Assessing the Diversity of Antarctic and New Zealand Arthropods through DNA Barcoding

A thesis

Submitted in partial fulfilment

of the requirements for the degree

of

Master of Science (Research)

in Biological Sciences

at

The University of Waikato

by

CLARE ROSE BEET

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2016

Thesis Abstract

Diversity is the foundation of all biological and ecological studies. Globally however, biodiversity is under threat from the varied and cumulative impacts of humans on the environment. It is thus necessary to continually develop tools that are able to assess diversity at the scales now required. DNA barcoding has become an increasingly common approach for species identification as it is efficient and can facilitate high throughput analyses without the routine need of taxonomic experts. This thesis examines the genetic diversity of invertebrate groups from Antarctica and New Zealand to gain an understanding of current baseline levels of variability and to facilitate their use as indicators of environmental change.

The juvenile stages of Ephemeroptera (mayflies), Plecoptera (stoneflies) and Trichoptera (caddisflies) (EPT taxa) are key components of aquatic food webs and are frequently used as bioindicators of water quality. However, challenges in identifying juveniles to species levels are one factor limiting their more routine use. New Zealand has over 244 caddisfly species 106 species of stonefly and 50 mayfly species all of which are endemic. Here, my primary aim was to obtain COI sequence coverage for the New Zealand EPT taxa using expertly identified collections of adult specimens. A second aim was to examine levels of sequence diversity within taxa and to test the endemicity of the New Zealand fauna through comparison with international records. Thus far, 225 caddisfly BINs, 48 stonefly and 37 mayfly BINs have been generated. Average intraspecific divergences were between 1.2-1.4% for all three orders while average interspecific distances ranged from 24-32%. The

designation of all New Zealand EPT species as endemic was supported with interspecific divergences generally above 13%. These data can facilitate the rapid and accurate assessment of larval specimens and can furthermore be used to facilitate research into the phylogenetic and phylogeographic patterns that have shaped the EPT fauna worldwide.

I also examined mitochondrial DNA (COI) sequences for three Antarctic springtail (Collembola) species collected from sites in the vicinity, and to the north of, the Mackay Glacier (77^{o}S) within the Ross Sea region. This area represents a transitional zone between two biogeographic regions (North and South Victoria Land). Here, I assessed levels of genetic variability within and among populations of the three putative springtail species. Each of the three recognised species had multiple highly divergent intraspecific populations (5-11.3% sequence divergence). Based on molecular clock estimates, these divergent lineages were likely to have been isolated for 3-5 million years, a time when the Western Antarctic Ice Sheet (WAIS) was thought to have completely collapsed. Given the current isolation of these genetically distinct populations, any future changes in species' distributions can be easily tracked through the DNA barcoding of individual springtails across the Mackay Glacier ecotone.

Collectively these two studies have established baseline levels of COI diversity for New Zealand and Antarctic invertebrates. Furthermore, they have revealed hidden (cryptic) diversity in both regions and presented opportunities to incorporate DNA barcoding into future studies of New Zealand aquatic and terrestrial Antarctic ecosystems.

Acknowledgements

First and foremost, I would like to thank my supervisor Ian Hogg for emailing me one day and asking if I wanted to go to Antarctica, and then saying if so, the proposal is due tomorrow. I am so grateful for all of the opportunities you have offered me including the trip to Canada. Thank you for all the help you've given me over the past four years and for the countless emails you've answered at all hours, especially these last few weeks.

Thank you to the New Zealand Antarctic Research Institute (NZARI) for the project funding that got our K024 team to Antarctica and Antarctica New Zealand for logistical support and making sure we were appropriately equipped. I am also grateful to New Zealand Post for awarding me the New Zealand Post Antarctic Scholarship and to the Environmental Research Institute for the ERI masters scholarship. Thank you to everyone at Scott Base for looking after us so well and in particular Andrew Hefford for superb flying in less than ideal conditions and for being the first one to find springtails at Tiger Island. Thank you to Don Cowan for helping out in the field and sitting out in the cold to read us stories. I am grateful to Diana Wall, Byron Adams and the rest of the US Antarctic 2014/15 team for sorting all of the soil samples we collected. I would also like to thank Diana, Byron and Don for all helping with revisions to the springtail manuscript.

Thank you to New Zealand Terrestrial and Freshwater Biodiversity Information System (TFBIS) for funding the EPT project. I am also grateful for being awarded the Hillary Jolly masters scholarship along with the Waikato Masters fees award and the Waikato Graduate Women Educational Trust masters scholarship which have all helped make these last two years possible.

Thank you to everyone in PBRL for the coffee breaks, the crosswords and the yarns, I have loved working in this lab. In particular, I would like to thank Stacey for always keeping the lab and financials in order and Gemma for being an amazing tent buddy to go a little insane with. I would also like to thank Gemma for all of your help throughout the year with analyses and for being there to talk through problems whenever I got stuck.

A special thanks to my favourite person, my niece Shyla for always making me smile, and to her amazing parents Sam and Henry for the frequent trips to Raglan so I could see her. Thank you to my brother Jack for the support and editing, love you ow.

Lastly, I would like to thank Mum and Dad. I would not have been able to do this without you and I am so very grateful for all of the support and love you have always given me, especially over the past two years. Thank you for all the roasts, they were a welcome change from takeaways and toast.

Table of Contents

List of Figures

Chapter II

Chapter III

List of Supplementary Materials

Chapter I

Thesis Introduction

Accurate assessments of biological diversity are the foundation of all ecological studies (Wilson 1985a; Hebert et al. 2003). Globally however, biodiversity is under threat from a variety of anthropogenic sources including climate change (Thomas et al. 2004), pollution (Freedman 1995), habitat disruption and fragmentation (Wilcox & Murphy 1985) along with invasive pests and pathogens (Simberloff & Van Holle 1999) (see Diamond 1989; Brook et al. 2008). These factors have collectively led to the beginning of a new epoch known as the anthropocene which is characterised by the current so-called 'sixth mass extinction' (Steffen et al. 2007; Wake & Vredenburg 2008).

Traditional morphological approaches for identification have been integral to advancing our knowledge of biodiversity to the present day. However, we currently lack the ability to assess biodiversity at the magnitude and global scale that is now urgently required in the face of mass extinctions (Valentini et al. 2008; Wake & Vredenburg 2008). DNA barcoding, which uses short standardised sequence fragments from the mitochondrial COI gene (Hebert et al. 2003), can provide a viable alternative. These sequences are replicable and accurate, can facilitate high throughput analyses and are available to anyone with access to a DNA sequencing facility. They also minimise the need for taxonomic experts in routine identifications, thus freeing them to undertake more meaningful biodiversity research (Hajibabaei et al. 2007; Valentini et al. 2008). The utility of DNA barcoding has been further enhanced through the development of the Barcode of Life Data Systems (BOLD) (boldsystems.org; Ratnasingham & Hebert 2007). This database aids in the archiving of DNA sequences and associated specimen data (e.g. taxonomy, voucher location,

collection information). These data then form part of the reference database by which future specimens can be queried and identified.

One of the key benefits of DNA barcoding has been the appreciation that the genetic and species diversity of eukaryotes has been seriously underestimated, with factors such as phenotypic plasticity and cryptic speciation complicating traditional methods (Valentini et al. 2008). The latest applications for DNA barcoding data have been the use of next generation sequencing and metabarcoding approaches in order to rapidly document entire faunas (see Shokralla et al. 2012; Taberlet et al. 2012; Yu et al. 2012). However, these methods still require the development of local and global reference sequence libraries in order to be fully effective. Development of these reference libraries and increasing application of DNA barcoding data has led to a considerable rise in the number of known species. The number of new-to-science or undescribed species has now massively outstripped classical naming capabilities and as such, new groups are assigned barcode index numbers (BINs), which act as a surrogate taxonomic designation in the interim (Ratnasingham & Hebert 2013). The BOLD database currently houses over 170,000 formally recognised animal species and more than 440,000 BINs (boldsystems.org).

With 1.9 million species formally described and overall estimates suggesting the number of species worldwide could range between 10-100 million (Wilson 1985b), documenting the planet's eukaryotic diversity will be time consuming even with new technologies such as metabarcoding and next generation sequencing (Hajibabaei et al. 2011). Given the enormity of such a task, it can be beneficial to focus on taxa that provide insight into the diversity and ecological integrity of specific ecosystems.

Bioindicators are organisms resident in natural systems that can provide information on the quality of a system and how it changes over time in response to environmental and climatic perturbations (McGeoch 1998; Hajibabaei et al. 2011). Bioindicators are particularly valuable as they integrate a number of stochastic chemical and physical variables to provide an overall view of their environment (McGeoch 1998; Duggan et al. 2002). A biomonitoring approach can furthermore be helpful for identifying potential biodiversity hotspots and areas of conservation importance (Myers et al. 2000). Examples of bioindicators include aquatic insects in freshwater environments and microarthropods in soil environments (Stark 1993; Hopkin 1997). Developing comprehensive DNA barcode databases for such indicator species is a crucial first step in enabling the rapid and routine use of such organisms.

In this thesis I focus on the molecular diversity of four orders of invertebrates, which have previously been used as bioindicators in order to further expand our knowledge of the biodiversity of New Zealand aquatic and terrestrial Antarctic ecosystems. Terrestrial Antarctic ecosystems have traditionally been under-explored and under-sampled while the genetic diversity of New Zealand's aquatic macroinvertebrate fauna has not been assessed at the comprehensive levels required for meaningful use as a bioindicators. Consequently, I examined the genetic diversity of 1.) New Zealand freshwater mayflies (Ephemeroptera), stoneflies (Plecoptera) and caddisflies (Trichoptera) and 2.) Antarctic springtails (Collembola) to gain an understanding of current baseline levels of diversity and variability within their respective habitats and to facilitate their potential use as indicators of environmental quality and climate changes.

The first research chapter (Chapter II) focusses on an assessment of three of the most environmentally sensitive groups of freshwater macroinvertebrates Ephemeroptera, Plecoptera and Trichoptera, otherwise known as 'EPT' taxa. Within New Zealand there are over 50 mayfly, 106 stonefly and 244 caddisfly species, all of which are endemic (Collier 1993; Hitchings 2003; Ward 2003; Fochetti & de Figueroa 2008; McLellan 2008). The larval/nymphal stages of EPT taxa are key components of aquatic food webs and exhibit varying degrees of environmental sensitivity (Pilgrim et al. 2011; Sweeney et al. 2011). However, their more widespread application as bioindicators has been hindered by the fact that species are identified based on adult male morphology while juveniles are the focus of environmental assessments (Zhou et al. 2009; Webb et al. 2012; Jackson et al. 2014). My research aimed to address this disconnect by developing a DNA database of New Zealand EPT taxa using curated collections of adult specimens and to test the endemicity of the New Zealand fauna through comparison with international records. Ultimately, these data will streamline macroinvertebrate surveys, facilitate whole community analyses via next-generation sequencing and highlight any taxonomic inconsistencies or phylogeographic structuring within and among species.

My second research chapter (Chapter III) assessed the fine scale diversity and distribution of three species of Antarctic Collembola found throughout and to the north of, the Mackay Glacier located within the Ross Sea region. Antarctica is subject to some of the harshest conditions on earth, which have resulted in the development of a highly adapted yet species-poor fauna (Convey 2011). This limited diversity is also underestimated with logistical difficulties leaving many areas unexplored and

under-sampled (see Peat et al. 2007). Genetic diversity in particular has been underestimated (Stevens & Hogg 2003; McGaughran et al. 2008). There have been a number of cases where morphologically identified species were found to represent multiple distinct clades or cryptic species (see McGaughran et al. 2010 and Mortimer et al. 2011).

Within the Ross Sea Region of Antarctica there are three currently recognised biogeographic zones; North Victoria Land, South Victoria Land and the Queen Maud Mountains (Terauds et al. 2012), each with three (and in one case four) unique and endemic species of springtail (Salmon 1965; Wise 1967; Adams et al. 2006; Hogg et al. 2014). The Mackay Glacier is situated to the north of the McMurdo Dry Valleys and acts as a transitional zone or ecotone (*sensu* Risser 1993). For example, in the McMurdo Dry Valleys there is only a single species of springtail (*Gomphiocephalus hodgsoni*), whereas an additional two rarer species (*Antarcticinella monoculata* and *Cryptopygus nivicolis*) are found within and to the north of, the Mackay Glacier (Hogg et al. 2014). Ecotones are potentially valuable sites for studying climate change as they are thought to be where alterations in species distributions or the introductions of 'new' species are most likely to occur (Gosz 1993; Risser 1993)

Springtails are the largest year-round terrestrial Antarctic animals and also have among the narrowest environmental tolerance levels of the terrestrial fauna, making them useful bioindicators of climate change (Hopkin 1997; Hogg et al. 2014; Collins & Hogg 2015). Climate changes within the Antarctic are dynamic. There has been a rapid warming of Western Antarctica by 2.5°C over the past 50 years (Turner et al. 2005; Steig et al. 2009). The aim of this study was to assess springtail distribution

and genetic (COI) diversity within the vicinity of Mackay Glacier. This research is essential for establishing baseline levels of diversity and distributional limits against which any future changes in diversity can be detected. Furthermore these data can be used to evaluate the evolution of the Antarctic landscape and in particular, the influence of past glaciations on present day species distributions and patterns of genetic diversity.

This thesis concludes with a summary chapter, which outlines the key findings of both research chapters (Chapters II and III). This final chapter also suggests possible avenues of research that would be profitable in the future.

LITERATURE CITED

Adams BJ, Bardgett RD, Ayres C, Wall DH, Aislabie J, Bamforth S, Bargagli R, Cary SC, 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.

Brook BW, Sodhi NS, Bradshaw CJ (2008) Synergies among extinction drivers under global change. *Trends in ecology & evolution*, **23**:8, 453-460.

Collier K (1993) Review of the status, distribution, and conservation of freshwater invertebrates in New Zealand. *New Zealand Journal of Marine and Freshwater Research*, **27**: 339-356.

Collins GE, Hogg ID (2015) Temperature-related activity of *Gomphiocephalus hodgsoni* (Collembola) mitochondrial DNA (COI) haplotypes in Taylor Valley, Antarctica. *Polar Biology*, **39**:2, 379-389

Convey P (2011) Antarctic terrestrial biodiversity in a changing world. *Polar Biology*, **34**, 1629–1641.

Diamond J (1989) Overview of recent extinctions. *In* Conservation for the twenty-first century, *Edited by* D Western and MC Pearl. Oxford University press, New York, pp. 37-41.

Duggan IC, Green JD, Thomasson K (2002) Do rotifers have potential as bioindicators of lake trophic state? *Verhandlungen des Internationalen Verein Limnologie*, **27**, 3497-3502.

Fochetti R, De Figueroa JMT (2008) Global diversity of stoneflies (Plecoptera; Insecta) in freshwater. *Hydrobiologia*, **595**, 365-377.

Freedman B (1995) Environmental ecology: the ecological effects of pollution, disturbance, and other stresses. Academic Press, London, UK.

Gosz JR (1993). Ecotone hierarchies. *Ecological applications*, **3**:3, 370-376.

Hajibabaei M, Singer GA, Hebert PD, Hickey DA (2007) DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *TRENDS in Genetics*, **23**:4, 167-172.

Hajibabaei M, Shokralla S, Zhou X, Singer GA, Baird DJ (2011) Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. *PLoS one*, **6**:4, e17497.

Hebert PD, Cywinska A, Ball SL (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences*, **270**, 313-321.

Hogg ID, Stevens MI, Wall DH (2014) Invertebrates. *In* Antarctic terrestrial microbiology. *Edited by* D. A. Cowan. Springer Heidelberg, New York, pp.55-78.

Hopkin SP (1997) Biology of the springtails (Insecta: Collembola). Oxford University Press, New York.

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**:2, 606-618.

McGaughran A, Convey P, Stevens MI, Chown SL (2010) Metabolic rate, genetic and microclimate variation among springtail populations from sub-Antarctic Marion Island. *Polar Biology*, **33**, 909–918.

McGeoch MA (1998) The selection, testing and application of terrestrial insects as bioindicators. *Biological Reviews of the Cambridge Philosophical Society*, **73**:02, 181-201.

McLellan I (2008) Checklist of the New Zealand Plecoptera (stoneflies). National Center for aquatic biodiversity and security. https://www.niwa.co.nz/sites/niwa.co.nz/files/import/attachments/stoneflies_june_2008.pdf

Mortimer E, Jansen van Vuuren B, Lee JE, Marshall DJ, Convey P, Chown SL (2011) Mite dispersal among the Southern Ocean Islands and Antarctica before the last glacial maximum. *Proceedings of the Royal Society of London B: Biological Sciences*, **278**, 1247–1255.

Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GA, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**:6772, 853-858.

Peat HJ, Clarke A, Convey P (2007) Diversity and biogeography of the Antarctic flora. *Journal of Biogeography*, **34**:1, 132-146.

Pilgrim EM, Jackson SA, Swenson S, Turcsanyi I, Friedman E, Weigt L, Bagley MJ (2011) Incorporation of DNA barcoding into a large-scale biomonitoring program: opportunities and pitfalls. *JNABS Journal*, **30**:1, 217-231.

Ratnasingham S, Hebert PDN (2007) BOLD: the barcode of life data system (www.barcodinglife.org). *Molecular Ecology Notes*, **7**, 355–364.

Ratnasingham S, Hebert PD (2013) A DNA-based registry for all animal species: The Barcode Index Number (BIN) System. *PloS one*, **8**:7, e66213.

Risser PG (1993). Ecotones at local to regional scales from around the world. *Ecological Applications*, **3**, 367–368.

Salmon JT (1965) Two new genera of Antarctic Collembola. *Journal of Pacific Insects*, **7**:3, 468-472.

Shokralla S, Spall JL, Gibson JF, Hajibabaei M (2012) Next-generation sequencing technologies for environmental DNA research.*Molecular ecology*, **21**:8, 1794-1805.

Simberloff D, Von Holle B (1999) Positive interactions of nonindigenous species: invasional meltdown?. *Biological invasions*, **1**:1, 21-32.

Stark JD (1993) Performance of the Macroinvertebrate Community Index: Effects of sampling method, sample replication, water depth, current velocity, and substratum on index values. *New Zealand Journal of Marine and Freshwater Research*, **27**:4, 463-478

Steffen W, Crutzen PJ, McNeill JR (2007) The Anthropocene: are humans now overwhelming the great forces of nature. *AMBIO: A Journal of the Human Environment*, **36**:8, 614-621.

Steig EJ, Schneider DP, Rutherford SD, Mann ME, Comiso JC, Shindell DT (2009) Warming of the Antarctic ice-sheet surface since the 1957 international geophysical year. *Nature*, **457**, 459–462

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**:9, 2357-2369.

Sweeney BW, Battle JM, Jackson JK, Dapkey T (2011). Can DNA barcodes of stream macroinvertebrates improve descriptions of community structure and water quality? *Journal of the North American Benthological Society*, **30**, 195-216.

Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012) Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, **21**:8, 2045-2050.

Terauds A, Chown SL, Morgan F, Peat HJ, Watts DJ, Keys H, Convey P, Bergstrom DM (2012). Conservation biogeography of the Antarctic. *Diversity and Distributions*, **18**:7, 726-741.

Thomas CD, Cameron A, Green RE, Bakkenes M, Beaumont LJ, Collingham YC, Erasmus BFN, de Siqueira MF, Grainger A, Hannah L, Hughes L, Huntley B, van Jaarsveld AD, Midgley GF, Miles L, Ortega-Huerta MA, Peterson AT, Phillips OL, Williams SE (2004). Extinction risk from climate change. *Nature*, **427**:6970, 145-148.

Yu DW, Ji Y, Emerson BC, Wang X, Ye C, Yang C, Ding Z (2012) Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution*, **3**:4, 613-623.

Valentini A, Pompanon F, Taberlet P (2008). DNA barcoding for ecologists. *Trends in Ecology and Evolution*, **24**:2, 110-117.

Wake DB, Vredenburg VT (2008) Are we in the midst of the sixth mass extinction? A view from the world of amphibians. *Proceedings of the National Academy of Sciences*, **105**, 11466-11473.

Ward JB (2003) Checklist of the New Zealand Trichoptera (caddis). National Center for aquatic biodiversity and security. https://www.niwa.co.nz/sites/niwa.co.nz/files/import/attachments/caddis.pdf

Wilcox BA, Murphy DD (1985) Conservation strategy: the effects of fragmentation on extinction. *The American Naturalist*, **125**:6, 879-887.

Wilson EO a (1985) The biological diversity crisis. *BioScience*, 700-706.

Wilson EO b (1985) Time to Revive Systematics. *Science*, **230**, 1227.

Wise KAJ (1967) Collembola (springtails). *Antarctic Research Series*, **10**, 123–148.

Zhou X, Adamowicz SJ, Jacobus LM, DeWalt RE, Hebert PDN (2009) Towards a comprehensive barcode library for arctic life - *Ephemeroptera*, *Plecoptera*, and *Trichoptera* of Churchill, Manitoba, Canada. *Frontiers in Zoology*, **6**.

Chapter II

Assessing diversity of the New Zealand "EPT" (Ephemeroptera, Plecoptera, Trichoptera) taxa using mitochondrial DNA (COI) sequences*

*To be submitted under the same title as, Beet CR, Hogg ID, Smith BJ, Bennett KR, Collins GC

ABSTRACT

The larval/nymphal stages of Ephemeroptera (mayflies), Plecoptera (stoneflies) and Trichoptera (caddisflies) (EPT taxa) are key components of aquatic food webs and are frequently used as bioindicators of water quality. However, challenges in identifying larvae/nymphs to species levels are one factor limiting their more routine use. New Zealand has over 50 described species of mayfly, 106 stonefly and 244 caddisfly all of which are endemic. The aim of this study was to provide comprehensive mitochondrial DNA (COI) sequence (DNA barcode) coverage for the New Zealand EPT taxa using curated adult specimens housed in national collections. To date, COI sequences have been generated for 41 mayfly, 41 stonefly and 180 caddisfly species. For stoneflies and caddisflies this has resulted in over 48 and 225 putative taxonomic units (BINs), respectively. Average intraspecific divergences were between 1.2-1.4% for all three orders while, average interspecific distances ranged from 24-32%. COI sequences were largely congruent with existing taxonomic designations and also highlighted some potential misidentifications within museum collections as well as several taxa in need of further review (e.g. *Deleatidium* spp, *Zelandobius* spp., *Tiphobiosis* spp.). Comprehensive DNA barcode data will facilitate the rapid and accurate assessment of larval specimens and additionally provide the opportunity to incorporate EPT taxa into next generation sequencing approaches, which seek to analyse entire aquatic insect communities.

INTRODUCTION

Globally, waterways and their inhabitants are becoming increasingly threatened by surrounding land use changes, the introduction of invasive species, pollution and climate changes (Collier 1993; Holzenthal 2007; Steffen et al. 2007). Understanding the environmental integrity of streams and rivers is critical for mitigating the effects of anthropogenic activities and identifying areas most at risk (Sweeney et al. 2011). The larval/nymphal stages of Ephemeroptera (mayflies), Plecoptera (stoneflies) and Trichoptera (caddisflies) (EPT taxa) are key components of aquatic food webs and are frequently used as bioindicators of water quality (Collier 1993; De Moor & Ivanov 2008; Hogg et al. 2009; Sweeney et al. 2011). Bioindicators are particularly useful as they integrate a number of stochastic physical and chemical variables to provide a more comprehensive view of the ecological state of the system (McGeoch 1998; Duggan et al. 2001).

Macroinvertebrate surveys using classical taxonomy rely on the assessment of juvenile specimens despite most species designations being based on the morphology of adult males (Zhou et al. 2009; De Moor & Ivanov 2008; Webb et al. 2012; Jackson et al. 2014). This disconnect has limited their more routine use and has resulted in a focus on higher taxonomic levels (e.g. family), even though environmental and pollution sensitivity levels have been shown to differ among species within a family or even genus (Sweeney et al. 2011; Macher et al. 2016).

Alternative approaches to species' identification such as DNA sequencing of the mitochondrial COI gene are becoming increasingly popular as they can facilitate a rapid and accurate identification of specimens without the routine need of a taxonomic expert (Hebert et al. 2003; Valentini et al. 2008). Further, such assessments are not confounded by phenotypic plasticity or morphologically cryptic species (Jackson et al. 2014). Molecular methods are also effective in identifying partial or damaged specimens including those lacking diagnostic morphological features. Accordingly, molecular-based approaches can provide an additional level of standardization to macroinvertebrate community assessments (Webb et al. 2012).

