Molecular Ecology (2004) **13**, 3329–3344 doi: 10.1111/j.1365-294X.2004.02340.x

Phylogeography recapitulates topography: very fine-scale local endemism of a saproxylic 'giant' springtail at Tallaganda in the Great Dividing Range of south–east Australia

R. C. GARRICK,* C. J. SANDS,* D. M. ROWELL,† N. N. TAIT,‡ P. GREENSLADE† and P. SUNNUCKS*
Department of Genetics and Evolution, Biological Sciences Building 1, La Trobe University, Plenty Road, Bundoora, VIC 3806, Australia; †School of Botany and Zoology, Australian National University, Canberra, ACT 0200, Australia; ‡Department of Biological Sciences, Macquarie University, Sydney, NSW 2109, Australia.

Abstract

Comparative phylogeography can reveal processes and historical events that shape the biodiversity of species and communities. As part of a comparative research program, the phylogeography of a new, endemic Australian genus and species of log-dependent (saproxylic) collembola was investigated using mitochondrial sequences, allozymes and anonymous single-copy nuclear markers. We found the genetic structure of the species corresponds with five a priori microbiogeographical regions, with population subdivision at various depths owing to palaeoclimatic influences. Closely related mtDNA haplotypes are co-distributed within a single region or occur in adjacent regions, nuclear allele frequencies are more similar among more proximate populations, and interpopulation migration is rare. Based on mtDNA divergence, a late Miocene–late Pliocene coalescence is likely. The present-day distribution of genetic diversity seems to have been impacted by three major climatic events: Pliocene cooling and drying (2.5–7 million years before present, Mybp), early Pleistocene wet-dry oscillations (*c.* 1.2 Mybp) and the more recent glacial-interglacial cycles that have characterized the latter part of the Quaternary (< 0.4 Mybp).

Keywords: Collembola, comparative phylogeography, conservation, Great Dividing Range, saproxylic, Tallaganda

Received 10 April 2004; revision received 3 August 2004; accepted 3 August 2004

Introduction

A primary objective of conservation biology is to ensure the maintenance of evolutionary processes, underpinned by genetic diversity (Frankham *et al*. 2002). Through revealing cryptic, deeply divergent evolutionary lineages that are otherwise overlooked by traditional taxonomy, and by elucidating processes of biotic diversification, phylogeography can direct conservation priorities (Arbogast & Kenagy 2001).

In contrast to vertebrates, invertebrates often exhibit extremely large abundances and sedentary taxa may show population structure over very fine spatial scales (Peterson & Denno 1998). Thus, the predominant focus on vertebrates and vascular plants in conservation research and planning (Stiling 1999) is likely to result in management strategies that fail to cater for a large proportion of biodiversity. For example, Australian vertebrates and vascular plants in the wet tropics were found to be inefficient surrogates for invertebrates, while invertebrates were excellent surrogates for other biota (Moritz *et al*. 2001). Similar conclusions have been reached for ground-dwelling arthropods in temperate Australian forests (Ferrier *et al*. 1999).

It has been proposed that ecological characteristics of certain fauna will tend to increase their efficiency in capturing geographical patterning in biota that results from long-acting processes such as climatic cycles. Specifically, Moritz *et al*. (2001) proposed that while many vertebrate taxa may show structure determined by Quaternary climatic

Correspondence: Ryan Garrick, Fax: + 61 3 94792480; E-mail: r.garrick@latrobe.edu.au

oscillations, invertebrates that experience the landscape at fine scales might not have had their more ancient patterning so completely overwritten. Similar arguments are well-established for freshwater phylogeography (Beheregaray *et al*. 2002). We propose that saproxylic biota in forests may be an exceptional ecological community in their ability to capture landscape history, since for many of these organisms even individual creek lines and gullies may represent refuges during cold dry periods.

The saproxylic (dead-wood dependent) community

Decomposing logs on the forest floor support a rich invertebrate fauna (Key 1993). Log-dwellers comprise a large proportion of the entire forest invertebrate fauna and play critical roles in nutrient cycling (Grove 2002). Many saproxylic organisms possess characteristics that make them highly dependent on their local environment and which may thus make them effective indicators of past environmental change. First, most species are ecologically specialized and inhabit narrow niches (Yee *et al*. 2001). Second, many saproxylic organisms are dependent on niches created by the activities of other saproxylic species (Warren & Key 1991). Third, many saproxylic organisms have poor dispersal ability (Hammond 1984).

Unfortunately, the characteristics that make saproxylic organisms fit candidates for comparative phylogeography and biodiversity indication may also render them susceptible to human impacts. Indeed, despite their diversity and ecological significance, the saproxylic community in boreal forest systems has a disproportionate fraction of species of conservation concern (Jonsson & Kruys 2001). Although production forestry in Australia has a shorter history, impacts seen in the northern hemisphere are becoming evident, including altered dynamics of coarse woody debris recruitment and the decline of many species (Grove 2002). However, basic questions critical in conservation management remain largely unanswered (Thomas & Morris 1994).

The Tallaganda model

We have been conducting comparative phylogeography and population biology of a suite of saproxylic invertebrates over a section of Tallaganda State Forest (SF) and the adjoining Badja SF (herein collectively referred to as Tallaganda), in south-east New South Wales (NSW) (Scott & Rowell 1991; Sunnucks & Wilson 1999; Barclay *et al*. 2000a,b; Sunnucks *et al*. 2000a,b; Sunnucks & Tait 2001). This site has a range of characteristics that render it unusually informative about processes of biological diversification in temperate Australian eucalypt forests. The following section derives from CSIRO (1969), Bowler (1982), Singh (1982), Frakes *et al*. (1987), Heatwole (1987), Hope (1994), State Forests of NSW (1995) and White (1990, 1994).

Tallaganda is an ecologically and topographically isolated but heterogeneous section of the Great Dividing Range (GDR) *c*. 100 km long by 3–17 km wide. It is part of a three-tier landscape system comprising a geologically recent coastal plain, an escarpment that has been physically stable for at least 20 million years (My), and the ancient GDR ≥ 50–70 My old. The GDR and its biota have been impacted by climate change at a range of timescales, at least during the Oligocene Refrigeration 25–36 million years before present (Mybp), the Mid-Miocene and Terminal Miocene cooling events 7–15 Mybp, extended cooling and drying in the Pliocene 2.5–7 Mybp, and during the 20 cycles of the Pleistocene (last 1.8 My). From the perspective of moisture-dependent fauna, Tallaganda can be viewed as an episodically expanding and contracting collection of sometimes tiny forested refuges, embedded in a periodically peri-glaciated inhospitable mosaic of cold, dry woodland or steppe.

Based on topographical and hydrological characteristics of Tallaganda, and an assessment of the likely palaeoclimatological impacts upon the distribution of moist forest habitats in southeast Australia, we recognize five microbiogeographical regions (see Materials and methods).

Saproxylic invertebrate phylogeography at Tallaganda: a 'giant' collembolon

A difficulty with studying saproxylic species at Tallaganda is that a large proportion are undescribed, necessitating detailed examination of morphological variation before phylogeographical studies can proceed. An unusually large and extremely dorso-ventrally flattened neanurid collembolon from Tallaganda, which has been clearly characterized, represents a new genus and species (P. Greenslade, taxonomic description in preparation), but its biology is largely unknown. Based on morphology, the species appears to be log-adapted. Vagility is presumed to be low given that the animal lacks a spring organ, is soft-bodied (thus extremely susceptible to desiccation), and is rarely found in leaf litter or pitfall traps. This species has a low-density and patchy distribution across Tallaganda, and occupies only relatively large moist logs at an advanced decomposition stage, often in excess of 50 years on the ground (Barclay *et al*. 2000a). This collembolon is believed to be part of a relictual fauna, possibly Gondwanan, persisting in upland, cool, moist refugia.

We used sequence data, allozymes and anonymous single-copy nuclear markers to test the hypothesis that the spatial distribution of genetic diversity and evolutionary relationships in the focal species corresponds with five a priori microbiogeographical regions.

Materials and methods

Identification of five microbiogeographical zones within Tallaganda

We base our a priori microbiogeographical zones on information from CSIRO (1969), Frakes *et al*. (1987), Heatwole (1987), Hope (1994), State Forests of NSW (1995) and G. Hope (personal communication). The regions correspond with the following locations reported in Sunnucks & Wilson (1999) and Sunnucks *et al*. (2000a): (1) Forbes Creek/Harolds Cross; (2) North Tallaganda SF; (3) Central Tallaganda SF; (4) South TSF; (5) Badja SF. They are outlined below in latitudinal order from north to south.

Microbiogeographical region 1. The Harolds Cross Region (HCR) is characterized by low altitude broad and undulating hills, low annual rainfall, and dry sclerophyll eucalypt forest. The region lies within the Lower Shoalhaven subcatchment (Fig. 1). During cold, dry periods in the Pleistocene and earlier, HCR would have been dominated by treeless steppe or woodland and thus largely inhospitable to moisture-dependent saproxylic fauna.

Microbiogeographical region 2. The Eastern Slopes Region (ESR) is dominated by high altitude, steep, heavily dissected topography. There are at least 15 east–west oriented creeks, each potentially serving as refugia for moisture-dependent biota during cold, dry periods. Owing to the presence of many sheltered areas and orographic rainfall, this region supports tall wet sclerophyll eucalypt forest. The region lies within the Ballalaba and Jerrabatgulla subcatchments (Fig. 1).

