Original Article.

High levels of intra-specific 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], 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

Running Head: Genetic divergences among Antarctic Springtails

Correspondence:

Ian Hogg

School of Science

University of W*aikato*

Private Bag 3105

Hamilton 3240, New Zealand

email: hogg@waikato.ac.nz

Abstract word count: 233. Whole text word count: 6,631.

Abstract

Aim To examine the levels of genetic variability within and among populations of three Antarctic springtail species (Arthropoda: Collembola) and test the hypothesis that genetic divergences occur among glacially-isolated habitats.

Location Southern Victoria Land, Ross Dependency, Antarctica.

Methods Samples were collected from locations in the vicinity of the Mackay Glacier. We analysed mitochondrial DNA (COI) sequence variability for 97 individuals representing three species (*Gomphiocephalus hodgsoni* n=67; *Cryptopygus nivicolus*, n=20; *Antarcticinella monoculata*, n=8). Haplotype diversity and genetic divergences were calculated and used to indicate population variability as well as infer divergence times of isolated populations using molecular clock estimates.

Results 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 *Cryptopygus nivicolus* species showed 4.0% divergence among populations. The third species, *Antarcticinella monoculata*, was found in only one location. Molecular clock estimates based on sequence divergences suggest that populations separated within the last 4 Mya.

Main Conclusions Habitat fragmentation resulting from Pliocene (5 Mya) and Pleistocene (2 Mya - 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 has 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.

Key Words Collembola, glaciation, Ross Sea region, population genetics, springtails, refugia.

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 comprised of 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 allow survival in sub-zero temperatures (McGaughran *et al.*, 2011a).

The terrestrial arthropods are represented primarily by springtails (Collembola) and mites (Acari) and are the largest year-round taxa on the continent (Gressitt, 1967; Hogg & Stevens, 2002; Adams *et al*., 2006). These taxa, which lack survival and dispersal strategies possessed by other invertebrate groups such as nematodes (Adhikari *et* al.,2010; Nkem *et al*., 2006), have been restricted to these fragmented, ice-free zones since the Middle Miocene (14-11 Mya; Stevens & Hogg, 2003; Stevens *et al.*, 2006; McGaughran *et al.*, 2010). At this time, 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). Since then, 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 at most one or two arthropod taxa (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, once 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 *et al*., 2001; Stevens & Hogg, 2003; McGaughran *et al*., 2008; Hawes *et al*., 2010; Stevens & D'Haese, 2014). Here, our aim was to extend these studies by focussing on 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 our study, consist of three species covering a 3^{o} 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 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).

In order to determine the geographic 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.

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 lab where they were suspended a 10% sucrose solution. Invertebrates were then removed from the solution surface under a dissecting microscope (10X 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 *et al*., (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 analyses:

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 *et al*., 2006) at CCDB, and Red Extract n Amp (Sigma-Aldrich) using 10 µl extraction solution and 2.5 µl tissue prep, following manufacture's protocol. Polymerase Chain Reactions (PCRs) were comprised of a 15 µl reaction containing 5.7 µl MQH_20, 7.5 µl 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 COI gene was amplified using the primers HCO2198 (sequence 5'-TAAACTTCAGGGTGACCAAAAAATCA-3') and an altered LCO1490 (sequence: 5'-AGTTCTAATCATTAARGATATYGG-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'-GGTCAACAAATCATAAAGATATTGG-3') was used for both species (in place of the altered LCO1490 and LepF1) at CCDB. Primers were used at 1.0 mM concentration. PCR conditions at CCDB were: initial denaturing at 94ºC for 1 min; 5 cycles of 94ºC for 1 min, 45ºC for 1.5 min and 72ºC for 1.5 min; 35 cycles of 94ºC for 1 min, 50ºC for 1.5 min and 72ºC for 1 min followed by a final 72ºC for 5 min. PCR conditions were: initial denaturing at 94ºC for 5 minutes; 36 cycles of 94ºC for 1 min, 52ºC for 1.5 min and 72ºC for 1 min, followed by a final 72ºC for 5 min.

PCR products were cleaned using Sephadex (CCDB) or 0.2 µl ExonucleaseI (EXO) and 0.1 µl Shrimp Alkaline Phosphate (SAP) with 2.7 µl MQH_20 following manufactures protocol (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 ABI3730x1 at CCDB. Sequences from the University of Waikato were aligned using Geneious, ver 5.4.2, and confirmed as the target species using the Barcode of Life DataSystems (BOLD; www.boldsystems.org) ver 3 COI

157 animal identification searches. Previous analyses of Antarctic springtails (e.g. Stevens &
158 Hogg 2003), have shown that allozyme analyses were congruent with COI data and that the
159 latter can be used as a reliable indicator of genomic differences occurring among populations.
160 Primer sequences were trimmed from sequence fragments for further analyses. All sequences
161 were uploaded to the BOLD project Antarctic Terrestrial Arthropods (ANTSP) and cross-
162 referenced to GenBank.

163 *Phylogenetic Analysis*

164 COI sequence fragments of 658 bp (219 codons) were obtained for 67 *G. hodgsoni* specimens
165 and 20 *C. nivicolus* specimens. Approximately 560 bp were obtained from single direction
166 reads (using primer LepF1) for eight *A. monoculata* specimens. No stop codons were
167 detected. Sequences of *G. hodgsoni* were unambiguous at 658 bp (no insertions or deletions).
168 However, sequences of *C. nivicolus* and *A. monoculata* contained ambiguous base pair
169 assignments which could not be easily resolved, so sequences were further trimmed at both
170 ends, resulting in sequence fragments of 547 bp (181 codons) for *C. nivicolus* and 527 bp
171 (175 codons) for *A. monoculata*. Two additional *C. nivicolus* sequences were also obtained
172 from GenBank (Accession numbers DQ285403 and DQ285404).