In New Zealand there are roughly 50 species of mayfly, 106 stonefly and 250 caddisfly and all species are considered endemic (Hitchings 2003, Ward 2003; Hitchings 2009; Fochetti & de Figueroa 2008; McLellan 2008; Smith 2014). New Zealand mayflies are divided into eight families, dominated by the Leptophlebiidae, which encompass 30 of the 50 species present. Worldwide, stoneflies comprise two sub-orders Antarctoperlaria and Arctoperlaria. Antarctopolaria are restricted to the southern hemisphere, with three of the four families present within New Zealand, the most common of which is Gripopterygidae (Fochetti & de Figueroa 2008; McLellan 2008). Arctopolaria in New Zealand are conversely only represented by a single family Notonemouridae with a further 11 families found globally, mostly within the northern hemisphere (Fochetti & de Figueroa 2008). The caddisflies are divided into 15 families, the most speciose of which is Hydrobiosidae. Within Hydrobiosidae alone there are over 86 currently recognised and 18 as yet unnamed species (Collier 1993; Ward 2003). To date, a number of studies have used COI sequences to assess

diversity of New Zealand mayflies (Macher et al. 2016), stoneflies (McCulloch et al. 2010) and caddisflies (Hogg et al. 2009; Smith & Smith 2009; Bennett 2013). However a comprehensive COI reference library for representatives of all New Zealand EPT taxa is currently lacking.

Here, our primary aim was to obtain COI sequence coverage for the New Zealand EPT taxa using curated and expertly identified collections of adult specimens. A secondary aim was to examine levels of sequence diversity within taxa and to test the endemicity of the New Zealand fauna through comparison with available international records.

METHODS

Macroinvertebrates were sourced from museums and research institutions including Canterbury Museum, Auckland Museum and the National Institute of Water and Atmospheric research (NIWA). Collections represented individuals from across all three main New Zealand islands (North, South and Stewart Island), a number of near-shore islands (e.g. Great Barrier Island), as well as individuals from Chatham and Kermadec Islands. A single leg was removed from individual specimens and each added to a single well on a 96-well plate for eventual genetic analysis at the Canadian Centre for DNA Barcoding (CCDB). Genomic DNA was extracted via the AcroPrep^{TM} PALL Glass Fibre plate method (Ivanova et al. 2006). A 658bp region of the mitochondrial COI gene was then amplified in accordance with standard CCDB protocols (see Ivanova et al. 2006) using a variety of primers including C_LepFolF (5'ATTCAACCAATCATAAAGATATTGG-3') and C_LepFolR (5'-TAAACTTCTGGATGTCCAAAAAATCA-3') (Folmer et al. 1994; Hebert et al. 2004; Ivanova et al. 2006). Amplification and sequencing of stonefly and mayfly specimens was occasionally unsuccessful and necessitated the use of alternative approaches such as mini-lep or degenerate primers to obtain partial sequences. Other primer combinations thus included C_LepFolF and MEPTR1_t1R (5'CAGGAAACAGCTATGACGGTGGRTATACIGTTCAICC-3') (Zhou et al. 2009), mLepF1 (5' GCTTTCCCACGAATAAATAATA-3') (Hajibabaei et al. 2006) and C_LepFolR, LepR1 (5' TAAACTTCTGGATGTCCAAAAAATCA -3') (Hebert et al. 2004) and mLepF1, as well as LCO1490_t1F (5'TGTAAAACGACGGCCAGTGGTCAACAAATCATAAAGATATTGG-3')

(Footit et al. 2009) and M13R (5' CAGGAAACAGCTATGAC-3') (Messing 1983). Successfully amplified products were then cleaned using Sephadex® before being sequenced in both directions on an ABI 3730xl DNA analyser. Sequencing primer combinations included M13R and C_LepFolF or mLepF1 and C_LepFolR. All photographs, collection, primer combinations and sequence data have been added to Barcode of Life Datasystems (boldsystems.org) and housed in the projects Caddisflies of New Zealand (NZCAD), Caddisflies of New Zealand II (NZTRI), Mayflies of New Zealand (NZMAY) and Stoneflies of New Zealand (NZSTO).

Sequences were assembled into three separate alignments (one for each order) using MUSCLE in Geneious 7.1.9 (Dummond et al. 2010). The caddisfly and stonefly alignments were then reduced to unique haplotypes to save computational time and used in all subsequent analyses. However, the numbers of individuals possessing each unique haplotype were noted and presented in all figures. Chi square (x^2) tests conducted in PAUP* 4.0 (Swofford, 2002) were used to determine whether base frequencies were equal among all sites, identify parsimony-informative sites, and designate first, second or third codon positions. The most appropriate model of evolution for all three alignments was determined using jModelTest 2.1.1 (Posada 2008). Bayesian trees were generated using BEAST software v1.7.5 (Drummond et al. 2007). A log normal relaxed clock model and speciation yule process as the tree prior were employed in BEAUTI v1.7.5, with the Markov chain Monte Carlo (MCMC) set at 50,000,000 generations, sampling trees every 5,000 generations. The Bayesian analyses were run in BEAST, with the quality of the results evaluated in TRACER v1.5. A burn in of 500 trees was entered into Tree Annotator v1.7.5 with

the final trees visualised in FigTree v1.4.0. Neighbour Joining (NJ) and Maximum Likelihood (ML) analyses were conducted in MEGA v5.05 (Tamura et al. 2011). ML and NJ settings both included 1000 bootstrap replicates with GTR+I+G used as the model of evolution for ML and Tamura-Nei in NJ. Tamura-Nei was used as it allows unequal base frequencies and multiple substitution types (Simon et al. 2006). All other settings were set to default options in MEGA. Maximum parsimony (MP) analyses were carried out in MEGA with 1000 bootstrap replicates employed and all other settings set to default options. Due to the large sample size, caddisflies were only analysed using Neighbour Joining and Bayesian while the mayflies and stoneflies were assessed using all four phylogenetic analyses. MEGA was also utilized to create pairwise distance matrices to calculate intraspecific and interspecific divergences while Barcode Index Numbers (BINs) were assigned by BOLD (Ratnasingham and Hebert 2013) and used as a measure of Molecular Operational Taxonomic Units (MOTUs). Pairwise distance matrices were also used to test the similarity of the New Zealand fauna with all publicly available sequences on BOLD particularly from Australia and South America.

RESULTS

Of the 618 nucleotide positions analysed in the mayfly alignment, 351 were constant, 24 were variable but uninformative and 243 were parsimony informative. The stonefly alignment comprised 658 nucleotide positions, 379 positions of which were constant with 16 positions variable but uninformative and the remaining 263 positions parsimony informative. Similarly the 615bp caddisfly alignment comprised 248 constant positions, 22 that were variable but uninformative and 345 positions that were parsimony informative. No insertions, deletions or stop codons were detected in any of the three alignments. Patterns of base frequencies were similar across the three ordinal alignments. Mayflies, stoneflies and caddisflies all had overall A-T biases of 58.7%, 60.2% and 68% respectively (Mayflies: A=24.4%, C=21.3%, G=20%, T=34.4%, Stoneflies: A=25.4%, C=21.4%, G=18.4%, T=34.8%, Caddisflies: A=30.3%, C=17.9%, G=14.1%, T=37.6%). Base frequencies were homogeneous across all sites for each of the three alignments (Mayflies: χ^2=169.2, df=189, *p*=0.85, Stoneflies: χ^2=291.5, df=264, *p*=0.12, Caddisflies: χ^2=1232.8, df=1188, *p*=0.18). However, this homogeneity was rejected at variable sites (Mayflies: χ^2 = 429.2, df =189, *p*= <0.001, Stoneflies: χ^2 =692.1, df=264, *p*= >0.001, Caddisflies: χ^2 =2335.7, df=1188, *p*= .0.001), informative sites (Mayflies: χ^2 =448.5, df=189, *p*= <0.001, Stoneflies: χ^2 =727.8, df=264, *p*= >0.001, Caddisflies: χ^2 =2530.4, df=1188, *p*= >0.001) and the third codon sites (Mayflies: χ^2 = 438.7, df=189, *p*= <0.001, Stoneflies: χ^2 =713.8, df=264, p= >0.001, Caddisflies: χ^2 =2836.3, df=1188, *p*= >0.001). The most appropriate model of evolution for all three alignments was

determined as GTR+I+G (mayfly –lnL=7953.9, stonefly –lnL=9482.7, caddisfly -lnL = 33823.3).

Sequences were obtained from 41 morphological species of mayfly from five of the eight families (Fig. 1). Of these, 32 had full-length sequences (>618bp) and were represented by 37 BINs. Nine species from six genera only had partial sequences between 300-500 nucleotides and did not generate BIN designations as sequences < 500 nucleotides are not assigned BINs. For stoneflies, sequences were obtained from 41 recognised species, 35 of which had full, 658bp sequences (Fig. 2). Of the 41 species, six species from three genera had sequences between 300-500bp (and hence not assigned BINs). Stonefly sequences consisted of 48 BINs, with 34 new to BOLD. Current caddisfly sequences covered 180 morphological species and included 225 BINs (Fig. 3). Of these, 168 species (n=553) had full length sequences (>615bp) and were included in further phylogenetic analyses.

All methods of phylogenetic analysis produced concordant results, with COI sequences reliably resolving relationships to the family level for each of the three orders (Figs 1, 2, 3). In the case of the stoneflies, analyses showed the two sub orders (Antarctopolaria and Arctopolaria) as well as differences among the finer scale tribes (Fig. 2). Results were thus generally in accordance with current taxonomic designations while also highlighting those taxa potentially requiring further attention. The average intraspecific divergences for mayflies were 1.2% (range: 0-9.4%). The highest intraspecific divergence (9.4%) was found among *Nesameletus ornatus* individuals indicating the possibility of cryptic species or phylogeographic differences. Average interspecific divergences for the mayflies were 26% with a

range of 3.6-43.3%. Genera within Leptophlebiidae were not monophyletic with *Zephlebia* spp. appearing with two other genera (*Isothraulus* and *Austronella*). Individuals identified as *N. flavitinctus* were separated by up to 26.2% and appeared in three separate BINs, indicating the presence of misidentifications. The other group that was potentially misidentified were individuals within the ACL2069 BIN, which had specimens designated as both *Zephlebia inconspicua* and *Z. tuberculata.* However, it was not clear if, or which of, these museum specimens were identified correctly. *Deleatidium* species would also benefit from re-evaluation as two clusters (Fig. 1) had individuals from three or four different morphological species with <1.3% divergence separating them (BOLD: ACL1763, BOLD: AAI9577).

For the stoneflies, the average intraspecific distance was 1.2% (range: 0-7.4) while average interspecific was 23.8% (2-35.8). The highest intraspecific value (7.4%) was found among *Holcoperla magna* individuals and could indicate the presence of cryptic species. Interspecific divergences were generally above 6%. Species within *Zelandobius* are most in need of further review. Specimens morphologically identified as *Z. macburneyi* appeared in three different clusters with a maximum of 22.8% divergence between them, although the sequence in the *Z. uniramus* cluster was likely a misidentification. The specimen identified as *Zelandoperla agnetis* which shared a BIN with *Acroperla trivacuata* (BOLD: AAZ3700) was also likely misidentified. Individuals identified as *Zelandobius pilosus* and *Z. truncus* both occurred in the same BIN with a maximum of 2.2% divergence between them, suggesting a need for morphological re-assessment. *Zelandoperla* (BOLD: AAJ9165), consisted of four putatively morphologically distinct species although

only had an average of 1% divergence between individuals (range: 0-3.37%). Upon further sequence and BIN analysis, it was found that three of the four species had been misidentified and should have been attributed as *Z. fenestrata.*

All caddisfly families were monophyletic with most genera also appearing to be monophyletic with a few exceptions (e.g. *Philanisus, Chathamia, Zelandopsyche*) (Fig. 3). *Philanisus* and *Chathamia* are marine genera and are closely related, with distances between all specimens ranging from 0-3.29%. Overall, average intraspecific divergences were 1.3% (range: 0-9.6%) while interspecific divergences averaged 31.3% (range 2-69%). Interspecific divergences were generally over 8%, with the few low values attributed to the presence of taxonomic inconsistencies. For example *Pycnocentrodes aureolis* and *P. aeris* appear to have been morphologically confused while multiple *Psilochorema* species appeared closely related (*P. bidens* and *P. donaldsoni* as well as *P. mimicum* and *P. tautoru*) with divergences between 2-3%. Five species had deep intraspecific divergences >5.5% including *Hydropsyche rarururaru* (6.8%), *Tiphobiosis childi* (9.6%), *Tiphobiosis intermedia* (8.6%) and *Pycnocentria evecta* (5.5%) indicating the presence of potential phylogeographic structuring, species complexes or cryptic species. A further eight species also had intraspecific divergences ranging from 2.5-4.9%. Two main genera that would benefit from further morphological attention are *Tiphobiosis* and *Paroxyethira*. *Tiphobiosis* has seven as yet unnamed species, a number of which appeared in multiple BINs, with individuals designated as *Tiphobiosis* sp. A, B and C all occurring in more than one BIN (e.g. BOLD: ACK5538, BOLD: AAY9307). A number of *Paroxyethira* specimens with differing morphological identifications also corresponded to unique

BINs (e.g. BOLD: ACK6505, BOLD: ABY2579). Furthermore, six different *Paroxyethira* 'species' had overlapping intraspecific and within-genus genetic distances, suggesting the need for a taxonomic re-assessment. For example, the mean intraspecific divergence among the six *Paroxyethira* species was 1.64% (range: 0-4.76%) while the mean intra-generic divergence was 2.92% (range: 0.77-5.47%).

Interspecific distances between New Zealand and international sequence records available on BOLD showed the minimum interspecific distance between New Zealand and Australian caddisflies was 6.8% among species of *Oecetis*. When this value was omitted, minimum interspecific distances ranged from 13-21.7%. Minimum interspecific distances between mayflies ranged from 23.9-25.6% with stonefly distances similarly ranging from 20.5-22.5%. The smallest stonefly genetic divergence was between the Australian *Dinotoperla* and New Zealand *Zelandoperla* species at 20.5 %.

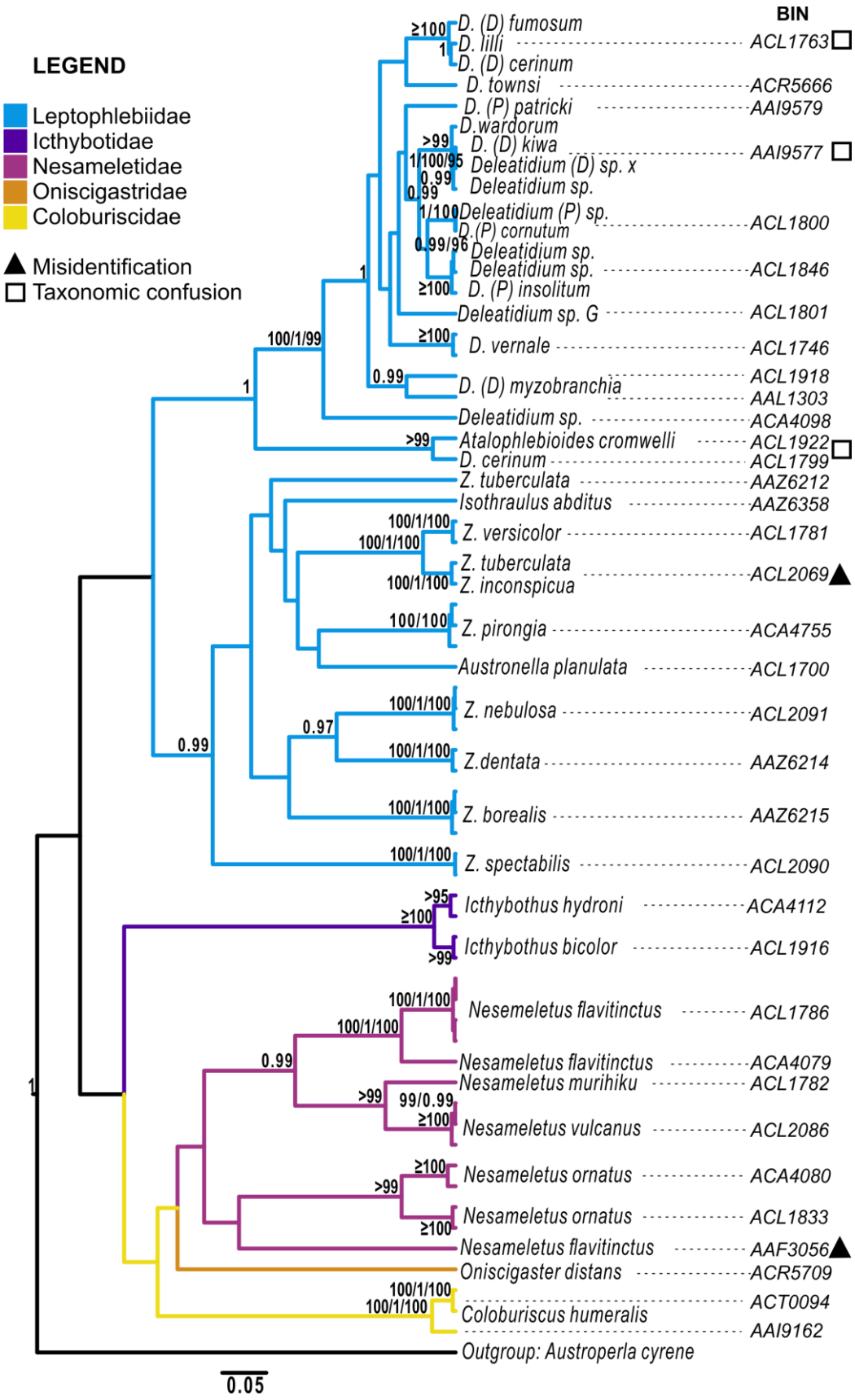

LEGEND
Leptophlebiidae
Icthybotidae
Nesameletidae
Oniscigastridae
Coloburiscidae
Misidentification
Taxonomic confusion
BIN
D. (D) fumosum
D. lilli
D. (D) cerinum
D. townsi
ACL1763
ACR5666
D. (P) patricki
D.wardorum
D. (D) kiwa
Deleatidium (D) sp. x
Deleatidium sp.
Deleatidium (P) sp.
D.(P) cornutum
Deleatidium sp.
Deleatidium sp.
D. (P) insolitum
AAI9579
AAI9577
ACL1800
ACL1846
Deleatidium sp. G
D. vernale
D. (D) myzobranchia
Deleatidium sp.
Atalophlebioides cromwelli
D. cerinum
ACL1801
ACL1746
ACL1918
AAL1303
ACA4098
ACL1922
ACL1799
Z. tuberculata
Isothraulus abditus
Z. versicolor
Z. tuberculata
Z. inconspicua
Z. pirongia
Austronella planulata
Z. nebulosa
Z.dentata
Z. borealis
Z. spectabilis
AAZ6212
AAZ6358
ACL1781
ACL2069
ACA4755
ACL1700
ACL2091
AAZ6214
AAZ6215
ACL2090
Icthybothus hydroni
Icthybothus bicolor
ACA4112
ACL1916
Nesemeletus flavitinctus
Nesameletus flavitinctus
Nesameletus murihiku
Nesameletus vulcanus
ACL1786
ACA4079
ACL1782
ACL2086
Nesameletus ornatus
Nesameletus ornatus
Nesameletus flavitinctus
ACA4080
ACL1833
AAF3056
Oniscigaster distans
ACR5709
Coloburiscus humeralis
ACT0094
AAI9162
Outgroup: Austroperla cyrene
0.05

Figure 1: Collated phylogenetic tree of NZ mayflies. Bayesian base tree with support values over 95/0.95 displayed. Support values are recorded in order of ML/Bayesian/NJ/MP and where all values are present they are displayed as >x. Tree is coloured according to family with misidentifications and areas of taxonomic confusion are respectively indicated by triangles and squares. Shortened Barcode Index Numbers are displayed on the right with each alpha numeric code preceded by 'BOLD:' e.g. BOLD: ACL1763.

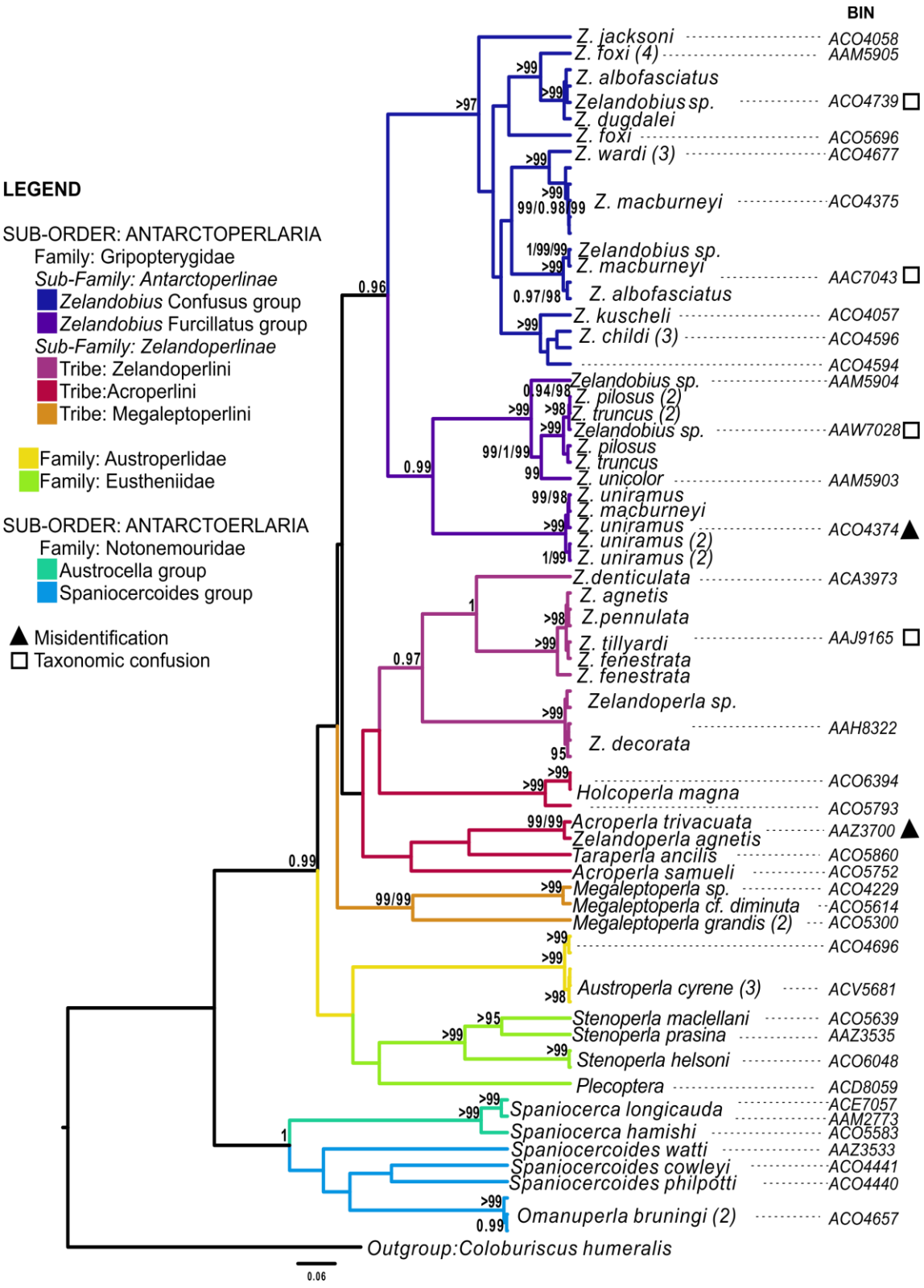

BIN
LEGEND
SUB-ORDER: ANTARCTOPERLARIA
Family: Gripopterygidae
Sub-Family: Antarctoperlinae
Zelandobius Confusus group
Zelandobius Furcillatus group
Sub-Family: Zelandoperlinae
Tribe: Zelandoperlini
Tribe:Acroperlini
Tribe: Megaleptoperlini
Family: Austroperlidae
Family: Eustheniidae
SUB-ORDER: ANTARCTOERLARIA
Family: Notonemouridae
Austrocella group
Spaniocercoides group
▲ Misidentification
☐ Taxonomic confusion
Z. jacksoni ACO4058
Z. foxi (4) AAM5905
Z. albofasciatus
Zelandobius sp. ACO4739
Z. dugdalei
Z. foxi ACO5696
Z. wardi (3) ACO4677
Z. macburneyi ACO4375
Zelandobius sp.
Z. macburneyi AAC7043
Z. albofasciatus
Z. kuscheli ACO4057
Z. childi (3) ACO4596
ACO4594
Zelandobius sp. AAM5904
Z. pilosus (2)
Z. truncus (2)
Zelandobius sp. AAW7028
Z. pilosus
Z. truncus
Z. unicolor AAM5903
Z. uniramus
Z. macburneyi
Z. uniramus ACO4374
Z. uniramus (2)
Z. uniramus (2)
Z.denticulata ACA3973
Z. agnetis
Z.pennulata
Z. tillyardi AAJ9165
Z. fenestrata
Z. fenestrata
Zelandoperla sp.
AAH8322
Z. decorata
ACO6394
Holcoperla magna
ACO5793
Acroperla trivacuata AAZ3700 ▲
Zelandoperla agnetis
Taraperla ancilis ACO5860
Acroperla samueli ACO5752
Megaleptoperla sp. ACO4229
Megaleptoperla cf. diminuta ACO5614
Megaleptoperla grandis (2) ACO5300
ACO4696
Austroperla cyrene (3) ACV5681
Stenoperla maclellani ACO5639
Stenoperla prasina AAZ3535
Stenoperla helsoni ACO6048
Plecoptera ACD8059
Spaniocerca longicauda ACE7057
Spaniocerca hamishi AAM2773 ACO5583
Spaniocercoides watti AAZ3533
Spaniocercoides cowleyi ACO4441
Spaniocercoides philpotti ACO4440
Omanuperla bruningi (2) ACO4657
Outgroup:Coloburiscus humeralis
0.06

Figure 2: Collated phylogenetic tree of NZ stoneflies. Bayesian base tree with support values over 95/0.95 displayed. Support values are recorded in order of ML/ Bayesian/NJ/MP and where all values are present they are displayed as >x. Tree is coloured according to tribe/ group with higher taxonomic levels indicated in the legend. Misidentifications and areas of taxonomic confusion are respectively indicated by triangles and squares. Shortened Barcode Index Numbers (BINs) are displayed on the right with each alpha numeric code preceded by 'BOLD:' e.g. BOLD: ACO4058.