Microbiogeographical region 3. The Anembo Region (AR) on the western side of the GDR at central Tallaganda is characterized by broad, undulating topography. The western slopes remain relatively dry during summer owing to rain shadow effects. Vegetation is dry sclerophyll at lower altitudes, with subalpine woodlands occupying the highest areas. The region lies within the Queanbeyan subcatchment (Fig. 1). In cold dry periods, treeless steppe is expected to have dominated.

Microbiogeographical region 4. The Pikes Saddle Region (PSR) is characterized by steep slopes that support moist eucalypt forest, particularly in gullies and sheltered areas. It is drained by the Upper Shoalhaven subcatchment (Fig. 1). To the east of PSR, the adjoining Deua NP (DNP) is comparatively low and flat. The region is predicted to have been mostly treeless during cold, dry periods.

Microbiogeographical region 5. The Badja Region (BR) is low-lying and relatively flat (except for some steep topography in the south-east), and forms part of the tablelands flanking the GDR. Tall moist eucalypt forests dominate the west, while the east supports dry sclerophyll forest and some small, isolated pockets of old growth vegetation in sheltered gullies. The region is in the Tuross catchment (Fig. 1). During cold, dry periods there was probably an abrupt boundary between treeless tablelands and adjoining high-elevation GDR, and BR is predicted to have been treeless steppe.

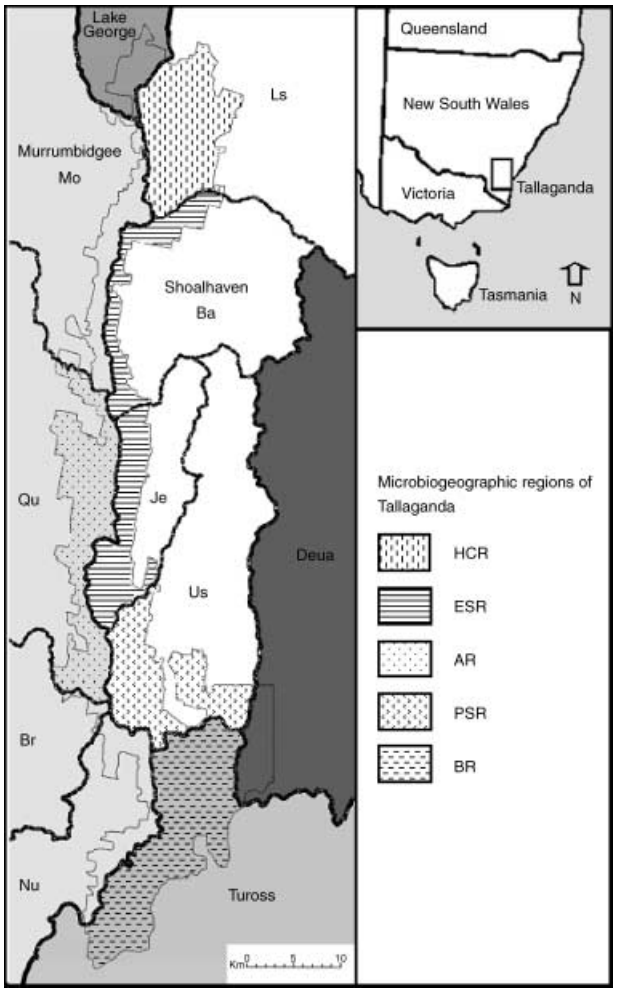

Fig. 1 Hydrological divisions of Tallaganda (modified from State Forests of NSW 1995) and locations of five a priori microbiogeographical regions. Major catchments are differentially shaded and labelled in full. Subcatchments are delineated by thick black lines and labelled as follows: Ls, Lower Shoalhaven; Ba, Ballalaba; Je, Jerrabatgulla; Us, Upper Shoalhaven; Mo, Molonglo; Qu, Queanbeyan; Br, Bredbo; Nu, Numeralla. A thin black line represents the perimeter of the study area, which is bisected by the spine of the Great Dividing Range (situated where the Shoalhaven and Murrumbidgee catchments meet). Inset: map of south-east Australia showing the location of Tallaganda.

Taxon sampling

The new genus and species of Neanuridae occurs at low densities and is cryptic. Collection is labour-intensive requiring logs to be dismantled by hand. Nevertheless, between 1997 and 2002, a total of 129 specimens were collected from 26 logs along an 80 km north–south transect. Specimens were stored at –80 °C or preserved in 100% ethanol. Voucher specimens have been deposited at the South Australian Museum, Adelaide.

Allozyme electrophoresis

Of 129 specimens, 72 were frozen and thus suitable for allozyme analysis. Individuals were homogenized in 30 μL grinding buffer (pH 6.7) (Peakall & Beattie 1991), microfuged for 3 min at 13 000 *g*, and the residual pellet retained for DNA extraction. Cellulose acetate electrophoresis of the supernatant followed Hebert & Beaton (1993). Of the six enzyme systems assayed, two, glucose phosphate isomerase (GPI, EC 5.3.1.9) and lactate dehydrogenase (LDH, EC 1.1.1.27), showed adequate resolution and variation to be scored. Activity was variable, and not all individuals were successfully scored for both loci.

DNA isolation and amplification

Genomic DNA was extracted using a simplified 2× cetyltrimethylammonium bromide (CTAB) protocol, modified after Murray & Thompson (1980), from whole specimens or tissue pellets (above). Approximately 50 ng of DNA was used as template in polymerase chain reaction (PCR) amplifications of nuclear loci, and 10 ng of DNA for amplifications of the mitochondrial cytochrome oxidase subunit I (COI) gene. The latter was necessary to prevent coamplification of putative nuclear mitochondrial pseudogenes.

Anonymous single-copy nuclear (scn) DNA marker development and genotyping

Codominant anonymous scnDNA genetic markers were developed using a modified 'Sequencing With Arbitrary Primer Pairs' protocol (Burt *et al*. 1994). Briefly, anonymous segments of the nuclear genome were amplified under low-stringency PCR using arbitrary primer pairs selected from a bank of random amplified polymorphic DNA (RAPD), microsatellite, and exon-priming intron crossing (EPIC) primers (Jarman *et al*. 2002). Sequence- and/or size-variable DNA fragments that had a different primer on each end were identified via single-stranded conformation polymorphism (SSCP) and/or standard electrophoresis, and sequenced from gel slices. Primers were then designed internal to the sequenced fragment (Table 1).

Table 1 Primers used for amplification of anonymous scnDNA loci and mtDNA cytochrome oxidase subunit I (COI) from the new species of Collembola (suffixes -F and -R refer to forward and reverse direction, respectively)

Primer	Primer sequence (5′ to 3′)	Source
Sm2-F[a]	GAAACGGGTGCTGGTTSRAGG	This study
Sm2-R	GGGTAACGCCRTTGGAAACAG	This study
Sm4-F	GAATTGGTGGGAGATCTCTC	This study
Sm4-R	TGTCGTCCGTCTATGATTCG	This study
Sm6-F	CTGAATGCCGTCGAAACGTAAAC	This study
Sm6-R	GTTGGTTTACCTGTTTTAAATG	This study
Sm8-F[a]	AGTGGGATTTTAGGATGGCAGG	This study
Sm8-R	CCAAGACTAAGATTGAGAAGAAGTC	This study
C1-J-1718[a]	GGAGGATTTGGAAATTGATTAGTTCC	Simon *et al.* (1994)
C1-N-2329	ACTGTAAATATATGATGAGCTCA	Simon *et al.* (1994)

[a]Primer used for sequencing.

All nuclear loci were amplified using [α^{33}P]-dATP incorporation PCR (described below) using an FTS-960 (Corbett Research, Mortlake, NSW, Australia) and the following protocol: 2 min at 94 °C (1 cycle), 30 s at 94 °C, 30 s at 48–60 °C, 1 min at 72 °C (35 cycles) and final extension for 2 min at 72 °C. Specimens were genotyped for size-variable alleles at loci Sm4 and Sm6 via electrophoretic separation of PCR products on a denaturing acrylamide gel visualized via autoradiography. Individuals were genotyped for sequence-variable alleles at loci Sm2 and Sm8 via SSCP (Sunnucks *et al*. 2000b). Samples with the same gel phenotype were run again side-by-side. Multiple representatives of each putative allele were sequenced (LI-COR automated DNA sequencer Lincoln, NE, USA, following the manufacturer's recommendations). Where different genotypes produced similar SSCP phenotypes, relevant samples were scored using a restriction enzyme assay (PCR-RFLP).

Single-stranded conformation polymorphism and sequencing of mtDNA

A 611 bp fragment of the COI gene was amplified using primers C1-J-1718 and C1-N-2329 (Simon *et al*. 1994) (Table 1). [α^{33}P]-dATP amplifications were performed in 10 μL reactions: 16 μM ammonium sulphate, 68 mM Tris-HCl (pH 8), 10 mM β-mercaptoethanol, 5% bovine serum albumin 10 mg/mL (Progen, Darra, QLD, Australia), 2 mM magnesium chloride, 200 μM each dNTP, 0.05 μL [α^{33}P]-dATP (10 mCi/mL), 0.5 μM each primer, 0.5 units of *Taq* DNA Polymerase (Promega, Madison, WI, USA), and 1 μL template DNA (above). COI was amplified using an FTS-960 and the following protocol: 2 min at 94 °C (1 cycle), 20 s at 94 °C, 30 s at 50 °C, 45 s at 72 °C (35 cycles) and final extension for 2 min at 72 °C. Screening for individual sequence variation was by SSCP (Sunnucks *et al*. 2000b). Multiple representatives of putative COI haplotypes were sequenced (as above).

