The LINNEAN SOCIETY of London BIOLOGICAL Journal of the Linnean Society

Biological Journal of the Linnean Society, 2016, **119**, 166–178. With 8 figures.

High levels of intraspecific genetic divergences revealed for Antarctic springtails: evidence for small-scale isolation during Pleistocene glaciation

KRISTI R. BENNETT[1], IAN D. HOGG[1]*, BYRON J. ADAMS[2] and PAUL D. N. HEBERT[3]

[1]*School of Science, University of Waikato, Private Bag 3105, Hamilton, 3240, New Zealand*
[2]*Department of Biology and Evolutionary Ecology Laboratories, Brigham Young University, 4102 LSB, Provo, UT 84602-5253, USA*
[3]*Biodiversity Institute of Ontario, University of Guelph, 50 Stone Road East, Guelph, ON N1G 2W1, Canada*

Received 29 October 2015; revised 4 February 2016; accepted for publication 4 February 2016

We examined levels of genetic variability within and among populations of three Antarctic springtail species (Arthropoda: Collembola) and tested the hypothesis that genetic divergences occur among glacially-isolated habitats. The study was conducted in southern Victoria Land, Ross Dependency, Antarctica, and samples were collected from locations in the vicinity of the Mackay Glacier. We analyzed mtDNA (cytochrome *c* oxidase subunit I; COI) sequence variability for 97 individuals representing three species (*Gomphiocephalus hodgsoni*, $N = 67$; *Cryptopygus nivicolus*, $N = 20$; and *Antarcticinella monoculata*, $N = 8$). Haplotype diversity and genetic divergences were calculated and used to indicate population variability and also to infer divergence times of isolated populations using molecular clock estimates. Two of the three species showed high levels of genetic divergence. *Gomphiocephalus hodgsoni*, a widespread and common species, showed 7.6% sequence divergence on opposite sides of the Mackay Glacier. The more range restricted *C. nivicolus* showed 4.0% divergence among populations. The third species, *A. monoculata*, was found in only one location. Molecular clock estimates based on sequence divergences suggest that populations separated within the last 4 Mya. We conclude that habitat fragmentation resulting from Pliocene (5 Mya) and Pleistocene (2 Mya to 10 Kya) glaciations has promoted and maintained high levels of diversity among isolated springtail populations on relatively small spatial scales. The region surrounding the Mackay Glacier is likely to have provided refugia for springtail populations during glacial maxima and remains an area of high genetic and species diversity for Collembola within the Ross Sea region.

KEYWORDS: Collembola – glaciation – population genetics – refugia – Ross Sea region – springtails.

INTRODUCTION

With only 0.34% (46 200 km²) of the total 14 million km² ice free and even marginally habitable, the Antarctic continent represents one of the most extreme environments for terrestrial life (Convey *et al.*, 2009; Hogg & Wall, 2012). The majority of these ice-free areas lie within the Dry Valleys and Transantarctic Mountains of the Ross Dependency (Janetschek, 1967a; Levy, 2012). Even here, exposed ground is often highly fragmented and comprises small, rocky outcrops separated by permanent snow fields and glaciers. Suitable habitat is then further restricted by the availability of liquid water necessary to support life (Hogg *et al.*, 2006). This latter requirement is relevant for the soil arthropod fauna, particularly the Antarctic springtails, which lack a desiccation-resistant life stage and, instead, use avoidance and super-cooling methods to enable survival in sub-zero temperatures (McGaughran, Hogg & Convey, 2011a).

The terrestrial arthropods are represented primarily by springtails (Collembola) and mites (Acari) and are the largest year-round taxa on the continent

*Corresponding author. E-mail: hogg@waikato.ac.nz

(Gressitt, 1967; Hogg & Stevens, 2002; Adams *et al.*, 2006). These taxa, which lack the survival and dispersal strategies possessed by other invertebrate groups such as nematodes (Nkem *et al.*, 2006; Adhikari, Wall & Adams, 2010), have been restricted to these fragmented, ice-free zones ever since the Middle Miocene (11–14 Mya) (Stevens & Hogg, 2003; Stevens *et al.*, 2006; McGaughran *et al.*, 2010). At this time, the glaciation of the whole continent reached its fullest extent and the polar ice cap overflowed the Transantarctic Mountains (Lewis *et al.*, 2007). Small oases of ice-free ground existed around the edge of the polar cap, the largest of which (the Dry Valleys) is still located within the Transantarctic Mountain on the western edge of the Ross Ice Shelf (Clapperton & Sugden, 1990). Subsequently, the East Antarctic Ice Sheet (EAIS) has undergone numerous glacial cycles, with the last glacial maximum ending 17 Kya (Suggate, 1990). This extensive glacial history has resulted in extremely low species richness for the Antarctic fauna, with many habitats containing one or two arthropod taxa at most (Janetschek, 1967a). Species are also rarely shared between regions (Gressitt, 1967; Wise, 1971; Sinclair & Stevens, 2006), suggesting limited inter-habitat dispersal. Consequently, the current arthropod taxa are likely to be long-term inhabitants and remnants of a previously more widespread species (Convey *et al.*, 2009). Even within regions, most species show high levels of genetic divergence across their distributional ranges, suggesting the effects of long-term isolation and/or survival in glacial refugia (Frati, Spinsanti & Dallai, 2001; Stevens & Hogg, 2003; McGaughran, Hogg & Stevens, 2008; Hawes, Torricelli & Stevens, 2010; Stevens & D'Haese, 2014). The present study aimed to extend these studies by focusing on the small-scale differences that might occur within faunally-diverse, yet heavily fragmented, landscapes.

Ten species of springtail are currently known from the Ross Dependency: four in northern Victoria Land, three in southern Victoria Land, and three in the southern Transantarctic Mountains. All species are range-restricted. Species from southern Victoria Land, the focus of the present study, consist of three species covering a 3° latitudinal range. Within this region, *Gomphiocephalus hodgsoni* is the only relatively widespread species and is common throughout southern Victoria Land (McGaughran *et al.*, 2011b). Two additional species, *Cryptopygus nivicolus* [recently redescribed from *Neocryptopygus nivicous* by Greenslade (2015)] and *Antarcticinella monoculata*, are extremely range-restricted and known only from one or two locations near the Mackay Glacier to the north of the Dry Valleys (Wise, 1971) (Fig. 1), suggesting the possibility of a glacial refugium. Recent

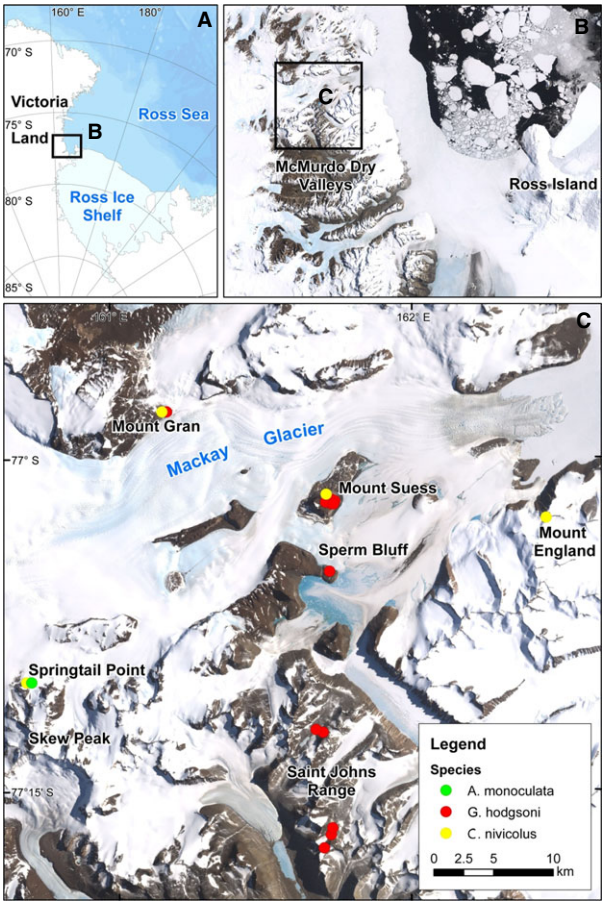

Figure 1. Sampling sites and Collembola species' locations in the Mackay Glacier vicinity. Two *Cryptopygus nivicolus* specimens were taken from GenBank and were collected from Mt England in 2005. Map adapted from the SCAR Antarctic Digital Database and the Landsat Image Mosaic of Antarctica (LIMA) project.

studies of lichens and mosses also near the Mackay Glacier (Green *et al.*, 2011), as well as haplotype diversity for springtail (*G. hodgsoni*) and mite (*Stereotydeus. mollis*) taxa, have further suggested this area as a likely refugial zone (Stevens & Hogg, 2003, 2006; McGaughran *et al.*, 2008; Demetras *et al.*, 2010).