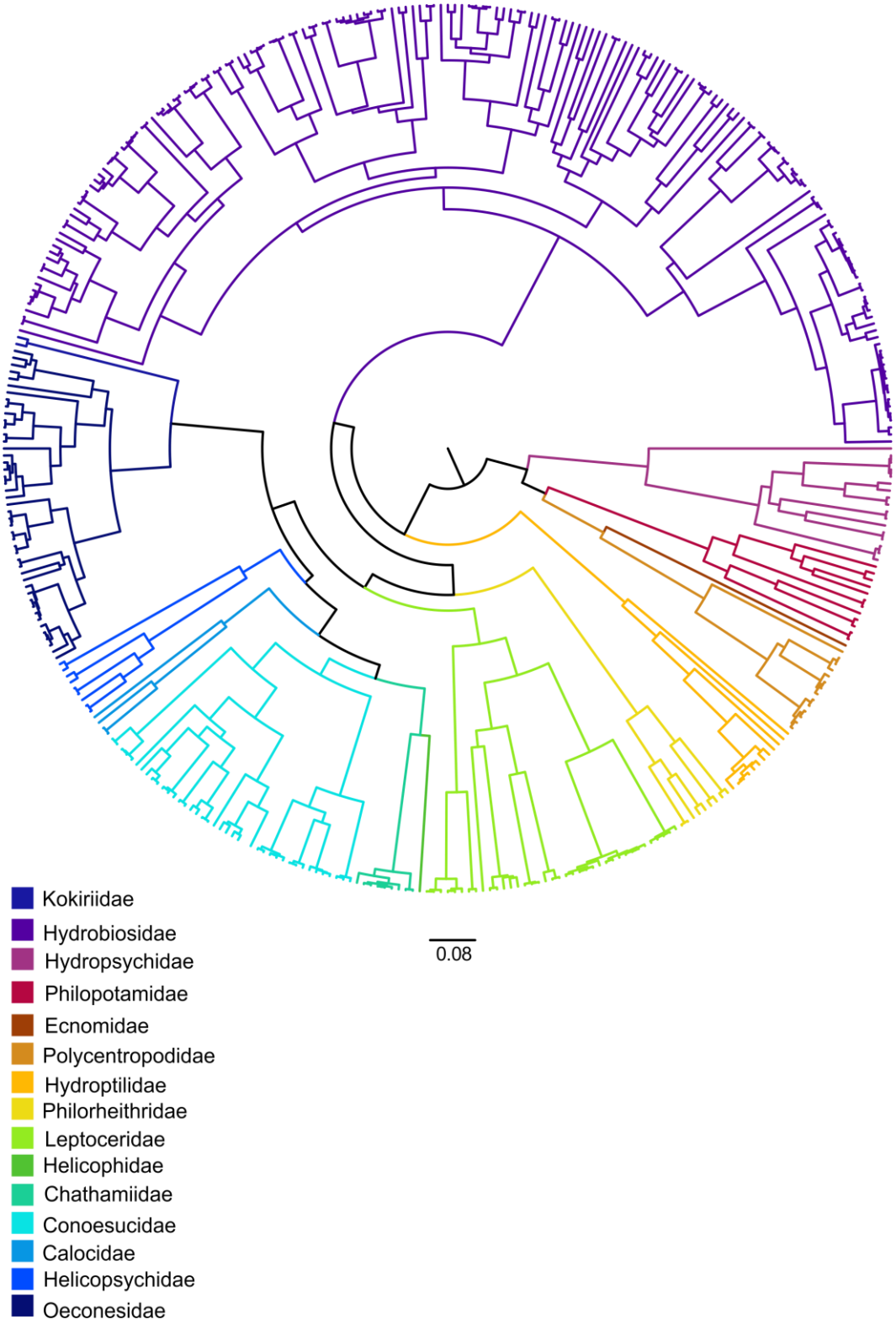

Kokiriidae
Hydrobiosidae
Hydropsychidae
Philopotamidae
Ecnomidae
Polycentropodidae
Hydroptilidae
Philorheithridae
Leptoceridae
Helicophidae
Chathamiidae
Conoesucidae
Calocidae
Helicopsychidae
Oeconesidae
0.08

Figure 3: Bayesian circle tree of NZ caddisflies coloured according to family. Posterior probabilities have been omitted, although all terminal branches had high support (over 0.7). For full, straight Bayesian tree see Supplementary Material 1 (pp. 59).

DISCUSSION

This study has begun the assembly of a complete reference library for the New Zealand EPT taxa. Including shorter sequences (300-500bp), 41 recognised and unnamed mayfly species representing have been recorded from the 71 sequences generated. From the 94 stonefly sequences 41 recognised and unnamed species and 48 BINs have been produced. The 565 sequences generated as part of the caddisfly analyses, represented a total of 180 morphological species, 225 BINs and provided coverage of all the New Zealand caddisfly genera. Many of the species that are yet to be sequenced represent rare species and many are absent or exist only as type specimens in museum collections. There are seven mayfly, 13 stonefly and 17 caddisfly species considered naturally uncommon (Grainger et al. 2013). Fortunately, such taxa are unlikely to be important for routine macroinvertebrate assessments. For the purposes of water quality assessments the caddisfly database can thus be considered functionally complete as the New Zealand macroinvertebrate community index (MCI) designates sensitivity values according to genera and not species (Stark 1993). However, the usefulness of these databases goes beyond the scope of streamlining macroinvertebrate assessments, it also presents opportunities to examine phylogenies, resolve taxonomic issues, and evaluate levels of diversity within and among geographic regions.

The study has highlighted a number of misidentifications within museum collections in addition to suggesting groups in need of further taxonomic review. The paraphyly of *Zephlebia* was likely an artefact of limited sequences as the genera *Isothraulus* and

Austronella are currently only represented by a single sequence. Three of the eight mayfly families are also only represented by a single species. Further individuals and additional sequences will help to resolve such issues. Individuals morphologically identified as *Deleatidium wardorum* and *D. kiwa* both appeared in the same BIN. Both are recently described species that resemble each other and are closely related (Hitchings 2010). Our genetic data suggest that these individuals may need to be re-examined to determine whether their species designations are valid or whether their appearance in the same BIN is simply a case of misidentification or resulting from a relatively recent divergence event. A recent study by Macher et al (2016) recorded 12 divergent *Deleatidium* COI clades within the Southland region of New Zealand and also suggested that this genus is in need of taxonomic revision. Similarly, *Zelandobius pilosus* and *Z. truncus* appeared to be the same species genetically despite being described as highly distinct morphologically (Death 1990; McLellan 1993). This could represent a possible case of phenotypic plasticity arising from differing morphological responses to environmental stressors (*sensu* Benard 2004) or alternatively another case of misidentification. Sequencing of additional genes such as 18S, 28S and H3 could aid in resolving these taxonomic issues as if they supported the COI results it would suggest that type specimens need to be re-examined and ideally barcoded. Nuclear genes may also be useful as they are slower evolving and better able to assess deeper phylogenetic relationships (Vonnemann et al. 2005; Macher et al. 2016). These examples of potential misidentifications highlight the difficulties encountered during morphological identifications, which are not limited to freshwater macroinvertebrates or invertebrates in general (Valentini et al. 2008; Webb et al. 2012; Jackson et al. 2014). However, morphological identifications

remain necessary for providing support and meaning to species boundaries suggested by genetics (Hajibabei et al. 2007; Zhou et al. 2010). Molecular techniques can also aid in identifying morphological traits that are taxonomically informative (Carew et al. 2011).

Mitochondrial COI sequences reliably resolved phylogenetic relationships to the family level for all three orders and additionally identified the finer scale relationships between tribes in the case of the stoneflies. The average intraspecific divergence values for mayflies (1.2%) stoneflies (1.2%) and caddisflies (1.3%) were all similar to those reported by Webb et al. (2012) for North American mayflies. Webb et al. (2012) found overall average intraspecific divergences of 1.97% and minimum interspecific distances between 0.3-24.7% which were also similar to the New Zealand mayfly values (average = 26%, range: 3.6-43.3%) stoneflies (average = 23.5%, range: 2-35.8%) and caddisflies (average = 31.3%, range: 2-69%). For North American mayflies examined (264 nominal and 90 provisional species) almost one third had intraspecific divergences higher than 2.2% (Webb et al. 2012). High maximal intraspecific divergences and potential cryptic species were also observed among Mediterranean mayflies (Cardoni et al. 2015), Australian and Chilean caddisflies (Baker et al. 2004; Pauls et al. 2010) and North American mayflies, stoneflies and caddisflies (Zhou et al. 2009; Zhou et al. 2010; Sweeney et al. 2011; Jackson et al. 2014).

Evidence for phylogeographic structuring, species complexes or cryptic species was found within each of the three New Zealand EPT orders. The species with the highest intraspecific divergence values included *Tiphobiosis childi* (9.6%), *T. intermedia*

(8.6%), *Holcoperla magna* (7.4%) and *Nesameletus ornatus* (9.4%). The divergent groups of *T. childi*, *T. intermedia* and *H. magna* all currently exist in sympatry while the *N. ornatus* clades appeared to represent a north, south island split, similar to findings of Bennett (2013) for New Zealand caddisflies. Bennett (2013) suggested that the high intraspecific divergences found between north and south island caddisfly populations reflected the isolation of the islands during the Pleistocene. While phylogeographic coverage was not the primary aim of this study, in the future it would be informative to further examine phylogeographic structuring among mayflies and stoneflies to determine whether Pleistocene glaciations had a similar influence.

The comparison of New Zealand specimens with available international sequences supported the designation of all New Zealand EPT taxa as endemic. Minimum interspecific distances were all above 20% for mayflies and stoneflies. Minimum interspecific divergences were also above 13% for all caddisflies with the exception of *Oecetis* species. New Zealand and Australian species of *Oecetis* were only separated by 6.8%, which although above the nominal 2% "species" threshold (Hebert et al. 2003; Zhou et al. 2009), is similar to levels of cryptic diversity found within New Zealand. A study by Johanson et al. (2009) demonstrated that New Zealand and Australian Conoesuciade are divergent from each other but do not form monophyletic clades according to country of origin.

The continued growth and development of the international BOLD database will facilitate finer-scale resolution of macroinvertebrate communities and thereby allow

accurate comparisons at a range of scales (Sweeney et al. 2011). This is likely to lead to the identification of further genetically distinct species resources, which could have implications for conservation management strategies (Smith & Smith 2009). Accurate assessments of species (including cryptic species) are critical, especially as 24 caddisfly and 7 mayfly species are designated as endangered, threatened or declining in New Zealand while a further 14 stonefly species are recognized as nationally critical and 22 considered data deficient (Grainger et al. 2013). Improved macroinvertebrate surveys could further enable the detection of any non-native or invasive species.

Our current data will facilitate the routine use of EPT taxa in molecular-based water quality assessments. Sweeney et al. (2011) also showed that using DNA barcoding approaches to species' identifications increased the number of macroinvertebrate species identified at sites by up to 70% compared to assessments made by taxonomic experts. Likewise, Jackson et al. (2014) identified an additional 100 taxa using DNA barcoding compared to morphological methods and was able to identify individuals to species level 93% of the time. Data obtained from comprehensive macroinvertebrate COI libraries can also be used to facilitate research into the phylogenetic and phylogeographic patterns that have shaped the EPT fauna worldwide (e.g. Kjer et al. 2001; 2002; Whitefield & Kjer 2008). These data provide the opportunity to incorporate EPT taxa into next generation sequencing and environmental DNA approaches to analyse entire aquatic insect communities (e.g. Dowle et al. 2015). Such methods rely on the presence of a reliable and robust reference library that once optimised, can enable large-scale biodiversity assessments

via the rapid processing of complex environmental samples (Hajibabei et al. 2011; Shokralla et al. 2012).

ACKNOWLEDGEMENTS

We are grateful for New Zealand Terrestrial and Freshwater Biodiversity Information System (TFBIS) funding. CB was the recipient of the Hillary Jolly Masters Scholarship. Sequencing at the Canadian Centre for DNA barcoding was supported through funding to the International Barcode of Life Project (iBOL) from the Ontario Genomics Institute (2008-OGI-ICI-03), Genome Canada, the Ontario Ministry of Research and Innovation, and the Natural Sciences and Engineering Research Council of Canada. Canterbury Museum (C. Vink) and Auckland Museum (J. Early) kindly provided access to their collections

REFERENCES

Baker AM, Hughes JM, Dean JC, Bunn SE (2004) Mitochondrial DNA reveals phylogenetic structuring and cryptic diversity in Australian freshwater macroinvertebrate assemblages. *Marine and Freshwater Research*, **55**:6, 629-640.

Benard MF (2004) Predator-induced phenotypic plasticity in organisms with complex life histories. *Annual Review of Ecology, Evolution, and Systematics*, **35**, 651-673.

Bennett KR (2013) Impacts of the Pliocene and Pleistocene glaciations on genetic diversity among New Zealand and Antarctic arthropods. M.Sc. thesis, Department of Biological Sciences, University of Waikato, Hamilton, NZ.

Cardoni S, Tenchini R, Ficulle I, Piredda R, Simeone MC, Belfiore C (2015) DNA barcode assessment of Mediterranean mayflies (Ephemeroptera), benchmark data for a regional reference library for rapid biomonitoring of freshwaters. *Biochemical Systematics and Ecology*, **62**, 36-50.

Carew ME, Marshall SE, Hoffmann AA (2011) A combination of molecular and morphological approaches resolves species in the taxonomically difficult genus *Procladius Skuse* (Diptera: Chironomidae) despite high intra-specific morphological variation. *Bulletin of entomological research*, **101**:05, 505-519.

Collier K (1993) Review of the status, distribution, and conservation of freshwater invertebrates in New Zealand. *New Zealand Journal of Marine and Freshwater Research*, **27**, 339-356.

Death RG (1990). A new species of *Zelandobius* (Plecoptera: Gripopterygidae: Antarctoperlinae) from New Zealand. *New Zealand natural sciences*, *17*, 23-28.

De Moor FC, Ivanov VD (2008) Global diversity of caddisflies (Trichoptera: Insecta) in freshwater. *Hydrobiologia*, **595**:1, 393-407.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC evolutionary biology*, **7**:214.

Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Duran C, Field M, Heled J, Kearse M, Markowitz S, Moir R, Stones-Havas S, Sturrock S, Thierer T, Wilson A (2010) Geneious v5.1. Biomatters Ltd, Auckland. http://www.geneious.com

Dowle EJ, Pochon X, Banks JC, Shearer K, Wood SA (2015) Targeted gene enrichment and high throughput sequencing for environmental biomonitoring: a case study using freshwater macroinvertebrates. *Molecular Ecology Resources*. doi: 10.1111/1755-0998.12488.

Duggan IC, Green JD, Thomasson K (2002) Do rotifers have potential as bioindicators of lake trophic state? *Verhandlungen des Internationalen Verein Limnologie*, **27**, 3497-3502.

Fochetti R, De Figueroa JMT (2008) Global diversity of stoneflies (Plecoptera; Insecta) in freshwater. *Hydrobiologia*, 595, 365-377.

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.

Foottit RG, Maw HEL, Havill NP, Ahern RG, Montgomery ME (2009) DNA barcodes to identify species and explore diversity in the Adelgidae (Insecta: Hemiptera: Aphidoidea). *Molecular Ecology Resources*, **9**, 188-195.

Grainger N, Collier K, Hitchmough R, Harding J, Smith B, Sutherland D (2013). Conservation status of New Zealand freshwater invertebrates. *New Zealand Threat Classification Series*, **8**, 28pp.

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.

Hajibabaei M, Singer GA, Hebert PD, Hickey DA (2007) DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *TRENDS in Genetics*, **23**:4, 167-172.

Hajibabaei M, Shokralla S, Zhou X, Singer GA, Baird DJ (2011). Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. *PLoS one*, **6**:4, e17497.

Hebert PD, Cywinska A, Ball SL (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences*, **270**, 313-321.

Hebert PDN, Penton EH, Burns J, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly, *Astraptes fulgerator*. *PNAS*, **101**, 14812-14817.

Hitchings TR (2003) Checklist of the New Zealand Ephemeroptera (mayflies). National Center for aquatic biodiversity and security. http://www.niwa.co.nz/sites/niwa.co.nz/files/import/attachments/mayflies.pdf

Hitchings TR (2009) Leptophlebiidae (Ephemeroptera) of the alpine region of the Southern Alps, New Zealand. *Aquatic Insects: International Journal of Freshwater Entomology*, **31**, 595-601.

Hitchings T (2010). Two new species of *Deleatidium* (*Deleatidium*) (Ephemeroptera: Leptophlebiidae) from the South Island, New Zealand. *Records of the Canterbury Museum*, **24**, 27-38.

Hogg ID, Smith BJ, Banks JC, deWaard JR, Hebert PD (2009) Testing use of mitochondrial COI sequences for the identification and phylogenetic analysis of New Zealand caddisflies (Trichoptera). *New Zealand Journal of Marine and Freshwater Research*, **43**, 1137-1146.

Holzenthal RW, Blahnik RJ, Prather AL, Kjer KM (2007) Order Trichoptera Kirby, 1813 (Insecta), Caddisflies. *Zootaxa*, **1668**, 639-698.

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

Jackson JK, Battle JM, White BP, Pilgrim EM, Stein ED, Miller PE, Sweeney BW (2014). Cryptic biodiversity in streams: a comparison of macroinvertebrate communities based on morphological and DNA barcode identifications. *Freshwater Science*, **33**, 312-324.

Johanson KA, Kjer K, Malm T (2009) Testing the monophyly of the New Zealand and Australian endemic family Conoesucidae Ross based on combined molecular and morphological data (Insecta: Trichoptera: Sericostomatoidea). *Zoologica Scripta*, **38**:6, 563-573.

Kjer KM, Blahnik RJ, Holzenthal RW (2001) Phylogeny of Trichoptera (caddisflies): characterization of signal and noise within multiple datasets. *Systematic Biology*, **50**(6), 781-816.

Kjer KM, Blahnik RJ, Holzenthal RW (2002) Phylogeny of caddisflies (Insecta, Trichoptera). *Zoologica Scripta*, **31**(1), 83-91.

Macher JN, Salis RK, Blakemore KS, Tollrian R, Matthaei CD, Leese F (2016) Multiple-stressor effects on stream invertebrates: DNA barcoding reveals contrasting responses of cryptic mayfly species. *Ecological Indicators*, **61**, 159-169.

McCulloch GA, Wallis GP, Waters J M (2010) Onset of glaciation drove simultaneous vicariant isolation of alpine insects in New Zealand. *Evolution*, **64**, 2033-2043.

McGeoch MA (1998) The selection, testing and application of terrestrial insects as bioindicators. *Biological Reviews of the Cambridge Philosophical Society*, **73**:02, 181-201.

McLellan ID (1993) Antartctoperlinae (Insecta: Plecoptera). *Fauna of New Zealand*, **27**, 70 pp.

McLellan I (2008) Checklist of the New Zealand Plecoptera (stoneflies). National Center for aquatic biodiversity and security. https://www.niwa.co.nz/sites/niwa.co.nz/files/import/attachments/stoneflies_june_2008.pdf

Messing J (1983) New M13 vectors for cloning. *Methods in Enzymology*, **101**, 20-78.

Pauls SU, Blahnik RJ, Zhou X, Wardwell CT, Holzenthal RW (2010). DNA barcode data confirm new species and reveal cryptic diversity in Chilean *Smicridea* (Smicridea) (Trichoptera: Hydropsychidae). *Journal of the North American Benthological Society*, **29**:3, 1058-1074.

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

Ratnasingham S, Hebert PD (2013) A DNA-based registry for all animal species: The Barcode Index Number (BIN) System. *PloS one*, **8**:7, e66213.

Shokralla S, Spall JL, Gibson JF, Hajibabaei M (2012) Next-generation sequencing technologies for environmental DNA research.*Molecular ecology*, **21**:8, 1794-1805.

Simon C, Buckley TR, Frati F, Stewart JB, Beckenbach AT (2006) Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics*, **37**, 545-579.

Smith BJ (2014) Trichoptera: A guide to the adult caddisflies from New Zealand. *NIWA Information Series*, No. 83. National Institute of water and Atmospheric Research Ltd: 247.

Smith PJ, Smith BJ (2009). Small-scale population-genetic differentiation in the New Zealand caddisfly *Orthopsyche fimbriata* and the crayfish *Paranephrops planifrons*. New Zealand Journal of Marine and Freshwater Research, 43(3), 723-734.

Stark JD (1993) Performance of the Macroinvertebrate Community Index: Effects of sampling method, sample replication, water depth, current velocity, and substratum on index values, *New Zealand Journal of Marine and Freshwater Research*, **27**:4, 463-478,

Steffen W, Crutzen PJ, McNeill JR (2007) The Anthropocene: are humans now overwhelming the great forces of nature. *AMBIO: A Journal of the Human Environment*, **36**:8, 614-621.

Sweeney BW, Battle JM, Jackson JK, Dapkey T (2011). Can DNA barcodes of stream macroinvertebrates improve descriptions of community structure and water quality? *Journal of the North American Benthological Society*, **30**, 195-216.

Swofford D (2002) PAUP*: phylogenetic analysis using parsimony, version 4.0b10 for Macintosh. Sinauer Associates Inc Publishers, Sunderland.

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

Valentini A, Pompanon F, Taberlet P (2008). DNA barcoding for ecologists. *Trends in Ecology and Evolution*, **24**:2, 110-117.

Vonnemann V, Schrödl M, Klussmann-Kolb A, Wägele H (2005) Reconstruction of the phylogeny of the Opisthobranchia (Mollusca: Gastropoda) by means of 18S and 28S rRNA gene sequences. *Journal of Molluscan Studies*, **71**:2, 113-125.

Ward JB (2003) Checklist of the New Zealand Trichoptera (caddis). National Center for aquatic biodiversity and security. https://www.niwa.co.nz/sites/niwa.co.nz/files/import/attachments/caddis.pdf

Webb JM, Jacobus LM, Funk DH, Zhou X, Kondratieff B, Geraci CJ, DeWalt E, Baird DJ, Richard B, Phillips I, Hebert PDN (2012) A DNA Barcode Library for North American Ephemeroptera: Progress and Prospects. *PLoS ONE,* **7**, e38063.