Sequence alignment

Sequences were edited with reference to digital gel images and aligned using ALIGNIR version 2.0 (LI-COR). Areas of ambiguous alignment were omitted from the dataset. Both nucleotide and amino acid mtDNA sequences aligned well with corresponding sequences of a morphologically similar unidentified Neanuridae from Mount Donna Buang, Victoria. Amino acid sequences were consistent with Lunt *et al*.'s (1996) model of insect COI. No nuclear DNA sequences contained putative open reading frames and were presumed to be noncoding.

Identification of population structure and assignment of individuals

A Bayesian clustering method using multilocus genotype data, implemented in STRUCTURE version 2.1 (Pritchard *et al*. 2000) was used to demonstrate the presence of population structure, estimate the number of populations (K), and assign individuals probabilistically to 'genetic' populations. This information delineated 'geographic' populations (i.e. groups of logs) used in all subsequent analyses. The method assumes that, within genetic populations, loci are in Hardy–Weinberg and linkage equilibria, and individuals are assigned to achieve this.

The Collembola dataset consisted of six nuclear loci: two allozymes (GPI and LDH) and four anonymous scnDNA markers (Sm2, Sm8, Sm4 and Sm6). The proportions of missing data were 59.7%, 49.6%, 0.8%, 20.9%, 2.3% and 0.8%, respectively. Separate runs were conducted using the independent allele frequency model and the correlated allele frequency model. All runs employed the admixture ancestry model. Estimated log likelihoods were obtained for $K = 1$ to $K = 8$, with five replicates of each K, for each of the two allele frequency models. The smallest value of K that captured the major structure in the data was considered 'correct'. In all cases a 'burn-in' phase of 10^5 Markov chain Monte Carlo (MCMC) generations and a run length of 10^6 MCMC generations was employed. All other parameters were default.

The application of novel genetic markers in the present study prompted us to test for Mendelian inheritance patterns. For each nuclear locus, where sample sizes permitted, we calculated F_{IS} (Weir & Cockerham 1984) for individuals from one log per geographical population, using GENEPOP version 3.3 (Raymond & Rousset 1995).

Partitioning of molecular variation, population divergence, and comparison of F_{ST} *vs.* Φ_{ST}

The distribution of genetic variation was determined via analysis of molecular variance (AMOVA) (Excoffier *et al*. 1992) implemented in ARLEQUIN version 2.000 (Schneider *et al*. 2000). Partitioning of variation within and among geographical populations was assessed (a hierarchical AMOVA was not appropriate owing to sample size constraints at the level of individual logs). Genetic divergence was estimated using F_{ST} (Weir & Cockerham 1984), calculated for each locus independently, allozyme loci only, anonymous scnDNA loci only and all nuclear loci combined.

To investigate the impact of incorporating information on nucleotide divergence (i.e. long-term, genealogical processes *sensu* Sunnucks 2000) upon estimates of population divergence based solely on frequency data (i.e. short-term genic processes), Φ_{ST} (Excoffier *et al*. 1992) was calculated for three loci (mtDNA, Sm2 and Sm8) in ARLEQUIN, then plotted against corresponding pair-wise and mean F_{ST} values. For each locus, the best fit model of molecular evolution was determined by means of likelihood ratio tests using MODELTEST version 3.06 (Posada & Crandall 1998), then a genetic distance matrix was generated using PAUP* version 4.0b10 (Swofford 2002). Significance of F_{ST} and Φ_{ST} was evaluated by permuting 16 000 times.

Alignment gaps were incorporated into the genetic distance matrices for loci Sm2 and Sm8. Adjacent multiple-base gaps were treated as a single indel following Retief *et al*.'s (1995) 'homologous gap coding' strategy. This is conservative, since a single deletion event is the most parsimonious explanation for contiguous alignment gaps. Given that it is presently computationally intractable to perform likelihood calculations that treat multiple-site indels as a single-event fifth-character state (McGuire *et al*. 2001), each indel was weighted equal to a base substitution (transversion was chosen). To achieve inclusion of indel data, one nucleotide position per indel position was added to the end of each sequence, with presence/absence of a given indel coded using 'T'/'A', respectively. Indel coding was used only in Φ_{ST} analyses.

GENECONV version 1.81 (Sawyer 1999) was employed to evaluate evidence for recombination. Following the method of Posada (2002), global permutation P-values based on BLAST-like global scores (10 000 replicates) smaller than 0.05 were considered evidence of recombination; the default mismatch penalty (*gscale* = 0) was used. Multiple-comparison correction for all sequence pairs in the alignment is built into global P-values, and contiguous alignment gaps are treated as a single segregating site. Hudson & Kaplan's (1985) bound R_m, the minimum number of recombination events in the history of a sample, was calculated using DNASP version 3.5 (Rozas & Rozas 1999). While recombination appears to be easier to detect in longer, more diverse sequences (Posada 2002), we note that positive identification of recombinant haplotypes has been reported for loci comparable to

those employed here (Ibrahim *et al*. 2002; Broughton & Harrison 2003).

Tajima's *D* (Tajima 1989) and Fu and Li's *D** (Fu & Li 1993) were calculated in DNASP to test for departures from neutrality. Significance was assessed via a two-tailed test assuming *D* follows a beta distribution, and Fu & Li's (1993) critical values, respectively.

Phylogenetic analysis

Genealogical relationships among mtDNA sequences were estimated via a haplotype network using the statistical parsimony method (Templeton *et al*. 1992) in TCS version 1.13 (Clement *et al*. 2000). By using haplotype frequency and divergence information, the method estimates haplotype outgroup probabilities, which correlate with haplotype age (Donnelly & Tavaré 1986). When relationships among mtDNA haplotypes make a multifurcating star-shaped network, one may tentatively assign polarity where the oldest haplotypes are taken as those at the centre (Beheregaray & Sunnucks 2001). Nested clade analysis (Templeton 1998) was not appropriate given that the spatial distribution of sampling sites was relatively sparse, and sampling effort was not exhaustive.

Estimating mtDNA divergence times assuming a molecular clock

To permit inference of the timing of divergence events, the hypothesis of mtDNA rate constancy was assessed via Tajima's (1993) relative rate test, implemented in MEGA version 2.1 (Kumar *et al*. 2001). A morphologically similar neanurid collembolon (above) was used an outgroup. To approximate absolute ages of divergence among haplotypes, we used the substitution rates 1.71% and 2.3%/My, based on those previously reported for invertebrate mitochondrial cytochrome oxidase genes (Brower 1994). These rates were applied to nucleotide divergences among haplotypes estimated using the best-fit model of molecular evolution selected by MODELTEST. Minimum and maximum divergences within and among major haplotype groups identified in the mtDNA network (above) were used to estimate divergence times.

Results

The newly developed nuclear markers and two allozymes are apparently codominant, single locus nuclear markers without appreciable null allele frequencies (Table 2). The largest available collections taken from single logs showed no evidence of consistent deviations from Hardy–Weinberg equilibrium, and there were no cases of putative homozygous nulls.

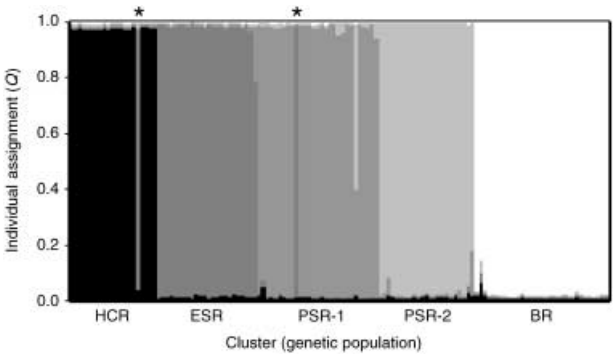

Fig. 2 Summary plot of *Q*, the estimated membership coefficient for each individual in each of the $K = 5$ clusters. Individuals are arranged sequentially from northern (left) to southern (right) geographical sampling localities (i.e. logs). Each individual is represented by a single vertical line divided into five shaded segments, where area is proportional to *Q*. Putative recent migrants are marked by asterisks.

Identification of population structure and assignment of individuals

Bayesian clustering analysis of the multilocus nuclear dataset strongly supported five genetic populations: the estimated log likelihood of the data progressively increased towards $K = 5$, then plateaued thereafter (not shown). The majority of individuals (94.8%) were assigned to one of the five genetic populations with high probability ($Q > 0.90$), and assignment of admixed individuals to the most likely population of origin was reasonably unambiguous ($Q > 0.75$). Only one individual was weakly assigned ($Q = 0.59$), but this membership coefficient was considerably larger than those obtained for alternative population designations (Fig. 2). All admixed individuals occurred in close geographical proximity to genetic contact zones. In all but one instance, admixed individuals possessed a multilocus genotype most characteristic of the two genetic populations flanking the contact zone nearest to where these individuals occurred (Fig. 2). Two individuals clearly originated from a genetic population adjacent to the one from which they were sampled (Fig. 2). This indicates a low level of dispersal; genetic populations are essentially allopatric.

The five genetic populations correspond with a priori microbiogeographical regions (see Materials and methods), although AR was not distinguished from PSR, possibly because the former was poorly represented in the dataset (a single log with two individuals). In addition PSR had two distinct genetic populations: PSR-1 encompassed most logs in this microbiogeographical region, while PSR-2 comprised the two easternmost logs sampled in DNP. Nonetheless, the spatial distribution of genetic populations does not conflict with any of the regional subdivisions recognized from physical landscape features and the palaeoclimatic history of Tallaganda (Fig. 1).