To determine the geographical scales on which genetic diversity may have been promoted and/or maintained, we focused on small-scale genetic variability in a region of comparatively high species diversity (Mackay Glacier, southern Victoria Land). This glacier is one of only a few outlet glaciers that connect the EAIS with the Ross Ice Shelf in southern Victoria Land (Clapperton & Sugden, 1990). Accordingly, we tested the hypothesis that this region would support genetically divergent springtail populations among isolated habitats. We predicted that high levels of

both genetic variability and genetic divergence would exist among these habitats, potentially indicating refugial zones from the Pleistocene glaciations.

MATERIAL AND METHODS

Study sites and sample collection

Samples were collected from St John's Ranges near Victoria Valley and on the northern and southern sides of the Mackay Glacier in the northern Dry Valleys region of the Ross Dependency (Fig. 1). Specimens were collected from the undersides of rocks using modified aspirators (Stevens & Hogg, 2002). Soil samples were also taken from each site and returned to the laboratory where they were suspended a 10% sucrose solution. Invertebrates were then removed from the solution surface under a dissecting microscope (× 10 magnification) using a fine wire loop. All specimens were stored in 95% ethanol and returned to the University of Waikato for further processing. All specimens were morphologically identified to species level using Gressitt, Leech & Wise (1963) and Salmon (1965). Specimens not used for DNA analyses were archived at the School of Science, University of Waikato, under the care of IDH.

Genetic analysis

Genetic analyses were jointly carried out at the University of Waikato and at the Canadian Centre for DNA Barcoding (CCDB) at the University of Guelph. At the University of Waikato, total genomic DNA was extracted from the tissue of entire specimens using a Glassfiber Plate DNA Extraction (AcroPrep) method (Ivanova, deWaard & Hebert, 2006) at CCDB, and Red Extract n Amp (Sigma-Aldrich) using 10 μL of extraction solution and 2.5 μL of tissue preparation, in accordance with the manufacturer's instructions. Polymerase chain reactions (PCRs) comprised a 15-μL reaction containing 5.7 μL of MQH_20, 7.5 μL of PCR Master Mix Solution (i-Taq; Intron Biotechnology), 0.4 μL of each primer and 1 μL of template DNA. A 658-bp fragment of the mitochondrial cytochrome *c* oxidase subunit I (COI) gene was amplified using the primers HCO2198 (sequence 5′-TAAACTTCAGGGTGACCAAAAAATCA-3′) and an altered LCO1490 (sequence: 5′-AGTTCTAAT CATTAARGATATYGG-3′) (Folmer *et al.*, 1994) for the *G. hodgsoni* specimens. HCO and LepF1 (sequence: 5′-ATTCAACCAATCATAAAGATATTGG-3′) (Hajibabaei *et al.*, 2006) were used to amplify the *C. nivicolus* and *A. monoculata* specimens. The standard LCO1490 (sequence: 5′-GGTCAACAAATCA-TAAAGATATTGG-3′) was used for both species (in place of the altered LCO1490 and LepF1) at CCDB. Primers were used at a concentration of 1.0 mM. PCR conditions at CCDB were: initial denaturing at 94 °C for 1 min; five cycles of 94 °C for 1 min, 45 °C for 1.5 min and 72 °C for 1.5 min; 35 cycles of 94 °C for 1 min, 50 °C for 1.5 min and 72 °C for 1 min followed by a final 72 °C for 5 min. PCR conditions at the University of Waikato were: initial denaturing at 94 °C for 5 min; 36 cycles of 94 °C for 1 min, 52 °C for 1.5 min and 72 °C for 1 min, followed by a final 72 °C for 5 min.

PCR products were cleaned using Sephadex (CCDB) or 0.2 μL of ExonucleaseI and 0.1 μL shrimp alkaline phosphate with 2.7 μL of MQH_20 in accordance with the manufacturer's instructions (Global Science & Tech Ltd) at Waikato. DNA was sequenced in both directions on an ABI3130 sequencer at the University of Waikato DNA sequencing facility using the same primers used for amplification, or on an ABI3730 × 1 at CCDB. Sequences from the University of Waikato were aligned using GENEIOUS, version 5.4.2, and confirmed as the target species using the Barcode of Life Data Systems (BOLD; www.boldsystems.org), version 3, COI animal identification searches. Previous analyses of Antarctic springtails (Stevens & Hogg, 2003) have shown that allozyme analyses were congruent with COI data and that the latter can be used as a reliable indicator of genomic differences occurring among populations. Primer sequences were trimmed from sequence fragments for further analyses. All sequences were uploaded to the BOLD dataset DS-SPMACK (Springtails of Mackay Glacier Region; dx.doi.org/10.5883/DS-SPMACK) and cross-referenced to GenBank (accession numbers KU876787 - KU876880).

Phylogenetic analysis

COI sequence fragments of 658 bp (219 codons) were obtained for 67 *G. hodgsoni* specimens and 20 *C. nivicolus* specimens. Approximately 560 bp were obtained from single direction reads (using primer LepF1) for eight *A. monoculata* specimens. No stop codons were detected. Sequences of *G. hodgsoni* were unambiguous at 658 bp (no insertions or deletions). However, sequences of *C. nivicolus* and *A. monoculata* contained ambiguous base-pair assignments that could not be easily resolved, and so the sequences were further trimmed at both ends, resulting in sequence fragments of 547 bp (181 codons) for *C. nivicolus* and 527 bp (175 codons) for *A. monoculata*. Two additional *C. nivicolus* sequences were also obtained from GenBank (Accession numbers DQ285403 and DQ285404).

Sequences for all species were initially examined in the context of generating a single Neighbour-joining tree using a Kimura two-parameter distance

model (Kimura, 1980). All duplicate sequences were identified and removed to include only unique haplotypes in subsequent analyses. As a result of the lack of publically available sequence data for taxa that share a recent common ancestor with our ingroup taxa (and that did not approach saturation), analyses were run unrooted among the ingroup taxa. No significant changes were noted in topography between these analyses and those run previously using *Podura aquatic*a as a test. Chi-squared tests as implemented in PAUP* 4.0 (Swofford, 2002) were used to determine whether the assumption of equal base frequencies among sites was violated on all sites and on third codon positions only. JMODELTEST, version 2.1.2 (Posada, 2008) was used to determine the most appropriate substitution model for maximum likelihood (ML) analysis. The settings were: 11 substitution schemes (88 models), base frequencies +F, rate variation +I, +Γ, set to BioNJ. The model selected for the dataset was GTR+I+Γ (–lnL = 1590.9). Maximum likelihood heuristic searches were conducted using this model in MEGA, version 5.10 (Tamura *et al.*, 2011) using 1000 bootstrap replicates. Maximum parsimony (MP) analyses were performed in PAUP* using 1000 full-heuristic search bootstrap replicates.

