Dispersal patterns and population connectivity of New Zealand stream insects: integrating landscape and population genetics

A thesis

submitted in partial fulfilment

of the requirements for the degree

of

Doctor of Philosophy in Biological Sciences

at

The University of Waikato

by

VANESSA DE ARAUJO BARBOSA

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2022

ABSTRACT

Increasing anthropogenic habitat fragmentation in riverine ecosystems, in particular the loss of natural riparian vegetation, has profound consequences for dispersal and connectivity of aquatic insects. Aquatic insects depend on dispersal to colonise new sites and expand their ranges, but colonisation is only effective if followed by successful reproduction of the colonising individual within the receiving population. Populations that are connected by dispersal are likely to possess and maintain higher levels of genetic diversity through gene flow, which enhances their long-term viability. Conversely, altered and/or restricted movement of individuals in fragmented landscapes may disrupt the connectivity of populations across habitats, reducing gene flow and decreasing genetic diversity of remnant populations. Thus, the ability to successfully disperse through fragmented and disturbed patches can be a key determinant of the long-term viability of populations.

Genetic markers provide an indirect approach to estimate dispersal potential and infer a species' gene flow, genetic diversity, and population connectivity. A variety of different genetic markers have been used over the past decades, each with advantages and disadvantages. Mitochondrial DNA (mtDNA) is commonly used in aquatic insect research and provides insight into historic dispersal, connectivity and isolation, while nuclear markers identified through next-generation sequencing approaches can reflect changes on a more contemporary time scale. The application of such markers in landscape genetics research further enables quantification of the effects of landscape and environmental features on gene flow, and identification of potential barriers to dispersal. This thesis is a collection of individual research papers that together provide new knowledge of functional connectivity and dispersal patterns of stream insect populations in altered landscapes. The research findings could ultimately assist with conservation and restoration planning.

An overview of the most commonly-used genetic markers, including their main features as applied to studies of aquatic insect dispersal, is provided as a global review in **Chapter 2**. Traditional markers, such as allozymes and mtDNA, are the most popularly applied and studies assessing the effects of specific landscape features on shaping population connectivity among habitats, especially in

fragmented landscapes, are rare. Higher resolution markers (e.g. single nucleotide polymorphisms; SNPs), have recently become available, providing high-throughput genome-wide data at low cost. With their potential to detect finer-scale genetic structure, these markers are expected to become more common in future studies, providing more accurate estimates of contemporary patterns of dispersal.

Chapter 3 compares the resolution power of mitochondrial cytochrome c oxidase subunit I (COI) and genome-wide SNP markers in estimating fine-scale genetic differentiation for three endemic New Zealand aquatic insects. Both markers provided comparable results: a general lack of strong population structure within each species, and fine-scale genetic differentiation among some populations. These results indicated substantial connectivity of the three analysed species within and between proximate streams separated up to 11 km. However, findings were considered preliminary as small sample sizes and limited data quality for the SNP datasets may have compromised their power to uncover genetic structuring. Nevertheless, an important overall finding was that either COI or SNP markers can provide suitable initial estimates of fine-scale population genetic differentiation in stream insects.

Finally, **Chapter 4** provides a detailed analysis of functional connectivity and dispersal patterns for the three stream insect species at multiple spatial scales, increasing the sample size and geographic coverage of sequenced individuals from Chapter 3. For all three species, clear spatial genetic structure marked by Isolation by Distance (IBD) was only observed at larger spatial scales (among mountain regions separated by ~30 and 170 km), whereas most gene flow occurred locally (up to 11 km). At the local spatial scale, landscape genetic analyses revealed that Isolation by Resistance (IBR) - particularly the influence of land cover - generally provided a better prediction of spatial genetic structure. Species-specific findings further highlighted the potential influence of continuous forest in the riparian zone in enhancing population connectivity and dispersal within the stream channel. Meanwhile, dispersal over pastoral land may be more common when insects must search for suitable habitat that cannot be found locally. These key findings are likely to be fundamental for future colonisation and persistence of these populations.

Understanding in-stream and overland dispersal, and how these affect the gene flow of species, is important for successful implementation of stream restoration measures. This research collectively elucidated patterns of population

connectivity and dispersal potential in a fragmented landscape that will provide valuable knowledge for conservation efforts aimed at enhancing restoration of stream insect biodiversity.

ACKNOWLEDGMENTS

Coloburiscus humeralis and Ngakoaohia Stream, Mount Pirongia
Photos: Vanessa Barbosa

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE

GENERAL INTRODUCTION

Aquatic insects are a common focus of freshwater ecology research and are often included in conservation and monitoring programmes. They have been a primary tool for studies of ecology, population genetics, evolution and many other areas of freshwater biology (Bilton et al., 2001; Hershey & Lamberti, 2001). However, their essential functions in aquatic ecosystems, and therefore their ecological and economical values, are often not well understood by decision makers or the public. Thus, I open this thesis with a brief overview of the significant roles of aquatic insects in riverine ecosystems.

1.1 The role of aquatic insects in riverine ecosystems

Aquatic insects are abundant in most freshwater habitats and often exhibit high species diversity throughout riverine systems. In New Zealand, they make up a large part of the total biodiversity and a very high proportion are endemic (Winterbourn et al., 2000). Most aquatic insect species spend their juvenile stage as larvae or nymphs in the water and then emerge onto land after metamorphosing into winged adults. Lifespan varies among taxa, but the aquatic stage can last at least a year, while the terrestrial adult stage may live for only a few days. During the aquatic phase (i.e. most) of their life cycle, aquatic insects perform various essential functions for the maintenance of the stream and river ecosystems they inhabit (Suter Ii & Cormier, 2015). They serve as a direct food source for most freshwater fish species, as well as for some vertebrate predators found in aquatic systems, such as amphibians and birds (Grant, 2001; Nakano & Murakami, 2001). Aquatic insects also play an important role in the cycling of materials and in trophic transfers. They retain nutrients in their biomass that directly improve the water quality by reducing downstream eutrophication (Newbold et al., 1982). In addition, their highly variable feeding traits (e.g. brushers, filterers, shredders, scrapers, predators) aid litter decomposition and periphyton reduction, and improve water quality (Cummins et al., 2005). Because most aquatic insect species are likely to reflect changes in their environment, they also act as indicators of water quality and ecosystem health of aquatic ecosystems. As greater demands have been placed on water resources,

many countries, including New Zealand, use aquatic insects for water quality monitoring (Boothroyd & Stark, 2000; Cormier & Messer, 2004). Therefore, these insects are an integral part of the aquatic ecosystem.

1.2 Dispersal in aquatic insects affects gene flow

The dispersal of aquatic insects is crucial for maintaining healthy aquatic ecosystems. The term dispersal has been used broadly in ecology, with the definition varying between subareas of research (Dingle, 1996). The Dictionary of Ecology, Evolution and Systematics defines dispersal as "[the] outward spreading of organisms or propagules from their point of origin or release; one-way movement of organisms from one home site to another" (Lincoln et al., 1998). Defined in this way, dispersal of aquatic insects equates to the movement of individuals between spatially discrete localities or populations, which may or may not result in migration, colonisation, or gene flow (Bilton et al., 2001).

Aquatic insects disperse using a range of mechanisms that can be broadly classified into active and passive dispersal. In streams, passive transport usually occurs by water currents or downstream drift, a dispersal mode prevalent in the aquatic larval or nymph stages (Bilton et al., 2001; Mackay, 1992). However, passive dispersal to new water bodies can also occur by wind, notably as adults, or by animal vectors (Maguire Jr, 1963). Active dispersal results predominantly from flight in adult insects that show varying degrees of dispersal ability. Larval swimming and benthic crawling can also occur as forms of active dispersal, albeit typically over short distances (Bilton et al., 2001; Graham et al., 2017; Mackay, 1992).

Considerable information is available on aquatic insect dispersal based on mark-recapture studies that use either physical marking methods or stable isotope tags. These studies have found that many adult aquatic insects prefer flying along stream corridors rather than overland between adjacent catchments (Collier & Smith, 1997; Petersen et al., 2004), even when the distance within a stream channel is longer than the overland distance between sites (Campbell Grant et al., 2007; Tonkin et al., 2014). Direct studies of dispersal often report that lateral dispersal of adults may be relatively uncommon for many species (Griffith et al., 1998),

2

although some strong fliers have been observed to travel from hundreds of metres to several kilometres (Briers et al., 2004; Kovats et al., 1996; Macneale et al., 2005).

Dispersal can have significant effects on the evolution and population genetics of a species. Aquatic insects depend on dispersal to colonise new sites and expand their ranges, but colonisation is only successful if followed by successful reproduction of the colonising individual within the receiving population. Such colonisation will then result in gene flow— the exchange of genes between individuals and populations, and the associated genetic contribution to subsequent generations — which can enable the spread of new adaptations and reduce inbreeding depression within populations (Slatkin, 1985). In other words, dispersal may reduce the probability of extinction within local populations by introducing new colonists and potentially increasing genetic diversity (Brown & Kodric-Brown, 1977). In the absence of gene flow, however, evolutionary divergence of populations may result, potentially leading to reproductive isolation and speciation or extinction (Tallmon et al., 2004). Therefore, populations that are connected by dispersal are likely to possess and maintain higher levels of genetic diversity through gene flow, which enhances their long-term viability (Bohonak, 1999).

In streams, dispersal capabilities interact with the biological traits and life history of a given species, as well as with the dendritic structure of the stream network, the spatial position of individuals, and the surrounding landscape structure (Alp et al., 2012; Hughes et al., 2009). Together, such dynamic processes determine the degree of connectivity among populations. When approached from a genetic perspective, these interactions will result in different potential models of gene flow that may vary from 'widespread' for species with high dispersal capacity and no particular habitat requirements (leading to low genetic structure), to 'limited' for those with low dispersal potential and/or habitat specialisations (creating high genetic structure) (Alp et al., 2012; Finn et al., 2007; Hughes et al., 2009). Likewise, gene flow among populations can be influenced by habitat loss and fragmentation that may result in dispersal constraints for organisms moving between patches (Crook et al., 2015). Thus, dispersal plays a vital role in ensuring population connectivity and persistence across degraded landscapes (Galic et al., 2013).

1.3 Genetic markers give an indirect measure of dispersal

The ecological significance of aquatic insect dispersal in streams has been well recognised, but methodological difficulties associated with direct observation limit our ability to understand processes and modes of dispersal (Hassall & Thompson, 2012). Gene flow also generally cannot be observed directly. Genetic markers provide an indirect approach to estimate dispersal potential and infer a species' gene flow, genetic diversity, and population connectivity (Hughes et al., 2008). Studies using genetic markers generally apply genetic differentiation parameters (e.g. pairwise F_{ST}-values; Wright, 1951) to estimate dispersal potential, assuming that widespread dispersal leads to genetic homogeneity, whereas restricted dispersal leads to a pattern of genetic differentiation among populations. Furthermore, the application of these parameters is commonly supplemented by several other analytical methods (e.g. haplotype analyses and multilocus genotype assignments to populations) in order to maximise the accuracy of gene flow estimates (Dufresne et al., 2014; Meirmans, 2015; Rousset, 1997).

Most investigations of aquatic insect dispersal using genetic markers focus on dispersal at a regional scale to estimate gene flow over evolutionary timescales (Miller et al., 2002). These studies have furthered our understanding of species' isolation, colonisation, and range expansion processes resulting from historic dispersal events, and the associated partitioning of genetic diversity in space (i.e. phylogeography) (Hotaling et al., 2019; Takenaka et al., 2019; Taylor et al., 2020), as well as helped to identify cryptic species, which show no or only subtle morphological differences (Pauls et al., 2009). Genetic markers can also be used to assess dispersal patterns at multiple spatial and temporal scales (Hughes, 2007; Hughes et al., 2009). With the increasing availability and popularity of higher resolution markers over the last decade (e.g. microsatellites and genome-wide single nucleotide polymorphisms, SNPs), genetic studies have contributed to a more comprehensive understanding of fine-scale dispersal and contemporary gene flow processes (Dussex et al., 2016; Geismar et al., 2015). The field of landscape genetics emerged during this time, and focuses on quantifying the effects of landscape and environmental features on gene flow. Although such studies are still very limited in aquatic insect dispersal research, they have enhanced our knowledge of potential natural and anthropogenic barriers to dispersal, offering valuable

implications for conservation and management of stream habitats (Keller & Holderegger, 2013; Phillipsen et al., 2015; Polato et al., 2017; Wilcock et al., 2007).

1.4 Thesis aims and objectives

Aquatic insects are one of the most threatened invertebrate groups in freshwater ecosystems due to their sensitivity to multiple stressors, including anthropogenic habitat modification, pollution, and climate change (Sánchez-Bayo & Wyckhuys, 2019; Stepanian et al., 2020). Given the increasing fragmentation in stream ecosystems associated with human development, research is required to enhance knowledge of functional connectivity and dispersal of stream insect populations in altered landscapes that will assist with conservation and restoration planning. Such knowledge can be facilitated by population genetics analysis, which has become a fundamental part of the tool kit for decision-makers in conservation biology. This thesis aims to assess dispersal and population connectivity patterns in stream insects using a combination of mitochondrial DNA and genome-wide nuclear genetic markers (i.e. SNPs). Using different population genetics approaches, this thesis used aquatic insect populations and species to address four specific goals:

1) To identify generalities of population genetic structure and dispersal patterns at varying spatial scales for aquatic insect species through a global review of published literature on the contribution of genetic markers to relevant aquatic insect research;

2) To compare the resolution of mitochondrial and genome-wide nuclear markers for detecting genetic differentiation among populations at small spatial scales;

3) To determine spatial patterns of population genetic structure at small and large spatial scales; and

4) To examine the influence of landscape features in shaping the spatial genetic structure of populations in a small fragmented landscape.

I investigated species-specific patterns for representative taxa from each of the orders Ephemeroptera, Plecoptera, and Trichoptera (EPT); these are the most

commonly studied of all benthic macroinvertebrate orders due to their widespread geographic distribution, their systematic responses to environmental change, and their diversity of functional groups in freshwater ecosystems. I selected three species, endemic to New Zealand and widely distributed in streams across the North Island: the mayfly *Coloburiscus humeralis* (Ephemeroptera: Coloburiscidae), the stonefly *Zelandobius confusus* (Plecoptera: Gripopterygidae) and the caddisfly *Hydropsyche fimbriata* (Trichoptera: Hydropsychidae).

1.5 Thesis Outline

This thesis comprises five chapters, of which research chapters 2-4 were developed as a series of three stand-alone studies. Each of these has been submitted, or is in preparation for submission, to peer-reviewed scientific journals. As the chapters are written as individual papers, there may be some repetition in the methodological details and context provided in the introductions.

> **Chapter 2** presents a global literature review of the contribution of genetic markers in assessing dispersal and spatial connectivity of aquatic insect populations in lotic ecosystems. This research provides an overview of the most commonly-used genetic markers, describing their main features as they apply to studies of dispersal. It particularly explores the application of genetic markers to studies of aquatic insects, and their contributions and limitations to our understanding of dispersal. This chapter further examines published data analysing genetic differentiation at different spatial scales to identify generalities in dispersal patterns and levels of population connectivity across a range of taxa. Finally, the review provides recommendations for avenues of future research, including the application of landscape-genetic studies to elucidate constraints and barriers to dispersal.

> **Chapter 3** examines fine-scale genetic differentiation among stream insect populations for the three target species. Here, I compare the resolution of mitochondrial cytochrome *c* oxidase subunit I (COI) and nuclear SNP markers for detecting population structure among sampling sites separated by up to 11 km. This is a preliminary study on the performance of both

marker types and provides initial insights into the population connectivity and dispersal patterns of the three analysed species.

Chapter 4 extends the research in Chapter 3 by increasing the sample size and geographic coverage of sequenced individuals and using a different sequencing technology to provide a detailed analysis of the connectivity and dispersal patterns of the studied stream insects. Here, I (i) analyse spatial genetic structure at both smaller and larger spatial scales and (ii) apply a fine-scale landscape genetic approach to examine the influence of landscape elements on genetic differentiation, dispersal, and gene flow. In particular, I explore whether local topography and land cover (forest versus open pasture) better predict spatial genetic structure than purely spatial scaling—often referred to "Isolation by Resistance" versus "Isolation by Distance". I discuss species-specific population connectivity and patterns of dispersal within the stream channel and among the studied streams and provide suggestions for conservation and stream restoration planning.

Chapter 5 provides a synthesis of the key findings from the preceding research chapters, discussing their main research limitations and identifying the main conclusions about the influence of the landscape on dispersal and gene flow of stream insect populations. Finally, I focus on the potential for future research to analyse fine-scale dispersal and deliver concrete suggestions for conservation and restoration measures in stream habitats.

In addition to the chapters described above, I contributed (sampling and identification) to a biodiversity survey in the Mount Priongia study area that is not included in this thesis, but is available as a preliminary online report titled *"Little bugs, big surprises — hidden freshwater treasures of Aotearoa"* (Drinan 2021). The records of this survey, which include adult aquatic insects collected from the malaise net by-catch, were submitted as a report to The Department of Conservation — DOC as a National Institute of Water and Atmospheric Research — NIWA client report (Smith, BJ., Barbosa, V., Graham, E. (2021): *A contribution to freshwater biodiversity values of Mount Pirongia, Waikato: Adult aquatic insects from Malaise trap by-catch*). Over the three streams, the survey identified 57 caddisfly species (23% of New Zealand's caddisfly fauna), 25 mayfly

species (45% of the mayfly fauna), and 15 stonefly species (14% of the stonefly fauna), found a micro-caddisfly potentially new to science, and generated new distribution records for the Waikato for numerous caddisfly species and stonefly species (Smith et al., 2021).

1.6 References

Alp, M., Keller, I., Westram, A. M., & Robinson, C. T. (2012). How river structure and biological traits influence gene flow: a population genetic study of two stream invertebrates with differing dispersal abilities. *Freshwater Biology*, **57**, 969-981.

Bilton, D. T., Freeland, J. R., & Okamura, B. (2001). Dispersal in freshwater invertebrates. *Annual Review of Ecology and Systematics*, **32**, 26.

Bohonak, A. J. (1999). Effect of insect-mediated dispersal on the genetic structure of postglacial water mite populations. *Heredity*, **82**, 451-461.

Boothroyd, I. K. G., & Stark, J. D. (2000). Use of invertebrates in monitoring. In K. J. Collier & M. J. Winterbourn (Eds.), *New Zealand Stream Invertebrates: Ecology and Implications for Management* (pp. 344-373). Christchurch, New Zealand: New Zealand Limnological Society.

Briers, R. A., Gee, J. H. R., Cariss, H. M., & Geoghegan, R. (2004). Inter-population dispersal by adult stoneflies detected by stable isotope enrichment. *Freshwater Biology*, **49**, 425-431.

Brown, J. H., & Kodric-Brown, A. (1977). Turnover rates in insular biogeography: effect of immigration on extinction. *Ecology*, **58**, 445-449.

Campbell Grant, E. H., Lowe, W. H., & Fagan, W. F. (2007). Living in the branches: population dynamics and ecological processes in dendritic networks. *Ecology Letters*, **10**, 165-175.

Collier, K. J., & Smith, B. J. (1997). Dispersal of adult caddisflies (Trichoptera) into forests alongside three New Zealand streams. *Hydrobiologia*, **361**, 53-65.

Cormier, S. M., & Messer, J. J. (2004). Opportunities and challenges in surface wate and quality monitoring. In G. B. Wiersma (Ed.), *Environmental monitoring* (pp. 217-238): Boca Raton (FL).

Crook, D. A., Lowe, W. H., Allendorf, F. W., Erős, T., Finn, D. S., Gillanders, B. M., . . . Hughes, J. M. (2015). Human effects on ecological connectivity in aquatic ecosystems: Integrating scientific approaches to support management and mitigation. *Science of the Total Environment*, **534**, 52-64.

Cummins, K. W., Merritt, R. W., & Andrade, P. C. N. (2005). The use of invertebrate functional groups to characterize ecosystem attributes in selected streams and rivers in south Brazil. *Studies on Neotropical Fauna and Environment*, **40**, 69-89.

Dingle, H. (1996). Migration. The biology of life on the move. Oxford, UK: Oxford University Press.

Dufresne, F., Stift, M., Vergilino, R., & Mable, B. K. (2014). Recent progress and challenges in population genetics of polyploid organisms: an overview of current state-of-the-art molecular and statistical tools. *Molecular Ecology*, **23**, 40-69.

Dussex, N., Chuah, A., & Waters, J. M. (2016). Genome-wide SNPs reveal fine-scale differentiation among wingless alpine stonefly populations and introgression between winged and wingless forms. *Evolution*, **70**, 38-47.

Finn, D. S., Blouin, M. S., & Lytle, D. A. (2007). Population genetic structure reveals terrestrial affinities for a headwater stream insect. *Freshwater Biology*, **52**, 1881-1897.

Galic, N., Hengeveld, G. M., Van den Brink, P. J., Schmolke, A., Thorbek, P., Bruns, E., & Baveco, H. M. (2013). Persistence of aquatic insects across managed landscapes: Effects of landscape permeability on re-colonization and population recovery. *PLoS One*, **8**, e54584.

Geismar, J., Haase, P., Nowak, C., Sauer, J., & Pauls, S. U. (2015). Local population genetic structure of the montane caddisfly Drusus discolor is driven by overland dispersal and spatial scaling. *Freshwater Biology*, **60**, 209-221.

Graham, E., Storey, R., & Smith, B. (2017). Dispersal distances of aquatic insects: upstream crawling by benthic EPT larvae and flight of adult Trichoptera along valley floors. *New Zealand Journal of Marine and Freshwater Research*, **51**, 146-164.

Grant, P. (2001). Mayflies as food. In D. E. (Ed.), *Trends in research in Ephemeroptera and Plecoptera* (pp. 107-124). New York (NY): Kluwer Academic.

Griffith, M. B., Barrows, E. M., & Perry, S. A. (1998). Lateral dispersal of adult aquatic insects (Plecoptera, Trichoptera) following emergence from headwater streams in forested Appalachian catchments. *Annals of the Entomological Society of America*, **91**, 195-201.

Hassall, C., & Thompson, D. J. (2012). Study design and mark recapture estimates of dispersal: a case study with the endangered damselfly Coenagrion mercuriale. *Journal of Insect Conservation*, **16**, 111-120.

Hershey, A. E., & Lamberti, G. A. (2001). Aquatic insect Ecology. In J. H.Thorp and A. P. Covich (Eds), *Ecology and Classification of North American Freshwater Invertebrates* (Vol. 2, pp. 733-775): Academic Press.

Hotaling, S., Giersch, J. J., Finn, D. S., Tronstad, L. M., Jordan, S., Serpa, L. E., . . . Weisrock, D. W. (2019). Congruent population genetic structure but differing depths of divergence for three alpine stoneflies with similar ecology and geographic distributions. *Freshwater Biology*, **64**, 335-347.

Hughes, J. M. (2007). Constraints on recovery: using molecular methods to study connectivity of aquatic biota in rivers and streams. *Freshwater Biology*, **52**, 616-631.

Hughes, J. M., Schmidt, D. J., & Finn, D. S. (2009). Genes in streams: Using DNA to understand the movement of freshwater fauna and their riverine habitat. *BioScience*, **59**, 10.

Hughes, J. M., Schmidt, D. J., McLean, A., & Wheatley, A. (2008). Population genetic structure in stream insects: what have we learned? In J. Lancaster & R. A. Briers (Eds.), *Aquatic insects: challenges to populations* (pp. 268-288): CAB International.

Keller, D., & Holderegger, R. (2013). Damselflies use different movement strategies for short- and long-distance dispersal. *Insect Conservation and Diversity*, **6**, 590-597.

Kovats, Z., Ciborowski, J. A. N., & Corkum, L. (1996). Inland dispersal of adult aquatic insects. *Freshwater Biology*, **36**, 265-276.

Lincoln, R., Boxshall, G., & Clark, P. (1998). *A dictionary of ecology, evolution and systematics* (R. Lincoln, G. Boxshall, P. Clark., Eds). Cambridge, UK: Cambridge University Press.

Mackay, R. J. (1992). Colonization by lotic macroinvertebrates: A review of processes and patterns. *Canadian Journal of Fisheries and Aquatic Sciences*, **49**, 617-628.

Macneale, K. H., Peckarsky, B. L., & Likens, G. E. (2005). Stable isotopes identify dispersal patterns of stonefly populations living along stream corridors. *Freshwater Biology*, **50**, 1117-1130.

Maguire Jr, B. (1963). The passive dispersal of small aquatic organisms and their colonization of isolated bodies of water. *Ecological Monographs*, **33**, 161-185.

Meirmans, P. G. (2015). Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology*, **24**, 3223-3231.

Miller, M. P., Blinn, D. W., & Keim, P. (2002). Correlations between observed dispersal capabilities and patterns of genetic differentiation in populations of four aquatic insect species from the Arizona White Mountains, USA. *Freshwater Biology*, **47**, 1660-1673.

Nakano, S., & Murakami, M. (2001). Reciprocal subsidies: Dynamic interdependence between terrestrial and aquatic food webs. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 166-170.

Newbold, J. D., O'Neill, R. V., Elwood, J. W., & Van Winkle, W. N. (1982). Nutrient spiraling in streams: implications for nutrient limitation and invertebrate activity. *American naturalist*, **120**, 628-652.

Pauls, S. U., Theissinger, K., Ujvarosi, L., Balint, M., & Haase, P. (2009). Patterns of population structure in two closely related, partially sympatric caddisflies in Eastern Europe: Historic introgression, limited dispersal, and cryptic diversity. *Journal of The North American Benthological Society*, **28**, 517-536.

Petersen, I., Masters, Z., Hildrew, A. G., & Ormerod, S. J. (2004). Dispersal of adult aquatic insects in catchments of differing land use. *Journal of Applied Ecology*, **41**, 934-950.

Phillipsen, I. C., Kirk, E. H., Bogan, M. T., Mims, M. C., Olden, J. D., & Lytle, D. A. (2015). Dispersal ability and habitat requirements determine landscape-level genetic patterns in desert aquatic insects. *Molecular Ecology*, **24**, 54-69.

Polato, N. R., Gray, M. M., Gill, B. B., Becker, C. G., Casner, K. L., Flecker, A. S., …, Zamudio, K. R. (2017). Genetic diversity and gene flow decline with elevation in montane mayflies. *Heredity*, **119**, 107-116.

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, **145**, 1219-1228.

Sánchez-Bayo, F., & Wyckhuys, K. A. G. (2019). Worldwide decline of the entomofauna: A review of its drivers. *Biological Conservation*, **232**, 8-27.

Slatkin, M. (1985). Gene flow in natural populations. *Annual Review of Ecology and Systematics*, **16**, 393-430.

Stepanian, P. M., Entrekin, S. A., Wainwright, C. E., Mirkovic, D., Tank, J. L., & Kelly, J. F. (2020). Declines in an abundant aquatic insect, the burrowing mayfly, across major North American waterways. *Proceedings of the National Academy of Sciences*, **6**, 2987-2992.

Suter Ii, G. W., & Cormier, S. M. (2015). Why care about aquatic insects: uses, benefits, and services. *Integrated Environmental Assessment and Management*, **11**, 188-194.

Takenaka, M., Tokiwa, T., & Tojo, K. (2019). Concordance between molecular biogeography of Dipteromimus tipuliformis and geological history in the local fine scale (Ephemeroptera, Dipteromimidae). *Molecular Phylogenetics and Evolution*, **139**, 106547.

Tallmon, D. A., Luikart, G., & Waples, R. S. (2004). The alluring simplicity and complex reality of genetic rescue. *Trends in Ecology & Evolution*, **19**, 489-496.

Taylor, C. L., Barker, N. P., Barber-James, H. M., Villet, M. H., & Pereira-da-Conceicoa, L. L. (2020). Habitat requirements affect genetic variation in three species of mayfly (Ephemeroptera, Baetidae) from South Africa. *ZooKeys*, **936**, 1-24.

Tonkin, J. D., Stoll, S., Sundermann, A., & Haase, P. (2014). Dispersal distance and the pool of taxa, but not barriers, determine the colonisation of restored river reaches by benthic invertebrates. *Freshwater Biology*, **59**, 1843-1855.

Wilcock, H. R., Bruford, M. W., Nichols, R. A., & Hildrew, A. G. (2007). Landscape, habitat characteristics and the genetic population structure of two caddisflies. *Freshwater Biology*, **52**, 1907-1929.

Winterbourn, M., Gregson, K. L. D., & Dolphin, C. H. (2000). Guide to the aquatic insects of New Zealand. *Bulletin of the Entomological Society of New Zealand*, No. 13.

Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics*, **15**, 323-354.

CHAPTER TWO
USING GENETIC MARKERS TO ASSESS DISPERSAL PATTERNS IN LOTIC INSECTS: A REVIEW

Caddisflies emerge en mass (top), adapted from video footage with permission from Brian Smith, 2020. Tawhiwhiti Stream, photo by Vanessa Barbosa.

This chapter has been submitted as a review article to Freshwater Biology and permission has been granted to use this article as part of the thesis. First submitted on 14/07/2021, first revised on 9/12/2021, currently undergoing second revision.

2.1 Abstract

1. Advances in sequencing technology led to the development of various genetic marker types that have been increasingly used in aquatic insect studies. The indirect assessment of dispersal abilities and patterns, and their potential influences, is a central goal of many of these studies. Likewise, they provide concrete suggestions for stream conservation management and restoration.

2. Here, we review the use of genetic markers in assessing dispersal and functional connectivity in aquatic insects. From 79 published papers, We examine the temporal changes on the use of different types of markers as they apply to the studies of dispersal, identifying generalities on the dispersal patterns across a range of taxa and the landscape features influencing connectivity among populations.

3. Until the early 2000s, allozyme markers were widely used in population genetic studies of the dispersal in aquatic insects, although the following advances in PCR technology resulted in the introduction of cost-effective mitochondrial DNA (mtDNA) markers, especially the cytochrome c oxidase subunit 1 (COI) gene, which are still the most used marker today. The subsequent development of next generation sequencing, has allowed genome-wide analysis the Single Nucleotide Polymorphisms (SNPs) markers to be performed for some species and enhanced understanding of population genetic structure on finer spatial and temporal scales. We highlight the resolution power varies among the most used markers, an aspect that should be carefully accounted when estimating dispersal ability based on genetic data.

4. Descriptive data based on estimates of genetic differentiation from widely-studied Ephemeroptera, Plecoptera, and Trichoptera, suggest that dispersal is more common within catchments for these taxa. For mayflies, a patchy oviposition was a frequently reported mechanism reducing in-stream dispersal in many species. Genetic data on stoneflies species suggest dispersal is potentially limited to neighbouring streams separated by tens of kilometers. Caddisflies showed isolation-by-distance for many species, suggesting mainly local dispersal within and among neighbouring streams.

5. Landscape genetics studies has increased our understating of how topography, and land cover features are also important for explaining spatial patterns of genetic variation.

6. We suggest that for future population genetic studies, combining mitochondrial and nuclear markers, will strengthen inferences about dispersal and connectivity of aquatic insects and the incorporation of a landscape genetics approach will further understanding of the interaction between dispersal and landscape barriers.

2.2 Introduction

Dispersal is an essential process contributing to the ecology and evolution of populations, facilitating gene flow, preventing local extinction, and enabling adaptation or recolonization following habitat disturbance (Bohonak, 1999; Waters et al., 2020). The definition of dispersal varies between subareas of research. Here, we call dispersal the one-way movement of individuals from one home site to another (Lincoln et al., 1998), between discrete localities or populations, that may result, or not, in colonization or gene flow (Bilton et al., 2001). For aquatic insects in particular, which often inhabit modified landscapes, dispersal is crucial to maintain population connectivity — the exchange of individuals among habitats, and to promote their long-term survival (Bilton et al., 2001). Understanding dispersal dynamics should therefore be a key consideration for conservation and restoration planning (Barton et al., 2015; Heino et al., 2017; (Parkyn & Smith, 2011)).

Trapping and mark-recapture studies and stable isotope labeling are typical direct approaches to estimate dispersal in aquatic insects, although the efforts required to undertake these methods limit their use (Hassall & Thompson, 2011). Research applying direct methods focus on the actual movement of the organism and have facilitated the understanding of the dispersal within the stream channel vs lateral and overland dispersal for many species and across different habitat types (e.g. Collier & Smith, 1997; Petersen et al., 2004). Molecular markers, on the other hand, have increasingly been used to address ecological questions, providing indirect estimates of dispersal. Levels of gene flow among populations is inferreded from individual genetic data, enabling assessment of the pattern and degree of connectivity among populations (Young et al., 2013). The application of molecular markers in aquatic insect studies has provided considerable insight into many aspects related to dispersal, including the role of larval versus adult dispersal, determine the spatial extent of dispersal, identify barriers to gene flow, and help differentiate influences from historical from contemporary processes (Hughes et al., 2009). However, genetic methods only measure 'effective dispersal' when an organism's movement is accompanied by a transfer of genes to future generations. Genetic approaches also rely on taxonomic certainty. Thus, genetic data should be ideally complemented by empirical studies for a comprehensive assessment of the

degrees of population connectivity, and dispersal patterns (Miller et al., 2002; Wubs et al., 2016).

One of the central questions of genetic studies in aquatic insects is to understand how populations are spatially connected (Hughes, 2009). In riverine systems, dispersal interact partly on the attributes of each species, including, for example, life-history and phenotypic traits (body size, wing shape), in addition to the hydrological and surrounding landscape structure: water flow, the position of populations in the stream network, for example, whether populations occur in headwaters or occupies an extent of the stream gradient; and the distance separating populations along and between stream channels, influencing their degree of connectivity (Bilton, 2001; Mazzucco et al., 2015). Research on aquatic insects often examines one, or a combination, of these factors to determine patterns of dispersal and connectivity among populations, and from a population genetics perspective, these interactions will result in different potential models of gene flow (e.g. Miller et al., 2002; Chester et al., 2015; Short & Caterino 2009; Sproul et al., 2014).These models were first described for fishes (Meffe and Vrijenhoek, 1988) and widen by Hughes et al. (2009), that have been used to describe dispersal by stream-dwelling insects: Death Valley Model (DVM), Stream Hierarchy Model (SHM), Headwater Model (HM) and panmixia (PAN) (Meffe & Vrijenhoek, 1988; Finn, et al., 2007; Hughes et al., 2009; Hughes et al., 2013). In the DVM, strong isolation, combined with small local population size and no terrestrial movement, as occurs in spring pools, for example, is expected to cause high genetic variation and lack of spatial genetic structure. In the SHM, the degrees of connectivity and gene flow vary depending on the distance between populations within the hierarchical network structure. This model should hold for aquatic insects that are not habitat-specialized and that spend most of their lifespan in the aquatic environment, with a short flight phase mostly confined to the stream channel (e.g. Wishart and Hughes, 2003; Saito and Tojo, 2016). In the HM, significant genetic structure occurs in the mountaintops reaches for headwater specialist species with low dispersal capacity among streams (e.g. Engelhardt et al., 201; Finn et al., 2007; Finn et al., 2011). Panmixia, expected to occur for strong-flying insect species, is characterized by widespread gene flow with little spatial structure. Isolation by distance (IBD) may be observed in these species, whereby populations that are

closer together are more similar genetically than those further apart (e.g. Phillipsen et al., 2014).

Over the past few decades, the technical progress made in sequencing methods and the appearance of high resolution markers, as well as advances in statistical approaches including in the field of landscape genetics, more detailed inferences about dispersal and gene flow could be made. It has helped, in particular, to infer gene flow in heterogeneous and fragmented landscapes providing estimates of functional connectivity. With such advances, it is now possible to identify landscape and environmental influences to dispersal as well as to better distinguish contemporary and historical processes that influence dispersal in a particular species. Local topography and recent habitat fragmentation (e.g. loss of vegetation, land-use change, urbanisation), for example have been revealed as important dispersal constrains in some species (Phillipsen et al., 2015; Didham et al., 2012; Wilcock et al., 2007). The increasing popularity of landscape genetic studies is strongly stimulated by new, higher resolution markers in the era of genomics research. Short-read and whole genome sequencing for population and community approaches are replacing traditional methods, aiming to increase the accuracy of estimate genome-wide parameter estimation and improve its applicability to conservation management (Fuentes-Pardo and Ruzzante, 2017; Supple and Shapiro, 2018). Notably, when compared to other species within the entomofauna, aquatic insects are profoundly underrepresented in genomics research (Hotaling et al., 2020), highlighting the need for progress in this field that will be essential for future conservation of freshwater ecosystems. In the meantime, concrete suggestions and solutions for conservation management today, and restoration of stream systems can be provided to some degree by population genetic studies (Hohenlohe et al., 2020; Willi et al., 2021).

In this review, we present a synthesis of population genetics research focused on dispersal-related questions in aquatic insects, in order to address the following objectives: 1) provide an overview of the most commonly-used genetic markers, describing their main features, their advancements over time, and their contributions and limitations to our understanding of aquatic insect dispersal; 2) use genetic variation data collected from a number of published studies across a range of taxa, to determine general patterns of dispersal at different spatial scales; 3) discuss the influence of different landscape features on dispersal and population

connectivity based on landscape genetics studies. We then suggest perspectives for future work that could be useful for conservation management of riverine ecosystems.

2.3 Literature search

Our literature search was based on published population genetic studies that addressed dispersal in riverine insect species (see searched keywords in Table A.2.1). This resulted in 179 articles published from 1987 to June 2021 obtained by searching keywords across title, abstract and text the Web of Science database, including searches of citations within each published paper. Relevant studies were included in our database if they met the following selection criteria: First, based on titles and abstracts, we excluded studies when the research was not conducted in streams or rivers, when focal species lacked a life stage reliant on freshwater, and when genetic data analysis did not included estimates genetic variation among populations. This resulted in 125 articles, which were read in full, and included in the final database only studies that used analysis of population genetic variation attempting to answer ecological questions about dispersal in riverine species. Seventy-nine publications (1987-2021) fit the final selection criteria (Table A.2.2). Their research was predominantly performed in North America (n=24), and Australasia (n=22) (Fig. 1a), and conducted on caddisflies (Trichoptera; 34%), mayflies (Ephemeroptera; 27%) and stoneflies (Plecoptera; 15%) (Fig. 2.1).
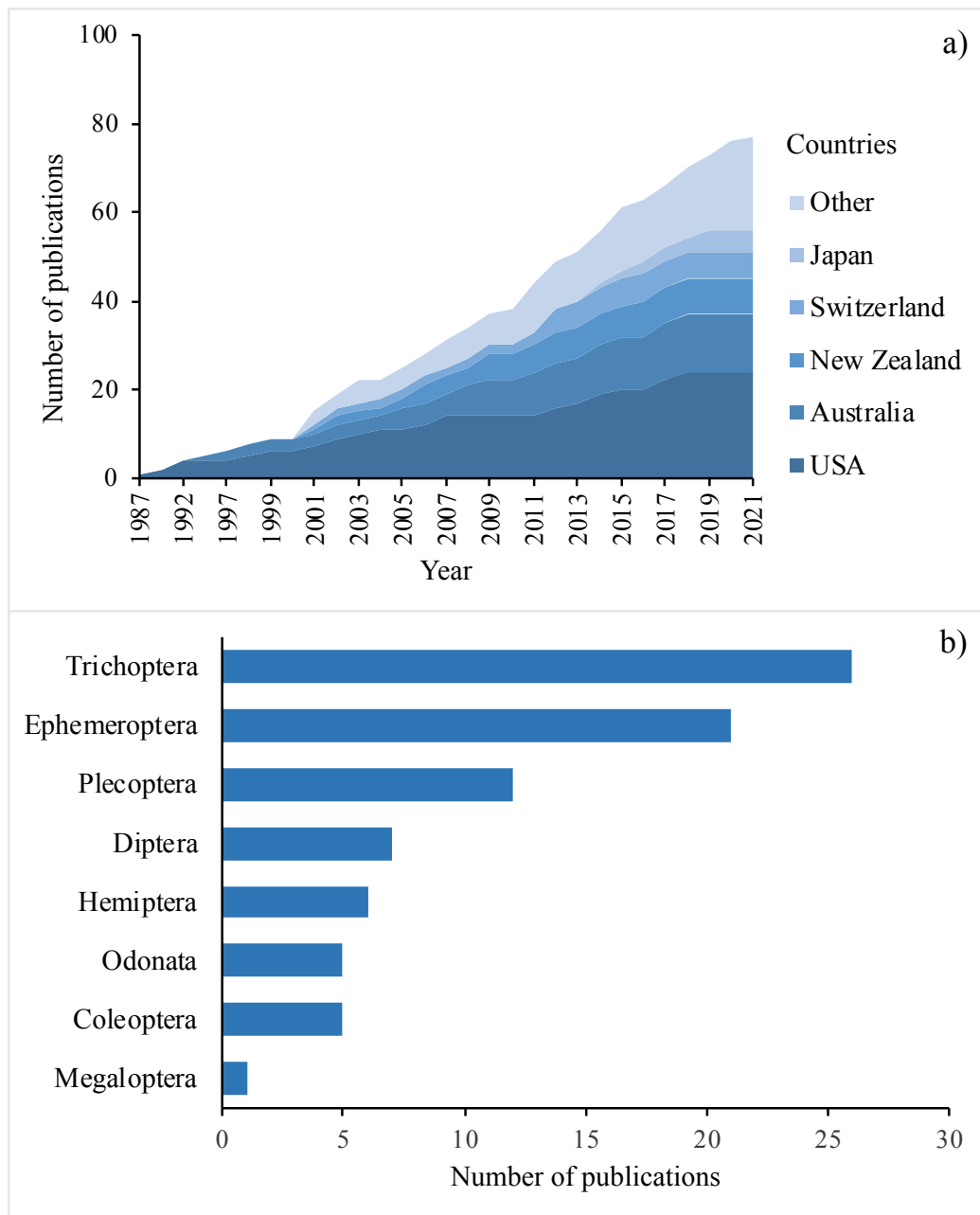
Figure 2.1 Based on articles reviewed (n=77): a) cumulative number of publications by country from 1987 to 2021, 'Other' includes countries representing <3% of publications; and b) number of publications by taxonomic order of investigated species.

2.4 Overview of molecular markers and their application on studies of aquatic insect dispersal

There are different types of genetic markers used in aquatic insect populations studies and their attributes can ultimately influence interpretation of

dispersal patterns. However, key considerations include: 1) nature of inheritance (maternal or bi-parental); 2) mode of gene action (dominant or co-dominant); 3) ability to resolve genetic variability; 4) reproducibility; and 5) costs and expertise required for development and analysis (Chenuil, 2006). Final marker selection is thus often based on a combination of biological and logistical constraints (Table 2.1). Of the 79 studies we reviewed, 43% (n=33) used mtDNA, 23% (n=18) allozymes, 14% (n=11) microsatellites, 3% (n=2) AFLPs, and 4% (n=5) SNPs. Eleven studies combined two marker types, using mtDNA together with either allozymes, AFLPs, microsatellites or another nuclear marker (Fig. 2.2, Table A.2.2). There was a move from allozyme-based to DNA-based markers around 2003, after which mtDNA became the most widely used. Microsatellites were first applied in the late 1990s and gradually gained more widespread use. AFLPs have remained uncommon and SNPs are a much more recent introduction (e.g. Dussex et al., 2016; Fig. 2.2). We follow with a brief description of the five main markers that have been used to examine population genetic structure of aquatic insects, and discuss the temporal changes AND contributions and limitations in their application for dispersal research. Detailed descriptions of each molecular marker type can be found elsewhere (Kirk & Freeland, 2011b; Semagn et al., 2006; Vignal et al., 2002; Zhang & Hewitt, 2003).

Table 2.1 Summary of the attributes of the different marker types, with corresponding properties about the mode of inheritance and nature of information (i.e. dominance), and rankings of their level of variability (i.e. polymorphism), reproducibility, and costs for development and sequencing: +, low; + +, intermediate; + + +, high.

	Allozymes	Microsatellites	mtDNA	AFLPs	SNPs
Mode of inheritance	Bi-parental	Bi-parental	Maternal	Bi-parental	Bi-parental
Dominance	Co-dominant	Co-dominant	Haploid	Dominant	Co-dominant
Polymorphism	+	+ + +	+ +	+ + +	+ + +
Reproducibility	+ +	+	+ + +	+ + +	+ + +
Costs (development and application)	+	+ + +	+	+ + +	++
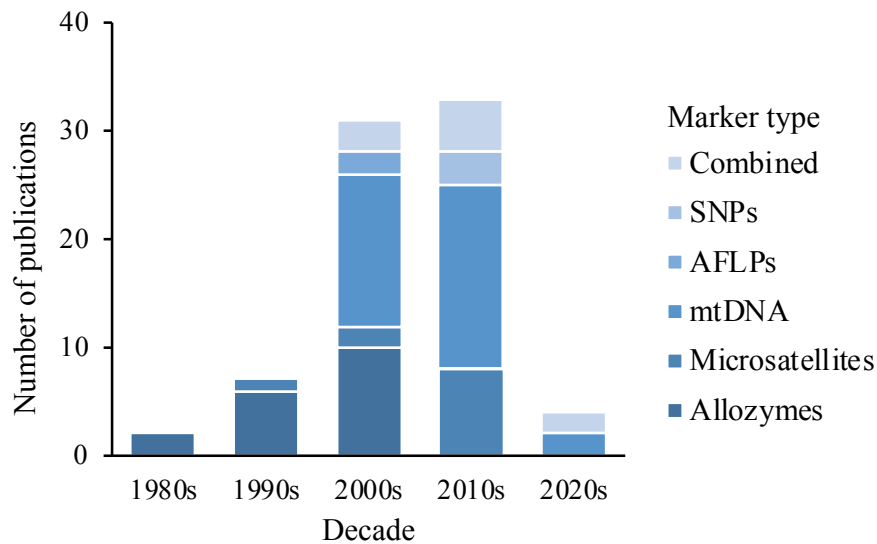
Figure 2.2 Molecular markers applied by decade to assess dispersal-related questions in aquatic insects.

Allozymes. Indirect estimates of dispersal in aquatic insects was first introduced analysing the distribution of allozyme variation across populations. Allozymes are protein-based markers that are allelic variants of enzymes. They exhibit co-dominant inheritance, meaning that both alleles of diploid individuals can be detected and heterozygotes can be distinguished from homozygotes (Ouborg et al., 1999). Allozymes can be used without extensive technical development and visualized through electrophoresis. They are cost-effective, fast, and relatively easy to use (Al-Samarai & Al-kazaz, 2015). The main limitation of this marker type, compared to other techniques now used is that they may not produce sufficient polymorphic signal to identify all potential genotypes in a population (Hughes, Hillyer, et al., 2003), limiting their resolution in resolving recent genetic divergence (Spitzer, 2014), or to measure gene flow at fine spatial scales (Turlure et al., 2014).

The introduction of allozyme analyses to aquatic insect research in the 1980s became possible due to the development of electrophoresis technology. Their application revealed that allozyme polymorphisms are maintained in aquatic insect species, resulting in the first genetic studies aimed to understand the extent of dispersal at catchment and regional scales, as well as the expansion history of populations (e.g. Preziosi & Fairbairn, 1992; Sweeney et al., 1987). Limited dispersal ability in Yoraperla brevis (Peltoperlidae) stoneflies, for example, was indicated by unexpectedly high genetic differentiation among closed-canopy stream populations, also suggesting that recolonization of this widespread species may be

25

limited if physical barriers exist between neighbouring channels (Hughes et al., 1999). Despite their initial widespread use, many allozyme studies reported a lack of genetic differentiation within and among catchments and potentially wrong attributed to strong dispersal capacity for many taxa (e.g. Hogg et al., 2002; Hughes et al., 2000; Wilcock et al., 2001). With the subsequent advance on DNA sequencing and increase of published studies, the general pattern of lack of genetic differentiation in allozyme data could be alternatively explained by their slow evolutionary rate, and allozymes were eventually replaced with more sensitive markers in the early 2000s (e.g. Hughes, Mather, et al., 2003).

Mitochondrial DNA (mtDNA). In the late 1990s and early 2000s, universal primers to amplify fragments of DNA began to be developed and DNA polymorphism analysis using Polymerase Chain Reaction (PCR) was commonly performed (e.g. Folmer et al. 1994). This progress led to the fast increase of the application of mtDNA markers, becoming the most convenient and cost-effective solution for genetic measures of aquatic insect dispersal (Cameron, 2014). It also resulted in a relevant improvement of taxonomic methods, with fragments of cytochrome c oxidase subunit I (COI) gene in particular, also used for species identification (Hebert et al., 2003), with databases such as the Barcode of Life Data systems (Ratnasingham & Hebert, 2007) providing a comprehensive repository of DNA sequences. The elevated mutation rate coupled with maternal inheritance, make mtDNA markers a sensitive indicator of population bottleneck and founder events (Moritz et al., 1987). Mitochondrial DNA is relatively easy to amplify because it appears in multiple copies in a cell. It is also characterized by the high rate of polymorphisms and mutations (Nabholz et al., 2008), making it well-suited to detecting genetic differentiation over short time frames (Addison et al., 2015). Combined, these characteristics have made mtDNA a popular genetic marker for a range of aquatic insect studies (Múrria et al., 2020) . However, due to their maternal inheritance, mtDNA analyses can be confounded by sex-biased dispersal, recombination, selection, or erratic evolutionary rates (Galtier et al., 2009). Other limitations include hybridization, introgression, and incomplete lineage sorting, processes which may result in underestimation of genetic diversity and inaccurate dispersal inferences (Al-Samarai & Al-Kazaz, 2015; Zhang & Hewitt, 2003).

Mitochondrial DNA markers, especially the COI gene region, gained popularity in the early 2000s and have since become the most used marker for assessing population structure and dispersal patterns in a range of aquatic insect species (Fig. 2.2). Published studies using this marker type, have increased knowledge of the evolutionary and phylogeographic history for a range of aquatic insect species (Hotaling et al., 2019; Schröder, et al., 2021; Stefanello et al., 2020; Taylor et al., 2020; Takenaka et al., 2019). The extensive use of COI markers have also contributed insights for management and conservation. Vulnerability to climate change, for example, was indicated by high genetic differentiation and limited dispersal ability in high-elevation Annitella (Limnephilidae) caddisflies (Múrria et al., 2020). Genetic isolation in the desert spring caddisfly Lepidostoma ojanum (Lepidostomatidae) indicated that disturbance in these habitats could result in their permanent loss owing to limited dispersal capacity (Myers, 2001). Similar effects of deforestation on population connectivity have been reported for other species (Múrria, 2010; Young, 2013).

Amplified Fragment Length Polymorphisms (AFLPs). Amplified Fragment Length Polymorphisms are a PCR-based technique that uses restriction enzymes for digestion of genomic DNA, followed by the ligation of adaptors to the end of DNA fragments, amplification and gel analysis for visualization (Vos et al., 1995). This results in hundreds of highly replicable markers for high-resolution genotyping (Mendelson & Shaw, 2005). The high resolution, reproducibility, and ability to discriminate from individuals to species are major advantages of this technique (Bonin et al., 2005). However, AFLP methods are costly and laborious (Semagn et al., 2006). Further, because of the dominant inheritance of AFLP loci, homozygotes cannot be distinguished from heterozygotes making it difficult to estimate allele frequency (Vos et al., 1995). Another limitation of AFLPs is that a higher number of markers are required to achieve a comparable level of resolution with other multi-allelic methods (Vignal et al., 2002).

The use of Amplified Fragment Length Polymorphisms (AFLPs) in aquatic insect studies is currently very limited (3% of the articles reviewed) and likely a result of the high costs and difficult data interpretation involved. Where AFLPs are used, inferences based on resolution power varied among studies. In general, AFLPs showed the potential for obtaining high resolution data within a catchment,

as observed in the calopterygid damselfly Calopteryx splendens, where individuals were mostly found to disperse along watercourses (Chaput-Bardy et al., 2008). The combined use of AFLPs with mtDNA have further shown that comparisons between mitochondrial and nuclear DNA data can provide a more confident interpretation of population structure patterns (Engelhardt et al., 2011; Sabando et al., 2011).