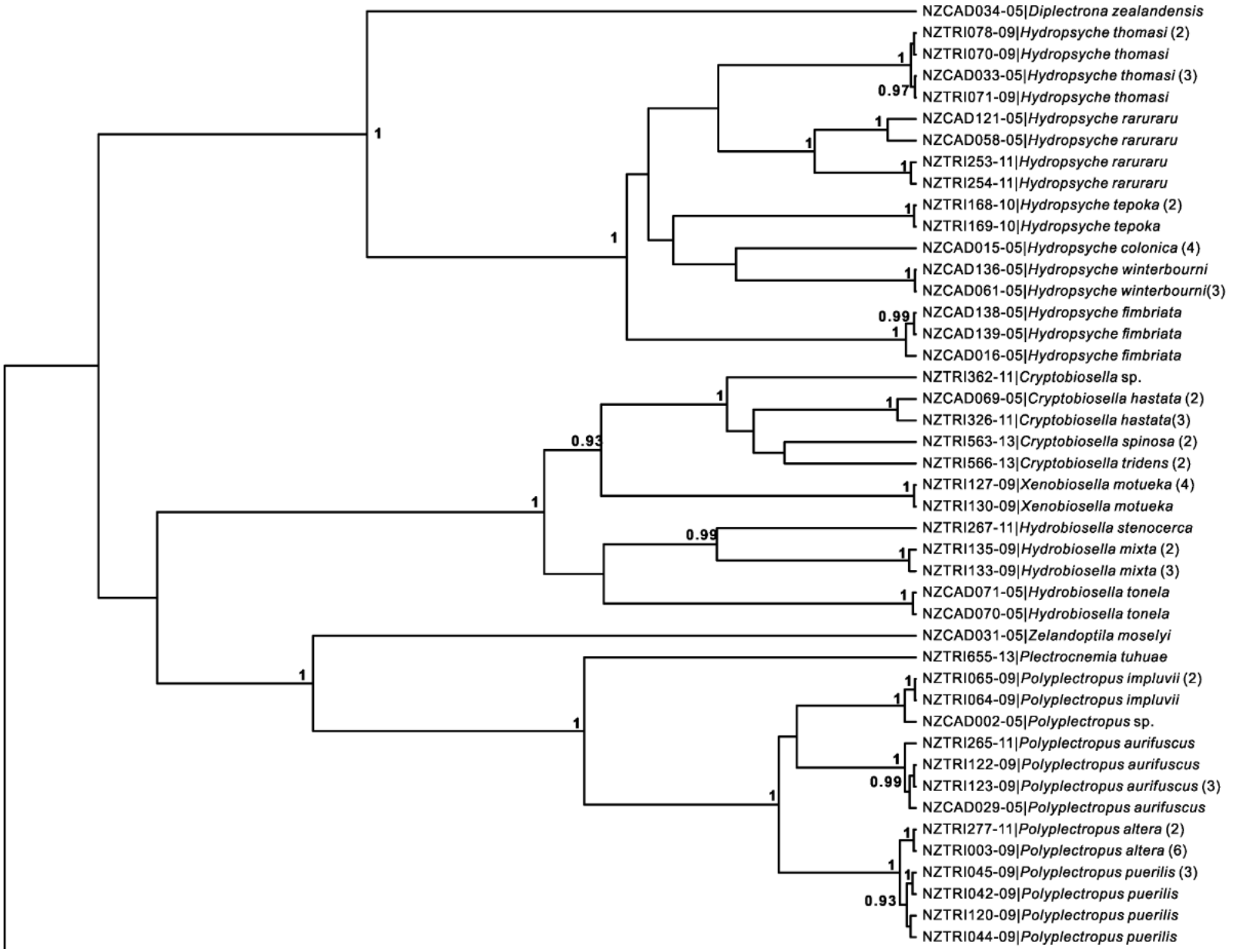
Whitfield JB, Kjer KM (2008) Ancient rapid radiations of insects: challenges for phylogenetic analysis. *Annual Reviews Entomology,* **53**, 449-472.

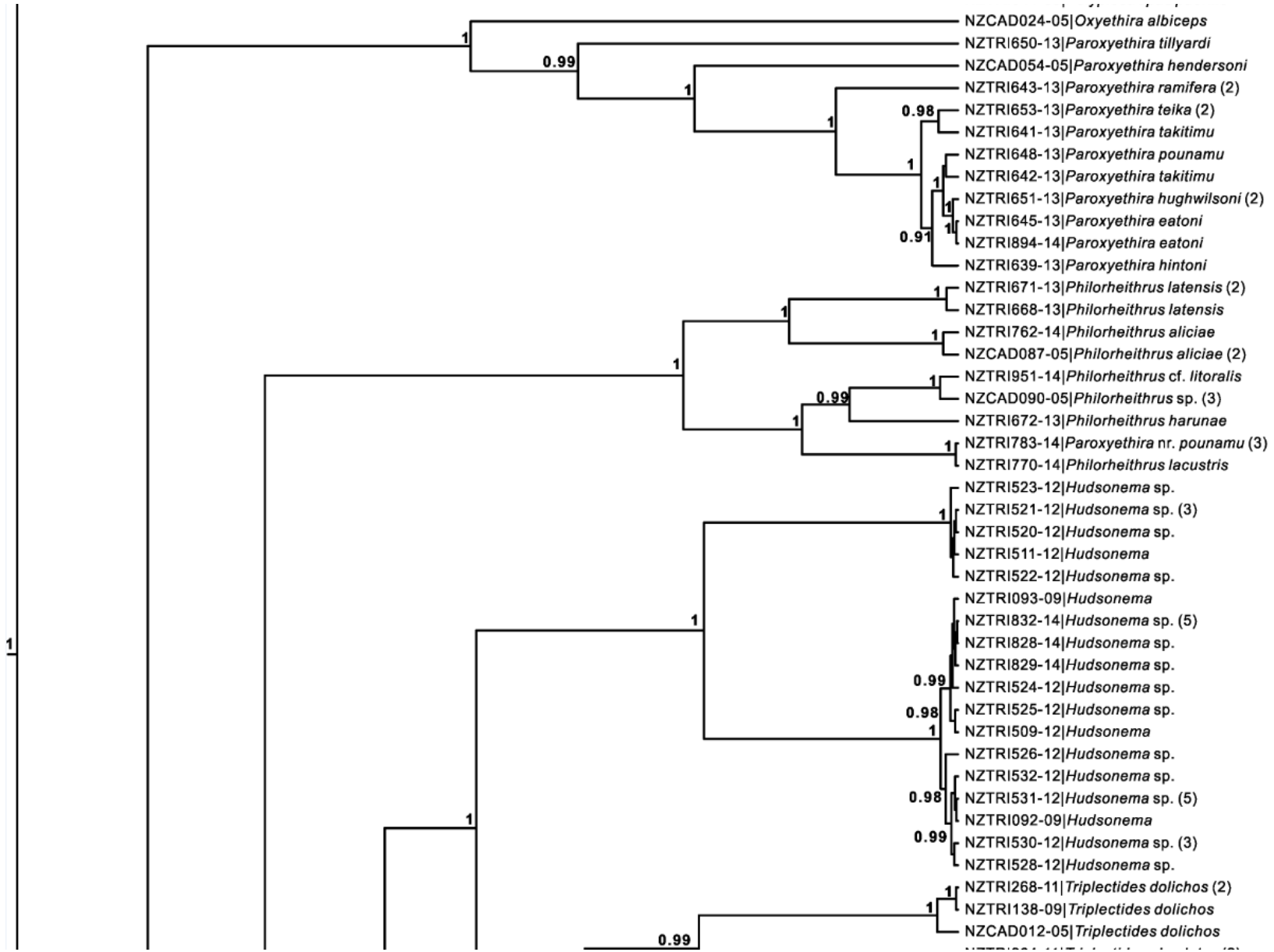
Zhou X, Adamowicz SJ, Jacobus LM, DeWalt RE, Hebert PDN (2009) Towards a comprehensive barcode library for arctic life - *Ephemeroptera*, *Plecoptera*, and *Trichoptera* of Churchill, Manitoba, Canada. *Frontiers in Zoology,* **6**.

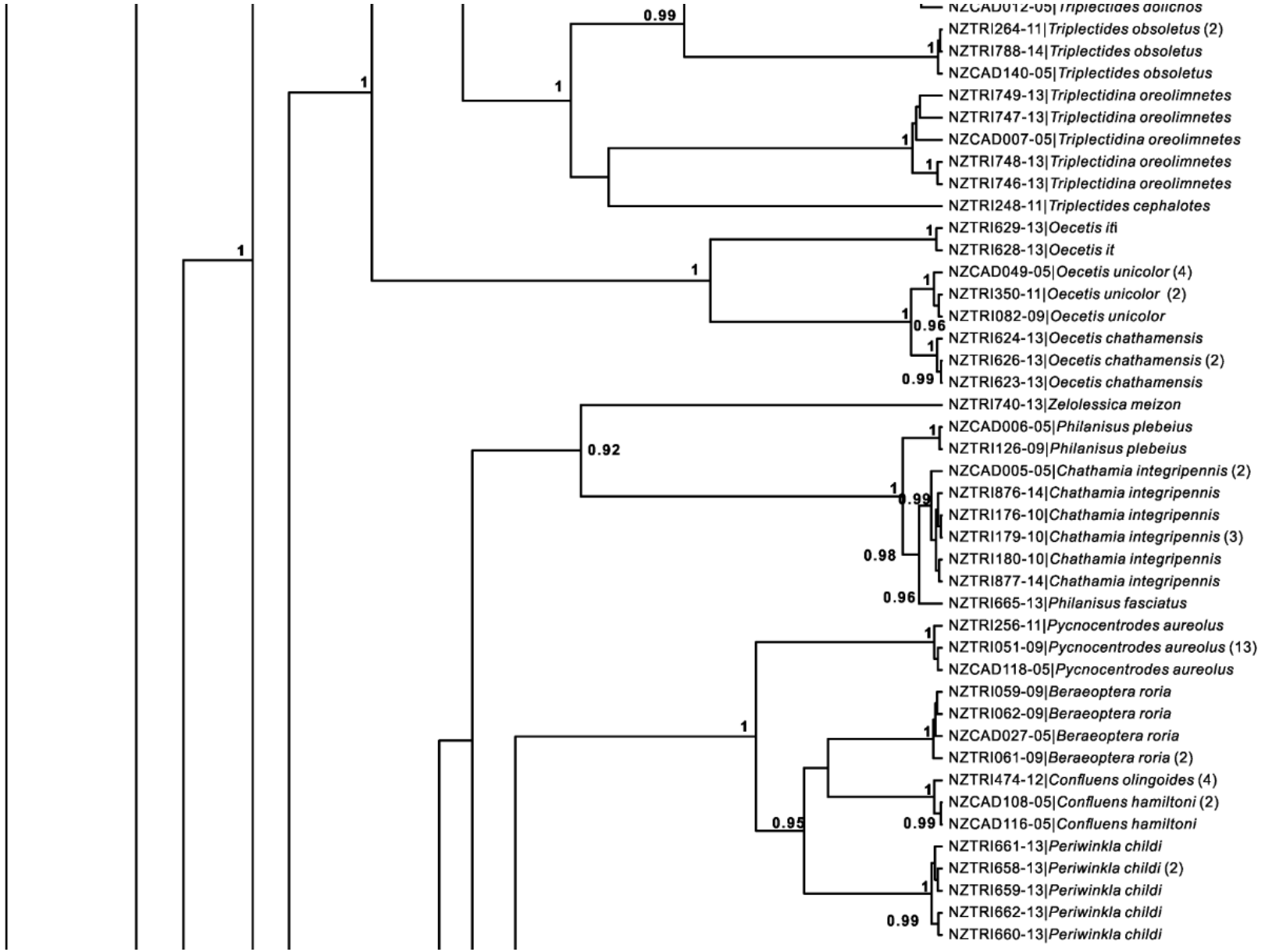
Zhou X, Jacobus LM, DeWalt RE, Adamowicz SJ, Hebert PD (2010) Ephemeroptera, Plecoptera, and Trichoptera fauna of Churchill (Manitoba, Canada): insights into biodiversity patterns from DNA barcoding. *Journal of the North American Benthological Society*, **29**:3, 814-837.

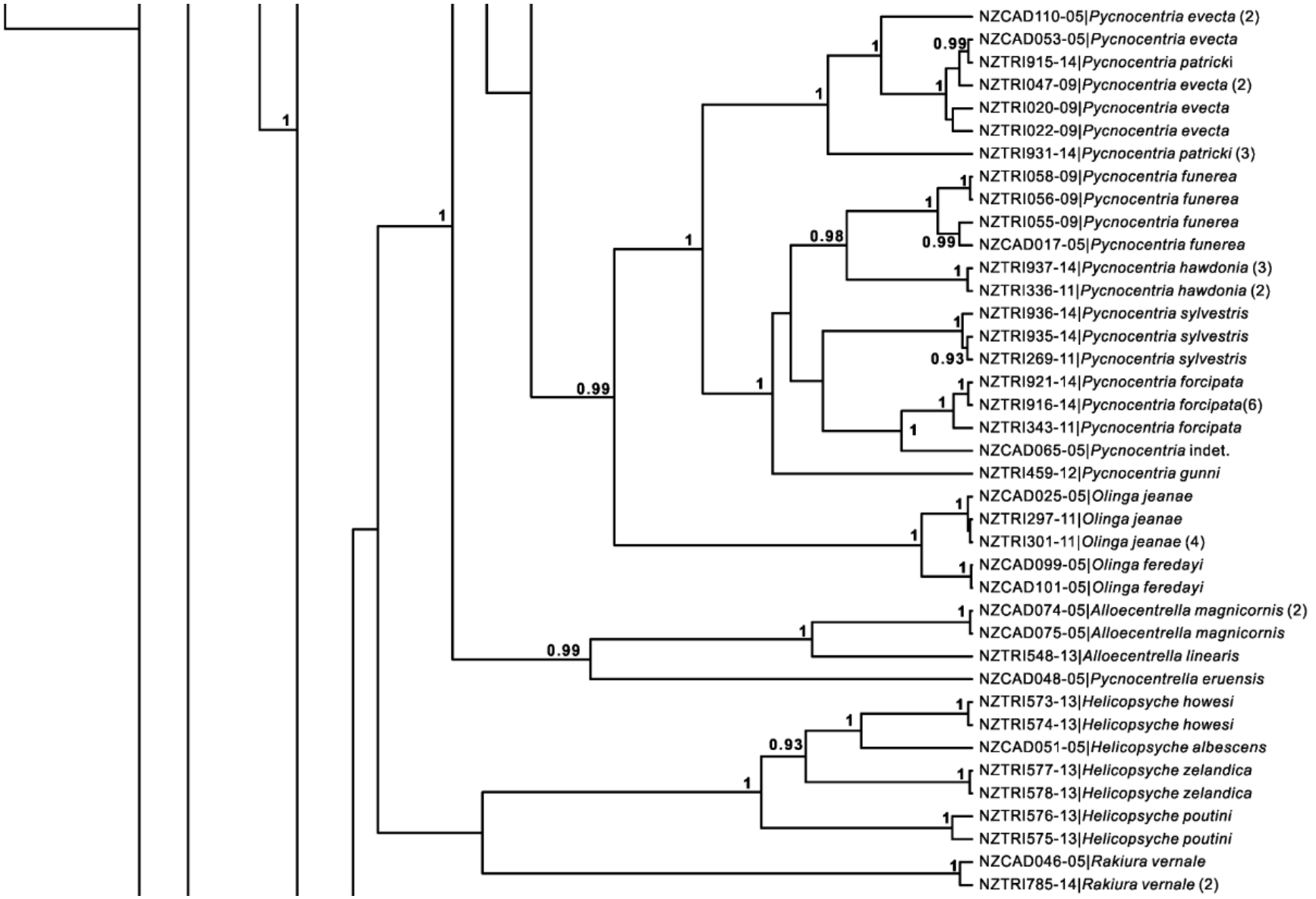
SUPPLEMENTARY MATERIALS

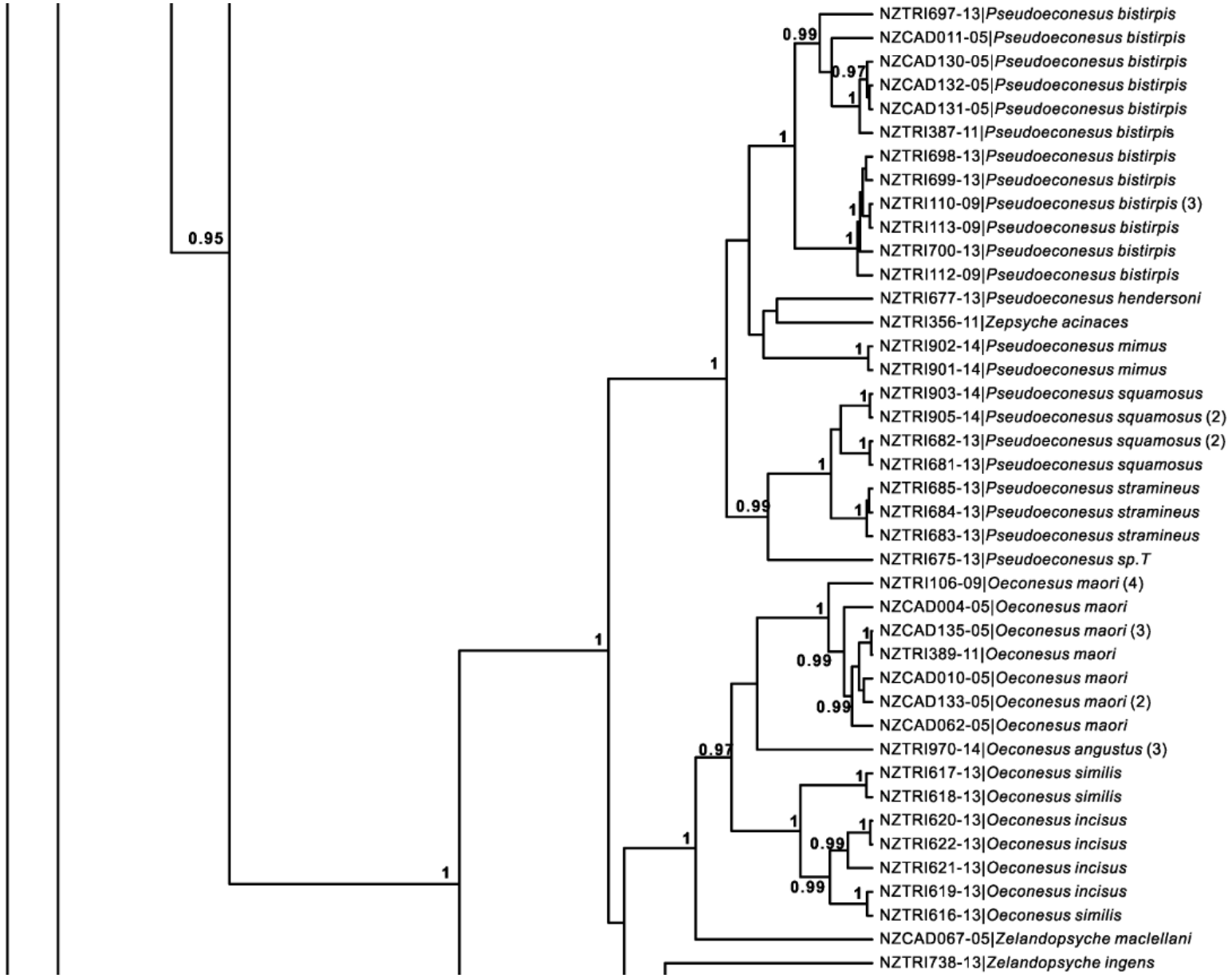
SM1 Bayesian phylogenetic tree of unique New Zealand caddisfly sequences, with the number of sequences indicated in brackets. Posterior probabilities above 0.9 are displayed.

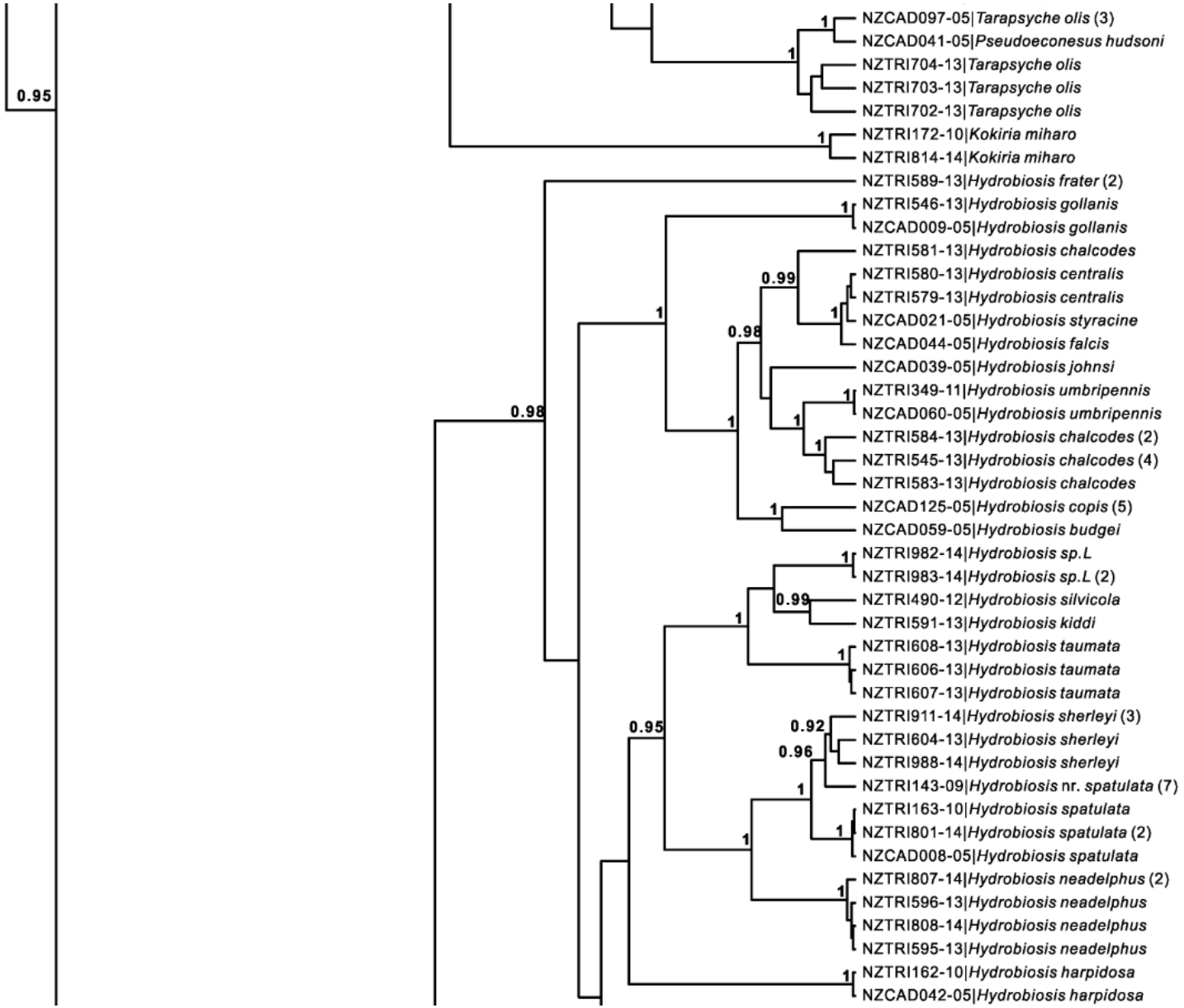

NZCAD034-05|Diplectrona zealandensis
NZTRI078-09|Hydropsyche thomasi (2)
NZTRI070-09|Hydropsyche thomasi
NZCAD033-05|Hydropsyche thomasi (3)
NZTRI071-09|Hydropsyche thomasi
NZCAD121-05|Hydropsyche rarurara
NZCAD058-05|Hydropsyche rarurara
NZTRI253-11|Hydropsyche rarurara
NZTRI254-11|Hydropsyche rarurara
NZTRI168-10|Hydropsyche tepoka (2)
NZTRI169-10|Hydropsyche tepoka
NZCAD015-05|Hydropsyche colonica (4)
NZCAD136-05|Hydropsyche winterbourni
NZCAD061-05|Hydropsyche winterbourni(3)
NZCAD138-05|Hydropsyche fimbriata
NZCAD139-05|Hydropsyche fimbriata
NZCAD016-05|Hydropsyche fimbriata
NZTRI362-11|Cryptobiosella sp.
NZCAD069-05|Cryptobiosella hastata (2)
NZTRI326-11|Cryptobiosella hastata(3)
NZTRI563-13|Cryptobiosella spinosa (2)
NZTRI566-13|Cryptobiosella tridens (2)
NZTRI127-09|Xenobiosella motueka (4)
NZTRI130-09|Xenobiosella motueka
NZTRI267-11|Hydrobiosella stenocerca
NZTRI135-09|Hydrobiosella mixta (2)
NZTRI133-09|Hydrobiosella mixta (3)
NZCAD071-05|Hydrobiosella tonela
NZCAD070-05|Hydrobiosella tonela
NZCAD031-05|Zelandoptila moselyi
NZTRI655-13|Plectrocnemia tuhuae
NZTRI065-09|Polyplectropus impluvii (2)
NZTRI064-09|Polyplectropus impluvii
NZCAD002-05|Polyplectropus sp.
NZTRI265-11|Polyplectropus aurifuscus
NZTRI122-09|Polyplectropus aurifuscus
NZTRI123-09|Polyplectropus aurifuscus (3)
NZCAD029-05|Polyplectropus aurifuscus
NZTRI277-11|Polyplectropus altera (2)
NZTRI003-09|Polyplectropus altera (6)
NZTRI045-09|Polyplectropus puerilis (3)
NZTRI042-09|Polyplectropus puerilis
NZTRI120-09|Polyplectropus puerilis
NZTRI044-09|Polyplectropus puerilis