173 Sequences for all species were initially examined in the context of generating a single
174 neighbour-joining tree using a Kimura 2-parameter distance model (Kimura, 1980). All
175 duplicate sequences were identified and removed to include only unique haplotypes in
176 subsequent analyses. Due to the lack of publically available sequence data for taxa that share
177 a recent common ancestor with our ingroup taxa (and that did not approach saturation),
178 analyses were run unrooted among the ingroup taxa. No significant changes were noted in
179 topography between these analyses and ones run previously using *Podura aquatic* as a test.
180 Chi-square tests (X^2) 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. JModel test 2.1.2 (Posada, 2008) was used to determine the most appropriate substitution model for Maximum Likelihood (ML) analysis. Settings were as follows: 11 substitution schemes (88 models), base frequencies +F, rate variation +I, +Γ, set to BioNJ. The model selected for the data set was GTR + I +Γ (-lnL = 1,590.9). Maximum Likelihood heuristic searches were conducted using this model in MEGA 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 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) 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 Tree Annotator (Drummond *et al.*, 2012).

Sequences for *G. hodgsoni* and *C. nivicolus* were split into separate data sets for analysis in the program TCS 1.21 (Clement *et al*., 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 as they were only collected from a single site and consisted of only two similar haplotypes (<0.2% divergence).

Uncorrected pair-wise genetic distances between COI sequences for populations at different locations were also calculated for the *G. hodgsoni* and *C. nivicolus* data sets in MEGA 5.10. The likelihood ratio test did not detect evidence of significant rate heterogeneity for *G. hodgsoni* (X^2=113.06; p<0.001; d.f=14) or *C. nivicolus* (X^2=141.15; p<0.001; d.f=10). Approximate geological timing of isolation for the populations was estimated through 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 set up as used for the BI tree analysis. A strict clock model with a fixed rate of 0.0115 was used to simulate 2.3% sequence divergence per million years, as determined using insect mitochondrial data (Brower 1994; Juan *et al.*, 1996; Quek *et al.*, 2004; McGaughran *et al.*, 2010). Despite being calibrated for insects, the 2.3% sequence divergence per million years was considered a suitable estimate for Collembola as both taxa have similar life cycles (McGaughran *et al.*, 2010).

Results

Of the 658 bp analysed 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 (X^2 = 2.19, p = 1.0, d.f = 48) or for third codon positions only (X^2 = 7.18, p = 1.0, d.f = 48). Of the 549 bp analysed 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 (X^2

= 1.41, p = 1.0, d.f = 36) or for third codon positions only (X^2 = 5.77, p =1.0, d.f = 36). Of the 527 bp (175 codons) analysed for *A. monoculata,* 408 characters were constant, 1 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 (X^2 = 3.39, p = 1.0, d.f = 21) or for third codon positions only (X^2 = 11.55, p = 0.95, d.f = 21).

Phylogenetic Analysis

A Maximum Likelihood (ML) tree is shown in Fig. 2. Tree constructions for Maximum Parsimony (Fig.3) and Neighbour Joining (data not 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 for the *G. hodgsoni* haplotypes differed in the ML from both the MP and BI trees. Two clusters were apparent, with 0.99 bootstrap support for the node. Collection locations of haplotypes were mixed between both clusters. The topology of the BI tree was also similar to all other trees for *G. hodgsoni*, *C. nivicolus* and *A. monoculata.* 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 geographic 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 in order to connect haplotype G16 to the rest of the haplotypes (Fig. 5). This network revealed 10 1-step haplotypes, three 2-step haplotypes, two 3-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, 7). The remainder of the network which included haplotypes from the St John's Range and Mt Seuss did not show high geographic 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 30 in order to connect the Mt Gran and Mt England haplotypes to the Springtail point haplotypes (Fig. 6). This network revealed nine 1-step haplotypes, two 3-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 2-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 0.0-8% for *G. hodgsoni* and 0.00-4.2% for *C. nivicolus* (Fig.7). 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, and those at Springtail Point. The St John's Range and Mt Seuss *G. hodgsoni* haplotypes showed an average divergence of 0.6% within the group (Fig.7). The single haplotype, G16, at Mt Gran showed an average of 7.6% sequence divergence from the other haplotypes.

The average 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 an average of 4.0% divergent from haplotypes at Mt England. Similarly, Springtail Point haplotypes were an average 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.

Based on a strict molecular clock rate of 2.3% sequence divergence per million years, these populations are all likely to have diverged within the last 4 My (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 mitochondrial DNA (COI) analysis of 97 Antarctic springtails from three taxonomic species revealed highly divergent populations across 65 km within the Mackay Glacier. Populations of *Gomphiocephalus hodgsoni* and *Cryptopygus nivicolus* on the lower slopes of Mt Gran were shown to be an average of 7.6% and 3.8% divergent from their nearest neighbours. For *G. hodgsoni*, this represents a considerably greater genetic divergence among

299 populations than the 2.4% divergence previously found for this species throughout the
300 McMurdo Dry Valleys (Stevens & Hogg, 2003; Nolan *et al*., 2006; McGaughran *et al*.,
301 2008). High genetic structure, within both putative species, suggests that populations may
302 have survived *in situ* since the Antarctic continent became fully glaciated. Given the
303 elevations of surrounding mountains it is possible that several locations such as Mt Gran
304 (2235 m) and Mt Seuss (1190 m) protruded above the advancing Mackay Glacier, and
305 remained so since the early Pliocene (Janetschek, 1967a; Clapperton & Sugden, 1990). In
306 particular, this area is known to contain the highest species diversity of springtails in southern
307 Victoria Land, with *G. hodgsoni, C. nivicolus* and *A. monoculata* all known from this area
308 (Gressitt *et al*., 1963). The species diversity of mites, lichens and mosses have also been
309 shown to be high in the Mackay Glacier region relative to other nearby areas such as the Dry
310 Valleys (Demetras *et al*., 2010; Green *et al*., 2011). This suggests that this area has served as
311 a glacial refuge for multiple taxa during the last 5 Mya.

