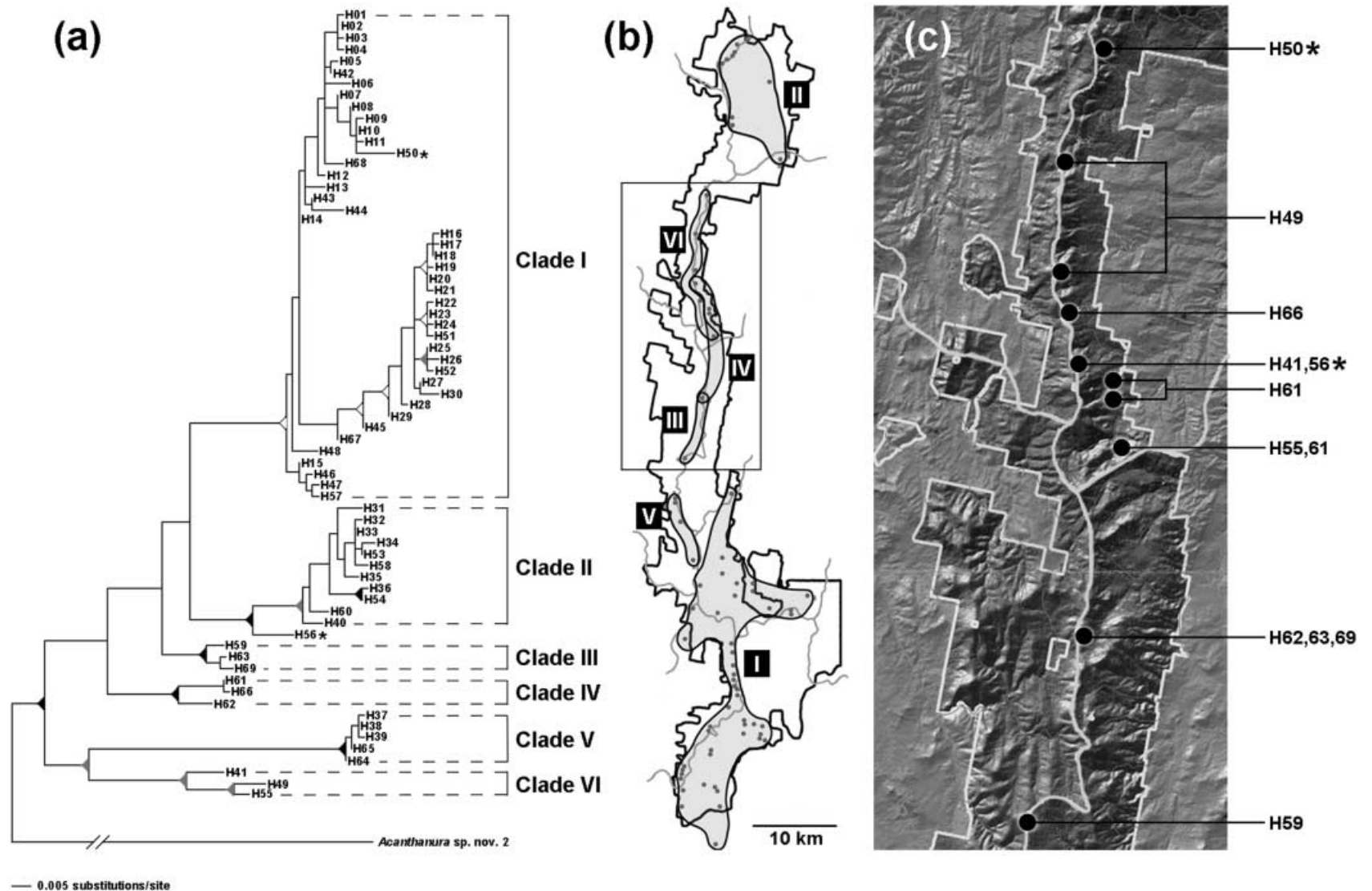


Fig. 2 Phylogenetic analysis of mitochondrial cytochrome oxidase I (mtCOI) haplotypes and geographical distribution of major clades. (a) Maximum-likelihood phylogeny showing relationships among 69 mtCOI haplotypes (H01 to H69) identified from 203 individuals of *Acanthanura* sp. n., estimated using the GTR + I + G substitution model, selected by MODELTEST. This tree had the best score ($-\ln L = 2770.32422$), and was recovered after two iterations of successive approximation. Triangles at nodes indicate bootstrap support (white, 71–80%; grey, 81–90%; black, 91–100%). (b) Map of Tallaganda showing geographical distribution of major clades (shaded and labelled). Grey lines extending outside the perimeter of the study area represent catchment divisions that define *a priori* regions (see Fig. 1). (c) Relief map of Eastern Slopes Region showing geographical distribution of haplotypes. Black circles represent sampling localities, and asterisks identify singleton haplotypes with closest relatives that occur in other *a priori* regions. Relief image modified from State Forests of New South Wales (1995).

of Clade V (AR) 5.0–3.5 Mybp; (ii) Late Pliocene separation of Clade VI (ESR) 3.2–2.5 Mybp; and (iii) Early Pleistocene separation of the four remaining clades 2.6–1.9 Mybp.

Nested clade analysis

Of the 25 nested clades with significant phylogeographical structure, 15 yielded biological inferences (Table 2). ESR mtCOI sequences were highly divergent, failing to satisfy the SP criterion of < 5% chance of homoplasy among haplotypes, thus producing several disconnected networks. Neither the ESR cladograms, nor the AR cladogram, had sufficient genetic/geographical variation for NCA. Accordingly, mtCOI phylogeographical inferences were possible for HCR, PSR and BR only (Clades V and I, Figs 2a, b and 3). A single cladogram was obtained for nuclear EF-1 α , whereas Uc180 and Uc3 produced multiple disconnected cladograms (Fig. 4).

Integrating inferences across loci was unambiguous. NCA detected historical events or population processes on four temporal scales (Phase 1–4, Fig. 5). The oldest inferred

events (Phase 1) were past fragmentation separating PSR and BR, followed by range expansions within each of these two regions. In Phase 2, restricted gene flow within HCR, PSR and BR was detected. Phase 3 was characterized by a second cycle of past fragmentation, marked by the separation of HCR, followed by range expansions within HCR, PSR and BR. At the most recent timescale (Phase 4), restricted gene flow was inferred within BR. Where applicable, most supplementary tests for population growth using FLUCTUATE corroborated NCA-inferred range expansions (Fig. 5). If geographical origins of range expansions are considered indicative of Pleistocene forest refuges (which are often sources of dispersal/recolonization, Heatwole 1987), northwest/central HCR, eastern/central PSR and north/central BR may have retained genetic diversity during recent periglaciation. Consistent with expectations for low-mobility invertebrates, restricted gene flow with isolation by distance was inferred over multiple spatial scales. Conversely, for nuclear genes, several inferences of restricted gene flow were coupled with unexpected rare long-distance dispersal over large spatial scales (Table 2).

Table 2 Biological inferences for nested clades with significant phylogeographical structure, as determined by a χ^2 nested contingency test. *P* value represents the frequency with which 10 000 χ^2 -statistics generated by random permutation were greater than or equal to the observed χ^2 . Abbreviations are: frag., fragmentation; CRE, contiguous range expansion; RE, range expansion; RGF, restricted gene flow; LDD, long-distance dispersal; IBD, isolation-by-distance. Spatial scale indicates the maximum geographical distance between localities where DNA sequences from a given nested clade were sampled.