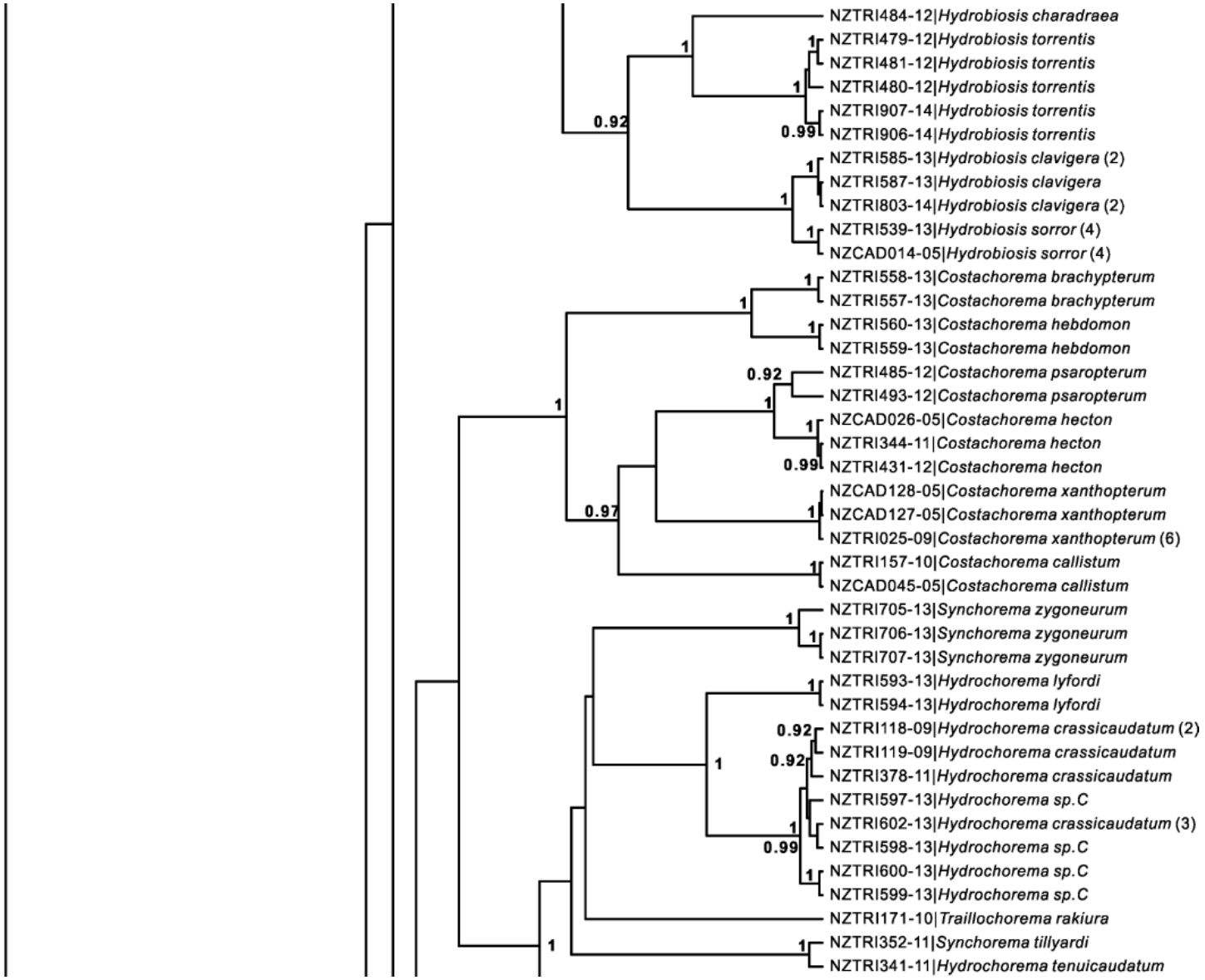

NZCAD024-05|Oxyethira albiceps
NZTRI650-13|Paroxyethira tillyardi
NZCAD054-05|Paroxyethira hendersoni
NZTRI643-13|Paroxyethira ramifera (2)
NZTRI653-13|Paroxyethira teika (2)
NZTRI641-13|Paroxyethira takitimu
NZTRI648-13|Paroxyethira pounamu
NZTRI642-13|Paroxyethira takitimu
NZTRI651-13|Paroxyethira hughwilsoni (2)
NZTRI645-13|Paroxyethira eatoni
NZTRI894-14|Paroxyethira eatoni
NZTRI639-13|Paroxyethira hintoni
NZTRI671-13|Philorheithrus latensis (2)
NZTRI668-13|Philorheithrus latensis
NZTRI762-14|Philorheithrus aliciae
NZCAD087-05|Philorheithrus aliciae (2)
NZTRI951-14|Philorheithrus cf. litoralis
NZCAD090-05|Philorheithrus sp. (3)
NZTRI672-13|Philorheithrus harunae
NZTRI783-14|Paroxyethira nr. pounamu (3)
NZTRI770-14|Philorheithrus lacustris
NZTRI523-12|Hudsonema sp.
NZTRI521-12|Hudsonema sp. (3)
NZTRI520-12|Hudsonema sp.
NZTRI511-12|Hudsonema
NZTRI522-12|Hudsonema sp.
NZTRI093-09|Hudsonema
NZTRI832-14|Hudsonema sp. (5)
NZTRI828-14|Hudsonema sp.
NZTRI829-14|Hudsonema sp.
NZTRI524-12|Hudsonema sp.
NZTRI525-12|Hudsonema sp.
NZTRI509-12|Hudsonema
NZTRI526-12|Hudsonema sp.
NZTRI532-12|Hudsonema sp.
NZTRI531-12|Hudsonema sp. (5)
NZTRI092-09|Hudsonema
NZTRI530-12|Hudsonema sp. (3)
NZTRI528-12|Hudsonema sp.
NZTRI268-11|Triplectides dolichos (2)
NZTRI138-09|Triplectides dolichos
NZCAD012-05|Triplectides dolichos

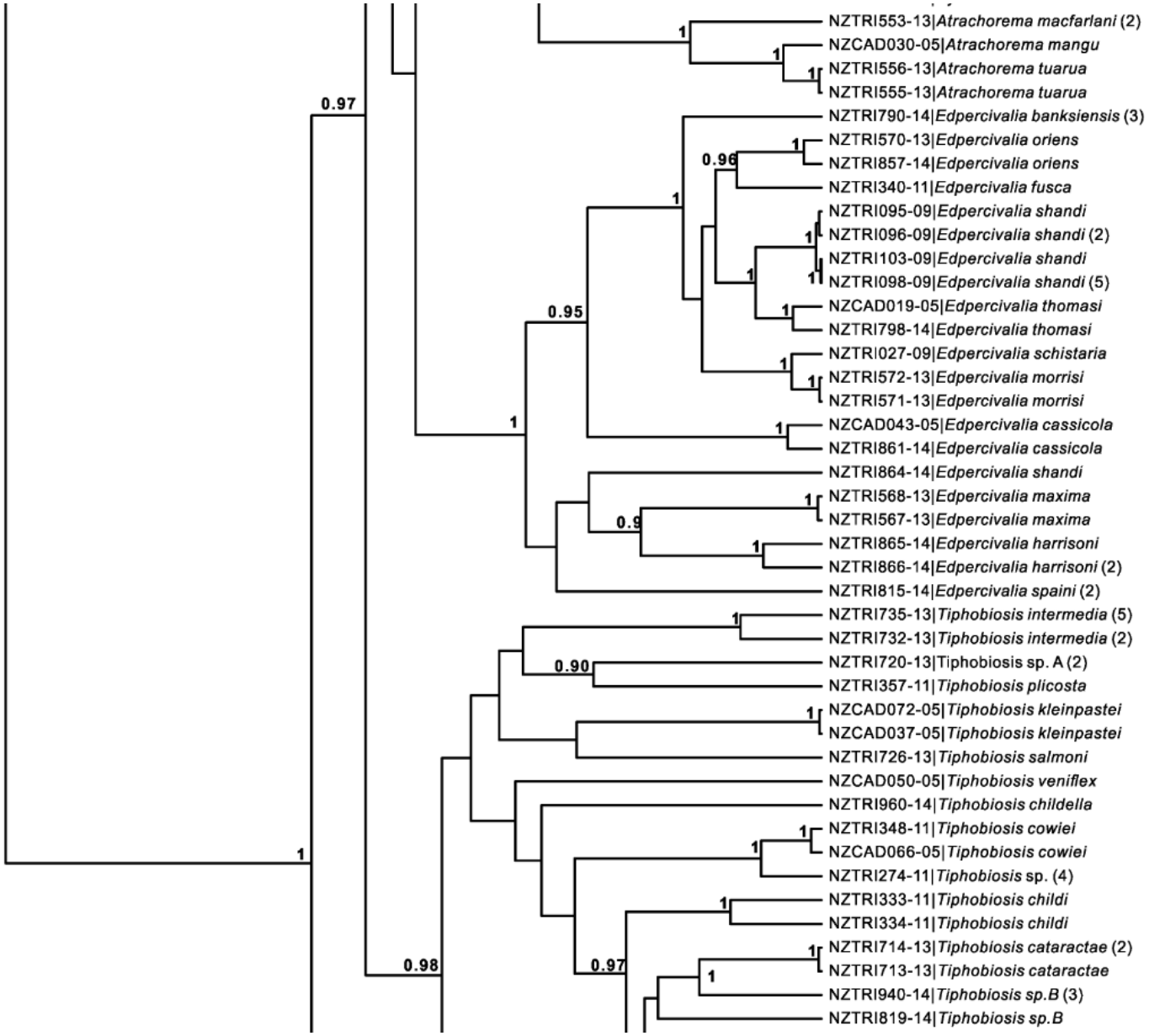

NZCAD012-05|Triplectides dolichos
NZTRI264-11|Triplectides obsoletus (2)
NZTRI788-14|Triplectides obsoletus
NZCAD140-05|Triplectides obsoletus
NZTRI749-13|Triplectidina oreolimnetes
NZTRI747-13|Triplectidina oreolimnetes
NZCAD007-05|Triplectidina oreolimnetes
NZTRI748-13|Triplectidina oreolimnetes
NZTRI746-13|Triplectidina oreolimnetes
NZTRI248-11|Triplectides cephalotes
NZTRI629-13|Oecetis iti
NZTRI628-13|Oecetis it
NZCAD049-05|Oecetis unicolor (4)
NZTRI350-11|Oecetis unicolor (2)
NZTRI082-09|Oecetis unicolor
NZTRI624-13|Oecetis chathamensis
NZTRI626-13|Oecetis chathamensis (2)
NZTRI623-13|Oecetis chathamensis
NZTRI740-13|Zelolessica meizon
NZCAD006-05|Philanisus plebeius
NZTRI126-09|Philanisus plebeius
NZCAD005-05|Chathamia integripennis (2)
NZTRI876-14|Chathamia integripennis
NZTRI176-10|Chathamia integripennis
NZTRI179-10|Chathamia integripennis (3)
NZTRI180-10|Chathamia integripennis
NZTRI877-14|Chathamia integripennis
NZTRI665-13|Philanisus fasciatus
NZTRI256-11|Pycnocentrodes aureolus
NZTRI051-09|Pycnocentrodes aureolus (13)
NZCAD118-05|Pycnocentrodes aureolus
NZTRI059-09|Beraeoptera roria
NZTRI062-09|Beraeoptera roria
NZCAD027-05|Beraeoptera roria
NZTRI061-09|Beraeoptera roria (2)
NZTRI474-12|Confluens olingoides (4)
NZCAD108-05|Confluens hamiltoni (2)
NZCAD116-05|Confluens hamiltoni
NZTRI661-13|Periwinkla childi
NZTRI658-13|Periwinkla childi (2)
NZTRI659-13|Periwinkla childi
NZTRI662-13|Periwinkla childi
NZTRI660-13|Periwinkla childi

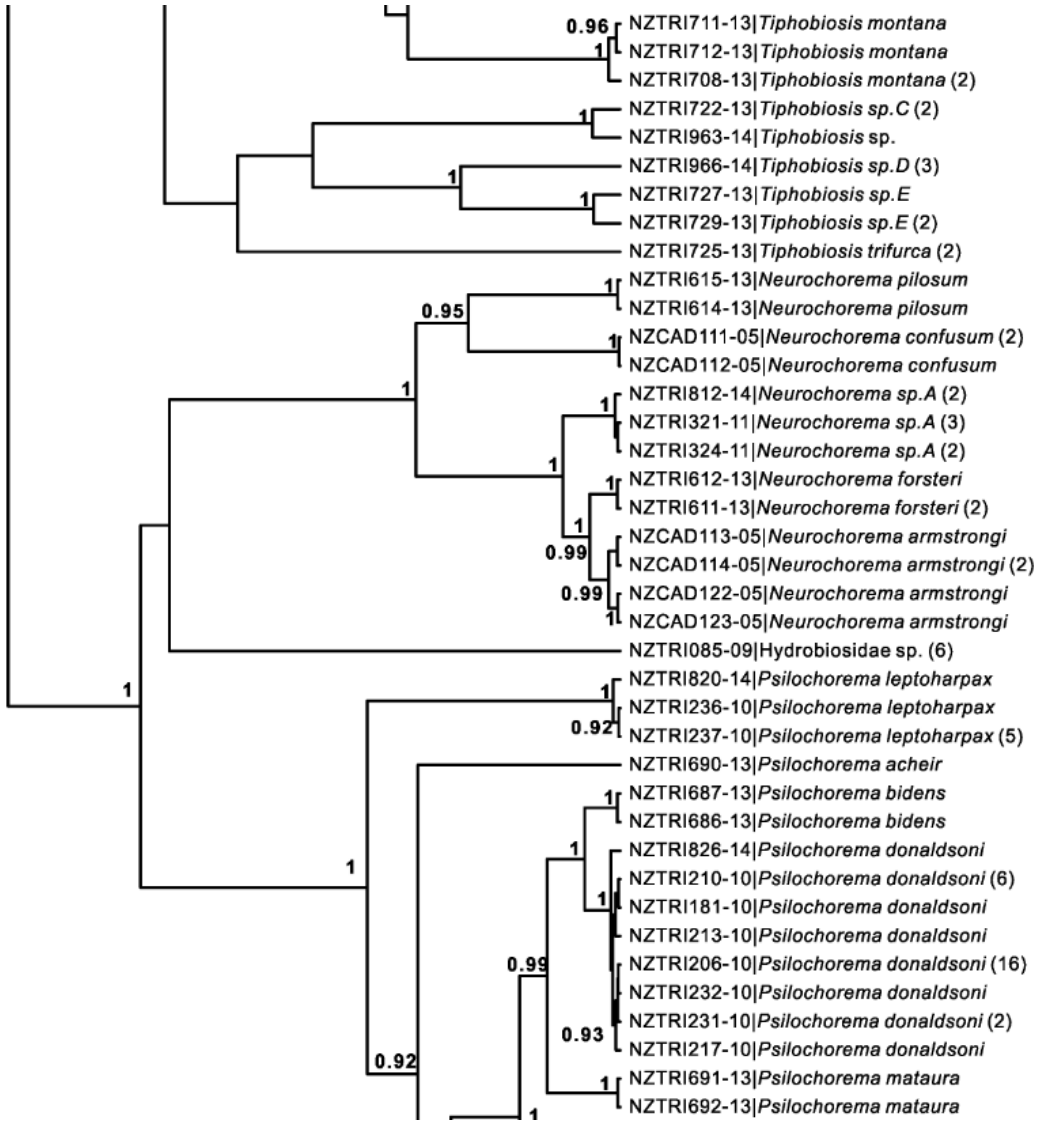

NZCAD110-05|Pycnocentria evecta (2)
NZCAD053-05|Pycnocentria evecta
NZTRI915-14|Pycnocentria patricki
NZTRI047-09|Pycnocentria evecta (2)
NZTRI020-09|Pycnocentria evecta
NZTRI022-09|Pycnocentria evecta
NZTRI931-14|Pycnocentria patricki (3)
NZTRI058-09|Pycnocentria funerea
NZTRI056-09|Pycnocentria funerea
NZTRI055-09|Pycnocentria funerea
NZCAD017-05|Pycnocentria funerea
NZTRI937-14|Pycnocentria hawdonia (3)
NZTRI336-11|Pycnocentria hawdonia (2)
NZTRI936-14|Pycnocentria sylvestris
NZTRI935-14|Pycnocentria sylvestris
NZTRI269-11|Pycnocentria sylvestris
NZTRI921-14|Pycnocentria forcipata
NZTRI916-14|Pycnocentria forcipata(6)
NZTRI343-11|Pycnocentria forcipata
NZCAD065-05|Pycnocentria indet.
NZTRI459-12|Pycnocentria gunni
NZCAD025-05|Olinga jeanae
NZTRI297-11|Olinga jeanae
NZTRI301-11|Olinga jeanae (4)
NZCAD099-05|Olinga feredayi
NZCAD101-05|Olinga feredayi
NZCAD074-05|Alloecentrella magnicornis (2)
NZCAD075-05|Alloecentrella magnicornis
NZTRI548-13|Alloecentrella linearis
NZCAD048-05|Pycnocentrella eruensis
NZTRI573-13|Helicopsyche howesi
NZTRI574-13|Helicopsyche howesi
NZCAD051-05|Helicopsyche albescens
NZTRI577-13|Helicopsyche zelandica
NZTRI578-13|Helicopsyche zelandica
NZTRI576-13|Helicopsyche poutini
NZTRI575-13|Helicopsyche poutini
NZCAD046-05|Rakiura vernale
NZTRI785-14|Rakiura vernale (2)

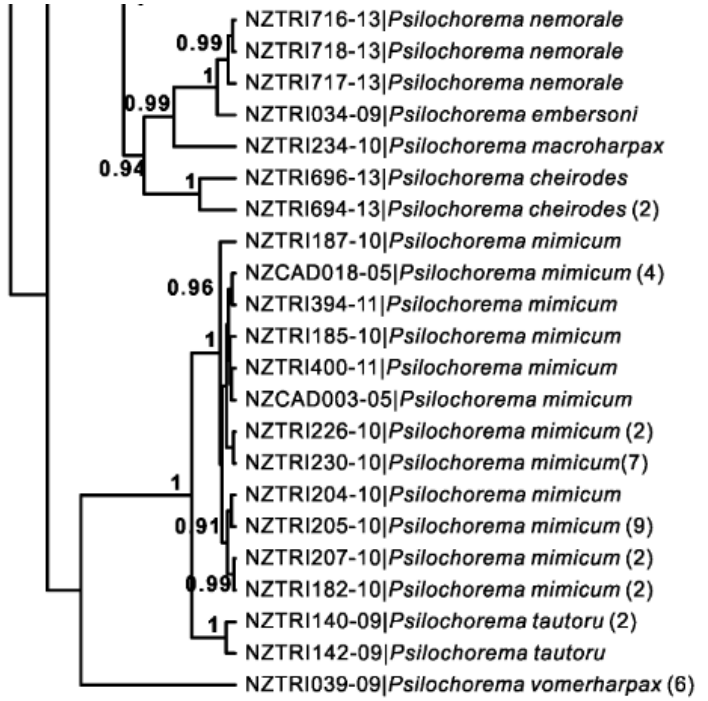

NZTRI697-13|Pseudoeconesus bistirpis
NZCAD011-05|Pseudoeconesus bistirpis
NZCAD130-05|Pseudoeconesus bistirpis
NZCAD132-05|Pseudoeconesus bistirpis
NZCAD131-05|Pseudoeconesus bistirpis
NZTRI387-11|Pseudoeconesus bistirpis
NZTRI698-13|Pseudoeconesus bistirpis
NZTRI699-13|Pseudoeconesus bistirpis
NZTRI110-09|Pseudoeconesus bistirpis (3)
NZTRI113-09|Pseudoeconesus bistirpis
NZTRI700-13|Pseudoeconesus bistirpis
NZTRI112-09|Pseudoeconesus bistirpis
NZTRI677-13|Pseudoeconesus hendersoni
NZTRI356-11|Zepsyche acinaces
NZTRI902-14|Pseudoeconesus mimus
NZTRI901-14|Pseudoeconesus mimus
NZTRI903-14|Pseudoeconesus squamosus
NZTRI905-14|Pseudoeconesus squamosus (2)
NZTRI682-13|Pseudoeconesus squamosus (2)
NZTRI681-13|Pseudoeconesus squamosus
NZTRI685-13|Pseudoeconesus stramineus
NZTRI684-13|Pseudoeconesus stramineus
NZTRI683-13|Pseudoeconesus stramineus
NZTRI675-13|Pseudoeconesus sp.T
NZTRI106-09|Oeconesus maori (4)
NZCAD004-05|Oeconesus maori
NZCAD135-05|Oeconesus maori (3)
NZTRI389-11|Oeconesus maori
NZCAD010-05|Oeconesus maori
NZCAD133-05|Oeconesus maori (2)
NZCAD062-05|Oeconesus maori
NZTRI970-14|Oeconesus angustus (3)
NZTRI617-13|Oeconesus similis
NZTRI618-13|Oeconesus similis
NZTRI620-13|Oeconesus incisus
NZTRI622-13|Oeconesus incisus
NZTRI621-13|Oeconesus incisus
NZTRI619-13|Oeconesus incisus
NZTRI616-13|Oeconesus similis
NZCAD067-05|Zelandopsyche maclellani
NZTRI738-13|Zelandopsyche ingens
0.95
0.99
0.97
1
0.99
0.99
0.97
0.99
0.99


0.95
NZCAD097-05|Tarapsyche olis (3)
NZCAD041-05|Pseudoeconesus hudsoni
NZTRI704-13|Tarapsyche olis
NZTRI703-13|Tarapsyche olis
NZTRI702-13|Tarapsyche olis
NZTRI172-10|Kokiria miharo
NZTRI814-14|Kokiria miharo
NZTRI589-13|Hydrobiosis frater (2)
NZTRI546-13|Hydrobiosis gollanis
NZCAD009-05|Hydrobiosis gollanis
NZTRI581-13|Hydrobiosis chalcodes
NZTRI580-13|Hydrobiosis centralis
NZTRI579-13|Hydrobiosis centralis
NZCAD021-05|Hydrobiosis styracine
NZCAD044-05|Hydrobiosis falcis
NZCAD039-05|Hydrobiosis johnsi
NZTRI349-11|Hydrobiosis umbripennis
NZCAD060-05|Hydrobiosis umbripennis
NZTRI584-13|Hydrobiosis chalcodes (2)
NZTRI545-13|Hydrobiosis chalcodes (4)
NZTRI583-13|Hydrobiosis chalcodes
NZCAD125-05|Hydrobiosis copis (5)
NZCAD059-05|Hydrobiosis budgei
NZTRI982-14|Hydrobiosis sp.L
NZTRI983-14|Hydrobiosis sp.L (2)
NZTRI490-12|Hydrobiosis silvicola
NZTRI591-13|Hydrobiosis kiddi
NZTRI608-13|Hydrobiosis taumata
NZTRI606-13|Hydrobiosis taumata
NZTRI607-13|Hydrobiosis taumata
NZTRI911-14|Hydrobiosis sherleyi (3)
NZTRI604-13|Hydrobiosis sherleyi
NZTRI988-14|Hydrobiosis sherleyi
NZTRI143-09|Hydrobiosis nr. spatulata (7)
NZTRI163-10|Hydrobiosis spatulata
NZTRI801-14|Hydrobiosis spatulata (2)
NZCAD008-05|Hydrobiosis spatulata
NZTRI807-14|Hydrobiosis neadelphus (2)
NZTRI596-13|Hydrobiosis neadelphus
NZTRI808-14|Hydrobiosis neadelphus
NZTRI595-13|Hydrobiosis neadelphus
NZTRI162-10|Hydrobiosis harpidosa
NZCAD042-05|Hydrobiosis harpidosa