Microsatellites. Microsatellites, also called simple sequence repeats, were discovered in the 1980s and were quickly considered the most versatile genetic markers in conservation genetic research (Shamjana et al., 2015). Many microsatellites contain a variable number of repeats throughout the genome, resulting in length polymorphism of the amplified fragment (Kim & Sappington, 2013). Like AFLP, microsatellites are amplified using conventional PCR techniques, although microsatellite-based methods are more readily automated (Shariflou et al., 2001). Microsatellites are multi-locus, codominant markers which are bi-parentally inherited, allowing the determination of genotypes (Tautz & Renz, 1984). They have a higher mutation rate than mtDNA, and are highly polymorphic within the genome (Abdul-Muneer, 2014). Usually found in non-coding DNA, they are less affected by natural selection than allozymes (Kirk & Freeland, 2011) and mtDNA (Galtier et al., 2009). Accordingly, the combination of these properties has led the use of microsatellites as an ideal marker for assessment of neutral genetic diversity, parentage and relatedness, as well as fine-scale population structure (Yaegashi et al., 2014). Commonly cited limitations of microsatellites are their high development cost, the need for species-specific primers, and relatively low throughput when compared to SNPs (described in the next section). However, the same Next-Generation Sequencing (NGS) technologies that have widened the use of SNPs have also improved microsatellites in the development phase (Churbanov et al., 2012; Gardner et al., 2011; Fernandez-Silva et al., 2003; Wei et al., 2014). Nevertheless, given their advantages, the use of SNPs is widely expected to dominate the field of population genetics in aquatic insect research in the near future (Putman and Carbone, 2014; Zimmerman et al., 2020).

Microsatellites have been used to detect differentiation between geographically proximate sites and to estimate fine-scale dispersal patterns in aquatic insects. For example, dispersal in the stenopsychid caddisflies Stenopsyche

marmorata (Yaegashi et al., 2014) and Drusus discolour (Limnephilidae) (Geismar et al., 2015), was shown to be reduced at distances above ~20 km. Fine-scale genetic structure was also observed in other species (Gibbs et al., 1998; Spitzer, 2014; Wilcock et al., 2007). Microsatellites have further provided an indication of recent dispersal/gene flow (Phillipsen & Lytle, 2013; Schröder et al., 2021), local effects of bottlenecks (Shama et al., 2011), and habitat constraints on population structure when used in a landscape genetics context (Phillipsen et al., 2015). Thus, they have contributed to a more comprehensive understanding of fine-scale dispersal and connectivity.

Single nucleotide polymorphisms (SNPs). Single-nucleotide polymorphisms are variants at a single nucleotide base in a DNA sequence (Vignal et al., 2002). They are biparental markers, highly abundant in the genome, relatively cost-effective, and are easy to genotype (Morin et al. 2004). Technological advancements in high-throughput genotyping with Next-Generation Sequencing (NGS) platforms have allowed creation of much larger genome-wide SNP genotype datasets (Mardis, 2008). SNPs are more abundant and uniformly distributed across the genome than microsatellites, and an increased number of loci can be sampled with less effort and lower cost in comparison to microsatellite genotyping (Schlötterer, 2004). The simultaneous genotyping of thousands of loci consequently increases the statistical power to resolve population structure and evolutionary processes, with associated bioinformatics tools minimizing genotyping errors and maximizing retrieval of informative loci (e.g. Catchen et al., 2013; Mastretta-Yanes et al., 2015). The development of reduced representation methods to obtain SNP genotypes without a reference genome has broadened the application of these markers to numerous species (Baird et al., 2008; Davey et al., 2011). Among these, restriction enzymes associated with DNA sequencing, or simply RAD-seq, is now a popular technique in population genetic studies (Davey & Blaxter, 2010). A major advantage is that RADseq-based SNPs can produce large quantities of sequence data quickly, at low cost, and they can be applied to any species, regardless of genome size and pre-existing sequence data (Davey & Blaxter, 2010; Jaccoud et al., 2001). As a result, the application of genome-wide SNP techniques has enabled detailed examination of contemporary processes driving dispersal, including the identification of fine-scale spatial genetic structure, recent population-level differentiation (Walton e al.,

2021), and recent barriers to gene flow (Brauer and Beheregaray, 2020). This highlights the benefits of a genome-wide approach for understanding the effects of habitat fragmentation on population connectivity (Trense et al., 2020). Bioinformatics challenges in managing, analysing, and storing large datasets, are the main bottleneck of NGS (Daber et al., 2013), with high computational resources and specialised staff required for the data analysis and interpretation, which can be costly (Scholz et al., 2012). Addressing these challenges is an active area of ongoing research (Bianchi et al., 2016; Pfeiffer et al., 2018).

The rapid progress of SNPs genotyping technologies in the last decade has led to the emergence of these markers in recent population genomics studies of aquatic insects to resolve fine-scale population structure, relatedness, and evolutionary history (e.g. Lam et al., 2018). For example, substantial genetic structure at small spatial scales supported the hypothesis that wing reduction associated with low dispersal has driven diversification in alpine Zelandoperla (Gripopterygidae) stoneflies (Dussex et al., 2016). By integrating SNP data in a landscape genetics context, Polato et al. (2017) explained how the complex topography of montane regions may influence the genetic diversity of mayfly populations. These authors found that local landscape barriers between streams resulted in limited gene flow and apparent adaptive divergence in the higher elevation species. Other studies using SNP data from caddisfly species revealed high population connectivity within the catchment scale (Weigand et al., 2017; Weigand et al., 2018), a dispersal pattern often suggested for a number of caddisflies that supports their strong flight capability. Although the use of genome-wide data remains limited in studies of aquatic insects (Hotaling et al., 2020), they reveal the potential of SNPs an appropriate marker to assess dispersal.

2.5 Spatial scales of dispersal identified using molecular data

Genetic differentiation is the most common metric used for estimating population connectivity (i.e. gene flow), and therefore elucidate aquatic insects dispersal abilities and patterns across spatial scales. To assess generalities on the spatial scales of dispersal for a range of taxa, we synthesized genetic results from published studies, including only those that (1) estimated genetic differentiation using F-statistics (i.e. F_{ST} or Φ_{ST} values; Weir & Cockerham, 1984; Wright, 1951),

and (2) analyzed at least two of the potential spatial levels of genetic differentiation (e.g. within stream channel, between streams, catchments or regions). Fifty-six of the 79 studies (75%) estimated levels of genetic differentiation (F_{ST}) at different spatial scales, allowing an assessment of gene flow (and hence, dispersal). Below, we focus on the "EPT" taxa (Ephemeroptera, Plecoptera and Trichoptera), owing to their extensive coverage in the published literature (42 of the 56 studies). Data for the remaining taxa are summarized in Table A.2.3.

Ephemeroptera

Measures of genetic differentiation (FST) were greater at the smallest spatial scales for most species of mayfly (Fig. 2.3).

Allozymes. For example, allozyme data for *Acanthoplebia cruentata* (Leptophlebiidae) showed limited genetic differentiation over a range of distances (Smith & Collier, 2001), a somewhat unexpected result, given their short lifespan and adult behaviour of flying close to stream channels (McLean, 1967). However, evidence for population structure was found for populations separated by >100 km within the same catchment, and >25 km between catchments, suggesting greater dispersal occurring within catchments (Smith & Collier, 2001). This research suggested that historical events, rather than dispersal ability alone, may have influenced population genetic structure in *A. cruentata*. However, it is also possible that the limited genetic differentiation detected was partially due to the use of allozymes, highlighting the need for more variable markers to detect finer-scale patterns of differentiation (Smith et al., 2006). Low genetic differentiation was also found in *Rhithrogena loyolaea* (Heptageniidae), except between major drainages, suggesting moderate dispersal within and between streams (Monaghan et al., 2002). For *R. loyolaea*, low genetic differentiation was expected due to previous observations, suggesting they are strong flyers (Thomas, 2011). In contrast, *Baetis alpinus* (Baetidae) showed substantial differentiation within and among close streams, though populations were homogeneous at larger spatial scales (Monaghan et al., 2001, 2002). Such patterns are possibly the result of patchy oviposition, in which larval individuals represent the offspring of only a few matings (Bunn & Hughes, 1997), and limited dispersal within the stream (Schmidt et al., 1995).

mtDNA. Using mitochondrial DNA (mtDNA) data, Sabando et al. (2011) found significant genetic differentiation for populations of Andesiops torrens within individual rivers, similar to that previously reported in allozyme studies for species of Baetidae. In contrast, other studies have found opposing patterns from that previously detected using allozyme analysis (Hughes, Mather, et al., 2003; Schmidt et al., 1995). For example, when analysing mtDNA data for the baetids *Baetis bicaudatus* or *Bungona narilla*, higher genetic differentiation was found between streams versus higher differentiation within streams found using allozymes (Hughes, Hillyer, et al., 2003). The combination of two marker types, in this case, was essential to identifying sex-biased dispersal, suggesting that males dispersed among streams, while females were restricted to the stream channel. Thus such approach was demonstrated to be essential in uncovering connectivity patterns and thus recolonization potential among populations. Accordingly, recolonization between streams was unlikely as previously indicated by the allozyme data (Hughes, 2007). For other mayfly taxa, studies using mtDNA data inferred a range of different dispersal patterns. For example, *Drunella grandis* (Ephemerellidae) showed patterns of gene flow consistent with the isolation by distance (IBD; Rousset, 1997) model of genetic structure, indicating that dispersal is more common between proximate drainage basins (Sproul et al., 2014). For *Ulmerophlebia* sp. (Leptophlebiidae), high genetic differentiation was observed at distances >15 km, suggesting a restricted dispersal range between streams and sub-catchments (Young et al., 2013). In contrast, a lack of genetic differentiation was observed in the baetid Fallceon quilleri and *Isonychia japonica* (Isonychiidae) at multiple spatial scales across a range of hundreds of kilometres, suggesting that dispersal in both species is widespread (Zickovich & Bohonak, 2007, Saito & Tojo, 2016).

Microsatellites. Microsatellite data also showed contrasting patterns across the studied species. For example, *Siphlonisca aerodromia* (Siphlonuridae) showed low genetic differentiation among populations in catchments located ~100 km apart, suggesting that adult flight may be an important and sufficiently frequent mode of dispersal to maintain gene flow (Gibbs et al., 1998). Similarly, very weak genetic structure was found in *Baetis rhodani* (Baetidae) from alpine streams at all spatial scales, including among populations in-stream and subdrainages separated by ~100

km (Alp et al., 2012). This results show how widespread gene flow generally occur from strong fliers (Hughes, 2009). However, when analysing the same putative species from tropical stream habitats, Rebora et al. (2005) found higher differentiation on smaller spatial scales but limited differentiation across larger spatial scales. This is similar to previous studies of baetid mayflies that used allozymes or mtDNA and provides further evidence that patchy oviposition is likely to be a key determinant in the observed population structure for these taxa.
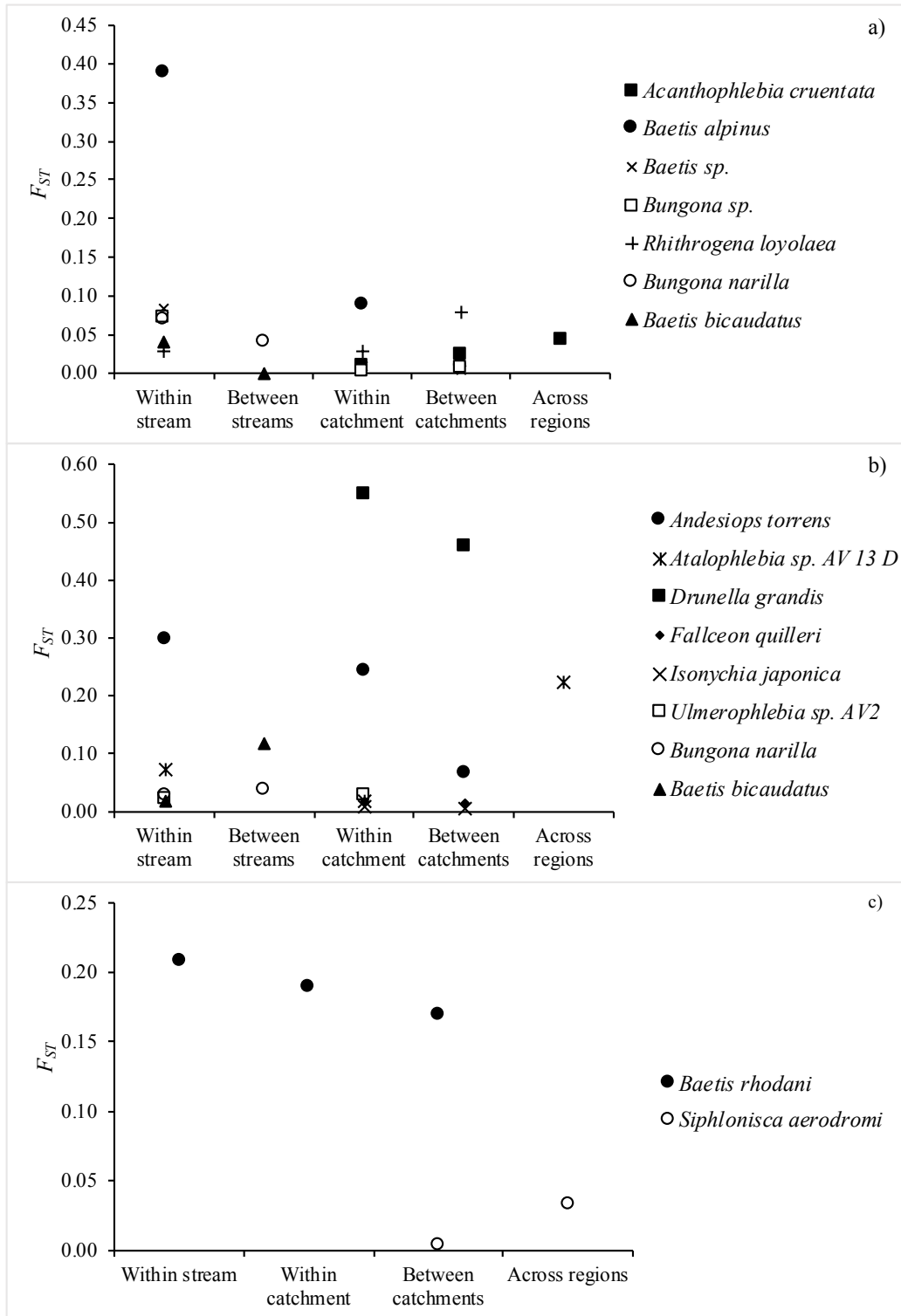
Figure 2.3 F_{ST} values among populations of mayfly species (Ephemeroptera) at different spatial scales. Data are presented for: a) allozymes; b) mtDNA; and c) microsatellites. Original references are provided in Appendix Table A.2.2.

Plecoptera

Genetic data for stonefly species were more limited compared to the other EPT taxa (Fig. 2.4). Overall, high genetic differentiation among populations detected in each of the studies indicate that stoneflies are not efficient dispersers (Hughes et al., 1999; Hughes, 2007). However, based on mtDNA and microsatellites, low genetic differentiation was observed among *Dinocras cephalotes* (Perlidae) populations, suggesting that this species is capable of dispersing between different headwater streams separated by up to ~ 50 km and confirming effective gene flow among populations at small spatial scales (Elbrecht et al., 2014). Populations of *Peltoperla tarteri* (Peltoperlidae) were significantly differentiated from one another using mtDNA data, except for those in the same stream, suggesting that adult dispersal is rare and larval movement is the primary mode of dispersal (Schultheis et al., 2002). Similar results obtained with mtDNA data for *Pteronarcella badia* (Peltoperlidae) revealed high differentiation among populations cross a mountain range (tens to hundreds of kilometers between sampling sites) and extended genetic isolation across the landscape, indicating that dispersal is very limited in this species and revealing some support to the headwater model (Sproul et al., 2014). To date, the application of SNPs markers was identified in only one stonefly *Zelandoperla* species, and observed genetic patterns were consistent with their wing morphology (Dussex et al., 2016). The wingless *Z. fenestrata* showed clear population structuring among neighbouring streams, associated with a loss of flying ability and reduced dispersal potential, whereas little differentiation was observed in *Z. decorata* populations, confirming that flight is a key dispersal mechanism in this species (Dussex et al., 2016; McCulloch et al. 2009).
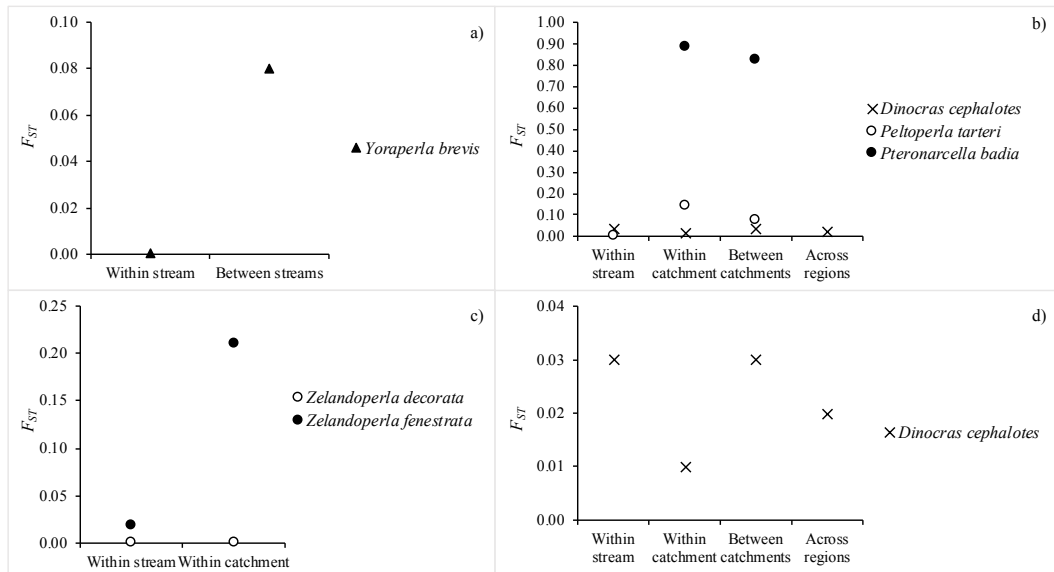
Figure 2.4 F_{ST} values among populations of stonefly species (Plecoptera) at different spatial scales. Data are presented for: a) allozymes; b) mtDNA; c) SNPs and d) microsatellites. Original references are provided in Appendix Table A.2.2.

Trichoptera

Results on the genetic variation among populations within a range of published studies indicate that dispersal abilities varied among caddisfly species although dispersal within catchments was important for maintaining population connectivity. This pattern is a result of a general widespread gene flow with IBD found for most species, that varied mainly with the space scale analyzed (Fig. 2.5).
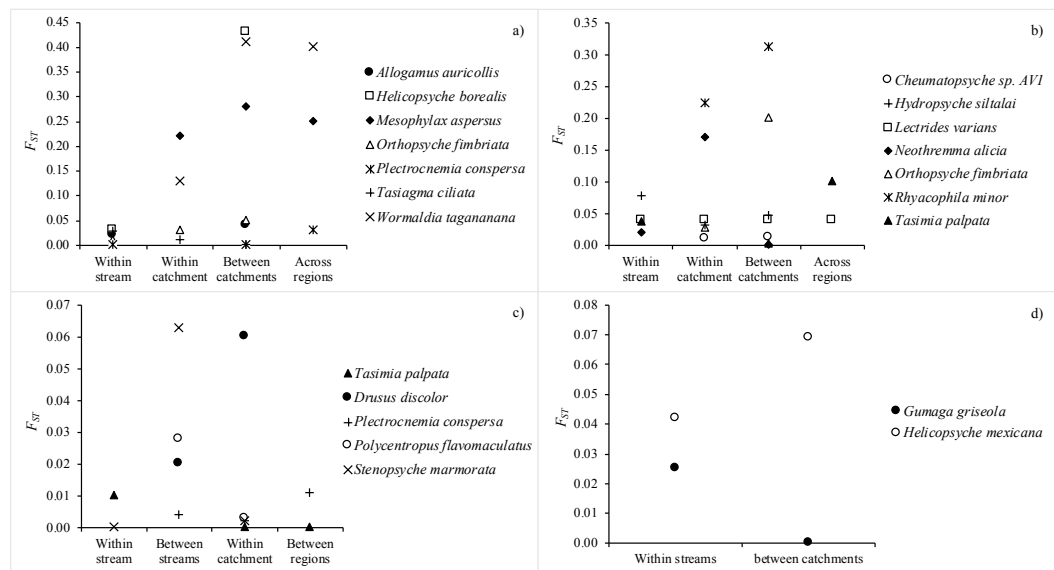
Figure 2.5 F_{ST} values among populations of caddisfly species (Trichoptera) at different spatial scales. Data are presented for: a) allozymes; b) mtDNA; c)

microsatellites; and d) AFLPs. Original references are provided in Appendix Table A.2.2.

Allozymes: Allozyme data showed low levels of genetic differentiation, suggesting that caddisflies are strong fliers and that dispersal occurs across multiple spatial scales (Hughes et al., 1998; Hughes, 2007). This pattern was observed in *Allogamus auricollis* (Limnephilidae) (Monaghan et al., 2002) and *Plectrocnemia conspersa* (Polycentropodidae) (Wilcock et al., 2003), in which major genetic differences were observed only at distances >100 km. Populations of *Mesophylax aspersus* (Limnephilidae) on the Canary Islands were not significantly different, indicating occasional dispersal of individuals between populations across the archipelago, despite their geographical isolation and the dispersal barrier of the sea (Kelly et al., 2001). Significant population structure, across a range of spatial scales, was only recorded in *Wormaldia tagananana* (Philopotamidae), suggesting that dispersal was lower than in *M. aspersus* on the same islands (Kelly et al., 2002). A hierarchical structure was evident in *Helicopsyche* (Helicopsychidae) (Jackson & Resh, 1992) and *O. fimbriata* (Smith & Collier, 2001), with smaller differences observed among sites within a stream and larger differences observed among catchments.

mtDNA: Inferences of dispersal based on mtDNA data also varied among Trichoptera species. No significant structure was detected within or between catchments in *Hydropsyche siltalai* (Hydropsychidae) (Múrria et al., 2010), *Lectrides varians* (Leptoceridae) (Wickson et al., 2014), and other species (Baker et al., 2003; Schultheis & Hughes 2005; Hughes, 2007; Fig. 2.5) indicating widespread adult dispersal across catchment boundaries. Hierarchical genetic structure was found for *O. fimbriata* (Smith & Smith, 2009), corroborating previously-published allozyme data and indicating that dispersal in this species is more likely to occur within catchments. A similar hierarchy was found for *Rhyacophila minor* (Rhyacophilidae) (Addison et al., 2015), indicating limited dispersal at the small spatial scale of the study landscape (45 km). Analysis of *Neothremma alicia* (Uenoidae) found that tributaries were the only level of stream hierarchy to show significant differentiation among populations, with dispersal confined to natal tributaries (Rader et al., 2019).

Overall, F_{ST} values were low for the Trichoptera species analysed either by microsatellites or AFLPs, although different patterns of differentiation were observed depending on the marker type. For example, microsatellite data for *Drusus discolor* (Limnephilidae) indicated high levels of overland dispersal but reduced levels at distances above 20 km (Geismar et al., 2015). Using microsatellite data in a landscape genetics approach, Wilcock et al. (2007) found contrasting results between two confamilial caddisflies *Plectrocnemia conspersa* and *Polycentropus flavomaculatus* (Polycentropodidae). For *P. conspersa*, genetic structure was only detected at a spatial scale of ~40 km in upland populations, and at a spatial scale of around 100 km in lowland populations. They suggested that geological formation, urbanization and the sea had a more pronounced effect on gene flow than altitude. In contrast, *P. flavomaculatus*, which was only analysed in lowland areas, had high genetic differentiation among most populations and no discernible effect of geographic distance or landscape features. Accordingly, even species with similar biological and ecological traits, can show marked differences in their adult dispersal and potentially influenced by their different habitat distributions (small versus large streams). In *Stenopsyche marmorata* (Stenopsychidae), microsatellite data showed that fine-scale estimates of dispersal suggested that gene flow was more pronounced along water courses (Yaegashi 2014). AFLPs analysis among populations within two adjacent catchments, showed different patterns for four species inhabiting the same geographic range, which was mainly attributed to their different dispersal abilities (Miller et al., 2002). For instance, *Helicopsyche mexicana* (Helicopsychidae) displayed significant differentiation at all spatial scales (up to few kilometers), indicating that individuals do not move far from oviposition or pupation sites. In contrast, patterns of differentiation for *Gumaga grioseola* (Sericostomatidae) were more influenced by adult flight across the terrestrial environment than the within-stream channel (Miller et al., 2002).

2.6 The influence of landscape features on gene flow

Over the past two decades, the field of landscape genetics has stimulated the combination of ecology, spatial statistics and population genetics, increasing our

understanding of the effect of landscape variables on genetic patterns. For aquatic insects, this method can be used to examine the relationship between gene flow and landscape features in shaping spatial population structure, helping defining the potential constraints and barriers to dispersal. Such approach is an important component of functional connectivity assessments, especially for stream insects populations, often inhabiting heterogenous and fragmented landscapes. Among the 79 published studies we reviewed, 55% (n=43) examined, at some degree, the spatial genetic structure of populations (Fig. 2.6; Table A.2.4). Basic analysis of gene flow in a purely theoretical space characterized by either Euclidean (straight-line distance between two points) or geographical distance (measured along the earth surface) was predominant, identified in 98% (n=41) of the articles. For an in-depth analysis of functional connectivity aiming a better understanding of the role of the landscape in species' dispersal and gene flow, 33% (n=14) of studies included the variable stream distance, 33% (n=15) included landscape topography, and 12% (n=6) included other landscape elements (e.g. canopy cover and/or environmental variables)
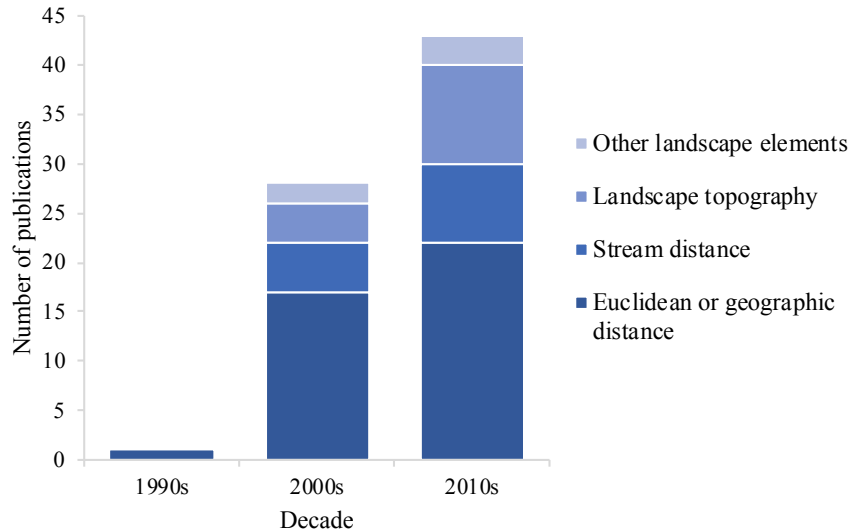
Figure 2.6 Summary of spatial measures of genetic structure applied by decade: Euclidean distance, stream distance, landscape topography, and other landscape elements.

Correlation analyses between Euclidean/geographical and genetic distance are commonly used to examine conformity to the isolation by distance (IBD) model. IBD proposes that more geographically distant populations will tend to be

more genetically distinct (Wright, 1943), assuming that closer populations exchange individuals more often than distant ones. Such analyses are often used to examine gene flow patterns in species with potential strong flight ability assuming their dispersal is not restricted by landscape features (Hughes et al., 2009). Commonly, studies of gene flow only uses a theoretical space when assessing connectivity and suggesting the dispersal ability of a particular species, without the integration of the habitat heterogeneity into the spatial-based analysis (e.g. Kelly et al., 2001; Zickovich and Bohonak, 2007; Table A.2.4). There are several of IBD examples reported in aquatic insects with a flying adult stage, elucidating dispersal patterns in a range of taxa, as discussed in the previous section. For some species, IBD is only detected at large spatial scales. For others, analysis at finer spatial scale, genetic distance increased significantly over a distance of only 4 km, suggesting that gene flow and hence dispersal is restricted to small spatial scales, as shown for the caddisflies *H. mexicana* (Miller et al., 2002) and *R. minor* (Addison et al., 2015). A small number of studies used stream network distances to resolve whether individuals mostly disperse along the stream channel via drift or overland by adult flight (Hughes, Mather, et al., 2003; Schultheis et al., 2002). For example, an IBD pattern was found in *C. splendens* populations, where Euclidean distance was used to test for overland dispersal and stream network distance analysis showed that in-stream dispersal was the main pathway for this species in the habitat analysed (Chaput-Bardy et al., 2008).

A prerequisite in landscape genetic analysis is the inclusion of Geographic information system (GIS) data. One important application of GIS is the integration of land use, land cover, or topographical maps with genetic data. The resulting combination, provides a powerful tool for evaluating the influence of different landscape features on functional connectivity and identifying potential barriers to dispersal. This approach has been used on a range of species and habitats, including headwater blackflies (Finn et al., 2006), widely-distributed caddisflies (Wilcock et al., 2007), damselflies (Chaput-Bardy et al., 2008; Keller et al., 2012), chironomids (Krosch et al., 2011), and stoneflies (Elbrecht et al., 2014). For example, Wilcock et al. (2007) attributed patterns of genetic structure in the caddisfly *P. conspersa* to be an effect of limited dispersal over large urban areas or natural barriers, such as the sea. Testing different landscape variables, Phillipsen and Lytle (2013) showed

that population connectivity of the flightless water bug *Abedus herberti* (Belostomatidae) may depend more on the shape of local topography than on direct connectivity within the stream network. Euclidean distance, topography, and intermittent habitat were identified as important factors for explaining landscape-level genetic patterns for the stonefly *Mesocapnia arizonensins* (Capniidae) (Phillipsen et al., 2015). Polato et al. (2017) assessed population structure in montane mayflies along an elevational gradient and also showed that dispersal was restricted at higher elevations by topographical barriers, resulting in greater genetic isolation. Recently, Weigand et al. (2018) analysed the effect of freshwater metal pollutants from wastewater and ore mining effluents on populations of the caddisfly *Glossosoma conformis* (Glossosomatidae), although the authors found low support for such environmental stressors being correlated with the identified pattern of gene flow.

The landscape genetics field is evolving to have a more impact on conservation management, by providing fundamental knowledge on dispersal ecology, functional connectivity and effects of habitat fragmentation. Few aquatic insect studies have been recently used this method to identify dispersal potential and its implications to recolonize restored stream habitats. For instance, sites containing isolated populations, due to limited dispersal of the species or landscape barriers, are much less likely to be recolonized following an extinction event (Rader et al., 2019). Alternatively, populations that are connected by widespread dispersal have more potential to naturally recolonize restored habitats, as reported for the stonefly *Dinocras cephalotes* (Elbrecht et al., 2014). With the continuous advancements in analyzing both high resolution genetic and landscape data, there is a great expectation that more detailed landscape genetic studies will soon provide valuable knowledge for a wider range of taxa, and used as component of stream conservation and restoration projects.

2.7 General discussion

Our review of genetic markers, as they have been used for research on aquatic insect dispersal, suggested a reduced ability to detect fine-scale structure using allozymes, mtDNA, and AFLPs compared with microsatellites and SNPs. The dominance of

mtDNA markers in aquatic insect studies to date indicates that cost-effectiveness is a major determinant of marker choice. Recent advances in Next-Generation Sequencing technologies and data analysis are enhancing the availability of highly informative genome-wide SNPs for non-model taxa at an affordable cost and should rapidly grow in freshwater research (Pauls et al., 2014), enabling a deeper understanding of patterns of genetic variation at finer spatial scales in a range of taxa (Hotaling et al., 2018; Rahman et al., 2018; Schmidt et al., 2017). Despite the increasing popularity of landscape genetics approaches in the last decade, we found that for aquatic insects, the influence of specific landscape features on population genetic structure has received limited attention. However, this approach has been beneficial for informing landscape connectivity models to predict patterns of dispersal among sites, identifying topographical barriers, and testing the effects of different configurations of fragmentation on dispersal and connectivity processes (e.g. Phillipsen et al., 2015; Polato et al., 2017).

The synthesis of the genetic differentiation data across the published studies indicate that the capacity of aquatic insects to disperse in-stream and overland varies considerably among species. For example, molecular data suggested that caddisflies are strong fliers, although dispersal is more common at the within-catchment scale (e.g. Geismar et al., 2015; Smith & Smith, 2009). Contrasting results were found for several mayfly taxa, whereby the highest genetic differentiation was often found at the smallest spatial scale (among reaches within a stream), suggesting a patchy oviposition mechanism could affect both larval and adult dispersal (e.g. Bunn & Hughes, 1997; Schmidt et al., 1995). Furthermore, in-stream dispersal may be important for maintaining connectivity within populations. For example, the caddisflies *S. marmorata* and *N. alicia* and the stonefly *P. tarteri* all showed significant differentiation at all spatial scales, indicating limited dispersal close to natal sites (Rader et al., 2019; Schultheis et al., 2002; Yaegashi et al., 2014). A majority (73%) of the reviewed studies incorporated a hierarchical analysis of genetic differentiation into their study designs, in order to assess patterns of dispersal at multiple spatial scales. Most of these studies found limited genetic differentiation among populations across a range of geographic scales and support previous suggestions that adult flight between neighbouring catchments is a major mechanism of dispersal (Hughes et al., 2008). While many species demonstrated strong dispersal abilities, connectivity within catchments often plays an important

role in maintaining high levels of gene flow among populations. Dispersal is one aspect of aquatic insect ecology and population connectivity. However, life history, habitat requirements, and geographic distributions associated with local environmental characteristics also play key roles, and may influence the observed patterns of genetic variation and should also be considered when interpreting population genetic data (e.g. Watanabe et al., 2014).

The potential effects of small sample sizes on genetic estimates were seldom discussed in the reviewed studies. However, it has been shown that increasing the number of individuals as well as amplifying more loci are likely to increase the precision of genetic estimates (e.g. Dubois et al., 2017; Landguth et al., 2012). Previous population genetics studies analysing the effect of sample size on commonly used measures of genetic variation, suggested that 20 to 30 individuals per population is ideal for inferences based on microsatellites (Hale et al., 2012; Pruett and Winker, 2008). For population genomic studies— when high-throughut sequencing methods are used, a mininum sample size of 8 individuals with a large number of SNPs (>1000) provide accurate genetic diversity estimates (Li et al., 2020; Nazareno et al., 2017). We suggest a careful consideration of sample size and its implications with the marker type used, in the study design of population genetic studies of dispersal. Further, studies examining spatial-temporal genetic structure can also be helpful for assessing dispersal and changes in connectivity. For example, Shama et al. (2011) showed that, when monitoring populations over a six year period, greater genetic differentiation was found following an extreme climatic event — the European heat wave in 2003. Similarly, Poff et al. (2018), showed that persistence and genomic diversity across a flooding gradient varied among species, suggesting that interaction between taxon-specific life cycle traits and timing of disturbance is an important predictor of how populations respond to extreme events. The results suggested more persistence and resilience via re-colonisation for species with mobile larvae and terrestrial adult stages, than taxa that lacked terrestrial adults. Unfortunately, collecting samples over multiple years for a temporal genetic study may not always be feasible due to logistic and funding constraints. However, this would be a valuable approach for investigating dispersal and monitoring populations that are subject to selection pressures which result in genetic changes (Hughes et al., 2003). Resolving fine temporal and spatial scales of dispersal remains a key purpose for genetic studies (Monaghan & Sartori, 2009; Pauls et al.,

2014). In this review, we found that many of the examined published research used either allozymes or mtDNA, thus contemporary patterns of dispersal could be biased by the limited inferential power of the markers used. Accordingly, genetic markers with high polymorphism and mutability, such as microsatellites and SNPs, are ideal for providing more accurate estimates of dispersal.

2.8 Future research

The accurate estimation of the patterns and mechanisms of dispersal and connectivity among aquatic insect populations is critical for the conservation, management and restoration of streams and rivers. To achieve this, we recommend that future studies include: 1) high polymorphic nuclear markers, such as microsatellites and SNPs, to better resolve fine-scale dispersal patterns and to determine the effects of any sex-biased dispersal; 2) the application of at least two different marker types, which will potentially allow differentiation of historical and contemporary patterns of gene flow and thereby aid in understanding how past natural events and recent fragmentation have together shaped present-day patterns; and 3) the use of a landscape genetics approach to facilitate a more accurate assessment of dispersal barriers and likely recolonization of habitats. The results from such integrated studies can then be applied to conservation management decisions and stream restoration planning.

2.9 References

Abdul-Muneer, P.M. (2014). Application of microsatellite markers in conservation genetics and fisheries management: recent advances in population structure analysis and conservation strategies. *Genetics Research International*, 691759.

Addison, J.A., Einfeldt, A.L., Kang, N.N., & Walde, S.J. (2015). Small-scale genetic structure in a stream-dwelling caddisfly in eastern Canada. *Marine and Freshwater Research*, **66**, 459-468.

Al-Samarai, F. R., & Al-kazaz, A. K. A. (2015). Molecular markers: an introduction and applications. *European Journal of Molecular Biotechnology*, **9**, 13187.

Alp, M., Keller, I., Westram, A.M., & Robinson, C.T. (2012). How river structure and biological traits influence gene flow: a population genetic study of two stream invertebrates with differing dispersal abilities. *Freshwater Biology*, **57**, 969-981.

Baker, A.M., Williams, S.A. & Hughes, J.M. (2003), Patterns of spatial genetic structuring in a hydropsychid caddisfly (*Cheumatopsyche* sp. AV1) from southeastern Australia. *Molecular Ecology*, **12**, 3313-3324.

Barton, P.S., Lentini, P.E., Alacs, E., Bau, S., Buckley Y.M., Burns, E.M., …, Smith, A.L. (2015). Guidelines for using movement science to inform biodiversity policy. *Environmental Management* **56,** 791-801.

Bianchi, V., Ceol, A., Ogier, A.G.E., de Pretis, S., Galeota, E., Kishore, K., …, Pelizzola, M. (2016). Integrated systems for NGS data management and analysis: Open issues and available solutions. *Frontiers in Genetics*. 7, 75.

Bilton, D.T., Freeland, J.R., & Okamura, B. (2001). Dispersal in freshwater invertebrates. *Annual Review of Ecology and Systematics*, **32**, 159-181.

Birky, C. W., Jr., Fuerst, P., & Maruyama, T. (1989). Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to

equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics*, **121**, 613-627.

Bohonak A.J. (1999). Dispersal, gene flow, and population structure. *The Quarterly Review of Biology*, **74**, 21-45.

Bonin, A., Pompanon, F., & Taberlet, P. (2005). Use of amplified fragment length polymorphism (AFLP) markers in surveys of vertebrate diversity. *Methods in Enzymology*, **395**, 145-161.

Brauer, C.J., Beheregaray, L.B (2020). Recent and rapid anthropogenic habitat fragmentation increases extinction risk for freshwater biodiversity. *Evolutionary Applications*, **13**, 2857-2869.

Bunn, S.E., & Hughes, J.M. (1997). Dispersal and recruitment in streams: Evidence from genetic studies. *Journal of the North American Benthological Society*, **16**, 338-346.

Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., & Cresko, W.A. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology*, **22**, 3124-3140.

Chaput-Bardy, A., Lemaire, C., Picard, D., & Secondi, J. (2008). In-stream and overland dispersal across a river network influences gene flow in a freshwater insect, *Calopteryx splendens*. *Molecular Ecology*, **17**, 3496-3505.

Chenuil, A. (2006). Choosing the right molecular genetic markers for studying biodiversity: from molecular evolution to practical aspects. *Genetica*, **127**, 101-120.

Da Silveira Queiroga D., Moura R.F., Ware J. (2019). Genetic connectivity in conservation of freshwater insects. Chapter 16. In: *Aquatic Insects*. (Eds: Del-Claro K., Guillermo R.). Springer.

Davey, J.W., & Blaxter, M.L. (2010). RADSeq: next-generation population genetics. *Briefings in Functional Genomics*, **9**, 416-423.

Didham, R., Blakely, T., Ewers, R., Hitchings, T., Ward, J., & Winterbourn, M. (2012). Horizontal and vertical structuring in the dispersal of adult aquatic insects in a fragmented landscape. *Fundamental and Applied Limnology,* **180**, 27-40.

Dubois, Q., Lebigre, C., Schtickzelle, N., Turlure, C. (2018). Sex, size and timing: Sampling design for reliable population genetics analyses using microsatellite data. *Methods in Ecology and Evolution*, **9**, 1036-1048.

Dussex, N., Chuah, A., & Waters, J.M. (2016). Genome-wide SNPs reveal fine-scale differentiation among wingless alpine stonefly populations and introgression between winged and wingless forms. *Evolution*, **70**, 38-47.

Elbrecht, V., Feld, C.K., Gies, M., Hering, D., Sondermann, M., Tollrian, R., & Leese, F. (2014). Genetic diversity and dispersal potential of the stonefly *Dinocras cephalotes* in a central European low mountain range. *Freshwater Science*, **33**, 181-191.

Engelhardt, C.H.M., Haase, P., & Pauls, S.U. (2011). From the Western Alps across Central Europe: Postglacial recolonisation of the tufa stream specialist *Rhyacophila pubescens* (Insecta, Trichoptera). *Frontiers in Zoology*, **8**, 10.

Finn, D.S., Theobald, D.M., Black, W.C., & Poff, N.L. (2006). Spatial population genetic structure and limited dispersal in a Rocky Mountain alpine stream insect. *Molecular Ecology*, **15**, 3553-3566.

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.

Fuentes-Pardo, A.P., Ruzzante, D.E. (2017). Whole-genome sequencing approaches for conservation biology: Advantages, limitations and practical recommendations. *Molecular Ecology*, **26**, 5369– 5406.

Galtier, N., Nabholz, B., GlÉMin, S., & Hurst, G.D.D. (2009). Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular Ecology,* **18**, 4541-4550.

Gardner, M.G., Fitch, A.J., Bertozzi, T., & Lowe, A.J. (2011). Rise of the machines- -recommendations for ecologists when using next generation sequencing for microsatellite development. *Molecular Ecology Resources*, **11**, 1093-1101.

Geismar, J., Haase, P., Nowak, C., Sauer, J., & Pauls, S.U. (2015). Local population genetic structure of the montane caddisfly *Drusus discolor* is driven by overland dispersal and spatial scaling. *Freshwater Biology*, **60**, 209-221.

Gibbs, H.L., Gibbs, K.E., Siebenmann, M., & Collins, L. (1998). Genetic differentiation among populations of the rare mayfly *Siphlonisca aerodromia* Needham. *Journal of the North American Benthological Society*, **17**, 464-474.

Hale, M.L., Burg, T.M., and Steeves, T.E. (2012). Sampling for microsatellite- based population genetic studies: 25 to 30 Individuals per population is enough to accurately estimate allele frequencies. *PLoS One*, 7, e45170.

Hebert, P. D., Cywinska, A., Ball, S.L., & deWaard, J.R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, **270**, 313-321.

Heino, J., Alahuhta, J., Ala-Hulkko, T., Antikainen, H., Bini, L. M., Bonada, N., …, Soininen, J. (2017). Integrating dispersal proxies in ecological and environmental research in the freshwater realm. *Environmental Reviews*. **25**: 334-349.

Helyar, S.J., Hemmer-Hansen, J., Bekkevold, D., Taylor, M.I., Ogden, R., Limborg, M.T., …, Nielsen, E.E. (2011). Application of SNPs for population genetics

of nonmodel organisms: new opportunities and challenges. *Molecular Ecology Resources*, **11**, 123-136.

Hohenlohe, PA, Funk, WC, Rajora, OP. (2021). Population genomics for wildlife conservation and management. *Molecular Ecology*, **30**, 62-82.

Hotaling, S., Muhlfeld, C.C., Joseph Giersch, J., Ali O.A., Jordan S., Miller, M.R., …, Weisrock, D.W. (2018) Demographic modelling reveals a history of divergence with gene flow for a glacially tied stonefly in a changing post-Pleistocene landscape. *Journal of Biogeography*, **45**: 304-317.

Hotaling, S., Giersch, J.J., Finn, D.S., Tronstad, L.M., Jordan, S., Serpa L.E., …, Weirosck, D.W. (2019). Congruent population genetic structure but differing depths of divergence for three alpine stoneflies with similar ecology and geographic distributions. *Freshwater Biology*, **64**, 335-347.

Hotaling, S., Kelley, J. L., & Frandsen, P. B. (2020). Aquatic insects are dramatically underrepresented in genomic research. *Insects,* **11**, 601.

Hughes, J.M. (2007). Constraints on recovery: using molecular methods to study connectivity of aquatic biota in rivers and streams. *Freshwater Biology*, **52**, 616-631.

Hughes, J.M., Bunn, S.E., Cleary, C., & Hurwood, D.A. (2000). A hierarchical analysis of the genetic structure of an aquatic insect *Bungona* (Baetidae: Ephemeroptera). *Heredity*, **85**, 561-570.

Hughes J.M., Bunn S.E., Hurwood D.A. & Cleary C. (1998). Dispersal and recruitment of *Tasiagma ciliata* (Trichoptera: Tasimiidae) in rainforest streams, south east Queensland, Australia. *Freshwater Biology*, **39**, 117-127.

Hughes, J.M., Hillyer, M., & Bunn, S.E. (2003). Small-scale patterns of genetic variation in the mayfly *Bungona narilla* (Ephemeroptera: Baetidae) in rainforest streams, south-east Queensland. *Freshwater Biology*, **48**, 709-717.

Hughes, J.M., Mather, P.B., Hillyer, M.J., Cleary, C., & Peckarsky, B. (2003). Genetic structure in a montane mayfly *Baetis bicaudatus* (Ephemeroptera: Baetidae), from the Rocky Mountains, Colorado. *Freshwater Biology*, **48**, 2149-2162.

Hughes, J.M., Mather, P.B., Sheldon, A.L., & Allendorf, F.W. (1999). Genetic structure of the stonefly, *Yoraperla brevis*, populations: the extent of gene flow among adjacent montane streams. *Freshwater Biology*, **41**, 63-72.

Hughes, J.M., Schmidt, D.J., & Finn, D.S. (2009). Genes in streams: Using DNA to understand the movement of freshwater fauna and their riverine habitat. *BioScience*, **59**, 573-583.

Hughes, J.M., Schmidt, D.J., McLean, A., & Wheatley, A. (2008). Population genetic structure in stream insects: what have we learned? In J. Lancaster & R.A. Briers (Eds.), *Aquatic insects: challenges to populations* (pp. 268-288): CAB International.

Jaccoud, D., Peng, K., Feinstein, D., & Kilian, A. (2001). Diversity arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Research*, **29**, E25.

Jackson, J.K., & Resh, V.H. (1992). Variation in genetic structure among populations of the caddisfly *Helicopsyche borealis* from three streams in northern California, USA. *Freshwater Biology*, **27**, 29-42.

Lam, A., Toussaint, E.F.A., Kindler, C., Dam M.V., Panjaitan, R., Roderick, G.K., Balke, M. (2018). Stream flow alone does not predict population structure of diving beetles across complex tropical landscapes. *Molecular Ecology*, **27**, 3541-3554.

Landguth, E.L., Fedy, B.C., Oyler-Mccance, S.J., Garey, A.L., Emel, S.L., Mumma, M., …, Cushman, S.A. (2012). Effects of sample size, number of markers, and allelic richness on the detection of spatial genetic pattern. *Molecular Ecology Resources*, **12**, 276-284.

Keller, D., Van Strien, M.J. and Holderegger, R. (2012). Do landscape barriers affect functional connectivity of populations of an endangered damselfly? *Freshwater Biology*, **57**, 1373-1384.

Kelly, L.C., Bilton, D.T., & Rundle, S.D. (2001). Population structure and dispersal in the Canary Island caddisfly *Mesophylax aspersus* (Trichoptera, Limnephilidae). *Heredity*, **86**, 370-377.

Kelly, L.C., Rundle, S.D., & Bilton, D.T. (2002). Genetic population structure and dispersal in Atlantic Island caddisflies. *Freshwater Biology*, **47**, 1642-1650.

Kim K.S., Sappington T.W. (2013) Microsatellite data analysis for population genetics. Chapter 19. In: Kantartzi S. (eds) *Microsatellites*. Methods in Molecular Biology (Methods and Protocols). Humana Press, Totowa, NJ.

Kirk, H., & Freeland, J. (2011). Applications and implications of neutral versus non-neutral markers in molecular ecology. *International Journal of Molecular Sciences*, **12**, 3966-3988.

Krosch, M.N., Baker, A.M., Mather, P.B., & Cranston, P.S. (2011). Spatial population genetic structure reveals strong natal site fidelity in *Echinocladius martini* (Diptera: Chironomidae) in northeast Queensland, Australia. *Freshwater Biology*, **56**, 1328-1341.

Mackay, R.J. (1992). Colonization by lotic macroinvertebrates: a review of processes and patterns. *Canadian Journal of Fisheries and Aquatic Sciences*, **49**, 617-628.

Manel, S., Schwartz, M.K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution*, **18**, 189-197.

Mardis, E.R. (2008). The impact of next-generation sequencing technology on genetics. *Trends in Genetics*, **24**, 133-141.

Mastretta-Yanes, A., Arrigo, N., Alvarez, N., Jorgensen, T.H., Piñero, D., & Emerson, B.C. (2015). Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Molecular Ecology Resources*, **15**, 28-41.

Mazzucco, R., Van Nguyen, T., Kim, D.H., Chon, T.S., & Dieckmann, U. (2015). Adaptation of aquatic insects to the current flow in streams. *Ecological Modelling*, **309-310**, 143-152.

McCulloch, G.A., Wallis, G.P., & Waters, J.M. (2009). Do insects lose flight before they lose their wings? Population genetic structure in subalpine stoneflies. *Molecular Ecology*, **18**, 4073-4087.

McLean, A. (1967). Studies of Ephemeroptera in the Auckland area I: light trapping in Cascade Kauri Park. *Tane*, 13, 99-105.

Mendelson, T.C., & Shaw, K.L. (2005). Use of AFLP markers in surveys of arthropod diversity. *Methods in Enzymology*, **395**, 161-177.

Miller, M.P., Blinn, D.W., & Keim, P. (2002). Correlations between observed dispersal capabilities and patterns of genetic differentiation in populations of four aquatic insect species from the Arizona White Mountains, USA. *Freshwater Biology*, **47**, 1660-1673.

Monaghan, M.T., Spaak, P., Robinson, C.T., & Ward, J.V. (2001). Genetic differentiation of *Baetis alpinus* Pictet (Ephemeroptera: Baetidae) in fragmented alpine streams. *Heredity*, **86**, 395-403.

Monaghan, M.T., Spaak, P., Robinson, C.T., & Ward, J.V. (2002). Population genetic structure of three alpine stream insects: influences of gene flow, demographics, and habitat fragmentation. *Journal of the North American Benthological Society*, **21**, 114-131.

Monaghan, M.T., & Sartori M. (2009) Genetic contributions to the study of taxonomy, ecology, and evolution of mayflies (Ephemeroptera): review and future perspectives. *Aquatic Insects*, **3**, 19-39.

Moritz, C., Dowling, T.E., & Brown, W.M. (1987). Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Annual Review of Ecology and Systematics*, **18**, 269-292.