MrBayes, version 3.2.6 (Huelsenbeck & Ronquist, 2001) was used to conduct a Bayesian inference analysis. A general time reversal model (GTR+I+Γ) was used, with a log normal relaxed clock model and speciation Yule process as the tree prior. The Markov chain Monte Carlo (MCMC) procedure was set to 1 100 000 generations, sampling trees every 200. A burn-in of 100 000 trees was determined by plotting log-likelihood values against generation time in TRACER (Rambaut & Drummond, 2007) and checking for the point at which normalization occurred. The majority rule tree was acquired from the 11 004 trees sampled after the burn-in period. The tree was then visualized in TREEANNOTATOR (Drummond *et al.*, 2012).

Sequences for *G. hodgsoni* and *C. nivicolus* were split into separate datasets for analysis in TCS, version 1.21 (Clement, Posada & Crandall, 2000) and to construct networks of sequence haploytpes. Single representatives of each haplotype were used in the final analysis to simplify files, and sequences of *C. nivicolus* were trimmed at 547 bp to avoid anomalies, as described above. The *A. monoculata* sequences were not included in these analyses because they were only collected from a single site and consisted of only two similar haplotypes (< 0.2% divergence).

Uncorrected pairwise genetic distances between COI sequences for populations at different locations were also calculated for the *G. hodgsoni* and *C. nivicolus* datasets in MEGA, version 5.10. The likelihood ratio test did not detect evidence of significant rate heterogeneity for *G. hodgsoni* ($\chi^2 = 113.06$, d.f. = 14, $P < 0.001$) or *C. nivicolus* ($\chi^2 = 141.15$, d.f. = 10, $P < 0.001$). Approximate geological timing of isolation for the populations was estimated via molecular clock analyses in BEAST 1.8.2 (Drummond *et al.*, 2012). Files generated in BEAUti used a general time reversal model (GTR+I+Γ) with speciation Yule Processes as the tree prior and the same MCMC setup as that used for the Bayesian inference tree analysis. A strict clock model with a fixed rate of 0.0115 was used to simulate 2.3% sequence divergence per Myr, as determined using insect mitochondrial data (Brower, 1994; Juan, Oromi & Hewitt, 1996; Quek *et al.*, 2004; McGaughran *et al.*, 2010). Despite being calibrated for insects, the 2.3% sequence divergence per Myr was considered a suitable estimate for Collembola because both taxa have similar life cycles (McGaughran *et al.*, 2010).

RESULTS

Of the 658 bp analyzed for *G. hodgsoni*, 515 characters were constant, 22 were parsimony informative, and the remaining 121 were parsimony uninformative. The nucleotide composition averaged across all sequences showed an A-T bias of 64.0% (A = 27.7%, T = 36.7%, C = 19.3%, G = 16.7%). Nucleotide frequencies were not significantly different among sequences for all codon positions ($\chi^2 = 2.19$, d.f. = 48, $P = 1.0$) or for third codon positions only ($\chi^2 = 7.18$, d.f. = 48, $P = 1.0$). Of the 549 bp analyzed for *C. nivicolus*, 433 characters were constant, 22 were parsimony informative, and the remaining 94 were parsimony uninformative. The nucleotide composition averaged across all sequences showed an A-T bias of 61.4% (A = 25.8%, T = 35.6%, C = 20.4%, G = 18.2%). Base pair frequencies for *C. nivicolus* were not significantly different among sequences for all codon positions ($\chi^2 = 1.41$, d.f. = 36, $P = 1.0$) or for third codon positions only ($\chi^2 = 5.77$, d.f. = 36, $P = 1.0$). Of the 527 bp (175 codons) analyzed for *A. monoculata*, 408 characters were constant, one was parsimony informative, and the remaining 118 were parsimony uninformative. The nucleotide composition averaged across all sequences showed an A-T bias of 59.0% (A = 23.9%, T = 35.1%, C = 22.3%, G = 18.7%). Base pairs were not significantly different among sequences for all codon positions ($\chi^2 = 3.39$, d.f. = 21, $P = 1.0$) or for third codon positions only ($\chi^2 = 11.55$, d.f. = 21, $P = 0.95$).

Phylogenetic analysis

The ML tree is shown in Fig. 2. Tree constructions for MP (Fig. 3) and Neighbour-joining (data not

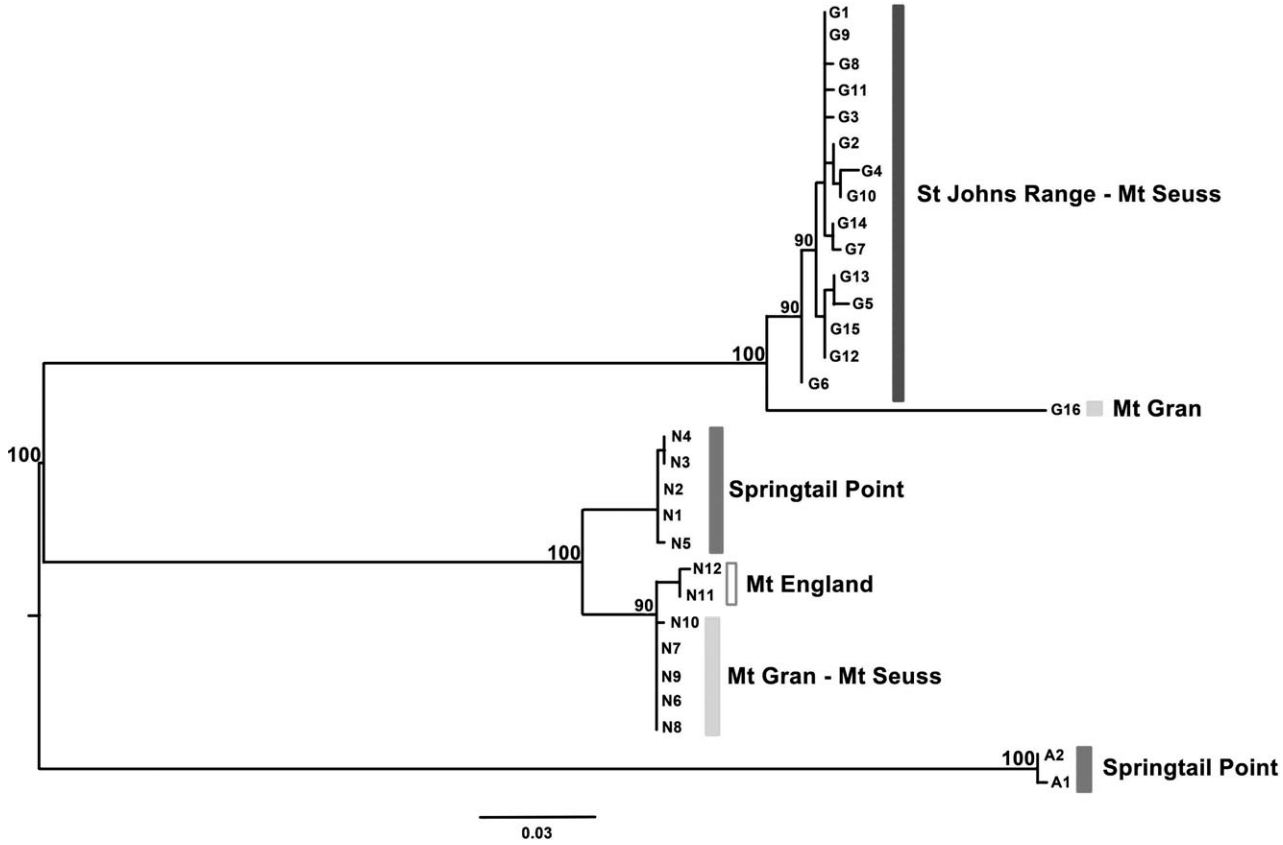

Figure 2. Maximum Likelihood phylogram constructed in MEGA, version 5.10, based on the GTR+I+Γ model derived from JMODELTEST, using 97 individual cytochrome *c* oxidase subunit I (COI) sequences reduced to unique haplotypes. Haplotype codes correspond to Table 1 and bootstrap values > 50 are shown. Tree is drawn to scale and branch lengths are the number of substitutions per site. Vertical bars indicate collection locations.

shown) showed similar topology and node support. Linking nodes between the haplotype G16 and the rest of the *G. hodgsoni* haplotypes had 100% bootstrap support in the ML and MP trees. The linking node between the *C. nivicolus* haplotypes at Springtail Point and at Mt Gran also received 100% bootstrap support in the ML and MP trees. Bootstrap values for the Mt England *C. nivicolus* haplotypes indicated high support from both the ML and MP trees. The topology of the *G. hodgsoni* haplotyes showed greater structure in the ML and BI trees. Two clusters were apparent in the ML tree, with node support of 99%. One cluster with the same haplotypes was apparent in the BI tree, with 0.55 posterior probability. Posterior probability values between *C. nivicolus* haplotypes at Springtail Point and at Mt Gran was 1.00, and also 1.00 between the Mt England and Mt Gran group (Fig. 4). The topology and node support of these trees supports the presence of high genetic structuring across the Mackay Glacier.