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

320 The differences in divergence estimates for *G. hodgsoni* (3.8 Mya) and *C. nivicolus* (1.4
321 Mya) may be the result of different evolutionary histories (e.g. later isolation) or possibly
322 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). *G. hodgsoni* is known to survive floating on both sea and fresh water, and dispersal events through wind or accidental carriage by birds is also possible (Stevens & Hogg, 2002; Hawes, 2011; McGaughran *et al*., 2011a, 2011b).

As Mackay Glacier is an outlet glacier for the EAIS, it is unlikely to have undergone significant retreat during the interglacial periods of the Pleistocene as many of the alpine glaciers did (Clapperton & Sugden, 1990; Sugden *et al*., 1999). This appears to have isolated the Mt Gran population of *G. hodgsoni* from the populations on Mt Seuss in the centre of the glacier, and 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 potentially recent dispersal from Mt Gran across the glacier. Hawes, (2011) suggested that potential dispersal mechanisms may work in concert, whereby 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 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 'up-glacier' position, making dispersal through

temporary melt water to more sea-ward 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 indicate they have changed little since 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 UV light and 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 thank M. Knox, U. Nielsen, D. Wall, and D. McKnight (US Antarctic LTER Programme) for helpful advice and/or assistance in the field/lab, and members of the Pacific Biosystematics Research group (University of Waikato) for discussion and comments during manuscript preparation. Glen Stichbury provided help in 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.

Literature Cited

Adams B, Bardgett R., Ayres E, Wall DH, Aislabie J, Bamforth S, Bargagli R, Cary C, Cavacini P, Connell L. 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.

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

413 Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. 1994. DNA primers for amplification of
414 mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates.
415 *Molecular Marine Biology and Biotechnology* **3**: 294–9.

416 Frati F, Spinsanti G, Dallai R. 2001. Genetic variation of mtCOII gene sequences in the
417 collembolan *Isotoma klovstadi* from Victoria Land, Antarctica: evidence for population
418 differentiation. *Polar Biology* **24**: 934–940.

419 Green TGA, Sancho LG, Türk R, Seppelt RD, Hogg ID. 2011. High diversity of lichens at
420 84°S, Queen Maud Mountains, suggests preglacial survival of species in the Ross Sea
421 region, Antarctica. *Polar Biology*, **34**, 1211–1220.

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

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

426 Gressitt J. 1967. The Fauna. *Terrestrial life of Antarctica* (ed. by V. Bushnell), pp. 17-24.
427 Antarctic map folio series, New York.

428 Hajibabaei M, Janzen DH, Burns JM, Hallwachs W, Hebert PDN. 2006. DNA barcodes
429 distinguish species of tropical Lepidoptera. *Proceedings of the National Academy of*
430 *Sciences of the United States of America* **103**: 968–71.

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, Wall DH. 2012. Extreme Habitats: Polar Deserts. *Life at Extremes: Environments, Organisms and Strategies for Survival* (ed. by E. Bell). CAB International, Cambridge, UK.

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.

Hogg ID, Stevens MI. 2002. Soil Fauna of Antarctic Coastal Landscapes. *Ecological Studies* **154**: 265–282.

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

Ivanova NV, Deewaard 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.

450 Janetschek H. 1967b. Growth and maturity of the springtail *Gomphiocephalus hodgsoni*
451 Carpenter, from south Victoria Land and Ross Island. *Antarctic Research Series* **10**: 295
452 –305.

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

456 Kimura M. 1980. A simple method for estimating evolutionary rate of base substitutions
457 through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*
458 **16**: 111-120.

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

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

464 Magalhães C, Stevens MI, Cary SC, Ball BA, Storey BC, Wall DH, Türk R, Ruprecht U.
465 2012. At limits of life: multidisciplinary insights reveal environmental constraints on
466 biotic diversity in continental Antarctica. *PloS one* **7**: e44578.

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

470 McGaughran A, Stevens MI, Hogg ID, Carapelli A. 2011b. Extreme Glacial Legacies: A
471 Synthesis of the Antarctic Springtail Phylogeographic Record. *Insects* **2**: 62–82.

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**: 606–18.

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.

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**: 1253–6.

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 v1.4, Available from http://beast.bio.ed.ac.uk/Tracer.

Salmon JT. 1965. An index to the Collemnola. *Bulletin, Royal Society of New Zealand* **7**.

493 Sinclair BJ, Stevens MI. 2006. Terrestrial microarthropods of Victoria Land and Queen Maud
494 Mountains, Antarctica: Implications of climate change. *Soil Biology and Biochemistry*
495 **38**: 3158–3170.

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

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

501 Stevens MI, Hogg ID. 2003. Long-term isolation and recent range expansion from glacial
502 refugia revealed for the endemic springtail *Gomphiocephalus hodgsoni* from Victoria
503 Land, Antarctica. *Molecular Ecology* **12**: 2357–2369.

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

508 Stevens MI, Greenslade P, Hogg ID, Sunnucks P. 2006. Southern hemisphere springtails:
509 could any have survived glaciation in Antarctica? *Molecular Biology and Evolution* **23**:
510 874-882.

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

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

516 Swofford D. 2002. PAUP*: Phylogenetic Analysis Using Parsimony, version 4.0b10 for
517 Macintosh.

518 Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: Molecular
519 Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and
520 Maximum Parsimony Methods. *Molecular Biology and Evolution* **28**: 2731-2739.