NZTRI484-12|Hydrobiosis charadraea
NZTRI479-12|Hydrobiosis torrentis
NZTRI481-12|Hydrobiosis torrentis
NZTRI480-12|Hydrobiosis torrentis
NZTRI907-14|Hydrobiosis torrentis
NZTRI906-14|Hydrobiosis torrentis
NZTRI585-13|Hydrobiosis clavigera (2)
NZTRI587-13|Hydrobiosis clavigera
NZTRI803-14|Hydrobiosis clavigera (2)
NZTRI539-13|Hydrobiosis sorror (4)
NZCAD014-05|Hydrobiosis sorror (4)
NZTRI558-13|Costachorema brachypterum
NZTRI557-13|Costachorema brachypterum
NZTRI560-13|Costachorema hebdomon
NZTRI559-13|Costachorema hebdomon
NZTRI485-12|Costachorema psaropterum
NZTRI493-12|Costachorema psaropterum
NZCAD026-05|Costachorema hecton
NZTRI344-11|Costachorema hecton
NZTRI431-12|Costachorema hecton
NZCAD128-05|Costachorema xanthopterum
NZCAD127-05|Costachorema xanthopterum
NZTRI025-09|Costachorema xanthopterum (6)
NZTRI157-10|Costachorema callistum
NZCAD045-05|Costachorema callistum
NZTRI705-13|Synchorema zygoneurum
NZTRI706-13|Synchorema zygoneurum
NZTRI707-13|Synchorema zygoneurum
NZTRI593-13|Hydrochorema lyfordi
NZTRI594-13|Hydrochorema lyfordi
NZTRI118-09|Hydrochorema crassicaudatum (2)
NZTRI119-09|Hydrochorema crassicaudatum
NZTRI378-11|Hydrochorema crassicaudatum
NZTRI597-13|Hydrochorema sp.C
NZTRI602-13|Hydrochorema crassicaudatum (3)
NZTRI598-13|Hydrochorema sp.C
NZTRI600-13|Hydrochorema sp.C
NZTRI599-13|Hydrochorema sp.C
NZTRI171-10|Traillochorema rakiura
NZTRI352-11|Synchorema tillyardi
NZTRI341-11|Hydrochorema tenuicaudatum


NZTRI553-13|Atrachorema macfarlani (2)
NZCAD030-05|Atrachorema mangu
NZTRI556-13|Atrachorema tuarua
NZTRI555-13|Atrachorema tuarua
NZTRI790-14|Edpercivalia banksiensis (3)
NZTRI570-13|Edpercivalia oriens
NZTRI857-14|Edpercivalia oriens
NZTRI340-11|Edpercivalia fusca
NZTRI095-09|Edpercivalia shandi
NZTRI096-09|Edpercivalia shandi (2)
NZTRI103-09|Edpercivalia shandi
NZTRI098-09|Edpercivalia shandi (5)
NZCAD019-05|Edpercivalia thomasi
NZTRI798-14|Edpercivalia thomasi
NZTRI027-09|Edpercivalia schistaria
NZTRI572-13|Edpercivalia morrisi
NZTRI571-13|Edpercivalia morrisi
NZCAD043-05|Edpercivalia cassicola
NZTRI861-14|Edpercivalia cassicola
NZTRI864-14|Edpercivalia shandi
NZTRI568-13|Edpercivalia maxima
NZTRI567-13|Edpercivalia maxima
NZTRI865-14|Edpercivalia harrisoni
NZTRI866-14|Edpercivalia harrisoni (2)
NZTRI815-14|Edpercivalia spaini (2)
NZTRI735-13|Tiphobiosis intermedia (5)
NZTRI732-13|Tiphobiosis intermedia (2)
NZTRI720-13|Tiphobiosis sp. A (2)
NZTRI357-11|Tiphobiosis plicosta
NZCAD072-05|Tiphobiosis kleinpastei
NZCAD037-05|Tiphobiosis kleinpastei
NZTRI726-13|Tiphobiosis salmoni
NZCAD050-05|Tiphobiosis veniflex
NZTRI960-14|Tiphobiosis childella
NZTRI348-11|Tiphobiosis cowiei
NZCAD066-05|Tiphobiosis cowiei
NZTRI274-11|Tiphobiosis sp. (4)
NZTRI333-11|Tiphobiosis childi
NZTRI334-11|Tiphobiosis childi
NZTRI714-13|Tiphobiosis cataractae (2)
NZTRI713-13|Tiphobiosis cataractae
NZTRI940-14|Tiphobiosis sp.B (3)
NZTRI819-14|Tiphobiosis sp.B
0.97
0.96
0.95
0.9
0.90
0.98
0.97


NZTRI711-13|Tiphobiosis montana
NZTRI712-13|Tiphobiosis montana
NZTRI708-13|Tiphobiosis montana (2)
NZTRI722-13|Tiphobiosis sp.C (2)
NZTRI963-14|Tiphobiosis sp.
NZTRI966-14|Tiphobiosis sp.D (3)
NZTRI727-13|Tiphobiosis sp.E
NZTRI729-13|Tiphobiosis sp.E (2)
NZTRI725-13|Tiphobiosis trifurca (2)
NZTRI615-13|Neurochorema pilosum
NZTRI614-13|Neurochorema pilosum
NZCAD111-05|Neurochorema confusum (2)
NZCAD112-05|Neurochorema confusum
NZTRI812-14|Neurochorema sp.A (2)
NZTRI321-11|Neurochorema sp.A (3)
NZTRI324-11|Neurochorema sp.A (2)
NZTRI612-13|Neurochorema forsteri
NZTRI611-13|Neurochorema forsteri (2)
NZCAD113-05|Neurochorema armstrongi
NZCAD114-05|Neurochorema armstrongi (2)
NZCAD122-05|Neurochorema armstrongi
NZCAD123-05|Neurochorema armstrongi
NZTRI085-09|Hydrobiosidae sp. (6)
NZTRI820-14|Psilochorema leptoharpax
NZTRI236-10|Psilochorema leptoharpax
NZTRI237-10|Psilochorema leptoharpax (5)
NZTRI690-13|Psilochorema acheir
NZTRI687-13|Psilochorema bidens
NZTRI686-13|Psilochorema bidens
NZTRI826-14|Psilochorema donaldsoni
NZTRI210-10|Psilochorema donaldsoni (6)
NZTRI181-10|Psilochorema donaldsoni
NZTRI213-10|Psilochorema donaldsoni
NZTRI206-10|Psilochorema donaldsoni (16)
NZTRI232-10|Psilochorema donaldsoni
NZTRI231-10|Psilochorema donaldsoni (2)
NZTRI217-10|Psilochorema donaldsoni
NZTRI691-13|Psilochorema mataura
NZTRI692-13|Psilochorema mataura


NZTRI716-13|Psilochorema nemorale
0.99
NZTRI718-13|Psilochorema nemorale
1
NZTRI717-13|Psilochorema nemorale
0.99
NZTRI034-09|Psilochorema embersoni
NZTRI234-10|Psilochorema macroharpax
0.94
NZTRI696-13|Psilochorema cheirodes
1
NZTRI694-13|Psilochorema cheirodes (2)
NZTRI187-10|Psilochorema mimicum
0.96
NZCAD018-05|Psilochorema mimicum (4)
NZTRI394-11|Psilochorema mimicum
1
NZTRI185-10|Psilochorema mimicum
NZTRI400-11|Psilochorema mimicum
NZCAD003-05|Psilochorema mimicum
NZTRI226-10|Psilochorema mimicum (2)
NZTRI230-10|Psilochorema mimicum(7)
1
NZTRI204-10|Psilochorema mimicum
0.91
NZTRI205-10|Psilochorema mimicum (9)
NZTRI207-10|Psilochorema mimicum (2)
0.99
NZTRI182-10|Psilochorema mimicum (2)
1
NZTRI140-09|Psilochorema tautoru (2)
NZTRI142-09|Psilochorema tautoru
NZTRI039-09|Psilochorema vomerharpax (6)

Chapter III

Genetic diversity among populations of springtails (Collembola) along the Mackay Glacier ecotone*

*Submitted under the same title as, Beet CR, Hogg ID, Collins GE, Cowan DA, Wall DH, Adams BJ

ABSTRACT

Past and present climate changes are likely to have had major influences on the distribution and abundance of Antarctic terrestrial biota. To assess arthropod distribution and diversity within the Ross Sea region, we examined mitochondrial DNA (COI) sequences for three currently recognised springtail (Collembola) species collected from sites in the vicinity, and to the north of, the Mackay Glacier (77^{o}S). This area represents a transitional zone between two biogeographic regions (North and South Victoria Land). We assessed levels of genetic variability within and among populations of three currently recognised springtail species. We found populations with highly divergent individuals (5-11.3% sequence divergence) for each of the three putative springtail species, suggesting the possibility of cryptic species. Based on molecular clock estimates, these divergent lineages are likely to have been isolated for 3-5 million years. It was during this time that the Western Antarctic Ice Sheet (WAIS) was thought to have completely collapsed potentially facilitating dispersal via rafting. The reformation of the WAIS would have likely led to the isolation of newly established populations with dispersal currently restricted by glaciers and ice-covered areas. Given the currently limited distributions for these genetically distinct populations, any future changes in species' distributions can be easily tracked through the DNA barcoding of springtails from within the Mackay Glacier ecotone.

INTRODUCTION

Antarctica's terrestrial biota are exposed to cold temperatures, low moisture levels and steep chemical gradients (Convey et al. 2008; Convey 2011; Hogg et al. 2006; Hogg et al. 2014; Velasco-Castrillon et al. 2014). These conditions, coupled with repeated glacial cycles over the past 80 million years have shaped a terrestrial landscape characterised by low biodiversity (Adams et al. 2006; Convey 2011). Furthermore, this limited diversity is spread across the highly fragmented 0.32% of Antarctica that is currently considered ice-free (Pugh and Convey 2008). Accelerating rates of climate change, anthropogenic activity and a heightened risk of invasive species introductions all have the potential to severely disrupt current Antarctic biodiversity and thereby ecosystem functioning (Wall 2005; Hogg & Wall 2011; Chown et al. 2012; Hogg et al. 2014). To understand and monitor changing biodiversity patterns, Terrestrial Observation Networks have been proposed to track environmental changes and consequences for the biota (e.g. Levy et al. 2013). Such networks will require a range of suitable locations to adequately characterise any changes.

Ecotones, transitions from one biological community to another, may be ideal sites to study shifts in species distributions as a consequence of global climate changes (Gosz 1993; Risser 1993). The Mackay Glacier in South Victoria Land is situated immediately to the north of the McMurdo Dry Valleys. Here, the landscape changes from one dominated by the expansive Dry Valleys to one of small rocky outcrops isolated by glaciers. These changes are also reflected in the communities of

invertebrates, particularly springtails and mites. For example, only one species of springtail (*Gomphiocephalus hodgsoni*) is known from the McMurdo Dry Valleys, whereas an additional two species (*Antarcticinella monoculata* and *Cryptopygus nivicolis*) are found north of the Dry Valleys (Adams et al. 2006; Hogg et al. 2014). Accordingly, the Mackay Glacier vicinity can be considered an ecotone or transitional zone (*sensu* Risser 1993).

Within the Ross Sea region, three biogeographic zones have been designated (North Victoria Land, South Victoria Land and the Transantarctic mountains) each with three, and in one case four, unique species of springtail for a total of 10 reported species within the region (Salmon 1965; Wise 1967; Adams et al. 2006). However, the basis of this designation was derived from early work utilizing classical morphological taxonomy and had been undertaken primarily in proximity to existing bases or camps, leaving vast areas still largely unexplored (e.g. see Peat et al. 2007). Springtails in particular are key features of Antarctic terrestrial ecosystems as they cycle available below ground nutrients and are also highly sensitive to environmental disturbances, making them ideal bioindicators of climate change (Hopkin 1997; Hogg et al. 2014; Collins & Hogg 2015). Within the wider Mackay Glacier region there are three currently recognised species of springtail: *Gomphiocephalus hodgsoni* (Carpenter 1921), *Antarcticinella monoculata* (Salmon 1965) and the recently reclassified *Cryptopygus nivicolus* (Salmon 1965), formerly *Neocryptopygus* (Greenslade 2015). Of these, *G. hodgsoni* is relatively widespread and well-studied while *C. nivicolus* and *A. monoculata* have received relatively less attention (Adams

et al. 2006; Bennett 2013; Hogg et al. 2014). For example, *A. monoculata* was previously known only from two extremely restricted locations (Salmon 1965; Bennett 2013; Hogg et al. 2014).

The aim of this study was to assess springtail distribution and genetic (COI) diversity within the vicinity of Mackay Glacier. This is a critical first step in establishing baseline levels of diversity and distributional limits against which any future changes in diversity can be detected.

MATERIALS AND METHODS

Individuals of the three species of springtail were collected from 12 sites in the vicinity of the Mackay Glacier covering a north-south range of approximately 120 km (Fig. 1) (SM2). Animals were collected directly using modified aspirators (Stevens and Hogg 2002) or from soil samples that were returned to the laboratory at McMurdo Station with individuals removed via floatation in a sucrose solution (Jenkins 1964). All individuals were preserved in 95% ethanol for subsequent genetic analyses.

Whole individuals were placed into 95 well plates and sent to the Canadian Centre for DNA barcoding (CCDB) for DNA barcoding. Genomic DNA was extracted via the AcroPrep™ PALL Glass Fibre plate method (Ivanova et al. 2006). A 658bp region of the mitochondrial COI gene was then amplified in accordance with standard CCDB protocols (see Ivanova et al. 2006) using the primers C_LepFolF (5'ATTCAACCAATCATAAAGATATTGG-3') and C_LepFolR (5'-TAAACTTCTGGATGTCCAAAAAATCA-3') (Folmer et al. 1994; Hebert et al. 2004; Ivanova et al. 2006). Successfully amplified products were then cleaned using Sephadex® before being sequenced in both directions on an ABI 3730xl DNA analyser. Of the 139 springtails extracted, 95 produced complete 658bp sequences. All photographs, collection and sequence data have been added to Barcode of Life Datasystems (boldsystems.org) under project Antarctic Terrestrial Arthropods (ANTSP).

Sequences were aligned using MUSCLE in Geneious 7.1.9 and trimmed to 527bp to incorporate 69 sequences from a previous study by Bennett (2013) and a *Deuterosminthurus* outgroup (BOLD accession number MHCLB366-09). This alignment was then reduced to the 40 unique haplotypes to save computational time and used in all subsequent analyses except for the haplotype networks. Chi square (x^2) tests conducted in PAUP* 4.0 (Swofford, 2002) were used to determine whether base frequencies were equal among all sites, identify parsimony-informative sites, and designate first, second or third codon positions. The most appropriate model of evolution was determined using jModelTest 2.1.1 (Posada 2008). Bayesian trees were generated using BEAST software v1.7.5 (Drummond et al. 2007). A log normal relaxed clock model and speciation yule process as the tree prior were employed in BEAUTI v1.7.5, with the Markov chain Monte Carlo (MCMC) set at 50,000,000 generations, sampling trees every 5,000 generations. The Bayesian analysis was run in BEAST, with the quality of the results evaluated in TRACER v1.5. A burn in of 500 trees was entered into Tree Annotator v1.7.5 with the final tree visualised in FigTree v1.4.0. Neighbour Joining (NJ) and Maximum Likelihood (ML) analyses were conducted in MEGA v5.05 (Tamura et al. 2011). ML and NJ settings both included 1000 bootstrap replicates. All other settings were set to default options in MEGA. Maximum parsimony (MP) analyses were carried out in MEGA with 1000 bootstrap replicates employed and all other settings set to default options. MEGA was also utilized to create a pairwise distance matrix to calculate intraspecific and interspecific divergences while Barcode Index Numbers (BINs) were assigned by BOLD (Ratnasingham and Hebert 2013) and used as a measure of Molecular

Operational Taxonomic Units (MOTUs). Haplotype networks were developed in TCS v.1.21 using all 164 sequences split into three separate alignments, representing each of the three morphologically identified species.

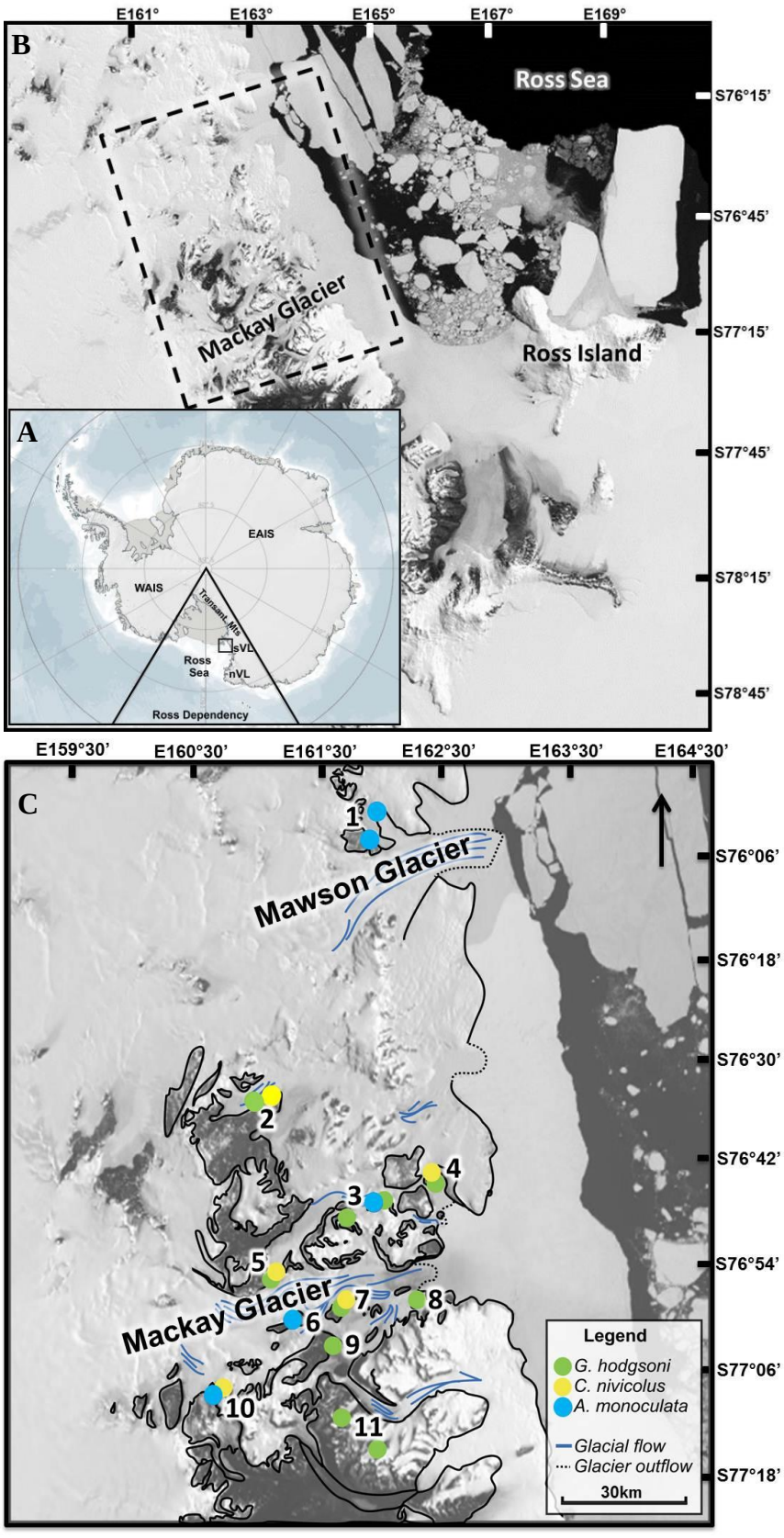

B
E161°
E163°
E165°
E167°
E169°
Ross Sea
S76°15'
S76°45'
S77°15'
S77°45'
S78°15'
S78°45'
Mackay Glacier
Ross Island
A
EAIS
WAIS
Transant. Mts
sVL
nVL
Ross Sea
Ross Dependency
C
E159°30'
E160°30'
E161°30'
E162°30'
E163°30'
E164°30'
S76°06'
S76°18'
S76°30'
S76°42'
S76°54'
S77°06'
S77°18'
Mawson Glacier
Mackay Glacier
1
2
3
4
5
6
7
8
9
10
11
Legend
G. hodgsoni
C. nivicolus
A. monoculata
Glacial flow
Glacier outflow
30km

Figure 1: A = general map of Antarctica showing the study area (black box) within the Ross Sea region of the Ross Dependency. Also pictured are the Western and Eastern Antarctic Ice Sheets (WAIS, EAIS) in addition to North and SouthVictoria Land (nVL, sVL) and the Transantarctic Mountains. Map adapted from the Antarctic Digital Database v6.0, British Antarctic Survey (http://www.add.scar.org/home/add6). B = Map showing the location of the study area around the Mackay Glacier in relation to Ross Island. C = Map of exact collection sites and distribution of the three springtail species. 1= Mt Murray/ Cliff Nunatak, 2= Towle Glacier, 3=Benson Glacier, 4= Tiger Island, 5= Mt Gran, 6= Pegtop Mt, 7= Mt Seuss, 8= The Flatiron, 9= Sperm Bluff, 10= Springtail Pt, 11= St Johns Range.

RESULTS

Of the 527 nucleotide positions used, 347 positions were constant, 14 were variable but uninformative and 166 were parsimony informative. No insertions, deletions or stop codons were detected. There was an overall A-T bias of 62.2% (A = 25.6%, C = 20.9%, G = 16.9%, T = 36.6%). Base frequencies were homogeneous across sequences for all sites (χ^2 = 42.4, df =120, p = 1.0). However, this homogeneity was rejected at variable sites (χ^2 = 173.6, df =120, p = <0.001), informative sites (χ^2 = 183.4, df =120, p = <0.001) and the third codon sites (χ^2 = 148.3, df =120, p = 0.04) which all indicated the presence of an A-T bias. The most appropriate model of evolution was determined as GTR+I+G (-lnL 2,486.799) for Bayesian and maximum likelihood analyses while Tamura-Nei was used for neighbour joining as it allows for unequal base frequencies and multiple substitution types (Simon et al. 2006).

Phylogenetic analyses revealed three previously unreported genetic lineages (MOTUs), one for each of the three known species within the region. NJ, ML, MP and Bayesian analyses all produced concordant results, with all within-species branches having high support (Fig. 2). Maximum intraspecific divergences ranged from 5-11.2%, resulting in a total of seven Barcode Index Numbers (BINs) for the region. Two of these BINs, one each for *G. hodgsoni* and *C. nivicolus*, were found at the same site of Towle Glacier, indicating a potentially important site for arthropod diversity. *A. monoculata* was found at two new locations: (Benson Glacier, Pegtop Mt) and *C. nivicolus* at an additional two locations (Tiger Island, Towle Glacier). The

widespread *G. hodgsoni* previously reported as far north as Mt Murray was not detected or collected at this location.

We examined haplotype networks in order to visualise the fine scale geographic relationships between these 'new' populations (MOTUs) and other Mackay Glacier groups. From the 121 *G. hodgsoni* sequences, we found 27 unique haplotypes. Overall, these haplotypes exhibited high levels of genetic connectivity between populations throughout the study area, albeit with some subtle geographic structuring (Fig. 3A). Four of the 27 haplotypes occurred in more than one location, with the most common haplotype (Gh1) found at four sites up to an estimated 55 km apart. However, two populations of *G. hodgsoni* were found to be highly distinct with the newly identified population at Towle Glacier (Gh23, 24) 28 mutational steps (7.3% sequence divergence) and more than 45 km away from Mackay Glacier populations. In contrast, the other divergent population (Gh25), previously identified at Mt Gran (within the Mackay Glacier) by Bennett (2013), was 32 mutational steps and 7.3% divergent from other nearby populations.