Locus	Significant nested clades	χ^2 permutation <i>P</i> value	Chain of inference	Inferred process or event	Spatial scale (km)
UcEF-1 α	Clade 1-10	< 0.001	1-2-11-RE-12	CRE	21–50
	Clade 1-11	< 0.001	1-2-11-RE-12-13	Past frag. followed by RE	> 50
	Clade 2-1	< 0.001	1-2-3-5-6-7	RGF with some LDD	> 50
	Clade 2-3	< 0.001	1-2-3-5-6-7	RGF with some LDD	> 50
	Total cladogram	< 0.001	1-2-11-RE-12-13	Past frag. followed by RE	> 50
Uc180	Clade 1-1	0.045	1-2-3-5-6-7	RGF with some LDD	> 50
	Clade 1-3	0.017	1-2-11-17	Inconclusive outcome	5–20
	Clade 2-1	< 0.001	1-2-3-5-6-7	RGF with some LDD	> 50
Uc3	Clade 1-1	< 0.001	1-2-3-5-6-7	RGF with some LDD	> 50
	Clade 2-1	0.004	1-2-3-4	RGF with IBD	> 50
	Total cladogram	< 0.001	No I-T contrast	Inconclusive outcome	> 50
mtCOI	Clade 1-16	< 0.001	1-2-3-5-6-7	RGF with some LDD	< 5
	Clade 1-25	< 0.001	1-2-3-4	RGF with IBD	< 5
	Clade 2-7	0.001	1-19-20	Inadequate sampling	5–20
	Clade 2-12	< 0.001	1-2-11-RE-12-13-14	Inadequate sampling	< 5
	Clade 3-5	0.001	1-2-3-5-6-7-8	Inadequate sampling	5–20
	Clade 3-6	0.039	1-2-3-5-6-N/A	Inconclusive outcome	21–50
	Clade 3-10	< 0.001	1-19-20-2-11-RE-12	CRE	5–20
	Clade 4-1	0.002	1-2-11-RE-12	CRE	5–20
	Clade 4-2	0.006	1-19-20	Inadequate sampling	5–20
	Clade 4-3	< 0.001	1-19-20	Inadequate sampling	21–50
	Clade 4-5	0.001	1-19-20-2-3-4	RGF with IBD	21–50
	Clade 5-1	< 0.001	No I-T contrast	Inconclusive outcome	5–20
	Clade 5-2	< 0.001	1-2-11-RE-12-13	Past Frag. followed by RE	> 50
	Total cladogram	< 0.001	No I-T contrast	Inconclusive outcome	> 50

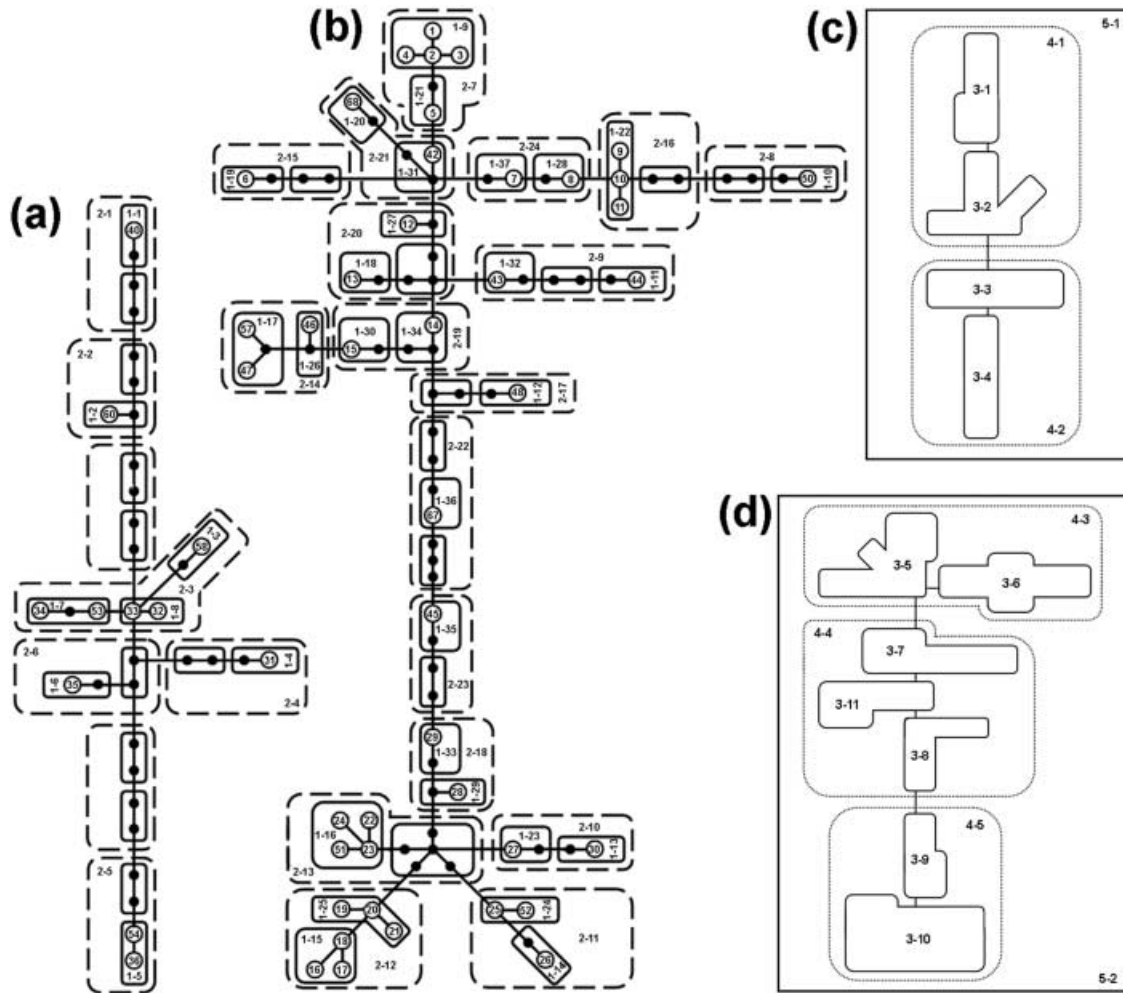


Fig. 3 Statistical parsimony cladograms and nesting design for mitochondrial cytochrome oxidase I haplotypes amenable to nested clade analysis. Zero-, one- and two-step clades: (a) Clade II and (b) Clade I; three-step and higher: (c) Clade II and (d) Clade I. Numbers inside circles identify unique haplotypes (zero-step clades, coded as in Fig. 2). Each single line represents one mutational step between haplotypes, and small black circles dividing single lines represent inferred haplotypes that were not sampled or are extinct. Other clades are labelled using nomenclature where the prefix indicates nesting level, and the suffix indicates clade number (corresponding with Table 2).

Contemporary population structure and partitioning of molecular variation

Bayesian clustering of six-locus nuclear genotypes strongly supported four genetic populations. A monophyletic mtCOI clade (Clade II, Fig. 2) nested within a Northern cluster identified a fifth distinct gene pool. These spatially cohesive, nonoverlapping groups delineated geographical populations (Fig. 6). Population-genetic subdivisions coincide with catchment boundaries, the exception being that the PSR population extends further north than predicted from the model described by Garrick *et al.* (2004) and patterns seen in mtCOI (Fig. 2), possibly indicating male-biased dispersal (and/or among-locus variability).