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

522

523 **Biosketch**

524 KRB is an MSc graduate from the University of Waikato with interests in animal
525 conservation and connectivity among natural populations. Her primary target taxa are
526 terrestrial Collembola and freshwater macroinvertebrates. She is also interested in the
527 evolution of the New Zealand and Antarctic landscapes.

528 Author contributions:

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

533

Table 1: Haplotypes, collection locations, coordinates and sequences (BOLD Process Id) associated with each haplotype for three species of Antarctic springtail. Two Mt England *C. nivicolus* sequences (N11, N12) were retrieved from GenBank.

Haplotype #	Location	Co-ordinates (south – east)	Process Id's
G. hodgsoni			
G1	St John's Range	-77.280 161.731	ANTSP131 ANTSP134 ANTSP136 ANTSP137 ANTSP138 ANTSP140 ANTSP141 ANTSP143 ANTSP129 ANTSP193 ANTSP151
G2			ANTSP133 ANTSP135 ANTSP139 ANTSP211 ANTSP212 ANTSP132
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	Mt Seuss	-77.280 161.731 -77.034 161.731	ANTSP144 ANTSP145 ANTSP147 ANTSP148 ANTSP149 ANTSP191 ANTSP192 ANTSP207 ANTSP219 ANTSP214 ANTSP128 ANTSP218
G12			ANTSP154 ANTSP157 ANTSP158 ANTSP159 ANTSP160 ANTSP163 ANTSP164 ANTSP165 ANTSP168 ANTSP169 ANTSP172 ANTSP174 ANTSP175 ANTSP220 ANTSP222 ANTSP221 ANTSP223 ANTSP224 ANTSP152 ANTSP225
G13			ANTSP162 ANTSP173

G14				ANTSP166 ANTSP153 ANTSP167
G15		77.034	161.731	ANTSP161
G16	Mt Gran	-76.966	161.179	ANTSP201 ANTSP202 ANTSP200

C. nivicolus

N1	Springtail Point	-77.167	160.710	ANTSP121 ANTSP188 ANTSP190 ANTSP230 ANTSP119
N2				ANTSP2234 ANTSP228
N3				ANTSP231 ANTSP226
N4				ANTSP227
N5				ANTSP118
N6	Mt Gran	-76.966	161.179	ANTSP233
N7				ANTSP199 ANTSP197
N8	Mt Seuss	-77.034	161.731	ANTSP156 ANTSP124
N9				ANTSP155
N10				ANTSP170
N11	M England	-77.046	162.450	DQ285403
N12				DQ285404

A. monoculata

A1	Springtail Point	-77.168	160.710	ANTSP196 ANTSP235
A2				ANTSP204 ANTSP205 ANTSP194 ANTSP195 ANTSP203

List of Figures

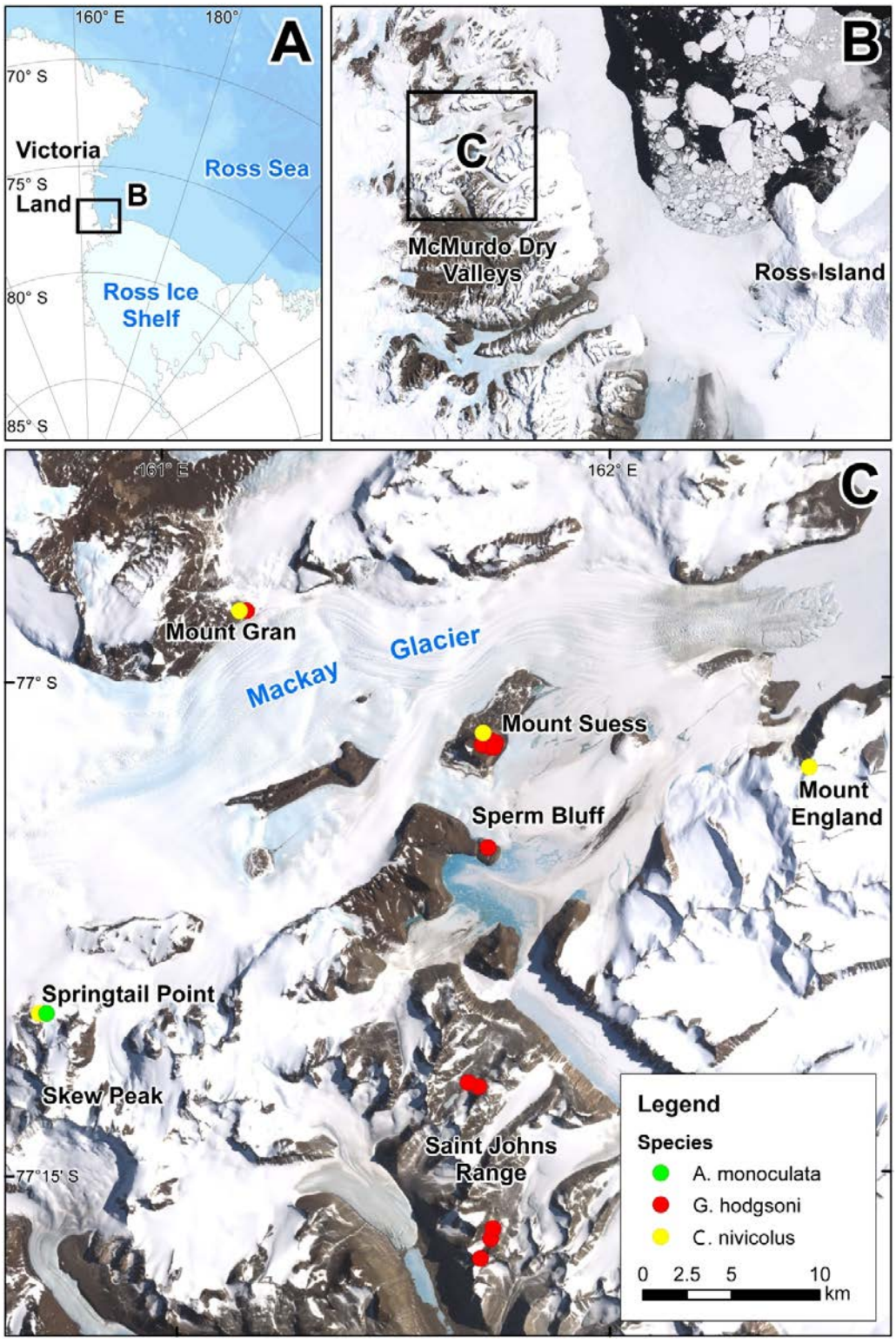
Figure 6: Haplotype network analysis for 12 haplotypes from 22 individuals of *C. nivicolus.* Haplotypes are indicated by their codes as referred to in Table 1. Missing haplotypes or mutational steps are indicated by black dots, or are collapsed into a count of missing steps as in the white square.