Cryptopygus nivicolus haplotypes exhibited broadly similar patterns to that of *G. hodgsoni*, although on a smaller scale owing to their more restricted distribution (Fig. 3B). A total of seven haplotypes and three BINs were identified from the 21 *C. nivicolus* sequences obtained. The main haplotype (Cn1) was found in three locations separated by distances of up to 30 km with the distinct Towle Glacier population (Cn5) 24 mutational steps (5% sequence divergence) away from Mackay haplotypes.

Antarcticinella monoculata were found in very low numbers at all sites where they were present with the exceptions of Mt Murray and Cliff Nunatak (Fig. 3C). Of the 23 individuals we sequenced from these two nearby locations (approximately 6 km apart), all had the same haplotype (Am1). This haplotype was also 50 mutational steps away from the *A. monoculata* collected from the more southern Mackay Glacier locations. The phylogenetic analyses estimated that these groups were 11.2% divergent. *G. hodgsoni* and *C. nivicolus* both had high levels of genetic variability within the immediate Mackay Glacier area (77°S) including a distinct haplotype for each species (Mt Gran haplotype Gh25 at 37 steps, Springtail Point haplotypes Cn6 and Cn7 at 17 steps). In contrast, *A. monoculata* populations found throughout the Mackay Glacier area (excluding northern haplotype) were all very similar, only separated by a maximum of six base pairs.

The application of standard molecular clock calibrations of 1.5 to 2.3 sequence divergence per million years (*sensu* Knowlton et al. 1993; Brower et al. 1994; Quek et al. 2004) suggested that individuals from three of the populations from Towle Glacier and Mt Murray (Gh23 & 24, Cn5, Am1) diverged within the last 5 my. The *C. nivicolus* haplotypes diverged the most recently (2.2 - 3.3 mya) followed by *G. hodgsoni* at 3.2 – 4.9 mya with the most divergent group of *A. monoculata* estimated to have been isolated between 4.9 and 7.5 mya.

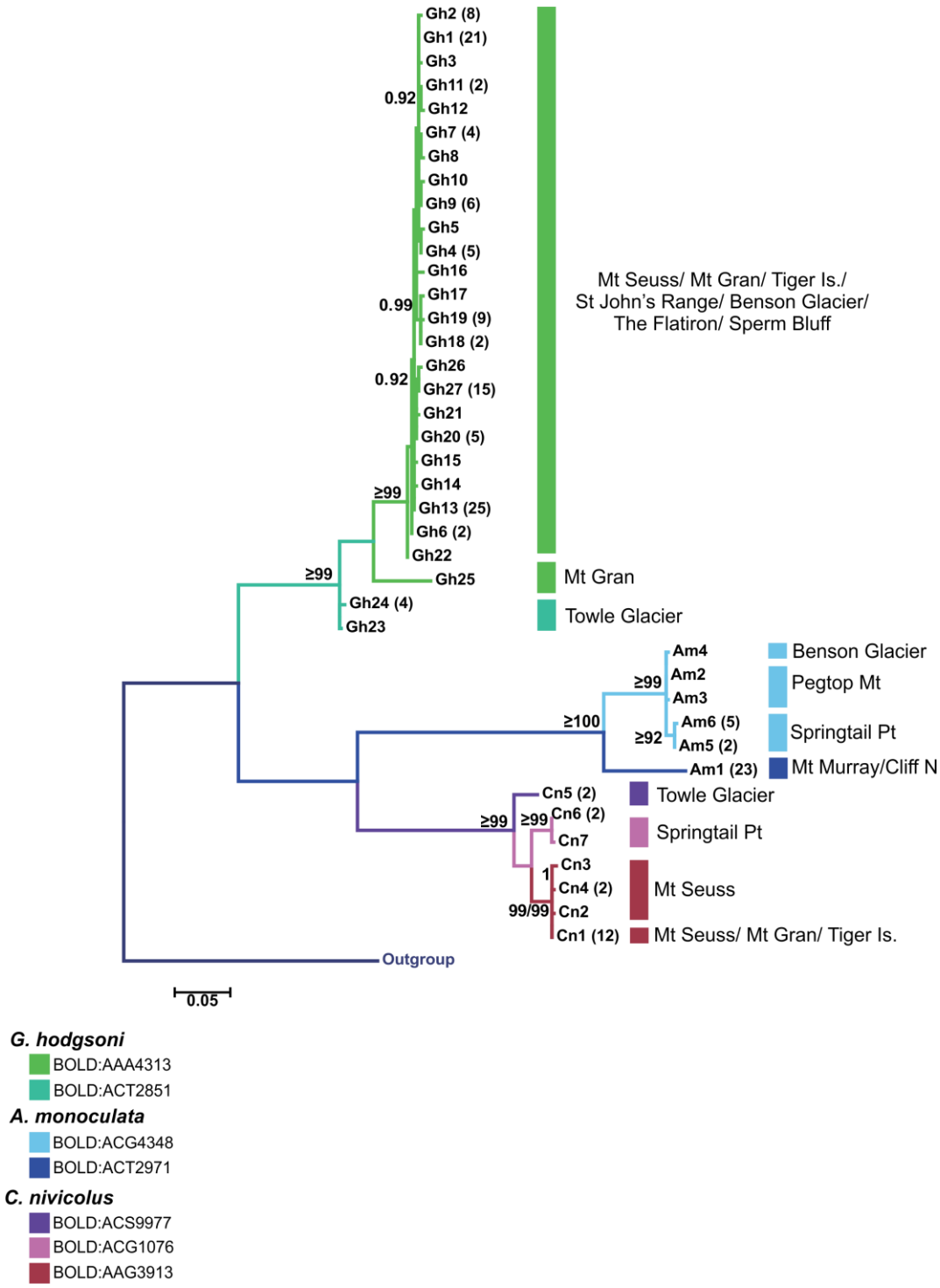
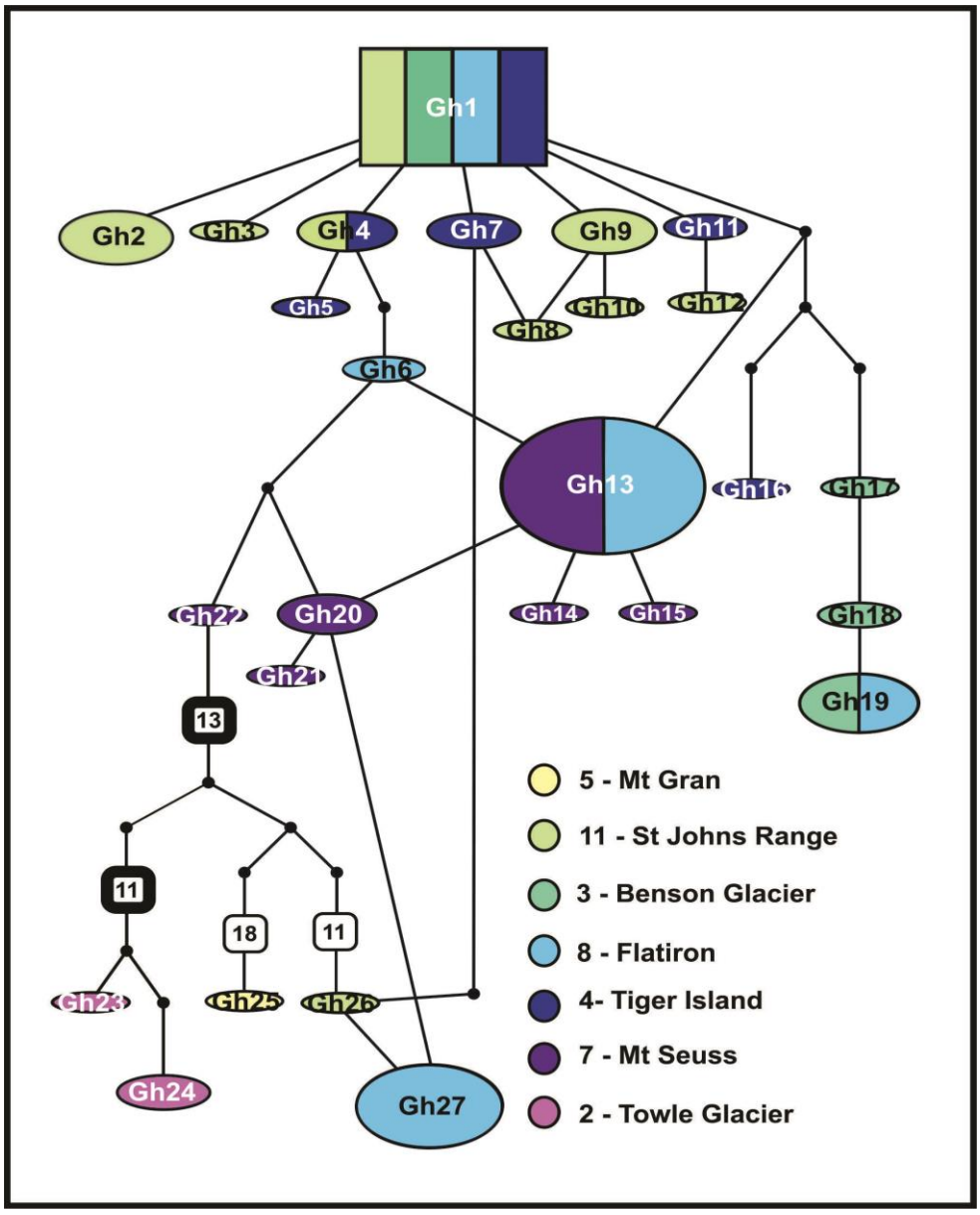

Gh2 (8)
Gh1 (21)
Gh3
Gh11 (2)
0.92
Gh12
Gh7 (4)
Gh8
Gh10
Gh9 (6)
Gh5
Gh4 (5)
Gh16
Gh17
0.99
Gh19 (9)
Gh18 (2)
Gh26
0.92
Gh27 (15)
Gh21
Gh20 (5)
Gh15
Gh14
≥99
Gh13 (25)
Gh6 (2)
Gh22
≥99
Gh25
Gh24 (4)
Gh23
Mt Seuss/ Mt Gran/ Tiger Is./
St John's Range/ Benson Glacier/
The Flatiron/ Sperm Bluff
Mt Gran
Towle Glacier
Am4
Am2
≥99
Am3
≥100
Am6 (5)
≥92
Am5 (2)
Am1 (23)
Benson Glacier
Pegtop Mt
Springtail Pt
Mt Murray/Cliff N
Cn5 (2)
Towle Glacier
≥99
≥99
Cn6 (2)
Cn7
Springtail Pt
Cn3
1
Cn4 (2)
Mt Seuss
99/99
Cn2
Cn1 (12)
Mt Seuss/ Mt Gran/ Tiger Is.
Outgroup
0.05
G. hodgsoni
BOLD:AAA4313
BOLD:ACT2851
A. monoculata
BOLD:ACG4348
BOLD:ACT2971
C. nivicolus
BOLD:ACS9977
BOLD:ACG1076
BOLD:AAG3913

Figure 2: Collated phylogenetic tree (NJ base tree) of 40 unique haplotypes for three springtail species including a *Deuterosminthurus* outgroup (BOLD accession number MHCLB366-09). The total number of sequences per haplotype is indicated in brackets. Support values over 90/ 0.9 are displayed in order of ML bootstrap values/ Bayesian posterior probabilities/ NJ bootstrap values/ MP bootstrap values. Where all support values are present ≥ is used. Tree is coloured according to the seven BINs present with bars indicating the location where specific haplotypes were collected.


Gh1
Gh2
Gh3
Gh4
Gh7
Gh9
Gh11
Gh5
Gh10
Gh12
Gh8
Gh6
Gh13
Gh16
Gh17
Gh22
Gh20
Gh14
Gh15
Gh18
Gh21
Gh19
13
11
18
11
Gh23
Gh25
Gh26
Gh24
Gh27
5 - Mt Gran
11 - St Johns Range
3 - Benson Glacier
8 - Flatiron
4- Tiger Island
7 - Mt Seuss
2 - Towle Glacier

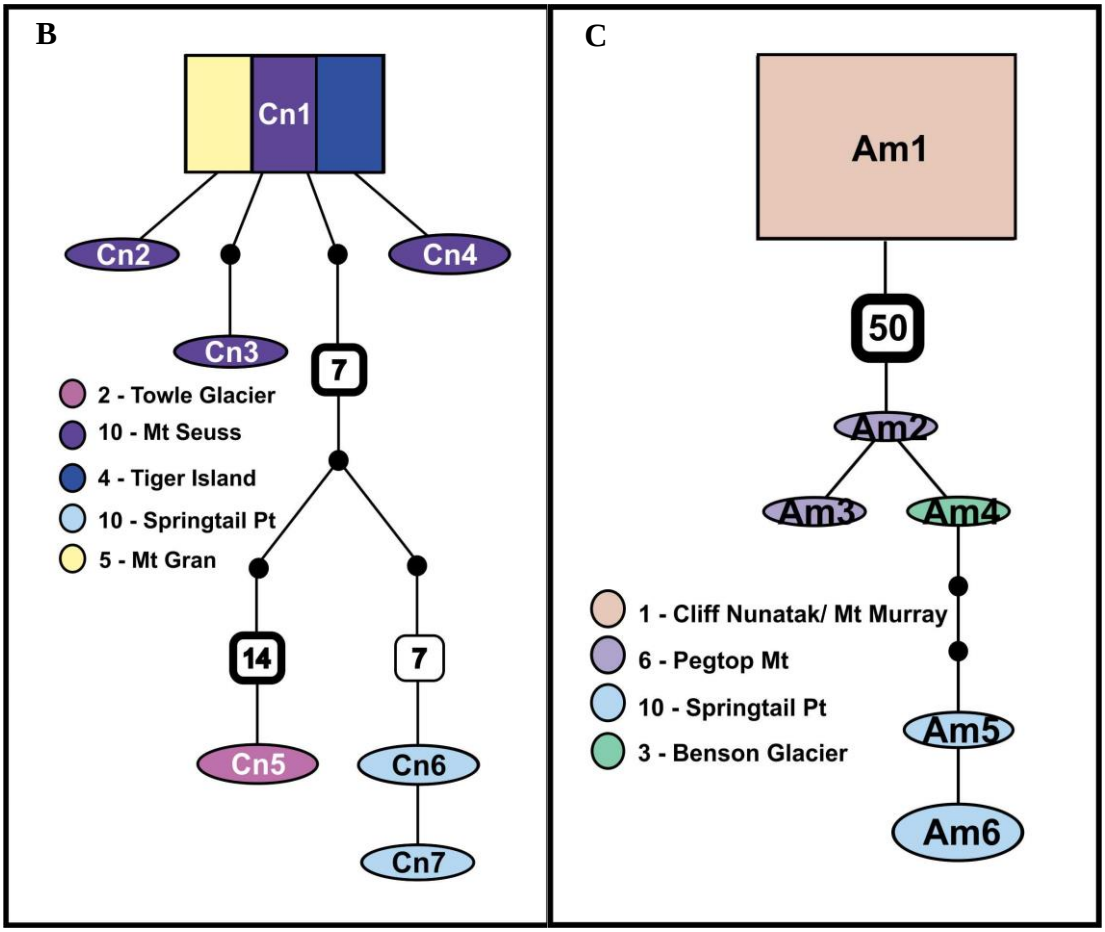


Figure 3: A = *Gomphiocephalus hodgsoni* network of 27 unique haplotypes coloured according to collection locations. B = *Cryptopygus nivicolus* network of seven unique haplotypes coloured according to collection locations. C = *Antarcticinella monoculata* network of six unique haplotypes coloured according to collection locations. Numbers beside location names refer to those pictured on Figure 1C. Numbers within boxes indicate the number of mutational steps separating haplotypes while black dots indicate a single mutational step. Haplotypes within squares (Gh1, Cn1, and Am1) are the most derived.

DISCUSSION

The mtDNA COI sequences revealed considerably greater diversity than previously known in the Ross Sea region of Antarctica. In particular, we found several highly divergent lineages of springtails including three previously unrecognised BINs for the three putative species to give a total of seven BINs within the wider Mackay Glacier area. The high levels of divergence suggest the presence of potentially cryptic species and that diversity in Antarctica whilst comparatively depauperate in the global sense, is much more diverse than once thought (Griffiths et al. 2010; Chown et al. 2015). The use of molecular techniques such as DNA barcoding to study diversity has been particularly helpful in detecting genetically distinct populations (e.g. Stevens and Hogg 2003, 2006). A study by McGaughran et al. (2010) found two very distinct clades within the springtail *Cryptopygus antarcticus* while Mortimer et al. (2011) identified two separate lineages within Western Antarctic Peninsula populations of the mite *Alaskozetes antarcticus.* Similarly, Pisa et al. (2014) found three lineages within the moss *Bryum argenteum*. Such high levels of variability amongst springtails and other Antarctic biota could be the result of a number of factors including genetic drift, environmental heterogeneity and potential selection for genetic polymorphisms as a mechanism for surviving extreme conditions (Fanciulli et al. 2001).

Habitat fragmentation restricts gene flow, leading to allopatric speciation, as was likely for the unique BINs from Towle Glacier. The genetic uniformity of all of the northern *A. monoculata* individuals could indicate that the population has also been derived allopatrically with genetic drift leading to a reduction in diversity and fixation of the haplotype (Costa et al. 2013). Alternatively, the population could be comparatively new, resulting from a founder effect from a more diverse population nearby. Although there was a general trend of increasing divergence with distance, similar to observations of Fanciulli et al. (2001), there were a number of *G. hodgsoni* haplotypes that were shared between locations up to 55 km apart demonstrating that populations currently exist in both sympatry and allopatry.

The genetically distinct populations (BINs) were estimated to have diverged within the last 5 mya (range: 2.2-7.5 my). Whilst there is always some controversy surrounding rates of molecular evolution, a number of studies have similarly reported clades of Antarctic invertebrates diverging within the Pliocene (Hawes et al. 2010; McGaughran et al. 2010; Bennett 2013). It was during this time that the Western Antarctic Ice Sheet (WAIS) was thought to have completely collapsed (Pollard and Deconto 2009). This would have resulted in sea levels rising and increased dispersal opportunities for Collembola via rafting in meltwater streams and open sea-ways before the WAIS eventually reformed and glaciers isolated newly established populations (Pollard and Deconto 2009; Hawes 2011; McGaughran et al. 2010; Hogg et al. 2014). This could explain why some *G. hodgsoni* haplotypes are shared in locations that are up to 55 km apart and currently isolated by glaciers. Springtails

generally have poor dispersal capabilities. Unlike Antarctic microinvertebrates such as nematodes and tardigrades, which can enter an anhydrobiotic phase that facilitates aeolian transport, Collembola are very prone to desiccation due to their permeable cuticle (Freckman and Virginia 1998; Hogg et al. 2014).

The northern Mackay region had been under-sampled in the past with this study therefore providing the first records of species from many of the sites visited. The sampling of *A. monoculata* at Mt Murray and Cliff Nunatak represented the first time they have been collected there since the original collections in 1958 (Salmon 1965; Wise 1967). Interestingly, *G. hodgsoni* was not collected from the previously recorded northern location (Mt Murray). This absence suggests that species may have altered their distributions within the past 50 years (e.g. Stevens and Hogg 2002). It is possible that some springtails have been essentially living at their environmental limit and that recent climatic changes have led to detectable range contractions. Alternatively, their absence from Mt Murray could have been due to stochastic events. The distribution of springtails is largely governed by the paucity of ice-free habitat, with abiotic factors such as the bioavailability of water further structuring what available habitat is occupied (Hogg et al. 2006; Sinclair & Stevens 2006).

Historical glacial cycles thus appear to have influenced the current distribution of springtail populations with habitat fragmentation and population dynamics refining these assemblages. Improving our understanding of genetic variability is fundamental to preserving maximal springtail diversity and thereby the evolutionary potential of natural systems and populations to respond to an ever changing environment (Adams et al. 2006; Hogg et al. 2006; Hogg et al. 2014). Given the presence of numerous unique and currently geographically restricted BINs, any future changes in species' distributions can be easily tracked through the DNA barcoding of individuals. This will thereby enhance our capacity to detect subtle biological responses resulting from gradual climate changes.

ACKNOWLEDGEMENTS

We are grateful to Antarctica New Zealand for logistic support and the New Zealand Antarctic Research Institute (NZARI) for financial support. CB was the recipient of a New Zealand Post Antarctic Scholarship, a University of Waikato Environmental Research Institute Scholarship and a Waikato Graduate Women Educational Trust Scholarship. Sequencing at the Canadian Centre for DNA barcoding was supported through funding to the International Barcode of Life Project (iBOL) from the Ontario Genomics Institute (2008-OGI-ICI-03), Genome Canada, the Ontario Ministry of Research and Innovation, and the Natural Sciences and Engineering Research Council of Canada. BJA and DHW were supported by McMurdo LTER NSF OPP grant 1115245. DAC gratefully acknowledges the financial support of the South African NRF SANAP program.

LITERATURE CITED

Adams, B.J., Bardgett, R.D., Ayres, C., Wall, D.H., Aislabie, J., Bamforth, S., Bargagli R., Cary, S.C., Cavacini, P., Connell, L., Convey, P., Fell, J.W., Frati, F., Hogg, I.D., Newsham, K.K., O'Donnell, A., Russell, N., Seppelt, R.D., and Stevens, M.I. 2006. Diversity and distribution of Victoria Land biota. Soil Biol. Biochem., **38**: 3003–3018. doi:10.1016/j.soilbio.2006.04.030

Bennett, K.R. 2013. Impacts of the Pliocene and Pleistocene glaciations on genetic diversity among New Zealand and Antarctic arthropods. M.Sc. thesis, Department of Biological Sciences, University of Waikato, Hamilton, NZ.

Brower, A. 1994. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. Proc. Natl. Acad. Sci. U.S.A., **91**: 6491–5.

Carpenter, G.H. 1921. Insecta Part 1. Collembola British Antarctic ("Terra Nova") Expedition, 1910, Natural History Reports, Zoology, Volume 3. British Museum (Natural History), London.

Chown, S.L., Huiskes, A.H.L., Gremmen, N.J. M., Lee, J.E., Terauds, A., Crosbie, K., Frenot, Y., Hughes, K.A., Imura, S., Kiefer, K., Lebouvier, M., Raymond, B., Tsujimoto, M., Ware, C., Van de Vijver, B., and Bergstrom, D.M. 2012. Continent-wide risk assessment for the establishment of nonindigenous species in Antarctica. PNAS, **109** (13): 4938-4943. doi:10.1073/pnas.1119787109.

Chown, S.L., Clarke, A., Fraser, C.I., Cary, C.S., Moon, K.L., and McGeoch, M.A. 2015. The changing form of Antarctic biodiversity. Nature, **522**: 431–438. doi:10.1038/nature14505.

Collins, G.E. and Hogg, I.D. 2015. Temperature-related activity of *Gomphiocephalus hodgsoni* (Collembola) mitochondrial DNA (COI) haplotypes in Taylor Valley, Antarctica. Polar Biol., 1-11. doi: 10.1007/s00300-015-1788-7

Convey, P., Gibson, J.A.E., Hillenbrand, C.D., Hodgson, D.A., Pugh, P.J.A., Smellie, J.L. and Stevens, M.I. 2008. Antarctic terrestrial life – challenging the history of the frozen continent? Biol. Rev., **83**: 103–117. doi: 10.1111/j.1469185X.2008.00034.x

Convey, P. 2011. Antarctic terrestrial biodiversity in a changing world. Polar Biol., **34**:1629–1641.

Costa, D., Timmermans, M.J.T.N., Sousa, J.P., Ribeiroa, R., Roelofsd, D., and Van Straalend, N.M. 2013. Genetic structure of soil invertebrate populations: Collembolans, earthworms and isopods. Appl. Soil Ecol., **68**: 61–66. doi:10.1016/j.apsoil.2013.03.003.

Drummond, A.J., Ashton, B., Buxton, S., Cheung, M., Cooper, A., Duran, C., Field, M., Heled, J., Kearse, M., Markowitz, S., Moir, R., Stones-Havas, S., Sturrock, S., Thierer, T. and Wilson, A. 2010. Geneious v5.1. Biomatters Ltd, Auckland. http://www.geneious.com

Fanciulli, P.P., Summa, D., Dallai, R., and Frati, F. 2001. High levels of genetic variability and population differentiation in *Gressittacantha terranova* (Collembola, Hexapoda) from Victoria Land, Antarctica. Antarct. Sci., **13**: 246-254. doi:10.1017/S0954102001000360.