The majority of individuals (167/205) showed strong membership ($Q \geq 0.90$) in a single population. Of these,

four animals were grouped with a geographical population adjacent to the one where they were sampled, but did not appear to be recent migrants: their multilocus genotypes contained no more than two alleles characteristic of the putative source population (and mtCOI haplotypes were either typical of the local population or unique). In these cases, it appears *STRUCTURE*'s Q -values were disproportionately influenced by very few alleles. Twenty-nine animals had moderate membership ($Q = 0.80$ – 0.89). Eight of nine admixed animals ($Q < 0.80$) were collected at population interfaces or at Tallaganda's southern extremity. Because no admixed individuals had multilocus genotypes consistent with first-generation (F_1) hybrids, we tentatively interpret admixture at genetic contact zones to indicate reproductive compatibility between those neighbouring populations (i.e. F_1 offspring are fertile), although this

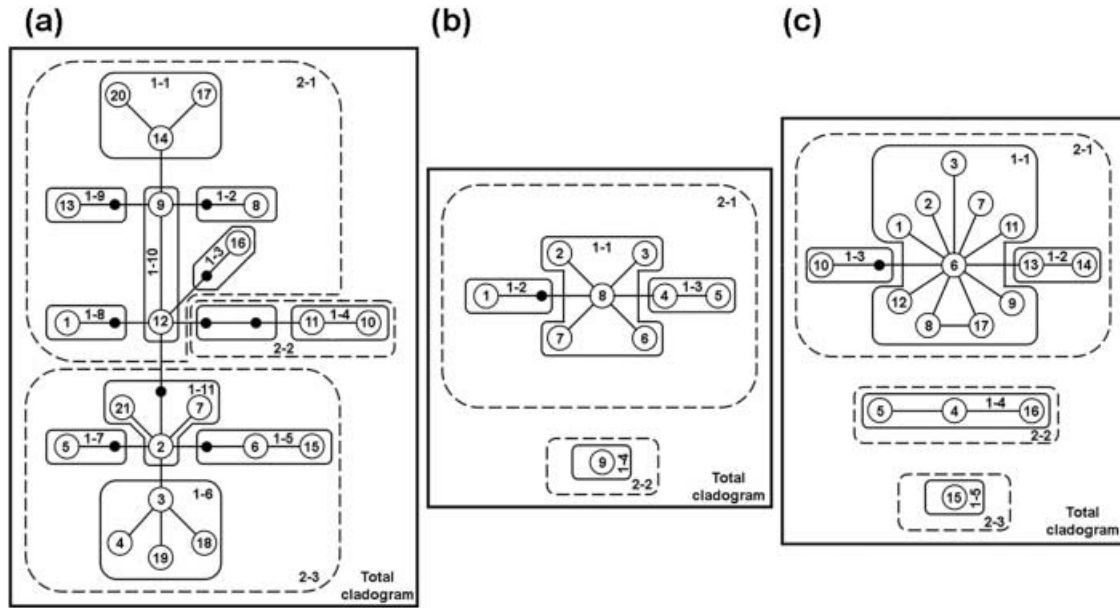


Fig. 4 Statistical parsimony cladograms and nesting design for three nuclear loci. (a) UcEF-1 α (b) Uc180 (c) Uc3. In satisfying the 95% confidence criterion, multiple disconnected cladograms were produced for Uc180 and Uc3. Numbers inside circles identify unique alleles (coded as in Table 6, see also Garrick & Sunnucks 2006). Each single line represents one mutational step between alleles, and small black circles dividing single lines represent inferred alleles. Nested clade nomenclature follows Fig. 3.

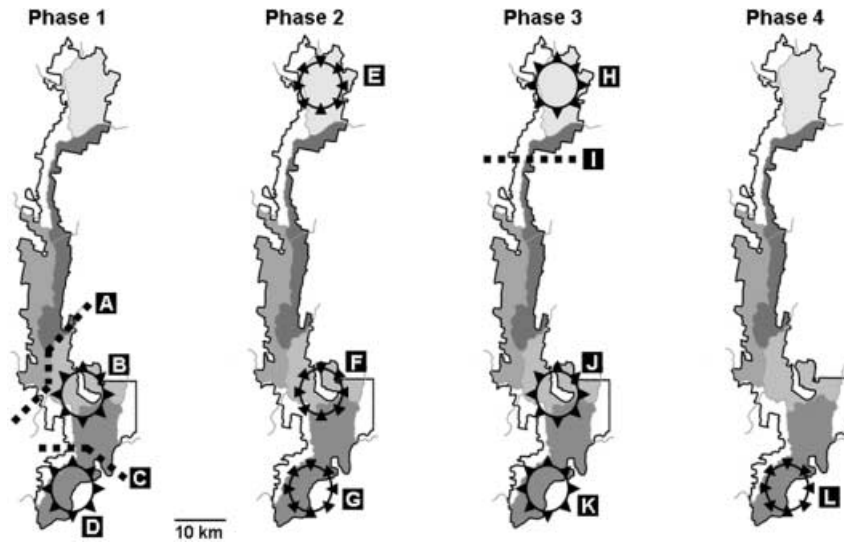


Fig. 5 Nested clade analysis reconstruction of *Acanthanura* sp. n. population history at Tallaganda, cross-validated using four independent loci. Temporally concordant primary phylogeographical inferences are shown from oldest (Phase 1) to most recent (Phase 4). Past fragmentation events are indicated by thick broken lines. Range expansion and restricted gene flow are represented by outward-pointing and inward-pointing small black triangles, respectively, with a solid black circle indicating the *a priori* microbiogeographical region (differentially shaded) to which the inference relates. The gene(s) and associated nested clade that yielded each primary inference are as follows (where inferred range expansion events corroborated by tests for exponential demographic growth are marked '*'): (A) UcEF-1 α total cladogram; (B) UcEF-1 α total cladogram, mtCOI clade 5-2*; (C) mtCOI clade 5-2; (D) mtCOI clade 5-2*; (E) UcEF-1 α clade 2-3, Uc180 clade 2-1; (F) UcEF-1 α clade 2-1, Uc180 clade 1-1, Uc3 clade 1-1; (G) UcEF-1 α clade 2-1, Uc180 clade 1-1, Uc3 clade 2-1 and 1-1; (H) UcEF-1 α clade 1-11, mtCOI clade 4-1*; (I) UcEF-1 α clade 1-11; (J) UcEF-1 α clade 1-10; (K) UcEF-1 α clade 1-10, mtCOI clade 3-10*; (L) mtCOI clade 1-16 and 1-25.

Table 3 Analysis of molecular variance showing partitioning of variation based on Φ_{ST} and F_{ST} . Substitution models selected by MODELTEST for DNA sequence loci were: GTR + I + G, mtCOI; HKY85, UcEF-1 α F81, Uc180 and Uc3. All variances were significant ($P < 0.01$).

Locus/Source of variation	Φ_{ST}				F_{ST}
	mtCOI	UcEF-1 α	Uc180	Uc3	Six nuclear loci
Within populations	20.58	39.00	15.35	61.77	46.79
Among populations	79.42	61.00	84.65	38.23	53.21

Table 4 Tests of demographic change within geographical populations. Maximum-likelihood point estimates of Θ were obtained with g held at zero (no growth), and with g allowed to vary (growth), using FLUCTUATE. Mean values are reported, with standard deviation given in parentheses. Significance of g was assessed following Lessa *et al.* (2003), where positive g indicates growth, and negative g indicates decline. Demographic change was also inferred when R_2 or F_5 was significantly different from zero: significantly small (S) R_2 or negative F_5 indicates growth, whereas significantly large (L) R_2 or positive F_5 indicates decline. Following Fu (1997), F_5 was considered significant at the 5% level only if $P < 0.02$. Values in bold are significant, 'ns' = not significant. Population abbreviations follow Fig. 6. N_{GC} is number of gene copies sampled, $N_{H/A}$ is number of unique haplotypes/alleles in the sample; dashed line indicates test not performed owing to insufficient genetic variation.