Figure 7: Genetic distances based on mitochondrial 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.

Figure 8: Estimated divergence times for populations of *G. hodgsoni* (circle) and *C. nivicolus* (squares). The timeline on the left is in millions of years. Overarching geologic events are presented in the appropriate time zones. Each bar indicates the divergence range between populations as indicated by the associated number pair. Each number refers to haplotypes from geographic locations as follows: 1 = *G. hodgsoni* haplotypes from the St John's range and Mt Seuss; 2 = the *G. hodgsoni* haplotype at Mt Gran; 3 = *C. nivicolus* haplotypes from Springtail Point; 4 = *C. nivicolus* haplotypes from Mt Gran and Mt Seuss; 5 = *C. nivicolus* haplotypes from Mt England.

580

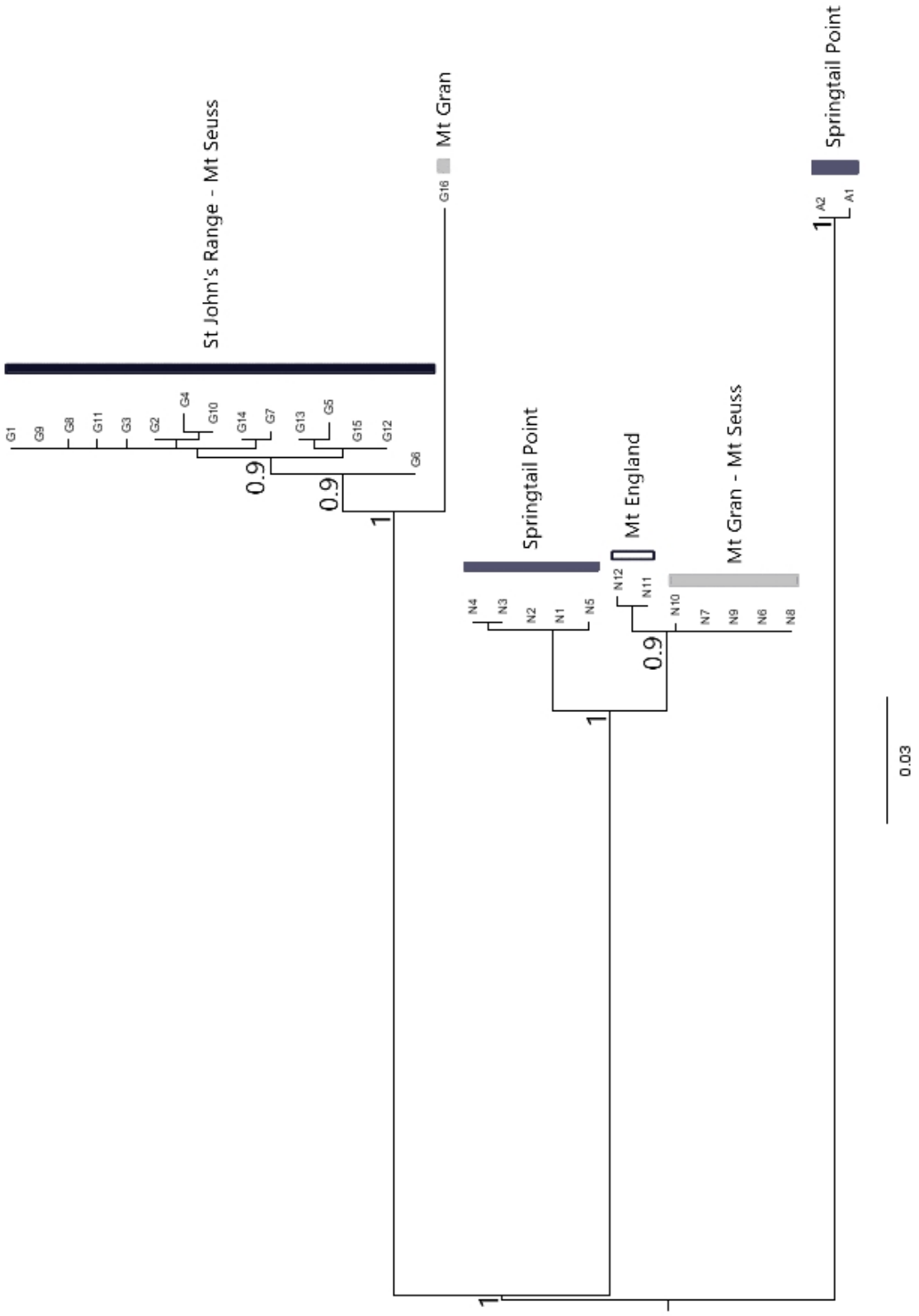
581



582

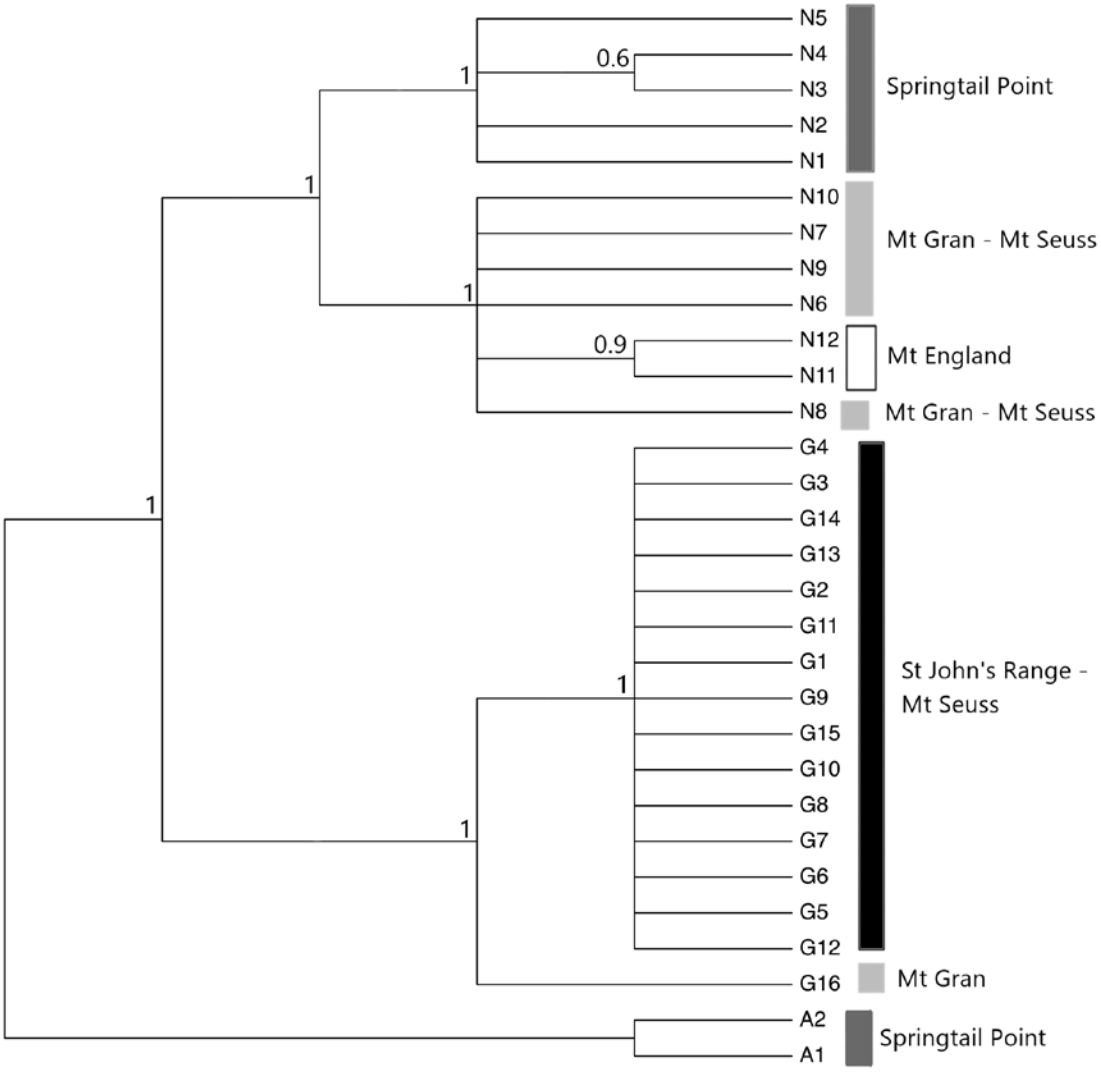
583 Fig. 1

584