Haplotype networks

The geographical distribution of sequence haplotypes for *G. hodgsoni* and *C. nivicolus* was investigated using haplotype-joining networks. Subsequent haplotype assignments and their collection locations are shown in Table 1. Sixteen haplotypes were found from 67 *G. hodgsoni* sequences. Maximum connection steps were fixed at 40 to connect haplotype G16 to the rest of the haplotypes (Fig. 5). This network revealed 10 one-step haplotypes, three two-step haplotypes, two three-step haplotypes, and one 35-step haplotype. The most divergent haplotype shown by this analysis was G16, representing three individuals from Mt Gran. This difference was also supported by divergence values and phylogenetic trees (Figs 2, 3, 4). The remainder of the network, which included haplotypes from the St John's Range and Mt Seuss, did not show high geographical structure, similar to that observed in the tree-based approaches.

Twelve haplotypes were found from 22 *C. nivicolus* sequences. Maximum connection steps were fixed at

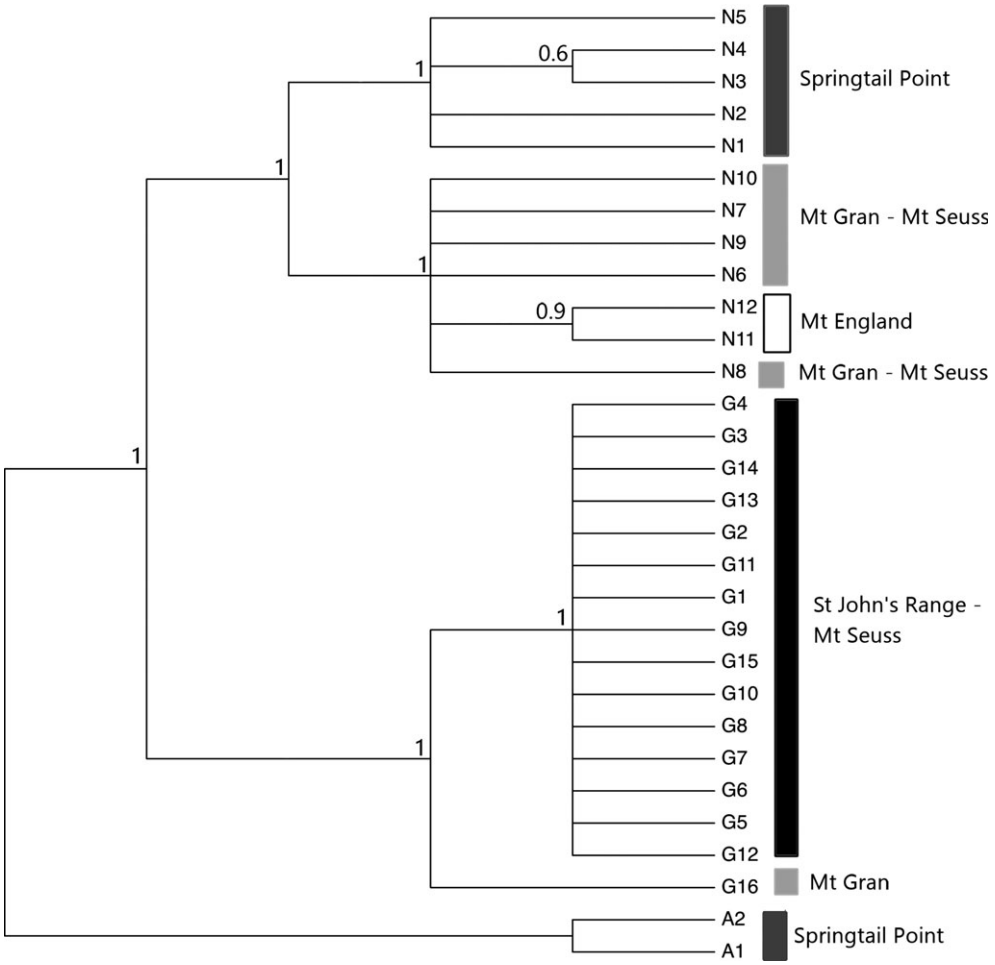

Figure 3. Maximum parsimony phylogram constructed in PAUP*, using 97 individual cytochrome *c* oxidase subunit I (COI) sequences reduced to unique haplotypes. Bootstrap values > 50 are shown. Tree is drawn to scale and branch lengths are the number of changes over the whole sequence. Haplotype codes correspond to Table 1 and vertical bars indicate collection locations.

30 to connect the Mt Gran and Mt England haplotypes to the Springtail point haplotypes (Fig. 6). This network revealed nine one-step haplotypes, two three-step haplotypes and one 16-step haplotype. This network analysis showed two groups of haplotypes that were connected by 16 missing mutational steps. These two groups corresponded to populations at Springtail Point on the south edge of Mackay Glacier, and Mt Gran and Mt Seuss to the north and in the centre of the glacier, respectively. This difference was supported by divergence values and phylogenetic trees. The two-step link to haplotypes at Mt England was also supported by divergence values and phylogenetic trees.

COI SEQUENCE DIVERGENCE AND MOLECULAR CLOCK ESTIMATES

Genetic distances ranged from 0.0% to 8% for *G. hodgsoni* and from 0.00% to 4.2% for *C. nivicolus* (Fig. 7). The greatest differences were found between haplotype G16 at Mt Gran and the remainder of the *G. hodgsoni* haplotypes, and the genetic distance between *C. nivicolus* haplotypes at Mt Gran and Mt England, as well as those at Springtail Point. The St John's Range and Mt Seuss *G. hodgsoni* haplotypes showed a mean divergence of 0.6% within the group (Fig. 7). The single haplotype, G16, at Mt Gran showed a mean of 7.6% sequence divergence from the other haplotypes.

The mean sequence divergences among *C. nivicolus* haplotypes within each location were 0.1% at Mt Gran, 0.2% at Springtail Point, and 0.2% at Mt England. Sequence divergences between locations showed the haplotypes at Mt Gran to be a mean of 4.0% divergent from haplotypes at Mt England. Similarly, Springtail Point haplotypes were a mean of 3.8% divergent from those found at Mt Gran. The Mt Gran and Mt England haplotypes were the most similar, with 0.8% sequence divergence between them.