Locus	Population	N_{GC}	$N_{H/A}$	Θ		g	R_2	Significance	F_5	Significance
				no growth	growth					
mtCOI	HCR	25	11	0.018 (\pm 0.001)	0.021 (\pm 0.001)	79.139 (\pm 5.571)	0.123	ns	0.641	ns
	ESR	19	7	0.044 (\pm 0.001)	0.030 (\pm 0.001)	-35.838 (\pm 1.502)	0.170	ns	12.224	ns
	AR	15	5	0.003 (\pm 0.000)	0.108 (\pm 0.085)	9665.158 (\pm 579.963)	0.105	$P < 0.05^S$	-2.677	$P < 0.02$
	PSR	66	25	0.038 (\pm 0.001)	0.037 (\pm 0.001)	-0.688 (\pm 1.663)	0.085	ns	-0.313	ns
	BR	78	21	0.020 (\pm 0.002)	0.019 (\pm 0.002)	48.057 (\pm 17.772)	0.076	ns	-2.655	ns
UcEF-1 α	HCR	50	6	0.005 (\pm 0.000)	0.008 (\pm 0.001)	676.036 (\pm 89.040)	0.101	ns	0.321	ns
	ESR	38	4	0.005 (\pm 0.000)	0.005 (\pm 0.000)	-77.835 (\pm 18.184)	0.118	ns	-1.811	ns
	AR	30	1	—	—	—	—	—	—	—
	PSR	134	10	0.011 (\pm 0.001)	0.013 (\pm 0.001)	37.257 (\pm 32.001)	0.086	ns	0.701	ns
	BR	156	10	0.011 (\pm 0.001)	0.023 (\pm 0.001)	575.153 (\pm 67.898)	0.033	$P < 0.05^S$	-3.990	ns
Uc180	HCR	46	4	0.009 (\pm 0.001)	0.010 (\pm 0.000)	-8.030 (\pm 5.977)	0.133	ns	2.817	ns
	ESR	38	2	—	—	—	0.160	ns	-0.657	ns
	AR	30	1	—	—	—	—	—	—	—
	PSR	134	3	0.003 (\pm 0.001)	0.015 (\pm 0.006)	2072.120 (\pm 1498.680)	0.061	ns	-4.475	$P < 0.01$
	BR	156	2	—	—	—	0.080	ns	-2.595	$P < 0.02$
Uc3	HCR	50	3	0.004 (\pm 0.001)	0.006 (\pm 0.002)	975.000 (\pm 467.621)	0.103	ns	-0.385	ns
	ESR	38	1	—	—	—	—	—	—	—
	AR	30	9	0.027 (\pm 0.002)	0.020 (\pm 0.001)	-26.224 (\pm 2.386)	0.217	$P < 0.05^L$	3.088	ns
	PSR	134	8	0.011 (\pm 0.001)	0.043 (\pm 0.003)	2206.794 (\pm 413.150)	0.036	ns	-4.321	ns
	BR	156	7	0.008 (\pm 0.001)	0.019 (\pm 0.008)	620.758 (\pm 415.337)	0.067	ns	-2.665	ns

might not be contemporary. The three population interfaces identified by STRUCTURE are qualitatively different (Fig. 6): although sparsely sampled, ESR and PSR are allopatric yet apparently interbreeding; AR and PSR, also allopatric, appear reproductively isolated; and the well-sampled PSR/BR interface has a narrow zone of parapatry (overlapping distributions of population range limits), and several admixed individuals indicate recurrent gene flow.

The majority of molecular variation was partitioned among geographical populations (Table 3). Global Φ_{ST} values differed across the four DNA sequence markers, with Uc3 showing a reversal from the trend of more variance among populations (probably caused by unusually high allelic diversity in AR). When averaged across nuclear

loci, F_{ST} partitioned only slightly more variation among (cf. within) populations (Table 3). Omission of Uc3 gave a similar result (not shown). Nonetheless, all Φ_{ST} and F_{ST} values were significant ($P < 0.01$) and clearly very large.

Demographic changes

Multi-locus evidence for population growth was detected in HCR, PSR and BR populations, whereas decline was inferred for ESR (Table 4). Interpretation of demographic change in AR was ambiguous owing to conflicting signal (mtCOI, growth; Uc3, decline). Although Uc3 may show some evidence for balancing or diversifying selection in AR (Garrick & Sunnucks 2006), growth was not cross-validated.

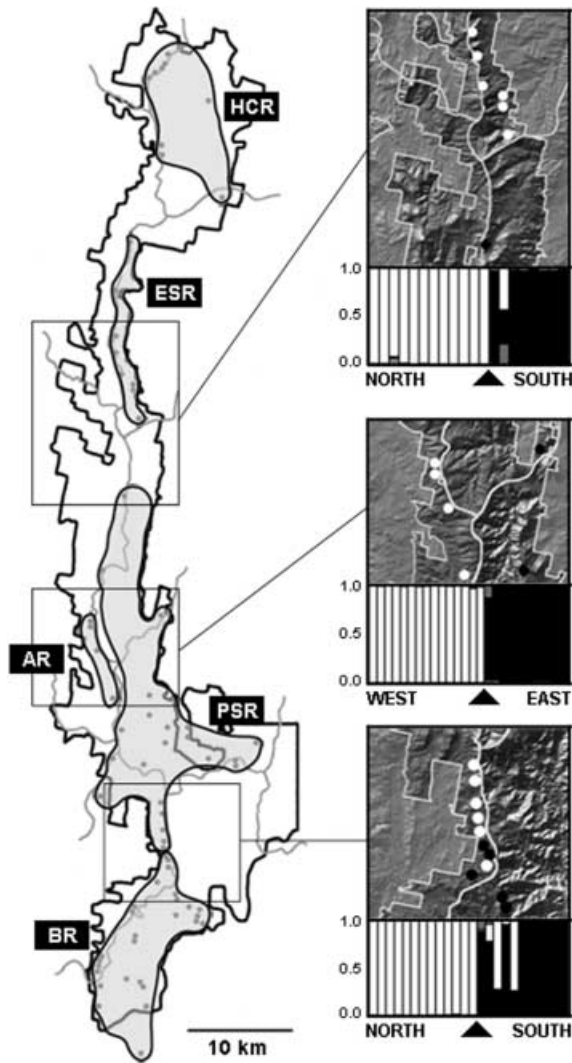


Fig. 6 Distribution of five *Acanthanura* sp. n. geographical populations (shaded and labelled) identified *a posteriori* using STRUCTURE, supplemented by phylogenetic analysis of mitochondrial cytochrome oxidase I. Population abbreviations are: HCR, Harolds Cross Region; ESR, Eastern Slopes Region; AR, Anembo Region; PSR, Pikes Saddle Region; BR, Badja Region. Grey lines extending outside the perimeter of the study area represent catchment divisions that define *a priori* regions (see Fig. 1). Relief maps show transition zones: ESR/PSR (top), AR/PSR (middle) and PSR/BR (bottom). Sampling localities are represented by dots coloured black or white, indicating which of the two populations under consideration a given rotting log was assigned (based on membership of resident individuals). Plots show estimated membership coefficient (Q -value, y -axis) for each individual (columns, x -axis) collected from logs depicted on the relief map, arranged sequentially from north to south or west to east. Columns are horizontally divided into three shaded segments, where area is proportional to Q in each geographical population. Grey shading indicates membership in populations other than the two under consideration. Black triangles indicate association between geographical populations (logs) and genetic populations (individuals): the point marked on the x -axis indicates the transition between neighbouring geographical populations. Relief images modified from State Forests of New South Wales (1995).