Múrria, C., Bonada, N., Ribera, C., & Prat, N. (2010). Homage to the Virgin of Ecology, or why an aquatic insect unadapted to desiccation may maintain populations in very small, temporary Mediterranean streams. *Hydrobiologia*, **653**, 179-190.

Múrria, C., Sáinz-Bariáin, M., Vogler, A.P., Viza, A., González, M. and Zamora-Muñoz, C. (2020), Vulnerability to climate change for two endemic high-elevation, low-dispersive *Annitella* species (Trichoptera) in Sierra Nevada, the southernmost high mountain in Europe. *Insect Conservation and Diversity*, **13**, 283-295.

Nabholz, B., Mauffrey, J.F., Bazin, E., Galtier, N., & Glemin, S. (2008). Determination of mitochondrial genetic diversity in mammals. *Genetics*, **178**, 351-361.

Ouborg, N.J., Piquot, Y., & Van Groenendael, J.M. (1999). Population genetics, molecular markers and the study of dispersal in plants. *Journal of Ecology*, **87**, 551-568.

Parkyn, S.M., & Smith, B.J. (2011). Dispersal constraints for stream invertebrates: setting realistic timescales for biodiversity restoration. *Environmental Management*, **48**, 602-614.

Pauls, S.U., Alp, M., Bálint, M., Bernabò, P., Čiampor, F., Jr, Čiamporová-Zaťovičová, Z., … Monaghan, M.T. (2014). Integrating molecular tools into freshwater ecology: developments and opportunities. *Freshwater Biology*, **59**, 1559-1576.

Pessino, M., Chabot, E.T., Giordano, R., & DeWalt, R.E. (2014). Refugia and postglacial expansion of *Acroneuria frisoni* Stark & Brown (Plecoptera: Perlidae) in North America. *Freshwater Science*, **33**, 232-249.

Pfeiffer, F., Gröber, C., Blank, M., Händler, K., Beyer, M., Schultze, J.L., & Mayer, G. (2018). Systematic evaluation of error rates and causes in short samples in next-generation sequencing. *Scientific Reports*, **8**, 10950.

Phillipsen, I.C., Kirk, E.H., Bogan, M.T., Mims, M.C., Olden, J.D., & Lytle, D.A. (2015). Dispersal ability and habitat requirements determine landscape-level genetic patterns in desert aquatic insects. *Molecular Ecology*, **24**, 54-69.

Phillipsen, I.C., & Lytle, D.A. (2013). Aquatic insects in a sea of desert: population genetic structure is shaped by limited dispersal in a naturally fragmented landscape. *Ecography*, **36**, 731-743.

Polato, N.R., Gray, M.M., Gill, B.B., Becker, C.G., Casner, K.L., Flecker, A.S., …, Zamudio, K.R. (2017). Genetic diversity and gene flow decline with elevation in montane mayflies. *Heredit*y, **119**, 107-116.

Poff, N.L., Larson, E.I., Salerno, P.E., Morton, S.G., Kondratieff, B.C., Flecker, A.S., Zamudio, K.R. and Funk, W.C. (2018). Extreme streams: species persistence and genomic change in montane insect populations across a flooding gradient. *Ecology Letters*, **21**, 525-535.

Preziosi, R.F., & Fairbairn, D.J. (1992). Genetic population structure and levels of gene flow in the stream dwelling Waterstrider, *Aquarius (=Gerris) remigis* (Hemiptera, Gerridae). *Evolution*, **46**, 430-444.

Rader, R.B., Unmack, P.J., Christensen, W.F., & Jiang, X.B. (2019). Connectivity of two species with contrasting dispersal abilities: a test of the isolated tributary hypothesis. *Freshwater Science*, **38**, 142-155.

Rahman, S., Schmidt, D., Hughes, J. (2018). De novo SNP discovery and strong genetic structuring between upstream and downstream populations of *Paratya australiensis* Kemp, 1917 (Decapoda: Caridea: Atyidae). *Journal of Crustacean Biology*, **38**, 166172.

Ratnasingham, S., & Hebert, P.D.N. (2007). BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). *Molecular Ecology Notes*, **7**, 355-364.

Rebora, M., Lucentini, L., Palomba, A., Panara, F., & Gaino, E. (2005). Genetic differentiation among populations of *Baetis rhodani* (Ephemeroptera, Baetidae) in three Italian streams. *Italian Journal of Zoology*, **72**, 121-126.

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from F-Statistics under isolation by distance. *Genetics*, **145**, 1219-1228.

Sabando, M.C., Vila, I., Penaloza, R., & Veliz, D. (2011). Contrasting population genetic structure of two widespread aquatic insects in the Chilean high-slope rivers. *Marine and Freshwater Research*, **62**, 1-10.

Saito, R., & Tojo, K. (2016). Comparing spatial patterns of population density, biomass, and genetic diversity patterns of the habitat generalist mayfly *Isonychia japonica* Ulmer (Ephemeroptera: Isonychiidae) in the Chikuma-Shinano River basin. *Freshwater Science*, **35**, 724-737.

Schmidt, S.K., Hughes, J.M., & Bunn, S.E. (1995). Gene flow among conspecific populations of *Baetis* sp. (Ephemeroptera) - adult flight and larval drift. *Journal of the North American Benthological Society*, **14**, 147-157.

Schmidt, D., Huey, J. A., Hughes, J.M. (2018). Genome-Wide SNPs identify limits to connectivity in the extreme freshwater disperser, spangled perch *Leiopotherapon unicolor* (Terapontidae). *Journal of Heredity*, **109**, 320-325.

Schröder, O., Cavanaugh, K.K., Schneider, J. V., Schell T., Bonada N., Seifert L., Pauls S.U. (2021). Genetic data support local persistence in multiple glacial refugia in the montane net-winged midge *Liponeura cinerascens cinerascens* (Diptera, Blephariceridae). *Freshwater Biology*, **66**, 859-868.

Schultheis, A.S., Booth, J.Y., Perlmutter, L.R., Bond, J.E., & Sheldon, A.L. (2012). Phylogeography and species biogeography of montane Great Basin stoneflies. *Molecular Ecology*, **21**, 3325-3340.

Schultheis, A.S., Hendricks, A.C., & Weigt, L.A. (2002). Genetic evidence for 'leaky' cohorts in the semivoltine stonefly *Peltoperla tarteri* (Plecoptera: Peltoperlidae). *Freshwater Biology*, **47**, 367-376.

Schultheis A.S. & Hughes J.M. (2005). Spatial patterns of genetic structure among populations of a stone-cased caddis (Trichoptera: Tasmiidae) in south-east Queens- land, Australia. *Freshwater Biology*, **50**, 2002-2010.

Semagn, K., Bjørnstad, Å., & Ndjiondjop, M. (2006). An overview of molecular marker methods for plants. *African Journal of Biotechnology*, **525**, 2540-2568.

Shama, L.N.S., Kubow, K.B., Jokela, J., & Robinson, C.T. (2011). Bottlenecks drive temporal and spatial genetic changes in alpine caddisfly metapopulations. *BMC Evolutionary Biology*, **11**, 278.

Slatkin, M. (1985). Gene flow in natural populations. *Annual Review of Ecology and Systematics*, **16**, 393-430.

Smith, P.J., & Collier, K.J. (2001). Allozyme diversity and population genetic structure of the caddisfly *Orthopsyche fimbriata* and the mayfly *Acanthophlebia cruentata* in New Zealand streams. *Freshwater Biology*, **46**, 795-805.

Smith, P.J., McVeagh, S.M., & Collier, K.J. (2006). Genetic diversity and historical population structure in the New Zealand mayfly *Acanthophlebia cruentata*. *Freshwater Biology*, **51**, 12-24.

Smith, P.J., & Smith, B.J. (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**, 723-734.

Spitzer, B. (2014). Genetic structure in populations of *Baetis tricaudatus* in the San Bernardino mountains. *Western North American Naturalist*, **74**, 434-445.

Sproul, J.S., Houston, D.D., Davis, N., Barrington, E., Oh, S.Y., Evans, R.P., & Shiozawa, D.K. (2014). Comparative phylogeography of codistributed aquatic insects in western North America: insights into dispersal and regional patterns of genetic structure. *Freshwater Biology*, **59**, 2051-2063.

Stefanello, F., Menezes, R.S.T., Ribeiro, J.R., Almeida, E.A.B. (2020). Widespread gene flow model explains the genetic–morphological variation in a giant water bug species under fine-scale spatial sampling. *Annals of the Entomological Society of America*, **113**, 160-170.

Supple, M.A., Shapiro, B. (2018). Conservation of biodiversity in the genomics era. *Genome Biology,* **19**, 131.

Sweeney, B.W., Funk, D.H., & Vannote, R.L. (1987). Genetic variation in stream mayfly (Insecta: Ephemeroptera) populations of eastern North America. *Annals of the Entomological Society of America*, **80**, 600-612.

Takenaka, M., Tokiwa, Tojo, K. (2019). Concordance between molecular biogeography of *Dipteromimus tipuliformis* and geological history in the local fine scale (Ephemeroptera, Dipteromimidae). *Molecular Phylogenetics and Evolution*, **139**, 106547.

Tautz, D., & Renz, M. (1984). Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Research*, **12**, 4127-4138.

Taylor C.L., Barker N.P., Barber-James H.M., Villet M.H., Pereira-da-Conceicoa L.L. (2020). Habitat requirements affect genetic variation in three species of mayfly (Ephemeroptera, Baetidae) from South Africa. *ZooKeys*, **936**, 1-24.

Thomas, A. (2011). Ephéméroptères du Sud-Ouest de la France. Migrations d'imagos à haute altitude. *Annales de Limnologie*, **11**, 47-66.

Trense, D., Schmidt, T.L., Yang, Q., Chung, J., Hoffmann, A.A., Fischer, K. (2021). Anthropogenic and natural barriers affect genetic connectivity in an Alpine butterfly. *Molecular Ecology*, **30**, 114-130.

Turlure, C., Vandewoestijne, S., & Baguette, M. (2014). Conservation genetics of a threatened butterfly: comparison of allozymes, RAPDs and microsatellites. *BMC Genetics*, **15**, 114.

Vieira, N.K.M., LeRoy Poff, N., Carlisle, D., Moulton II, S.R., Koski, M.L., Kondratieff, B.C. (2006). A database of lotic invertebrate traits for North America. *U.S. Geological Survey Data Series*, 187.

Vignal, A., Milan, D., SanCristobal, M., & Eggen, A. (2002). A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics Selection Evolution*, **34**, 275-305.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee, T.A.J., Hornes, M., …, Kuiper, M. (1995). AFLPs a new technique for DNA fingerprinting. *Nucleic Acids Research*, **23**, 4407-4414.

Walton, Z., Hagenlund, M., Østbye, K. Samelius, G., Odden., G., Norman, A., …, Spong, G. (2021). Moving far, staying close: red fox dispersal patterns revealed by SNP genotyping. *Conservation Genetics*, **22**, 249-257.

Watanabe K., Kazama S., Omura T., Monaghan M.T. (2014). Adaptive Genetic Divergence along narrow environmental gradients in four stream insects. *PLoS One*, **9**, e93055.

Waters, J.M., Emerson, B.C., Arribas, P., McCulloch, G.A. (2020). Dispersal reduction: causes, genomic mechanisms, and evolutionary consequences. *Trends in Ecology and Evolution*, **35**, 512-522.

Weigand, H., Weiss, M., Cai, H., …, Leese, F. (2017). Deciphering the origin of mito-nuclear discordance in two sibling caddisfly species. *Molecular Ecology*, **26**, 5705-5715.

Weigand, H., Weiss, M., Cai, H., Li, Y., Yu, L., Zhang, C., & Leese, F. (2018). Fishing in troubled waters: Revealing genomic signatures of local adaptation in response to freshwater pollutants in two macroinvertebrates. *Science of the Total Environment,* **633**, 875-891.

Weir, B.S., & Cockerham, C.C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358-1370.

Wickson, S.J., Chester, E.T., Valenzuela, I., Halliday, B., Lester, R. E., Matthews, T. G., & Miller, A. D. (2014). Population genetic structure of the Australian caddisfly *Lectrides varians* Mosely (Trichoptera: Leptoceridae) and the identification of cryptic species in south-eastern Australia. *Journal of Insect Conservation*, **18**, 1037-1046.

Wilcock, H.R., Bruford, M.W., Nichols, R.A., & Hildrew, A.G. (2007). Landscape, habitat characteristics and the genetic population structure of two caddisflies. *Freshwater Biology*, **52**, 1907-1929.

Wilcock, H.R., Hildrew, A.G., & Nichols, R.A. (2001). Genetic differentiation of a European caddisfly: past and present gene flow among fragmented larval habitats. *Molecular Ecology*, **10**, 1821-1834.

Wilcock, H.R., Nichols, R.A., & Hildrew, A.G. (2003). Genetic population structure and neighbourhood population size estimates of the caddisfly *Plectrocnemia conspersa. Freshwater Biology*, **48**, 1813-1824.

Willi, Y., Kristensen, T. N., Sgrò, C. M., Weeks, A. R., Ørsted, M., & Hoffmann, A. A. (2022). Conservation genetics as a management tool: The five best-supported paradigms to assist the management of threatened species. *Proceedings of the National Academy of Sciences of the United States of America,* **119**, e2105076119.

Wright, S. (1943). Isolation by distance. *Genetics*, **28**, 114-138.

Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics*, **15**, 323-354.

Wubs, J., Fraaije, R., Groot, A., Erkens, R., Garssen, A., Kleyheeg, E., …, Soons, M. (2016). Going against the flow: A case for upstream dispersal and detection of uncommon dispersal events. *Freshwater Biology,* **61**, 580-595.

Yaegashi, S., Watanabe, K., Monaghan, M.T., & Omura, T. (2014). Fine-scale dispersal in a stream caddisfly inferred from spatial autocorrelation of microsatellite markers. *Freshwater Science*, **33**, 172-180.

Young, B.A., Schmidt, D.J., & Sheldon, F. (2013). Small-scale patterns of genetic variation in a headwater specialist mayfly: No influence of selective forest harvesting on diversity. *Austral Ecology*, **38**, 504-515.

Zane, L., Bargelloni, L., & Patarnello, T. (2002). Strategies for microsatellite isolation: a review. *Molecular Ecology*, **11**, 1-16.

Zhang, D.X., & Hewitt, G.M. (2003). Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology*, **12**, 563-584.

Zickovich, J.M., & Bohonak, A.J. (2007). Dispersal ability and genetic structure in aquatic invertebrates: a comparative study in southern California streams and reservoirs. *Freshwater Biology*, **52**, 1982-1996.

CHAPTER THREE

ASSESSING POPULATION GENETIC STRUCTURE OF THREE NEW ZEALAND STREAM INSECTS USING MITOCHONDRIAL AND NUCLEAR DNA MARKERS

Zelandobius confusus nymph (top) and *Hydropsyche fimbriata* larvae.
Photos: Vanessa Barbosa

3.1 Abstract

Assessing genetic differentiation among natural populations can aid understanding of dispersal patterns and connectivity among habitats. Several molecular markers have become increasingly popular in determining population genetic structure for this purpose. Here, we compared the resolution of mitochondrial cytochrome c oxidase subunit I (COI) and nuclear single nucleotide polymorphism (SNP) markers for detecting population structure among stream insects at small spatial scales. Individuals of three endemic taxa - *Coloburiscus humeralis* (Ephemeroptera), *Zelandobius confusus* (Plecoptera), and *Hydropsyche fimbriata* (Trichoptera) - were collected from forested streams that flow across open pasture in the North Island of New Zealand. Both COI and SNP data indicated limited population structure across the study area, and small differences observed among these species were likely related to their putative dispersal abilities. For example, fine-scale genetic differentiation between and among neighbouring stream populations for *H. fimbriata* suggests that gene flow, and hence dispersal, may be more limited for this species relative to the others. Based on the generally similar results provided by both markers, we suggest that either COI or SNP markers can provide suitable initial estimates of fine-scale population genetic differentiation in stream insects.

3.2 Introduction

Understanding how individuals move between populations across the landscape is an important component of the conservation and management of natural populations and their habitats (Alp et al., 2012, Venables et al., 2021, Weston et al., 2020). Genetic markers can be used to infer levels of gene flow, providing an indirect measure of dispersal and connectivity among populations (Driscoll et al., 2014, Saastamoinen et al., 2018). This approach assumes that when genetic differences among populations are high (measured as genetic differentiation), gene flow/dispersal is restricted. In contrast, where populations show lower levels of genetic differentiation, they are assumed to have higher rates of dispersal (gene flow) among habitats (Hughes, 2007, Slatkin, 1985). In stream ecosystems, the extent of gene flow occurring within and among aquatic populations is determined by the dispersal abilities and life histories of resident taxa, the dendritic structure of stream networks, and physical features of the local landscape such as riparian cover (Hughes et al., 2009). Analysis of spatial patterns of genetic differentiation can further determine processes of population connectivity and the effects of landscape features on population genetic structure (Alexander et al., 2011, Engelhardt et al., 2011, Wilcock et al., 2007).

Different genetic markers and associated sequencing techniques offer a range of options to investigate population genetic structure. Single locus, maternally inherited mitochondrial DNA (mtDNA) markers, for example, are among the most widely used to assess evolutionary and contemporary processes driving population connectivity in stream insects. In particular, extensive use of the cytochrome c oxidase subunit 1 (COI) gene has broadened understanding of species diversity, phylogeography, and dispersal patterns (Monaghan & Sartori, 2009). COI data are often obtained through traditional Sanger sequencing (Sanger et al., 1977, Smith et al., 1986), a method characterised by ease-of-use and rapid turnaround (Heather & Chain, 2016) that remains useful for initial investigations of genetic differentiation and/or analysis of a small number of samples (De Cario et al., 2020).

However, current methods offering higher resolution are more suitable for the detection of fine-scale population structure, for example within watersheds or proximate stream networks (Hotaling et al., 2018; Taylor et al., 2020). Among these, genome-wide single nucleotide polymorphisms (SNPs) are a popular marker

of choice following the advent and reduced costs of next generation sequencing (NGS) technologies. Genome-wide SNPs represent mainly neutral, bi-parentally inherited markers and show high polymorphism among individuals (Mardis, 2017). Restriction-site Associated DNA sequencing, or RADseq (Davey & Blaxter, 2010) has been an increasingly-used NGS technique for SNP discovery as part of population genetic studies (e.g. Polato et al., 2017, Trense et al., 2020). RADseq uses restriction enzymes to cut DNA into short fragments, followed by simultaneous sequencing and SNP discovery (Mardis, 2008). The main advantages of this method include: the massively parallel sequencing across many individuals at a reduced cost; the production of highly reproducible data; and the application to species with limited, or no existing sequence data (Davey & Blaxter, 2010, Baird et al., 2008). Among natural populations, RADseq data has enabled resolution of fine spatial patterns of genetic variation and recent population-level differentiation (Vendrami et al., 2017, Wang et al., 2020, Dussex et al., 2016), as well as the identification of recent barriers to gene flow or changes in population structure (Devlin-Durante & Baums, 2017).

For aquatic insects, the incorporation of both mitochondrial and nuclear DNA markers can improve understanding of species boundaries. For example, in caddisflies within the *Apatania zonella* group, similarity of COI sequences within species and slightly diverged genome-wide SNP data between species indicated close relationships among nevertheless genetically distinct taxa (Salokannel et al., 2021). In contrast, distinct COI lineages without differentiated SNP patterns among *Limnephilus* species suggested a lack of cryptic diversity in this taxa (Salokannel et al., 2021). Such combined use of genetic markers may provide a more comprehensive assessment of dispersal and connectivity more generally, particularly since single organelle mtDNA markers can provide limited inference on smaller spatial scales (e.g. McGaughran et al., 2019, Pazmiño et al., 2017).

Here, we examined population connectivity and dispersal patterns for three New Zealand endemic stream insects, comparing the resolution of COI and SNP data for detecting population structure at small spatial scales (~11 km). We selected one species from each of the commonly encountered freshwater insect orders, Ephemeroptera, Plecoptera and Trichoptera, based on their putative dispersal abilities: the mayfly *Coloburiscus humeralis* (Coloburiscidae), the stonefly *Zelandobius confusus* (Gripopterygidae)*,* and the caddisfly *Hydropsyche fimbriata*

(Hydropsychidae). Previous studies for *C. humeralis* using allozymes and nuclear AFLP data have provided conflicting results, with low genetic differentiation among populations even across distant catchments using allozyme data (Hogg et al., 2002), versus significant differentiation detected using AFLP between populations <2 km (Wallace, 2013). Population genetic structure in *Z. confusus* has yet to be assessed, although recent research on other winged New Zealand stonefly taxa suggests that they have moderate flight capabilities and dispersal among streams at regional scales (Dussex et al., 2016). For *H. fimbriata*, a pattern of isolation by distance has been found with increasing levels of genetic differentiation at increasing spatial scales (Smith & Collier, 2001, Smith et al., 2006, Smith and Smith, 2009). Our aim was to compare the resolution of COI and nuclear SNP markers for detecting population structure among streams over small spatial scales.

We hypothesised that patterns of population genetic structure would reflect differential dispersal abilities among the studied species and that the genome-wide SNP data would detect finer-scale genetic differentiation relative to the single-locus mtDNA marker.

3.3 Methods

3.3.1 Study area and insect collection

Caddisfly larvae and stonefly and mayfly nymphs were collected between December 2017 and January 2019 from 11 sites within three streams (Ngakoaohia, Tawhitiwhiti, Te Pahu streams), of which two (Tawhitiwhiti and Te Pahu) were in neighbouring catchments in Pirongia Forest Park (Fig. 3.1; Table 3.1), under collection authorisation from the New Zealand Department of Conservation (permit number 68083-FAU). Part of the Alexandra Volcanic Group, Mount Pirongia was formed by a succession of eruptions between ~2.5-1.6 million years ago (Kear, 1960, McLeod et al., 2020). The local landscape is characterised by forest fragments surrounded by intense agricultural production, predominantly dairy and sheep farmland. Samples were collected from first and second order permanent and stony-bottom streams that all flowed from indigenous forest before entering open farmland.
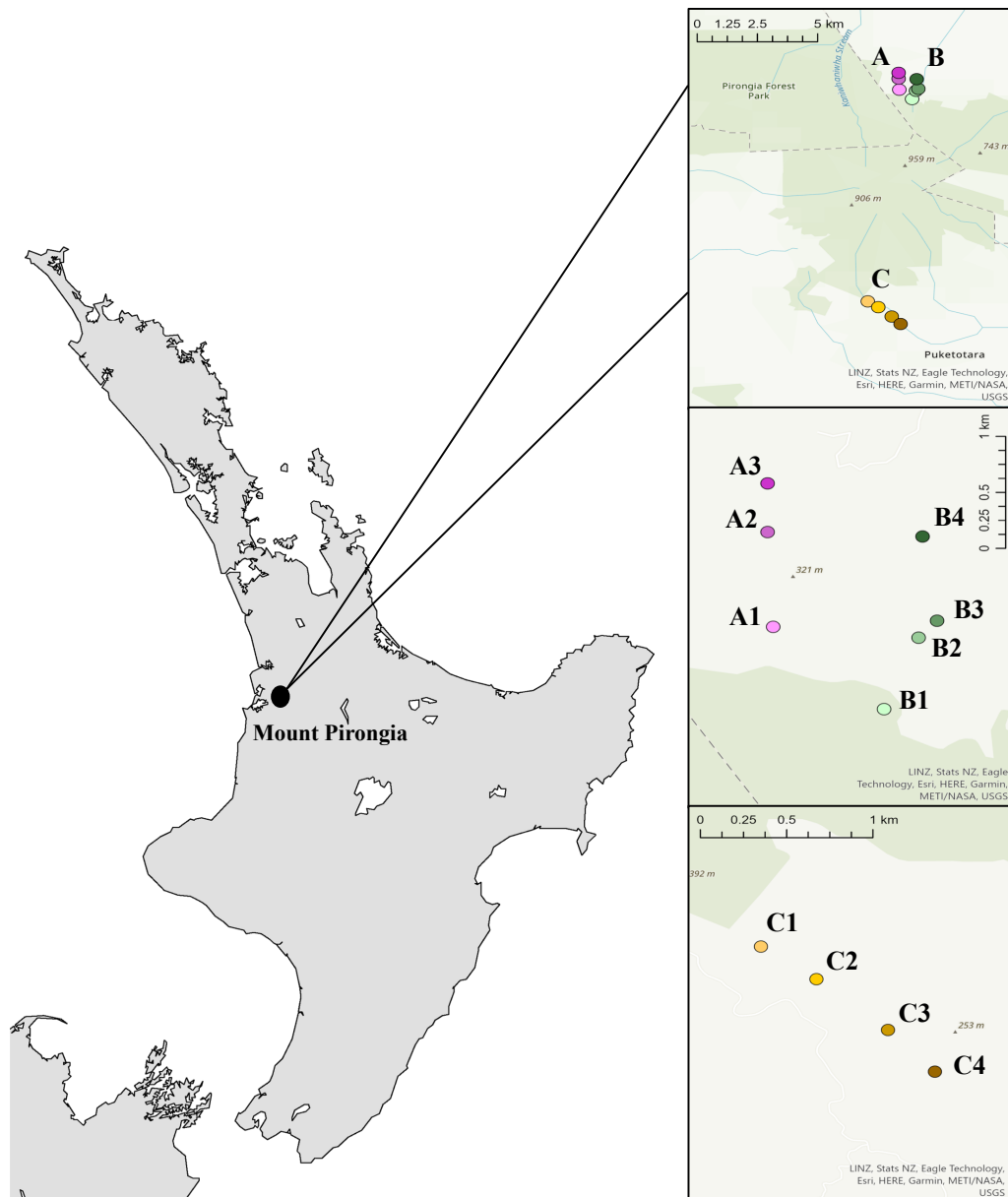
Figure 3.1 Maps showing the 11 sampling sites distributed across three streams in Pirongia (North Island, New Zealand). Stream A is Tawhitiwhiti, Stream B is Te Pahu, and Stream C is Ngakoaohia.

For each species, approximately 50 individuals were collected from three to four sites within each of the three streams. Samples were collected using a kick-net or directly hand-picked from the substrate, and immediately preserved in 95% ethanol. *Zelandobius confusus* nymphs were identified using McLellan (1993) and *C. humeralis* and *H. fimbriata* were confirmed using Winterbourn et al. (2000). The nomenclature of *H. fimbriata* follows Geraci et al. (2010). Following morphological identification, samples were transferred to vials with fresh 95%

ethanol and stored at -20°C until the left rear leg was dissected from each individual using sterilised forceps and added to a single well of a 96-well PCR plate for DNA extraction.

3.3.2 Study taxa

The mayfly *C. humeralis* is widespread in New Zealand, inhabiting the underside of stones predominantly in riffles (Wisely, 1961). The nymph stage ranges from 12–27 months (Harding & Winterbourn, 1993). Although no specific information is available on adult longevity for *C. humeralis*, most adult mayfly taxa are comparatively short-lived (with 14 days the maximum recorded for New Zealand species; Smith B., unpublished data).

The stonefly *Z. confusus* is widely distributed throughout New Zealand (McLellan 1993). The nymph stage lasts 9–12 months (McLellan, 1993), while adults can live for a few days to weeks (Collier & Smith, 2000), and are often found on riparian vegetation (Smith & Collier, 2000; Winterbourn, 2005). In our study, *Z. confusus* nymphs were usually found in leaf packs that had accumulated or become trapped against rocks or woody debris within the stream.

The caddisfly *H. fimbriata* is restricted to the North Island of New Zealand. Larvae are found mostly in native forest streams, where they build fixed retreats attached to stable substrates and spin nets to filter drifting food particles/detritus (Cowley, 1978). The larval stage is 9–12 months (Cowley, 1978), while adult longevity is approximately 13 days (Smith, B., unpublished data).

3.3.3 DNA extraction, amplification and sequencing

DNA sequencing of both mtDNA COI and genome-wide SNP markers was conducted at the Canadian Centre for DNA Barcoding (University of Guelph, Canada). Extraction and sequencing of mtDNA followed standard protocols (Ivanova et al., 2006). In brief, DNA was extracted following the AcroPrepTM PALL Glass Fibre plate method using a total mix of 5 ml insect lysis buffer (0.5 ml of Proteinase K, 20 mg/ml per 96-well plate). A 658 bp region of the COI gene was PCR amplified using the primer pair LepF1 and LepR1 (Hebert et al., 2004, Wilson, 2012) and 5 µl of the DNA extraction product. PCR thermal cycling conditions were: initial denaturation of samples at 94°C for 1 min, followed by five cycles of

94°C for 30 s, 48°C for 1.5 min, and 72°C for 1 min. This was followed by 35 cycles of 94°C for 30 s, 52°C for 1 min, and 72°C for 1 min; with a final extension of 72°C for 10 min. PCR products were cleaned using Sephadex® and then sequenced using an ABI3730xl DNA analyser. All DNA sequence data have been added to the Barcode of Life Datasystems (BOLD) database (Ratnasingham & Hebert, 2007) and are available under dataset DS-EPTNZNI and cross referenced to GenBank under concession numbers 0K502554: 0K502876.

For RADseq library construction, 9–10 samples were selected from each sampling site for which COI data were also available. DNA was extracted using DNeasy Blood and Tissue Kits (Quiagen Inc., Valencia, CA, USA) following the manufacturer's instructions. All extractions were visualised on a 1% agarose gel and quantified using an AccuClear UltraHigh Sensitivity dsDNA Quantification Kit (Biotium, Inc.) and SpectraMax M2 microplate reader. Extracted DNA was incubated with the GATC cut site restriction enzyme, *DpnII*, in 50 µl reactions, for 3 h at 37°C followed by 20 min at 65°C (Knapp et al., 2016). Samples were then cleaned with Ampure XP beads in a 1:1.8 ratio of DNA:beads and libraries were generated using KAPA HyperPrep Library Preparation Kit (Roche), including a 350–700 bp size-selection and PCR steps based on manufacturers recommendations. All libraries were sequenced as 300 bp paired-end reads on the Illumina MiSeq platform at the Biodiversity Genomics Facility at the University of Guelph. Raw and demultiplexed reads for each sample were then used for downstream analysis.

3.3.4 Mitochondrial DNA (COI) data analysis

Sequences were aligned in ClustalX v. 2.1 (Larkin et al., 2007) and, if necessary, ends were trimmed in Jalview v. 2.11.1.3 (Waterhouse et al., 2009). Haplotype (h) and nucleotide (π) diversity indices (Nei, 1987) were calculated in ARLEQUIN v. 3.0 (Sd et al., 2000). For visualisation of genetic structure, haplotype networks assigning individual sequences to their respective sampling site locations were constructed. Analyses were conducted in R v. 1.3.1 (R core team, 2020) using the packages ape v. 5.5 (Paradis and Schliep, 2018) and pegas v. 1.0 (Goudet, 2005). Global and pairwise measures of genetic distance (Φ_{ST}) among populations were estimated between all 11 sampling sites to infer patterns of

population structure for each species. These statistics were calculated using pairwise comparisons and analysis of molecular variance (AMOVA; Excoffier et al., 1992) in ARLEQUIN, with the statistical significance of variance components obtained from 10,000 permutations.

3.3.5 Single Nucleotide Polymorphism (SNP) data analysis

Before processing, the quality of raw Illumina data was checked using FastQC v. 0.11.8 (Andrews, 2010). Example FastQC reports for raw and processed data are shown in Appendix Figure A.3.1. IPYRAD v. 0.7.28 (Eaton and Overcast, 2020) was used to filter and remove low quality data, identify homology among reads through *de novo* assembly, make SNP calls, and format output files for each species dataset. Only forward reads (R1) were used in the pipeline to enhance computational performance. Additionally, to determine final filtering settings in IPYRAD, certain default settings were changed to explore the effect of several parameters on the final amount and quality of reads, including: filter_min_trim_len (which sets the final minimum length after filtering; we tested values of 60 and 150, default: 35), and trim_reads (trims raw read edges; we tested values of 10, 150, 0, 0; 10, -140, 0, 0; and 10, 150, 0, 0 to set a minimum length of 150bp, default: 0, 0, 0, 0). Although these filters had little effect in terms of removing low quality data or changing preliminary phylogenetic tree structure, the number of SNPs retained varied when restricting the minimal read length or trimming all reads to a uniform length. Thus, to obtain the maximum number of SNPs while removing low quality data, reads were processed with the following non-default parameter settings: filter_adapters (2, where adapters are removed), filter_min_trim_len (60), and trim_reads (10, -140, 0, 0). The total number of retained reads and filtered data for the three species are summarised in Appendix Table A.3.1.

The SNP dataset was filtered using VCFTOOLS v. 0.1.13 (Danecek et al., 2011), with --missing-indv, --max-missing-count and --maf parameters applied to explore the effects of missing genotypes and minor allele frequency (MAF) on a per-individual basis. We proceeded with two datasets: (1) a less conservative dataset created by filtering individuals with >98% missing data; and (2) a more conservative dataset created by setting missing genotype data across all individuals to 20% and applying a MAF cut-off of 5% in addition to applying the 98% missing

data threshold. As no significant differences were observed in population structure analysis among the two filtered datasets, we hereafter present results for (2).

Principal Component Analysis (PCA) was performed using PLINK v. 3 (Purcell et al., 2007) to examine population differentiation based on orthogonal transformation of the SNP data. PCA output files were then used to plot results using the R package tidyverse v. 1.3.1 (Wickham et al., 2019). PLINK was used to convert VCF files to a format suitable for use in the clustering software, fastSTRUCTURE v. 1.0 (Raj et al., 2014), which was used to identify admixture proportions among individuals and populations. This software indentifies groups of individuals based on their allele frequency profiles by placing individuals into groups whose members share similar patterns of genetic variation (Pritchard et al., 2000). In fastSTRUCTURE, we tested 1–11 genetic clusters (depending on the availability of data from each sampling site/population) by specifying the K parameter (where K=number of genetic clusters), and ran five replicates for each K-value, using a simple prior. Results for each K-value were visualised, and model complexity chosen using the distruct.py and chooseK.py scripts from the fastSTRUCTURE program, respectively.

Finally, pairwise comparisons of genetic distance between all sampling sites were calculated using Φ_{ST} (Weir and Cockerham, 1984) in the R package StAMPP v. 1.6.2 (Pembletom et al., 2013). Global Φ_{ST} over all populations was estimated using AMOVA in the R package poppr v. 2.8.1 (Kamvar et al., 2014).

3.4 Results

Coloburiscus humeralis (Ephemeroptera)

COI: Trimmed sequences of 549 bp were obtained for 78 (46%) of the 171 processed specimens of *C. humeralis* (Table 3.2). A total of 10 haplotypes were identified and haplotype diversity ranged from 0.181 to 0.809 across the 11 sampling sites. Nucleotide diversity was low for all sites (0.001—0.004; Table 3.2). Seven of the 10 haplotypes were singletons (found in a single individual), with only three haplotypes found in five or more individuals (Fig. 3.2). Haplotype 1 (H1) was the most frequent (n=59) and was present at all sampling sites (Table 3.4; Fig. 3.2). Consistent with the lack of population structure observed in the haplotype network, AMOVA analysis revealed limited genetic differentiation across the study area

(global Φ_{ST}=0.010, P >0.05; Table 3.3). Similarly, most of the pairwise comparisons of Φ_{ST} between sampling sites did not differ significantly from zero (Table 3.5).

SNPs: The RADseq library produced a total of 43,363,740 reads from 19 specimens collected at seven sampling sites within the Pirongia range (Table 3.5). After SNP identification and the application of a conservative filtering pipeline, 41,029 SNPs from all 19 individuals were recovered. A visual representation of the sequencing depth per sample and the location of missing data in the SNP matrix (72.8%) is provided in Appendix Fig. A.3.2. Pairwise comparisons of Φ_{ST} were not calculated for most populations due to small sample sizes (n<3, Table 3.5). However, genetic differentiation among the remaining populations was low (global Φ_{ST}=0.017, P >0.05), and consistent with the COI data (Table 3.3). The fastSTRUCTURE analysis indicated that values of K between 1—2 maximized the marginal likelihood. Results are provided for K=2, showing two major genetic groups across seven sampling sites, with each genetic group present in all streams and a lack of admixture found within individuals (Fig. 3.3). The first two components of the PCA explained 18.1% of the cumulative total variance and no distinct clustering was observed between sampling sites or streams (Fig. 3.2).

Zelandobius confusus (Plecoptera)
COI: Sequences of 658 bp were obtained for 101 (92%) of the 110 processed specimens from all 11 sites (Table 3.2). Analysis of *Z. confusus* revealed 10 haplotypes and haplotype diversity was moderate to high (0.377—1.000), while nucleotide diversity was relatively low (0.002—0.010; Table 3.2). Six of the 10 haplotypes were singletons and four haplotypes were found in five or more individuals (Fig. 3.2; Table 3.4). The central and putatively ancestral haplotype in the network (H1, n=29) occurred in all three streams (but not all sampling sites), while the most frequent haplotype (H2, n=49) occurred at all sampling sites across the study area (Fig 2b; Table 4). AMOVA revealed low genetic differentiation among all sampling sites (global Φ_{ST}=0.038, P>0.05; Table 3.3), and most of the pairwise comparisons were not significant (Table 3.5).

SNPs: The RADseq library produced a total of 40,549,912 reads from 73 specimens collected at 11 sampling sites within Pirongia (Table 3.5). Filtering resulted in 932

SNPs from all 73 individuals. Graphics showing the sequencing depth per sample and the location of missing data in the SNP matrix (65%) are available in Appendix Fig. A.3.3. Global Φ_{ST} analysis revealed a lack of population differentiation (Φ_{ST}=0.003; Table 3.3). Pairwise comparisons of Φ_{ST} identified just two population pairs with significant (but weak) genetic differentiation (B1 versus C2 Φ_{ST}=0.041, P <0.05; C1 versus C2 Φ_{ST}=0.042, P<0.05), suggesting high gene flow within Pirongia, as per the COI data. The fastSTRUCTURE analysis indicated that values of K between 2—3 maximized the marginal likelihood. For K=2, genetic clusters spanned all sampling sites, with admixture shown in some individuals. For K=3, the third group was shared as a small proportion between a few individuals from sites B1, B2 and C4 (Fig. 3.3). Two clusters were observed in the PCA analysis, with the first two components explaining 38.5% of cumulative variance and showing a lack of strong structure between sampling sites or streams (Fig. 3.2). These results were consistent with the lack of geographic structure observed in the mtDNA haplotype network.

Hydropsyche fimbriata (Trichoptera)

COI: Sequences of 552 bp were obtained for 107 (43%) of the 249 processed specimens from the 11 sites (Table 3.2). Analysis of *H. fimbriata* revealed 18 haplotypes, high haplotype diversity (0.400—1.000), and moderate nucleotide diversity (0.004—0.010). Both of these diversity measures were highest for *H. fimbriata* compared to the other two species. Eight of the 18 haplotypes identified were singletons, with six haplotypes found in more than three individuals (Fig. 3.2; Table 3.4). Among the most frequent haplotypes, H3 (n=20) and H16 (n=18) were distributed across sampling sites from all three streams, whereas H8 (n=15) and H12 (n=23) were only found in populations from the adjacent Streams A and B. AMOVA analysis showed significant but weak genetic differentiation between all sampling sites (Φ_{ST}=0.092, P<0.01; Table 3.3). Pairwise comparisons of Φ_{ST} between 19 population pairs were significantly different from zero, including the most distant sampling sites within each of two streams: A1—A3 and B1—B4 (Table 3.5).

SNPs: The RADseq library produced a total of 22,946,236 reads from 93 specimens collected at ten sampling sites (Table 3.5). After filtering, 202 SNPs from 92

specimens were retained. Graphics showing the sequencing depth per sample and the location of missing data in the SNP matrix (46.8%) are available in Appendix Fig A.3.4. AMOVA revealed low but significant population differentiation (global Φ_{ST}=0.096, P<0.01; Table 3.3). Pairwise comparisons of Φ_{ST} values were significantly high for ten pairs of populations, and significant but low for 13 population pairs (Table 3.5), suggesting that gene flow might be more limited in this species. Consistent with the COI data, pairwise comparisons indicated significant genetic differentiation between the most distant within-stream pairs (A1 versus A3, B1 versus B4). Significant, low genetic differentiation was also observed between the most distant sites within stream C (C1 versus C4; Φ_{ST}=0.079, P<0.01). fastSTRUCTURE analysis indicated that a value of K=5 maximised the marginal likelihood. The prevalence of each genetic group differed between the three streams and low levels of admixture were observed within individuals across all sampling sites (Fig. 3.3). The first two PCA components explained 54.7% of the total cumulative variance and, unlike the PCAs for the other two species, clear clusters could be distinguished (Fig. 3.2). In particular, individuals from Stream C sampling sites tended to group together more in one cluster than the others, but the four clusters showed a lack of clear geographic partitioning overall.
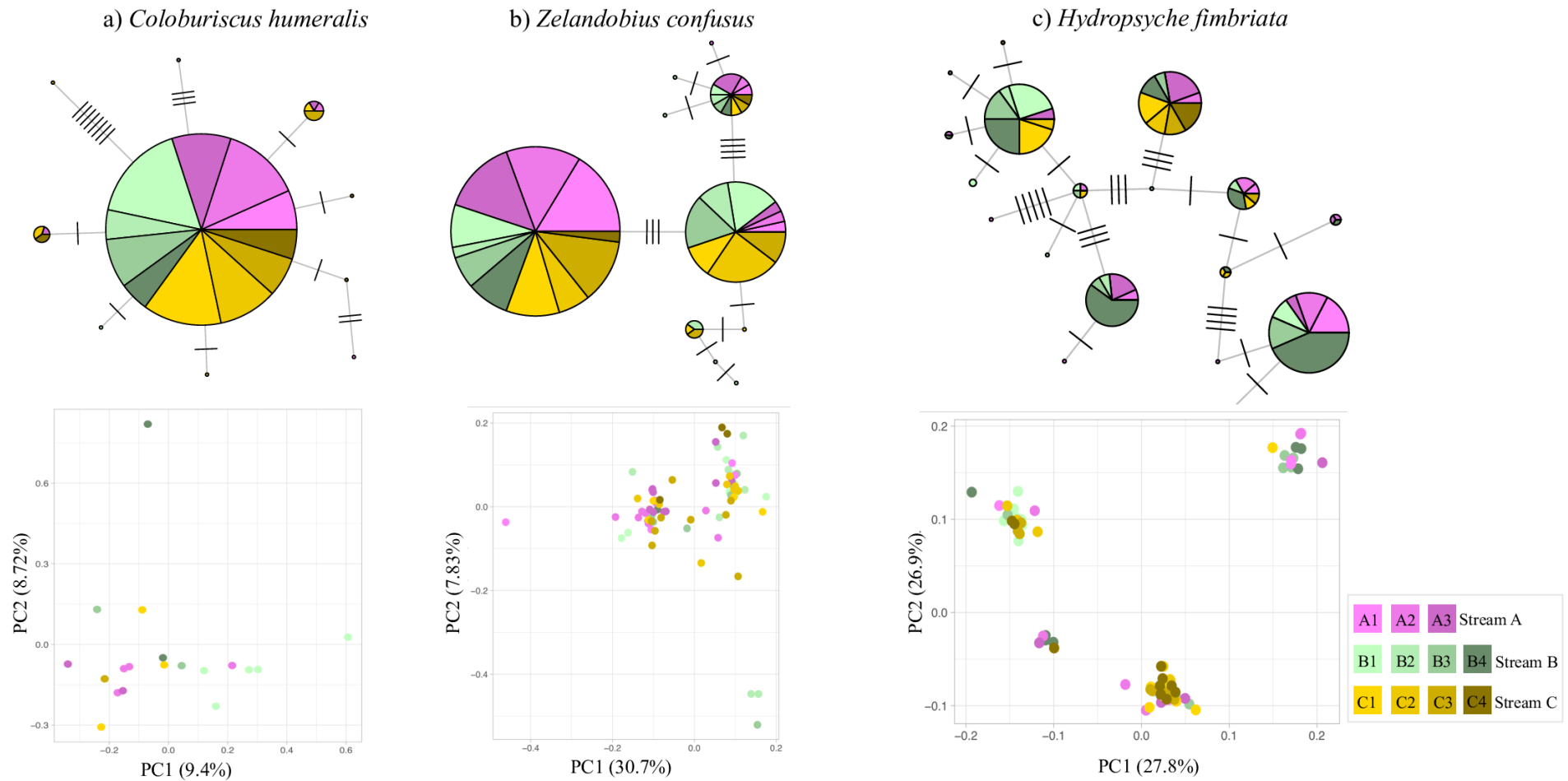
Figure 3.2 Visual representation of genetic relationships among individuals and sampling locations for (a) *Coloburiscus humeralis*, (b) *Zelandobius confusus,* and (c) *Hydropsyche fimbriata*. Individuals are colour-coded according to collection locality and population label. Top panel shows haplotype networks based on COI sequence data — each pie chart represents a single haplotype, with the size proportional to the frequency of individuals containing that particular haplotype at each site where samples were collected; dashes indicate missing mutational steps between haplotypes. Lower panel shows a Principal Component Analysis (PCA) based on orthogonal transformation of SNP data — the percentage of variation explained by each principal component is indicated on the axes.
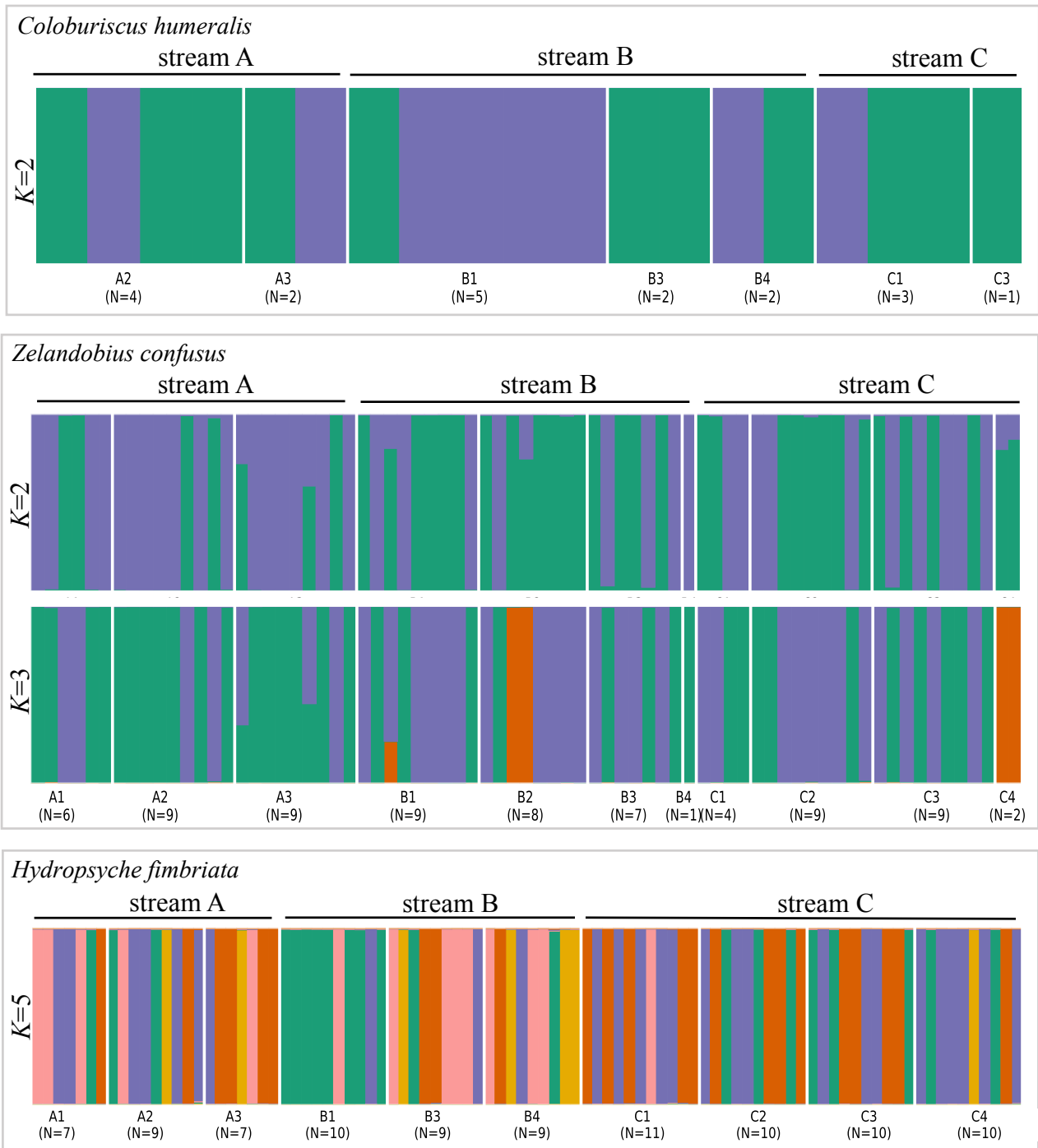
Figure 3.3 Visualisation of population structure and admixture of the study species from fastSTRUCTURE software. Results are presented for *K-value*s of 2 for *C. humeralis*, 2 and 3 for *Z. confusus*, and 5 for *H. fimbriata*, with colours in each panel representing the assignment of individuals to each genetic cluster. Single bars with >1 colour indicate admixture of that particular individual (i.e. sharing of genetic structure across more than one genetic group). Each sampling site is designated by its label (Stream A1, A2, etc).

Table 3.1 Population locations and site codes, downstream distance within each stream (from the top to the bottom sampling site), and Euclidean distance between streams for all sites analysed in this study.

Stream	Population code	Downstream distance	Euclidean distance	Riparian land cover	Coordinates
Tawhitiwhiti Stream	Stream A		-		
	A1	0 m		forest	37° 57' 33.0" S, 175° 05' 44.8" E
	A2	700 m		pasture	37°57' 15.5" S, 175° 05' 43.9" E
	A3	1100 m		pasture	37°57' 06.5" S, 175° 05' 43.9" E
Te Pahu Stream	Stream B		~600 m from stream A		
	B1	0 m		forest	37° 57' 48.2" S, 175° 06' 02.1" E
	B2	490 m		forest fragment	37° 57' 35.0" S, 175°06' 07.5" E
	B3	620 m		pasture	37° 57' 31.9" S, 175° 06' 10.4" E
	B4	1140 m		replanted	37° 57' 16.3" S, 175° 06' 08.1" E
Ngakoaohia Stream	Stream C		~11 km from stream A, B		
	C1	0 m		forest	38° 03' 10.8" S, 175° 05' 02.9" E
	C2	550 m		forest	38° 03' 20.2" S, 175° 05' 16.9" E
	C3	1500 m		forest	38° 03' 35.0" S, 175° 05' 34.9" E
	C4	3500 m		forest	38°03' 47.1" S, 175° 05' 46.7" E

Table 3.2 Sample sizes and genetic diversity of COI sequences for each of three study species *Coloburiscus humeralis*, *Zelandobius confusus* and *Hydropsyche fimbriata*. *n*, sample size; *x*, number of haplotypes; *h*, haplotype diversity; π nucleotide diversity.