585

586 Fig. 2



587

588

589

590

591 Fig. 3

592

593

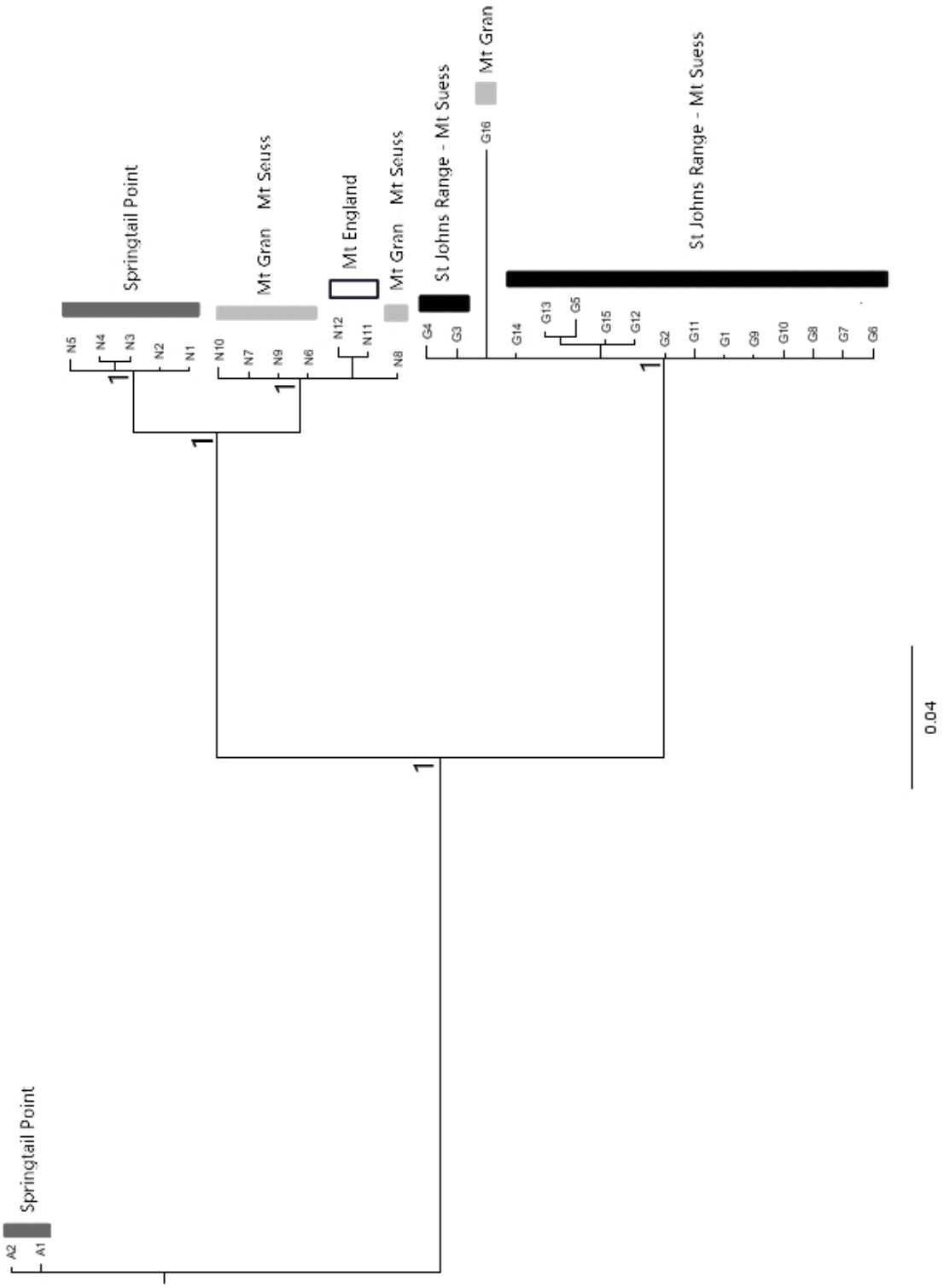

Springtail Point
Mt Gran
Mt Seuss
Mt England
Mt Gran
Mt Seuss
St Johns Range - Mt Suess
Mt Gran
St Johns Range - Mt Suess
N5
N4
N3
N2
N1
N10
N7
N9
N6
N12
N11
N8
G4
G3
G16
G14
G13
G5
G15
G12
G2
G11
G1
G9
G10
G8
G7
G6
1
1
1
1
1
A2
A1
Springtail Point
0.04

Fig. 4

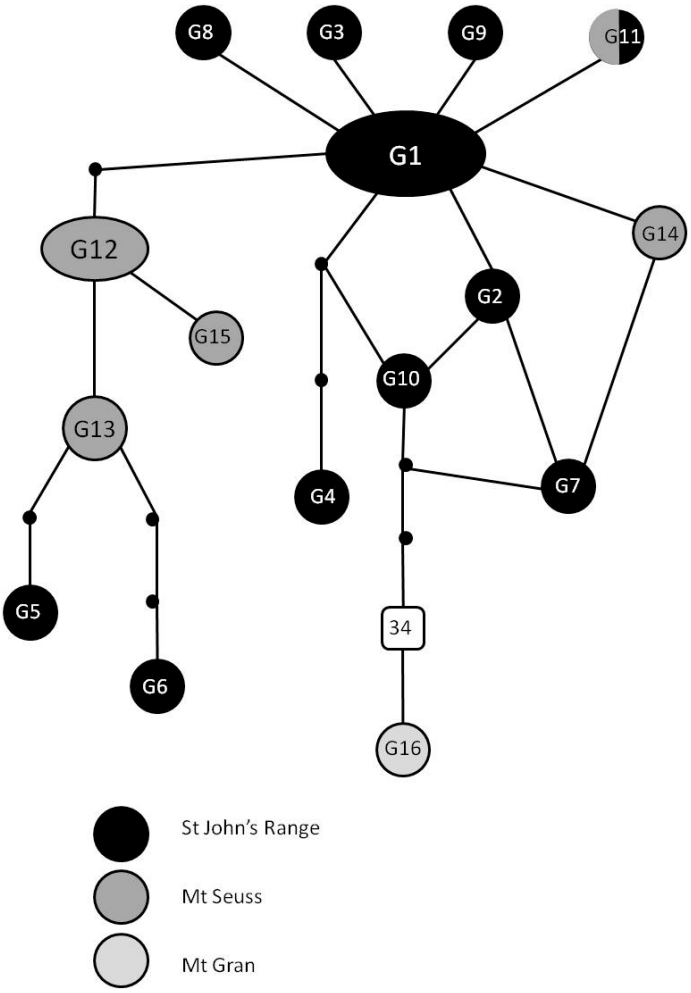
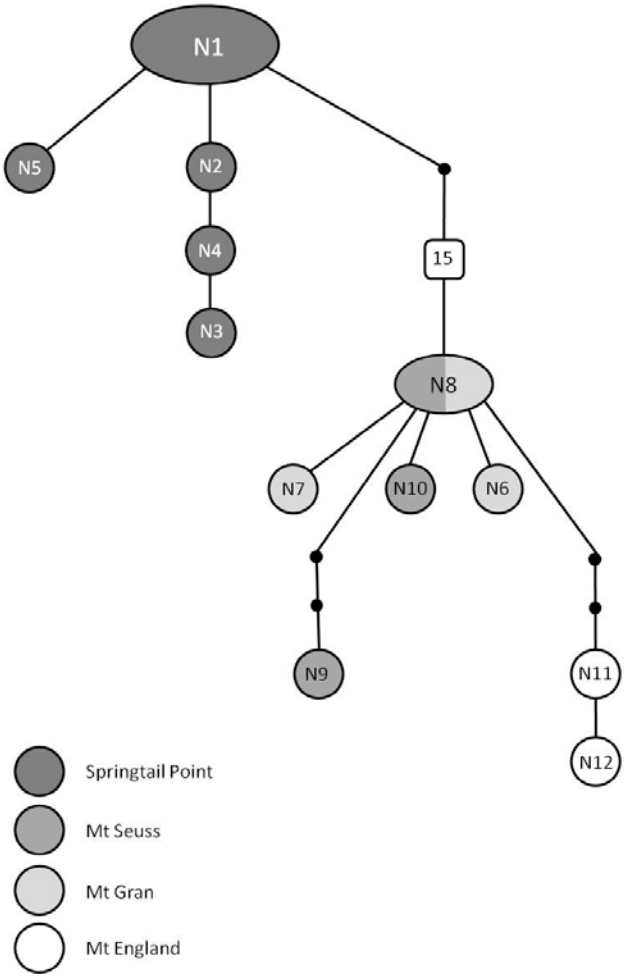