Table 2 Genetic diversity of mtDNA cytochrome oxidase subunit I (COI) and nuclear loci. Haplotype and allele frequencies are given per 'geographic' population, as identified by the software STRUCTURE (Pritchard *et al*. 2000). All F_{IS} values were significant ($P > 0.05$) unless indicated by ns.

Population/ locus		HCR	ESR	PSR-1	PSR-2	BR
mtDNA	n[a]	21	22	29	23	32
	A	0.71	—	—	—	—
	B	0.10	—	—	—	—
	C	0.05	—	—	—	—
	D	0.14	—	—	—	—
	E	—	0.45	0.03	—	—
	F	—	0.41	—	—	—
	G	—	0.05	—	—	—
	H	—	0.09	—	—	—
	I	—	—	0.07	—	—
	J	—	—	0.07	—	—
	K	—	—	0.83	0.91	—
	L	—	—	—	0.04	—
	M	—	—	—	0.04	—
	N	—	—	—	—	0.09
	O	—	—	—	—	0.09
	P	—	—	—	—	0.81
GPI	n[a]	3	2	12	23	12
	1	1.00	1.00	—	0.65	0.25
	2	—	—	1.00	0.02	—
	3	—	—	—	—	0.75
	4	—	—	—	0.33	—
	H_E	—	—	—	0.48	0.39
	H_O	—	—	—	0.35	—
	F_{IS}[b]	N/A	N/A	N/A	+0.30ns	N/A
LDH	n[a]	4	2	12	23	24
	1	—	—	0.96	0.13	—
	2	—	—	—	0.04	0.04
	3	1.00	1.00	0.04	0.83	—
	4	—	—	—	—	0.85
	5	—	—	—	—	0.10
	H_E	—	—	0.08	0.31	0.26
	H_O	—	—	0.08	0.26	0.21
	F_{IS}[b]	N/A	N/A	N/A	−0.14ns	−0.11ns
Sm2	n[a]	21	24	29	23	31
	1	0.05	1.00	0.04	—	—
	2	—	—	0.19	0.85	—
	3	—	—	0.78	0.15	—
	4	—	—	—	—	1.00
	5	0.95	—	—	—	—
	H_E	0.09	—	0.37	0.26	—
	H_O	—	—	0.17	0.22	—
	F_{IS}[b]	N/A	N/A	+1.00	+0.25ns	N/A
Sm4	n[a]	21	24	29	23	32
	1	—	0.64	0.02	—	—
	2	—	0.02	0.22	—	—
	3	1.00	0.34	0.76	1.00	1.00
	H_E	—	0.49	0.38	—	—
	H_O	—	0.68	0.48	—	—
	F_{IS}[b]	N/A	−0.27ns	−0.44ns	N/A	N/A
Sm6	n[a]	21	24	29	23	32
	1	1.00	0.96	0.35	0.46	—
	2	—	—	—	—	0.13
	3	—	0.04	0.02	—	—
	4	—	—	0.64	0.54	0.87
	H_E	—	0.08	0.48	0.51	0.23
	H_O	—	0.08	0.34	0.48	0.26
	F_{IS}[b]	N/A	N/A	+0.13ns	+0.07ns	−0.07ns
Sm8	n[a]	11	17	22	23	29
	1	—	—	0.16	0.52	—
	2	—	—	—	0.48	—
	3	—	—	—	—	0.07
	4	—	—	—	—	0.81
	5	—	—	—	—	0.09
	6	0.59	0.32	—	—	—
	7	—	0.06	—	—	—
	8	—	0.62	0.05	—	—
	9	0.41	—	0.25	—	—
	10	—	—	0.21	—	—
	11	—	—	0.34	—	—
	12	—	—	—	—	0.04
	H_E	0.51	0.53	0.77	0.51	0.34
	H_O	0.09	0.41	0.64	0.43	0.10
	F_{IS}[b]	N/A	−0.33ns	−0.17ns	+0.70ns	+0.35ns
	Mean H_E[c]	0.10 (0.20)	0.18 (0.25)	0.35 (0.28)	0.34 (0.20)	0.20 (0.17)
	Mean H_O[c]	0.02 (0.04)	0.20 (0.29)	0.29 (0.25)	0.29 (0.17)	0.10 (0.12)

[a]n = sample size per locus per geographic population.
[b]F_{IS} calculated for individuals from one log per geographic population (ESR n = 8; PSR-1 n = 11; PSR-2 n = 21; BR n = 26), except where indicated by 'N/A' (not applicable: sample size constraints or insufficient polymorphism).
[c]Standard deviation given in parentheses.

Table 3 Analysis of molecular variance (AMOVA) showing partitioning of variation based on F_{ST}

Source of variation	mtDNA	All nuclear loci	Allozyme loci	scnDNA loci
Within populations	41.15	39.44	29.86	36.55
Among populaitons	58.85	60.56	70.14	63.45

All variances were significant ($P < 0.01$).

Partitioning of molecular variation and population divergence

The majority of molecular variance occurred among geographical populations (Table 3). Haplotype/allele frequency data from mitochondrial COI, all nuclear loci combined, and all scnDNA loci combined yielded similar partitioning of molecular variation and F_{ST} estimates, while the two allozyme loci combined produced slightly higher estimates of population divergence (Table 3). The F_{ST} values differed considerably among loci, but mean F_{ST} values are clearly very large and significant for all loci ($P < 0.01$). Differentiation was most pronounced for Sm2 ($F_{ST} = 0.85$), followed by the two allozymes, then COI ($F_{ST} = 0.75$, 0.65, 0.59, respectively). The remaining three loci (Sm6, Sm4 and Sm8) produced the lowest estimates of population divergence ($F_{ST} = 0.56$, 0.47 and 0.46, respectively).

Comparison of F_{ST} *and* Φ_{ST}*: short- and long-term population divergence*

There was no significant evidence for recombination affecting the mtDNA COI locus or scnDNA loci Sm2 and Sm8 ($P > 0.05$, Table 4; Appendix 1). For all three loci, neither Tajima's D nor Fu & Li's D^* differed significantly from the neutral expectation of zero ($P < 0.1$). Accordingly, we considered these genetic markers largely unaffected by recombination and selection.

MtDNA COI, and nuclear loci Sm2 and Sm8 showed different relationships between Φ_{ST} and F_{ST} (Fig. 3). For all three loci, mean Φ_{ST} values exceeded those of corresponding F_{ST} values, but the magnitude of the difference was approximately 4.5 times greater for the mtDNA locus relative to scnDNA loci (Fig. 3a). This is consistent with the maternally inherited mtDNA locus being more predisposed to stochastic lineage sorting owing to its lower effective population size, which is one-fourth that of the nuclear genome if operational sex ratios are equal.

Pairwise geographical population comparisons for COI suggested that F_{ST} underestimated at least some population divergences (Fig. 3b). For example, for comparisons involving HCR, F_{ST} progressively increased in unison with increasing geographical distance but reached a maximum of $F_{ST} = 0.68$ (HCR verses PSR-2). Conversely, corresponding Φ_{ST} values increased almost linearly towards $\Phi_{ST} = 1$. Thus, except for the nonsignificant PSR-1/PSR-2 split (Fig. 3b), population subdivisions have occurred on timescales over which new mtDNA protein coding sequences evolve.

Nuclear loci Sm2 and Sm8 showed Φ_{ST} values generally larger than their corresponding F_{ST} values. However, in contrast to mtDNA, F_{ST} vs. Φ_{ST} pair-wise geographical population comparisons tended to mirror one another in their relative estimates of population differentiation (Fig. 3c,d). These differences between mtDNA and nuclear loci are consistent with Palumbi *et al.*'s (2001) proposed deeper coalescence times of diploid nuclear loci than haploid uniparental ones through the process of lineage sorting (caveats outlined by Hudson & Turelli 2003; Ballard & Whitlock 2004).

Phylogenetic relationships among mtDNA haplotypes

In satisfying the 95% confidence criterion of Templeton *et al.*'s (1992) statistical parsimony, two disconnected networks were produced. Phylogenetic relationships among mtDNA haplotypes mirror their geographical distribution with remarkable congruence, and indicate a deep phylogeographical structure (Fig. 4). Haplotypes cluster in accordance with *a priori* microbiogeographical regions (see Materials and methods; Fig. 1) and the five geographical populations inferred from Bayesian clustering (Figs 2 and 4b) with one exception: the PSR-1/PSR-2 distinction seen in nuclear loci was not detected in terms of mtDNA

Table 4 Tests of recombination affecting mtDNA cytochrome oxidase subunit I (COI), Sm2 and Sm8

Locus	Total sites	Polymorphic sites	Indels	Segregating sites	Inner fragment *P*-value	Outer fragment *P*-value	R_m
mtDNA	522	50	0	50	0.22	1.00	6
Sm2	201	9	1	10	0.13	0.13	0
Sm8	235	21	3	24	0.36	0.52	1

P-values are multiple-comparison corrected and based on 10 000 permutations; R_m is the minimum number of recombination events per locus based on the presence of possible recombination products.