Population code	*Coloburiscus humeralis* n(x) h	π	*Zelandobius confusus* n(x) h	π	*Hydropsyche fimbriata* n(x) h	π
A1	4(1) -	-	10(3) 0.377 ± 0.181	0.002	7(4) 0.714 ± 0.180	0.007
A2	10(3) 0.377 ± 0.181	0.001	10(4) 0.533 ± 0.180	0.004	10(7) 0.911 ± 0.077	0.010
A3	8(4) 0.642 ± 0.184	0.001	11(3) 0.563 ± 0.134	0.004	12(8) 0.893 ± 0.077	0.010
Total stream A	22(5) 0.407 ± 0.128	0.001	31(4) 0.475 ± 0.096	0.004	29(13) 0.889 ± 0.038	0.010
B1	11(2) 0.181 ± 0.143	0.001	10(3) 0.644 ± 0.101	0.003	11(5) 0.781 ± 0.107	0.005
B2	3(1) -	-	10(7) 0.911 ± 0.077	0.006	3(3) 1.000 ± 0.272	0.009
B3	5(1) -	-	10(4) 0.711 ± 0.117	0.004	10(6) 0.866 ± 0.085	0.009
B4	4(2) 0.500 ± 0.265	0.001	5(2) 0.400 ± 0.237	0.004	33(9) 0.820 ± 0.107	0.009
Total stream B	23(3) 0.170 ± 0.103	0.001	35(9) 0.751 ± 0.049	0.004	57(13) 0.835 ± 0.026	0.008
C1	9(2) 0.222 ± 0.166	0.001	11(5) 0.763 ± 0.106	0.004	9(4) 0.750 ± 0.112	0.006
C2	8(2) 0.428 ± 0.168	0.001	10(2) 0.466 ± 0.131	0.002	5(4) 0.900 ± 0.161	0.006
C3	9(4) 0.750 ± 0.878	0.004	12(4) 0.712 ± 0.105	0.004	3(2) 0.666 ± 0.314	0.004
C4	7(4) 0.809 ± 0.129	0.002	2(2) 1.000 ± 0.500	0.010	4(2) 0.500 ± 0.265	0.005
Total stream C	33(7) 0.580 ± 0.093	0.002	35(5) 0.682 ± 0.048	0.003	21(6) 0.728 ± 0.079	0.005

- indicates genetic diversity is not available (for populations that contain only one haplotype).

Table 3.3 Hierarchical analysis of molecular variance (AMOVA) for COI sequences for three species *Coloburiscus humeralis*, *Zelandobius confusus* and *Hydropsyche fimbriata*.

	COI data		SNP data	
	Variance %	Φ_{ST}	Variance	Φ_{ST}
Coloburiscus humeralis				
Among all sites	1.01	0.010	1.21	0.017
Within all sites	98.99		98.79	
Zelandobius confusus				
Among all sites	3.82	0.038	1.12	0.003
Within all sites	96.18		98.88	
Hydropsyche fimbriata				
Among all sites	9.23	0.092*	9.61	0.096*
Within all sites	90.77		90.38	

*P-value<0.01

Table 3.4 Haplotype frequencies based on COI data for each of three study species *Coloburiscus humeralis, Zelandobius confusus* and *Hydropsyche fimbriata.*

Coloburiscus humeralis

	Population											
	A1	A2	A3	B1	B2	B3	B4	C1	C2	C3	C4	*n*
H1	4	8	6	10	3	5	3	8	6	4	3	60
H2		1										1
H3			1						2		2	5
H4				1								1
H5										1		1
H6											1	1
H7											1	1
H8		1	1					1		3		6
H9						1						1
H10										1		1

Zelandobius confusus

	Population											
	A1	A2	A3	B1	B2	B3	B4	C1	C2	C3	C4	*n*
H1	1	1	1	5	3	5		3	7	3		29
H2	8	7	7	4	1	3	4	5	3	6	1	49
H3					1							1
H4						1						1
H5					2			1		2		5
H6								1				1
H7	1	1	3	1	1	1	1	1		1	1	12
H8					1							1
H9					1							1
H10		1										1

Hydropsyche fimbriata

	Population											
	A1	A2	A3	B1	B2	B3	B4	C1	C2	C3	C4	*n*
H1				1								1
H2		1		1		1			1			4
H3			1	5	1	3	5	4	1			20
H4											1	1
H5			1				1					2
H6							1					1
H7				2								2
H8		1	3		1	1	9					15
H9		1										1
H10	1											1
H11	1											1
H12	4	3	1	2		3	10					23
H13			1									1
H14	1	2			1		3	1		1		9
H15						1						1
H16		1	4		1		2	3	2	2	3	18
H17		1	1				1					3
H18							1	1	1			3

Grey shaded haplotypes are private. *n*, number of individuals sharing each haplotype

Table 3.5 Pairwise genetic distances (Φ_{ST}) between sampling sites for each of the three study species. Results for SNP and COI data presented in the upper and lower diagonals, respectively. SNP data sample size for each site is provided in parentheses.

Coloburiscus humeralis

	A1(0)	A2(4)	A3(2)	B1(5)	B2(0)	B3(2)	B4(2)	C1(3)	C2(0)	C3(1)	C4(0)
A1	-	**0.052***	NA	NA	NA	NA	NA	NA	NA	NA	NA
A2	0.000	-	NA	NA	NA	NA	NA	0.006	NA	NA	NA
A3	0.000	0.000	-	NA	NA	NA	NA	NA	NA	NA	NA
B1	0.000	0.006	0.018	-	NA	NA	NA	0.000	NA	NA	NA
B2	0.000	0.000	0.000	0.000	-	NA	NA	NA	NA	NA	NA
B3	0.000	0.000	0.000	0.000	0.000	-	NA	NA	NA	NA	NA
B4	0.000	0.007	0.079	0.189	0.000	0.063	-	NA	NA	NA	NA
C1	0.116	0.000	0.000	0.002	0.000	0.000	0.145	-	NA	NA	NA
C2	0.020	0.046	0.000	0.124	0.000	0.063	0.127	0.105	-	NA	NA
C3	0.000	0.003	0.000	**0.082***	0.000	0.000	0.000	0.007	0.063	-	NA
C4	0.000	0.027	0.000	**0.124***	0.000	0.023	0.059	**0.097***	0.000	0.054	-

Zelandobius confusus

	A1(6)	A2(9)	A3(9)	B1(9)	B2(8)	B3(7)	B4(1)	C1(4)	C2(9)	C3(9)	C4(2)
A1	-	0.000	0.000	0.025	0.019	0.000	NA	0.017	0.007	0.007	NA
A2	0.000	-	0.000	0.024	0.016	0.000	NA	0.025	0.014	0.000	NA
A3	0.000	0.000	-	0.000	0.021	0.000	NA	0.000	0.000	0.000	NA
B1	0.111	0.018	0.005	-	0.000	0.000	NA	0.000	**0.041***	0.000	NA
B2	**0.319*****	0.195	**0.145***	0.074	-	NA	NA	NA	NA	NA	NA
B3	0.167	0.064	0.044	0.000	0.014	-	NA	0.000	0.004	0.000	NA
B4	0.000	0.000	0.000	0.053	0.216	0.104	-	NA	NA	NA	NA
C1	0.067	0.000	0.000	0.000	0.066	0.000	0.011	-	**0.042***	0.000	NA
C2	**0.277***	0.162	0.146	0.000	0.121	0.000	**0.258***	0.000	-	0.011	NA
C3	0.041	0.000	0.000	0.000	0.090	0.000	0.000	0.000	0.213	-	NA
C4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	-

Hydropsyche fimbriata

	A1(7)	A2(9)	A3(7)	B1(10)	B2(0)	B3(9)	B4(9)	C1(11)	C2(10)	C3(10)	C4(10)
A1	-	**0.080****	**0.327*****	**0.248*****	NA	0.075	**0.486*****	**0.440*****	**0.273*****	**0.320*****	**0.252*****
A2	0.106	-	0.016	0.000	NA	0.000	**0.143*****	**0.106*****	**0.087*****	**0.085****	0.000
A3	**0.276*****	0.000	-	0.000	NA	0.075	0.000	0.015	**0.054***	0.000	0.050
B1	**0.391****	**0.168***	0.127	-	NA	0.000	**0.086****	**0.077*****	**0.092*****	**0.089****	0.015
B2	**0.341***	0.000	0.000	0.070	-	NA	NA	NA	NA	NA	NA
B3	**0.168***	0.000	0.000	0.005	0.000	-	**0.295*****	**0.281*****	**0.220****	**0.164****	0.000
B4	**0.166***	0.000	0.008	**0.081***	0.000	0.000	-	0.000	0.020	0.000	**0.105***
C1	**0.432*****	0.099	0.000	**0.156***	0.000	0.036	**0.108***	-	0.013	0.000	**0.079****
C2	**0.394****	0.030	0.000	**0.182***	0.000	0.000	0.068	0.000	-	0.000	0.042
C3	**0.467***	0.118	0.033	**0.494*****	0.182	0.195	**0.219*****	0.063	0.000	-	0.047
C4	**0.509****	**0.200***	0.030	**0.411*****	0.195	0.182	**0.232*****	0.023	0.000	0.000	-

Significant values are shown in bold (*$P<0.05$; **$P<0.01$; ***$P<0.001$). NA indicates values that were not calculated (for populations with ≤2 individuals).

3.5 Discussion

Using both COI and SNP markers, we identified a lack of strong differentiation among populations for the three aquatic insect species analysed, suggesting high population connectivity (gene flow) between sampling sites and stream localities within Pirongia at a spatial scale of ~11 km. However, the genetic markers detected minor differences in the degree and pattern of population genetic differentiation observed for each of the study species that may reflect their different dispersal abilities.

For the common mayfly *C. humeralis*, COI analysis showed a homogeneous distribution of haplotypes across all populations within and between the three study streams in Pirongia. Results of the SNP analysis were consistent with the COI data, with no population structure revealed by the Bayesian analysis (FastStructure) or PCA. This suggests high population connectivity at small spatial scales across a variety of land use types, consistent with a previous allozyme study that showed very low levels of differentiation between *C. humeralis* populations at different spatial scales in the North and South Islands of New Zealand (Hogg et al., 2002). The historic distribution of *C. humeralis* across both islands of New Zealand suggests widespread dispersal is possible across distant catchments, although ongoing habitat fragmentation could restrict future connectivity among populations. Analysis across distant catchments is necessary to confirm the inter-regional dispersal suggested by the previous allozyme data.

For the stonefly *Z. confusus*, both COI and SNP data analysed showed a general pattern of low genetic differentiation between sampling sites, indicating substantial population connectivity across the study area. These results are consistent with SNP data for another New Zealand (medium-sized) stonefly species *Zelandoperla decorata*, which shows low genetic differentiation within and between parallel streams separated by ~10 km, indicating overland dispersal at small spatial scales over fragmented or non-forested landscapes (Dussex et al., 2016). Our results also support previous direct studies of dispersal using stable isotope enrichment in stoneflies, which indicate dispersal between streams (Briers et al., 2004), although these taxa are generally considered weak fliers (Sproul et al., 2014; Briers et al., 2002).

For the caddisfly *H. fimbriata*, common COI haplotypes were not found at all sites and individual membership in genetic clusters differed among locations for the SNP analysis. Significant pairwise Φ_{ST} for both markers — even from the upper and lower sampling sites within each of the three Pirongia streams (as observed within Streams A and B for the COI data, and Streams A, B and C for the SNP data) — indicated limited connectivity within and between adjacent streams and suggests that dispersal in this species may be limited even within the same stream channel. Collectively, these data suggest more limited gene flow among populations of *H. fimbriata* at small spatial scales relative to that found for either *C. humeralis* or *Z. confusus*. Hydropsychid larvae, such as *H. fimbriata*, are sedentary filter-feeders that build and attach a "fixed retreat" shelter of silk and organic matter to a stable substrate. The larval investment in shelter construction may reduce the propensity for drift, and this may partially explain the high genetic differentiation between populations. Settlement onto substrata may also be related to limited larval downstream drift, as suggested by Downes & Lancaster (2010) for another Hydropsychid caddisfly.

Previous landscape genetics studies on different riverine species have shown that human-driven fragmentation for agricultural purposes can affect population structure and may restrict dispersal in these ecosystems (Blanchet et al., 2010; Lean et al., 2017; Wilcock et al., 2007). Accordingly, local landscape features which constrain dispersal could further explain the observed population genetic structure for this species. *Hydropsyche fimbriata* larvae are mostly restricted to cool streams in native forests and fragmented stream corridors provide less optimal microclimate conditions that may limit dispersal of adults outside of forest areas.

We found the lowest Φ_{ST} values between sites within Stream C (Te Pahu) and individuals from this stream, which is covered by forest riparian vegetation, tended to form a single genetic group in the PCA. In contrast, greater genetic differentiation was found within the other more fragmented streams, where riparian vegetation is mainly dominated by low growing pasture grasses and weed species. Further landscape genetics analyses are needed to determine the relevance of riparian land cover in shaping the spatial genetic structure in *H. fimbriata* populations (see Chapter 4).

Overall, our combined COI and SNP data suggests limited population structure and non-limited dispersal at small spatial scales within and between neighbouring catchments. The small differences we observed in patterns of genetic differentiation among the three species— suggesting higher population connectivity for *C. humeralis* and *Z. confusus* compared to *H. fimbriata*—highlight the potential influence of their dispersal abilities and/or landscape features, such as the riparian land cover (forested versus open pasture) in shaping connectivity among populations. However, we caution that smaller sample sizes and lower data quality for the SNP datasets may have limited their resolution in determining population genetic structure within our study area. We note that genetic differentiation within the stream channel was slightly more pronounced in the SNP than the COI dataset. However, the generally consistent results between markers suggests that either COI or SNP data can independently offer suitable estimates of population differentiation at small spatial scales. Used in combination, they could strengthen inferences of population structure for estimates of contemporary gene flow and elucidation of fine-scale dispersal patterns that might be otherwise confounded by the resolution power and specific characteristics of a given single marker. Extending this approach to a landscape genetics context (see Chapter 4) with increasing geographic coverage will further enhance our knowledge of genetic structure and landscape influences on aquatic insect dispersal and population connectivity.

3.6 References

Alexander, L. C., Hawthorne, D. J., Palmer, M. A., & Lamp, W. O. (2011). Loss of genetic diversity in the North American mayfly *Ephemerella invaria* associated with deforestation of headwater streams. *Freshwater Biology,* **56**, 1456-1467.

Alp, M., Keller, I., Westram, A. M., & Robinson, C. T. (2012). How river structure and biological traits influence gene flow: a population genetic study of two stream invertebrates with differing dispersal abilities. *Freshwater Biology,* **57**, 969-981.

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data [Online]. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., …, Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One,* **3**, e3376.

Bohonak, A. J. 1999. Effect of insect-mediated dispersal on the genetic structure of postglacial water mite populations. *Heredity,* **82**, 451-461.

Briers, R. A., Gee, J. H. R., Cariss, H. M., & Geoghegan, R. 2004. Inter-population dispersal by adult stoneflies detected by stable isotope enrichment. *Freshwater Biology,* **49**, 425-431.

Briers, R. A., Cariss, H. M., Gee, J. H. R. 2002. Dispersal of adult stoneflies (Plecoptera) from upland streams draining catchments with contrasting land-use. *Archiv für Hydrobiologie*, **4**, 627-644.

Brittain, J. 2003. Biology of mayflies. *Annual Review of Entomology,* **27**, 119-147.

Collier, K.J. & Smith, B.J., 2000. Interactions of adult stoneflies (Plecoptera) with riparian zones I. Effects of air temperature and humidity on longevity. *Aquatic Insects*, **22**, 275-284.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., …, Genomes Project Analysis, G. (2011). The variant call format and VCFtools. *Bioinformatics,* **27**, 2156-2158.

Davey, J. W., & Blaxter, M. L. (2010). RADSeq: next-generation population genetics. *Briefings in Functional Genomics,* **9**, 416-423.

De Cario, R., Kura, A., Suraci, S., Magi, A., Volta, A., Marcucci, R., …, Sticchi, E. (2020). Sanger validation of high-throughput sequencing in genetic diagnosis: Still the best practice? *Frontiers in genetics,* **11**, 592588-592588.

Devlin-Durante, M. K., & Baums, I. B. (2017). Genome-wide survey of single-nucleotide polymorphisms reveals fine-scale population structure and signs of selection in the threatened Caribbean elkhorn coral, *Acropora palmata. PeerJ,* **5**, e4077-e4077.

Downes, B.J. & Lancaster, J., (2010). Does dispersal control population densities in advection-dominated systems? A fresh look at critical assumptions and a direct test. *Journal of Animal Ecology*, **79**, 235-248.

Driscoll, D. A., Banks, S. C., Barton, P. S., Ikin, K., Lentini, P., Lindenmayer, D. B., …, Westgate, M. J. (2014). The trajectory of dispersal research in conservation biology. Systematic Review. *PLoS One,* **9**, e95053.

Dussex, N., Chuah, A., & Waters, J. M. (2016). Genome-wide SNPs reveal fine-scale differentiation among wingless alpine stonefly populations and introgression between winged and wingless forms. *Evolution,* **70**, 38-47.

Eaton, D., & Overcast, I. (2020). Ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics,* **36**, 2592-2594.

Ekblom, R. & Galindo, J. 2011. Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity,* **107**, 1-15.

Elbrecht, V., Feld, C. K., Gies, M., Hering, D., Sondermann, M., Tollrian, R., & Leese, F. (2014). Genetic diversity and dispersal potential of the stonefly

Dinocras cephalotes in a central European low mountain range. *Freshwater Science,* **33**, 181-191.

Engelhardt, C. H. M., Haase, P., & Pauls, S. U. (2011). From the Western Alps across Central Europe: Postglacial recolonisation of the tufa stream specialist *Rhyacophila pubescens* (Insecta, Trichoptera). *Frontiers in Zoology,* **8**.

Emerson, K. J., Merz, C. R., Catchen, J. M., Hohenlohe, P. A., Cresko, W. A., Bradshaw, W. E. & Holzapfel, C. M. 2010. Resolving postglacial phylogeography using high-throughput sequencing. *Proceedings of the National Academy of Sciences,* **107**, 16196.

Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics,* **131**, 479-491.

Geraci, C.J., Zhou, X., Morse, J.C. & Kjer, K.M., 2010. Defining the genus Hydropsyche (Trichoptera: Hydropsychidae) based on DNA and morphological evidence. *Journal of the North American Benthological Society*, **29**, 918-933.

Goudet, J. (2005). hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes,* **5**, 184-186.

Harding, J., & Winterbourn, M. (1993). Life history and production of *Coloburiscus humeralis* (Ephemeroptera: Oligoneuriidae) in two South Island high-country streams, New Zealand. *New Zealand Journal of Marine and Freshwater Research,* **27**, 445-451.

Head, S. R., Komori, H. K., LaMere, S. A., Whisenant, T., Van Nieuwerburgh, F., Salomon, D. R. & Ordoukhanian, P. 2014. Library construction for next-generation sequencing: overviews and challenges. *Biotechniques,* **56**, 61-passim.

Heather, J. M., & Chain, B. (2016). The sequence of sequencers: The history of sequencing DNA. *Genomics, * **107**, 1-8.

Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H., & Hallwachs, W. (2004). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator. Proceedings of the National Academy of Sciences of the United States of America,* **101**, 14812-14817.

Hogg, I. D., Willmann-Huerner, P., & Stevens, M. I. (2002). Population genetic structures of two New Zealand stream insects: *Archichauliodes diversus* (Megaloptera) and *Coloburiscus humeralis* (Ephemeroptera). *New Zealand Journal of Marine and Freshwater Research,* **36**, 491-501.

Hotaling, S., Kelley, J. L. & Frandsen, P. B. 2020. Aquatic insects are dramatically underrepresented in genomic research. *Insects,* **11**, 601.

Hughes, J. M. (2007). Constraints on recovery: using molecular methods to study connectivity of aquatic biota in rivers and streams. *Freshwater Biology,* **52**, 616-631.

Hughes, J. M., Schmidt, D. J., & Finn, D. S. (2009). Genes in streams: using DNA to understand the movement of freshwater fauna and their riverine habitat. *BioScience,* **59**, 10.

Ivanova, N. V., Dewaard, J. R., & Hebert, P. D. N. (2006). An inexpensive, automation-friendly protocol for recovering high-quality DNA. *Molecular Ecology Notes,* **6**, 998-1002.

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ,* **2**, e281-e281.

Kear, D. (Cartographer). (1960). Sheet 4-Hamilton. Geological map of New Zealand 1:250000.

Knapp, I., Puritz, J., Bird, C., Whitney, J., Sudek, M., Forsman, Z., & Toonen, R. (2016). ezRAD- an accessible next-generation RAD sequencing protocol suitable for non-model organisms_v3.2. *protocols.io*.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., …, Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics,* **23**, 2947-2948.

Mardis, E. (2017). DNA sequencing technologies: 2006-2016. *Nature Protocols,* **12**, 213-218.

McGaughran, A., Terauds, A., Convey, P., & Fraser, C. I. (2019). Genome-wide SNP data reveal improved evidence for Antarctic glacial refugia and dispersal of terrestrial invertebrates. *Molecular Ecology,* **28**, 4941-4957.

McLeod, E. O., Pittari, A., Brenna, M., & Briggs, R. (2020). *Geology of the Pirongia Volcano, Waikato: 1:30,000 Geological Map*: Geoscience Society of New Zealand.

Monaghan, M. T., & Sartori, M. (2009). Genetic contributions to the study of taxonomy, ecology, and evolution of mayflies (Ephemeroptera): review and future perspectives. *Aquatic Insects,* **31**, 19-39.

Paradis, E., & Schliep, K. (2018). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics,* **35**.

Pazmiño, D. A., Maes, G. E., Simpfendorfer, C. A., Salinas-de-León, P., & van Herwerden, L. (2017). Genome-wide SNPs reveal low effective population size within confined management units of the highly vagile Galapagos shark (*Carcharhinus galapagensis*). *Conservation Genetics,* **18**, 1151-1163.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., …, Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics,* **81**, 559-575.

Raj, A., Stephens, M., & Pritchard, J. K. (2014). fastSTRUCTURE: Variational inference of population structure in large SNP data sets. *Genetics,* **197**, 573.

Ratnasingham, S., & Hebert, P. D. N. (2007). BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). *Molecular Ecology Notes,* **7**, 355-364.

Saastamoinen, M., Bocedi, G., Cote, J., Legrand, D., Guillaume, F., Wheat, C. W., …, Del Mar Delgado, M. (2018). Genetics of dispersal. *Biological reviews of the Cambridge Philosophical Society,* **93**, 574-599.

Salokannel, J., Min Lee, K., Rinne, A., Mutanen, M. 2021. ddRAD sequencing sheds light on low interspecific and high intraspecific mtDNA Divergences in two groups of caddisflies. *Insect Systematics and Diversity*, **5**.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America,* **74**, 5463-5467.

Sd, S., Roessli, D., & Excoffier, L. (2000). ARLEQUIN: A software for population genetics data analysis. *Genetics and Biometry Lab, Dept. of Anthropology, Univ. of Geneva,* **1**.

Slatkin, M. (1985). Gene flow in natural populations. *Annual Review of Ecology and Systematics,* **16**, 393-430.

Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., …, Hood, L. E. (1986). Fluorescence detection in automated DNA sequence analysis. *Nature,* **321**, 674-679.

Smith, P. J., & Collier, K. J. (2001). Allozyme diversity and population genetic structure of the caddisfly *Orthopsyche fimbriata* and the mayfly *Acanthophlebia cruentata* in New Zealand streams. *Freshwater Biology,* **46**, 795-805.

Smith, P. J., McVeagh, S. M., & Collier, K. J. (2006). Population-genetic structure in the New Zealand caddisfly *Orthopsyche fimbriata* revealed with mitochondrial DNA. *New Zealand Journal of Marine and Freshwater Research,* **40**, 141-148.

Smith, P. J., & Smith, B. J. (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**, 723-734.

Sproul, J. S., Houston, D. D., Davis, N., Barrington, E., Oh, S. Y., Evans, R. P., & Shiozawa, D. K. 2014. Comparative phylogeography of codistributed aquatic insects in western North America: insights into dispersal and regional patterns of genetic structure. *Freshwater Biology,* **59**, 2051-2063.

Tan, M. P., Wong, L. L., Razali, S. A., Afiqah-Aleng, N., Mohd Nor, S. A., Sung, Y. Y., Van de Peer, Y., Sorgeloos, P. & Danish-Daniel, M. 2019. Applications of next-generation sequencing technologies and computational tools in molecular evolution and aquatic animals conservation studies: a short review. *Evolutionary Bioinformatics Online,* **15**, 1176934319892284-1176934319892284.

Trense, D, Schmidt, TL, Yang, Q, Chung, J, Hoffmann, AA, Fischer, K. 2021. Anthropogenic and natural barriers affect genetic connectivity in an Alpine butterfly. *Molecular Ecol*ogy, **30**: 114130.

Venables, S. K., Marshall, A. D., Armstrong, A. J., Tomkins, J. L., & Kennington, W. J. (2021). Genome-wide SNPs detect no evidence of genetic population structure for reef manta rays (*Mobula alfredi*) in southern Mozambique. *Heredity,* **126**, 308-319.

Vendrami, D. L. J., Telesca, L., Weigand, H., Weiss, M., Fawcett, K., Lehman, K., …, Hoffman, J. I. (2017). RAD sequencing resolves fine-scale population structure in a benthic invertebrate: implications for understanding phenotypic plasticity. *Royal Society Open Science,* **4**, 160548.

Wallace, F., Matthew. (2013). *Using population genetics to assess the dispersal patterns of the New Zealand mayfly Coloburiscus humeralis in a landscape context.* (Master of Science Dissertation), University of Canterbury.

Wang, P., Chen, B., Zheng, J., Cheng, W., Zhang, H., Wang, J., …, Mao, Y. (2020). Fine-scale population genetic structure and parapatric cryptic species of kuruma shrimp (*Marsupenaeus japonicus*), along the northwestern pacific coast of China. *Frontiers in genetics,* **11**, 118.

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., & Barton, G. J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics,* **25**, 1189-1191.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution,* **38**, 1358-1370.

Weston, M., Clarke, K., Maguire, G., & Sumner, J. (2020). Morphological and molecular evidence of population divergence in a widespread shorebird across its southern mainland Australian distribution. *Conservation Genetics,* **21**.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., …, Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software,* **4**.

Wilcock, H. R., Bruford, M. W., Nichols, R. A., & Hildrew, A. G. (2007). Landscape, habitat characteristics and the genetic population structure of two caddisflies. *Freshwater Biology,* **52**, 1907-1929.

Wilson, J. J. (2012). DNA barcodes for insects. *Methods in molecular biology (Clifton, N.J.),* **858**, 17-46.

Winterbourn, M., Gregson, K. L. D., & Dolphin, C. H. (2000). Guide to the aquatic insects of New Zealand. *Bulletin of the Entomological Society of New Zealand,* **13**.

Winterbourn, M. J. (2005). Dispersal, feeding and parasitism of adult stoneflies (Plecoptera) at a New Zealand forest stream. *Aquatic Insects,* **27**, 155-166.

Zhang, J., Chiodini, R., Badr, A. & Zhang, G. (2011). The impact of next-generation sequencing on genomics. *Journal of Genetics and Genomics,* **38**, 95-109.

Chapter Four

Fine-scale landscape genetics analyses for stream insects reveals species-specific patterns of population structure across a fragmented landscape

Contrast of the riparian zone cover between Tawhiwhiti Stream (top) and Ngakoaohia Stream, Mount Pirongia. Photos: Vanessa Barbosa

This chapter is in preparation as a research article to be submitted in Molecular Ecology.

4.1 Abstract

Dispersal ensures population connectivity across altered landscapes and understanding this process is essential for the conservation, restoration, and management of streams. We used mitochondrial DNA (mtDNA) and genome-wide single nucleotide polymorphism (SNP) markers to assess functional connectivity of stream insects in a fragmented landscape dominated by pasture land. We investigated species-specific patterns in three insects with a terrestrial winged adult stage: The mayfly *Coloburiscus humeralis*, the stonefly *Zelandobius confusus*, and the caddisfly *Hydropsyche fimbriata*. A landscape genetic approach was used to test the relative influence of pure spatial effects and landscape elements (i.e. topography and land cover) on genetic structure among sites separated by up to 11 km. Spatial genetic structure was also examined for populations separated by distinct mountain regions. We found clear spatial genetic structure with marked Isolation By Distance (IBD) at only the broadest spatial scales (mountain regions separated by ~30 and 170 km). For *C. humeralis,* population structure was marked by discordance between mtDNA and SNP data potentially related to recent isolation. At smaller spatial scales (up to 11 km), landscape influences generally best predicted genetic structuring: widespread gene flow among populations suggested highest dispersal potential across forested and pasture land for *Z. confusus*; for *C. humeralis,* fine-scale genetic differentiation in the SNP dataset and a weak positive correlation with land cover suggested higher population connectivity within the stream channel protected by forested riparian zone than within fragmented streams, and overland dispersal across pasture land; for *H. fimbriata*, overland dispersal may be reduced due to local habitat features, but this does not appear to prevent more widespread population connectivity. Collectively, our results highlight the importance of assessing landscape features when determining population connectivity in the stream riparian zone, and indicate that any restoration efforts should include estimates of connectivity.

4.2 Introduction

Habitat loss and fragmentation are major threats to biodiversity among natural populations (Fahrig, 2003; Lawler et al., 2013). As the development of natural environments expands, available habitats decrease, transforming the continuous landscape into isolated patches of varying size and connectivity. Anthropogenic habitat fragmentation is particularly prevalent—and increasing—in riverine ecosystems (Reid et al., 2019). Agricultural development, including the construction of dams, land-use change within river catchments, and removal of riparian vegetation, is recognised as having negative effects on many freshwater species and ecosystem health (Dala-Corte et al., 2016; Fuller et al., 2015). The loss of natural riparian vegetation, in particular, has profound consequences for stream function (Burrell et al., 2014; Hladyz et al., 2011), affecting the physical habitat conditions for aquatic insects. Forested riparian zones provide shade, cooler water temperatures, and reduced wind speeds, as well as cooler and more humid microclimates—all factors crucial for dispersal and persistence of populations across habitats (Carlson et al., 2016; Collier & Smith, 2000).

For most species, the ability to successfully disperse through fragmented and disturbed patches is a key determinant of the long-term viability of populations (Fuller et al., 2015). At the genetic level, altered and restricted movements of individuals in fragmented landscapes may disrupt the connectivity of populations across habitats, reducing gene flow and decreasing effective population size and genetic diversity of remnant populations through the processes of genetic drift and inbreeding (Pavlova et al., 2017; Schlaepfer et al., 2018). Species dispersal, therefore, plays a vital role in ensuring population connectivity and persistence across altered landscapes (Blanchet et al., 2010; Galic et al., 2013).

In stream insect populations, dispersal and colonisation typically follow the stream channel, making longitudinal connectivity essential (Petersen et al., 2004; Wiens, 2002). Aerial dispersal can be affected by many constraints, including topographic or anthropogenic physical barriers, weather, and land cover (Blakely et al., 2006; Parkyn & Smith, 2011; Phillipsen & Lytle, 2013). However, recent studies of population genetic structure have found evidence for lateral dispersal across landscapes between riverine systems (Geismar et al., 2015; Wilcock et al., 2007; Yaegashi et al., 2014). Therefore, lateral connectivity of intervening

landscape habitats and features is likely also important for dispersal and connectivity among populations (Alp et al., 2012; Hughes, 2007). Most restoration efforts rely on natural recolonisation of the restored habitat via dispersal of individuals from nearby areas (Blakely et al., 2006; Bond & Lake, 2003). However, for recolonisation to be successful, it is crucial that habitat patches facilitate movement and therefore increase the exchange of individuals and gene flow among populations (Christie & Knowles, 2015). Thus, understanding population connectivity and dispersal patterns in modified landscapes is vital for conservation and the success and management of restoration efforts.

Functional connectivity can be defined as "the degree to which the landscape facilitates or impedes movement along with resource patches" (Taylor et al., 1993), and is both species and landscape-specific (Tischendorf & Fahrig, 2000). In streams, the life history and dispersal traits of a particular species, combined with the dendritic structure of the stream network and the spatial location of individuals, together determine functional connectivity (Hughes et al., 2009). Accordingly, the pattern and scale of population genetic structure varies among taxa. Exchange of gene flow may be 'widespread' among populations for species with high dispersal capacity and no particular habitat requirements (leading to low genetic structure), or 'limited' for those with low dispersal potential and/or particular habitat requirements (creating high genetic structure) (Finn et al., 2007; Hughes et al., 2009).

Landscape genetic analysis can be used to assess functional connectivity in stream insects by investigating the correlation between population genetic differentiation and landscape and/or environmental features (Manel & Holderegger, 2013; Manel et al., 2003; Spear et al., 2010). When relationships between genetic and landscape parameters are significant, the analysed landscape resistance predicts the spatial genetic structure—a pattern referred to as Isolation By Resistance (IBR; McRae, 2006). Species-specific landscape features are taken into account with IBR analysis by testing whether portions of the landscape may affect gene flow (Shah & McRae, 2008). Therefore, IBR extends the Isolation by Distance model (IBD), where distance is the primary constraint on dispersal and the landscape is assumed to be homogeneous and have no effect on dispersal. Investigation of spatial genetic structure under the IBD model is commonly applied in aquatic insect dispersal research (see Chapter 2). However, more detailed analyses of functional

connectivity can lead to a better understanding of the influence of the landscape on populations when purely geographical delimitation is not clear, revealing natural and anthropogenic barriers to connectivity (Galic et al., 2013; Phillipsen & Lytle, 2013; Polato et al., 2017), with substantial implications for conservation and management (Keller et al., 2012).

We conducted a fine-scale landscape genetic analysis on three different endemic New Zealand stream insect species with terrestrial life-stages. We used mitochondrial DNA (mtDNA) and genome-wide single nucleotide polymorphism (SNP) markers to: 1) assess functional connectivity between stream insect populations by examining their spatial genetic structure; 2) test the effect of pure space i.e. Euclidean distance (IBD) versus the intervening landscape (IBR) on population connectivity among habitats; and 3) identify landscape elements (i.e. topography and land cover) that inhibit dispersal and gene flow. In addition, we included one to two outgroup populations per species from different regions to compare fine-scale spatial population structure within the main study area to that over a larger geographical scale. Our study was carried out in a fragmented agricultural landscape dominated by open pasture land that included native forest fragments, meaning gene flow between populations is vulnerable to ongoing local deforestation. We predicted that, for the three studied species, a pattern of IBD would be evident at larger spatial scales (among populations from different mountain regions). At finer spatial scales, we predicted that pasture land would act as a barrier to dispersal within and among streams. Our results provide insights into species-specific patterns of dispersal and connectivity within a fragmented landscape, and have broad implications for management and restoration planning.

4.3 Methods

4.3.1 Study area

This research was mainly conducted in the southeast of the Mount Pirongia area in the North Island of New Zealand, covering a spatial scale of ~11 km. Different mountain regions were also included in our broader spatial analysis: Wainui Stream in Mount Karioi (~30 km away from Mount Pirongia) and Katikara Stream and Patea stream in Mount Taranaki (~170 km away from Mount Pirongia) (Fig. 4.1). All three regions are characterised by volcanic fields. Mount Pirongia

and Mount Karioi are both parts of the Alexandra Volcanic Group, formed by a succession of eruptions beginning over 2.5 million years ago and last erupting approximately 1.6 million years ago (Kear, 1960; McLeod et al., 2020). Mount Taranaki is geologically younger, having commenced activity approximately 125,000 years ago, and has been dormant for the last 150 years (Turner et al., 2008). Intense deforestation in New Zealand has occurred since colonisation (~1000 years ago), and agricultural land use is one of the main drivers of forest loss (Brockerhoff et al., 2010). Within the North Island in particular, indigenous forest has decreased from 96% to 24% of the total land cover (Ewers et al., 2006). Today, the landscape structure in these regions is characterised by large and small forest fragments surrounded by intense agricultural production—predominantly dairy and sheep farmland. Thus, many resident streams and rivers flow through native forest before entering agricultural landscapes.
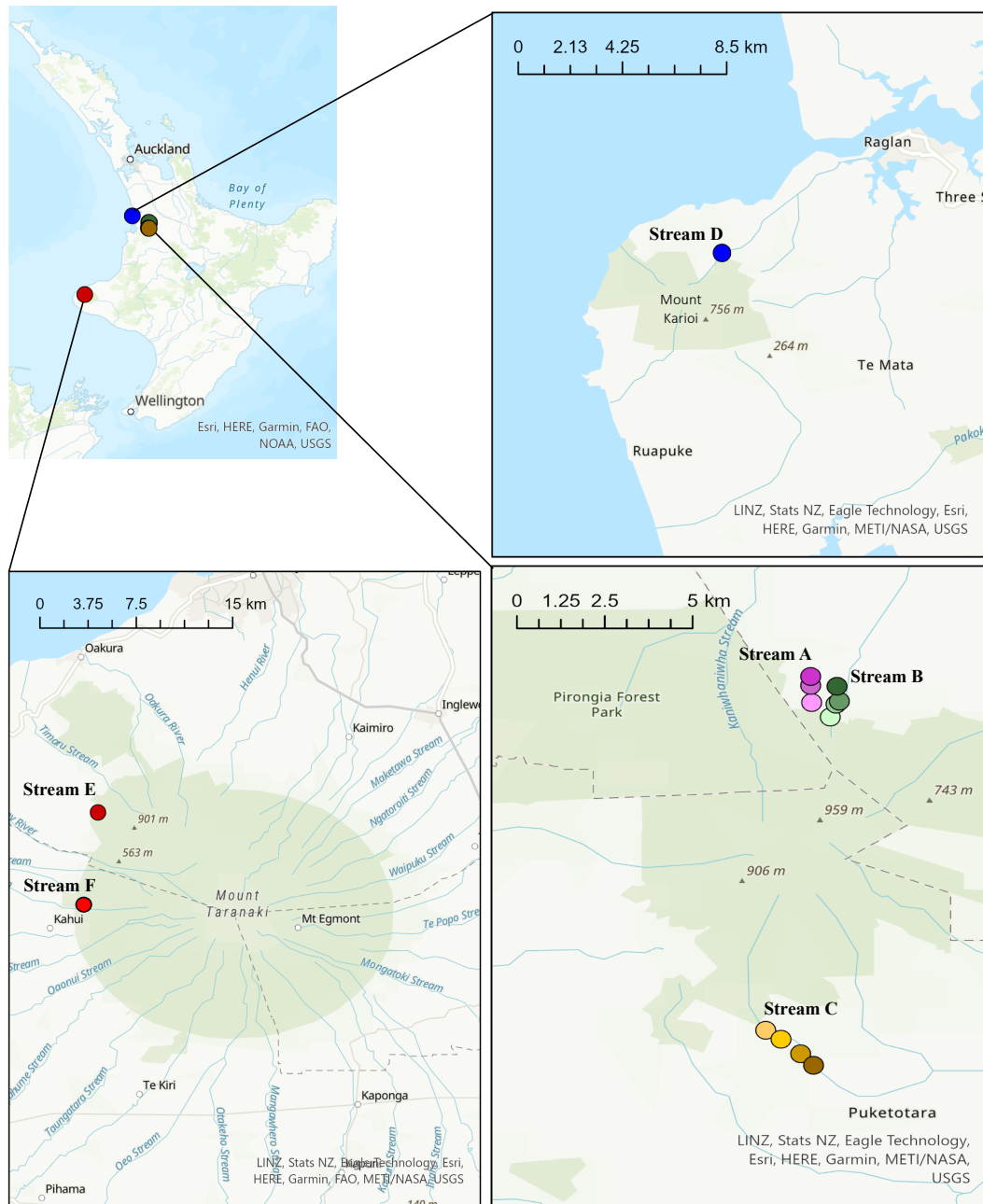
Figure 4.1 Location of 13 sampling sites distributed in five streams in the North Island of New Zealand. Dots show the smaller detail of sampling location within the stream. Further locality data are given in Appendix Table A.4.1.

4.3.2 Study species

We selected one species from each of the commonly encountered freshwater insect orders, Ephemeroptera, Plecoptera and Trichoptera. The mayfly *Coloburiscus humeralis* is widely distributed in New Zealand, with nymphs (living from 12 to 27 months) commonly inhabiting the underside of stones, predominantly

in riffles (Harding & Winterbourn, 1993). The stonefly *Zelandobius confusus* is widely distributed throughout New Zealand and nymphs (9—12 month lifespan; McLellan, 1993) are often found in accumulated leaf packs or woody debris within streams. The caddisfly *Hydropsyche fimbriata* is restricted to the North Island of New Zealand, where larvae (9—12 month lifespan; Cowley, 1978) can be mostly found in small, stony, forested streams (Winterbourn et al., 2000). All three species have a winged adult stage (living for a few days to weeks; Collier & Smith, 2000; Smith B., unpublished data), during which overland dispersal may occur.

4.3.3 Insect collection

We sampled 11 sites from three streams located in two neighbouring catchments. Samples were collected from 3—4 sites per stream at intervals of at least 490 m (Appendix Table A.4.1). All sampling sites were located at the base of the mountain at a similar elevation (Appendix Fig. A.4.2), but with different land cover (native forest, forest fragments, pasture; Appendix Fig. A.4.3). Tawhitiwhiti Stream (Stream A) included one forested site (upstream), and two sites covered by pasture; Te Pahu stream (Stream B) included one forested site, one forest fragment, one pasture site and one restored (riparian planted) site; Ngakoaohia Stream (Stream C) was mostly confined by steep limestone walls and was fully covered by indigenous forest in the riparian zone. Mount Karioi (Wainui Stream—Site D1) and Mount Taranaki (Katikara Stream—Site E1, and Patea Stream—Site F1) sites were both fully forested (Appendix Table A.4.1).

Caddisfly larvae and stonefly and mayfly nymphs were collected in austral summer periods between December 2017 and January 2020. Within each of the three locations, approximately 50 individuals were collected from first and second-order perennial and stony-bottom streams using a kick-net or hand-picked from the substrate and immediately preserved in 95% ethanol for further analysis. At Mount Taranaki, samples of *Z. confusus* were collected from site E1, whereas *C. humeralis* and *H. fimbriata* were collected from site F1. Nymphs of *Z. confusus* were identified (McLellan, 1993), and *C. humeralis* and *H. fimbriata* were confirmed, using Winterbourn et al. (2000). The nomenclature of *H. fimbriata* follows Geraci et al. (2010). Following morphological identification, samples were transferred to vials with fresh 95% ethanol and stored at -20°C until the left rear leg was dissected

from each individual using sterilised forceps and added to a single well of a 96-well PCR plate for DNA extraction.

4.3.4 DNA extraction, mtDNA and SNP sequencing

Genomic DNA was re-used or re-extracted from all 286 individuals reported in our final mtDNA data analysis (Chapter 3) and newly extracted from an additional 372 individuals, resulting in a final dataset of 658 individuals for SNP sequencing.

DNA extraction and amplification of mtDNA cytochrome *c* oxidase subunit I (COI) gene fragments was conducted by the Canadian Centre for DNA Barcoding following standard protocols (see Ivanova et al., 2006). In brief, DNA was extracted following the AcroPrepTM PALL Glass Fibre plate method using a total mix of 5 ml insect lysis buffer (0.5 ml of Proteinase K, 20 mg/ml per 96-well plate). A 658 bp region of the COI gene was PCR amplified using the primer pair LepF1 and LepR1 (Hebert et al., 2004, Wilson, 2012) and 5 μl of the DNA extraction product. PCR thermal cycling conditions were: initial denaturation of samples at 94°C for 1 min, followed by five cycles of 94°C for 30 s, 48°C for 1.5 min, and 72°C for 1 min. This was followed by 35 cycles of 94°C for 30 s, 52°C for 1 min, and 72°C for 1 min; with a final extension of 72°C for 10 min. PCR products were cleaned using Sephadex® and then sequenced using an ABI3730xl DNA analyser. All DNA sequence data have been added to the Barcode of Life Datasystems (BOLD) database (Ratnasingham & Hebert, 2007) and are available under dataset DS-EPTNZNI.

DNA extraction, sequencing and SNP genotyping were processed by Diversity Array Technology Pty Ltd (DarTseq™), Canberra, Australia. Genotyping was performed using a combination of DarTseq™ complexity reduction methods and next generation sequencing (NGS) to detect a large number of SNPs (Kilian et al., 2012). DNA samples were digested using the PstI-SphI restriction enzyme pair after a pilot study was performed to identify the enzyme combination most suitable for genome complexity reduction in the target species. The PstI-compatible forward adapter included the Illumina flow cell attachment sequence, sequencing primer and a barcode for sample identification within pooled libraries. The reverse adapter contained the Illumina flow cell attachment region and the SphI-compatible

overhang sequence. PstI-SphI ligated fragments were amplified by adapter-mediated PCR as follows: initial denaturation at 94°C for 1 min followed by 30 cycles of denaturation at 94°C for 20 s, annealing at 58°C for 30 s, and extension at 72°C for 45 s, and an additional final extension step at 72°C for 7 min. After PCR amplification, equimolar amounts of amplification product for each sample were pooled before 77 cycles of single-read NGS on the HiSeq2500 (Illumina) platform. Following sequencing, raw reads were processed using a proprietary DarT analytical pipeline, which performed filtering and variant calling, and generated final genotypes. Two technical replicates of each DNA sample were genotyped to calculate the reproducibility of the marker data. Finally, SNPs that were polymorphic across samples for each species were obtained from DarTseq™, with markers scored as binary data: '1' for presence, '0' for absence, and '-' for failure to score.

4.3.5 Spatial population structure analysis

mtDNA COI data. Haplotype networks were constructed, assigning individuals to their respective sampling site locations using the packages ape v. 5.5 (Paradis & Schliep, 2018) and pegas v.1.0 (Paradis, 2010) in R v. 4.1.2 (Team, 2020). Next, we partitioned total genetic variation into geographic hierarchies using independent Analysis of Molecular Variance (AMOVA) analyses in ARLEQUIN v. 3.5 (Excoffier & Lischer, 2010) with 10,000 permutations. For this analysis, populations were partitioned into: "between mountain regions", "between neighbouring catchments", "between adjacent streams", "within Stream A", "within Stream B", and "within Stream C". Simple Mantel tests were conducted to assess IBD using genetic (both raw F_{ST} and linearised F_{ST} (F_{ST} / (1 - F_{ST})), as per the suggestion of Rousset (1997) and geographic distances (Euclidean distance and shortest waterway distance). Mantel tests were conducted in three hierarchical groups: "among all sampling sites", "among sites within Mount Pirongia" and "among sites within each stream". These analyses were calculated using the *'mantel'* function in the R package vegan v. 2.5-7 (Oksanen et al., 2020), with *P*-values estimated from 10,000 permutations. Outcomes based on raw and linearised F_{ST} were qualitatively very similar, as were results considering either waterway or Euclidean distance, thus we present only the latter in our results.

Due to the low sample size or absence of sequence data in some locations, analyses of spatial genetic structure between *H. fimbriata* populations were only calculated for sites within the Pirongia region.

SNP data. Data quality control and filtering were performed using the R package DartR v. 1.9.9.1 (Gruber et al., 2018). After testing different parameter combinations (e.g. missing data 5%, 10%, 20%, and minor allele frequency — MAF<0.02, <0.05), showed no significant differences in the downstream population genetic data analysis (data not shown), I proceeded with the following parameter setting for SNP filtering: SNP markers with more than 20% missing data, MAF<0.05, and/or an unknown position were removed. Pairwise population genetic differentiation was estimated using F_{ST} (Weir & Cockerham, 1984) and Nei's distance (Nei, 1972), using the R package StAMPP v. 1.6.2 (Pembleton et al., 2013). Visualisation of population structure was performed using Principal Coordinates Analysis (PCoA) in DartR with associated functions from the ADEGENET package v. 2.1.4 (Jombart, 2008). Separate PCoA plots were also generated for each of the three streams to investigate whether structure between populations within streams varied with riparian land cover. fastSTRUCTURE v. 1.0 (Raj et al., 2014) was used to identify admixture proportions among individuals and populations. In fastSTRUCTURE, we tested 1–13 genetic clusters (depending on the availability of data from each sampling site/population) by specifying the *K* parameter, and ran five replicates for each *K*-value, using a simple prior. Results for each *K*-value were visualised using the distruct.py script, and model complexity was chosen using the chooseK.py script, both associated with the fastSTRUCTURE program. To aid visualisation of population structure, we investigated the association of the resulting genetic clusters to the type of riparian land cover at each sampling site. AMOVA was performed with the same geographic hierarchies as described for the COI data, using the poppr v. 2.9.3 package (Kamvar et al., 2014) in R. Mantel tests were performed on the SNP data as described for the COI data (above).

4.3.6 Fine-scale landscape genetics analysis for Mount Pirongia populations

Landscape data and rasters. We calculated two connectivity indices to test for IBR: land cover resistance and topographic resistance. For the land cover resistance (LCR), we used the Land Cover Database (10 m resolution) for mainland New Zealand (LCDB v. 5.0) from the Land Resource Information System (LRIS). We chose the updated map from 2018, as this was the closest time to our sampling period. The original shapefiles were processed to raster maps using Geopandas v. 0.10.2 (Jordahl et al., 2020) and Geocube v. 0.1.2 libriaries in Python v. 2.7 (Van Rossum & Drake Jr, 1995), and the data classes of interest were filtered. Seven forest land cover-related classes and four grassland cover-related classes were used to build a final raster with two classes, 'forest' and 'pasture' (Appendix Fig. A.4.2). The analysis of forest versus pasture was included to test whether open pasture land is a barrier to dispersal. We selected resistance costs based on the assessment of a range of resistance values (Richardson, 2012), rather than based on expert opinion alone—a practice commonly used in previous studies which has been questioned as open to subjectivity (Spear et al., 2010). Overall, six land cover rasters were generated using three different resistance values assigned to each of the two categorical landscape variables. By examining a range of potential values, we could also determine the potential resistance of either vegetation or pasture to connectivity (and therefore influence on dispersal) for each of the three insects.

For topographic resistance, we used an 8m resolution digital elevation model (DEM) sourced from the Land Information New Zealand (LINZ) to calculate a topographic complexity raster based on the topographic slope values (Appendix Fig. A.4.3), using the GDAL library v. 3.4.1 (GDAL/OGR contributors, 2020). Slope values were further interpolated to the same geographical points as represented by the LCR data. 'Slope' was chosen as a topographic variable based on the potential preference for adult flight within stream valley corridors confined by steep sloping sides that may also inhibit lateral movement (Hughes et al., 1999; Phillipsen et al., 2015; Winterbourn, 2007). We used the raw slope values of the map pixels as a continuous variable to assign resistance values to our maps, with values varying from 1 (indicating flat areas and lowest resistance), to 60 (deepest slope and highest resistance), rather than assigning relative costs as suggested by Spear et al (2010). This enabled us to test the prediction that higher slope resulted

in lower levels of gene flow or higher genetic differentiation. The final resistance raster was transformed to a 10 m resolution for subsequent analysis.

Estimating resistance distances from landscape rasters. We calculated resistance distances between populations for the topographic raster and each of the potential six land cover resistance rasters by: 1) estimating pairwise least-cost pathways between populations using the *'genleastcost'* function in the R package PopGenReport v. 3.0.4 (Adamack & Gruber, 2014), which calculates the shortest distance between pairs of populations based on the given landscape resistance; and 2) using CIRCUITSCAPE v. 4.0 (McRae, 2006). The CIRCUITSCAPE approach is based on circuit theory, and uses information on how current flows across the resistance landscape to estimate the resistance distance between focal points. It enables calculation of the resistance of the landscape to gene flow between pairs of populations, while assuming that movement of individuals is not optimal across the landscape, so that multiple paths contribute to effective dispersal.

Resistance model optimisation and exclusion of spurious variables. We next conducted an exploratory analysis using simple Mantel tests to optimise our resistance model. Each pairwise resistance distance from the least-cost and CIRCUITSCAPE analysis was correlated with genetic distance (F_{ST}) values to compare the fit and significance of the relationship for each of the competing models. Final land cover resistance costs and variables were selected based on the results of the simple Mantel tests: for each species, the land cover resistance representation with the highest correlation coefficient was chosen and any variables that did not lead to significant correlation were excluded from further analysis. We also used the Mantel test results to choose between least-cost or CIRCUITSCAPE distances matrices, with the highest coefficient retained.