Table 5 Slatkin & Maddison's (1989) s , the minimum number of dispersal events required to explain the geographical distribution of alleles with respect to their phylogeny. Values are given for each of three nuclear sequence loci, and averaged across loci, for all geographical population pairs (abbreviations follow Fig. 6). Where multiple most parsimonious allele phylogenies were obtained, s was calculated on each tree (in one case, yielding two s -values for a single comparison). SDD, short-distance dispersal events; and LDD, long-distance dispersal events implied. †, s -values excluded when calculating mean s owing to suspected non-neutrality of the Uc3 locus in the AR population.

Locus/ population pair		UcEF-1 α	Uc180	Uc3	Mean s
SDD	HCR v ESR	4	2	1	2.33
	ESR v PSR	4	2	1	2.33
	AR v PSR	1	1	5†	1.00
	PSR v BR	8	3	5	5.33
LDD	HCR v PSR	4	2	3	3.00
	HCR v AR	1	1	3†	1.00
	ESR v AR	1	1	1†	1.00
	ESR v BR	3	2	1	2.00
	AR v BR	1	1	4†, 5†	1.00
HCR v BR	3	2	3	2.66	

Long- vs. short-distance movements, and distinguishing incomplete lineage sorting from gene flow

There was good agreement between UcEF-1 α and Uc180 on the relative proportion of inferred dispersal events between all population pairs (Table 5). Uc3 yielded high s -values wherever AR contributed to the estimate (excluding ESR-AR). Because anomalies seen in other analyses also relate to Uc3 in AR (above; Garrick & Sunnucks 2006), we excluded these values when averaging s across loci. The PSR/BR interface is clearly the most permeable, whereas exchange of migrants and genes with AR appears highly limited. UcEF-1 α and Uc180 showed no evidence for any dispersal to or from AR beyond initial colonization ($s = 1$, Table 5). After accounting for the latter, there is no strong trend of higher mean s -values between neighbours vs. non-neighbours, suggesting that ILS may have affected results.

When incompletely sorted alleles observed in the 100 simulated datasets were considered together, the majority (89.3%) were interiors, whereas incompletely sorted tip alleles were few (6.3%). Although close, the latter result did not meet our *a priori* threshold. Nonetheless, because two of the four 25-tree datasets simulated under different degrees of ILS did satisfy the 5% criterion (Appendix II), we consider 6.3% to be sufficiently rare to warrant interpretation of empirical observations of tip alleles shared among populations as being consistent with gene

Table 6 Frequencies of alleles at six nuclear loci in five *Acanthanura* sp. n. geographical populations (abbreviations follow Fig. 6). Low frequency alleles (bold) that occur at high frequency in other populations (shaded) guided identification of individuals used for analysis of incomplete lineage sorting (DNA sequence loci only). *N* is sample size (diploid individuals) per locus per population.

	Allele	Population					Locus	Allele	Population					
		HCR	ESR	AR	PSR	BR			HCR	ESR	AR	PSR	BR	
UcEF-1 α	<i>N</i>	25	19	15	67	78	Uc3	<i>N</i>	25	19	15	67	78	
	A01	—	0.026	—	0.545	0.006		A01	—	—	—	0.015	—	
	A02	0.240	—	—	—	—		A02	0.180	—	0.033	0.015	—	
	A03	0.500	0.921	—	0.007	0.006		A03	—	—	—	0.007	—	
	A04	0.020	0.026	—	—	—		A04	—	—	0.233	—	—	
	A05	—	—	1.000	—	—		A05	—	—	0.033	—	—	
	A06	0.140	—	—	—	—		A06	0.800	1.000	0.267	0.776	0.692	
	A07	—	—	—	0.134	0.006		A07	—	—	—	0.022	—	
	A08	—	—	—	0.015	—		A08	—	—	0.100	0.030	—	
	A09	—	—	—	0.194	0.346		A09	0.020	—	—	0.127	0.006	
	A10	—	—	—	0.007	—		A10	—	—	—	0.007	0.006	
	A11	—	—	—	0.007	—		A11	—	—	—	—	0.032	
	A12	—	—	—	0.015	0.532		A12	—	—	—	—	0.173	
	A13	—	—	—	0.067	0.013		A13	—	—	0.067	—	0.071	
	A14	—	—	—	—	0.058		A14	—	—	—	—	0.019	
	A15	0.020	—	—	—	—		A15	—	—	0.133	—	—	
	A16	—	—	—	—	0.013		A16	—	—	0.067	—	—	
	A17	—	—	—	—	0.013		A17	—	—	0.067	—	—	
	A18	—	—	—	0.007	—								
	A19	—	0.026	—	—	—		UcANT	<i>N</i>	25	19	15	67	78
	A20	—	—	—	—	0.006			A01	1.00	1.00	—	0.500	0.026
A21	0.080	—	—	—	—	A02	—		—	—	0.500	0.974		
						A03	—	—	1.000	—	—			
Uc180	<i>N</i>	23	19	15	67	78	UcWnt	<i>N</i>	25	19	15	67	78	
	A01	—	—	1.000	—	—		A01	1.000	1.000	1.000	0.239	0.250	
	A02	—	—	—	0.007	—		A02	—	—	—	0.164	0.071	
	A03	—	—	—	—	0.006		A03	—	—	—	0.597	0.679	
	A04	0.478	—	—	—	—								
	A05	0.174	—	—	—	—								
	A06	—	0.974	—	0.007	—		Uc44	<i>N</i>	25	19	15	67	78
	A07	—	0.026	—	—	—			A01	0.700	0.316	0.033	—	0.006
	A08	0.239	—	—	0.985	0.994			A02	—	0.026	0.567	—	0.006
A09	0.109	—	—	—	—	A03	0.080		0.368	—	—	—		
						A04	—		—	—	0.007	—		
						A05	0.220		0.289	0.400	0.993	0.974		
						A06	—	—	—	—	0.013			

flow. None of the 10 individuals identified as possessing at least one nuclear sequence allele atypical of the local population (see Table 6) was recognized from multiple loci. In six individuals anomalous alleles were tips, indicating past gene flow rather than ILS. In three of these cases, long-distance gene flow was implicated (i.e. putative source population was a non-neighbour). Three animals shared interior alleles with a neighbouring population, and one individual from the PSR–BR interface had an interior allele characteristic of ESR. The latter four cases are parsimoniously explained by ILS. Our simulations also aided interpretation of unexpected phylogenetic placement of two mtCOI haplotypes collected from ESR (Fig. 2): H50 – a tip, because a large number of nodes separate it from the root of the tree – is probably the result of gene flow between PSR and ESR. Although H56 is separated from the root by

few nodes, clades basal to it (Clades III and IV) are found exclusively in PSR (Fig. 2a,b), suggesting past gene flow between HCR and ESR.

Discussion

Our study confirmed that genealogical information from multiple independent loci enables reconstruction of population history across a broad temporal spectrum, and permits cross-validation of phylogeographical inferences (Hare 2001; Templeton 2002). We detected impacts of historical climatic cycles on multiple timescales, and identified at least three putative moist forest refuges at Tallaganda. Our data revealed marked genetic structuring on very fine spatial scales that was consistent with microgeographical patterns predicted *a priori* on the basis of catchment divisions.