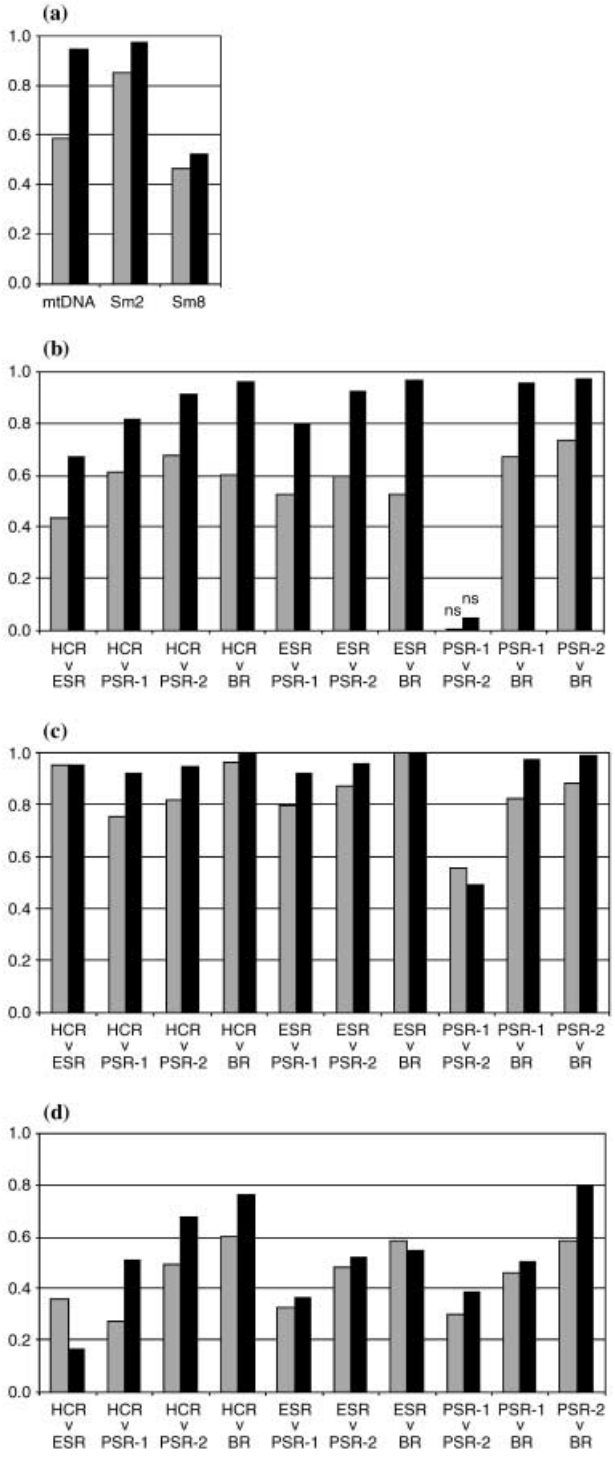

Fig. 3 Population divergences estimated by F_{ST} (light) and Φ_{ST} (dark). (a) Mean population divergence per locus; pair-wise population divergence for (b) mtDNA cytochrome oxidase subunit I (COI), (c) Sm2 and (d) Sm8. The best-fit models of nucleotide substitution selected to compute Φ_{ST} for the three loci were: GTR + I, F81, and TrN + I, respectively. All values were significant ($P < 0.01$) unless indicated by ns.

phylogeography. Except for the widespread haplotype K, haplotypes show considerable spatial localization.

Within the 'Northern network' (haplotypes A to M), most haplotypes derive from ESR haplotype F (Fig. 4). This haplotype also had the highest outgroup probability when mtDNA sequence data alone were considered. With the incorporation of frequency information, haplotype J was identified as the oldest haplotype, but we suggest this was spurious, possibly caused by the closely related haplotype K occurring at high frequency. The BR haplotypes form a separate 'Southern network' (haplotypes N to P; Fig. 4). Here, the presence of only three haplotypes precludes assignment of polarity.

Estimated mtDNA divergence times assuming a molecular clock

Significant departure from rate constancy was detected for mtDNA haplotype I ($P < 0.05$). It was consequently excluded from analyses involving the application of a molecular clock. Three main lineage-splitting events were evident, and their timing corresponds with major palaeoclimatological events affecting southeast Australia, irrespective of which molecular clock rate was applied (Table 5).

Discussion

Population structure, gene flow and dispersal

All analyses indicated significant genetic differentiation among geographical populations: closely related mtDNA haplotypes reside within a single population, or in adjacent ones; nuclear allele frequencies were more similar among proximate populations; genetic contact zones are abrupt, with little admixture across boundaries, and interpopulation migration is rare.

In general, Collembola are sedentary hexapods with limited long-distance movements. Rare long-distance dispersal may occur via wind (Farrow & Greenslade 1992), phoresy (King *et al*. 1985), or water (Moore 2002). Owing to the large size, highly hydrophobic cuticle and presumed desiccation susceptibility of the collembolon considered here, the latter mechanism is the most plausible explanation for two putative recent migrants (Fig. 2).

Patterns of population-genetic structuring over relatively small spatial scales have been reported for several collembola: *Isotoma klovstadi* (Frati *et al*. 2001), *Gressittacantha terranova* (Fanciulli *et al*. 2001), *Gomphiocephalus hodgsoni* (Stevens & Hogg 2003) and *Orchesella cincta* (van der Wurff *et al*. 2003). In the present study, substructuring was even more intense and interpopulation gene flow was apparently highly restricted. In forested areas, habitat patchiness has been reported to constitute a significant barrier to gene flow among invertebrates with poor vagility,

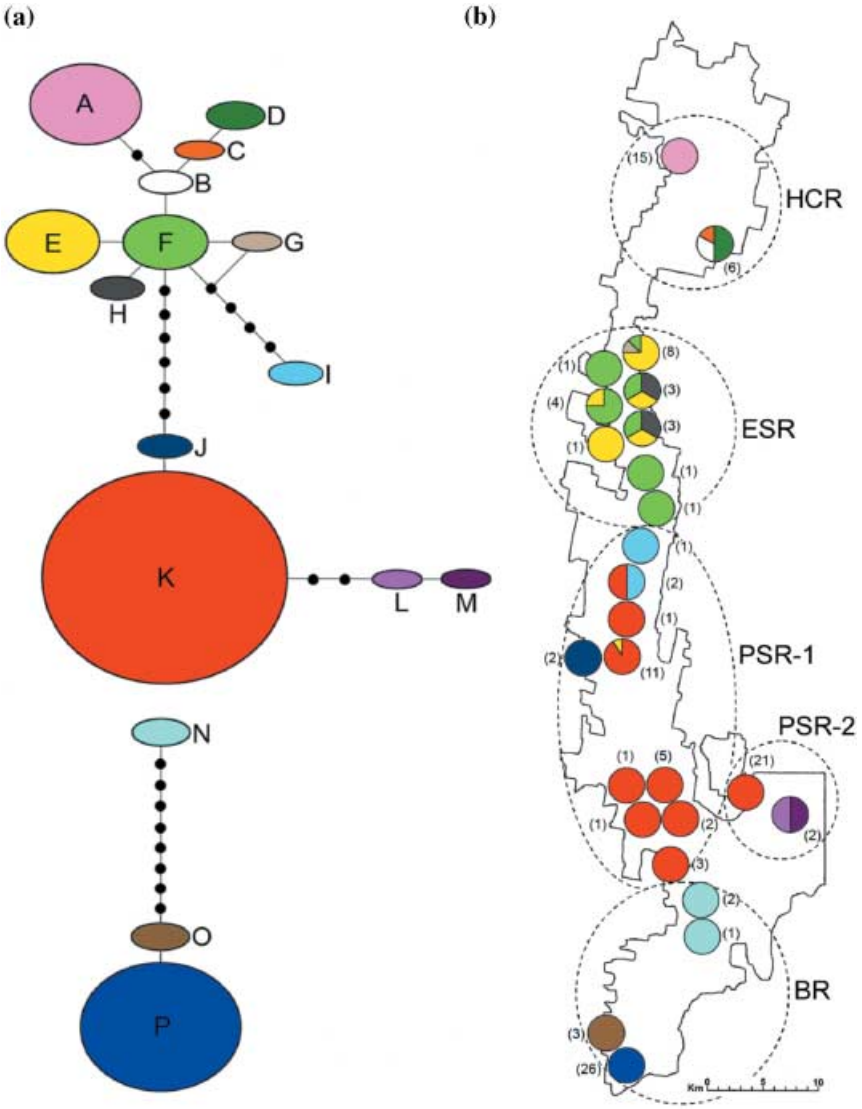

Fig. 4 Collembola mtDNA cytochrome oxidase subunit I (COI) phylogeographical patterning at Tallaganda. (a) Haplotype network estimated by statistical parsimony (Templeton *et al*. 1992). Letters designate unique haplotypes (also colour-coded). The areas of the ovals are proportional to frequency. Each single line represents one mutation between haplotypes, and small black circles dividing single lines represent missing haplotypes. Homoplasy, reflected by reticulation, was observed between haplotype G and a missing haplotype. (b) Geographic distribution of haplotypes. Each sampling unit (i.e. log) is represented by a pie chart showing haplotype frequency (colour-codes as (a)). Sample sizes per log are given in parentheses. Ellipses (broken lines) delineate logs comprising geographical populations, inferred from Bayesian clustering of multilocus nuclear genotypes (see Materials and methods and Fig. 2).

resulting in highly structured populations over very fine spatial scales (Keyghobadi *et al*. 1999). Considering the intrinsic features that limit dispersal ability of the focal collembolon (see Introduction), strong population structuring is expected. Nonetheless, the abrupt genetic contact zones uncovered here demand explanation.