Assessing landscape variables importance. To assess the relative importance of different landscape variables identified in the exploratory Mantel analysis, we used multiple regression on distances matrices (MRM; Balkenhol et al., 2009; Wang, 2013). This method is advantageous for analysis of landscape genetic data because it enables the use of distinct distance matrices simultaneously, allowing inferences to be made at the level of individual landscape variable (Lichstein, 2007). We conducted MRM analysis on seven potential models to investigate the importance

of geographic distance (IBD) versus landscape resistance (IBR) on shaping spatial genetic structure, namely whether the amount of genetic variation was: 1) solely due to spatial influence (i.e. IBD — Euclidean distances); 2) a correlation purely attributable to the heterogeneous landscape, (i.e. IBR, which is represented by resistance distances from topographic and land cover rasters); or 3) influenced by a shared component, which cannot be separated into purely spatial IBD versus purely landscape IBR contributions (e.g. because of correlations between Euclidean and resistance distances).

Analyses were conducted using the *'MRM'* function in the R package ECODIST v. 2.0.7 (Goslee & Urban, 2007). Because each effective distance was used for multiple statistical tests, inferences on both Mantel tests and MRM were based on Bonferroni-corrected *P*-values for multiple comparisons. We assessed multi-colinearity in the full models using variance inflation factors (VIF) with the *'vif'* function in the R package car v. 3.0-12 (Weisberg & Fox, 2011). All variables showed VIF scores <6 and were therefore retained in the candidate models analysed (Dormann et al., 2013; Zuur et al., 2010). Next, we used hierarchical partitioning to assess the relative importance of the three landscape variables for predicting spatial genetic structure. This approach, commonly applied in ecological studies, aims to determine which of the various independent variables included in the model has the strongest influence on the response variable by calculating the average variable's contribution to the response variable over all possible combinations of the independent variables (Murray & Conner, 2009). Hierarchical partitioning has recently been applied in landscape genetic studies (Balkenhol et al., 2020) as an alternative to other model selection procedures (e.g. based on Akaike Information Criterion, AIC). Procedures based on AIC values are discouraged for use with MRM because of a potential bias towards selecting unnecessarily complex models with spurious variables resulting in the erroneous ranking of resistance models (Franckowiak et al., 2017).

4.4 Results

4.4.1 Spatial genetic structure

Coloburiscus humeralis

COI. The haplotype network consisted of 16 haplotypes, with a few derived haplotypes (including 11 singletons) connecting to a centrally located dominant haplotype (H1). Haplotype H1 was found in 59 individuals located only in Pirongia. Two haplotypes were restricted to the Karioi sampling site, whereas most of the individuals from Taranaki shared the same haplotype with individuals from Pirongia (Fig. 4.2). Global F_{ST} revealed high genetic differentiation among all sampling sites (0.428, $P<0.001$). AMOVA across multiple geographic hierarchies only indicated significant differentiation only at the broadest spatial scale (among mountain regions; $F_{ST}=0.675$, $P<0.05$; Table 4.1). The Mantel test provided no evidence for a correlation between Euclidean distance and genetic differentiation across all sampling sites ($r=0.009$, $P=0.157$) (Appendix Fig. A.4.1).

SNPs. The filtered data set consisted of 4,609 SNPs (1.09% missing data) from 208 individuals (Appendix Table A.4.3). AMOVA analysis revealed low but significant differentiation among all sampling sites ($F_{ST}=0.074$, $P<0.001$, 7.4% variance) and across all spatial scales (Table 4.1), with the highest F_{ST} detected among mountain regions ($F_{ST}=0.060$, $P<0.05$) and the lowest detected within stream C ($F_{ST}=0.006$, $P<0.05$, 7.4%). This result contrasts the COI data, where hierarchical AMOVA analysis only detected significant genetic differentiation among regions. Pairwise F_{ST} values were slightly higher than pairwise Nei's D values, but both parameters showed similar patterns of divergence across pairs of populations (Table Appendix A.4.4). The highest pairwise F_{ST} value was 0.121 (B1/E1), which reflected geographic distance (170.8 km). Contrasting with the COI results, a strong IBD pattern was found within the overall study area (Euclidean distance, $r=0.689$, $P=0.001$; Appendix Fig. A.4.1).

In the PCoA, the two principal coordinates explained 6.1% and 2.5% of the total genetic variation, confirming a general pattern of low genetic differentiation between all populations (Fig. 4.2). E1 (Taranaki) clustered separately from the remaining populations, and D1 (Karioi) clustered together with most individuals

from stream A and B in Pirongia, but separately from stream C. When examining clustering patterns associated with land cover type (forested x pasture) in individual streams, the PCoA showed that individuals from forested sites (A1 and B1 in Fig. 4.3) tended to cluster together in fragmented streams. In the fully forested stream C, no clustering associated with a specific sampling site was observed.

fastSTRUCTURE analysis suggested that K values of 3 and 6 were most likely (for maximising marginal likelihood and explaining structure in the data, respectively; Fig. 4.4). For K=3, only individuals from Taranaki were assigned to a distinct population, and admixture between the remaining genetic clusters was observed across Pirongia and Karioi sites. For K=6, assignment of individuals to clusters was mainly region-specific, separating populations of Pirongia, Karioi and Taranaki (Fig. 4.4). Within Pirongia, clusters were not exclusive to sampling sites. Of particular note, for streams A and B, lower admixture between clusters was found for the source forest sites A1 and B1, and admixture increased for the remaining downstream sites within both streams (Fig. 4.5). These results are consistent with the PCoA showing separate groups for A1 and B1 sampling sites. Increased admixture between genetic clusters was also observed for sampling sites within the fully forested stream C when compared to the fragmented streams A and B.

Zelandobius confusus

COI. A total of 14 haplotypes were identified, with three haplotypes being common in the network (Fig. 4.2). The central haplotype (H1) was shared among the three different mountain regions while the dominant haplotype (H2) was separated from the central one by three mutational steps and was found only at Pirongia sites. The third most frequent haplotype (H3) was found in individuals from Karioi and Pirongia and was separated from the central one by four mutational steps. Most of the remaining haplotypes split off from the central one and were found in a few or single individuals. Global F_{ST} revealed significant genetic differentiation among all sampling sites (0.102, P<0.01). AMOVA analysis showed significant genetic differentiation when comparing populations among mountain regions (F_{ST}=0.214, P<0.05). Within Pirongia, the variation between populations at the remaining hierarchical levels was non-significant (Table 4.1). Mantel tests across all sampling

sites showed a significant correlation between Euclidean distance and genetic differentiation ($r=0.588$, $P=0.005$; Appendix Fig. A.4.1), indicating IBD.

SNPs. The filtered data set consisted of 6,388 SNPs (2.48% missing data) and 160 individuals (Appendix Table A.4.3). AMOVA analysis revealed low but significant genetic differentiation among all sampling sites ($F_{ST}=0.048$, $P<0.001$, 4.8% variance). Across different geographic hierarchies, significant differentiation was observed among mountain regions, as reported for COI, but also between neighbouring catchments ($F_{ST}=0.016$, $P<0.05$ and $F_{ST}=0.044$, $P<0.05$, respectively). Pairwise F_{ST} values were slightly lower than pairwise Nei's D values for most of the pairwise comparisons, but both parameters showed similar patterns of divergence across pairs of populations (Appendix Table A.4.4). Very low but significant genetic differentiation (F_{ST}) was observed for population comparisons separated by neighbouring catchments, consistent with the AMOVA results. The highest pairwise F_{ST} value was 0.104 (A1/E1), which reflected the geographic distance separating these sampling sites (170.8 km). As observed for COI, a strong IBD pattern was found within the overall study area ($r=0.795$, $P=0.001$).

In the PCoA, the two principal coordinates explained 2.6% and 1.7% of the total genetic variation, confirming a general pattern of low genetic differentiation between populations (Fig. 4.2). Two main clusters of individuals were observed, separating population E1 (Taranaki) from the remaining individuals located across Pirongia. Within individual streams, PCoA analysis showed no clustering of individuals associated with the riparian land cover (Fig. 4.3). The fastSTRUCTURE analysis suggested that *K* values of 2 and 4 maximised marginal likelihood and explained structure in the data, respectively (Fig. 4.4). The *K*=2 result was consistent with the PCoA results, with two distinct genetic clusters clearly assigning individuals to populations from two distinct mountain regions: Pirongia (A1 to C3) and Taranaki (E1). There was no evidence for finer spatial genetic structure, with *K*=4 mostly showing admixture between Pirongia sampling sites. As for the PCoA, we did not observe assignment of individuals to a particular genetic cluster that was associated with the land cover type (Fig. 4.5).

Hydropsyche fimbriata

COI. The haplotype network showed that most individuals were assigned to 4 of the 19 identified haplotypes. The most dominant haplotype (H13; n=23) and haplotype H8 (n=14) were shared only between the adjacent streams A and B, whereas H3 (n=0) and H17 (n=18) were shared among all three studied streams in Mount Pirongia (Fig. 4.2). AMOVA showed significant genetic differentiation between neighbouring catchments (F_{ST}=0.142, *P*<0.01) and within Stream A (F_{ST}=0.120, *P*<0.05; Table 4.1). The Mantel test indicated IBD within the Pirongia region (*r*=0.347, *P*=0.010).

SNPs. The filtered data set consisted of 1,789 SNPs (1.13% of missing data) from 290 individuals (Appendix Table A.4.3). AMOVA analysis revealed significant genetic differentiation among all sampling sites (F_{ST}=0.069, *P*<0.01, 6.9% variance), among mountain regions (F_{ST}=0.180, *P*<0.01) and between neighbouring catchments (F_{ST}=0.012, $P < 0.05$), but not within individual streams (F_{ST}=0.000), contrasting the COI data for stream A (Table 4.1). Pairwise F_{ST} values were similar to pairwise Nei's *D* values for most of the pairwise comparisons, and both parameters showed similar patterns of divergence across pairs of populations (Appendix Table A.4.4). Very low but significant genetic differentiation (F_{ST}) was observed for population comparisons within Pirongia. Higher F_{ST} was found when comparing sampling sites from distinct mountain regions; for example, the highest pairwise F_{ST} value was 0.238 (A1/F1), which reflected the geographic distance separating these sampling sites (170.8 km). A strong IBD pattern was found within the overall study area, including the three mountain regions (*r*=0.706, *P* 0.001).

Individuals fell into four clusters in the PCoA, where the first two principal coordinates explained 5.6% and 2.2% of the genetic variation (Fig. 4.2). Spatial structure was only observed for individuals from Taranaki (F1), which grouped together exclusively. Within individual streams, the PCoA showed no clustering of individuals that associated with riparian land cover (Fig. 4.3). fastSTRUCTURE analysis suggested that K values of 3 and 5 were most likely for maximising the marginal likelihood and explaining structure in the data, respectively (Fig. 4.4). Individuals from Taranaki that grouped together in the PCoA were assigned to a single genetic cluster in the fastSTRUCTURE analysis. This cluster also indicated a very low proportion of admixture for a few individuals from sampling sites in Pirongia. For *K*=3, Pirongia and Karioi shared two clusters with similar admixture

levels across all sampling sites, whereas at $K=5$ an additional frequent genetic cluster appeared that showed admixture of individuals mostly from the fully-forested Stream C (Fig. 4.5).

4.4.2 Fine-scale landscape genetic analysis within Pirongia populations

Simple Mantel correlations obtained for each landscape variable and each species dataset are shown in Table 4.2. Mantel correlations for resistance distances based on circuit theory versus least-cost analysis generally led to similar coefficients. The highest coefficient for the variable topography was obtained with least-cost ($r=0.686$; $r=0.640$ with circuit theory, $P=0.001$), whereas the contrary was found for the land cover variable ($r=0.709$ with circuit theory; $r=0.557$ with least-cost, $P=0.001$). Mantel correlations also revealed proximate coefficient values among the potential land cover costs included (2:5:10). Accordingly, we constructed the remaining landscape resistance models analyses using resistances distances generated with circuit theory and chose the land cover variable showing the highest coefficient as follows: forest:pasture 2:1 for *Z. confusus* and *H. fimbrita*, and forest:pasture 5:1 for *C. humeralis* (Table 4.2).

For all three studied species, Mantel correlations showed higher coefficients when resistance distances were analysed using the SNP dataset compared to the COI dataset (Table 4.3). These analyses showed weak support for IBD ($r=0.285$, $P<0.05$) and IBR (land-cover $r=0.410$, $P<0.02$) for the mayfly *C. humeralis*, although P-values were not significant after Bonferroni correction. In contrast, there was a significant relationship between each of the three variables (Euclidean distance, typography, and land cover) and genetic differentiation in both the stonefly *Z. confusus* and the caddisfly *H. fimbriata* SNP datasets (all P-values<0.005). In addition, the land cover variable had a significant relationship ($r=0.417$, $P=0.003$) with the COI dataset for *H. fimbriata*.

Results of the MRM analysis were consistent with the Mantel test results for *C. humeralis*, where a correlation between the landscape and genetic differentiation was only found for a single model using the land cover variable ($R^2=0.169$, $P=0.047$), and was not significant after Bonferroni correction (Table 4.3). In *Z. confusus*, all seven candidate models showed a significant correlation.

The highest was found for the full model including all three independent variables (R^2=0.338, P=0.001). All the candidate models also showed significant correlation with the SNP dataset for *H. fimbriata*, with the highest R^2 value found for the model combining the independent topography and land cover variables (R^2=0.524, P=0.001). Additionally, the land cover variable showed a significant correlation with the COI genetic distances for *H. fimbriata* (R^2=0.221, P=0.004), corroborating the Mantel tests results (Table 4.3). Evaluating individual variable importance to the full models (Fig. 4.6), land cover was the best predictor of genetic distances for *C. humeralis* (52% of the contribution to the model), whereas topography (39%) and Euclidean distance (35%) made important contributions to the full model for *Z. confusus*. Topography and land cover were the best predictors of genetic distance for the COI *H. fimbriata* dataset (45% and 37%, respectively), while land cover had the highest contribution (39%), followed by Euclidean distance (32%) and slope (29%) for the SNP dataset.

Table 4.1 Hierarchical analysis of molecular variance (AMOVA) from both markers for each species.

Hierarchical level	*C. humeralis* COI data Variance %	F_{ST}	SNP data Variance %	F_{ST}	*Z. confusus* COI data Variance %	F_{ST}	SNP data Variance %	F_{ST}	*H. fimbriata* COI data Variance %	F_{ST}	SNP data Variance %	F_{ST}
Among all sites	42.85	0.428***	7.40	0.074*	10.24	0.102**	4.83	0.048*	9.23	0.092**	6.97	0.069**
Within all sites	57.15		92.60		89.76		95.16		90.77		93.03	
Across large spatial scales												
Among mountain regions	67.58	0.675*	6.08	0.060*	21.45	0.214*	1.60	0.016*	-	-	18.01	0.180**
Within Pirongia (main study area)												
Between neighbouring catchments	0.90	0.009	2.50	0.025*	0.10	0.001	4.42	0.044*	14.23	0.142**	1.21	0.012*
Between adjacent streams	0.94	0.009	1.27	0.012*	9.21	0.092	0.12	0.001	0.00	0.000	0.39	0.003
Within stream A	0.00	0.000	4.83	0.048*	0.00	0.000	0.35	0.003	12.01	0.120*	0.00	0.000
Within stream B	6.10	0.061	2.88	0.028*	5.79	0.057	0.35	0.003	0.74	0.007	0.56	0.005
Within stream C	4.06	0.040	0.69	0.006	0.00	0.000	0.42	0.000	0.00	0.000	0.17	0.001

*$P<0.05$; **$P<0.01$; ***$P<0.001$. - indicates that values could not be obtained due to sample size constraints.

Table 4.2 Exploratory data analysis using Mantel test correlation coefficient (*r*) for each landscape variable representing both least-cost and CIRCUITSCAPE resistance distances.

Species	Category	Variable	Cost values	Least-cost COI	Least-cost SNPs	Circuitscape COI	Circuitscape SNPs
C. humeralis	Topography	Slope	min=1, max=60	0.055, *P*=0.370	0.263, *P*=0.0379	0.003, *P*=0.495	0.362, *P*=0.061
	Land cover	vegetation:pasture	2:1	0.056, *P*=0.357	0.228, *P*=0.067	0.148, *P*=0.197	0.382, *P*=0.029
			5:1	0.069, *P*=0.360	0.246, *P*=0.069	0.197, *P*=0.194	0.410, *P*=0.022
			10:1	0.087, *P*=0.294	0.236, *P*=0.074	0.221, *P*=0.169	0.409 *P*=0.033
		pasture:vegetation	2:1	0.043, *P*=0.391	0.225, *P*=0.051	-0.086, *P*=0.690	-0.041, *P*=0.575
			5:1	0.030, *P*=0.438	0.204, *P*=0.063	-0.196, *P*=0.816	-0.285, *P*=0.906
			10:1	0.016, *P*=0.460	0.164, *P*=0.109	-0.227, *P*=0.848	-0.333, *P*=0.930
Z. confusus	Topography	Slope	min=1, max=60	-0.060, *P*=0.0701	0.574, *P*=0.001	-0.011, *P*=0.513	0.538, *P*=0.001
	Land cover	vegetation:pasture	2:1	-0.075, *P*=0.754	0.543, *P*=0.004	-0.099, *P*=0.712	0.490, *P*=0.003
			5:1	-0.077, *P*=0.744	0.557, *P*=0.001	-0.126, *P*=0.749	0.404, *P*=0.008
			10:1	-0.073, *P*=0.734	0.553, *P*=0.001	-0.146, *P*=0.739	0.350, *P*=0.015
		pasture:vegetation	2:1	-0.080, *P*=0.734	0.524, *P*=0.004	-0.051, *P*=0.640	0.402, *P*=0.013
			5:1	-0.096, *P*=0.783	0.517, *P*=0.004	-0.040, *P*=0.572	0.138, *P*=0.247
			10:1	-0.105, *P*=0.801	0.500, *P*=0.004	-0.052, *P*=0.590	0.065, *P*=0.414
H. fimbriata	Topography	Slope	min=1, max=60	0.356, *P*=0.009	0.686, *P*=0.001	0.508, *P*=0.012	0.640, *P*=0.001
	Land cover	vegetation:pasture	2:1	0.335, *P*=0.018	0.667, *P*=0.001	0.417, *P*=0.003	0.709, *P*=0.001
			5:1	0.341, *P*=0.012	0.673, *P*=0.002	0.488, *P*=0.005	0.649, *P*=0.001
			10:1	0.325, *P*=0.027	0.665, *P*=0.002	0.480, *P*=0.018	0.605, *P*=0.001
		pasture:vegetation	2:1	0.336, *P*=0.009	0.661, *P*=0.002	0.001, *P*=0.494	0.354, *P*=0.013
			5:1	0.307, *P*=0.014	0.640, *P*=0.002	-0.317, *P*=0.902	-0.086, *P*=0.631
			10:1	0.250, *P*=0.040	0.597, *P*=0.003	-0.395, *P*=0.951	-0.215, *P*=0.835

Table 4.3 Correlation between landscape variables and genetic differentiation using simple Mantel tests and multiple regression on distances matrices (MRM).

	Coloburiscus humeralis				*Zelandobius confusus*				*Hydropsyche fimbriata*			
	COI		SNP		COI		SNP		COI		SNP	
(a) Mantel test	r	P-value	r	P-value	r	P-value	r	P-value	r	P-value	r	P-value
Euclidean distance (GEO)	0.104	0.213	0.285	0.043	0.000	0.803	0.675	0.002*	0.347	0.010	0.715	0.001*
Slope (TOPO)	0.003	0.495	0.362	0.061	-0.011	0.513	0.538	0.001*	0.508	0.012	0.640	0.002*
forest:pasture (LAND)	0.197	0.194	0.410	0.022	-0.099	0.712	0.490	0.003*	0.417	0.003*	0.709	0.001*
(b) MRM	R^2	P-value	R^2	P-value	R^2	P-value	R^2	P-value	R^2	P-value	R^2	P-value
1 Full model: TOPO + LAND + GEO	-	-	0.200	0.232	-	-	0.338	0.001*	0.300	0.119	0.513	0.002*
2 GEO	-	-	0.052	0.064	-	-	0.279	0.004*	0.117	0.018	0.444	0.003*
3 TOPO	-	-	0.136	0.052	-	-	0.290	0.001*	0.258	0.009	0.410	0.001*
4 LAND	-	-	0.169	0.047	-	-	0.240	0.003*	0.221	0.004*	0.502	0.001*
5 TOPO + GEO	-	-	0.138	0.186	-	-	0.340	0.002*	0.258	0.066	0.504	0.002*
6 LAND + GEO	-	-	0.177	0.109	-	-	0.282	0.005*	0.243	0.045	0.511	0.002*
7 TOPO + LAND	-	-	0.180	0.190	-	-	0.304	0.001*	0.273	0.058	0.524	0.001*

* variable/model still significant after Bonferroni P-value correction.

GEO: geographic distance (i.e. Euclidean distance); TOPO: Topography (i.e. slope); LAND: land cover (i.e. forest versus pasture).

NA: Not available. Results not shown for spurious landscape variables excluded after exploratory data analysis.
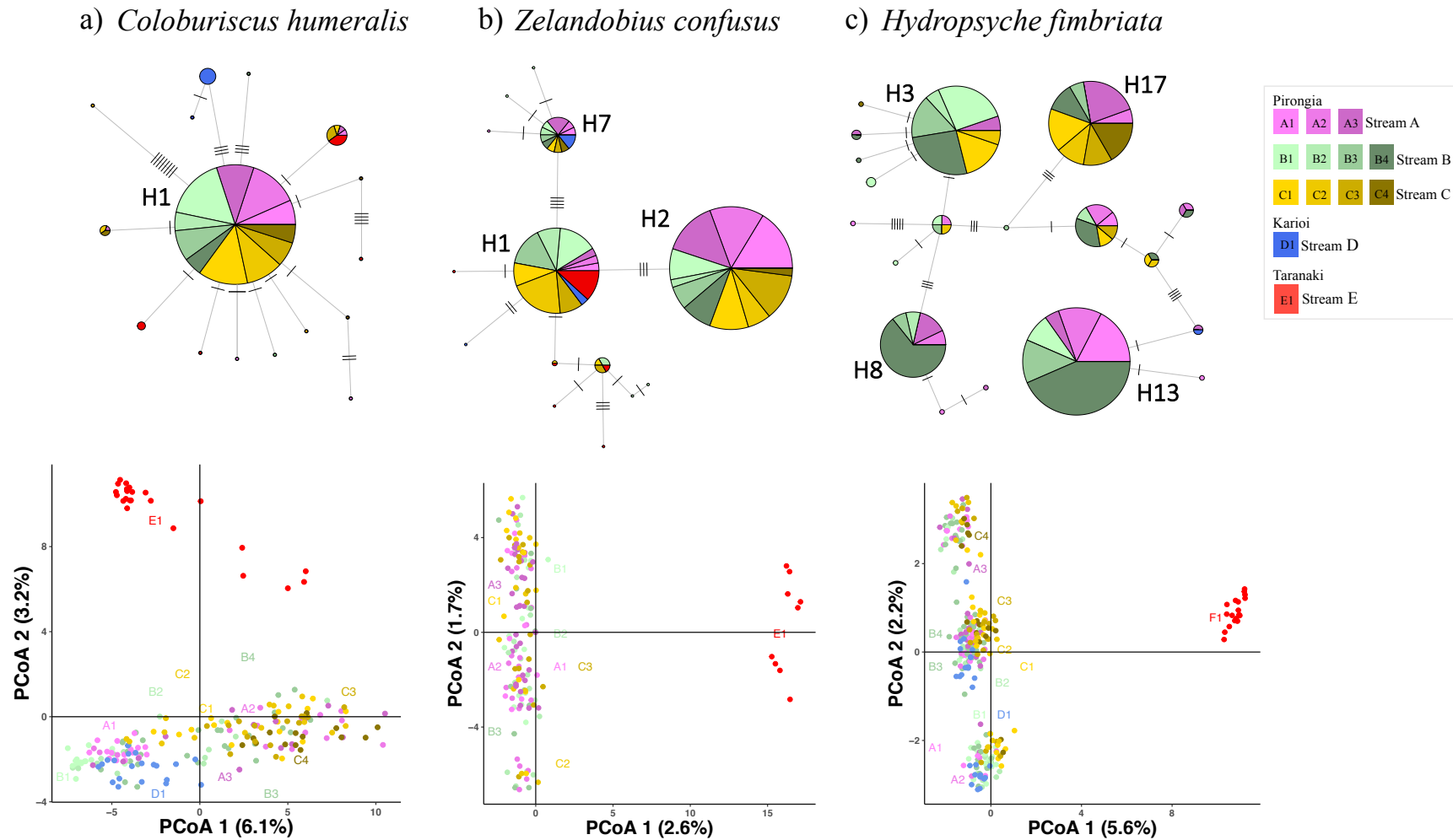
Figure 4.2 Genetic relationships among individuals and sampling locations for each species. Individuals are colour-coded according to collection locality and population label. Above: Haplotype network based on COI sequence data. Each pie chart represents a single haplotype, with the size proportional to the frequency of individuals containing that particular haplotype at each site where samples were collected. Dashes indicate missing mutational steps. Below: Principal Coordinate Analysis (PCoA) based on orthogonal transformation of SNP data. The percentage of variation explained by each principal coordinate is indicated on the axes.
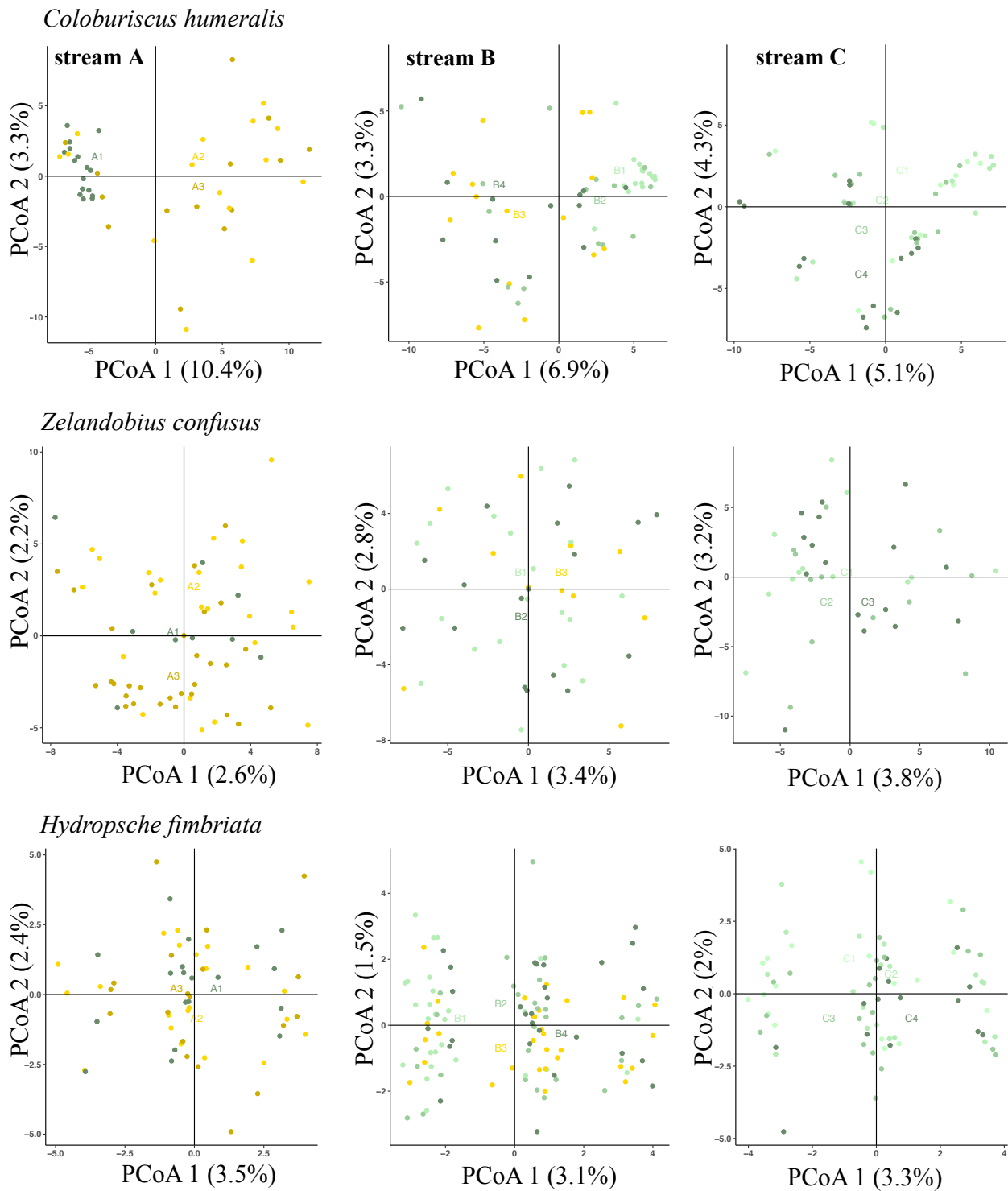
Figure 4.3 Two-dimensional plots of Principal Coordinate Analyses (PCoA) based on orthogonal transformation of SNP data for *C. humeralis*, *Z. confusus,* and *H. fimbriata*. Data were analysed for each stream in Pirongia separately. Individuals are colour-coded according to the land cover type (forest x pasture) of each collection locality: shades of green indicate sampling sites fully covered by riparian forest, and shades of yellow indicate sampling sites covered by pasture open land. The percentage of variation explained by each principal coordinate is indicated on the axes.

Coloburiscus humeralis

forest / pasture / pasture / forest / forest / pasture / replanted / forest / forest / forest / forest / forest / forest

K=3

K=6

A1 (N=15) / A2 (N=15) / A3 (N=14) / B1 (N=15) / B2 (N=15) / B3 (N=15) / B4 (N=15) / C1 (N=16) / C2 (N=15) / C3 (N=15) / C4 (N=15) / D1 (N=20) / E1 (N=23)

Stream A / Stream B / Stream C / KARIOI / TARANAKI

PIRONGIA

Zelandobius confusus

forest / pasture / pasture / forest / forest / pasture / forest / forest / forest / forest

K=2

K=4

A1 (N=9) / A2 (N=21) / A3 (N=33) / B1 (N=20) / B2 (N=18) / B3 (N=10) / C1 (N=12) / C2 (N=12) / C3 (N=16) / E1 (N=9)

Stream A / Stream B / Stream C / TARANAKI

PIRONGIA

Figure 4.4 fastSTRUCTURE bar plots representing population structure and admixture coefficients for each of the three species. Colours in each panel represent the assignment of individuals to each genetic cluster. Single bars with >1 colour indicate admixture of that particular individual (i.e. sharing of genetic ancestry across more than one genetic cluster). Sampling sites are designated by their sampling sites codes and locations, and riparian land cover type is indicated for each site.
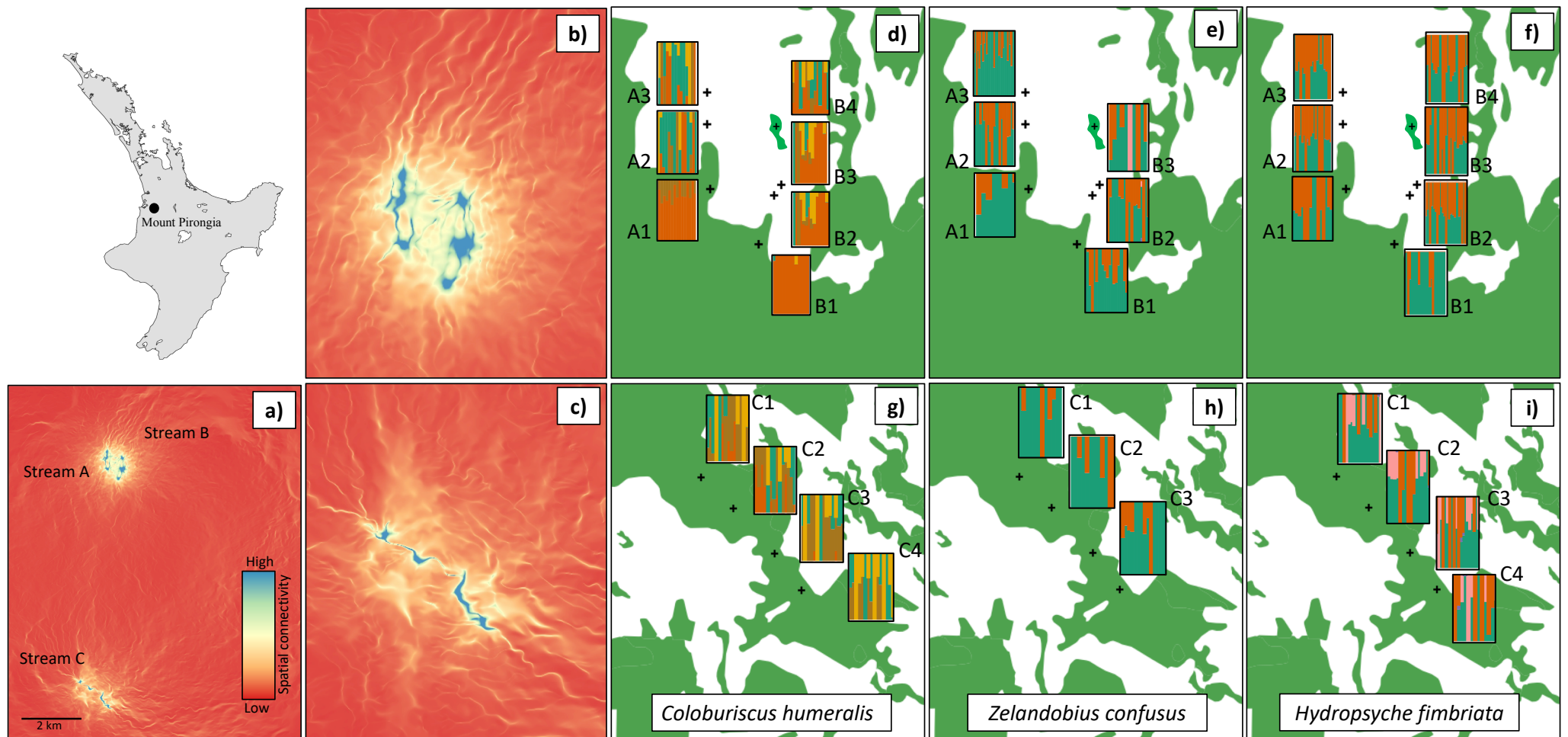
Figure 4.5 Spatial connectivity (a—c) on the left within Mount Pirongia study area based on topographic complexity calculated from a digital elevation model. Assignments of fastSTRUCTURE genetic clusters found for the 11 sampling sites for each of the studied species are shown in panels (d—i) on the right, with land cover type coloured with: green (forest) and white (pasture).
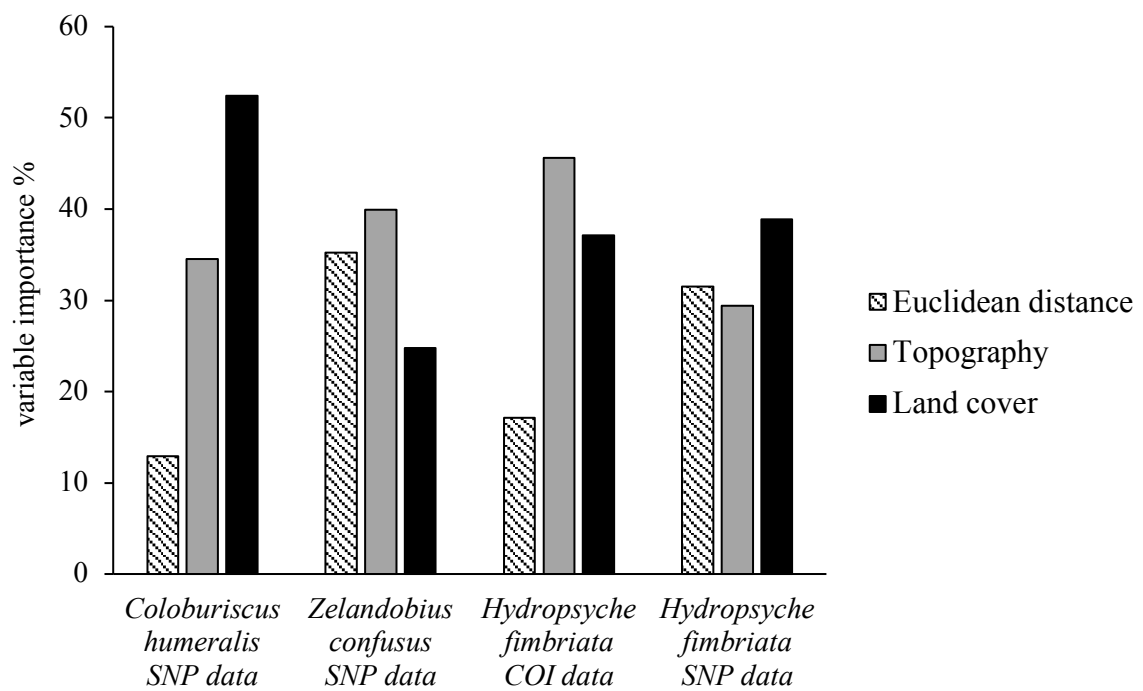
Figure 4.6 Relative variable importance for full MRM models based on hierarchical partitioning for each studied species. Variable importance is given as the percentage contribution of each variable to overall variation explained.

4.5 Discussion

Functional connectivity and the influence of landscape elements on gene flow was examined for three endemic aquatic insects within a fragmented landscape in the North Island of New Zealand. For all three species, we found an overall pattern of population structure followed by IBD among the three mountain regions (separated by up to ~170 km). However, within species, we found differing degrees of population connectivity at spatial finer scales that may correspond to different dispersal abilities and/or land cover types. We also found discordance between mtDNA and SNP results for *C. humeralis* and *H. fimbriata*, where greater genetic differentiation was observed in the mtDNA than in the SNP dataset.

4.5.1 Fine-scale population structure and landscape influence on connectivity

To date, few studies have investigated landscape-genetic relationships in aquatic insects (e.g. Finn et al., 2006; Phillipsen et al., 2015; Polato et al., 2017), and the effects of habitat fragmentation on populations is rarely addressed (e.g. Keller & Holderegger, 2013). However, variation in land use can affect abundance and aerial dispersal of aquatic insects. For example, Winterbourn (2007) found that different montane species were more abundant in forested habitats than in grasslands, while greater abundance of caddisflies has been found in Swedish agricultural stream systems compared to forested habitats (Carlson et al., 2016). Local dispersal and the degree of habitat fragmentation are key factors affecting connectivity between populations and are important considerations for stream restoration measures (Harding et al., 2006; Parkyn & Smith, 2011; Petersen et al., 2004). Here, we found that landscape features (IBR) were more important than pure space (IBD) in predicting genetic structure for all examined landscape resistance scenarios. However, we identified species-specific differences in the number and relative importance of the predictive landscape features.

The local pattern in *C. humeralis* showed low genetic differentiation (mtDNA and SNP), suggesting substantial population connectivity across the main study area. However, fine-scale differentiation in the SNP dataset indicated that gene flow may not be widespread between all sampling sites, and our PCoA,

fastSTRUCTURE and MRM analyses collectively suggested that land cover of the riparian zone may have a weak influence on genetic structure. In particular, populations in fragmented streams (Streams A and B) tended to be more differentiated, whereas weaker structure was found between sites connected by riparian forest within Stream C. In other words, population connectivity is higher when the stream channel is fully covered by riparian vegetation for *C. humeralis*. For some adult aquatic insects, dispersal within the stream corridor may be preferred when the riparian zone is protected by forest, offering better habitat conditions for survival, including cooler and moist microclimate and reduced wind speed (Collier & Smith, 1997; Harding et al., 2006; Petersen et al., 2004). Harding et al. (2006) showed that *C. humeralis* is more predominant in continuously forested reaches and is rarely found in agricultural sites, a pattern also commonly observed in other New Zealand species (Winterbourn, 2007). Conversely, lower genetic differentiation among adjacent fragmented streams linked via continuous open pasture indicate that lateral dispersal for short distances (<1 km) may occur when suitable habitat is not found locally for *C. humeralis*. Thus, for this species, riparian forest cover likely facilitates connectivity among populations within the stream channel, but lateral dispersal across agricultural land may still be possible when suitable habitat is not found locally.

In *Z. confusus*, we found evidence for widespread gene flow among populations within the studied area, with extremely low pairwise F_{ST} (F_{ST}<0.006) indicating high gene flow even among the most distant populations of the neighbouring catchments (~11 km). Meanwhile, subtle genetic differentiation in the SNP dataset related to a combined effect of topography and Euclidean distance. Stoneflies are generally considered to be weak fliers and frequent long-distance dispersal is not expected (Briers et al., 2002; Sproul et al., 2014). However, between-stream dispersal at small spatial scales (<1 km) has been directly observed in mark-recapture studies, (Briers et al., 2004; Macneale et al., 2005b). Likewise, Dussex et al. (2016) showed high gene flow among populations of the winged stonefly *Zelandoperla decorata* in different streams separated by up to 10 km. Our findings suggest that *Z. confusus* populations are functionally well connected within the study area, and that local dispersal in this species occurs, even across fragmented landscapes.

In *H. fimbriata*, substantial connectivity was found among most of the studied area, although local patterns of genetic differentiation were highly correlated with landscape elements. Topography and land cover were significantly correlated with genetic distances in both datasets, and these combined variables gave the best prediction of genetic structure for the SNP dataset. However, for both datasets there was no marked genetic clustering that could be related to the landscape, with COI haplotypes or SNP genetic clusters closely related to one another but not shared in a geographically-correlated pattern. For example, only a portion of Stream C individuals show genetic relationships with Streams A and B; sections of steep sloping sides along stream C and suitable habitat provided by the local forested riparian zone could limit lateral dispersal in this species. Topographic effects, such as slope, have been to shown to act as significant barriers to dispersal in montane mayflies, resulting in adaptive divergence, especially at higher elevation sites (Polato et al., 2017). Previous COI analysis of *H. fimbriata* indicated high genetic differentiation only at broad spatial scales (~ 100 km; Smith & Smith, 2009), but the association of this species with forested and cool streams suggests that connectivity of populations at smaller spatial scales is more likely within continuous forested habitat. Therefore, *H. fimbriata* is likely to be vulnerable to further habitat fragmentation (Smith & Smith, 2009). Overland dispersal by *H. fimbriata* adults from Stream C may be reduced due to local habitat features. However, this does not appear to limit population connectivity more widely.

4.5.2 Discordance between mtDNA and SNP data in *C. humeralis* and *H. fimbriata*

We observed discordance in two of the studied species. For *C. humeralis*, this discrepancy was strongest in the isolated Karioi population (D1), for which mtDNA data showed highest differentiation of this from other populations, independent of the geographic distance among them. Conversely, in the SNP dataset, genetic differentiation of populations among mountain regions followed a pattern of IBD. In *H. fimbriata*, greater mtDNA versus SNP genetic differentiation was predominant between populations from sampling sites within Pirongia. Discordance between mitochondrial and nuclear genomes has been observed in many taxa and can have a variety of causes, including *Wolbachia* infection (Toews

& Brelsford, 2012), evolutionary processes such as introgression and incomplete lineage sorting (Buckley et al., 2006), and sex-mediated processes (Hurst & Jiggins, 2005).

Wolbachia is an intracellular bacterium typically found in the reproductive tissues of insects that can manipulate reproduction of its female host to enhance its own vertical transmission in the host from mother to offspring (Werren et al., 2008). Here, we found no evidence for *Wolbachia* in the BOLD trace files of the analysed sequences, nor in the preliminary taxon id tree, in which contaminated sequences are easily identified as a false outgroup. In addition, the accuracy of DNA barcoding is unlikely to be compromised by the presence of this bacteria (Smith et al., 2012). Therefore, it is unlikely that the mito-nuclear discordance observed in our study has resulted from *Wolbachia* contamination of the mtDNA dataset.

Distinguishing between introgression (the exchange of genes between related species through hybridisation) and incomplete lineage sorting (ILS — recent gene divergence, where distinct and isolated lineages have not progressed sufficiently to allow gene sorting) is challenging because they each generate similar genetic signatures (Buckley et al., 2006). However, these processes have been shown to influence divergence between populations for a range of taxa (Dincă et al., 2019; Pavlova et al., 2013; Wang et al., 2019). Recent isolation with ILS is a possible hypothesis for the discordant results found in the *C. humeralis* Mount Karioi/D1 population because anthropogenic deforestation occurring in the Waikato region since colonisation (~ 1,000 years) has resulted in isolation and fragmentation of the Pirongia Forest Park into two large disconnected forest patches (Mount Pirongia and Mount Karioi). As a result, the Karioi population may be reproductively isolated from populations in Pirongia. Coupled with the maternal inheritance and haploid nature of mtDNA, which makes lineage sorting progress faster compared to nuclear DNA (Avise, 1994), ILS may explain our finding of COI differentiation in the face of limited but ongoing nuclear gene flow for this species.

Another possibility is that late Pliocene volcanic events caused historic isolation followed by subsequent contact between the Karioi and Pirongia populations for *C. humeralis*, preserving the mtDNA divergence while resulting in limited genetic differentiation in the nuclear SNPs. Historic isolation with secondary contact has been discussed as a force underlying mito-nuclear

discordance in other aquatic insects, including the European stonefly *Dinocras cephalotes* (Elbrecht et al., 2014). Further investigation, including detailed phylogenetic and demographic analyses, is necessary to explore the underlying causes of mito-nuclear discordance in *C. humeralis*. Extending population comparisons to other streams in the Karioi and Pirongia regions, as well to intermediate locations between the two mountains, would also help elucidate the historical and contemporary processes underlying the observed genetic variation in this species.

Finally, discordance between markers can arise if there are differences in how selection acts on the mitochondrial genome as compared to the nuclear genome, or if there is a biased movement of either marker type driven by sex-biased dispersal (Toews & Brelsford, 2012). Despite the established assumption of neutrality, a number of studies have attributed mtDNA variation to natural selection (Camus et al., 2017; Galtier et al., 2009). If such selection varies geographically, then discordance between mtDNA and nuclear DNA can be expected. In addition, sex-biased asymmetries, such as male-biased dispersal, can promote gene flow for nuclear DNA in the absence of concordant movement of mtDNA (Prugnolle & de Meeus, 2002), resulting in greater structure and/or narrower geographic clines for mtDNA versus nuclear DNA. Male-biased dispersal has driven such patterns in several insect species (Bluher et al., 2020; Johnstone et al., 2012; López-Uribe et al., 2014). In aquatic insects, this pattern has been reported, mainly for mayflies, indicating males tend to disperse more laterally between streams and further along the stream channel than females (Macneale et al., 2005a; Petersen et al., 2004; Sabando et al., 2011; Schultheis & Hughes, 2005), while the contrary has been indicated for mayflies inhabiting ponds (Caudill, 2003). Our results for *H. fimbriata* indicated substantially higher genetic differentiation for COI than SNPs markers (as indicated by pairwise F_{ST} comparisons and AMOVA) that may be consistent with a pattern of male-biased dispersal — higher COI genetic differentiation was found between reaches along the stream channel, indicating that females may not move as far from their natal sites. However, both markers indicated significant genetic differentiation at this small spatial scale and we cannot currently rule out other explanations (above). To obtain a clear understanding of the relationship between genetic structure and dispersal capacity differences between males and

females in *H. fimbriata*, direct studies of movement, such as mark-recapture, would be necessary.

4.5.3 Implications for conservation management

As well as detecting differentiation over broad spatial scales, we detected species-specific influences of landscape elements on genetic structure at finer spatial scales, suggesting different responses to habitat fragmentation in terms of dispersal and population connectivity. This result is consistent with previous studies showing that attributes affecting dispersal can vary between species and landscapes (Phillipsen et al., 2015), which are therefore not equally affected by the natural or anthropogenically-driven structure of the stream habitat (Harding et al., 2006). While the three aquatic insects studied here are common in the North Island of New Zealand, each are likely affected differently by ongoing habitat fragmentation. In particular, the higher dispersal capacity and lower dispersal constraints of *Z. confusus* will likely give this species a considerable future advantage in colonising disturbed habitats and restored sections to maintain connectivity among proximate streams. For *C. humeralis* and *H. fimbriata*, present populations will benefit most from restoration of contiguous stream reaches via riparian planting to enhance short-distance dispersal within the stream channel. Restored sites will provide suitable environmental conditions (e.g. cooler temperatures and shade) for colonisation, thereby enhancing connectivity to source populations upstream. Restorarion planting should occur proximate (<1 km) to potential source populations in native forested sites to best assure the long-term viability of populations. These suggestions are likely broadly applicable to other common winged aquatic insect species sharing similar dispersal capabilities and habitat types.

4.6 References

Adamack, A., & Gruber, B. (2014). PopGenReport: Simplifying basic population genetic analyses in R. *Methods in Ecology and Evolution,* **5**, 384-387.

Alp, M., Keller, I., Westram, A. M., & Robinson, C. T. (2012). How river structure and biological traits influence gene flow: a population genetic study of two stream invertebrates with differing dispersal abilities. *Freshwater Biology,* **57**, 969-981.

Avise, J. C. (1994). *Molecular Markers, Natural History and Evolution*: Springer, Boston, MA.

Balkenhol, N., Schwartz, M. K., Inman, R. M., Copeland, J. P., Squires, J. S., Anderson, N. J., & Waits, L. P. (2020). Landscape genetics of wolverines (*Gulo gulo*): scale-dependent effects of bioclimatic, topographic, and anthropogenic variables. *Journal of Mammalogy,* **101**, 790-803.

Balkenhol, N., Waits, L., & Dezzani, R. (2009). Statistical approaches in landscape genetics: An evaluation of methods for linking landscape and genetic data. *Ecography,* **32**, 818-830.

Blakely, T. J., Harding, J. S., Mcintosh, A. R., & Winterbourn, M. J. (2006). Barriers to the recovery of aquatic insect communities in urban streams. *Freshwater Biology,* **51**, 1634-1645.

Blanchet, S., Rey, O., Etienne, R., Lek, S., & Loot, G. (2010). Species-specific responses to landscape fragmentation: implications for management strategies. *Evolutionary Applications,* **3**, 291-304.

Bluher, S. E., Miller, S. E., & Sheehan, M. J. (2020). Fine-scale population structure but limited genetic differentiation in a cooperatively breeding paper wasp. *Genome Biology and Evolution,* **12**, 701-714.

Bond, N. R., & Lake, P. S. (2003). Local habitat restoration in streams: Constraints on the effectiveness of restoration for stream biota. *Ecological Management and Restoration, 4*, 193-198.

Briers, R., Cariss, H., & Gee, J. H. R. (2002). Dispersal of adult stoneflies (Plecoptera) from upland streams draining catchments with different land use. *Archiv Fur Hydrobiologie, 155*, 627-644.

Briers, R. A., Gee, J. H. R., Cariss, H. M., & Geoghegan, R. (2004). Inter-population dispersal by adult stoneflies detected by stable isotope enrichment. *Freshwater Biology, 49*, 425-431.

Brockerhoff, E., Shaw, W., Hock, B., Kimberley, M., Paul, T., Quinn, J., & Pawson, S. (2010). Re-examination of recent loss of indigenous cover in New Zealand and the relative contributions of different land uses. *New Zealand Journal of Ecology, 32*, 115-126.

Buckley, T. R., Cordeiro, M., Marshall, D. C., & Simon, C. (2006). Differentiating between hypotheses of lineage sorting and introgression in New Zealand alpine cicadas (*Maoricicada* Dugdale). *Systematic Biology, 55*, 411-425.

Burrell, T., O'Brien, J., Graham, E., Simon, K., Harding, J., & McIntosh, A. (2014). Riparian shading mitigates stream eutrophication in agricultural catchments. *Freshwater Science, 33*, 73-84.