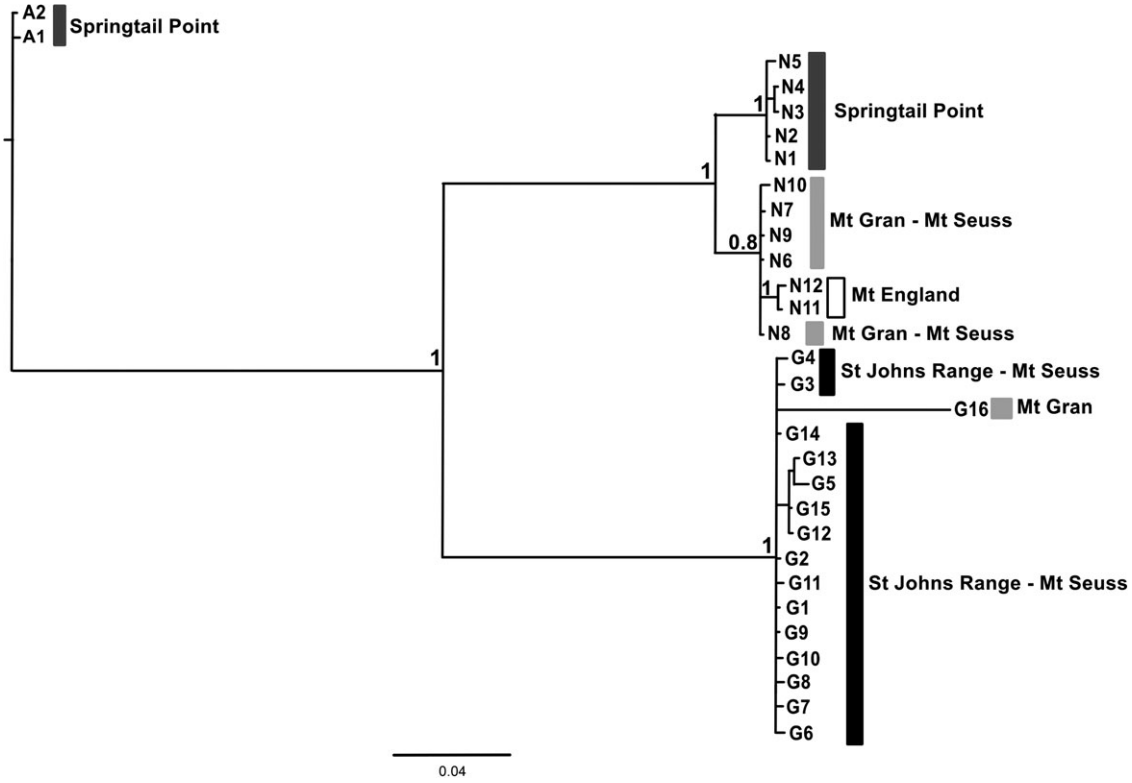

Figure 4. Bayesian inference phylogram constructed in MrBayes, version 3.2.6, based on the GTR+I+Γ model derived from JMODELTEST, using 97 individual cytochrome *c* oxidase subunit I (COI) sequences reduced to unique haplotypes. Posterior probabilities for haplotype group nodes are presented above 0.5. Tree is drawn to scale and branch lengths are measured in the number of changes per site. Haplotype codes correspond to Table 1 and vertical bars indicate collection locations.

Based on a strict molecular clock rate of 2.3% sequence divergence per Myr, these populations are all likely to have diverged within the last 4 Myr (Figs 7, 8). The oldest estimated divergence dated the genetic separation of *G. hodgsoni* haplotypes at Mt Gran (G16) and those in the St John's Range and at Mt Seuss at 3.8 Mya. Divergence dates between the three *C. nivicolus* populations suggested that the Springtail Point haplotypes diverged from the Mt Gran – Mt Seuss population 1.44 Mya. The difference between haplotypes from Mt Gran and Mt Seuss relative to those at Mt England is much more recent by comparison, estimated at 0.38 Mya.

DISCUSSION

Our mtDNA (COI) analysis of 97 Antarctic springtails from three taxonomic species revealed highly divergent populations across 65 km within the Mackay Glacier. Populations of *G. hodgsoni* and *C. nivicolus* on the lower slopes of Mt Gran were shown to be a mean of 7.6% and 3.8% divergent from their nearest neighbours. For *G. hodgsoni*, this represents a considerably greater genetic divergence among populations than the 2.4% divergence previously found for this species throughout the McMurdo Dry Valleys (Stevens & Hogg, 2003; Nolan *et al.*, 2006; McGaughran *et al.*, 2008). High genetic structure, within both putative species, suggests that populations may have survived *in situ* ever since the Antarctic continent became fully glaciated. Given the elevations of surrounding mountains, it is possible that several locations, such as Mt Gran (2235 m) and Mt Seuss (1190 m), protruded above the advancing Mackay Glacier, and remained so since the early Pliocene (Janetschek, 1967a; Clapperton & Sugden, 1990). In particular, this area is known to contain the highest species diversity of springtails in southern Victoria Land, with *G. hodgsoni*, *C. nivicolus*, and *A. monoculata* all known from this area (Gressitt *et al.*, 1963). The species diversity of mites, lichens, and mosses has also been shown to be high in the Mackay Glacier region relative to other nearby areas such as the Dry Valleys (Demetras *et al.*, 2010; Green *et al.*, 2011). This suggests that this area has served as a glacial refuge for multiple taxa during the last 5 Mya.

Table 1. Haplotypes, collection locations, coordinates, and sequences (BOLD Sample Id) associated with each haplotype for three species of Antarctic springtail

Haplotype number	Location	Co-ordinates (south – east)	Sample Id's
Gomphiocephalus hodgsoni			
G1	St John's Range	−77.280 161.731	ANTSP129 ANTSP131 ANTSP134 ANTSP136 ANTSP137 ANTSP138 ANTSP140 ANTSP141 ANTSP143 ANTSP151 ANTSP193
G2			ANTSP132 ANTSP133 ANTSP135 ANTSP139 ANTSP211 ANTSP212
G3		−77.208 161.700	ANTSP213 ANTSP215
G4		−77.285 161.726	ANTSP150
G5			ANTSP142
G6			ANTSP146
G7		−77.208 161.700	ANTSP209
G8			ANTSP210
G9			ANTSP216
G10		−77.285 161.726	ANTSP217
G11		−77.280 161.731	ANTSP144 ANTSP145 ANTSP147 ANTSP148 ANTSP149 ANTSP191
G11	Mt Seuss	−77.034 161.731	ANTSP128 ANTSP192 ANTSP207 ANTSP214 ANTSP218 ANTSP219
G12			ANTSP152 ANTSP154 ANTSP157 ANTSP158 ANTSP159 ANTSP160 ANTSP163 ANTSP164 ANTSP165 ANTSP168 ANTSP169 ANTSP172 ANTSP174 ANTSP175 ANTSP220 ANTSP221 ANTSP222 ANTSP223 ANTSP224 ANTSP225
G13			ANTSP162 ANTSP173
G14			ANTSP153 ANTSP166 ANTSP167
G15		−77.034 161.731	ANTSP161
G16	Mt Gran	−76.966 161.179	ANTSP200 ANTSP201 ANTSP202
Cryptopygus nivicolus			
N1	Springtail Point	−77.167 160.710	ANTSP119 ANTSP121 ANTSP188 ANTSP190 ANTSP230
N2			ANTSP228 ANTSP234
N3			ANTSP226 ANTSP231
N4			ANTSP227
N5			ANTSP118
N6	Mt Gran	−76.966 161.179	ANTSP233
N7			ANTSP197 ANTSP199
N8			ANTSP156
N8	Mt Seuss	−77.034 161.731	ANTSP124
N9			ANTSP155
N10			ANTSP170
N11	Mt England	−77.046 162.450	DQ285403
N12			DQ285404
Antarcticinella monoculata			
A1	Springtail Point	−77.168 160.710	ANTSP196 ANTSP235
A2			ANTSP194 ANTSP195 ANTSP203 ANTSP204 ANTSP205

Two Mt England *C. nivicolus* sequences (N11, N12) were retrieved from GenBank.