Folmer, O., Black, M., Hoeh, W., Lutz, R., and Vrijenhoek, R. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Bio. Biotechnol. **3**: 294-299.

Freckman, D.W., and Virginia, R.A. 1998. Soil Biodiversity and Community Structure in the Mcmurdo Dry Valleys, Antarctica. *In* Ecosystem Dynamics in a Polar Desert: the Mcmurdo Dry Valleys, Antarctica. *Edited by* J. C. Priscu, American Geophysical Union, Washington, D.C. doi: 10.1029/AR072p0323.

Greenslade, P. 2015. Synonymy of two monobasic *Anurophorinae* genera (Collembola: Isotomidae) from the Antarctic Continent. System. Entomol., **38** (2): 134-141. doi: 10.1080/00779962.2015.1033810.

Griffiths, H.J. 2010. Antarctic marine biodiversity – what do we know about the distribution of life in the southern ocean? PLoS ONE, **5** (8): e11683. doi:10.1371/journal.pone.0011683

Gosz, J.R. 1993. Ecotone Hierarchies. Ecol. Appl., **3** (3): 369-376.

Havermans, C., Nagy, Z.T., Sonet, G., De Broyer, C., and Martin, P. 2011. DNA barcoding reveals new insights into the diversity of Antarctic species of *Orchomene sensu lato* (Crustacea: Amphipoda: Lysianassoidea). Deep Sea Res. Part II Top. Stud. Oceanogr., **58**: 230–241.

Hawes, T.C. 2011. Rafting in the Antarctic springtail, *Gomphiocephalus hodgsoni*. Antarct. Sci., **23**: 456–460.

Hawes, T.C., Torricelli G and Stevens, M.I. 2010. Haplotype diversity in the Antarctic springtail *Gressittacantha terranova* at fine spatial scales—a Holocene twist to a Pliocene tale. Antarct. Sci., **22**: 766–773.

Hebert, P.D.N., Cywinska, A., and Ball, S.L. 2003. Biological identifications through DNA barcodes. Proc. R Soc. Lond. B Biol. Sci., **270**: 313–321

Hebert, P.D.N., Penton, E.H., Burns, J., Janzen, D.H., and Hallwachs, W. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly, *Astraptes fulgerator*. PNAS, **101**: 14812-14817.

Hogg, I.D., Cary, S.C., Convey, P., Newsham, K.K., O'Donnell, A.G., Adams, B.J., Aislabie, J., Frati, F., Stevens, M.I. and Wall, D.H. 2006. Biotic interactions in Antarctic terrestrial ecosystems: are they a factor? Soil Biol. Biochem., **38**: 3035–3040.

Hogg, I.D., Stevens, M.I. and Wall, D.H. 2014. Invertebrates. *In* Antarctic terrestrial microbiology. *Edited by* D. A. Cowan. Springer Heidelberg, New York, pp.55-78. doi: 10.1007/978-3-642-45213-0_4

Hopkin, S.P. 1997. Biology of the springtails (Insecta: Collembola). Oxford University Press, New York.

Ivanova, N.V., Dewaard, J.R. and Hebert, P.D.N. 2006. An inexpensive, automation-friendly protocol for recovering high-quality DNA. Mol. Ecol. Notes, **6**: 998–1002.

Jenkins W.R. 1964. A rapid centrifugal-flotation technique for separating nematodes from soil. Plant Dis.. Rep., **48** (492).

Knowlton, N., Weight, L.A., Solórzano, L.A., Mills, D.K., and Bermingham, E. 1993. Divergence in Proteins, Mitochondrial DNA, and Reproductive Compatibility Across the Isthmus of Panama. Science, **260** (5114): 1629-1632.

Levy, J., Lyons, W.B., and Adams, B.J. 2013. Understanding Terrestrial Ecosystem Response to Antarctic Climate Change. Eos, Trans. Amer. Geophys. Union, **94**: 33–33. doi: 10.1002/2013EO030009.

McGaughran, A., Convey, P., Stevens, M.I. and Chown, S.L. 2010. Metabolic rate, genetic and microclimate variation among springtail populations from sub-Antarctic Marion Island. Polar Biol., **33**: 909–918.

Mortimer, E., Jansen van Vuuren, B., Lee, J.E., Marshall, D.J., Convey, P., and Chown, S.L. 2011. Mite dispersal among the Southern Ocean Islands and Antarctica before the last glacial maximum. Proc. R. Soc. B, **278**. doi:10.1098/rspb.2010.1779.

Peat, H.J., Clarke, A., and Convey, P. 2007. Diversity and biogeography of the Antarctic flora. J. Biogeog., **34**(1): 132-146. doi: 10.1111/j.1365-2699.2006.01565.

Pisa, S., Biersma, E.M., Convey, P., Patiño, J., Vanderpoorten, A., Werner, O., and Ros, R.M. 2014. The cosmopolitan moss *Bryum argenteum* in Antarctica: recent colonisation or in situ survival? Polar Biol., **37**: 1469–1477. doi: 10.1007/s00300-014-1537-3.

Pollard, D., and DeConto, R.M. 2009. Modelling West Antarctic ice sheet growth and collapse through the past five million years. Nature, **458**(7236): 329-332. doi:10.1038/nature07809.

Posada, D. 2008. jModelTest: phylogenetic model averaging. Mol. Biol. Evol., **25**: 1253–1256.

Pugh, P.J.A. and Convey, P. 2008. Surviving out in the cold: Antarctic endemic invertebrates and their refugia. J. Biogeogr., **35**: 2176–2186. doi:10.1111/j.1365-2699.2008.01953.x

Quek, S., Davies, S., Ilino, T., and 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.

Ratnasingham, S., and Hebert, P.D.N. 2007. BOLD: the barcode of life data system (www.barcodinglife.org). Mol. Ecol. Notes, **7**: 355–364.

Ratnasingham, S,, Hebert, P.D. 2013. A DNA-based registry for all animal species: The Barcode Index Number (BIN) System. *PloS one*, **8**(7): e66213.

Risser, P.G. 1993. Ecotones at local to regional scales from around the world. Ecol. Appl., **3**: 367–368.

Salmon, J.T. 1965. Two new genera of Antarctic Collembola. Pac. Insects, **7** (3):468-472.

Simon, C., Buckley, T.R., Frati, F., Stewart, J.B., and Beckenbach, A.T. 2006. Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. Annu. Rev. Ecol. Evol. Syst., **37**:545-579.

Sinclair, B.J., and Stevens, M.I. 2006. Terrestrial microarthropods of Victoria Land and Queen Maud Mountains, Antarctica: implications of climate change. Soil Biol. Biochem., **38**:3158–3170.

Stevens, M.I., and Hogg, I.D. 2002. Expanded distributional records of Collembola and Acari in southern Victoria Land, Antarctica. Pedobiol., **46**: 485–495.

Stevens, M.I., and Hogg, I.D. 2003. Long-term isolation and recent range expansion from glacial refugia revealed for the endemic springtail *Gomphiocephalus hodgsoni* from Victoria Land, Antarctica. Mol. ecol., **12**(9): 2357-2369. doi:10.1046/j.1365-294X.2003.01907.

Stevens, M.I., and Hogg, I.D. 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 Biol. Biochem., **38**: 3171–3180.

Swofford, D. 2002. PAUP*: phylogenetic analysis using parsimony, version 4.0b10 for Macintosh. Sinauer Associates Inc Publishers, Sunderland.

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

Velasco-Castrillón, A., Gibson, J.A.E., Stevens, M.I. 2014. A review of current Antarctic limno-terrestrial microfauna. Polar Biol., **37** (10): 1517-1531. doi:10.1007/s00300-014-1544-4.

Wall, D.H. 2005. Biodiversity and ecosystem functioning in terrestrial habitats of Antarctica. Antarct. Sci., **17**: 523-531. doi:10.1017/S0954102005002944.

Wise, K.A.J. 1967. Collembola (springtails). Antarct. Res. Ser., **10**: 123–148.

SUPPLEMENTARY MATERIALS

SM 2 Table of springtail collection information showing the location, coordinates and BOLD sample IDs of sequences for each unique haplotype

Haplotype #	Location	Coordinates	BOLD Sample IDs
G. hodgsoni			
Gh1	Tiger Island	-76.784 162.452	ANTSP607
	Flatiron	-77.005 162.408	ANTSP585, ANTSP587, ANTSP589, ANTSP605, ANTSP606, ANTSP628, ANTSP629
	Benson Glacier	-76.822 162.107	ANTSP596, ANTSP597
	St Johns Range	-77.2801 161.731	ANTSP129, ANTSP131, ANTSP134, ANTSP136, ANTSP137, ANTSP138, ANTSP140, ANTSP141, ANTSP143, ANTSP151,
Gh2	St Johns Range	-77.2801 161.731	ANTSP128, ANTSP144, ANTSP145, ANTSP147, ANTSP148, ANTSP149, ANTSP218
Gh3	St Johns Range	-77.208 161.7	ANTSP210
Gh4	Flatiron	-77.005 162.408	ANTSP584, ANTSP586, ANTSP604
	St Johns Range	-77.208 161.7	ANTSP213, ANTSP215
Gh5	Flatiron	-77.005 162.408	ANTSP631

Gh6	Benson Glacier	-76.87 161.754	ANTSP564, ANTSP565
Gh7	Mt Seuss	-77.015 161.75	ANTSP153, ANTSP166, ANTSP167, ANTSP534
Gh8	St Johns Range	-77.208 161.7	ANTSP209
Gh9	St Johns Range	-77.2801 161.731	ANTSP132, ANTSP133, ANTSP135, ANTSP139, ANTSP211, ANTSP212
Gh10	St Johns Range	-77.285 161.726	ANTSP217
Gh11	Flatiron	-77.005 162.408	ANTSP634, ANTSP635
Gh12	St Johns Range	-77.2849 161.726	ANTSP150
Gh13	Mt Seuss	-77.023 161.738	ANTSP152, ANTSP154, ANTSP157, ANTSP158, ANTSP159, ANTSP160, ANTSP163, ANTSP164, ANTSP165, ANTSP168, ANTSP169, ANTSP172, ANTSP174, ANTSP175, ANTSP518, ANTSP520, ANTSP522, ANTSP524, ANTSP532, ANTSP533, ANTSP540, ANTSP541
	Benson Glacier	-76.87 161.754	ANTSP563
Gh14	Mt Seuss	-77.021 161.737	ANTSP521
Gh15	Mt Seuss	-77.015 161.75	ANTSP535
Gh16	Flatiron	-77.005 162.408	ANTSP627
Gh17	Tiger Island	-76.783 162.452	ANTSP639
Gh18	Tiger Island	-76.783 162.452	ANTSP590, ANTSP612
Gh19	Tiger Island	-76.783 162.452	ANTSP592, ANTSP637, ANTSP638, ANTSP640, ANTSP642, ANTSP644,

			ANTSP646
	Benson Glacier	-76.822 162.107	ANTSP623
Gh20	Mt Seuss	-77.011 161.768	ANTSP162, ANTSP173, ANTSP519, ANTSP525, ANTSP531
Gh21	Mt Seuss	-77.023 161.738	ANTSP539
Gh22	Mt Seuss	-77.2849 161.726	ANTSP142
Gh23	Towle Glacier	-76.655 161.093	ANTSP561
Gh24	Towle Glacier	-76.729 161.011	ANTSP556, ANTSP557, ANTSP558, ANTSP559
Gh25	Mt Gran	-76.982 161.16	ANTSP202
Gh26	St Johns Range	-77.2849 161.726	ANTSP146
Gh27	Benson Glacier	-76.822 162.107	ANTSP594, ANTSP595, ANTSP598, ANTSP599, ANTSP600, ANTSP601, ANTSP642, ANTSP649, ANTSP651, ANTSP652, ANTSP653, ANTSP654, ANTSP655, ANTSP656
A. monoculata			
Am1	Cliff Nunatak	-76.11 162.015	ANTSP552, ANTSP553, ANTSP554, ANTSP555, ANTSP566, ANTSP567, ANTSP568, ANTSP569, ANTSP579, ANTSP580, ANTSP581, ANTSP582

	Mt Murray	-76.11 162.014	ANTSP570, ANTSP571,ANTSP572, ANTSP573, ANTSP574, ANTSP575, ANTSP576, ANTSP577, ANTSP578, ANTSP613, ANTSP615
Am2	Pegtop Mt	-77.046 161.362	ANTSP618
Am3	Pegtop Mt	-77.046 161.362	ANTSP616
Am4	Benson Glacier	-76.822 162.107	ANTSP622
Am5	Springtail Pt	-77.168 160.71	ANTSP196, ANTSP235
Am6	Springtail Pt	-77.168 160.71	ANTSP194, ANTSP195, ANTSP205, ANTSP203, ANTSP204
C. nivicolus			
Cn1	Tiger Island	-76.784 162.452	ANTSP609
	Mt Seuss	-77.023 161.738	ANTSP124, ANTSP155, ANTSP156, ANTSP536, ANTSP543, ANTSP544, ANTSP548, ANTSP611
	Mt Gran	-76.972 161.148	ANTSP619, ANTSP620, ANTSP621
Cn2	Mt Seuss		ANTSP170
Cn3	Mt Seuss	-77.023 161.738	ANTSP547
Cn4	Mt Seuss	-77.023 161.746	ANTSP550, ANTSP537
Cn5	Towle Glacier	-76.655 161.093	ANTSP602, ANTSP603
Cn6	Springtail Pt	-77.1676 160.71	ANTSP119, ANTSP121

Chapter IV

Thesis Summary and Conclusions

In this thesis I undertook two studies, both of which examined the genetic diversity of invertebrates. The first study involved a nationwide assessment of three orders of aquatic invertebrates known as the EPT taxa (Ephemeroptera, Plecoptera, Trichoptera). The juveniles of these orders are key components of stream ecosystems and act as bioindicators of water quality. However, EPT larvae can be difficult to accurately identify using their morphological features. This study therefore aimed to provide complete barcode coverage for all New Zealand EPT taxa to streamline routine macroinvertebrate surveys. The second study addressed the fine scale haplotype diversity of three known species of terrestrial springtail (Collembola) found over a north-south range of 120 km within the vicinity of Mackay Glacier in the Ross Sea Region of Antarctica. Diversity within Antarctica has traditionally been understudied and, therefore, underestimated. Assessing the consequences global climate changes on the Antarctic biota are thus constrained by our limited knowledge of the biota. Springtails are sensitive to environmental changes and present an ideal indicator for monitoring the effects of any climatic shifts. This study examined current levels of diversity and distributional limits against which any future changes in diversity can be detected. Additionally, molecular clock approaches were used to reconstruct the evolutionary history of the Antarctic landscape and evaluate any influence on current patterns of biodiversity and species' distributions.

Both studies examined arthropod diversity across differing habitats and at different spatial scales. The resulting data have established baseline levels of diversity and provided coverage for a range of taxa to facilitate biomonitoring approaches. These in

turn can be used to improve estimates of water quality and predict the consequences of global climate changes. These projects have also demonstrated the utility of molecular-based approaches in the accurate assessment of diversity. Cases of a single morphologically identified species being represented by multiple genetically divergent populations were evident within each of the four orders. For example, the Antarctic springtail *Cryptopygus nivicolus* was represented by three distinct genetic groups or BINs (Barcode Index Numbers are used as a surrogate taxonomic designation). Similarly, four species of New Zealand stonefly were represented by more than one BIN. These studies have also highlighted groups that would benefit from further taxonomic review.

In Chapter II, I focussed on the assembly of a comprehensive reference library for the New Zealand EPT taxa. Thus far, sequence data has been generated for 41 morphologically-recognised mayfly, 41 stonefly and 180 caddisfly species. However, in the case of the stoneflies and caddisflies the sequence data actually indicated the presence of further species resources (48 stonefly and 225 caddisfly BINs). As a consequence of this research, all caddisfly families and genera have now been sequenced. The caddisfly database can therefore be considered functionally complete as the New Zealand macroinvertebrate community index (MCI) designates sensitivity values according to genus not species (Collier 1993; Stark 1993). Including short sequences, 83% of mayfly and 75% of stonefly genera have also been completed.

Mitochondrial COI sequences reliably resolved phylogenetic relationships to the family level for all three orders and additionally identified the finer scale

relationships between tribes in the case of the stoneflies. Average intraspecific divergences were between 1.2-1.3% for all orders while average interspecific divergences ranged from 23.8 to 31.3%. Barcoding also enabled the detection of misidentifications within museum collections in addition to identifying potential phylogeographic structuring and groups in need of further taxonomic review, such as *Deleatidium*, *Tiphobiosis* and *Zelandobius* species. Comparisons with international records supported the designation of all New Zealand EPT species as endemic with interspecific divergences generally above 13%. Accordingly, COI was an effective marker for discriminating between not only currently recognised species but also potential cryptic species and phylogeographic groups.

The findings of this study will enable the rapid and accurate identification of juvenile macroinvertebrates and facilitate their routine use in stream assessments (e.g. Dowle et al. 2015). These data also present the opportunity to conduct whole community analyses via next generation sequencing and environmental DNA protocols to further increase the scale and efficiency of macroinvertebrate assessments (Hajibabei et al. 2011; Shokralla et al. 2012). Data obtained from macroinvertebrate assessments can furthermore be used to facilitate research into the phylogenetic and phylogeographic patterns that have shaped the EPT fauna worldwide (Kjer et al. 2001; Kjer et al. 2002; Whitefield & Kjer 2008).

In Chapter III, I showed that mtDNA COI sequences have revealed considerably greater diversity than previously known in the Ross Sea region of Antarctica. In particular, we found several highly divergent lineages of springtails (5-11.3%

sequence divergence) including three previously unrecognised BINs for the three putative species to give a total of seven BINs within the wider Mackay Glacier area. Based on molecular clock estimates (*sensu* Knowlton et al. 1993; Brower et al. 1994; Quek et al. 2004), these divergent lineages were estimated to have been isolated within the last 5 million years (range: 2.2-7.5my), a time when the Western Antarctic Ice Sheet (WAIS) was thought to have completely collapsed (Pollard & Deconto 2009). This event would have facilitated dispersal via rafting (floating on the surface of water) due to increased liquid water availability (Hawes 2011; Hogg et al. 2014).

Improving our understanding of genetic variability within natural populations is fundamental to preserving diversity within natural systems and thereby the evolutionary potential of systems and populations to respond to a changing environment (Adams et al. 2006; Hogg et al. 2006; 2014). These data can be used to evaluate the evolution of the Antarctic landscape and in particular, the influence of past glaciations on present day species distributions and patterns of genetic diversity (Stevens & Hogg 2003; McGaughran et al. 2010). Given the presence of numerous unique and currently geographically restricted populations, any future changes in species' distributions can be easily tracked through the DNA barcoding of individuals. This will thereby enhance our capacity to detect subtle biological responses resulting from gradual climate changes. Understanding the effects of climate change within Antarctica could furthermore provide insight into the potential impacts of such changes, worldwide.

FUTURE WORK

The study described in Chapter II aimed to establish the first comprehensive COI reference library for the entire New Zealand EPT fauna with considerable progress being made in this effort. A primary aim for future work would be to complete the EPT databases and provide coverage for all mayfly and stonefly genera to facilitate macroinvertebrate surveys. Sequencing all of the EPT species and increasing the geographic coverage of specimens would further enable the more robust examination of phylogeographic patterns, the identification of populations and areas most at risk, and more accurate assessments of other biotic indices (%EPT abundance, quantitative MCI [QMCI] and semi-quantitative MCI [SQMCI]) (Stark 1998; Jackson et al. 2014). This study highlighted a number of groups that would benefit from future morphological re-examination to resolve taxonomic issues and confirm suspected misidentifications. Sequencing of additional genes such as the nuclear 28S and H3 genes would also likely aid in the resolution of taxonomic and cryptic speciation issues as well as provide support for any phylogeographic structuring (Vonnemann et al. 2005; Macher et al. 2016).

The study outlined in Chapter III demonstrated how much biodiversity remains undocumented within terrestrial Antarctic systems. In the future, multiple genes could be employed to provide further support for the relationships between the newly found springtail populations. Additionally, morphological re-examination of the genetically divergent, morphologically cryptic taxa could be used to determine whether they need to be formally reclassified. More extensive sampling between the Mackay Glacier

and northern locations would also likely reveal additional genetically distinct populations as well as help to clarify the current distributional limits of such populations. In particular, sampling within the Mackay Glacier and along the Clare Range could elucidate the relationship between Springtail Pt and Pegtop Mt *Antarcticinella monoculata* haplotypes and thus the evolutionary history of the lineages. Furthermore, determining how far sea levels rose within the Mackay Glacier during historical warming events and WAIS collapses could provide a more accurate view of past dispersal paths. Finally, the incorporation of genetic markers and gene frequencies into long term monitoring programmes could facilitate the detection of subtle biotic responses to climate changes and aid in the identification of areas and populations most at risk.

LITERATURE CITED

Adams BJ, Bardgett RD, Ayres C, Wall DH, Aislabie J, Bamforth S, Bargagli R, Cary SC, 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.

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 U.S.A*, **91**: 6491–5.

Collier K (1993) Review of the status, distribution, and conservation of freshwater invertebrates in New Zealand. *New Zealand Journal of Marine and Freshwater Research*, **27**, 339-356.

Dowle EJ, Pochon X, Banks JC, Shearer K, Wood SA (2015) Targeted gene enrichment and high throughput sequencing for environmental biomonitoring: a case study using freshwater macroinvertebrates. *Molecular Ecology Resources*. doi: 10.1111/1755-0998.12488

Hajibabaei M, Shokralla S, Zhou X, Singer GA, Baird DJ (2011). Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. *PLoS one*, **6**:4, e17497.

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

Hogg ID, Cary SC, 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, Wall DH (2014) Invertebrates. *In* Antarctic terrestrial microbiology. *Edited by* D. A. Cowan. Springer Heidelberg, New York, pp.55-78.

Jackson JK, Battle JM, White BP, Pilgrim EM, Stein ED, Miller PE, Sweeney BW (2014). Cryptic biodiversity in streams: a comparison of macroinvertebrate communities based on morphological and DNA barcode identifications. *Freshwater Science*, **33**, 312-324.

Kjer KM, Blahnik RJ, Holzenthal RW (2001) Phylogeny of Trichoptera (caddisflies): characterization of signal and noise within multiple datasets. *Systematic Biology*, **50**(6), 781-816.

Kjer KM, Blahnik RJ, Holzenthal RW (2002) Phylogeny of caddisflies (Insecta, Trichoptera). *Zoologica Scripta*, **31**(1), 83-91.

Knowlton N, Weight LA, Solórzano LA, Mills DK, Bermingham E (1993) Divergence in Proteins, Mitochondrial DNA, and Reproductive Compatibility Across the Isthmus of Panama. *Science*, **260** (5114): 1629-1632.

Macher JN, Salis RK, Blakemore KS, Tollrian R, Matthaei CD, Leese F (2016) Multiple-stressor effects on stream invertebrates: DNA barcoding reveals contrasting responses of cryptic mayfly species. *Ecological Indicators*, **61**, 159-169.

McGaughran A, Convey P, Stevens MI Chown SL (2010) Metabolic rate, genetic and microclimate variation among springtail populations from sub-Antarctic Marion Island. *Polar Biology*, **33**, 909–918.

Pollard D, DeConto RM (2009) Modelling West Antarctic ice sheet growth and collapse through the past five million years. *Nature*, **458**:7236, 329-332.

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.

Shokralla S, Spall JL, Gibson JF, Hajibabaei M (2012) Next‐generation sequencing technologies for environmental DNA research. *Molecular ecology*, **21**(8), 1794-1805

Stark JD (1993) Performance of the Macroinvertebrate Community Index: Effects of sampling method, sample replication, water depth, current velocity, and substratum on index values, *New Zealand Journal of Marine and Freshwater Research*, **27**:4, 463-478,

Stark JD (1998) SQMCI: A biotic index for freshwater macroinvertebrate coded-abundance data, *New Zealand Journal of Marine and Freshwater Research*, **32**:1, 55-66.

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**:9, 2357-2369.

Vonnemann V, Schrödl M, Klussmann-Kolb A, Wägele H (2005) Reconstruction of the phylogeny of the Opisthobranchia (Mollusca: Gastropoda) by means of 18S and 28S rRNA gene sequences. *Journal of Molluscan Studies*, **71**:2, 113-125.

Whitfield JB, Kjer KM (2008) Ancient rapid radiations of insects: challenges for phylogenetic analysis. *Annual Reviews Entomology*, **53**, 449-472.