Phylogeography

MtDNA haplotypes showed deep sequence divergences (up to 11.4% GTR + I corrected) and strong geographical structure among groups of haplotypes. This phylogeographical pattern typically results from isolation due to environmental or geographical barriers (Avise 2000). The approximate timing of divergences among major haplotype groupings coincides with palaeoclimatic events that affected southeast Australia (Table 5). The present-day distribution of genetic diversity in the focal collembolon appears to have been affected by three major climatic events throughout the species' evolutionary history.

In the late Miocene–early Pliocene (5–7 Mybp) there was a dramatic change from the previously warm and moist climate of southern Australia (Bowler 1982), when forest diversity was probably maximal (Martin 1994). The new conditions, which consisted of extended periods of intense seasonal aridity, imposed severe stresses upon flora and fauna (Bowler 1982), and coincided with the contraction of rainforests across southeast Australia (Singh 1982). The arid climatic regime prevailed until the late Pliocene (2.5 Mybp) (Bowler 1982). This event corresponds with the estimated time of divergence of haplotypes in the Northern network (haplotypes A–H and J–M) from those in the Southern network (haplotypes N–P) (Fig. 4a, Table 5). This initial split is corroborated by evidence from anonymous scnDNA locus

Table 5 Estimated mtDNA divergence times assuming a molecular clock

	Inferred timing of divergence (Mybp)					
	2.3%[b]		1.71%[b]			
	Min	Max	Min	Max	Palaeoclimatological event	Timing of event (Mybp)
Among groups[a]						
(A-H, J-M)/(N-P)	2.87	4.96	3.86	6.67	Terminal Miocene Event (Pliocene cooling and drying, extended period of aridity)	2.5–7
(A-H)/(J-M)	0.61	1.39	0.82	1.87	Late Pliocene — early Pleistocene (low amplitude and frequency wet-dry cycles)	> 0.7–2
(N)/(O-P)	0.78	0.91	1.05	1.23	Late Pliocene — early Pleistocene	> 0.7–2
Within groups[a]						
(A-H)	0.09	0.35	0.12	0.47	Mid — late Pleistocene (high amplitude and frequency glacial-interglacial cycles)	< 0.7–0.01
(J-M)	0.09	0.44	0.12	0.59	Mid — late Pleistocene	< 0.7–0.01
(O-P)	0.09	0.09	0.12	0.12	Mid — late Pleistocene	< 0.7–0.01

[a]MtDNA haplotype groupings indicated by parentheses, letters identify constituent haplotypes as per Fig. 4; '/' indicates splitting event.
[b]Molecular clock rate (% sequence divergence per My).
Haplotype groups identified from mtDNA genealogical relationships shown in Fig. 4a. Approximate timing of putative palaeoclimatological events taken from Bowler (1982) and Singh (1982).

Sm2: pairwise Φ_{ST}-based AMOVAs involving the BR geographical population consistently yielded the highest estimates of genetic differentiation (Fig. 3c). Although the BR is expected to have been dominated by woodland/steppe during cold dry periods, based on our findings the presence of refugia seems likely, possibly provided by deep gullies among the steep topography in the south-east.

While not as dramatic as impacts in the northern hemisphere (Hope 1994), the Quaternary saw 20 or more cycles of cooling and drying with periglacial conditions over much of southeast Australia above *c*. 1000 m. Early in the Pleistocene (*c*. 1.2 Mybp), the amplitude and frequency of wet–dry cycles increased and began to reflect the very large and rapid cyclic climatic fluctuations between cold dry glacial and warm humid interglacial periods that have characterized the latter part of the Quaternary (Bowler 1982). These early climatic oscillations coincide with divergence between two Northern network mtDNA subgroups (haplotypes A–H and J–M), and between two Southern network subgroups (haplotype N, and O to P) (Fig. 4a, Table 5). Based on mtDNA genealogical relationships, the oldest haplotype resides in the ESR. This is consistent with a priori expectations that ESR is likely to have served repeatedly as a refuge for moist forest-dependent organisms at Tallaganda during cold dry periods since the mid Miocene *c*. 15 Mybp, and perhaps as far back as 36 Mybp in the Oligocene epoch. Nonetheless, although deep east–west oriented gullies are most abundant in the ESR, high-elevation eastern slopes of PSR may also have retained refugia during glacial periods.

Ecologically, the last 0.4 My have been considered most important in shaping present-day floral and faunal biogeographical zonations because of a series of high-amplitude glacial-interglacial cycles occurring on *c*. 100 000-year rotations (Bowler 1982; Hope 1994). Such timescales appear to have been sufficiently short to limit species specialization to stable niches, and consequently plant populations and whole ecosystems were repeatedly wiped out. These relatively recent, very large and perhaps very rapid climatic oscillations coincide with evolutionary radiations within each of the two Northern network mtDNA subgroups (above) and within one of the Southern network subgroups (haplotypes O–P) (Fig. 4a, Table 5). Vicariance events that resulted in the present-day separation of HCR and ESR populations, and PSR-1 and PSR-2 populations, are clearly the most recent evolutionary episodes in the three-tiered history of this species at Tallaganda.

In contrast to other geographical populations, PSR-1 was dominated by a broadly distributed high-frequency haplotype (haplotype K), possibly indicating a relatively rapid and recent range expansion from a single glacial refuge. The high haplotypic diversity contained within HCR and ESR suggests that these microbiogeographical regions retained the largest population sizes and/or contained multiple refuges. However, the spatial distribution of nuclear genetic diversity contravenes the expectation of highest nuclear diversity co-occurring in those regions with the highest mtDNA diversity: PSR-1 and PSR-2 geographical populations have the greatest nuclear diversity (Table 2), yet their mtDNA diversity is low. There are

several plausible explanations for this pattern, but perhaps the most likely of these is illuminated when the spatial arrangement of the five geographical populations is considered: PSR-1 and PSR-2 are generally flanked by other populations to the north and south, and may act as sinks for nuclear genetic diversity. Gene flow into the PSR-1 and PSR-2 populations, sourced from both northern and southern origins, is evident from nuclear allele frequencies. These two geographical populations display a disproportionate number of cases where low-frequency alleles occur at high frequency in neighbouring populations (Table 2). Plausibility of the proposed source–sink relationship would be further strengthened if male-biased dispersal occurs in this animal. Currently, we only have anecdotal evidence to suggest male-biased dispersal (i.e. males are smaller in size and perhaps more active compared with females) since juveniles are very difficult to sex morphologically.

While the relative importance of Quaternary and earlier processes acting on biological diversification remains a contentious issue (Ribera & Vogler 2004) and probably varies considerably among regions and taxa, molecular studies have tended to lead to greater estimates of the ages of species, including vertebrates (Bermingham & Moritz 1998). Given the potential for single gullies to act as refuges for viable populations of moisture-dependent biota during glacial periods (Heatwole 1987; Hope 1994), and that Australia was not heavily impacted by glaciation compared with the northern hemisphere (Hope 1994) it is likely that population divergences of the focal collembolon species at Tallaganda could be of considerable antiquity. Based on the observed level of mtDNA divergence, and assuming a molecular clock, a late Pliocene — late Miocene coalescence of the present haplotypes seems likely (caveats outlined by Moritz *et al*. 1987). Results from comparisons of F_{ST} and Φ_{ST} support the notion that origins of population subdivision are of considerable age. The present data thus extend to temperate Australian eucalypt forests findings from studies of the Australian wet tropics showing that topographically heterogeneous upland areas can harbour, and indeed repeatedly generate, very high levels of biotic endemism (Moritz *et al*. 2001). Nuclear gene genealogies will undoubtedly aid in understanding the evolutionary history of this collembolon at Tallaganda, but will be presented elsewhere.

Conservation implications

Tallaganda is bounded by cleared grazing lands to the east, north and west. Forestry has been the predominant land use in the study area since the 1860s and remains so in Tallaganda and Badja SFs. Since the 1960s commercial forestry operations have intensified over progressively larger areas (State Forests of NSW 1995), and logging is planned to continue beyond 2005 (Australian Government 2001). Nearly all of the tall wet sclerophyll forests have been gazetted as State Forest, and may be subject to commercial forestry operations (Australian Government 2001). These areas encompass the majority of the putative glacial refuges, and the largest proportion of genetic diversity in the focal collembolon species.

The potential impacts upon terrestrial invertebrates from current and proposed forest management at Tallaganda remain to be established. However, the 1995 Environmental Impact Statement (EIS) relating to the study area stated 'Any terrestrial invertebrates totally dependent on old trees may be adversely affected if their populations are localized and not widespread in the EIS area' (State Forests of NSW 1995, EIS Main Report, Section 11 p. 66). The present new genus and species of saproxylic Collembola clearly fits this description.

Acknowledgements

We thank NSW National Parks and State Forests for permission to collect saproxylic invertebrates. Dave Runciman, Mark Blacket, Mel Lancaster and other members of our research groups gave invaluable assistance and discussions (see http://www.latrobe.edu.au/genetics/staff/sunnucks/homepage/). Dave Briscoe, Mark Westoby and Andy Beattie contributed ideas that inspired this program, and Geoff Hope inspired the environmental model. This work was funded by grants from the Australian Research Council (grant DP-0211156) to PS and DMR, and the Holsworth Wildlife Research Fund to RCG. The DEST Systematic Infrastructure Initiative grant funded the LI-COR. We thank Graham Wallis, Nick Porch, John Trueman, Mike Westerman and two anonymous reviewers for valuable comments on the manuscript.

References

Arbogast BS, Kenagy GJ (2001) Comparative phylogeography as an integrative approach to historical biogeography. *Journal of Biogeography*, **28**, 819–825.