G8
G3
G9
G11
G1
G12
G14
G2
G15
G10
G13
G4
G7
G5
34
G6
G16
St John's Range
Mt Seuss
Mt Gran

Fig. 5

631

632



633

634

635

636

637 Fig. 6

638

639

640

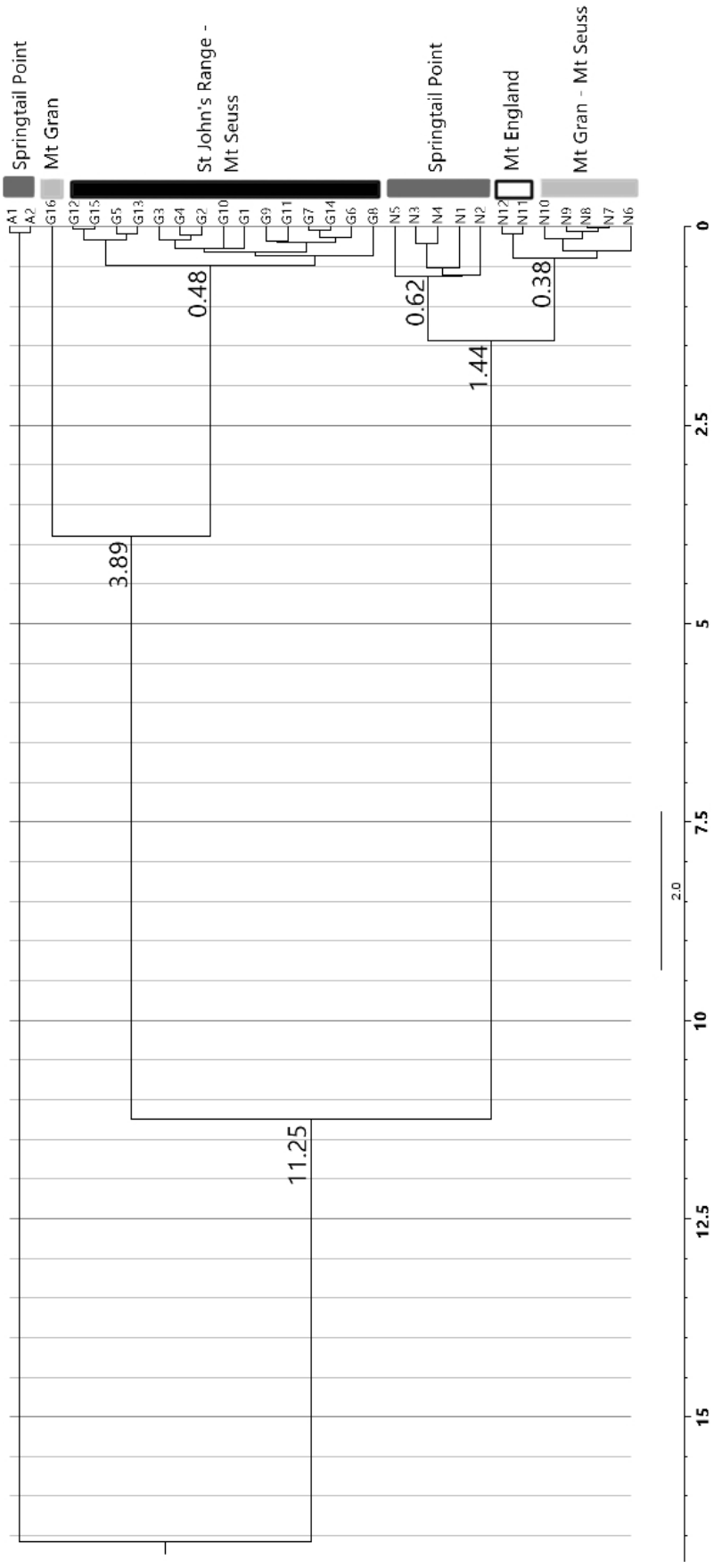
641

642

Haplotypes
G. hodgsoni
N. nivicolus
A. monoculata
St John's Range
St John's Range – Mt Seuss
Mt Seuss
Mt Gran
Springtail Point
Mt Gran – Mt Seuss
Mt England
A. monoculata
G. hodgsoni
N. nivicolus

643

644 Fig. 7


Springtail Point
Mt Gran
St John's Range - Mt Seuss
Springtail Point
Mt England
Mt Gran - Mt Seuss
A1
A2
G16
G12
G15
G5
G13
G3
G4
G2
G10
G1
G9
G11
G7
G14
G6
G8
N5
N3
N4
N1
N2
N12
N11
N10
N9
N8
N7
N6
0.48
0.62
0.38
1.44
3.89
11.25
0
2.5
5
7.5
10
12.5
15
2.0

645

646 Fig. 8