We now also highlight the potential for species-level genetic divergences within two springtail taxa for populations on opposite sides of the Mackay Glacier, which may indicate early stages of speciation. Our data suggest that the population of *G. hodgsoni* present on the lower slopes of Mt Gran has been isolated from other known *G. hodgsoni* populations since the Mid-Pliocene (4 Mya). Similarly, the population of *C. nivicolus* from the same location has been isolated from a neighbouring population at Springtail Point by as much as 1.4 Mya. The occurrence of *A. monoculata* at Springtail Point, coupled

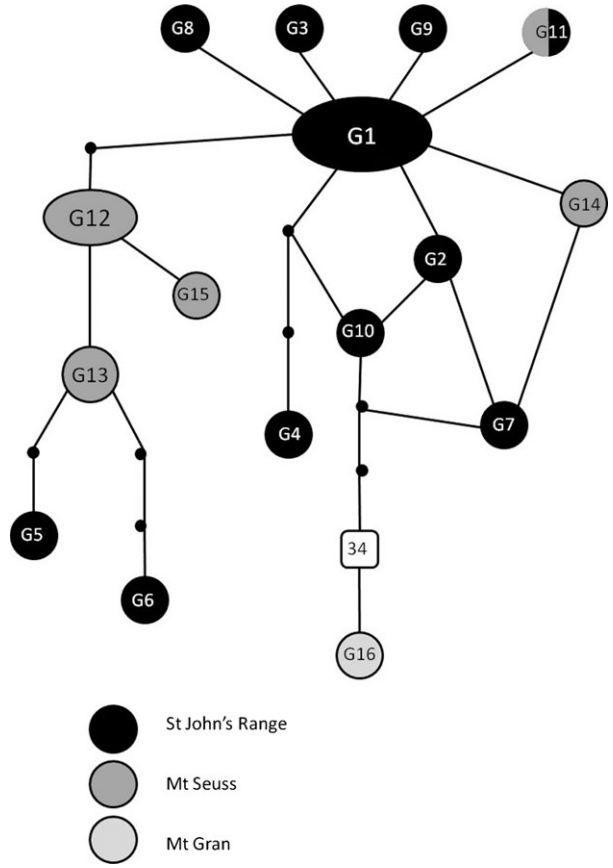

Figure 5. Haplotype network analysis for 16 haplotypes from 67 individuals of *Gomphiocephalus hodgsoni.* Haplotypes are indicated by their codes as referred to in Table 1. Missing haplotypes or mutational steps are indicated by black dots or are collapsed into a count of missing steps as in the single white square.

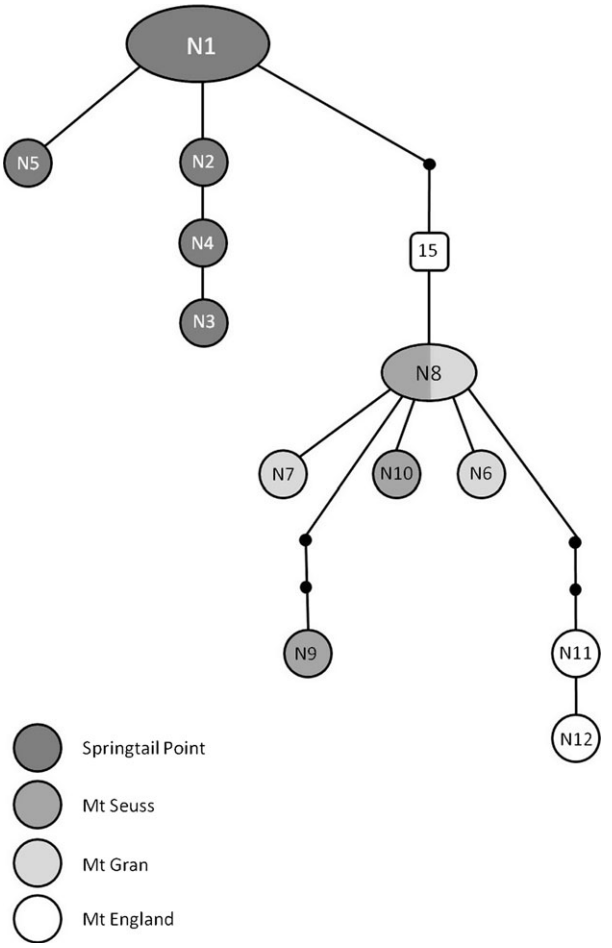

Figure 6. Haplotype network analysis for 12 haplotypes from 22 individuals of *Cryptopygus nivicolus.* Haplotypes are indicated by their codes as referred to in Table 1. Missing haplotypes or mutational steps are indicated by black dots or are collapsed into a count of missing steps as in the white square.

with the highly divergent populations at Mt Gran, supports the notion of high arthropod diversity for this area.

The differences in divergence estimates for *G. hodgsoni* (3.8 Mya) and *C. nivicolus* (1.4 Mya) may be the result of different evolutionary histories (e.g. later isolation) or possibly differences in mutation rates. For example, Stevens & Hogg (2006) suggested that differing mutation rates may exist between *G. hodgsoni* and the mite *Stereotydeus mollis.* However, little is known about the life history of *C. nivicolus.* The lack of ecological knowledge for *C. nivicolus* also makes it difficult to predict its dispersal abilities. Dispersal events in Antarctica are likely to be rare, and often accidental, making it difficult to attribute the presence of a species to ecological gradients (Janetschek, 1967b; Magalhães *et al.*, 2012). *Gomphiocephalus hodgsoni* is known to survive floating on both sea and fresh water, and dispersal events through wind or accidental carriage by birds are also possible (Stevens & Hogg, 2002; Hawes, 2011; McGaughran *et al.*, 2011a,b).

Many of the alpine glaciers underwent significant retreat during the interglacial periods of the Pleistocene (Clapperton & Sugden, 1990; Sugden *et al.*, 1999), although Mackay Glacier is unlikely to have done so because it is an outlet glacier for the EAIS. This appears to have isolated the Mt Gran population of *G. hodgsoni* from the populations on Mt Seuss in the centre of the glacier, as well as those in the St John's Range bordering Victoria Valley. It is possible that the presence of haplotypes from the St John's range in the Mt Seuss population relate to recent dispersal since the last glacial maximum. The sharing of *C. nivicolus* haplotypes between Mt Gran and Mt Seuss also indicates the potentially recent

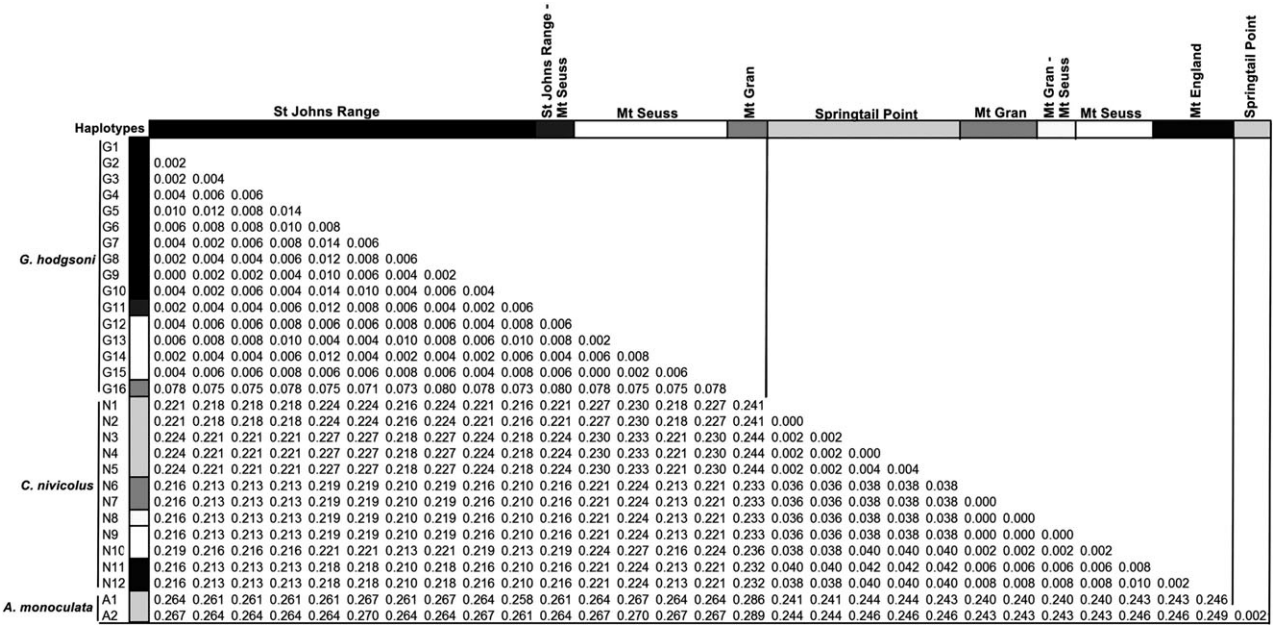

Figure 7. Genetic distances based on mitochondrial cytochrome *c* oxidase subunit I (COI) sequences of 97 springtails covering 30 unique haplotypes. Haplotype codes refer to those in Table 1. Collection locations for each haplotype are indicated in the bar at the top and side of the table.

dispersal from Mt Gran across the glacier. Hawes (2011) suggested that potential dispersal mechanisms may work in concert, where individuals could be wind-blown onto glaciers and then moved by glacial surface streams. The stochastic nature of dispersal events in Antarctica may explain why *G. hodgsoni* has yet to disperse from the Mt Gran population.