Australian Government (2001) *Regional Forest Agreement*. Department of Agriculture, Fisheries and Forestry, http://www.rfa.gov.au.

Avise JC (2000) *Phylogeography: the History and Formation of Species*. Harvard University Press, Cambridge, MA.

Ballard JMO, Whitlock MC (2004) The incomplete natural history of mitochondria. *Molecular Ecology*, **13**, 729–744.

Barclay S, Ash JE, Rowell DM (2000a) 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.

Barclay SD, Rowell DM, Ash JE (2000b) Pheromonally mediated colonization patterns in the velvet worm *Euperipatoides rowelli* (Onychophora). *Journal of the Zoological Society of London*, **250**, 437–446.

Beheregaray LB, Sunnucks P (2001) Fine-scale genetic structure, estuarine colonization and incipient speciation in the marine silverside fish *Odontesthes argentinensis*. *Molecular Ecology*, **10**, 2849–2866.

Beheregaray LB, Sunnucks P, Briscoe DA (2002) A rapid fish radiation associated with the last sea-level changes in southern

Brazil: the silverside *Odontesthes perugiae* complex. *Proceedings of the Royal Society of London Series B*, **269**, 65–73.

Bermingham E, Moritz C (1998) Comparative phylogeography: concepts and applications. *Molecular Ecology*, **7**, 367–369.

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.

Broughton RE, Harrison RG (2003) Nuclear gene genealogies reveal historical, demographic and selective factors associated with speciation in field crickets. *Genetics*, **163**, 1389–1401.

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 of the USA*, **91**, 6491–6495.

Burt A, Carter DA, White TJ, Taylor JW (1994) DNA sequencing with arbitrary primer pairs. *Molecular Ecology*, **3**, 523–525.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1660.

CSIRO (Commonwealth Scientific and Industrial Research Organisation) (1969) *The Lands of the Queanbeyan–Shoalhaven Area, ACT and NSW. Land Research Series* Number 24. CSIRO, Melbourne.

Donnelly P, Tavaré S (1986) The ages of alleles and a coalescent. *Advances in Applied Probability*, **18**, 1–9.

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.

Fanciulli PP, Summa D, Dallai R, Frati F (2001) High levels of genetic variability and population differentiation in *Gressittacantha terranova* (Collembola, Hexapoda) from Victoria Land, Antarctica. *Antarctic Science*, **13**, 246–254.

Farrow RA, Greenslade P (1992) A vertical migration of Collembola. *Entomologist*, **111**, 38–45.

Ferrier S, Gray MR, Cassis GA, Wilkie L (1999) Spatial turnover in species composition of ground-dwelling arthropods, vertebrates and vascular plants in north-east New South Wales: implications for selection of forest reserves. In: *The Other 99% The Conservation and Biodiversity of Invertebrates* (eds Ponder W, Lunney D), pp. 68–76. The Royal Zoological Society of New South Wales, Mosman.

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.

Frankham R, Ballou JD, Briscoe DA (2002) *Introduction to Conservation Genetics*, pp. 1–20. Cambridge University Press, Cambridge.

Frati F, Spinsanti G, Dallai R (2001) Genetic variation of mtCOII gene sequences in the collembolan *Isotoma klovstadi* from Victoria Land, Antarctica: evidence for population differentiation. *Polar Biology*, **24**, 943–940.

Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. *Genetics*, **133**, 693–709.

Grove SJ (2002) Tree basal area and dead wood as surrogate indicators of saproxylic insect faunal integrity: a case study from the Australian lowland tropics. *Ecological Indicators*, **1**, 171–188.

Hammond PM (1984) Changes in the British coleopterous fauna. In: *The Changing Flora and Fauna of Britain. Proceedings of a Symposium of the Systematics Association* (ed. Hawksworth DL), pp. 323–369. Academic Press, London.

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.

Hebert PDN, Beaton MJ (1993) *Methodologies for Allozyme Analysis Using Cellulose Acetate Electrophoresis*. Helena Laboratories, Beaumont, Texas.

Hope GS (1994) Quaternary vegetation. In: *History of the Australian Vegetation: Cretaceous to Recent* (ed. Hill RS), pp. 368–389. Cambridge University Press, Cambridge.

Hudson R, Kaplan N (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics*, **111**, 147–164.

Hudson RR, Turelli M (2003) Stochasticity overrules the 'three-times rule': genetic drift, genetic draft, and coalescence times for nuclear loci verses mitochondrial DNA. *Evolution*, **57**, 182–190.

Ibrahim KM, Cooper SJB, Hewitt GM (2002) Testing for recombination in a short nuclear DNA sequence of the European meadow grasshopper, *Chorthippus parallelus*. *Molecular Ecology*, **11**, 583–590.

Jarman SN, Ward RD, Elliot NG (2002) Oligonucleotide primers for PCR amplification of coelomate introns. *Marine Biotechnology*, **4**, 347–355.

Jonsson BG, Kruys N (2001) Ecology of woody debris in boreal forests. *Ecological Bulletins*, **49**, 1–281.

Key R (1993) What are saproxylic invertebrates?. In: *Dead Wood Matters: the Ecology and Conservation of Saproxylic Invertebrates in Britain* (eds Kirby KJ, Drake CM). English Nature, Peterborough.

Keyghobadi H, Roland J, Strobeck C (1999) Influence of landscape on population genetic structure of the alpine butterfly *Parnassius smintheus* (Papilionidae). *Molecular Ecology*, **8**, 1481–1495.

King KL, Greenslade P, Hutchinons KJ (1985) Collembolan associations in natural verses improved pastures of the New England Tableland, NSW: distribution of native and introduced species. *Australian Journal of Ecology*, **10**, 421–427.

Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA: *Molecular Evolutionary Genetics Analysis Software*, Version 2.1 Arizona. State University, Tempe, AZ.

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.

Martin HA (1994) Australian Tertiary phytogeography: evidence from palynology. In: *History of the Australian Vegetation: Cretaceous to Recent* (ed. Hill RS), pp. 104–142. Cambridge University Press, Cambridge.

McGuire G, Denham MC, Balding DJ (2001) Models of sequence evolution for DNA sequences containing gaps. *Molecular Biology and Evolution*, **18**, 481–490.

Moore PD (2002) Springboards for springtails. *Nature*, **418**, 381.

Moritz C, Dowling TE, Brown WM (1987) Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Annual Review of Ecology and Systematics*, **18**, 269–292.

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*, **268**, 1875–1881.

Murray MG, Thompson WF (1980) Rapid isolation of high-molecular-weight plant DNA. *Nucleic Acids Research*, **8**, 4321–4325.

Palumbi SR, Cipriano F, Hare MP (2001) Predicting nuclear gene coalescence from mitochondrial data: the three-times rule. *Evolution*, **55**, 859–868.

Peakall R, Beattie AJ (1991) The genetic consequences of worker ant pollination in a self-compatible clonal orchid. *Evolution*, **45**, 1837–1848.

Peterson MA, Denno RF (1998) The influence of dispersal and diet breadth on patterns of genetic isolation by distance in phytophagous insects. *American Naturalist*, **152**, 428–446.

Posada D (2002) Evaluation of methods for detecting recombination from DNA sequences: empirical data. *Molecular Biology and Evolution*, **19**, 708–717.

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Raymond M, Rousset F (1995) GENEPOP version 1.2.: population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Retief JD, Krajewski C, Westerman M, Winkfein RJ, Dixon GH (1995) Molecular phylogeny and evolution of marsupial protamine P1 genes. *Proceedings of the Royal Society of London Series B*, **259**, 7–14.

Ribera I, Vogler AP (2004) Speciation of Iberian diving beetles in Pleistocene refugia (Coleoptera, Dytiscidae). *Molecular Ecology*, **13**, 179–193.

Rozas J, Rozas R (1999) DNASP, version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics*, **15**, 174–175.

Sawyer SA (1999) *GENECONV: a Computer Package For The Statistical Detection Of Gene Conversion* (distributed by the author). Department of Mathematics. Washington University in St Louis, St Louis, MO. http://www.math.wustl.edu/~sawyer

Schneider S, Roessli D, Excoffier L (2000) *ARLEQUIN, Version 2.000: a Software for Population Genetic Data Analysis Genetics and Biometry Laboratory*. University of Geneva, Geneva.

Scott IAW, Rowell DM (1991) Population biology of *Euperipatoides leuckartii* (Onychophora: Peripatopsidae). *Australian Journal of Zoology*, **39**, 499–508.

Simon C, Frati F, Beckenbach A *et al.* (1994) Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America*, **87**, 651–701.

Singh G (1982) Environmental upheaval: vegetation of Australasia during the Quaternary. In: *A History of Australasian Vegetation* (ed. Smith JMB), pp. 90–108. McGraw-Hill, Sydney.

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, NSW, Australia.

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.

Stiling P (1999) *Ecology: Theories and Applications*, 3rd edn. Prentice Hall, Englewood Cliffs, NJ.

Sunnucks P (2000) Efficient genetic markers for population biology. *Trends in Ecology and Evolution*, **15**, 199–203.

Sunnucks P, Tait N (2001) What's so interesting about velvetworm sex? Tales of the unexpected. *Nature Australia Winter edition*, **27**, 61–69.

Sunnucks P, Wilson ACC (1999) Microsatellite markers for the onychophoran *Euperipatoides rowelli*. *Molecular Ecology Primer Notes*, **8**, 895–906.

Sunnucks P, Curach NC, Young A *et al.* (2000a) Reproductive biology of the onychophoran *Euperipatoides rowelli*. *Journal of the Zoological Society of London*, **250**, 447–460.