Repeated impacts of Pleistocene climatic cycles

The Pleistocene is considered the most influential period in shaping contemporary biogeographical patterns across southeast Australia (Keast 1981). Independent lines of evidence indicate repeated impacts of Pleistocene or earlier climatic cycles on genetic structure of *Acanthanura* sp. n. at Tallaganda. Phylogenetic analysis of mtCOI identified six major clades, each geographically localized within *a priori* microbiogeographical regions (Fig. 2). This pattern of highly divergent allopatric monophyletic clades restricted to HCR, AR and PSR + BR is indicative of past fragmentation, and typically results from isolation due to environmental or geographical barriers (Avice 2000). In ESR, multiple episodes of within-drainage fragmentation seem likely. Our data suggest the deepest phylogeographical breaks marked by mtCOI may have arisen over three different timescales, possibly dating back to the Early–Mid Pliocene (5.0–3.5 Mybp). In the saproxylic flatworm *Artioposthia lucasi* from Tallaganda, a Pliocene coalescence of mtCOI haplotypes was proposed (Sunnucks *et al.* in press). Comparable ages of lineage-splitting have been proposed for a codistributed saproxylic Collembolon, which also showed evidence of repeated climatic impacts on genetic structuring (Garrick *et al.* 2004).

Nested clade analysis of four unlinked loci detected historical events or recurrent population processes on four timescales (Fig. 5). The biological inferences are consistent with repeated impacts of climatic cycles. Alternation between fragmentation and range expansion most likely correspond with periods of climate-induced contraction and isolation of moist forest habitats during cool, dry periglacials, followed by expansion out of refugia during interglacials when forest connectivity was maximal (Heatwole 1987). Genetic isolation among drainages was perhaps most pronounced during persistence in refugia, such that temporal concordance of inferred restricted gene flow seen within *a priori* microbiogeographical regions is expected, at least with sufficient molecular data. Furthermore, phases of restricted gene flow were chronologically nested between phases of past fragmentation/range expansion (Fig. 5), possibly mirroring cycles of forest contraction and expansion. Accordingly, our data suggest that Tallaganda was not completely deforested during the ~19 glacial-interglacial cycles of the Pleistocene (Frakes *et al.* 1987).

Locations of montane forest refugia

Refugia are centres for generation and retention of biodiversity in fluctuating environments (Heatwole 1987). In the present study, *Acanthanura* sp. n. from ESR had the highest mtCOI phylogenetic diversity and the most basal haplotypes (Fig. 2) — a pattern expected for long-term stable refugia.

However, nuclear genetic diversity was relatively low in ESR, yet high in PSR and BR (Table 6). Garrick *et al.* (2004) reported a similar asymmetry between centres of mitochondrial and nuclear diversity in a codistributed Collembolon, suggesting PSR acts as a sink owing to geographical connectedness with other regions. But in the present case, co-occurrence of high mtCOI and nuclear diversity in PSR and BR indicates at least one southern refuge. NCA-inferred geographical origins of putative postglacial range expansions within HCR, PSR and BR (Fig. 5) provide corroborating evidence for refuges to the north and south of ESR. We suggest that areas of HCR, PSR and BR that experienced extensive forest loss during the Early Kosciuszko glaciation and last glacial maximum were probably recolonized from local sources, and ESR refuges (see below) are likely to be of considerable antiquity.

The phylogenetic distinctiveness of *Acanthanura* sp. n. from AR (Fig. 2) requires explanation. Because Tallaganda is isolated from most other montane forests, AR lineages are unlikely to have evolved *ex situ*, and then (re)colonized the region. AR is a region of topographical complexity, and although predicted to have been treeless steppe during cool dry periods (Garrick *et al.* 2004), a small local refuge seems plausible. If strong preferences for large-diameter rotting logs seen in the syntopic velvet-worm *Euperipatoides rowelli* (Barclay *et al.* 2000) also apply to the present Collembolon, high-quality habitats may be very limited along the high-elevation ridgeline of the Gourcock Range, impeding east–west gene flow. Small N_e combined with a potentially impermeable barrier to dispersal could explain the marked genetic differentiation of AR samples (Table 6) and apparently highly restricted gene flow across the AR boundary (Fig. 6, Table 5).

Demographic growth in periodically deforested areas

With the exception of AR, our analyses detected demographic growth in geographical populations from catchments likely to have experienced the greatest contraction or loss of moist forests during Pleistocene periglaciation (Table 4). Because very recent range expansion events are not expected to be marked by mutations at nDNA loci (Templeton 2002), population growth in HCR, PSR and BR probably predates the last glacial maximum, perhaps considerably. Demographic decline inferred for ESR may be an artefact of geographical structuring (multiple refuges), as indicated by presence of three divergent, parapatric mtCOI clades (Fig. 2). Subdivision can increase N_e , causing retention of ancestral sequences (Wakeley 2000), thus producing a genealogy consistent with population decline under the assumption of panmixia. Indeed, we expected the signal of exponential growth to be weakest in areas of long-term stable refugia and strongest where large tracts of previously unsuitable habitat became available following climatic amelioration.

Catchments catch phylogeographical patterning

For species-rich faunal assemblages, effective surrogates for biodiversity indication and prediction of spatial-genetic patterns across heterogeneous landscapes are essential when prioritizing areas for conservation (Moritz *et al.* 2001). Here, we have shown that drainage divisions of Tallaganda reflect ancient and contemporary genetic divisions within *Acanthanura* sp. n. with high precision (Figs 2, 5 and 6). Comparative phylogeographical data from a codistributed Collembolon (Garrick *et al.* 2004) and two terrestrial flatworms (Sunnucks *et al.* in press) show qualitatively similar spatial-genetic patterning, and the velvet worm *Euperipatoides rowelli* shows morphological variation and population-genetic subdivision consistent with *a priori* regions (Sunnucks & Wilson 1999; Sunnucks *et al.* 2000b). Together, these studies indicate that the interaction between topography and Pleistocene climatic cycles may have some predictable influences on phylogeographical structure in low-mobility, wet-adapted arthropods in montane temperate forests of southeast Australia.

Ecology of a previously unstudied taxon

The present work represents the first population-genetic investigation of any member of the subfamily Uchidanurinae. Marked population structure, deep molecular divergences, and abrupt phylogeographical breaks over distances on the order of tens of kilometres or less characterize *Acanthanura* sp. n. from Tallaganda. Despite a propensity for short-distance movements, some long-distance dispersal was inferred (Fig. 2; Table 2; Table 5). Of the mechanisms by which terrestrial Collembola have been reported to passively disperse over long distances, only wind (Farrow & Greenslade 1992) and phoresy (Stevens & Hogg 2003) can account for apparent upslope movements. However, gradual range expansion followed by extinction of intermediate populations is a plausible alternative explanation independent of direct long-distance dispersal (Masta *et al.* 2003). Under both scenarios, male-biased dispersal is a possible explanation (although others exist) for long-distance dispersal being evident from nuclear genes, but not mtCOI haplotypes.

Simulations showed that ILS alone cannot explain some locally high-frequency nuclear alleles occurring at very low frequency outside their apparent population of origin, indicating that despite considerable genetic differentiation (Table 3), reproductive compatibility among at least some populations has been maintained. Because no putative migrants or F_1 hybrids were recognized, recent interbreeding is considered rare: multilocus genotype composition of admixed individuals indicates multiple backcrosses with members of the local gene pool since successful reproduction of a migrant. Accordingly, effective migrants per generation

are probably too few to prevent divergence of populations via drift, such that contemporary processes probably reinforce (rather than overwrite) Pleistocene divergences of *Acanthanura* sp. n.