One species, *A. monoculata*, was found at only one location in our study area, although another isolated

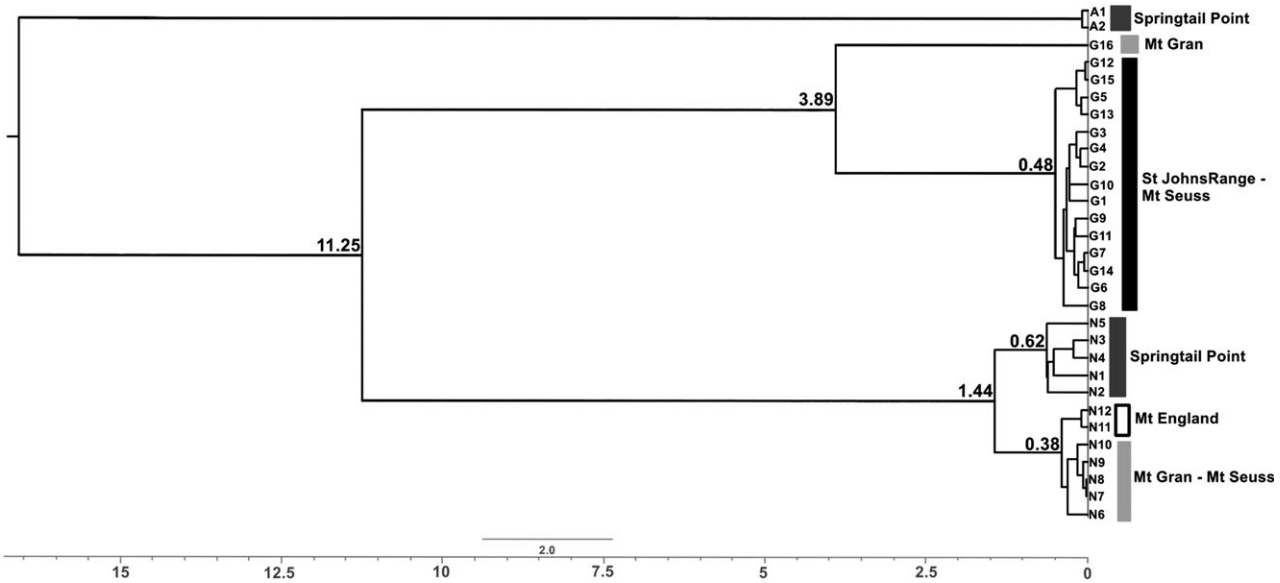

Figure 8. Estimated divergence times for populations of *Gomphiocephalus hodgsoni* and *Cryptopygus nivicolus*. The timeline on the x-axis is in millions of years. Values at nodes are estimated divergence times in millions of years. Haplotype codes correspond to Table 1 and vertical bars indicate locations of haplotypes.

population is known from Mt Murray 150 km to the north (Gressitt *et al.*, 1963). Similarly, haplotypes of *C. nivicolus* present at this site were not found elsewhere in our study area. Springtail Point is in an 'upglacier' position, making dispersal through temporary melt water to more seaward locations possible. However, there was no evidence of water courses being formed by temporary streams in this area, and visual assessment of snow banks that surround the site indicates they have changed little subsequent to a previous visit (Gressitt *et al.*, 1963). Even with surface water, the dispersal mechanisms used by other springtail species such as wind and stream flow may be limited for *A. monoculata*. The loss of pigmentation, limited tolerance of ultraviloet light, and a presence deeper in the soil profile (Janetschek, 1967a) make it less likely that *A. monoculata* would experience accidental dispersal by water or wind movement.

We conclude that the Mackay Glacier has provided a sufficient dispersal barrier to promote and maintain high levels of genetic divergence in two Antarctic springtail species endemic to southern Victoria Land. This isolation likely occurred around the early Pliocene (4 Mya) and has been maintained by on-going glaciations during the Pleistocene. The high genetic diversity, both at the population and species level, suggests that high altitude sites in this region have served as glacial refugia over the past 4 Mya. The isolation of these sites highlights the potential for high genetic diversity to be maintained on a small scale among the fragmented ice-free areas of Antarctica. Accordingly, we suggest that conservation efforts be directed toward maintaining and protecting the integrity of highly fragmented landscapes within the Transantarctic Mountains of the Ross Dependency.

ACKNOWLEDGEMENTS

We are grateful to three anonymous reviewers for their thoughtful and constructive suggestions that improved the quality of the manuscript. We thank M. Knox, C. Beet, U. Nielsen, D. Wall, and D. McKnight (US Antarctic LTER Programme) for helpful advice and/or assistance in the field/laboratory, as well as members of the Pacific Biosystematics Research group (University of Waikato) for discussion and comments during manuscript preparation. Glen Stichbury provided help preparing Fig. 1, which was derived from the SCAR Antarctic Digital Database and the Landsat Image Mosaic of Antarctica (LIMA) project. Financial support was provided to KRB through a New Zealand Post scholarship administered by Antarctica New Zealand, a University of Waikato Masters Research Scholarship, and The International Centre for Terrestrial Antarctic Research (ICTAR) Young Investigator Award. Field work was supported by Antarctica New Zealand and the US National Science Foundation through the McMurdo LTER NSF OPP grant 1115245. Sequencing at the Biodiversity Institute of Ontario was supported by funding from the Government of Canada through Genome Canada and the Ontario Genomics Institute.

AUTHOR CONTRIBUTIONS

IDH, KRB, BJA and PDNH conceived the research and obtained funding. KRB and IDH conducted the field work and KRB conducted the primary analyses and was lead author of the manuscript in conjunction with IDH BJA and PDNH. All authors reviewed and contributed revisions to the final version of the manuscript.

REFERENCES

Adams B, Bardgett R, Ayres E, Wall DH, Aislabie J, Bamforth S, Bargagli R, Cary C, Cavacini P, Connell L, Convey P, Fell JW, Frati F, Hogg ID, Newsham KK, O'Donnell A, Russell N, Seppelt RD, Stevens MI. 2006. Diversity and distribution of Victoria Land biota. *Soil Biology and Biochemistry* **38:** 3003–3018.

Adhikari BN, Wall DH, Adams BJ. 2010. Effect of slow desiccation and freezing on gene transcription and stress survival of an Antarctic nematode. *Journal of Experimental Biology* **213:** 1803–1812.

Brower A. 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 United States of America* **91:** 6491–6495.

Clapperton CM, Sugden D. 1990. Late Cenozoic glacial history of the Ross Embayment, Antarctica. *Quaternary Science Reviews* **9:** 253–272.

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

Convey P, Stevens MI, Hodgson DA, Smellie JL, Hillenbrand CD, Barnes DKA, Clarke A, Pugh PJA, Linse K, Cary SC. 2009. Exploring biological constraints on the glacial history of Antarctica. *Quaternary Science Reviews* **28:** 3035–3048.