Sunnucks P, Wilson ACC, Beheregaray LB *et al.* (2000b) 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.

Swofford DL (2002) *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods)*, Version 4. Sinauer Associates, Sunderland, MA.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Tajima F (1993) Simple methods for testing molecular clock hypothesis. *Genetics*, **135**, 599–607.

Templeton AR (1998) Nested clade analysis of phylogeographical data: testing hypotheses about gene flow and population history. *Molecular Ecology*, **7**, 381–397.

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.

Thomas JA, Morris MG (1994) Patterns, mechanisms and rates of extinction among invertebrates in the United Kingdom. *Philosophical Transactions of the Royal Society of London Series B*, **344**, 47–54.

Warren MS, Key RS (1991) Woodlands: past, present and potential for insects. In: *The Conservation of Insects and Their Habitats* (eds Collins NM, Thomas JA), pp. 155–212. Academic Press, London.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

White ME (1990) *The Nature of Hidden Worlds*, pp. 176–225. Reed Books, Balgowlah.

White ME (1994) *After the Greening*, pp. 41–245. Kangaroo Press, Kenthurst.

van der Wurff AWG, Isaaks JA, Ernsting G, van Straalen NM (2003) Population substructures in the soil invertebrate *Orchesella cincta*, as revealed by microsatellite and TE-AFLP markers. *Molecular Ecology*, **12**, 1349–1359.

Yee M, Yuan ZQ, Mohammed C (2001) Not just waste wood: decaying logs as key habitats in Tasmania's wet sclerophyll *Eucalyptus obliqua* production forests: the ecology of large and small logs compared. *Tasforests*, **13**, 119–128.

This work represents part of Ryan Garrick's PhD research, and is a component of our comparative phylogeography program at Tallaganda. Chester Sands is a fellow PhD student and has contributed to various aspects of this work. Dave Rowell is Senior Lecturer at ANU and has been conducting genetic and ecological research on Tallaganda invertebrates for nearly two decades. Noel Tait has a life-long passion for invertebrates, and has contributed throughout this program. Penny Greenslade is a morphotaxonomist and ecologist with particular interests in Collembola. Paul Sunnucks is Senior Lecturer in Genetics whose molecular population biology research programs encompass saproxylic invertebrates.

Appendix 1

Distribution of polymorphic sites: (a) mtDNA cytochrome oxidase subunit I (COI); (b) Sm2; and (c) Sm8 locus. Site designations are nucleotide positions relative to the end of the forward primer (excluding areas of ambiguous alignment; – deletion, and / multibase insertions)

(a)

Haplotype/	mtDNA COI locus																										
site	12	15	18	27	36	57	60	72	78	81	115	120	129	138	141	159	168	174	183	186	187	204	234	240	247	255	258
A	G	T	G	A	G	C	A	C	C	G	T	A	A	T	A	G	T	A	A	A	T	T	G	C	T	T	A
B	G	T	A	A	G	C	A	C	C	G	T	A	A	T	A	G	T	A	A	A	T	T	G	C	T	T	A
C	G	T	A	A	G	C	A	C	C	G	T	A	A	T	A	G	T	A	A	A	T	T	G	C	T	T	A
D	G	T	A	A	G	C	A	C	C	G	T	A	A	T	A	G	T	A	A	A	T	T	G	C	T	T	A
E	G	T	A	A	A	C	A	C	C	G	T	A	A	T	A	G	T	A	A	A	T	T	G	C	T	T	A
F	G	T	A	A	A	C	A	C	C	G	T	A	A	T	A	G	T	A	A	A	T	T	G	C	T	T	A
G	G	T	A	A	A	C	A	C	C	G	T	A	A	T	A	G	T	A	A	A	T	T	G	C	T	T	A
H	G	T	A	A	A	C	A	C	C	G	T	A	A	T	A	G	T	A	A	A	T	T	G	C	T	T	A
I	G	T	A	C	A	C	A	C	C	G	T	A	A	T	A	G	T	A	A	A	T	T	G	C	T	T	A
J	A	T	A	A	A	A	A	C	A	G	T	G	A	T	G	G	T	A	A	A	T	T	G	C	T	T	A
K	A	T	A	A	A	A	A	C	A	G	T	G	A	T	G	G	T	A	A	A	T	T	G	C	T	T	A
L	A	T	A	A	A	A	A	C	A	G	T	G	A	T	G	G	T	A	A	A	C	T	G	C	T	T	A
M	A	T	A	A	A	A	A	C	A	G	T	G	A	T	G	G	T	A	A	G	C	T	G	C	T	T	A
N	A	A	A	A	A	C	A	T	C	A	C	A	C	C	G	A	T	G	A	G	T	A	A	C	C	A	C
O	A	A	A	A	A	C	C	T	C	A	C	A	C	C	G	A	C	A	A	G	T	A	A	T	C	A	C
P	A	A	A	A	A	C	C	T	C	A	C	A	C	C	G	A	C	A	G	G	T	A	A	T	C	A	C

Haplotype/	mtDNA COI locus continued																							GenBank
site	261	270	315	327	333	336	339	342	357	363	381	390	396	399	408	417	418	447	468	471	498	504	510	accession number
A	A	C	T	G	T	T	T	T	A	A	T	T	T	T	A	T	A	C	G	G	A	G	A	AY694098
B	A	C	T	G	T	T	T	T	G	A	T	T	T	T	A	T	A	C	G	G	A	G	A	AY694099
C	A	C	T	G	T	T	T	T	G	A	T	T	T	T	A	T	A	C	G	A	A	G	A	AY694100
D	A	C	T	G	T	T	T	T	G	A	T	T	T	T	A	T	A	C	G	A	A	G	A	AY694101
E	A	C	T	G	T	T	T	T	G	A	T	T	T	T	A	T	A	C	G	G	A	G	G	AY694102
F	A	C	T	G	T	T	T	T	G	A	T	T	T	T	A	T	A	C	G	G	A	G	A	AY694103
G	A	C	T	A	T	T	T	T	G	A	T	T	T	T	A	T	A	C	G	G	A	G	A	AY694104
H	T	C	T	G	T	T	T	T	G	A	T	T	T	T	A	T	A	C	G	G	A	G	A	AY694105
I	A	C	T	C	T	T	T	C	G	A	T	C	T	T	A	T	G	C	G	G	G	G	A	AY694106
J	A	C	T	G	T	T	T	T	G	A	T	T	C	T	A	T	A	C	A	G	G	G	A	AY694107
K	A	C	T	G	T	T	T	C	G	A	T	T	C	T	A	T	A	C	A	G	G	G	A	AY694108
L	A	C	T	G	T	T	T	C	G	A	T	T	C	T	A	C	A	C	A	A	A	G	A	AY694109
M	A	C	T	G	T	T	T	C	G	A	T	T	C	T	A	C	A	C	A	A	A	G	A	AY694110
N	T	T	T	G	A	C	G	T	A	A	C	T	T	C	A	T	A	C	A	G	G	A	A	AY694111
O	T	T	C	G	A	C	G	T	A	G	C	T	T	C	C	T	A	C	G	A	G	A	A	AY694112
P	T	T	C	G	A	C	G	T	A	G	C	T	T	C	C	T	A	C	G	A	G	A	A	AY694113

(b)

Allele/	Sm2 locus														
site	16	23	37	68	83	88	123–124	125–136	137	138	139	140–150	156	163	177
1	T	T	G	A	T	A	/	/	T	A	A	/	G	A	C
2	T	T	G	A	T	A	/	/	T	T	A	/	T	T	C
3	C	T	G	A	T	A	/	/	T	T	A	/	T	T	C
4	T	G	A	G	C	G	/	—	T	T	T	/	G	T	T
5	T	T	G	A	T	A	—	—	—	—	—	—	G	A	C

Appendix 1 *Continued*
(c)

Allele/ site	Sm8 locus																								
	1	4	5	7	9	14	15	17	18	23	24	30	33	36	38	62	67	97	124	130	145	175	184	187	190
1	G	—	A	A	G	A	C	A	T	—	G	T	T	G	T	A	T	G	T	A	G	A	A	A	A
2	G	—	A	A	G	A	C	A	T	—	G	T	T	G	T	A	T	G	T	A	A	A	A	A	A
3	G	—	A	A	G	C	C	A	G	—	G	T	T	G	T	A	T	T	T	A	A	A	A	A	A
4	G	—	A	A	G	C	C	A	T	—	G	T	T	G	T	A	T	T	T	A	A	A	A	A	A
5	G	—	A	A	G	C	C	A	T	—	G	T	T	G	T	A	T	T	C	G	A	A	A	A	A
6	G	—	G	A	T	T	C	G	T	—	—	T	A	A	C	A	T	G	T	A	G	A	A	A	A
7	G	—	G	A	T	T	C	G	T	T	G	—	C	A	C	T	G	G	T	A	G	G	A	C	G
8	A	—	A	A	G	C	C	A	T	—	G	T	T	C	T	A	T	T	T	A	G	A	A	A	A
9	G	—	A	A	G	C	C	A	T	—	G	T	T	C	T	A	T	G	T	A	G	A	G	A	A
10	G	T	A	C	G	C	C	A	T	—	G	T	T	G	T	A	T	G	T	A	A	A	A	A	A
11	G	—	A	A	G	C	C	A	T	—	G	T	T	G	T	A	T	G	T	A	A	A	A	A	A
12	G	—	A	A	G	C	T	A	T	—	G	T	T	G	T	A	T	T	C	G	A	A	A	A	A