Conclusions

To conserve evolutionary processes and local co-evolutionary associations, habitat preservation and production forestry management needs to take account of microgeographical patterns of endemism in low-mobility taxa (Garrick *et al.* 2006). While potential impacts on sedentary terrestrial invertebrates are presently unquantified, congruence among species' refugia and areas of high endemism at Tallaganda lend support to the idea that production forestry could be tailored to conserve such diversity. Our study demonstrated the utility of low-mobility animals in detecting very fine-scale Pleistocene climatic impacts affecting the current structuring of biodiversity in Australian short-range endemics. We also showed that despite new analytical challenges faced in multilocus phylogeographical studies, assessing among-locus congruence is tractable. The present work represents a first step in this exciting research area.

Acknowledgements

We thank New South Wales National Parks and State Forests for permission to collect saproxylic invertebrates. Penny Greenslade, Bob Mesibov, Mark Blacket, Chris Smith, Jack Sullivan, Meredith Mahoney and Beckie Symula gave advice on aspects of this work. Members of the Sunnucks and Rowell labs helped with sampling, Mike Westerman provided some laboratory equipment, and Chris Simmons assisted with running the cluster. Noel Tait and Dave Briscoe helped initiate the research program, which was supported by grants from the Australian Research Council (DP-0211156) to P.S. and D.M.R., and the Holsworth Wildlife Research Fund to RCG. The cluster was funded by a National Science Foundation grant (DGE-0114387) to DMH. This manuscript was improved by comments from Shannon Hedtke, Meredith Mahoney, Mark Stevens, Graham Wallis and two anonymous reviewers.

References

- Avice JC (2000) *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, Massachusetts.
- Barclay S, Ash JE, Rowell DM (2000) Environmental factors influencing the presence and abundance of a log-dwelling invertebrate, *Euperipatoides rowelli* (Onychophora: Peripatopsidae). *Journal of the Zoological Society of London*, **250**, 425–436.
- Barrows TT, Stone JO, Fifield LK, Cresswell RG (2001) Late Pleistocene glaciation of the Kosciuszko Massif, Snowy Mountains, Australia. *Quaternary Research*, **55**, 179–189.
- Barrows TT, Stone JO, Fifield LK (2004) Exposure ages for Pleistocene periglacial deposits in Australia. *Quaternary Science Reviews*, **23**, 697–708.

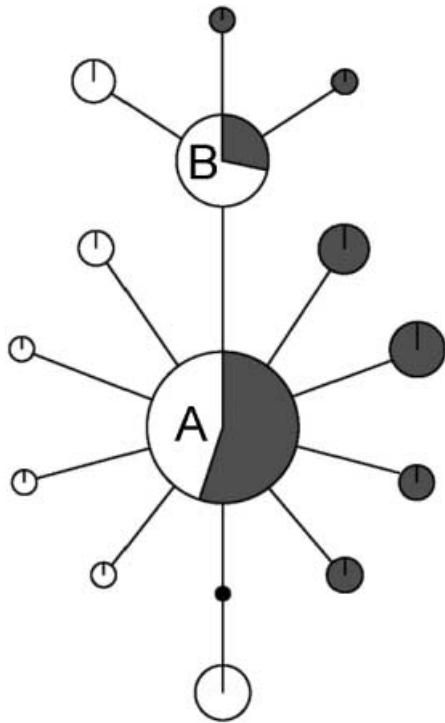
- Bowler JM (1982) Aridity in the late Tertiary and Quaternary of Australia. In: *Evolution of the Flora and Fauna of Arid Australia* (eds Barker WR, Greenslade PJM), pp. 35–45. Peacock Publications, Frewville, South Australia.
- Brower AVZ (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences, USA*, **91**, 6491–6495.
- Clement M, Posada D, Crandall KA (2000) rcs: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.
- Crandall KA (1996) Multiple interspecies transmissions of human and simian T-cell leukemia/lymphoma virus type I sequences. *Molecular Biology and Evolution*, **13**, 115–131.
- Crandall KA, Templeton AR (1993) Empirical tests of some predictions from coalescent theory with applications to intra-specific phylogeny reconstruction. *Genetics*, **134**, 959–969.
- Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.
- Excoffier L, Laval G, Schneider S (2005) ARLEQUIN (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.
- Farrow RA, Greenslade P (1992) A vertical migration of Collembola. *Entomologist*, **111**, 38–45.
- Frakes LA, McGowran B, Bowler JM (1987) Evolution of Australian environments. In: *The Fauna of Australia*, Vol. 1 (eds Dyne GR, Walton DW), pp. 1–16. Australian Government Publishing Service, Canberra.
- Fu Y-X (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.
- Garrick RC, Sunnucks P (2006) Development and application of three-tiered nuclear genetic markers for basal Hexapods using single-stranded conformation polymorphism coupled with targeted DNA sequencing. *BMC Genetics*, **7**, 11.
- Garrick RC, Sands CJ, Rowell DM *et al.* (2004) Phylogeography recapitulates topography: very fine-scale local endemism of a saproxylic 'giant' springtail at Tallaganda in the Great Dividing Range of south-east Australia. *Molecular Ecology*, **13**, 3329–3344.
- Garrick RC, Sands CJ, Sunnucks P (2006) The use and application of phylogeography for invertebrate conservation research and planning. In: *Insect Biodiversity and Dead Wood: Proceedings of a Symposium for the 22nd International Congress of Entomology* (eds Grove SJ, Hanula JL), pp. 15–22. General Technical Report SRS-93. U.S. Department of Agriculture Forest Service, Southern Research Station, Asheville, North Carolina.
- Greenslade P, Moore S, Farrow R (2002) Observations on the feeding behaviour of Uchidanurinae (Collembola: Neanuridae) in Australia. *Victorian Naturalist*, **119**, 221–223.
- Hare MP (2001) Prospects for nuclear gene phylogeography. *Trends in Ecology & Evolution*, **16**, 700–706.
- Harvey MS (2002) Short-range endemism among the Australian fauna: some examples from non-marine environments. *Invertebrate Systematics*, **16**, 555–570.
- Heatwole H (1987) Major components and distributions of the terrestrial fauna. In: *The Fauna of Australia*, Vol. 1 (eds Dyne GR, Walton DW), pp. 69–100. Australian Government Publishing Service, Canberra.
- Hewitt GM (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.
- Hugall A, Moritz C, Moussalli A, Stanisic J (2002) Reconciling paleodistribution models and comparative phylogeography in the Wet Tropics rainforest land snail *Gnarosophia bellendenkerensis* (Brazier 1875). *Proceedings of the National Academy of Sciences, USA*, **99**, 6112–6117.
- Keast A (1981) Distributional patterns, regional biotas, and adaptations in the Australian biota: a synthesis. In: *Ecological Biogeography of Australia*, Vol. 3 (ed. Keast A), pp. 1891–1997. Junk, The Hague, The Netherlands.
- Kershaw AP, D'Costa DM, McEwen Mason JRC, Wagstaff BE (1991) Palynological evidence for Quaternary vegetation and environments of mainland southeastern Australia. *Quaternary Science Reviews*, **10**, 391–404.
- Kuhner MK, Yamato J, Felsenstein J (1998) Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics*, **149**, 429–434.
- Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Briefings in Bioinformatics*, **5**, 150–163.
- Lessa EP, Cook JA, Patton JL (2003) Genetic footprints of demographic expansion in North America, but not Amazonia, during the late Quaternary. *Proceedings of the National Academy of Sciences, USA*, **100**, 10331–10334.
- Lunt DH, Zhang D-X, Szymura JM, Hewitt GM (1996) The insect cytochrome oxidase I gene: evolutionary patterns and conserved primers for phylogenetic studies. *Insect Molecular Biology*, **5**, 153–165.
- Maddison WP, Maddison DR (2005) MESQUITE: A Modular System for Evolutionary Analysis, Version 1.06. <http://mesquiteproject.org>.
- Masta SE, Laurent NM, Routman EJ (2003) Population genetic structure of the toad *Bufo woodhousii*: an empirical assessment of the effects of haplotype extinction on nested cladistic analysis. *Molecular Ecology*, **12**, 1541–1554.
- Moritz C, Patton JL, Schneider CJ, Smith TB (2000) Diversification of rainforest faunas: an integrated molecular approach. *Annual Review of Ecology and Systematics*, **31**, 533–563.
- Moritz C, Richardson KS, Ferrier S *et al.* (2001) Biogeographical concordance and efficiency of taxon indicators for establishing conservation priority in a tropical rainforest biota. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **268**, 1875–1881.
- Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research*, **8**, 4321–4325.
- Nei M, Li W-H (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences, USA*, **76**, 5269–5273.
- Pfenninger M, Posada D (2002) Phylogeographic history of the land snail *Candidula unifasciata* (Helicellinae, Stylommatophora): fragmentation, corridor migration, and secondary contact. *Evolution*, **56**, 1776–1788.
- Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.
- Posada D, Crandall KA, Templeton AR (2000) GEODIS: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology*, **9**, 487–488.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
- Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, **19**, 2092–2100.

- Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003) DNASP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.
- Slatkin M, Maddison WP (1989) A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics*, **123**, 603–613.
- State Forests of New South Wales (1995) *Proposed Forestry Operations in the Queanbeyan and Badja Management Area, Environmental Impact Statement*. State Forests of New South Wales, Southern Region. Pennant Hills, New South Wales.
- Stevens MI, Hogg ID (2003) Long-term isolation and recent range expansion from glacial refugia revealed for the endemic springtail *Gomphiocephalus hodgsoni* from Victoria Land, Antarctica. *Molecular Ecology*, **12**, 2357–2369.
- Stevens MI, Greenslade P, Hogg ID, Sunnucks P (2006) Southern Hemisphere springtails: could any have survived glaciation of Antarctica? *Molecular Biology and Evolution*, **23**, 874–882.
- Sunnucks P (2000) Efficient genetic markers for population biology. *Trends in Ecology & Evolution*, **15**, 199–203.
- Sunnucks P, Wilson ACC (1999) Microsatellite markers for the onychophoran *Euperipatoides rowelli*. *Molecular Ecology*, **8**, 895–906.
- Sunnucks P, Wilson ACC, Beheregaray LB *et al.* (2000a) SSCP is not so difficult: the application and utility of single-stranded conformation polymorphism in evolutionary biology and molecular ecology. *Molecular Ecology*, **9**, 1699–1710.
- Sunnucks P, Curach NC, Young A *et al.* (2000b) Reproductive biology of the onychophoran *Euperipatoides rowelli*. *Journal of the Zoological Society of London*, **250**, 447–460.
- Sunnucks P, Blacket MJ, Taylor JM *et al.* (in press) A tale of two flatties: different responses of two terrestrial flatworms to past environmental climatic fluctuations at Tallaganda in montane south-eastern Australia. *Molecular Ecology*.
- Swofford DL (2002) *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods)*, Version 4. Sinauer Associates, Sunderland, Massachusetts.
- Swofford DL, Olsen GJ, Waddell PJ, Hillis DM (1996) Phylogenetic inference. In: *Molecular Systematics*, 2nd edn (eds Hillis DM, Moritz C, Mable B), pp. 407–514. Sinauer Associates, Sunderland, Massachusetts.
- Tajima F (1993) Simple methods for testing molecular clock hypothesis. *Genetics*, **135**, 599–607.
- Takahata N, Nei M (1985) Gene genealogy and variance of interpopulational nucleotide differences. *Genetics*, **110**, 325–344.
- Templeton AR (1998) Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology*, **7**, 381–397.
- Templeton AR (2002) Out of Africa again and again. *Nature*, **416**, 45–51.
- Templeton AR (2004) Statistical phylogeography: methods for evaluating and minimizing inference errors. *Molecular Ecology*, **13**, 789–809.
- Templeton AR, Sing CF (1993) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. Nested analyses with cladogram uncertainty and recombination. *Genetics*, **134**, 659–669.
- Templeton AR, Boerwinkle E, Sing CF (1987) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. I. Basic theory and an analysis of alcohol dehydrogenase activity in *Drosophila*. *Genetics*, **117**, 343–351.
- Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, **132**, 619–633.
- Templeton AR, Routman E, Phillips CA (1995) Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the Tiger Salamander, *Ambystoma tigrinum*. *Genetics*, **140**, 767–782.
- Wakeley J (2000) The effects of subdivision on the genetic divergence of populations and species. *Evolution*, **54**, 1092–1101.
- Watterson GA (1975) On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, **7**, 256–276.
- Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.
- Yeates DK, Bouchard P, Monteith GB (2002) Patterns and levels of endemism in the Australian Wet Tropics rainforest: evidence from flightless insects. *Invertebrate Systematics*, **16**, 605–619.

This paper represents part of Ryan Garrick's PhD research, and is one component of a larger collaborative project centred at Tallaganda. He is interested in phylogeography and population structure, and will be working on Sonoran Desert biota as a postdoc in Rodney Dyer's lab at Virginia Commonwealth University. Chester Sands is Molecular Phylogeneticist at the British Antarctic Survey, and makes a critical contribution to the velvet-worm component of our research programme. Dave Rowell is Associate Professor in Botany and Zoology at the Australian National University and has worked on ecology, behaviour and evolutionary biology of Tallaganda invertebrates for 20 years. David Hillis is Professor in the Section of Integrative Biology at the University of Texas, with interests in phylogenetic reconstruction and its application for understanding mechanisms driving biotic diversification. Paul Sunnucks is Senior Lecturer in Zoology at Monash University and works on molecular population biology and ecology of diverse organisms to answer evolutionary questions at a range of spatial and temporal scales, and to inform conservation management. His research programs encompass saproxylic invertebrates.

Appendix I

Statistical parsimony cladogram estimated from one of the 100 DNA sequence datasets simulated via the coalescent process under a two-population isolation model with divergence time = $0.25 N_e$ and the HKY85 model of nucleotide substitution. Circles represent unique alleles, and the size of each circle is proportional to overall frequency. Relative frequencies of alleles in each population are indicated by shading (grey vs. white). Each single line represents one mutational step between alleles, and small black circles dividing single lines represent inferred alleles. In this example two incompletely sorted alleles, A and B, were present. Allele A was considered 'interior' because it has the greatest number of mutational connections ($N = 10$), allele B is 'intermediate' ($N = 4$), and all other observed alleles are 'tips' ($N = 1$). The coalescent tree on which these DNA sequence characters were overlaid had an s -value of 5.



Appendix II

Summary of computer simulation results used to guide interpretation of incomplete lineage sorting vs. gene flow from empirical data

Dataset	Proportion of incompletely sorted alleles (%)		
	Interior	Intermediate	Tip
$s = 3$	96.2	0.0	3.8
$s = 4$	83.3	6.7	10.0
$s = 5$	86.2	6.9	6.9
$s = 6$	92.6	3.7	3.7
Combined	89.3	4.5	6.3