Camus, M. F., Wolff, J. N., Sgrò, C. M., & Dowling, D. K. (2017). Experimental support that natural selection has shaped the latitudinal distribution of mitochondrial haplotypes in australian *Drosophila melanogaster*. *Molecular Biology and Evolution, 34*, 2600-2612.

Carlson, P. E., McKie, B. G., Sandin, L., & Johnson, R. K. (2016). Strong land-use effects on the dispersal patterns of adult stream insects: implications for transfers of aquatic subsidies to terrestrial consumers. *Freshwater Biology, 61*, 848-861.

Christie, M. R., & Knowles, L. L. (2015). Habitat corridors facilitate genetic resilience irrespective of species dispersal abilities or population sizes. *Evolutionary Applications, 8*, 454-463.

Collier, K. J., & Smith, B. J. (1997). Dispersal of adult caddisflies (Trichoptera) into forests alongside three New Zealand streams. *Hydrobiologia, 361*, 53-65.

Collier, K. J., & Smith, B. J. (2000). Interactions of adult stoneflies (Plecoptera) with riparian zones I. Effects of air temperature and humidity on longevity. *Aquatic Insects, 22*, 275-284.

Cowley, D. R. (1978). Studies on the larvae of New Zealand Trichoptera. *New Zealand Journal of Zoology, 5*, 639-750.

Dala-Corte, R. B., Giam, X., Olden, J. D., Becker, F. G., Guimarães, T. d. F., & Melo, A. S. (2016). Revealing the pathways by which agricultural land-use affects stream fish communities in South Brazilian grasslands. *Freshwater Biology, 61*, 1921-1934.

Dincă, V., Lee, K. M., Vila, R., & Mutanen, M. (2019). The conundrum of species delimitation: a genomic perspective on a mitogenetically super-variable butterfly. *Proceedings of the Royal Society B: Biological Sciences, 286*, 20191311.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., …, Lautenbach, S. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography, 36*, 27-46.

Dussex, N., Chuah, A., & Waters, J. M. (2016). Genome-wide SNPs reveal fine-scale differentiation among wingless alpine stonefly populations and introgression between winged and wingless forms. *Evolution, 70*, 38-47.

Elbrecht, V., Feld, C. K., Gies, M., Hering, D., Sondermann, M., Tollrian, R., & Leese, F. (2014). Genetic diversity and dispersal potential of the stonefly

Dinocras cephalotes in a central European low mountain range. *Freshwater Science,* **33**, 181-191.

Ewers, R., Kliskey, A., Walker, S., Rutledge, D., Harding, J., & Didham, R. (2006). Past and future trajectories of forest loss in New Zealand. *Biological Conservation,* **133**, 312-325.

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources,* **10**, 564-567.

Fahrig, L. (2003). Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology, Evolution, and Systematics,* **34**, 487-515.

Finn, D. S., Blouin, M. S., & Lytle, D. A. (2007). Population genetic structure reveals terrestrial affinities for a headwater stream insect. *Freshwater Biology,* **52**, 1881-1897.

Finn, D. S., Theobald, D. M., Black, W. C., & Poff, N. L. (2006). Spatial population genetic structure and limited dispersal in a Rocky Mountain alpine stream insect. *Molecular Ecology,* **15**, 3553-3566.

Franckowiak, R. P., Panasci, M., Jarvis, K. J., Acuña-Rodriguez, I. S., Landguth, E. L., Fortin, M.-J., & Wagner, H. H. (2017). Model selection with multiple regression on distance matrices leads to incorrect inferences. *PLoS One,* **12**, e0175194.

Fuller, M., Doyle, M., & Strayer, D. (2015). Causes and consequences of habitat fragmentation in river networks. *Annals of the New York Academy of Sciences,* **1355**, 31-55.

Galic, N., Hengeveld, G. M., Van den Brink, P. J., Schmolke, A., Thorbek, P., Bruns, E., & Baveco, H. M. (2013). Persistence of aquatic insects across managed landscapes: effects of landscape permeability on re-colonization and population recovery. *PLoS One,* **8**, e54584.

Galtier, N., Nabholz, B., GlÉMin, S., & Hurst, G. D. D. (2009). Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular Ecology,* **18**, 4541-50.

GDAL/OGR contributors. (2020). *GDAL/OGR Geospatial Data Abstraction software Library*. Retrieved from https://gdal.org.

Geismar, J., Haase, P., Nowak, C., Sauer, J., & Pauls, S. U. (2015). Local population genetic structure of the montane caddisfly *Drusus discolor* is driven by overland dispersal and spatial scaling. *Freshwater Biology,* **60**, 209-221.

Geraci, C. J., Zhou, X., Morse, J. C., & Kjer, K. M. (2010). Defining the genus (Trichoptera:Hydropsychidae) based on DNA and morphological evidence. *Journal of the North American Benthological Society,* **29**, 918-933, 916.

Goslee, S., & Urban, D. (2007). The ECODIST package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software,* **22**, 1-19.

Harding, J., & Winterbourn, M. (1993). Life history and production of *Coloburiscus humeralis* (Ephemeroptera: Oligoneuriidae) in two South Island high-country streams, New Zealand. *New Zealand Journal of Marine and Freshwater Research,* **27**, 445-451.

Harding, J. S., Claassen, K., & Evers, N. (2006). Can forest fragments reset physical and water quality conditions in agricultural catchments and act as refugia for forest stream invertebrates? *Hydrobiologia,* **568**, 391-402.

Hladyz, S., Åbjörnsson, K., Giller, P., & Woodward, G. (2011). Impacts of an aggressive riparian invader on community structure and ecosystem functioning in stream food webs. *Journal of Applied Ecology,* **48**, 443-452.

Hughes, J. M. (2007). Constraints on recovery: using molecular methods to study connectivity of aquatic biota in rivers and streams. *Freshwater Biology,* **52**, 616-631.

Hughes, J. M., Mather, P. B., Sheldon, A. L., & Allendorf, F. W. (1999). Genetic structure of the stonefly, *Yoraperla brevis*, populations: the extent of gene flow among adjacent montane streams. *Freshwater Biology, ***41***, 63-72.

Hughes, J. M., Schmidt, D. J., & Finn, D. S. (2009). Genes in streams: using DNA to understand the movement of freshwater fauna and their riverine habitat. *BioScience, ***59***, 573-583.

Hurst, G. D. D., & Jiggins, F. M. (2005). Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proceedings of the Royal Society B: Biological Sciences, ***272***, 1525-1534.

Ivanova, N. V., Dewaard, J. R., & Hebert, P. D. N. (2006). An inexpensive, automation-friendly protocol for recovering high-quality DNA. *Molecular Ecology Notes, ***6***, 998-1002.

Johnstone, R. A., Cant, M. A., & Field, J. (2012). Sex-biased dispersal, haplodiploidy and the evolution of helping in social insects. *Proceedings of the Royal Society B: Biological Sciences, ***279***, 787-793.

Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics, ***24***, 1403-1405.

Jordahl, K., Bossche, V. J., Fleischmann, M., Wasserman, J., McBride, J., Gerard, J., …, Leblanc, F. (2020). *geopandas/geopandas*: v0.8.1 (Version v0.8.1). Zenodo.

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ, ***2***, e281-e281.

Kear, D. (Cartographer). (1960). Sheet 4-Hamilton. Geological map of New Zealand 1:250000

Keller, D., & Holderegger, R. (2013). Damselflies use different movement strategies for short- and long-distance dispersal. *Insect Conservation and Diversity, 6*, 590-597.

Keller, D., Van Strien, M. J., & Holderegger, R. (2012). Do landscape barriers affect functional connectivity of populations of an endangered damselfly? *Freshwater Biology, 57*, 1373-1384.

Kilian, A., Wenzl, P., Huttner, E., Carling, J., Xia, L., Blois, H., …, Uszynski, G. (2012). Diversity Arrays Technology: A generic genome profiling technology on open platforms. In F. Pompanon & A. Bonin (Eds.), *Data production and analysis in population genomics: methods and protocols* (pp. 67-89). Totowa, NJ: Humana Press.

Lawler, J., Aukema, J., Grant, J., Halpern, B., Kareiva, P., Nelson, C., …, Zaradic, P. (2013). Conservation science: a 20-year report card. *Frontiers in Ecology and the Environment, 4*, 473-480.

Lichstein, J. (2007). Multiple regression on distance matrices: A multivariate spatial analysis tool. *Plant Ecology, 188*, 117-131.

López-Uribe, M., Zamudio, K., Cardoso, C., & Danforth, B. (2014). Climate, physiological tolerance and sex-biased dispersal shape genetic structure of Neotropical orchid bees. *Molecular Ecology, 23*, 1874-1890.

Macneale, K., Peckarsky, B., & Likens, G. (2005a). Stable isotopes identify dispersal patterns of stonefly populations living along stream corridors. *Freshwater Biology, 50*, 1117-1130.

Macneale, K. H., Peckarsky, B. L., & Likens, G. E. (2005b). Stable isotopes identify dispersal patterns of stonefly populations living along stream corridors. *Freshwater Biology, 50*, 1117-1130.

Manel, S., & Holderegger, R. (2013). Ten years of landscape genetics. *Trends in Ecology and Evolution, 28*, 614-621.

Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution, 18*, 189-197.

McLellan, I. D. (1993). *Antarctoperlinae (Insecta: Plecoptera)*: Fauna of New Zealand, No. 27.

McLeod, E. O., Pittari, A., Brenna, M., & Briggs, R. (2020). *Geology of the Pirongia Volcano, Waikato: 1:30,000 Geological Map*: Geoscience Society of New Zealand.

McRae, B. (2006). Isolation by resistance. *Evolution, 60*, 1551-1561.

Murray, K., & Conner, M. M. (2009). Methods to quantify variable importance: implications for the analysis of noisy ecological data. *Ecology, 90*, 348-355.

Nei, M. (1972). Genetic distance between populations. *The American Naturalist, 106*, 283-292.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., …, Wagner, H. (2020). *vegan community ecology package version 2.5-7 November 2020*.

Paradis, E. (2010). PEGAS: an R package for population genetics with an integrated–modular approach. *Bioinformatics, 26*, 419-420.

Paradis, E., & Schliep, K. (2018). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics, 35*.

Parkyn, S. M., & Smith, B. J. (2011). Dispersal constraints for stream invertebrates: setting realistic timescales for biodiversity restoration. *Environmental Management, 48*, 602-614.

Pavlova, A., Amos, J. N., Joseph, L., Loynes, K., Austin, J. J., Keogh, J. S., …, Sunnucks, P. (2013). Perched at the mito-nuclear crossroads: divergent mitochondrial lineages correlate with environment in the face of ongoing nuclear gene flow in an Australian bird. *Evolution, 67*, 3412-3428.

Pavlova, A., Beheregaray, L. B., Coleman, R., Gilligan, D., Harrisson, K. A., Ingram, B. A., …, Sunnucks, P. (2017). Severe consequences of habitat fragmentation on genetic diversity of an endangered Australian freshwater fish: A call for assisted gene flow. *Evolutionary Applications,* **10**, 531-550.

Pembleton, L. W., Cogan, N. O. I., & Forster, J. W. (2013). StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources,* **13**, 946-952.

Petersen, I., Masters, Z., Hildrew, A. G., & Ormerod, S. J. (2004). Dispersal of adult aquatic insects in catchments of differing land use. *Journal of Applied Ecology,* **41**, 934-950.

Phillipsen, I. C., Kirk, E. H., Bogan, M. T., Mims, M. C., Olden, J. D., & Lytle, D. A. (2015). Dispersal ability and habitat requirements determine landscape-level genetic patterns in desert aquatic insects. *Molecular Ecology,* **24**, 54-69.

Phillipsen, I. C., & Lytle, D. A. (2013). Aquatic insects in a sea of desert: population genetic structure is shaped by limited dispersal in a naturally fragmented landscape. *Ecography,* **36**, 731-743.

Polato, N. R., Gray, M. M., Gill, B. B., Becker, C. G., Casner, K. L., Flecker, A. S., …, Zamudio, K. R. (2017). Genetic diversity and gene flow decline with elevation in montane mayflies. *Heredity,* **119**, 107-116.

Prugnolle, F., & de Meeus, T. (2002). Inferring sex-biased dispersal from population genetic tools: a review. *Heredity,* **88**, 161-165.

Raj, A., Stephens, M., & Pritchard, J. K. (2014). fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics,* **197**, 573.

Reid, A. J., Carlson, A. K., Creed, I. F., Eliason, E. J., Gell, P. A., Johnson, P. T. J., …, Cooke, S. J. (2019). Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biological Reviews,* **94**, 849-873.

Richardson, J. (2012). Divergent landscape effects on population connectivity in two co-occurring amphibian species. *Molecular Ecology,* **21**, 4437-4451.

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics,* **145**, 1219-1228.

Sabando, M. C., Vila, I., Penaloza, R., & Veliz, D. (2011). Contrasting population genetic structure of two widespread aquatic insects in the Chilean high-slope rivers. *Marine and Freshwater Research,* **62**, 1-10.

Schlaepfer, D. R., Braschler, B., Rusterholz, H.-P., & Baur, B. (2018). Genetic effects of anthropogenic habitat fragmentation on remnant animal and plant populations: a meta-analysis. *Ecosphere,* **9**, e02488.

Schultheis, A. S., & Hughes, J. M. (2005). Spatial patterns of genetic structure among populations of a stone-cased caddis (Trichoptera: Tasimiidae) in south-east Queensland, Australia. *Freshwater Biology,* **50**, 2002-2010.

Shah, V., & McRae, B. (2008). Circuitscape: A tool for landscape ecology. *Proceedings of the 7th Python in Science Conference*, 62-66.

Smith, M. A., Bertrand, C., Crosby, K., Eveleigh, E. S., Fernandez-Triana, J., Fisher, B. L., …, Zhou, X. (2012). Wolbachia and DNA barcoding insects: patterns, potential, and problems. *PLoS One,* **7**, e36514.

Smith, P. J., & Smith, B. J. (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**, 723-734.

Spear, S. F., Balkenhol, N., Fortin, M.-J., Mcrae, B. H., & Scribner, K. (2010). Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. *Molecular Ecology,* **19**, 3576-3591.

Sproul, J. S., Houston, D. D., Davis, N., Barrington, E., Oh, S. Y., Evans, R. P., & Shiozawa, D. K. (2014). Comparative phylogeography of codistributed

aquatic insects in western North America: insights into dispersal and regional patterns of genetic structure. *Freshwater Biology, 59*, 2051-2063.

Taylor, P. D., Fahrig, L., Henein, K., & Merriam, G. (1993). Connectivity is a vital element of landscape structure. *Oikos, 68*, 571-573.

Team, R. C. (2020). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria.*

Tischendorf, L., & Fahrig, L. (2000). How should we measure landscape connectivity? *Landscape Ecology, 15*, 633-641.

Toews, D. P. L., & Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology, 21*, 3907-3930.

Turner, M. B., Cronin, S. J., Smith, I. E., Stewart, R. B., & Neall, V. E. (2008). Eruption episodes and magma recharge events in andesitic systems: Mt Taranaki, New Zealand. *Journal of Volcanology and Geothermal Research, 177*, 1063-1076.

Van Rossum, G., & Drake Jr, F. L. (1995). *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam.

Wang, I. J. (2013). Examining the full effects of landscape heterogeneity on spatial genetic variation: a multiple matrix regression approach for quantifying geographic and ecological isolation. *Evolution, 67*, 3403-3411.

Wang, W., Wang, Y., Lei, F., Liu, Y., Wang, H., & Chen, J. (2019). Incomplete lineage sorting and introgression in the diversification of Chinese spot-billed ducks and mallards. *Current Zoology, 65*, 589-597.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution, 38*, 1358-1370.

Weisberg, S., & Fox, J. (2011). *An R Companion to Applied Regression*. Sage, Thousand Oaks CA.

Werren, J. H., Baldo, L., & Clark, M. E. (2008). Wolbachia: master manipulators of invertebrate biology. *Nature Reviews Microbiology, 6*, 741-751.

Wiens, J. A. (2002). Riverine landscapes: taking landscape ecology into the water. *Freshwater Biology, 47*, 501-515.

Wilcock, H. R., Bruford, M. W., Nichols, R. A., & Hildrew, A. G. (2007). Landscape, habitat characteristics and the genetic population structure of two caddisflies. *Freshwater Biology, 52*, 1907-1929.

Winterbourn, M., Gregson, K. L. D., & Dolphin, C. H. (2000). Guide to the Aquatic insects of New Zealand. *Bulletin of the Entomological Society of New Zealand, 13*.

Winterbourn, M. J. (2007). Stable isotope analysis identifies morphologically indistinguishable caddisflies. *New Zealand Journal of Marine and Freshwater Research, 41*, 401-404.

Yaegashi, S., Watanabe, K., Monaghan, M. T., & Omura, T. (2014). Fine-scale dispersal in a stream caddisfly inferred from spatial autocorrelation of microsatellite markers. *Freshwater Science, 33*, 172-180.

Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution, 1*, 3-14.

CHAPTER FIVE
THESIS CONCLUSION

5.1 Contextual overview

Aquatic insects are one of the most vulnerable invertebrate groups in freshwater ecosystems, as evidenced by global species declines due to the effect of multiple stressors, including anthropogenic habitat modification, pollution and climate change (Gage et al., 2009; Sánchez-Bayo & Wyckhuys, 2019; Stepanian et al., 2020). Dispersal is a key process supporting aquatic insect biodiversity in stream habitats, as it maintains connectivity among populations and shapes the dynamics of metapopulation systems (Bilton et al., 2001). Dispersal and connectivity are crucial for reproduction and long term persistence of organisms, and determine how genetic diversity arises and is maintained within species, which in turn influences the adaptability and resilience of populations to disturbance. The conservation of aquatic insect populations is vital for ecosystem health and function, as aquatic insect larvae play crucial roles in nutrient cycling and organic matter decomposition, and are a primary food source for higher trophic levels.

Decisions regarding the management and conservation of populations are often informed to some degree by population genetics (Hohenlohe et al., 2021). Population genetic data can provide estimates of basic features of wildlife populations, such as effective population size, inbreeding, demographic history ,and population structure, that are critical for conservation efforts. The basic premise is that individual populations may be threatened if there is high genetic differentiation among them. When genetically isolated populations are small, they are more susceptible to the effects of random genetic drift and inbreeding, which will reduce genetic diversity (Jameson et al., 2008). The application of population genetics to conservation and management of natural populations can, for example, enable definition of conservation units, genetic monitoring of population size, and assisted gene flow by translocation of individuals (Willi et al., 2022). A fundamental measure sought by decision makers is the degree of connectivity between populations, which helps conservation biologists to understand the amount of gene flow among populations, particularly those isolated in fragmented landscapes

(Crooks & Sanjayan, 2006; Hohenlohe et al., 2021; Lowe & Allendorf, 2010). Thus, characterising the distribution of genetic variation within and between populations, and evaluating the amount of gene flow among them, is key for generating information that can help wildlife managers and conservation practitioners make difficult management decisions (Walters & Schwartz, 2020).

5.2 Thesis synthesis and key findings

Population genetic data on the dispersal and connectivity of aquatic insects is being increasingly documented, but major gaps still remain in our knowledge and understanding of how anthropogenic habitat modification influences patterns and rates of gene flow (Chapter 2). A few studies have attempted to address the issue through landscape-genetic methods (Keller & Holderegger, 2013; Phillipsen et al., 2015) but they are still very limited (Queiroga et al., 2019), despite the growing application of next generation sequencing technologies, which have the potential to provide fine-scale genetic data with different implications for conservation actions (Hohenlohe et al., 2021). Additionally, stream insects are deeply underrepresented in genomics research (Hotaling et al., 2020).

In order to enhance our understanding of population connectivity and dispersal patterns in stream insects, I combined mitochondrial (mt) cytochrome c oxidase subunit I (COI) and genome-wide single nucleotide polymorphism (SNP) markers to assess spatial patterns of population structure at multiple spatial scales (Chapter 3 and 4). This include included a fine-scale landscape genetic approach that examined the effects of landscape elements on gene flow and dispersal over a small fragmented landscape (Chapter 4). While RADseq (restriction associated digest sequencing) for SNP discovery has recently been used in aquatic insect studies to assess fine-scale population structure and delimit species (Dussex et al., 2016; Hotaling et al., 2019; Polato et al., 2017; Salokannel et al., 2021), to my knowledge, no published studies have focused either on applying SNP data or combining it with mitochondrial COI markers in a landscape genetics context to explore functional connectivity in stream insect populations. Below, I provide a synthesis of the key findings of each research chapter, as well as implications for conservation management and recommendations for future research, in hope that

these insights will be a useful aid for future conservation and restoration efforts of stream insects.

Many different molecular markers have been used to detect variation among aquatic insect populations as an indirect measure of dispersal in order to better understand connectivity among habitats. However, recent (i.e. over the last 20 years) developments in molecular genetics have paved the way for dramatic advances in DNA sequencing technologies, resulting in a new generation of molecular markers for use in understanding the distribution and differences of natural populations (Futschik & Schlötterer, 2010). In **Chapter 2**, I assessed the use over time of the different markers by conducting a global literature review on their use in studies of aquatic insect dispersal and connectivity among lotic habitats. I provided an overview of the most commonly-used traditional and contemporary molecular markers, their main features, and their contributions and limitations to our understanding of dispersal. The knowledge provided could be helpful for those choosing a molecular marker in future studies. The review also examined the spatial scale at which markers were applied. I analysed genetic data from 77 published studies to identify generalities in dispersal patterns across multiple spatial scales for a range of aquatic insect taxa. Most of these studies found limited genetic differentiation among populations across a range of geographic scales, suggesting high adult dispersal potential across catchments. However, based on pairwise genetic differentiation results, I suggest that connectivity within the catchment often plays an important role in maintaining high levels of gene flow among populations. Overall, I found that traditional markers (allozymes and mtDNA) are still those dominantly used, and little is known about the influence on specific landscape features on shaping population connectivity among habitats.

The mt COI gene is a traditional marker for standard DNA barcoding (Hebert et al., 2003). It is frequently used as the marker of choice for species-level discrimination, because no other genetic region can be found in taxonomically verified databases (e.g. BOLD), with sequences covering so many taxa (Ratnasingham & Hebert, 2007). Combined with its cost-effectiveness, it is not surprising that COI markers have also long been used in population genetic studies attempting to characterise population structure (key for inferring gene flow and indirectly, dispersal; Hughes et al., 2011). However, newer sequencing approaches, such as RADseq, provide fine-scale SNP data at a genome-wide scale and may

enhance our knowledge of natural populations (Davey & Blaxter, 2010; Peterson et al., 2012), providing increased statistical power and resolution, increased efficiency, and cost-effectiveness (Sunde et al., 2020; Walters & Schwartz, 2020). Thus, they have high potential for assessments of contemporary patterns of genetic variation and reductions in connectivity due to habitat fragmentation (Brauer & Beheregaray, 2020). **Chapter 3** addressed this by comparing the resolution power of COI and SNP data in estimating genetic differentiation among populations within the main study area, Mount Pirongia. For each of three aquatic insect target species, individuals that were successfully genotyped for COI markers were then genotyped with RADseq for SNP discovery. The key finding of this work was that estimated genetic differentiation and population structure for both markers provided generally comparable results. A general lack of strong population structure was found within each species, while fine-scale genetic differentiation among some populations was detected for both markers. Combined, these results suggested substantial connectivity of the three analysed species within the study area, but finer-scale genetic differentiation between populations for *H. fimbriata* suggested that gene flow, and hence dispersal, may be more limited for this species relative to the others. However, these results are yet not conclusive as small sample sizes (<10 individuals for most populations), low success of COI sequencing, and limited data quality for the SNP datasets may have collectively compromised the power to uncover genetic structure, making interpretation of the observed patterns of population structure diffcult. Nevertheless, the results broadly indicate that either COI or SNP markers can be used to provide suitable initial estimates of fine-scale population genetic differentiation in stream insects.

Chapter 4 extended the research of Chapter 3 by increasing the sample size and geographic coverage of sequenced individuals for a detailed analysis of the functional connectivity and dispersal patterns of stream insects at multiple spatial scales. Notably, extending the size and the quality of final dataset resulted in well-resolved patterns of populations structure, especially when visualising structure based on SNP data. For the three studied species, clear spatial genetic structure was only observed at larger spatial scales (among mountain regions separated by ~ 30 and 170 km), whereas most gene flow occurred locally (up to 11 km). This pattern—substantial local gene flow with isolation by distance found only at larger spatial scales—is expected in strong-flying insect species and is commonly found

across a range of taxa (Hughes, 2007; Hughes et al., 2009). Further analysis showed the influence of landscape features (land cover type and topography) on the observed patterns of genetic differentiation at small spatial scales. Thus, the key finding of this chapter was that dispersal and gene flow within the stream channel are likely constrained by a lack of riparian forest, whereas overland dispersal is not constrained by open pasture. In the latter case, populations maintain substantial connectivity, particularly between adjacent streams separated by ~ 1 km. However, while long-distance dispersal crossing agricultural land occurs and is important to connect more isolated populations, short-distance dispersal within the stream channel and along continuous forested habitats is more frequent and fundamental for colonisation and population connectivity (Collier & Smith, 1997; Keller & Holderegger, 2013; Petersen et al., 2004). Combining my results with previously reported studies, I suggest that population connectivity, and hence dispersal, is preferred within the stream channel covered with forested riparian zone; but overland dispersal between proximate streams is more common across pastoral land when suitable habitat cannot be found locally due the lack of riparian cover. Another key finding of this chapter was discordance between the COI and SNP datasets, suggesting potential recent isolation with incomplete lineage sorting in the Karioi population for *C. humeralis* and potentially higher male dispersal potential in *H. fimbriata*. Thus, the application of both marker types yielded inferences of spatial population structure and connectivity and associated evolutionary processes, providing a more comprehensive understanding of dispersal patterns that could be identified by using only a single marker type.

Collectively, this thesis elucidated patterns of population connectivity and dispersal potential that will provide valuable knowledge for conservation efforts aimed at enhancing restoration of stream insect biodiversity. Overall, genetically distinct groups were only found for populations located in different mountain regions and therefore separated by larger geographic distances. This supports the conclusion that populations of the studied species in the Mount Pirongia study area remain highly connected when separated by up to ~ 11 km. However, I also found species-specific patterns of fine-scale genetic differentiation, and of the influence of landscape features — particularly the degree of forest fragmentation — on population structure.

5.3 Implications for conservation management

Knowing how populations are functionally connected across a fragmented landscape is important for our understanding of population dynamics in altered stream systems, and has critical implications for conservation and restoration. In particular, understanding in-stream and overland dispersal, and how these affect the gene flow and genetic diversity of species, is important for successful implementation of stream restoration measures (Haase et al., 2013; Parkyn & Smith, 2011). Unfortunately, our understanding of terrestrial stages and overland dispersal in the face of growing anthropogenic habitat modification remains limited and requires more research effort (Lowe & Allendorf, 2010; Queiroga et al., 2019; Smith et al., 2009), to allow the incorporation of gene flow patterns into biodiversity and conservation-oriented applications (Chaput-Bardy et al., 2008; Hohenlohe et al., 2021).

My results demonstrate that, even at small spatial scales, landscape features can be more important in determining population differentiation than Euclidean distance. Thus, landscape genetics provides a valuable tool to better understand the distribution and variation of populations at small spatial scales, particularly in cases where their geographical delineation is not clear. The lack of pairwise differentiation and genetic structure among populations of the stonefly *Z. confusus* suggests that this species has the highest dispersal capacity across a small fragmented landscape dominated by pasture land. Such dispersal capacity may give *Z. confusus* an advantage in colonising disturbed habitats and restored sections over short distances. Therefore, *Z. confusus* is likely to be a valuable and reliable indicator species for monitoring the success of restoration if found in restored stream sections from which it was absent prior to restoration. Based on the results found for the mayfly *C. humeralis* and the caddisfly *H. fimbriata*, habitat maintenance and connection of present populations via restoration of contiguous streams through riparian planting will enhance short-distance dispersal within the stream channel.

My research indicates that restoration measures (e.g. riparian planting) should occur proximate (<1 km) to potential source populations in native forested sites and should continue along the stream channel to promote gene flow and assure the long-term viability of populations. Streams protected by riparian forest provide

suitable microhabitat conditions (e.g. cooler temperatures, increased humidity, reduced wind and shade) for colonisation (Turunen et al., 2021). Therefore enhancing connectivity within the stream channel is essential. I highlight that overland dispersal across pastoral land can maintain substantial population connectivity, especially between adjacent streams - in such cases dispersal may be more common between sites where suitable forested habitat cannot be found locally. Thus, enhancing riparian planting of proximate streams in a small fragmented landscape should help to promote more suitable and healthier local habitats for colonisation and persistence of populations, which are important considerations for restoration and recolonisation of insects in New Zealand streams (Harding et al., 2006; Parkyn & Smith, 2011). Broadly, these recommendations may apply to other common winged aquatic insect species sharing similar habitat types.

5.4 Future research

High resolution genetic markers offer the potential to infer contemporary gene flow and reflect historical dispersal events. SNP-based RADseq, as well as whole genome sequencing methods, have fueled studies in ecological and conservation genomics, and today they represent powerful tools for assessing connectivity questions and providing implications for conservation management (Andrews et al., 2016; Fuentes-Pardo & Ruzzante 2017). Nevertheless, it is important to highlight that these indirect genetic methods only assess dispersal that results in mating and gene flow. In contrast, direct methods, for instance trapping and stable isotope labelling studies, assess actual movement events. Both methods are complementary (Keller & Holderegger, 2013; Lowe & Allendorf, 2010), and therefore, future population genetic studies focusing on present-day dispersal patterns should integrate assessment methods that directly track contemporary movements when possible in order to enhance the understanding of population connectivity and dispersal patterns in recently or ongoing fragmented landscapes. Although the use of 'big' data is increasing, mitochondiral DNA still has an important role to play and the continuously growing barcoding databases of COI sequences is particularly valuable. Future work could consider optimisation of the COI sequencing approach to mitigate the high failure rate seen for some species.

Genetic patterns of population structure can result from a series of complex interactions between spatial and environmental heterogeneity, and ecological and evolutionary processes (Balkenhol et al., 2015; Polato et al., 2017). This thesis was mainly focused on the role of physical landscape features (forest versus pasture land cover) in shaping genetic differentiation to better understand the effect of local habitat fragmentation in within- and among-stream dispersal. However, when extrapolating results from basin-wide to metapopulations, increasing spatial scale and replication, i.e. the inclusion of multiple study areas with similar basic landscape elements, is an ideal consideration for future landscape genetics studies — particularly those aiming to reliably assess/distinguish isolation by distance and isolation by landscape resistance hypotheses, and to identify corridors and barriers of connectivity (Short Bull et al., 2011; Hand et al., 2015). Furthermore, landscape-level replication of genetic research in stream insects may improve understandings of species' habitat requirements/preferences for gene flow, and how continuous forest cover vs open land interact with it. Additionally, the presence or absence of forest cover in stream habitats directly interacts with environmental factors (e.g. temperature, precipitation, wind), which also have consequences on the dispersal of aquatic insects (Carlson et al., 2016; Collier & Smith, 2000; Turunen et al., 2021). Thus, it is possible that environmental factors, especially the variation of microclimate (e.g. mean temperature and seasonality, wind direction and speed) among sites that are protected and not protected by riparian forest, may have influenced the patterns of gene flow identified here. Unfortunately, weather and climatic data is not available at a resolution that would be compatible with the study area and landscape data, and climate data collection on site was out of the scope of this project. I suggest future landscape genetic studies should incorporate both local environmental variables and landscape resistance matrices where possible to better assess the ecological drivers of genetic differentiation (Davis et al., 2018; Queiroga et al., 2019) especially for species that are sensitive to environmental change (Bauernfeind & Moog, 2000; Polato et al., 2017).

The multi-species approach of this thesis allowed the examination of population structure among common aquatic insects with different potential flight ability. Future studies aimed at providing suggestions for conservation management could further benefit by investigating co-distributed species with different habitat

requirements, life histories, and dispersal abilities to better understand how species' distinct biological and ecological traits influence population structure.

5.5 References

Andrews, K.R., Good, J.M., Miller, M.R., Luikart, G., Hohenlohe, P.A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, **17**, 81-92.

Balkenhol, N., Cushman, S., Waits, L., & Storfer, A. (2015). Current status, future opportunities, and remaining challenges in landscape genetics. In (pp. 247-256).

Bauernfeind, E., & Moog, O. (2000). Mayflies (Insecta: Ephemeroptera) and the assessment of ecological integrity: A methodological approach. *Hydrobiologia,* **422-423**, 71-83.

Bilton, D. T., Freeland, J. R., & Okamura, B. (2001). Dispersal in freshwater invertebrates. *Annual Review of Ecology and Systematics,* **32**, 26.

Brauer, C. J., & Beheregaray, L. B. (2020). Recent and rapid anthropogenic habitat fragmentation increases extinction risk for freshwater biodiversity. *Evolutionary Applications,* **13**, 2857-2869.

Carlson, P. E., McKie, B. G., Sandin, L., & Johnson, R. K. (2016). Strong land-use effects on the dispersal patterns of adult stream insects: implications for transfers of aquatic subsidies to terrestrial consumers. *Freshwater Biology,* **61**, 848-861.

Chaput-Bardy, A., Lemaire, C., Picard, D., & Secondi, J. (2008). In-stream and overland dispersal across a river network influences gene flow in a freshwater insect, *Calopteryx splendens*. *Molecular Ecology,* **17**, 3496-3505.

Collier, K. J., & Smith, B. J. (1997). Dispersal of adult caddisflies (Trichoptera) into forests alongside three New Zealand streams. *Hydrobiologia,* **361**, 53-65.

Collier, K. J., & Smith, B. J. (2000). Interactions of adult stoneflies (Plecoptera) with riparian zones I. Effects of air temperature and humidity on longevity. *Aquatic Insects,* **22**, 275-284.

Crooks, K., & Sanjayan, M. (2006). *Connectivity Conservation* (K. R. Crooks & M. Sanjayan Eds.). Cambridge: Cambridge University Press.

Davey, J. W., & Blaxter, M. L. (2010). RADSeq: next-generation population genetics. *Briefings in Functional Genomics,* **9**, 416-423.

Davis, C. D., Epps, C. W., Flitcroft, R. L., & Banks, M. A. (2018). Refining and defining riverscape genetics: How rivers influence population genetic structure. *WIREs Water,* **5**, e1269.

Dussex, N., Chuah, A., & Waters, J. M. (2016). Genome-wide SNPs reveal fine-scale differentiation among wingless alpine stonefly populations and introgression between winged and wingless forms. *Evolution,* **70**, 38-47.

Fuentes-Pardo, A.P., Ruzzante, D.E. (2017). Whole-genome sequencing approaches for conservation biology: Advantages, limitations and practical recommendations. *Molecular Ecology*, **26**, 5369– 5406.

Futschik, A., & Schlötterer, C. (2010). The next generation of molecular markers from massively parallel sequencing of pooled DNA samples. *Genetics,* **186**, 207-218.

Gage, M., Spivak, A., & Paradise, C. (2009). Effects of land use and disturbance on benthic insects in headwater streams draining small watersheds north of Charlotte, NC. *Southeastern Naturalist,* **3**, 345-358.

Haase, P., Hering, D., Jähnig, S. C., Lorenz, A. W., & Sundermann, A. (2013). The impact of hydromorphological restoration on river ecological status: a

comparison of fish, benthic invertebrates, and macrophytes. *Hydrobiologia,* **704**, 475-488.

Hand, B.K., Muhlfeld, C.C., Wade, A.A., Kovach, R.P., Whited, D.C., Narum, S.R., Matala, A.P., Ackerman, M.W., Garner, B.A., Kimball, J.S., Stanford, J.A. and Luikart, G. (2016), Climate variables explain neutral and adaptive variation within salmonid metapopulations: the importance of replication in landscape genetics. *Molecular Ecology*, 25, 689-705.

Harding, J. S., Claassen, K., & Evers, N. (2006). Can forest fragments reset physical and water quality conditions in agricultural catchments and act as refugia for forest stream invertebrates? *Hydrobiologia,* **568**, 391-402.

Hebert, P. D. N., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological sciences,* **270**, 313-321.

Hohenlohe, P. A., Funk, W. C., & Rajora, O. P. (2021). Population genomics for wildlife conservation and management. *Molecular Ecology,* **30**, 62-82.

Hotaling, S., Giersch, J. J., Finn, D. S., Tronstad, L. M., Jordan, S., Serpa, L. E., …, Weisrock, D. W. (2019). Congruent population genetic structure but differing depths of divergence for three alpine stoneflies with similar ecology and geographic distributions. *Freshwater Biology,* **64**, 335-347.

Hotaling, S., Kelley, J. L., & Frandsen, P. B. (2020). Aquatic insects are dramatically underrepresented in genomic research. *Insects,* **11**, 601.

Hughes, J. M. (2007). Constraints on recovery: using molecular methods to study connectivity of aquatic biota in rivers and streams. *Freshwater Biology,* **52**, 616-631.

Hughes, J. M., Huey, J. A., McLean, A. J., & Baggiano, O. (2011). Aquatic insects in eastern Australia: a window on ecology and evolution of dispersal in streams. *Insects,* **2**, 447-461.

Hughes, J. M., Schmidt, D. J., & Finn, D. S. (2009). Genes in streams: using DNA to understand the movement of freshwater fauna and their riverine habitat. *BioScience,* **59**, 10.

Jameson, I. G., Grueber, C., Waters, J., & Gleeson, D. (2008). Managing genetic diversity in threatened populations: A New Zealand perspective. *New Zealand Journal of Ecology,* **32**, 130-137.

Keller, D., & Holderegger, R. (2013). Damselflies use different movement strategies for short- and long-distance dispersal. *Insect Conservation and Diversity,* **6**, 590-597.

Lowe, W., & Allendorf, F. (2010). What can genetics tell us about connectivity? *Molecular Ecology,* **19**, 3038-3051.

Parkyn, S. M., & Smith, B. J. (2011). Dispersal constraints for stream invertebrates: setting realistic timescales for biodiversity restoration. *Environmental Management,* **48**, 602-614.

Petersen, I., Masters, Z., Hildrew, A. G., & Ormerod, S. J. (2004). Dispersal of adult aquatic insects in catchments of differing land use. *Journal of Applied Ecology,* **41**, 934-950.

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double Digest RADseq: An inexpensive method for de novo snp discovery and genotyping in model and non-model species. *PLoS One,* **7**, e37135.

Phillipsen, I. C., Kirk, E. H., Bogan, M. T., Mims, M. C., Olden, J. D., & Lytle, D. A. (2015). Dispersal ability and habitat requirements determine landscape-level genetic patterns in desert aquatic insects. *Molecular Ecology,* **24**, 54-69.

Polato, N. R., Gray, M. M., Gill, B. B., Becker, C. G., Casner, K. L., Flecker, A. S., …, Zamudio, K. R. (2017). Genetic diversity and gene flow decline with elevation in montane mayflies. *Heredity,* **119**, 107-116.

Queiroga, D., Fernandes Moura, R., & Ware, J. (2019). Genetic connectivity in conservation of freshwater insects. In (pp. 381-399).

Ratnasingham, S., & Hebert, P. D. N. (2007). BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). *Molecular Ecology Notes, 7*, 355-364.

Salokannel, J., Lee, K. M., Rinne, A., & Mutanen, M. (2021). ddRAD sequencing sheds light on low interspecific and high intraspecific mtDNA divergences in two groups of caddisflies. *Insect Systematics and Diversity, 5*.

Sánchez-Bayo, F., & Wyckhuys, K. A. G. (2019). Worldwide decline of the entomofauna: A review of its drivers. *Biological Conservation, 232*, 8-27.

Short Bull, R.A., Cushman, S.A., Mace, R., Chilton, T., Kendall, K.C., Landguth, E.L., Schwartz, M., Mckelvey, K., Allendorf, F.W. And Luikart, G. (2011). Why replication is important in landscape genetics: American black bear in the Rocky Mountains. *Molecular Ecology, 20*, 1092-1107.

Smith, R. F., Alexander, L. C., & Lamp, W. O. (2009). Dispersal by terrestrial stages of stream insects in urban watersheds: a synthesis of current knowledge. *Journal of the North American Benthological Society, 28*, 1022-1037.

Stepanian, P. M., Entrekin, S. A., Wainwright, C. E., Mirkovic, D., Tank, J. L., & Kelly, J. F. (2020). Declines in an abundant aquatic insect, the burrowing mayfly, across major North American waterways. *Proceedings of the National Academy of Sciences, 6*, 2987-2992.

Sunde, J., Yıldırım, Y., Tibblin, P., & Forsman, A. (2020). Comparing the performance of microsatellites and RADseq in population genetic studies: analysis of data for pike (*Esox lucius*) and a synthesis of previous studies. *Frontiers in Genetics, 11*.

Turunen, J., Elbrecht, V., Steinke, D., & Aroviita, J. (2021). Riparian forests can mitigate warming and ecological degradation of agricultural headwater streams. *Freshwater Biology, 66*, 785-798.

Walters, A. D., & Schwartz, M. K. (2020). Population genomics for the management of wild vertebrate populations. In H. P. A. & R. O. P. (Eds.), *Population genomics* (pp. p. 419-436): Springer.

Willi, Y., Kristensen, T. N., Sgrò, C. M., Weeks, A. R., Ørsted, M., & Hoffmann, A. A. (2022). Conservation genetics as a management tool: The five best-supported paradigms to assist the management of threatened species. *Proceedings of the National Academy of Sciences, 119*, e2105076119.

APPENDICES

A1: Co-authorship Forms

A1.1 Co-authorship form for Chapter 2

Co-Authorship Form

THE UNIVERSITY OF WAIKATO
Te Whare Wānanga o Waikato

This form is to accompany the submission of any PhD that contains research reported in published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in your appendices for all the copies of your thesis submitted for examination and library deposit (including digital deposit).

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

co-authored work from Chapter Two
Paper Title: Assessing dispersal and connectivity of lotic insect populations using genetic markers
Journal: Freshwater Biology. Initial submission on 14/07/2021, first revised on 9/12/2021 and currently under Revision Two by the authors.

Nature of contribution by PhD candidate	Designed study with input from supervisors; conducted literature search, article reviews, analysis and discussion
Extent of contribution by PhD candidate (%)	80%

CO-AUTHORS

Name	Nature of Contribution
Angela McGaughran	contributed critically to drafts, revised and gave final approval for publication
S. Elizabeth Graham	contributed critically to drafts, revised and gave final approval for publication
Ian D. Hogg	contributed critically to drafts, revised and gave final approval for publication
Kevin J. Collier	contributed critically to drafts, revised and gave final approval for publication

Certification by Co-Authors

The undersigned hereby certify that:
❖ the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and

Name	Signature	Date
Vanessa De Araujo Barbosa		23/02/2022
Angela McGaughran		24/02/2022
Ian Hogg		24/02/2022
Elizabeth Graham		25/02/2022
Kevin J. Collier		24/02/2022

A1.2 Co-authorship form for Chapter 3

Co-Authorship Form

This form is to accompany the submission of any PhD that contains research reported in published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in your appendices for all the copies of your thesis submitted for examination and library deposit (including digital deposit).

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

co-authored work from Chapter Three
Paper Title: Assessing population genetic structure of three New Zealand stream insects using mitochondrial and nuclear DNA markers
Journal: Genome. Initial submission on 21/02/2022

Nature of contribution by PhD candidate	Study design, aquatic insect collection and identification, data analysis, manuscript writing and revision
Extent of contribution by PhD candidate (%)	60%

CO-AUTHORS

Name	Nature of Contribution
S. Elizabeth Graham	designed study, collected insects, contributed critically to drafts, revised and gave final approval for publication
Brian Smith	designed study, collected and identified insects, contributed critically to drafts, revised and gave final approval for publication
Ian D. Hogg	designed study, contributed critically to drafts, revised and gave final approval for publication
Angela McGaughran	Analysed data, contributed critically to drafts, revised and gave final approval for publication

Certification by Co-Authors

The undersigned hereby certify that:
* the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and

Name	Signature	Date
Vanessa De Araujo Barbosa		23/02/2022
Elizabeth Graham		23/02/2022
Angela McGaughran		24/02/2022
Brian Smith		24/02/2022
Ian Hogg		24/02/2022

A2: Supplementary material for Chapter 2

Table A.2.1 Keywords and search terms used in the systematic review

Dispersal	Stream	Aquatic insect	Genetic markers
Dispersal	Freshwater	"Stream insect*"	Genetic
OR	OR	OR	OR
Movement	Stream*	"Aquatic insect*"	"Population genetic*"
OR	OR	OR	OR
Migration	River*	Mayfl*	"Population structure"
	OR	OR	OR
	Lotic	Stonefl*	"Molecular markers"
			OR
		OR	"Genetic marker"
		Caddisfl*	OR
		OR	Allozyme*
		Dobsonfl*	OR
		OR	AFLP*
		Beetl*	OR
		OR	SNP*
		Dragonfl*	OR
		OR	Microsatellite*
		Damselfl*	OR
			mtDNA
			OR
			COI
			OR
			"mitochondrial DNA"
			OR
			"nuclear DNA"
			OR
			"DNA

Table A.2.2 Summary of species and type of genetic marker used in the analysed articles.

Order, common name	Species name	Allozymes	mtDNA	Microsatellites	AFLPs	SNPs	Other marker	References
Trichoptera, caddisflies	*Agapetus sp., Lectrides varians*	No	Yes	No	No	No	No	Chester et al. (2015)
	Allogamus auricollis	Yes	No	No	No	No	No	Monaghan et al. (2002)
	Allogamus uncatus	No	No	Yes	No	No	No	Shama et al. (2011)
	Annitella esparraguera, Annitella iglesiasi	No	Yes	No	No	No	Yes	Murria et al. (2019)
	Cheumatopsyche sp. AV1	No	Yes	No	No	No	No	Baker et al. (2003)
	Dicosmoecus gilvipes	No	Yes	No	No	No	Yes	Peterson et al. (2017)
	Drusus discolor	No	No	Yes	No	No	No	Geismar et al. (2015)
	Gumaga griseola, Helicopsyche mexicana	No	No	No	Yes	No	No	Miller et al. (2002)
	Glossosoma conformis	No	No	No	No	Yes	No	Weigand et al. (2018)
	Hesperophylax designatus, Lepidostoma ojanum	No	Yes	No	No	No	No	Myers et al. (2001)
	Helicopsyche borealis	Yes	No	No	No	No	No	Jackson & Resh (1992)
	Hydropsyche siltalai	No	Yes	No	No	No	No	Murria et al. (2010)
	Lectrides varians	No	Yes	No	No	No	No	Wickson et al. (2014)
	Mesophylax aspersus	Yes	No	No	No	No	No	Kelly et al. (2001, 2002)
	Neothremma alicia	No	Yes	No	No	No	No	Rader et al. (2019)
	Orthopsyche fimbriata	Yes	Yes	No	No	No	No	Smith & Collier (2001), Smith et al. (2006b), Smith & Smith (2009)
	Plectrocnemia conspersa	Yes	No	Yes	No	No	No	Wilcock et al. (2001, 2003, 2007)
	Polycentropus flavomaculatus	No	No	Yes	No	No	No	Wilcock et al. (2007)
	Rhyacophila minor	No	Yes	No	No	No	No	Addison et al. (2015)
	Rhyacophila pubescens	No	Yes	No	Yes	No	No	Engelhardt et al. (2011)
	Smicridea annulicornis						No	Sabando et al. (2011)

	Species	1	2	3	4	5	6	Reference
	Sericostoma flavicorne, S personatum	No	Yes	No	No	No	No	Weigand et al. (2017)
	Sericostoma spp.	No	Yes	No	No	Yes	No	Weigand et al. (2017)
	Stenopsyche marmorata	No	No	Yes	No	No	No	Yaegashi et al. (2014)
	Tasiagma ciliata	Yes	No	No	No	No	No	Hughes et al. (1998)
	Tasimia palpata	No	Yes	Yes	No	No	No	Schultheis & Hughes (2005), Schultheis et al. (2008)
	Wormaldia tagananana	Yes	No	No	No	No	No	Kelly et al. (2002)
Ephemeroptera, mayflies	*Acanthophlebia cruentata*	Yes	Yes	No	No	No	No	Smith & Collier (2001), Smith et al. (2006a)
	Afroptilum sudafricanum Demoreptus capensis Demoreptus natalensis	No	Yes	No	No	No	No	Taylor et al. (2020)
	Andesiops torrens	No	Yes	No	Yes	No	No	Sabando et al. (2011)
	Atalophlebia sp.	No	Yes	No	No	No	No	Baggiano et al. (2011)
	Baetis alpinus	Yes	No	No	No	No	No	Monaghan et al. (2001, 2002)
	Baetis bicaudatus	Yes	Yes	No	No	Yes	No	Hughes, Mather, et al. (2003), Polato et al. (2017)
	Baetis tricaudatus	No	No	No	No	Yes	No	Polato et al. (2017)
	Baetis rhodani	No	No	Yes	No	No	No	Rebora et al. (2005), Alp et al. (2012)
	Baetis sp.	Yes	No	No	No	No	No	Bunn & Hughes (1997), Schmidt et al. (2005)
	Bungona narilla	Yes	Yes	No	No	No	No	Schmidt et al. (1995), Hughes, Hillyer, et al. (2003), McLean et al. (2008)
	Bungona sp	Yes	No	No	No	No	No	Hughes et al. (2000)
	Cloeon sp., Tasmanocoenis sp.	No	Yes	No	No	No	No	Razeng et al. (2017)
	Coloburiscus humeralis	Yes	No	No	No	No	No	Hogg et al. (2002)
	Dipteromimus tipuliformis	No	Yes	No	No	No	Yes	Takenaka et al. (2019)
	Drunella grandis	No	Yes	No	No	No	No	Sproul et al. (2014)

Taxon	Species							Reference
	Ephemerella subvaria, Ephemerella aurivillii, Ephemerella septentrionalis, Eurylophella funeralis, Eurylophella verisimilis	Yes	No	No	No	No	No	Sweeney et al. (1987)
	Ephoron shigae	No	Yes	No	No	No	No	Sekine et al. (2015)
	Fallceon quilleri	No	Yes	No	No	No	No	Zickovich & Bohonak (2007)
	Isonychia japonica	No	Yes	No	No	No	No	Saito & Tojo (2016)
	Koorrnonga AV3, Nousia AV1	No	Yes	No	No	No	No	Chester et al. (2015)
	Rhithrogena loyolaea	Yes	No	No	No	No	No	Monaghan et al. (2002)
	Siphlonisca aerodromi	No	No	Yes	No	No	No	Gibbs et al. (1998)
	Ulmerophlebia sp. AV2	No	Yes	No	No	No	No	Young et al. (2013)
Plecoptera, stoneflies	*Acroneuria frisoni*	No	Yes	No	No	No	No	Pessino et al. (2014)
	Allocapnia recta, Leuctra tenuis	No	Yes	No	No	No	No	Yasick et al. (2007)
	Calineuria californica, Hesperoperla pacifica, Pteronarcys californica, Pteronarcys princeps	No	Yes	No	No	No	Yes	Peterson et al. (2017)
	Dinocras cephalotes	No	Yes	Yes	No	No	No	Elbrecht et al. (2014)
	Doroneuria baumanni	No	Yes	No	No	No	No	Schultheis et al. (2012)
	Lednia tumana, Lednia tetonica, Zapada glacier	No	Yes	No	No	No	No	Hotaling et al. (2018)
	Mesocapnia arizonensis						No	Phillipsen et al. (2015)
	Peltoperla tarteri	No	Yes	No	No	No	No	Schultheis et al. (2002)
	Pteronarcella badia	No	Yes	No	No	No	No	Sproul et al. (2014)
	Pteronarcys californica	No	Yes	No	No	No	No	Kauwe et al. (2004)
	Pteronarcys proteus	yes	No	No	No	No	No	White (1989)

	Yoraperla brevis	Yes	No	No	No	No	No	Hughes et al. (1999)
	Zelandoperla decorata	No	Yes	No	No	Yes	Yes	McCulloch et al. (2009), Dussex et al. (2016)
	Zelandoperla fenestrata	No	Yes	No	No	Yes	Yes	McCulloch et al. (2009), Dussex et al. (2016)
Odonata, dragonflies and damselflies	*Calopteryx splendens*	No	No	No	Yes	No	No	Chaput-Bardy et al.(2008)
	Coenagrion mercuriale	No	No	Yes	No	No	No	Keller et al. (2012), Keller & Holderegger (2013)
	Diplacodes haematodes, Orthetrum caledonicum	No	Yes	No	No	No	No	Razeng et al. (2017)
	Euphaea formosa	No	Yes	No	No	No	No	Huang & Ling (2011)
Coleoptera, water bettles	*Boreonectes aequinoctialis*	No	No	Yes	No	No	No	Phillipsen et al. (2015)
	Exocelina manokwariensis	No	Yes	No	No	Yes	No	Lam et al. (2018a)
	Philaccolilus ameliae	No	No	No	No	Yes	No	Lam et al. (2018b)
	Psephenus montanus	No	No	No	Yes	No	No	Miller et al. (2002)
	Sclerocyphon sp.	No	Yes	No	No	No	No	Chester et al. (2015)
Hemiptera, water bugs	*Abedus herberti*	No	No	Yes	No	No	No	Phillipsen & Lytle (2013), Phillipsen et al. (2015)
	Ambrysus thermarum	No	No	No	Yes	No	No	Miller et al. (2002)
	Aquarius remigis	Yes	No	No	No	No	No	Preziosi & Fairbairn (1992)
	Belostoma angustum	No	Yes	No	No	No	Yes	Stefanello et al. (2020)
	Rheumatometra sp.	Yes	No	No	No	No	No	Bunn & Hughes (1997)
Diptera, flies	*Echinocladius martini*	No	Yes	No	No	No	Yes	Krosch et al. (2009, 2011)
	Elporia barnardi	Yes	Yes	No	No	No	No	Wishart & Hughes (2001, 2003)*
	Ferringtonia patagonica, Naonella forsythi	No	Yes	No	No	No	No	Krosch et al. (2012)

Liponeura cinerascens cinerascens	No	Yes	Yes	No	No	No	Schröder, et al. (2021)
Prosimulium neomacropyga	No	Yes	No	No	No	No	Finn et al. (2006)

Table A.2.3 Summary of species and indicated dispersal pattern based on molecular data. Data are presented for species within Coleoptera, Odonata, Diptera, Hemiptera and Megaloptera.