Demetras NJ, Hogg ID, Banks JC, Adams BJ. 2010. Latitudinal distribution and mitochondrial DNA (COI) variability of *Stereotydeus* spp. (Acari: Prostigmata) in Victoria Land and the central Transantarctic Mountains. *Antarctic Science* **22:** 749–756.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29:** 1969–1973.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* **3:** 294–299.

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:** 934–940.

Green TGA, Sancho LG, Türk R, Seppelt RD, Hogg ID. 2011. High diversity of lichens at 84°S, Queen Maud Mountains, suggests preglacial survival of species in the Ross Sea region, Antarctica. *Polar Biology* **34:** 1211–1220.

Greenslade P. 2015. Synonymy of two monobasic Anurophorinae genera (Collembola: Isotomidae) from the Antartctic Continent. *New Zealand Entomologist* **38:** 134–141.

Gressitt J. 1967. The Fauna. In: Bushnell V, ed. *Terrestrial life of Antarctica*. Antarctic Map Folio Series, New York, NY: American Geographical Society, 17–24.

Gressitt J, Leech R, Wise KAJ. 1963. Entomological investigations in Antarctica. *Pacific Insects* **5:** 287–304.

Hajibabaei M, Janzen DH, Burns JM, Hallwachs W, Hebert PDN. 2006. DNA barcodes distinguish species of tropical Lepidoptera. *Proceedings of the National Academy of Sciences of the United States of America* **103:** 968–971.

Hawes TC. 2011. Rafting in the Antarctic springtail, *Gomphiocephalus hodgsoni*. *Antarctic Science* **23:** 456–460.

Hawes TC, Torricelli G, Stevens MI. 2010. Haplotype diversity in the Antarctic springtail *Gressittacantha terranova* at fine spatial scales – a Holocene twist to a Pliocene tale. *Antarctic Science* **22:** 766–773.

Hogg ID, Stevens MI. 2002. Soil fauna of antarctic coastal landscapes. *Ecological Studies* **154:** 265–282.

Hogg ID, Wall DH. 2012. Extreme Habitats: polar Deserts. In: E Bell, ed. *Life at extremes: environments, organisms and strategies for survival*. CAB International, Cambridge, 176–195.

Hogg ID, Cary CS, Convey P, Newsham KK, O'Donnell AG, Adams BJ, Aislabie J, Frati F, Stevens MI, Wall DH. 2006. Biotic interactions in Antarctic terrestrial ecosystems: are they a factor? *Soil Biology and Biochemistry* **38:** 3035–3040.

Huelsenbeck J, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* **17:** 754–755.

Ivanova NV, deWaard JR, Hebert PDN. 2006. An inexpensive, automation-friendly protocol for recovering high-quality DNA. *Molecular Ecology Notes* **6:** 998–1002.

Janetschek H. 1967a. Arthropod ecology of south Victoria Land. *Antarctic Research Series* **10:** 205–293.

Janetschek H. 1967b. Growth and maturity of the springtail *Gomphiocephalus hodgsoni* Carpenter, from south Victoria Land and Ross Island. *Antarctic Research Series* **10:** 295–305.

Juan C, Oromi P, Hewitt GM. 1996. Phylogeny of the genus *Hegeter* (Tenebrionidae, Coleoptera) and its colonization of the Canary Islands deduced from cytochrome oxidase I mitochondrial DNA sequences. *Heredity* **76:** 392–403.

Kimura M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* **16:** 111–120.

Levy J. 2012. How big are the McMurdo Dry Valleys? Estimating ice-free area using Landsat image data. *Antarctic Science* **25:** 119–120.

Lewis A, Marchant D, Ashworth A, Hemming S, Machlus M. 2007. Major middle Miocene global climate change: evidence from East Antarctica and the Transantarctic Mountains. *Geological Society of America Bulletin* **119:** 1449–1461.

Magalhães C, Stevens MI, Cary SC, Ball BA, Storey BC, Wall DH, Türk R, Ruprecht U. 2012. At limits of life: multidisciplinary insights reveal environmental constraints on biotic diversity in continental Antarctica. *PLoS ONE* **7:** e44578.

McGaughran A, Hogg ID, Stevens MI. 2008. Patterns of population genetic structure for springtails and mites in southern Victoria Land, Antarctica. *Molecular Phylogenetics and Evolution* **46:** 606–618.

McGaughran A, Torricelli G, Carapelli A, Frati F, Stevens MI, Convey P, Hogg ID. 2010. Contrasting phylogeographical patterns for springtails reflect different evolutionary histories between the Antarctic Peninsula and continental Antarctica. *Journal of Biogeography* **37:** 103–119.

McGaughran A, Hogg ID, Convey P. 2011a. Extended ecophysiological analysis of *Gomphiocephalus hodgsoni* (Collembola): flexibility in life history strategy and population response. *Polar Biology* **34:** 1713–1725.

McGaughran A, Stevens MI, Hogg ID, Carapelli A. 2011b. Extreme glacial legacies: a synthesis of the Antarctic springtail phylogeographic record. *Insects* **2:** 62–82.

Nkem J, Wall D, Virginia R, Barrett JE, Broos E, Porazinska DL, Adams B. 2006. Wind dispersal of soil invertebrates in the McMurdo Dry Valleys, Antarctica. *Polar Biology* **29:** 346–352.

Nolan L, Hogg ID, Stevens MI, Haase M. 2006. Fine scale distribution of mtDNA haplotypes for the springtail *Gomphiocephalus hodgsoni* (Collembola) corresponds to an ancient shoreline in Taylor Valley, continental Antarctica. *Polar Biology* **29:** 813–819.

Posada D. 2008. jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution* **25:** 1253–1256.

Quek S, Davies S, Ilino T, Pierce N. 2004. Codiversification in an ant-plant mutualism: stem testure and the evolution of host use in *Crematogaster* (Formicidae: Myrmicinae) inhabitants of *Macaranga* (Euphorbiaceae). *Evolution* **58:** 554–570.

Rambaut A, Drummond A. 2007. Tracer, Version 1.4. Available at: http://beast.bio.ed.ac.uk/Tracer.

Salmon JT. 1965. An index to the Collembola. *Bulletin, Royal Society of New Zealand* **7**, 645–651.

Sinclair BJ, Stevens MI. 2006. Terrestrial microarthropods of Victoria Land and Queen Maud Mountains, Antarctica: implications of climate change. *Soil Biology and Biochemistry* **38:** 3158–3170.

Stevens MI, D'Haese C. 2014. Islands in ice: isolated populations of *Cryptopygus sverdrupi* (Collembola) among nunataks in the Sor Rondane Mountains, Dronning Maud Land, Antarctica. *Biodiversity* **15:** 169–177.

Stevens MI, Hogg ID. 2002. Expanded distributional records of Collembola and Acari in Southern Victoria Land, Antarctica. *Pedobiologia* **46:** 485–495.

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, Hogg ID. 2006. Contrasting levels of mitochondrial DNA variability between mites (Penthalodidae) and springtails (Hypogastruridae) from the Trans-Antarctic Mountains suggest long-term effects of glaciation and life history on substitution rates, and speciation processes. *Soil Biology and Biochemistry* **38:** 3171–3180.

Stevens MI, Greenslade P, Hogg ID, Sunnucks P. 2006. Southern hemisphere springtails: could any have survived glaciation in Antarctica? *Molecular Biology and Evolution* **23:** 874–882.

Sugden DE, Summerfield MA, Denton GH, Wilch TI, McIntosh WC, Marchant DR, Rutford RH. 1999. Landscape development in the Royal Society Range, southern Victoria Land, Antarctica: stability since the mid-Miocene. *Geomorphology* **28:** 181–200.

Suggate RP. 1990. Late Pliocene and Quaternary glaciations of New Zealand. *Quaternary Science Reviews* **9:** 175–197.

Swofford D. 2002. PAUP*: Phylogenetic Analysis Using Parsimony, Version 4.0b10 for Macintosh.Sunderland, MA: Sinauer Associates.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* **28:** 2731–2739.

Wise KAJ. 1971. The Collembola of antarctica. *Pacific Insects Monograph* **25:** 57–74.