Order / Species	Location	Marker	Level Tested	Indicated dispersal pattern	Reference
Coleoptera					
Boreonectes aequinoctialis	Arizona, USA	Microsatellites	Catchments	Long-distance dispersal by flight	Phillipsen et al. (2014)
Psephenus montanus	Arizona, USA	AFLPs	Catchments	Dispersal limited within catchment	
Sclerocyphon sp.	Victoria, Australia	mtDNA	Catchments	Limited dispersal by crawling	Chester et al. (2015)
Odonata					
Calopteryx splendens	Western France	AFLPs	Catchments	Overland dispersal between streams	Chaput-Bardy et al. (2008)
Coenagrion mercuriale	Switzerland	Microsatellites	Catchments	Dispersal limited to natal site and occasional overland dispersal	Keller and Holderegger (2013)
Diplacodes haematodes, Orthetrum caledonicum	Australian arid zone	mtDNA	Regional	Strong dispersal potential across regions	Razeng et al. (2017)
Diptera					
Echinocladius martini	Northeast Queensland, Australia	mtDNA	Catchments	Dispersal is limited within a stream channel	Krosch et al. (2011)
Elporia barnardi	South-western Cape, South Africa	Allozymes, mtDNA	Catchments	Dispersal occurs within catchment	Wishart & Hughes (2001,2003)*
Ferringtonia patagonica, Naonella forsythi	Patagonia and New Zealand	mtDNA	Catchment	Overland dispersal between streams	Krosch et al. (2012)
Prosimulium neomacropyga	Colorado, USA	mtDNA	Catchments	Dispersal occurs among adjacent streams	Finn et al. (2006)
Hemiptera					
Abedus herberti	Arid southwest, USA and Northern Mexico	Microsatellites	Catchments	Dispersal occurs within stream	Phillipsen and Lytle (2013)
Ambrysus thermarum	Arizona, USA	AFLPs	Catchments	Dispersal occurs both within catchments and across watersheds	Miller et al. (2002)
Aquarius remigis	USA	Allozymes	Catchments	Dispersal mainly by drift within stream	Preziosi and Fairbairn (1992)
Rheumatometra sp.	Queensland, Australia	Allozymes	Catchments	Overland adult dispersal among catchments	Bunn and Hughes (1997)
Megaloptera					
Archichauliodes diversus	New Zealand	Allozymes	Regional	Dispersal is more likely within a catchment	Hogg et al. (2002)

* References not cited in the main text:

Wishart, M., J. & Hughes, J., M. (2001). Exploring patterns of population subdivision in the net-winged midge, *Elporia barnardi* (Diptera: Blephariceridae), in mountain streams of the south-western Cape, South Africa. *Freshwater Biology*, **46**, 479-490.

Wishart, M., J. & Hughes, J., M. (2003) Genetic population structure of the net-winged midge, *Elporia barnardi* (Diptera: Blephariceridae) in streams of the south- western Cape, South Africa: implications for dispersal. *Freshwater Biology*, **48**, 828-838.

Table A.2.4 Articles that adopted a spatial genetic structure approach for data analysis and spatial variables included in each of the studies.

Article	Euclidean or Geographic distance	Stream distance	Landscape topography	Other landscape/environmental elements e.g. canopy cover, wind, temperature
Preziozi, 1992	X			
Kelly et al., 2001	X			
Myers et al., 2001	X			
Wilcock et al., 2001	X		X	
Kelly et al., 2002	X			
Miller et al., 2002	X			
Schultheis et al., 2002		X		
Hughes, Mather et al., 2003	X	X		
Wilcock et al., 2003	X			
Wishart & Hughes, 2003	X		X	
Kauwe et al., 2004	X			
Finn et al., 2006	X	X	X	X
Smith et al., 2006	X			
Smith et al., 2006	X			
Wilcock et al., 2007	X	X	X	X
Zickovich & Bohonak, 2007	X			
Chaput-Bardy et al., 2008	X	X		
Mclean et al., 2008	X			
Smith & Smith, 2009	X			
Murria et al., 2010	X			
Baggiano et al., 2011	X			
Engelhardt et al., 2011	X			
Huang & Ling, 2011	X			
Krosch et al., 2011	X		X	
Sabando et al., 2011	X			
Alp et al., 2012	X	X	X	X
Keller et al., 2012	X		X	
Phillipsen, 2012	X	X	X	
Keller & Holderegger, 2013	X	X	X	
Elbrecth et al., 2014	X	X	X	
Pessino et al., 2014	X			
Sproul et al., 2014	X			
Wickson et al., 2014	X			
Yaegashi et al., 2014	X	X		
Addison et al., 2015	X			
Geismar et al., 2015	X		X	
Phillpsen et al., 2015	X	X	X	X
Sekine et al., 2015	X			
Dussex et al., 2016	X			
Polato et al., 2017	X	X	X	X
Weigand et al., 2018	X	X		X

Rader et al., 2019	X	X	X
Stefanello et al., 2020	X	X	X

A3: Supplementary material for Chapter 3

Figure A.3.1. Summary of the FastQC quality report for one individual of *Coloburiscus humeralis*. Report includes basic statistics, per base sequence quality, per base sequence content, sequence duplication level, adapter content, and over-represented sequences.

a) Raw RAD-seq data

b) Cleaned data following IPYRAD pipeline.

Basic Statistics

Measure	Value
Filename	AIWNZ038_L001001.trimmed_R1_.fastq.gz
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	970102
Sequences flagged as poor quality	0
Sequence length	60-150
%GC	34

Per base sequence quality

Quality scores across all bases (Sanger / Illumina 1.9 encoding)

Per base sequence content

Sequence content across all bases

Sequence Duplication Levels

Percent of seqs remaining if deduplicated 56.62%

Adapter Content

% Adapter

Overrepresented sequences

Sequence	Count	Percentage	Possible Source
AGATATCAACCGTGTAAGTTTGCATGTTTTACATGAGAAATAGCTATGAA	8183	0.8435195474290332	No Hit
AGATATCAACCGTGTAAGTTTGCATGTTTTACGTGAGAAATAGCTATGAA	3960	0.4082044980837067	No Hit
GCACACGTCTGAACTCCAGTCACACTTGAATCTCGTATGCCGTCTTCTGC	1857	0.1914231699347079	TruSeq Adapter, Index 8 (100% over 50bp)
AAAAAAAAATAAAAAAATTTTTTTGTGTGATTTTATAGCTATATCTCATT	1166	0.1201935466579803	No Hit

166

Table A.3.1 Summary statistics following use of IPYRAD pipeline for the complete dataset of the three target species.

a) *Coloburiscus humeralis*

sample	raw reads	trim_adapter_bp_read1	trim_quality_bp_read1	reads_filtered_by_Ns	reads_filtered_by_minlen	retained reads
AIWNZ002	1250730	144390	4319176	0	0	1237336
AIWNZ006	1787319	194944	4497704	0	0	1775140
AIWNZ008	978374	122154	2305894	0	0	968932
AIWNZ009	770808	80704	1939025	0	0	765944
AIWNZ010	1209658	124800	2550310	0	0	1203374
AIWNZ011	1624516	179093	2818549	0	0	1617133
AIWNZ013	2155454	274306	4544528	0	0	2142887
AIWNZ025	1793482	174892	3115827	0	0	1786866
AIWNZ028	884619	66858	1893826	0	0	880935
AIWNZ038	974056	74923	1427752	0	0	970102
AIWNZ039	627332	54424	1745074	0	0	624027
AIWNZ041	712507	69363	1191650	0	0	709580
AIWNZ043	593127	60426	1001441	0	0	590482
AIWNZ051	1077486	115391	1718402	0	0	1072827
AIWNZ052	688853	70493	1404937	0	0	685124
AIWNZ061	1070817	103186	2162314	0	0	1065788
AIWNZ064	715033	65877	1673982	0	0	711089
AIWNZ065	1357465	203023	3532701	0	0	1346276
AIWNZ080	1410234	146217	3217388	0	0	1403736
Total	21681870			0	0	21557578
Total %						99.42%

b) *Zelandobius confusus*

sample	raw reads	trim_adapter_bp_read1	trim_quality_bp_read1	reads_filtered_by_Ns	reads_filtered_by_minlen	retained reads
AIWNZ191	517434	22640	778693	0	0	515962
AIWNZ192	262593	21818	523506	0	0	262003
AIWNZ193	130051	12218	263161	0	0	129767
AIWNZ194	363935	47420	543372	0	0	363049
AIWNZ195	161310	10495	385567	0	0	160980
AIWNZ196	89563	9290	156752	0	0	89425
AIWNZ197	189440	12429	375112	0	0	189048
AIWNZ198	362364	22464	503176	0	0	361668
AIWNZ199	545824	52809	594314	0	0	544553
AIWNZ200	362717	96465	615109	0	0	361504
AIWNZ201	450496	26282	494098	0	0	449497
AIWNZ202	104378	8568	317156	0	0	104141
AIWNZ203	389202	12708	468458	0	0	388325
AIWNZ204	343803	33269	584970	0	0	343017
AIWNZ205	39845	3118	124126	0	0	39757
AIWNZ206	263775	13261	783812	0	0	263204
AIWNZ207	147364	9318	273124	0	0	147097
AIWNZ208	90293	6688	237820	0	0	90093
AIWNZ209	99572	7425	209439	0	0	99407
AIWNZ210	7545	1741	39845	0	0	7511
AIWNZ211	57800	3959	184118	0	0	57653
AIWNZ212	425219	33477	559520	0	0	424189
AIWNZ213	158380	13086	402197	0	0	158011
AIWNZ214	71359	5692	295005	0	0	71104
AIWNZ215	449593	19692	512080	0	0	448626
AIWNZ216	302339	32580	514347	0	0	301447
AIWNZ217	9331	1874	47992	0	0	9283

AIWNZ218	4982	944	34494	0	0	4958
AIWNZ219	91198	17965	204946	0	0	90989
AIWNZ220	344391	56405	593859	0	0	343228
AIWNZ222	6607	213	186820	0	0	6531
AIWNZ223	285063	33274	592576	0	0	284003
AIWNZ224	78073	27707	199666	0	0	77642
AIWNZ225	6229	2015	28503	0	0	6201
AIWNZ226	64595	18966	185830	0	0	64376
AIWNZ227	456888	21551	554214	0	0	455819
AIWNZ228	241878	27311	404320	0	0	241247
AIWNZ229	29944	3781	110479	0	0	29871
AIWNZ230	377468	42767	528438	0	0	376793
AIWNZ231	233089	35363	527411	0	0	232068
AIWNZ232	145419	20764	459069	0	0	144657
AIWNZ233	30136	2256	136701	0	0	30050
AIWNZ234	146641	22606	460364	0	0	145843
AIWNZ235	115603	9531	185176	0	0	115397
AIWNZ236	5545	499	20464	0	0	5533
AIWNZ237	149855	7232	285129	0	0	149583
AIWNZ238	125455	9983	265931	0	0	125226
AIWNZ239	182047	4768	626368	0	0	181732
AIWNZ240	388795	57552	815750	0	0	387538
AIWNZ241	174809	15087	377994	0	0	174345
AIWNZ242	56713	5950	164759	0	0	56584
AIWNZ243	52717	4470	157303	0	0	52613
AIWNZ244	329954	34111	442494	0	0	329258
AIWNZ245	107131	7350	242148	0	0	106901
AIWNZ246	212750	15656	425216	0	0	212330
AIWNZ247	200472	10631	419285	0	0	200119
AIWNZ248	335599	25551	678383	0	0	334839
AIWNZ249	408302	21329	607323	0	0	407483
AIWNZ250	2125	160	26328	0	0	2101
AIWNZ251	276556	11208	485735	0	0	276018
AIWNZ252	454420	47740	871652	0	0	453178
AIWNZ253	218186	20613	350700	0	0	217713
AIWNZ254	524791	32193	769796	0	0	523621
AIWNZ255	85895	3529	238409	0	0	85732
AIWNZ256	33506	4927	74930	0	0	33453
AIWNZ257	366565	19424	722452	0	0	365814
AIWNZ258	259869	19444	565009	0	0	259307
AIWNZ259	34419	1338	79926	0	0	34354
AIWNZ260	28428	1602	87304	0	0	28357
AIWNZ261	62097	2655	137740	0	0	61959
AIWNZ262	40258	3023	228662	0	0	39663
AIWNZ263	1576	92	20201	0	0	1559
AIWNZ264	157418	47556	377371	0	0	156293
AIWNZ265	22315	2214	84817	0	0	22242
AIWNZ266	22225	4608	81560	0	0	22178
AIWNZ267	361453	59013	608558	0	0	360520
AIWNZ268	515496	37191	766528	0	0	514246
AIWNZ269	416151	39720	697143	0	0	415022
AIWNZ270	63130	3658	221323	0	0	63002
AIWNZ271	194292	17096	270032	0	0	193910
AIWNZ272	405853	30094	754366	0	0	404760
AIWNZ273	403487	27220	637730	0	0	402480
AIWNZ274	428029	49336	564806	0	0	426891
AIWNZ275	337274	59508	661591	0	0	336199
AIWNZ276	224561	57925	509582	0	0	223633
AIWNZ277	424716	60343	644373	0	0	423179
AIWNZ278	110156	9813	269333	0	0	109913
AIWNZ279	422995	18898	756868	0	0	422094
AIWNZ280	391077	41743	636642	0	0	390151
AIWNZ281	220882	10099	518711	0	0	220417
AIWNZ282	184877	14985	365208	0	0	184510
AIWNZ283	389030	17509	896702	0	0	388243
AIWNZ284	185491	14518	387306	0	0	185101

sample	reads_raw					retained reads
AIWNZ285	221484	9891	629678	0	0	220096
Total	20274956					20221987
Total %						99.70%

c) *Hydropsyche fimbriata*

sample	reads_raw	trim_adapter_bp_read1	trim_quality_bp_read1	reads_filtered_by_Ns	reads_filtered_by_minlen	retained reads
AIWNZ096	171196	16558	105630	0	0	170967
AIWNZ097	168609	39090	93428	0	0	168131
AIWNZ098	166823	30358	128268	0	0	166330
AIWNZ099	140117	45705	110315	0	0	139159
AIWNZ100	51439	13881	67644	0	0	51335
AIWNZ101	123888	19983	174747	0	0	123535
AIWNZ102	156038	25376	162454	0	0	155373
AIWNZ103	154834	31416	139107	0	0	154025
AIWNZ104	141418	22860	102220	0	0	141156
AIWNZ105	148915	22645	162751	0	0	148447
AIWNZ106	177419	23459	110682	0	0	177115
AIWNZ107	100693	18995	187868	0	0	100242
AIWNZ108	113694	22524	83883	0	0	113355
AIWNZ109	156137	30854	93686	0	0	155717
AIWNZ110	104758	19275	102702	0	0	104384
AIWNZ111	25777	5280	39637	0	0	25727
AIWNZ112	149307	32301	144193	0	0	148447
AIWNZ113	107627	14016	196910	0	0	107272
AIWNZ114	125097	17405	233011	0	0	124674
AIWNZ115	156592	26865	128451	0	0	156130
AIWNZ116	147861	28359	142080	0	0	147309
AIWNZ117	100966	12879	155894	0	0	100767
AIWNZ118	98930	15764	187284	0	0	98556
AIWNZ119	115813	19637	171009	0	0	115339
AIWNZ120	149527	19720	95179	0	0	149150
AIWNZ121	96320	16809	88132	0	0	95915
AIWNZ122	90683	12352	137231	0	0	90455
AIWNZ123	156109	32706	152137	0	0	154965
AIWNZ124	125357	23717	160395	0	0	124835
AIWNZ125	171353	26911	151657	0	0	170849
AIWNZ126	111146	13173	194982	0	0	110835
AIWNZ127	104458	16112	142574	0	0	104183
AIWNZ129	128242	18660	131526	0	0	127813
AIWNZ130	99300	18861	122330	0	0	98779
AIWNZ131	104759	19928	229635	0	0	104152
AIWNZ132	150231	17536	121488	0	0	149917
AIWNZ133	97551	15908	72002	0	0	97323
AIWNZ134	103451	35431	95556	0	0	102925
AIWNZ135	112388	16953	110780	0	0	112108
AIWNZ136	155214	36408	199125	0	0	154104
AIWNZ137	60851	11211	95421	0	0	60644
AIWNZ138	144506	32134	188155	0	0	143401
AIWNZ139	86681	11842	112700	0	0	86509
AIWNZ140	173707	22226	118463	0	0	173360
AIWNZ141	152297	20333	189554	0	0	151787
AIWNZ142	148966	18810	160191	0	0	148640
AIWNZ143	96162	16872	237099	0	0	95608
AIWNZ144	105592	14434	136245	0	0	105357
AIWNZ145	16414	2207	62735	0	0	16341
AIWNZ146	18748	3333	39455	0	0	18680

AIWNZ147	54098	16116	84918	0	0	53680
AIWNZ148	34456	7666	73794	0	0	34229
AIWNZ149	79794	21940	131997	0	0	79280
AIWNZ150	10424	1693	31862	0	0	10388
AIWNZ151	57214	9173	176261	0	0	56896
AIWNZ152	155807	30048	228323	0	0	154918
AIWNZ153	125206	17133	187243	0	0	124839
AIWNZ154	152327	21221	145446	0	0	151983
AIWNZ155	80574	13935	181524	0	0	80300
AIWNZ156	120615	15030	87296	0	0	120451
AIWNZ157	160673	23977	120049	0	0	160348
AIWNZ158	112501	14012	153565	0	0	112308
AIWNZ159	123528	22948	212174	0	0	122822
AIWNZ160	159006	19981	144723	0	0	158705
AIWNZ161	124697	22335	110824	0	0	124362
AIWNZ162	148133	25803	161650	0	0	147719
AIWNZ163	149955	23043	156994	0	0	149402
AIWNZ164	124633	22677	198707	0	0	124014
AIWNZ165	170405	18834	150149	0	0	170023
AIWNZ166	130316	16675	163599	0	0	129989
AIWNZ167	133524	20721	164312	0	0	132946
AIWNZ168	135204	20438	100669	0	0	134885
AIWNZ169	134796	22689	120372	0	0	134469
AIWNZ170	130685	33744	130094	0	0	130108
AIWNZ171	172372	29655	187696	0	0	171766
AIWNZ172	151424	27685	176066	0	0	150924
AIWNZ173	181564	25827	185046	0	0	180982
AIWNZ174	95812	17755	155906	0	0	95377
AIWNZ175	118442	15777	160036	0	0	118174
AIWNZ176	185243	31397	186999	0	0	184486
AIWNZ177	93671	16586	108569	0	0	93440
AIWNZ178	138030	20295	143798	0	0	137547
AIWNZ179	113738	19201	151563	0	0	113227
AIWNZ180	144922	21012	135788	0	0	144526
AIWNZ181	127470	23159	121525	0	0	126926
AIWNZ182	116357	21583	155439	0	0	115666
AIWNZ183	87051	16699	132460	0	0	86496
AIWNZ184	125352	18562	131132	0	0	124846
AIWNZ185	132891	22914	139180	0	0	132529
AIWNZ186	108974	21758	198791	0	0	108310
AIWNZ187	153781	20523	132634	0	0	153416
AIWNZ188	139885	29236	166996	0	0	139423
AIWNZ189	151295	17916	86487	0	0	151111
AIWNZ190	62312	9534	119194	0	0	62149
Total	11473118					11432512
Total %						99.64%

Figure A.3.2 Coverage and missing data for *Coloburiscus humeralis* data set of 41,029 SNPs and 72.84% missing data. a) Depth of coverage per individual information for the more conservative dataset; and b) matrix of missing data. Individuals are listed in Appendix Table A.3.1.

a)
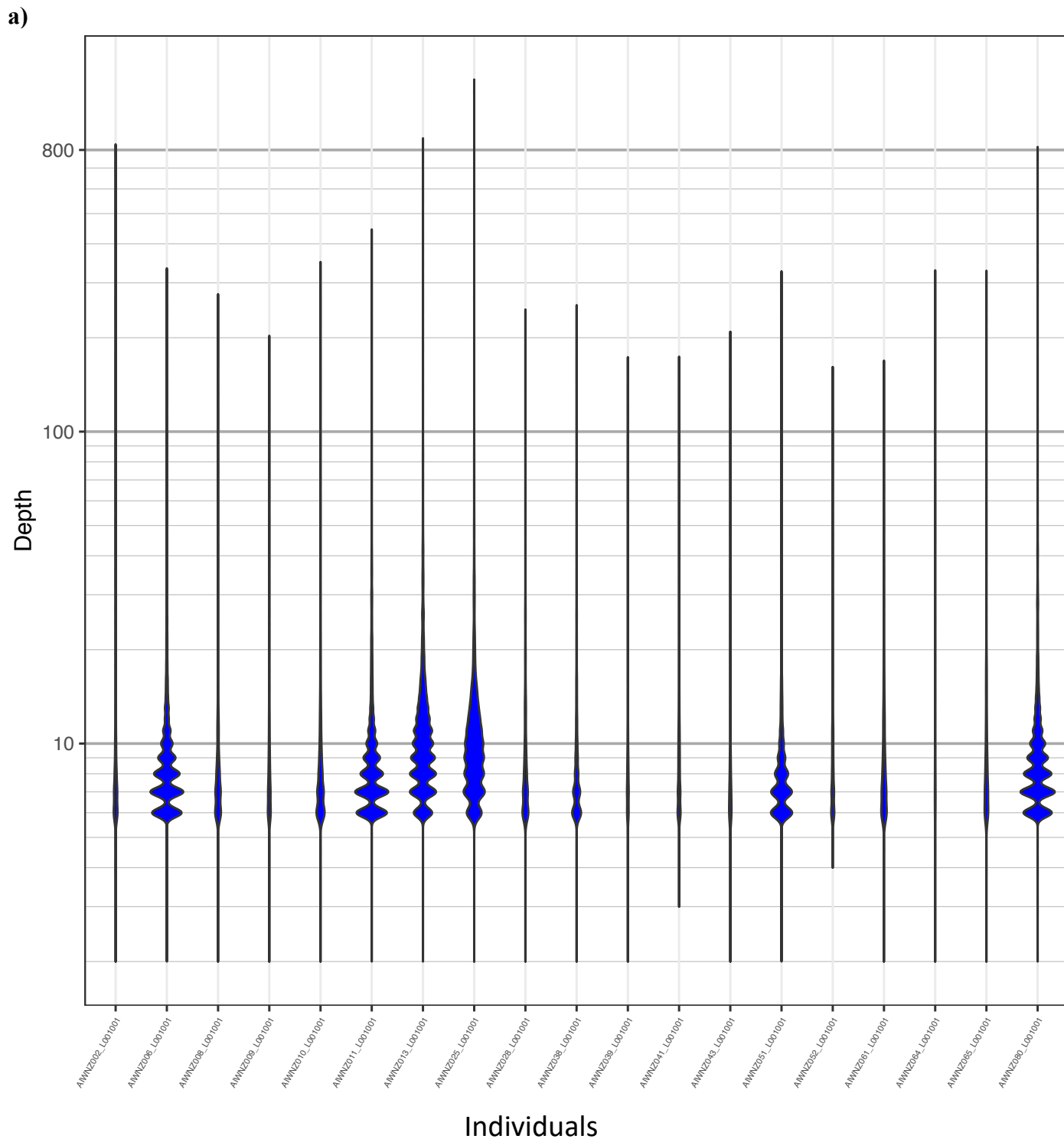
b)
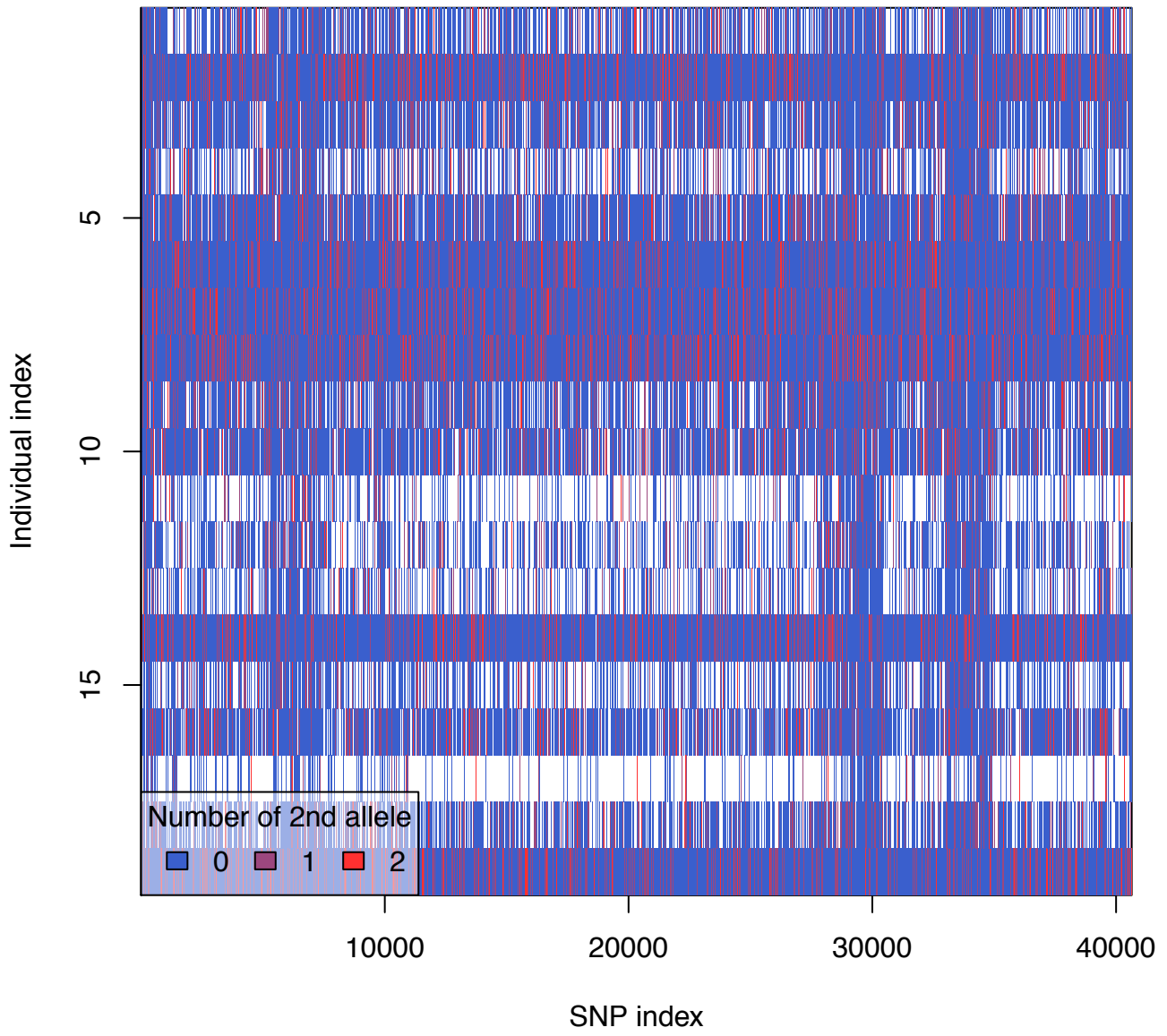
Figure A.3.3 Coverage and missing data for *Zelandobius confusus* set of 925 SNPs and 65% missing data. a) Depth of coverage information for the more conservative dataset; and b) matrix of missing data. Individuals are listed in Appendix Table A.3.1.

a)

b)

Number of 2nd allele

0   1   2
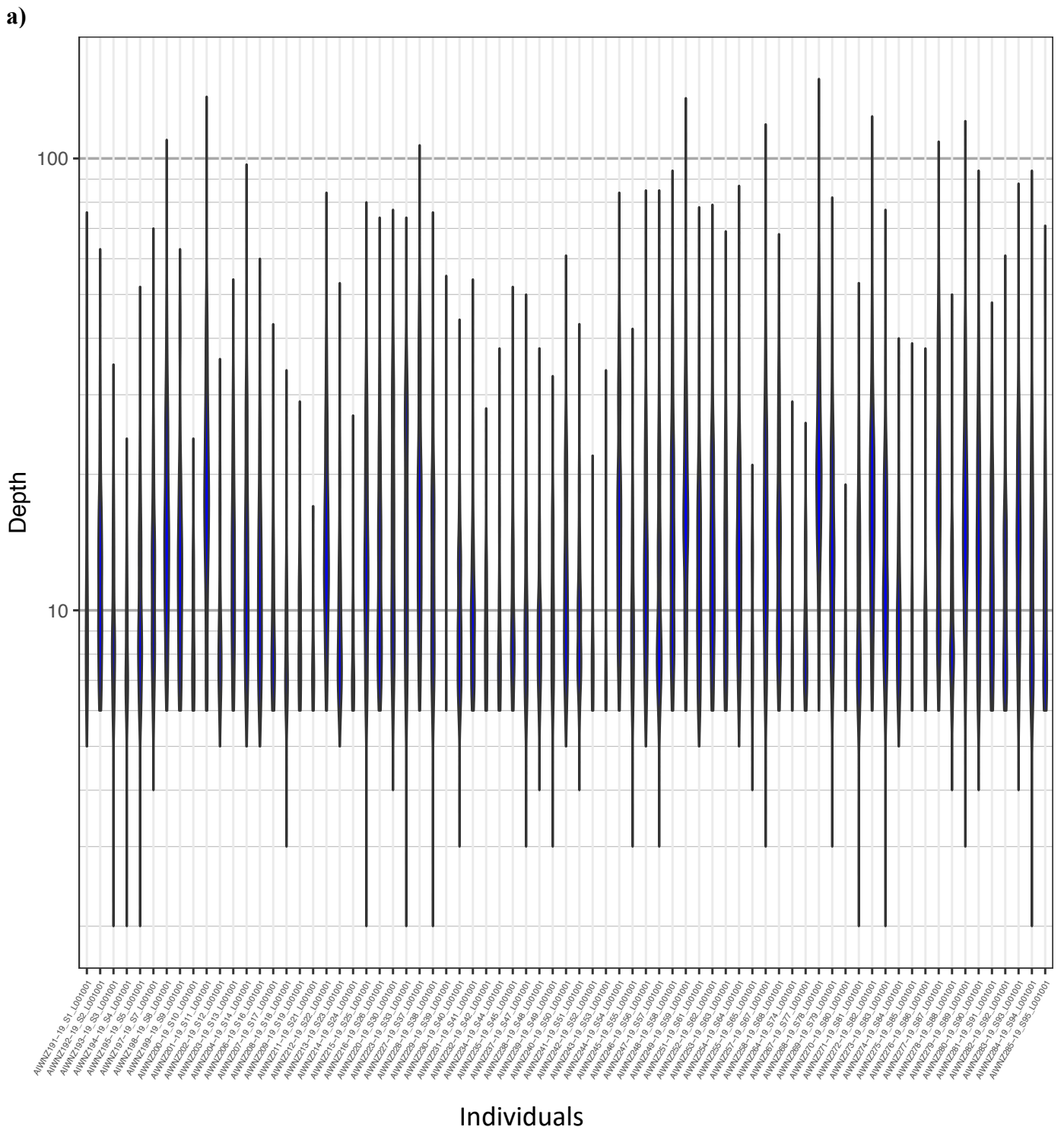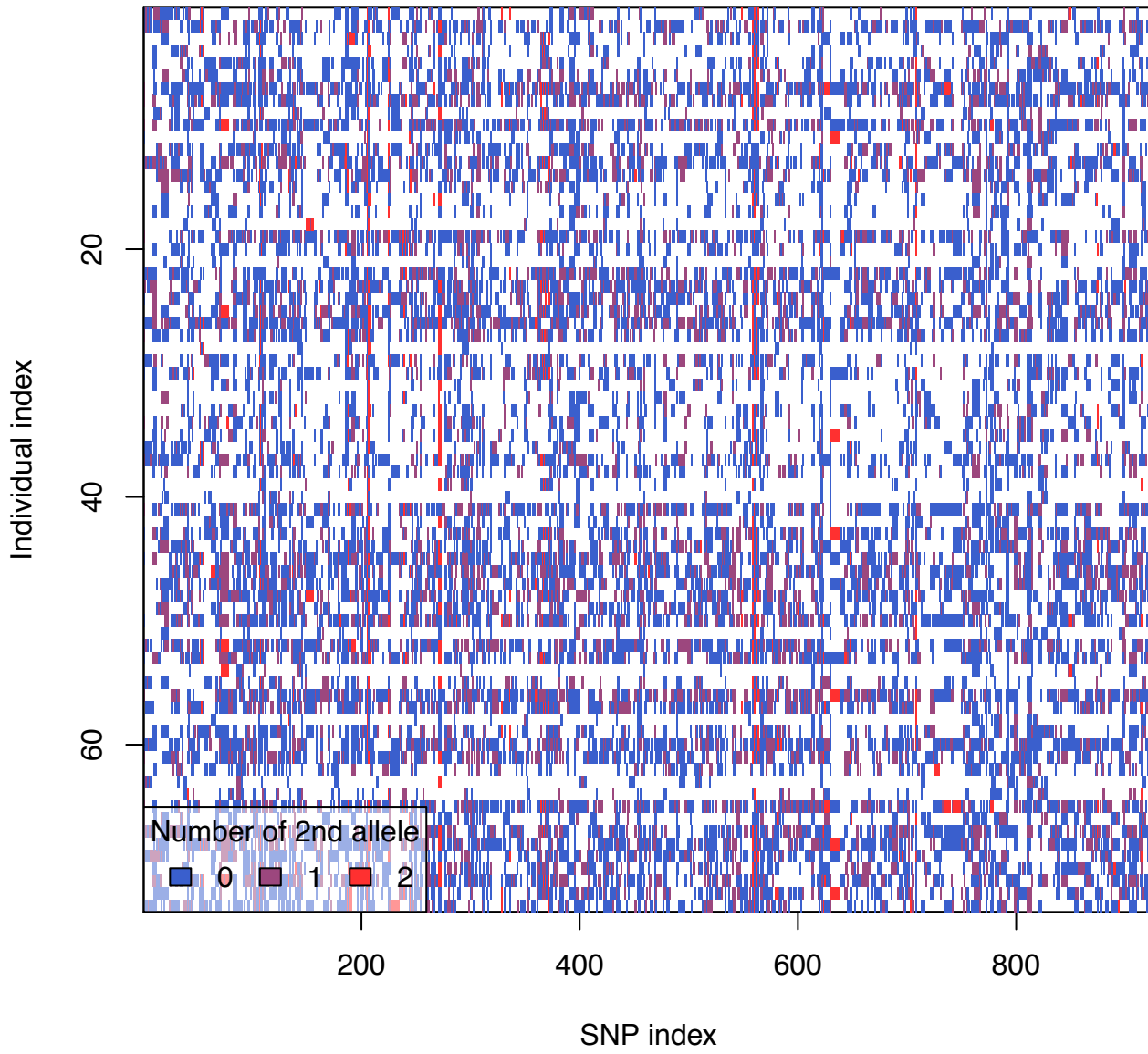
Figure A.3.4 Coverage and missing data for *Hydropsyche fimbriata* set of 202 SNPs and 48.86% missing data. (a) Depth of coverage information for the more conservative dataset; and (b) matrix of missing data. Individuals are listed in Appendix Table A.3.1.

a)
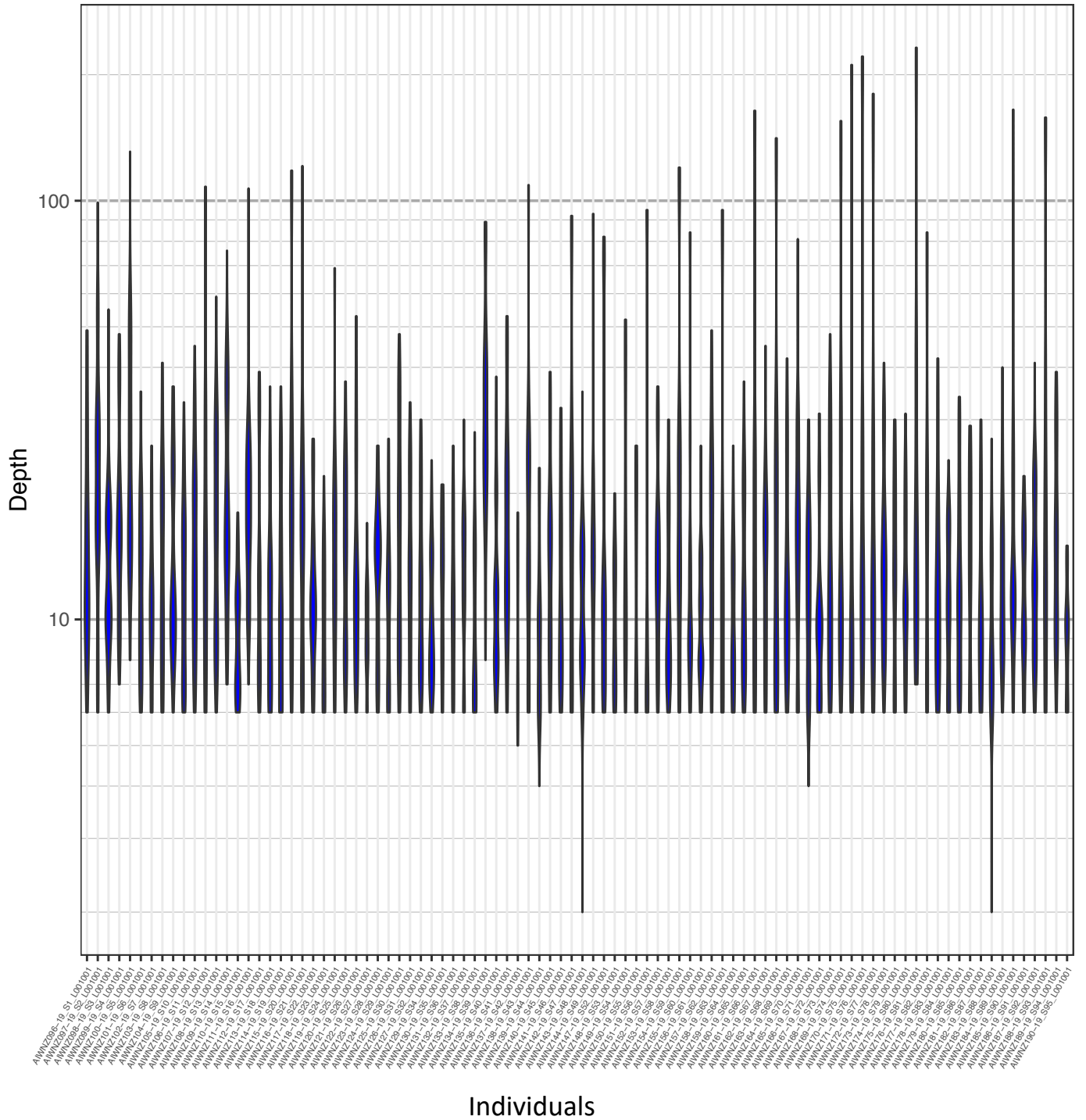
b)
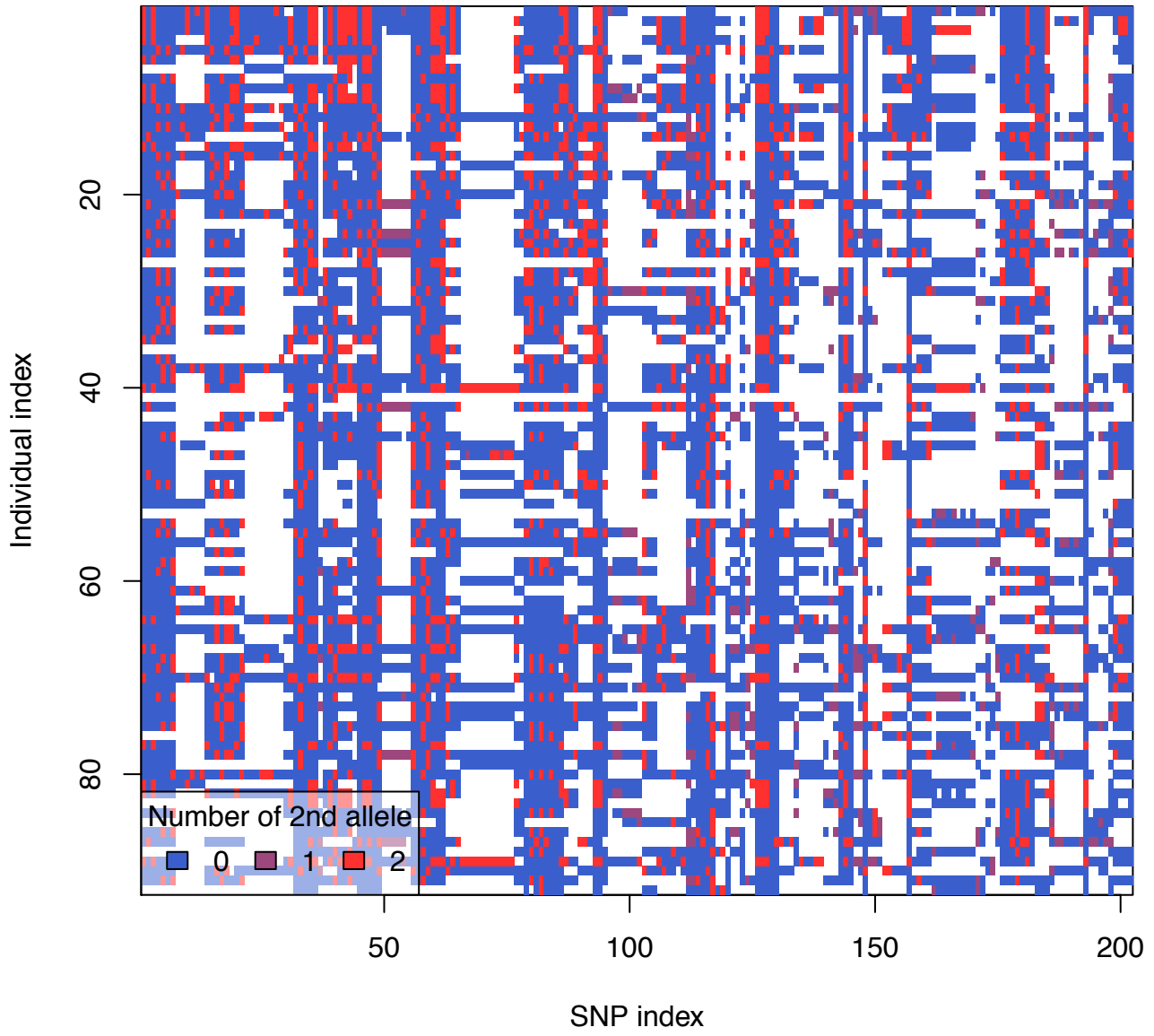
A4: Supplementary material for Chapter 4

Table A.4.1 Sampling sites localities, population codes, downstream distance (from the top to the bottom sampling site) within each stream, and riparian vegetation characteristics for each sampling site.

Locality	Population code	Waterway distance	Riparian land cover	Coordinates
Pirongia				
Tawhitiwhiti Stream	stream A			
	A1	0 m	indigenous forest	37° 57' 33.0" S, 175° 05' 44.8" E
	A2	700 m	pasture	37° 57' 15.5" S, 175° 05' 43.9" E
	A3	1100 m	pasture	37° 57' 06.5" S, 175° 05' 43.9" E
Te Pahu Stream	stream B			
	B1	0 m	indigenous forest	37° 57' 48.2" S, 175° 06' 02.1" E
	B2	490 m	indigenous forest fragment	37° 57' 35.0" S, 175° 06' 07.5" E
	B3	620 m	pasture	37° 57' 31.9" S, 175° 06' 10.4" E
	B4	1140 m	replanted	37° 57' 16.3" S, 175° 06' 08.1" E
Ngakoaohia Stream	stream C			
	C1	0 m	indigenous forest	38° 03' 10.8" S, 175° 05' 02.9" E
	C2	550 m	indigenous forest	38° 03' 20.2" S, 175° 05' 16.9" E
	C3	1500 m	indigenous forest	38° 03' 35.0" S, 175° 05' 34.9" E
	C4	3500 m	indigenous forest	38° 03' 47.1" S, 175° 05' 46.7" E
Karioi	stream D			
Wainui Stream	D1	-	indigenous forest	37° 50' 26.1" S, 174° 48' 27.1'' E
Taranaki				
Katikara Stream	stream E			
	E1	-	indigenous forest	39° 12' 07.4'' S, 173° 57' 36.1'' E
Patea Stream	stream F			
	F1	-	indigenous forest	39° 19' 26.1'' S, 174° 11' 25.9'' E

Table A.4.2 Results of Mantel tests across all populations for each of the three study species using linearised F_{ST} as a measure of genetic distance and Euclidean distance as a measure of geographic distance.

	COI-based analysis		SNP-based analysis	
	r	P-value	r	P-value
Coloburiscus humeralis (13, 13)				
Stream A	0.000	0.832	0.926	0.337
Stream B	0.589	0.130	0.637	0.082
Stream C	0.000	0.668	0.662	0.082
Zelandobius confusus (13, 10)				
Stream A	0.957	0.114	0.000	0.834
Stream B	0.122	0.330	0.377	0.673
Stream C	0.000	0.828	0.903	0.333
Hydropsyche fimbriata (11, 13)				
Stream A	0.983	0.168	0.767	0.336
Stream B	0.617	0.165	0.567	0.285
Stream C	0.767	0.167	0.516	0.175

Numbers in parentheses indicate the total number of sampling sites analysed for each species based on COI and SNP data sets, respectively. The approximate spatial scale in each of the analyses was: all sampling sites (170 km); Pirongia (12 km), stream A (0.8 km), stream B (1 km), stream C (1.5 km).

Table A.4.3 Sample sizes (*n*) of analysed data in each of the population sampling localities for the three study species.

Species	Locality	COI data (*n*)	SNP data (*n*)
Coloburiscus humeralis	A1	4	16
	A2	10	15
	A3	8	14
	B1	11	15
	B2	3	15
	B3	5	15
	B4	4	15
	C1	9	15
	C2	8	15
	C3	9	15
	C4	7	15
	D1	9	20
	E1	10	23
	F1	-	-
Total		97	208
Zelandobius confusus	A1	10	9
	A2	10	24
	A3	11	30
	B1	10	20
	B2	10	17
	B3	10	11
	B4	5	-
	C1	11	12
	C2	10	12
	C3	12	16
	C4	2	-
	D1	4	-
	E1	9	9
	F1	-	-
Total		114	160
Hydropsyche fimbriata	A1	**7**	19
	A2	10	21
	A3	12	20
	B1	11	27
	B2	3	29
	B3	10	29
	B4	33	28
	C1	9	19
	C2	5	19
	C3	3	20
	C4	4	15
	D1	-	26
	E1	-	2
	F1	-	19
Total		107	293

Table A.4.4 Pairwise genetic distances between sampling sites based on SNP data for each species. Upper diagonal: Nei's D (Nei, 1987) and lower diagonal: F_{ST} (Weir & Clark Cockerham, 1984).

C. humeralis

|	A1	A2	A3	B1	B2	B3	B4	C1	C2	C3	C4	D1	E1	F1
A1	-	0.022	0.025	0.017	0.011	0.012	0.019	0.028	0.023	0.031	0.041	0.023	0.042	-
A2	**0.038**	-	0.009	0.035	0.018	0.014	0.013	0.010	0.012	0.014	0.016	0.029	0.040	-
A3	**0.045**	0.000	-	0.038	0.020	0.015	0.013	0.011	0.014	0.015	0.016	0.032	0.042	-
B1	**0.029**	**0.074**	**0.080**	-	0.012	0.018	0.026	0.043	0.036	0.047	0.056	0.027	0.050	-
B2	**0.009**	**0.025**	**0.030**	**0.015**	-	0.009	0.014	0.024	0.020	0.028	0.034	0.021	0.039	-
B3	**0.012**	**0.013**	**0.016**	**0.028**	0.001	-	0.011	0.018	0.016	0.021	0.027	0.022	0.038	-
B4	**0.025**	**0.005**	**0.006**	**0.048**	**0.010**	0.001	-	0.016	0.016	0.020	0.024	0.027	0.040	-
C1	**0.053**	**0.004**	**0.005**	**0.094**	**0.041**	**0.024**	**0.014**	-	0.012	0.011	0.013	0.036	0.044	-
C2	**0.033**	**0.001**	**0.004**	**0.072**	**0.025**	**0.011**	**0.007**	0.001	-	0.015	0.019	0.031	0.040	-
C3	**0.056**	**0.007**	**0.008**	**0.100**	**0.044**	**0.025**	**0.019**	0.000	0.003	-	0.016	0.040	0.047	-
C4	**0.080**	**0.011**	**0.011**	**0.122**	**0.060**	**0.040**	**0.029**	**0.006**	**0.014**	**0.005**	-	0.046	0.053	-
D1	**0.041**	**0.054**	**0.060**	**0.057**	**0.034**	**0.037**	**0.044**	**0.072**	**0.055**	**0.076**	**0.091**	-	0.051	-
E1	**0.096**	**0.087**	**0.090**	**0.121**	**0.087**	**0.082**	**0.084**	**0.096**	**0.084**	**0.101**	**0.116**	**0.120**	-	-
F1	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Z. confusus

	A1	A2	A3	B1	B2	B3	B4	C1	C2	C3	C4	D1	E1	F1
A1	-	0.019	0.018	0.020	0.022	0.026	NA	0.026	0.025	0.023	NA	NA	0.072	-
A2	0.000	-	0.011	0.011	0.011	0.017	NA	0.018	0.017	0.014	NA	NA	0.063	-
A3	0.001	**0.00**	-	0.019	0.014	0.018	NA	0.016	0.015	0.012	NA	NA	0.062	-
B1	0.000	0.000	**0.003**	-	0.014	0.019	NA	0.018	0.017	0.015	NA	NA	0.063	-
B2	0.000	0.001	**0.004**	0.001	-	0.019	NA	0.020	0.020	0.017	NA	NA	0.065	-
B3	0.000	0.000	**0.003**	0.000	0.000	-	NA	0.025	0.023	0.021	NA	NA	0.070	-
B4	NA	NA	NA	NA	NA	NA	-	NA	NA	NA	NA	NA	NA	-
C1	**0.004**	**0.005**	**0.002**	0.003	**0.004**	**0.005**	NA	-	0.020	0.019	NA	NA	0.069	-
C2	0.002	0.003	**0.001**	0.002	**0.003**	0.002	NA	0.000	-	0.018	NA	NA	0.068	-
C3	**0.004**	**0.003**	0.001	**0.003**	**0.002**	**0.002**	NA	0.001	0.000	-	NA	NA	0.065	-
C4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	-	NA	NA	-
D1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	-	NA	-
E1	**0.104**	**0.097**	**0.096**	**0.097**	**0.098**	**0.101**	NA	**0.100**	**0.100**	**0.097**	NA	NA	-	-
F1	-	-	-	-	-	-	-	-	-	-	-	-	-	-

H. fimbriata

	A1	A2	A3	B1	B2	B3	B4	C1	C2	C3	C4	D1	E1	F1
A1	-	0.011	0.013	0.014	0.011	0.010	0.011	0.016	0.015	0.015	0.018	0.019	NA	0.120
A2	0.000	-	0.012	0.013	0.010	0.010	0.011	0.016	0.014	0.015	0.018	0.017	NA	0.119
A3	0.001	0.000	-	0.015	0.011	0.011	0.011	0.016	0.014	0.015	0.017	0.018	NA	0.118
B1	**0.009**	**0.009**	**0.011**	-	0.011	0.011	0.014	0.016	0.016	0.016	0.020	0.016	NA	0.114
B2	0.002	0.002	0.002	**0.006**	-	0.008	0.008	0.014	0.013	0.013	0.015	0.015	NA	0.116
B3	0.001	0.000	**0.002**	**0.007**	0.000	-	0.008	0.013	0.012	0.012	0.015	0.015	NA	0.117
B4	0.003	0.002	0.001	**0.009**	0.000	0.000	-	0.014	0.014	0.013	0.016	0.016	NA	0.117
C1	**0.011**	**0.010**	**0.009**	**0.010**	**0.008**	**0.008**	**0.009**	-	0.012	0.012	0.015	0.019	NA	0.107
C2	0.007	0.006	0.003	**0.014**	0.006	0.004	**0.008**	0.000	-	0.012	0.014	0.020	NA	0.112
C3	**0.009**	**0.008**	**0.006**	**0.015**	**0.006**	**0.006**	**0.007**	0.001	0.000	-	0.014	0.020	NA	0.111
C4	**0.012**	**0.011**	**0.006**	**0.018**	**0.007**	**0.009**	**0.010**	0.005	0.000	0.001	-	0.023	NA	0.114
D1	**0.020**	**0.018**	**0.019**	**0.019**	**0.163**	**0.017**	**0.019**	**0.020**	**0.022**	**0.022**	**0.027**	-	NA	0.122
E1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	-	NA
F1	**0.238**	**0.235**	**0.233**	**0.227**	**0.227**	**0.229**	**0.230**	**0.219**	**0.226**	**0.231**	**0.231**	**0.241**	NA	-

Significant F_{ST} values $P<0.001$ are showed in bold.

Table A.4.5 Variance inflation factors (VIF) for each variable (i.e. distance matrix) included in the candidate Multiple Regression Models (MRMs).

	VIF values
Euclidean distance	2.526
Topography (slope)	2.475
Land cover (forest:pasture, 2:1)	5.442
Land cover (forest:pasture, 5:1)	2.548

Figure A.4.1 Relationship between genetic (linearised F_{ST}) and geographic (Euclidean) distances between pairs of populations for *Coloburiscus humeralis*, *Zelandobius confusus* and *Hydropsyche fimbriata*. Black dots indicate pairs of populations within Pirongia and grey dots indicate pairs of populations among mountain regions. For *H. fimbriata*, data is only available for populations within the Pirongia region. Linear trend lines indicate IBD.
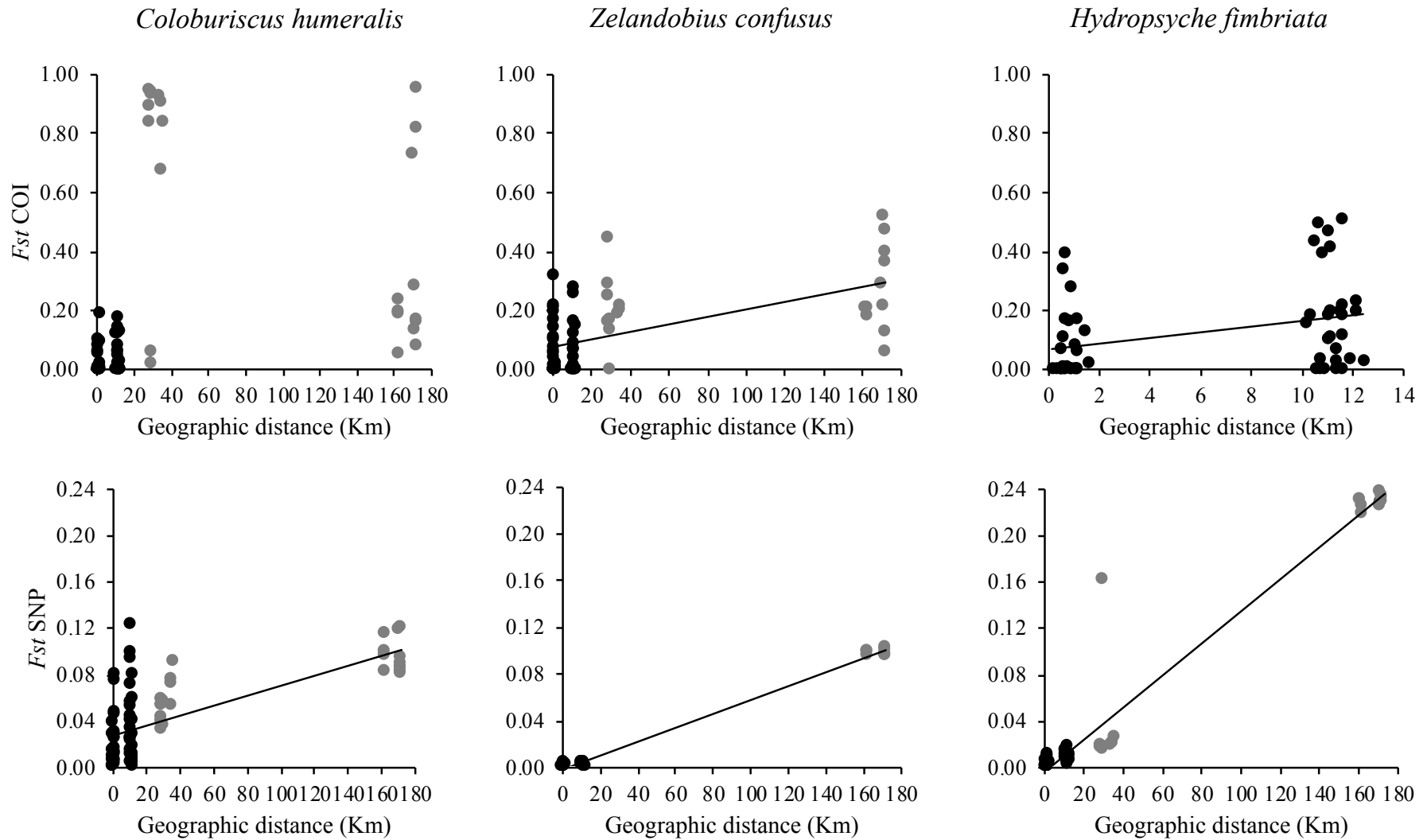
Figure A.4.2 Grassland/pasture-related classes and forest-related classes were used to build a final raster with two classes, 'forest' and 'pasture' and associated resistance cost matrix, using the Land Cover Database (10 m resolution) for mainland New Zealand (LCDB v. 5.0) from the Land Resource Information System (LRIS).

Mount Pirongia Land cover derived resistance cost
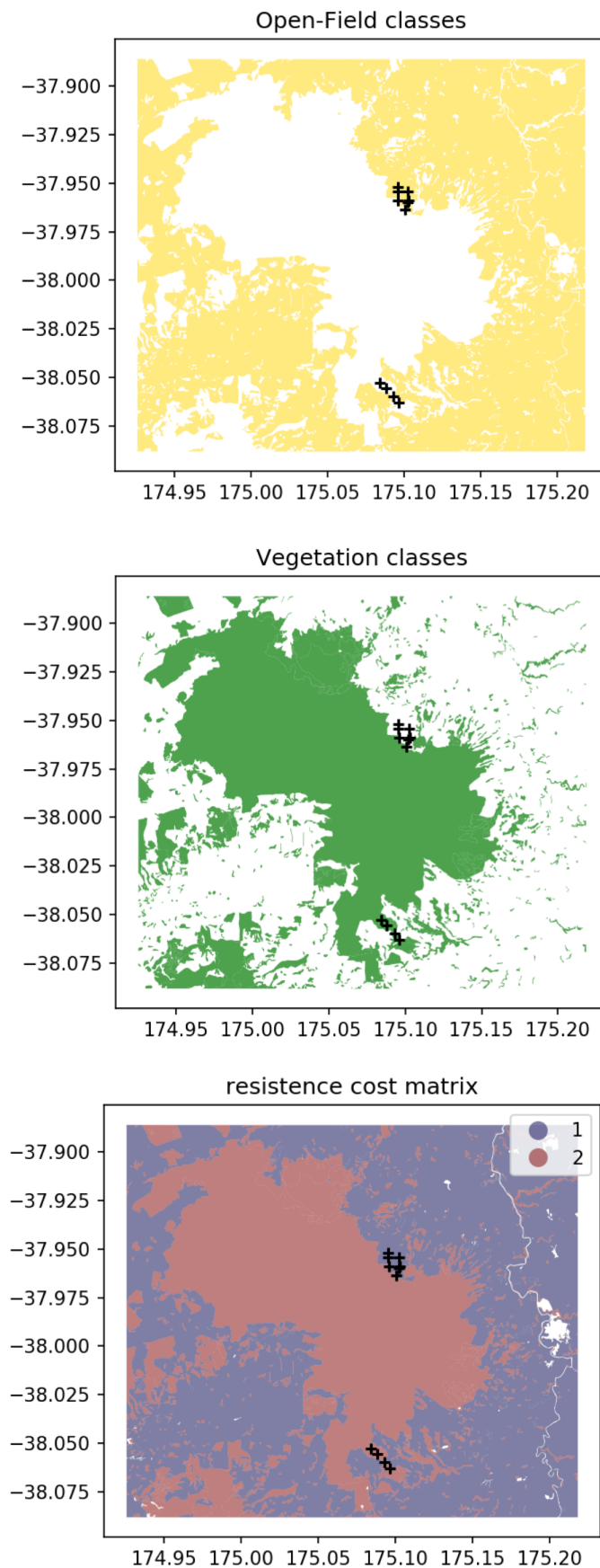
Figure A.4.3 Topographic complexity raster (10 m resolution) based on the slope values generated from a Digital Elevation Model (DEM, 8 m resolution) of Mount Pirongia region from the Land Information New Zealand (LINZ) Data Service.
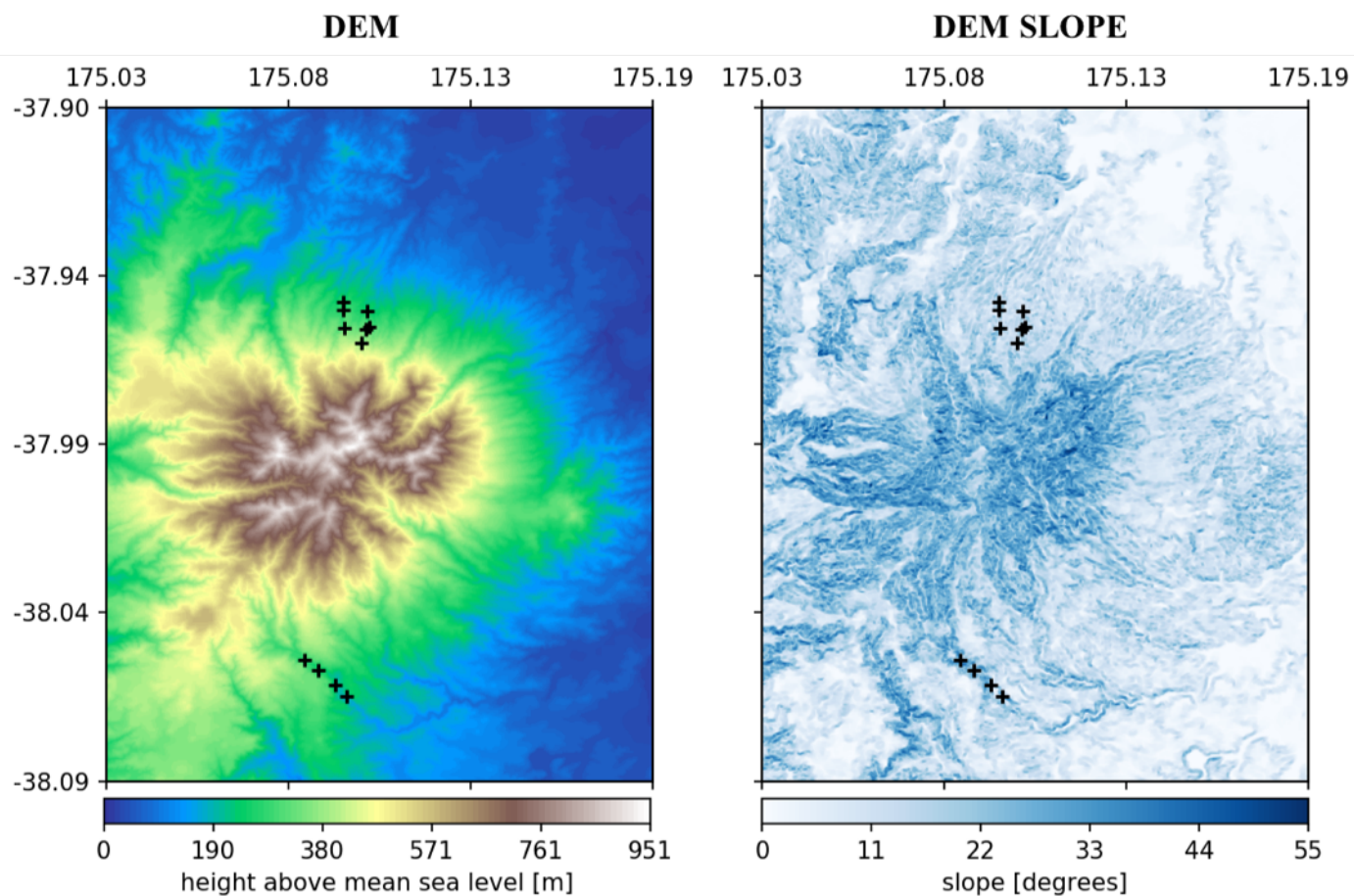
Assessing population genetic structure of three New Zealand stream insects using mitochondrial and nuclear DNA markers

Vanessa de Araujo Barbosa [a], S. Elizabeth Graham[b], Brian J. Smith[b], Ian D. Hogg[a,c], and Angela McGaughran[a]

[a]Te Aka Mātuatua, School of Science, University of Waikato, Hamilton, Waikato, New Zealand; [b]National Institute of Water and Atmospheric Research—NIWA, Hamilton, Waikato, New Zealand; [c]Polar Knowledge Canada, Canadian High Arctic Research Station, Cambridge Bay, NU, Canada

Corresponding author: **Vanessa de Araujo Barbosa** (email: van.araujobarbosa@gmail.com)

Abstract

Assessing genetic differentiation among natural populations can aid understanding of dispersal patterns and connectivity among habitats. Several molecular markers have become increasingly popular in determining population genetic structure for this purpose. Here, we compared the resolution of mitochondrial cytochrome *c* oxidase subunit I (COI) and nuclear single nucleotide polymorphism (SNP) markers for detecting population structure among stream insects at small spatial scales. Individuals of three endemic taxa—*Coloburiscus humeralis* (Ephemeroptera), *Zelandobius confusus* (Plecoptera), and *Hydropsyche fimbriata* (Trichoptera)—were collected from forested streams that flow across open pasture in the North Island of New Zealand. Both COI and SNP data indicated limited population structure across the study area, and small differences observed among these species were likely related to their putative dispersal abilities. For example, fine-scale genetic differentiation between and among neighbouring stream populations for *H. fimbriata* suggests that gene flow, and hence dispersal, may be more limited for this species relative to the others. Based on the generally similar results provided by both types of markers, we suggest that either COI or SNP markers can provide suitable initial estimates of fine-scale population genetic differentiation in stream insects.

Key words: dispersal, mtDNA, population structure, SNPs, stream insects

Résumé

Mesurer la différenciation génétique au sein de populations naturelles peut aider à comprendre leur dispersion et la connectivité entre habitats. Plusieurs marqueurs moléculaires sont devenus de plus en plus populaires pour déterminer la structure génétique des populations pour ces fins. Dans ce travail, les auteurs ont comparé la résolution de la structure des populations, parmi des insectes de ruisseaux à une échelle spatiale réduite, obtenue à l'aide de la sous-unité I de la cytochrome *c* oxydase (COI) mitochondriale et de marqueurs mononucléotidiques (SNP) nucléaires. Des individus de trois taxons endémiques—*Coloburiscus humeralis* (Ephemeroptera), *Zelandobius confusus* (Plecoptera) et *Hydropsyche fimbriata* (Trichoptera)—ont été collectés dans des ruisseaux boisés coulant au sein de prairies ouvertes de l'Île du Nord de la Nouvelle-Zélande. Tant les données COI que SNP ont indiqué une faible structure des populations au sein de la zone étudiée, et les petites différences observées parmi ces espèces étaient vraisemblablement dues à leur aptitude à la dispersion. Par exemple, une différenciation génétique à échelle fine entre et parmi les populations de ruisseaux voisins pour le *H. fimbriata* suggèrait un flux génique ; ainsi, la dispersion est peut-être plus limitée chez cette espèce par rapport aux autres. Sur la base des résultats généralement semblables obtenus avec les deux types de marqueurs, les auteurs suggèrent que les marqueurs COI ou SNP peuvent fournir des estimés initiaux valables de la différenciation génétique des populations à échelle réduite chez les insectes de ruisseaux. [Traduit par la Rédaction]

Mots-clés : dispersion, ADNmt, structure des populations, SNP, insectes de ruisseaux

Introduction

Understanding how individuals move between populations across the landscape is an important component of the conservation and management of natural populations and their habitats (Alp et al. 2012; Weston et al. 2020; Venables et al. 2021). Genetic markers can be used to infer levels of gene flow, providing an indirect measure of dispersal and connectivity among populations (Driscoll et al. 2014; Saastamoinen et al. 2018). This approach assumes that when genetic differences among populations are high (measured as genetic

differentiation), gene flow/dispersal is restricted. In contrast, where populations show lower levels of genetic differentiation, they are assumed to have higher rates of dispersal (gene flow) among habitats (Slatkin 1985; Hughes 2007). In stream ecosystems, the extent of gene flow occurring within and among aquatic populations is determined by the dispersal abilities and life histories of resident taxa, hydrological features and dendritic structure of the stream network, and physical nature of the local landscape such as riparian cover (Hughes et al. 2009). Likewise, historical climate change events including glacial periods and hydrological shifts can affect the extent of gene flow, and hence, the patterns of population genetic structure in stream insects (Hotaling et al. 2018; Múrria et al. 2020; Rader et al. 2019; Schröder et al. 2021). Analysis of spatial patterns of genetic differentiation can further determine processes of contemporary population connectivity, and the effects of landscape features on population genetic structure (Wilcock et al. 2007; Alexander et al. 2011; Engelhardt et al. 2011).

Different genetic markers and associated sequencing techniques offer a range of options to investigate population genetic structure. Single locus, maternally inherited mitochondrial DNA (mtDNA) markers, for example, are among the most widely used to assess evolutionary and contemporary processes driving population connectivity in stream insects. In particular, extensive use of the cytochrome *c* oxidase subunit 1 (COI) gene has broadened understanding of species diversity, phylogeography, and dispersal patterns (Monaghan and Sartori 2009). COI data are often obtained through traditional Sanger sequencing (Sanger et al. 1977; Smith et al. 1986), a method characterised by high accuracy, ease-of-use, and rapid turnaround (Hebert et al. 2003; Heather and Chain 2016) that remains useful for initial investigations of genetic differentiation and (or) analysis of a small number of samples (De Cario et al. 2020).

However, current methods offering higher resolution are more suitable for the detection of fine-scale population structure, for example, within watersheds or proximate stream networks (Hotaling et al. 2018; Taylor et al. 2020). Among these, genome-wide single nucleotide polymorphisms (SNPs) are a popular marker of choice following the advent and reduced costs of next generation sequencing (NGS) technologies. Genome-wide SNPs represent mainly neutral, biparental inherited markers and show high polymorphism among individuals (Mardis 2017). Restriction-site Associated DNA sequencing, or RADseq (Davey and Blaxter 2010) has been an increasingly used NGS technique for SNP discovery as part of population genetic studies (e.g., Polato et al. 2017; Trense et al. 2021). RADseq uses restriction enzymes to cut DNA into short fragments, followed by simultaneous sequencing and SNP discovery (Mardis 2008). The main advantages of this method include the: massively parallel sequencing across many individuals at a reduced cost, production of highly reproducible data, and application to species with limited, or no existing sequence data (Baird et al. 2008; Davey and Blaxter 2010). Among natural populations, RADseq data have enabled resolution of fine spatial patterns of genetic variation and recent population-level differentiation (Dussex et al. 2016; Vendrami et al. 2017; Wang et al. 2020), as well as the

identification of recent barriers to gene flow or changes in population structure (Devlin-Durante and Baums 2017).

The popularity of DNA barcoding in aquatic insects has also resulted in an increase in the number of studies exclusively using mtDNA sequences to test for spatial population structure to indirectly infer species' dispersal potential (Hughes et al. 2009). However, mtDNA accumulates mutations at a slower rate that makes them most suitable for examining evolutionary history over longer timescales (Teske et al. 2018). By contrast, SNP data represent mutations that are abundant and widespread across the genome and show a greater influence of processes such as genetic drift in populations (Coates et al. 2011). SNPs, therefore, have a higher information content for fine-scale population structure analysis, and have been recently applied in studies attempting to answer ecological questions about present-day dispersal patterns and influences (e.g., habitat, environmental) on population connectivity (Morin et al. 2004; Storfer et al. 2018). For aquatic insects, the incorporation of both mitochondrial and nuclear DNA markers can improve understanding of species boundaries. For example, in caddisflies within the *Apatania zonella* group, similarity of COI sequences within species and slightly diverged genome wide SNP data between species indicated close relationships among nevertheless genetically distinct taxa (Salokannel et al. 2021). In contrast, distinct COI lineages without differentiated SNP patterns among *Limnephilus* species suggested a lack of cryptic diversity in this taxon (Salokannel et al. 2021). Such combined use of genetic markers may provide a more comprehensive assessment of dispersal and connectivity more generally, particularly since single organelle mtDNA markers can provide limited inference on smaller spatial scales (e.g., Pazmiño et al. 2017; McGaughran et al. 2019).

Here, we examined population connectivity and dispersal patterns for three New Zealand endemic stream insects, comparing the resolution of COI and SNP data for detecting population structure at small spatial scales (∼11 km). We selected one species from each of the commonly encountered freshwater insect orders, Ephemeroptera, Plecoptera, and Trichoptera, based on their putative dispersal abilities: the mayfly *Coloburiscus humeralis* (Coloburiscidae), the stonefly *Zelandobius confusus* (Gripopterygidae), and the caddisfly *Hydropsyche fimbriata* (Hydropsychidae). Previous studies for *C. humeralis* using allozymes and nuclear amplified fragment length polymorphism (AFLP) data have provided conflicting results, with low genetic differentiation among populations even across distant catchments using allozyme data (Hogg et al. 2002), versus significant differentiation detected using AFLP between populations <2 km (Wallace 2013). For other mayflies, dispersal varies significantly among species and is typically associated with strong flight ability, which is in turn determined by species-specific traits, such as adult longevity and wing structure (Malmqvist 2000; Petersen et al. 2004). Local habitat and environmental features may also reduce dispersal capacity, thus influencing genetic structure (Monaghan et al. 2002; Hughes et al. 2009; Alp et al. 2012; Polato et al. 2017). While empirical and genetic studies on dispersal ability are lacking for *Z. confusus*, stoneflies are generally weak fliers (Nelson 1994; Hughes et al. 1999; Sproul et

al. 2014), although recent population genomic research on the New Zealand stonefly *Zelandoperla decorata* suggests that dispersal by flight can occur among streams at regional scales (Dussex et al. 2016). In contrast, fine-scale genetic differentiation (between populations separated by hundreds of meters within the stream) was found in another New Zealand stonefly (*Zelandoperla fenestrata* complex), although this was mainly associated with the evolutionary consequences and ecological drivers of wing reduction (Dussex et al. 2016; McCulloch et al. 2019). Compared to mayflies and stoneflies, empirical studies have shown that caddisflies have higher dispersal capacity, and individuals have been trapped both close to (20 m) and far away from (300 m) the stream channel (Collier and Smith 1997; Petersen et al. 2004; Winterbourn et al. 2007). For *H. fimbriata*, genetic studies indicated a pattern of isolation by distance, with increasing levels of genetic differentiation at increasing spatial scales (Smith and Collier 2001; Smith et al. 2006; Smith and Smith 2009).

Our aim was to compare the resolution of COI and nuclear SNP markers for detecting population structure among streams over small spatial scales. We hypothesised that patterns of population genetic structure would reflect differential dispersal abilities among the studied species and that the genome-wide SNP data would detect finer-scale genetic differentiation relative to the single-locus mtDNA marker. Based on the previous studies, the stonefly *Z. confusus* is expected to show higher genetic differentiation among populations, due to its potential limited dispersal ability when compared to the mayfly *C. humeralis* and the caddisfly *H. fimbriata*.

Materials and methods

Study area and insect collection

Caddisfly larvae and stonefly and mayfly nymphs were collected between December 2017 and January 2019 from 11 sampling sites within three streams (Ngakoaohia, Tawhitiwhiti, and Te Pahu streams), of which two were in neighbouring catchments in Pirongia Forest Park (Fig. 1; Table 1), under collection authorisation from the New Zealand Department of Conservation (permit number 68083-FAU). Part of the Alexandra Volcanic Group, Mount Pirongia was formed by a succession of eruptions between ~2.5 and 1.6 million years ago (Kear 1960; McLeod et al. 2020). The local landscape is characterised by forest fragments surrounded by intense agricultural production, predominantly dairy and sheep farmland. Samples were collected from first- and second-order permanent and stony-bottom streams that all flowed from indigenous forest before entering open farmland.

For each species, approximately 50 individuals per site were collected from three to four sites at intervals of at least 490 m within each of the three streams; and the distance between streams varied from 600 m to 11 km (Table 1). Samples were collected using a kick-net or directly hand-picked from the substrate, and immediately preserved in 95% ethanol. All samples were double-labelled (on the inside and outside of collection vials) and cross-checked on site. *Zelandobius confusus* nymphs were identified using McLellan (1993) and *C. humeralis* and *H. fimbriata* were confirmed using Winterbourn et al.

(2000). The nomenclature of *H. fimbriata* follows Geraci et al. (2010). Following morphological identification, samples were transferred to vials with fresh 95% ethanol and stored at −20 °C until the left rear leg was dissected from each individual using sterilised (flamed) forceps and added to a single well of a 96-well PCR plate for DNA extraction.

Study taxa

The mayfly *C. humeralis* is widespread in New Zealand, inhabiting the underside of stones predominantly in riffles (Wisely 1961). The nymph stage ranges from 12 to 27 months (Harding and Winterbourn 1993). Although no specific information is available on adult longevity for *C. humeralis*, most adult mayfly taxa are comparatively short-lived (with 14 days the maximum recorded for New Zealand species; B. Smith, unpublished data).

The stonefly *Z. confusus* is widely distributed throughout New Zealand (McLellan 1993). The nymph stage lasts 9–12 months (McLellan 1993), while adults can live for a few days to weeks (Collier and Smith 2000), and are often found on the riparian vegetation (Smith and Collier 2000; Winterbourn 2005). In our study, *Z. confusus* nymphs were usually found in leaf packs that had accumulated or become trapped against rocks or woody debris within the stream.

The caddisfly *H. fimbriata* is restricted to the North Island of New Zealand. Larvae are found mostly in native forest streams, where they build fixed retreats attached to stable substrates and spin nets to filter drifting food particles/detritus (Cowley 1978). The larval stage is 9–12 months (Cowley 1978), while adult longevity is approximately 13 days (B. Smith, unpublished data).

DNA extraction, amplification, and sequencing

DNA sequencing of both mtDNA COI and genome-wide SNP markers were conducted at the Canadian Centre for DNA Barcoding (University of Guelph, Canada). Extraction and sequencing of mtDNA followed standard protocols (Ivanova et al. 2006). In brief, DNA was extracted following the AcroPrepTM PALL Glass Fibre plate method using a total mix of 5 mL insect lysis buffer (0.5 mL of Proteinase K, 20 mg/mL per 96-well plate). A 658 bp region of the COI gene was PCR amplified using the primer pair LepF1 and LepR1 (Hebert et al. 2004; Wilson 2012) and 5 μL of the DNA extraction product. PCR thermal cycling conditions were: initial denaturation of samples at 94 °C for 1 min, followed by five cycles of 94 °C for 30 s, 48 °C for 1.5 min, and 72 °C for 1 min. This was followed by 35 cycles of 94 °C for 30 s, 52 °C for 1 min, and 72 °C for 1 min; with a final extension of 72 °C for 10 min. PCR products were cleaned using Sephadex® before being sequenced in forward and reverse directions on an ABI 3730xl DNA Analyser (Applied BiosystemsTM) using the same primers used for PCR amplification.

Specimen images, collection data, raw trace files, and sequence data were uploaded to the Barcode of Life Datasystems (BOLD) database (Ratnasingham and Hebert 2007) and sequences were checked for stop codons as part of the BOLD quality control process. In addition, preliminary neighbour joining phylogenetic trees were generated in the

Fig. 1. Sampling locations showing the 11 sampling sites distributed across three streams in Pirongia (North Island, New Zealand).

BOLD platform to identify and, if present, remove contaminated or misidentified individuals for both COI and SNP datasets (data not shown). All final data are available under dataset DS-EPTNZNI (dx.doi.org/10.5883/DS-EPTNZNI) and cross-referenced to GenBank (accession numbers: OK502554–OK502876).

Table 1. Population locations and site codes, downstream distance within each stream (from the top to the bottom sampling site), and Euclidean distance between streams for all sites analysed in this study.

Stream	Population code	Downstream distance (m)	Euclidean distance	Riparian land cover	Coordinates
Tawhitiwhiti Stream	Stream A		—		
	A1	0		Forest	37°57′33.0″S, 175°05′44.8″E
	A2	700		Pasture	37°57′15.5″S, 175°05′43.9″E
	A3	1100		Pasture	37°57′06.5″S, 175°05′43.9″E
Te Pahu Stream	Stream B		~600 m from Stream A		
	B1	0		Forest	37°57′48.2″S, 175°06′02.1″E
	B2	490		Forest fragment	37°57′35.0″S, 175°06′07.5″E
	B3	620		Pasture	37°57′31.9″S, 175°06′10.4″E
	B4	1140		Replanted	37°57′16.3″S, 175°06′08.1″E
Ngakoaohia Stream	Stream C		~11 km from Stream A, B		
	C1	0		Forest	38°03′10.8″S, 175°05′02.9″E
	C2	550		Forest	38°03′20.2″S, 175°05′16.9″E
	C3	1500		Forest	38°03′35.0″S, 175°05′34.9″E
	C4	3500		Forest	38°03′47.1″S, 175°05′46.7″E

For RADseq library construction, 9–10 individuals were selected from each sampling site for which COI data were also available. DNA was extracted using DNeasy Blood and Tissue Kits (Quiagen Inc., Valencia, CA, USA) following the manufacturer's instructions. DNA was extracted, processed, and sequenced from all individuals together to avoid any potential batch effects on the SNP data. All extractions were visualized on a 1% agarose gel and quantified using an AccuClear UltraHigh Sensitivity dsDNA Quantification Kit (Biotium, Inc.) and SpectraMax M2 microplate reader. Extracted DNA was incubated with the GATC cut site restriction enzyme, *Dpn*II, in 50 μL reactions, for 3 h at 37 °C followed by 20 min at 65 °C (Knapp et al. 2016). Samples were then cleaned with Ampure XP beads in a 1:1.8 ratio of DNA: beads and libraries were generated using KAPA HyperPrep Library Preparation Kit (Roche), including a 350–700 bp size-selection and PCR steps based on manufacturers' recommendations. All libraries were sequenced as 300 bp paired-end reads on the Illumina MiSeq platform at the Biodiversity Genomics Facility at the University of Guelph. Raw and demultiplexed reads for each sample were then used for downstream analysis.

Mitochondrial DNA (COI) data analysis

Sequences were aligned in ClustalX v. 2.1 (Larkin et al. 2007) and ends were trimmed in Jalview v. 2.11.1.3 (Waterhouse et al. 2009) so that all sequences were of equal length. Downstream data analysis was conducted both before and after trimming to ensure consistency of results (data not shown). Haplotype (h) and nucleotide (π) diversity indices (Nei 1987) were calculated in ARLEQUIN v. 3.0 (Sd et al. 2000). For visualization of the genetic structure, haplotype networks assigning individual sequences to their respective sampling site locations were constructed. Analyses were conducted in RStudio v 1.3.1 (Studio team 2020) using the packages ape v. 5.5 (Paradis and Schliep 2018) and pegas v. 1.0 (Goudet 2005). Global and pairwise measures of genetic distance (Φ_{ST}) among populations were estimated among all 11 sampling sites to infer patterns of population structure for each species. These statistics were calculated using pairwise comparisons and analysis of molecular variance (AMOVA; Excoffier et al. 1992) in ARLEQUIN, with the statistical significance of variance components obtained from 10 000 permutations.

Single nucleotide polymorphisms (SNPs) data analysis

Before processing, the quality of raw Illumina data was checked using FastQC v. 0.11.8 (Andrews 2010). Example FastQC reports for raw and processed data are shown in Fig. S1. IPYRAD v. 0.7.28 (Eaton and Overcast 2020) was used to filter and remove low-quality data, identify homology among reads through de novo assembly, make SNP calls, and format output files for each species dataset. Following recommendations of the software developer, only forward reads (R1) were used in the pipeline to enhance computational performance (due to the presence of a high amount of adapter contamination in combination with the reads length interfering with read processing). Additionally, to determine final filtering settings in IPYRAD, certain default settings were changed to explore the effect of several parameters on the final amount and quality of reads, including filter_min_trim_len (which sets the final minimum length after filtering; we tested values of 60 and 150, default: 35) and trim_reads (trims raw read edges; we tested values of 10, 150, 0, 0; 10, −140, 0, 0; and 10, 150, 0, 0 to set a minimum length of 150 bp, default: 0, 0, 0, 0). Although these filters had little effect in terms of removing low-quality data or changing preliminary phylogenetic tree structure, the number of SNPs retained varied when restricting the minimal read length or trimming all reads to a uniform length. Thus, to obtain the maximum number of SNPs while removing low-quality data, reads were processed with the following non-default parameter settings: filter_adapters (2, where adapters are removed), filter_min_trim_len (60), and trim_reads (10, −140, 0, 0). The total number of retained reads and filtered data for the three species are summarised in Table S1.

The SNP dataset was filtered (Table S2) using VCFTOOLS v. 0.1.13 (Danecek et al. 2011), with –missing-indv, –max-missing-count and –maf parameters applied to explore the effects of missing genotypes and minor allele frequency (MAF) on a per-individual basis. We proceeded with two datasets: (1) a less conservative dataset created by filtering individuals with >98% missing data (Supplementary material—VCF file A1); and (2) a more conservative dataset created by setting missing genotype data across all individuals to 20% and applying a MAF cut-off of 5% in addition to applying the 98% missing data threshold (Supplementary material—VCF file A2). As no significant differences were observed in population structure analysis among the two filtered datasets, we hereafter present results for dataset (2).

Principal component analysis (PCA) was performed using PLINK v. 3 (Purcell et al. 2007) to examine population differentiation based on orthogonal transformation of the SNP data. PCA output files were then used to plot results using the R package tidyverse v. 1.3.1 (Wickham et al. 2019). We also used PLINK to convert VCF files to a format suitable for use in the clustering software, FastStructure v. 1.0 (Raj et al. 2014), which was used to identify admixture proportions among individuals and populations. In FastStructure, we tested 1–11 genetic clusters (depending on the availability of data from each sampling site/population) by specifying the *K* parameter (where *K* is the number of genetic clusters), and ran five replicates for each *K*-value, using a simple prior. Results for each *K*-value were visualised, and model complexity chosen using the distruct.py and chooseK.py scripts from the FastStructure program, respectively.

Finally, pairwise comparisons of the genetic distance between all sampling sites were calculated using Φ_{ST} (Weir and Cockerham 1984) in the R package StAMPP v. 1.6.2 (Pembletom et al. 2013). Global Φ_{ST} over all populations was estimated using AMOVA in the R package poppr v. 2.8.1 (Kamvar et al. 2014).

Results

Coloburiscus humeralis (Ephemeroptera)

COI

Trimmed sequences of 549 bp were obtained for 78 (46%) of the 171 processed specimens of *C. humeralis* (Table 2). A total of 10 haplotypes were identified and haplotype diversity ranged from 0.181 to 0.809 across the 11 sampling sites. Nucleotide diversity was low for all sites (0.001–0.004; Table 2). Seven of the 10 haplotypes were singletons (found in a single individual), with only three haplotypes found in five or more individuals (Fig. 2a). Haplotype 1 (H1) was the most frequent ($n = 59$) and was present at all sampling sites (Table 4; Fig. 2a). Consistent with the lack of the population structure observed in the haplotype network, AMOVA analysis revealed limited genetic differentiation across the study area (global $\Phi_{ST} = 0.010$, $P = 0.31$; Table 3). Similarly, most of the pairwise comparisons of Φ_{ST} between sampling sites did not differ significantly from zero (Table 5).

SNPs

The RADseq library produced a total of 43 363 740 reads from 19 specimens collected at seven sampling sites within the Pirongia range (Table 5). After SNP identification and the application of a conservative filtering pipeline, 41 029 SNPs from all 19 individuals were recovered. A visual representation of the sequencing depth per sample and the location of missing data in the SNP matrix (72.8%) is provided in Fig. S2. Pairwise comparisons of Φ_{ST} were not calculated for most populations due to small sample sizes ($n < 3$; Table 5). However, genetic differentiation among the remaining populations was low (global $\Phi_{ST} = 0.017$, $P = 0.37$), and consistent with the COI data (Table 3). The FastStructure analysis indicated that values of K between 1 and 2 maximized the marginal likelihood. Results are provided for $K = 2$, showing two major genetic groups across seven sampling sites, with each genetic group present in all streams and a lack of admixture found within individuals (Fig. 3). The first two components of the PCA explained 18.1% of the cumulative total variance and no distinct clustering was observed between sampling sites or streams (Fig. 2a).

Zelandobius confusus (Plecoptera)

COI

Sequences of 658 bp were obtained for 101 (92%) of the 110 processed specimens from all 11 sites (Table 2). Analysis of *Z. confusus* revealed 10 haplotypes and haplotype diversity was moderate to high (0.377–1.000), while nucleotide diversity was relatively low (0.002–0.010; Table 2). Six of the 10 haplotypes were singletons and four haplotypes were found in five or more individuals (Fig. 2b; Table 4). The central and putatively ancestral haplotype in the network (H1, $n = 29$) occurred in all three streams (but not all sampling sites), while the most frequent haplotype (H2, $n = 49$) occurred at all sampling sites across the study area (Fig 2b; Table 4). AMOVA re-

vealed low genetic differentiation among all sampling sites (global $\Phi_{ST} = 0.038$, $P = 0.10$; Table 3), and most of the pairwise comparisons were not significant (Table 5).

SNPs

The RADseq library produced a total of 40 549 912 reads from 73 specimens collected at 11 sampling sites within Pirongia (Table 5). Filtering resulted in 932 SNPs from all 73 individuals. Graphics showing the sequencing depth per sample and the location of missing data in the SNP matrix (65%) are available in Fig. S3. Global Φ_{ST} analysis revealed a lack of population differentiation ($\Phi_{ST} = 0.003$; Table 3). Pairwise comparisons of Φ_{ST} identified just two population pairs with significant (but weak) genetic differentiation (B1 vs. C2 $\Phi_{ST} = 0.041$, $P = 0.03$; C1 vs. C2 $\Phi_{ST} = 0.042$, $P = 0.03$), suggesting high gene flow within Pirongia, as per the COI data. The FastStructure analysis indicated that values of K between 2 and 3 maximized the marginal likelihood. For $K = 2$, genetic clusters spanned all sampling sites, with admixture shown in some individuals. For $K = 3$, the third group was shared in a small proportion between a few individuals from sites B1, B2, and C4 (Fig. 3). Two clusters were observed in the PCA analysis, with the first two components explaining 38.5% of cumulative variance and showing a lack of strong structure between sampling sites or streams (Fig. 2b). These results were consistent with the lack of geographic structure observed in the mtDNA haplotype network.

Hydropsyche fimbriata (Trichoptera)

COI

Sequences of 552 bp were obtained for 107 (43%) of the 249 processed specimens from the 11 sites (Table 2). Analysis of *H. fimbriata* revealed 18 haplotypes, high haplotype diversity (0.400–1.000), and moderate nucleotide diversity (0.004–0.010). Both of these diversity measures were highest for *H. fimbriata* compared to the other two species. Eight of the 18 haplotypes identified were singletons, with six haplotypes found in more than three individuals (Fig. 2c; Table 4). Among the most frequent haplotypes, H3 ($n = 20$) and H16 ($n = 18$) were distributed across sampling sites from all three streams, whereas H8 ($n = 15$) and H12 ($n = 23$) were only found in populations from the adjacent Streams A and B. AMOVA analysis showed significant but weak genetic differentiation between all sampling sites ($\Phi_{ST} = 0.092$, $P = 0.009$; Table 3). Pairwise comparisons of Φ_{ST} between 19 population pairs were significantly different from zero, including the most distant sampling sites within each of two streams: A1–A3 and B1–B4 (Table 5).

SNPs

The RADseq library produced a total of 22 946 236 reads from 93 specimens collected at ten sampling sites (Table 5). After filtering, 202 SNPs from 92 specimens were retained. Graphics showing the sequencing depth per sample and the location of missing data in the SNP matrix (46.8%) are avail-

Table 2. Sample sizes and genetic diversity of COI sequences for each of three study species *Coloburiscus humeralis*, *Zelandobius confusus,* and *Hydropsyche fimbriata*.

Population code	*Coloburiscus humeralis* n(x) h	π	*Zelandobius confuses* n(x) h	π	*Hydropsyche fimbriata* n(x) h	π
A1	4(1) —	—	10(3) 0.377 ± 0.181	0.002	7(4) 0.714 ± 0.180	0.007
A2	10(3) 0.377 ± 0.181	0.001	10(4) 0.533 ± 0.180	0.004	10(7) 0.911 ± 0.077	0.010
A3	8(4) 0.642 ± 0.184	0.001	11(3) 0.563 ± 0.134	0.004	12(8) 0.893 ± 0.077	0.010
Total Stream A	22(5) 0.407 ± 0.128	0.001	31(4) 0.475 ± 0.096	0.004	29(13) 0.889 ± 0.038	0.010
B1	11(2) 0.181 ± 0.143	0.001	10(3) 0.644 ± 0.101	0.003	11(5) 0.781 ± 0.107	0.005
B2	3(1) —	—	10(7) 0.911 ± 0.077	0.006	3(3) 1.000 ± 0.272	0.009
B3	5(1) —	—	10(4) 0.711 ± 0.117	0.004	10(6) 0.866 ± 0.085	0.009
B4	4(2) 0.500 ± 0.265	0.001	5(2) 0.400 ± 0.237	0.004	33(9) 0.820 ± 0.107	0.009
Total Stream B	23(3) 0.170 ± 0.103	0.001	35(9) 0.751 ± 0.049	0.004	57(13) 0.835 ± 0.026	0.008
C1	9(2) 0.222 ± 0.166	0.001	11(5) 0.763 ± 0.106	0.004	9(4) 0.750 ± 0.112	0.006
C2	8(2) 0.428 ± 0.168	0.001	10(2) 0.466 ± 0.131	0.002	5(4) 0.900 ± 0.161	0.006
C3	9(4) 0.750 ± 0.878	0.004	12(4) 0.712 ± 0.105	0.004	3(2) 0.666 ± 0.314	0.004
C4	7(4) 0.809 ± 0.129	0.002	2(2) 1.000 ± 0.500	0.010	4(2) 0.500 ± 0.265	0.005
Total Stream C	33(7) 0.580 ± 0.093	0.002	35(5) 0.682 ± 0.048	0.003	21(6) 0.728 ± 0.079	0.005

Notes: *n*, sample size; *x*, number of haplotypes; *h*, haplotype diversity; *π*, nucleotide diversity. A dash(—) indicates genetic diversity is not available (for populations that contain only one haplotype).

Table 3. Hierarchical analysis of molecular variance (AMOVA) for COI sequences for three species *Coloburiscus humeralis*, *Zelandobius confusus*, and *Hydropsyche fimbriata*.

	COI data Variance (%)	Φ_{ST}	SNP data Variance (%)	Φ_{ST}
Coloburiscus humeralis				
Among all sites	1.01		1.21	
Within all sites	98.99	0.010	98.79	0.017
Zelandobius confusus				
Among all sites	3.82		1.12	
Within all sites	96.18	0.038	98.88	0.003
Hydropsyche fimbriata				
Among all sites	9.23		9.61	
Within all sites	90.77	0.092*	90.38	0.096*

*P-value < 0.01.

able in Fig. S4. AMOVA revealed low but significant population differentiation (global Φ_{ST} = 0.096, P = 0.007; Table 3). Pairwise comparisons of Φ_{ST} values were significantly high for ten pairs of populations, and significant but low for 13 population pairs (Table 5), suggesting that gene flow might be more limited in this species. Consistent with the COI data, pairwise comparisons indicated significant genetic differentiation between the most distant within-stream pairs (A1 vs. A3, B1 vs. B4). Significant, low genetic differentiation was also observed between the most distant sites within Stream C (C1 vs. C4; Φ_{ST} = 0.079, P = 0.009). FastStructure analysis indicated that a value of K = 5 maximized the marginal likelihood. The prevalence of each genetic group differed between the three streams and low levels of admixture were observed within individuals across all sampling sites (Fig. 3). The first two PCA components explained 54.7% of the total cumulative variance and, unlike the PCAs for the other two species, clear

clusters could be distinguished (Fig. 2). In particular, individuals from Stream C sampling sites tended to group together more in one cluster than the others, but the four clusters showed a lack of clear geographic partitioning overall.

Discussion

Using both COI and SNP markers, we identified a lack of strong differentiation among populations for the three aquatic insect species analysed, suggesting high population connectivity (gene flow) between sampling sites and stream localities within our study area at a spatial scale of ~11 km. However, the genetic markers detected minor differences in the degree and pattern of population genetic differentiation observed for each of the study species that may reflect their different dispersal abilities.

For the common mayfly *C. humeralis*, COI analysis showed a homogeneous distribution of haplotypes across all populations within and between the three study streams. Results of the SNP analysis were consistent with the COI data, with no population structure revealed by the Bayesian analysis (FastStructure) or PCA. This suggests high population connectivity at small spatial scales across a variety of land use types, consistent with a previous allozyme study that showed very low levels of differentiation between *C. humeralis* populations at different spatial scales in the North and South Islands of New Zealand (Hogg et al. 2002). The historic distribution of *C. humeralis* across both islands of New Zealand suggests widespread dispersal is possible across distant catchments, although ongoing habitat fragmentation could restrict future connectivity among populations. Analysis across distant catchments is necessary to confirm the inter-regional dispersal suggested by the previous allozyme data.

For the stonefly *Z. confusus*, both COI and SNP data showed a general pattern of low genetic differentiation between sam-

Fig. 2. Visual representation of genetic relationships among individuals and sampling locations for (*a*) *Coloburiscus humeralis*, (*b*) *Zelandobius confusus,* and (*c*) *Hydropsyche fimbriata*. Individuals are colour-coded according to collection locality and population label (Streams A1, A2, etc.). The top panel shows haplotype networks based on cytochrome oxidase subunit 1 (COI) sequence data—each pie chart represents a single haplotype, with the size proportional to the frequency of individuals containing that particular haplotype at each site where samples were collected; dashes indicate missing mutational steps between haplotypes. Lower panel shows a principal component analysis (PCA) based on orthogonal transformation of SNP data—the percentage of variation explained by each principal component is indicated on the axes.

pling sites, indicating substantial population connectivity across the study area. These results are consistent with SNP data for another New Zealand stonefly species *Zelandoperla decorata*, which shows low genetic differentiation within and between parallel streams separated by ∼10 km, indicating overland dispersal at small spatial scales over fragmented or non-forested landscapes for this medium-sized stonefly (Dussex et al. 2016). Our results also support previous direct studies of dispersal using stable isotope enrichment in stoneflies, which indicate dispersal between streams (Briers et al. 2004), although these taxa are generally considered weak fliers (Briers et al. 2002; Sproul et al. 2014).

For the caddisfly *H. fimbriata*, common COI haplotypes were not found at all sites and individual membership in genetic clusters differed among locations for the SNP analysis. Significant pairwise Φ_{ST} for both markers—even from the upper and lower sampling sites within each of the three study streams (as observed within Streams A and B for the COI data and Streams A, B, and C for the SNP data)—indicated limited connectivity within and between adjacent streams and suggests that dispersal in this species may be limited even within the same stream channel. Collectively, these data suggest more limited gene flow among populations of *H. fimbriata* at

small spatial scales relative to that found for either *C. humeralis* or *Z. confusus*. Hydropsychid larvae, such as *H. fimbriata*, are sedentary filter-feeders that build and attach a "fixed retreat" shelter of silk and organic matter to a stable substrate. The larval investment in shelter construction may reduce the propensity for drift, and this may partially explain the high genetic differentiation between populations. Settlement onto substrata may also be related to limited larval downstream drift, as suggested by Downes and Lancaster (2010) for another Hydropsychid caddisfly. Previous landscape genetics studies on different riverine species have shown that human-driven fragmentation for agricultural purposes can affect population structure and may restrict dispersal in these ecosystems (Wilcock et al. 2007; Blanchet et al. 2010; Lean et al. 2017). Accordingly, local landscape features which constrain dispersal could further explain the observed population genetic structure for this species. *Hydropsyche fimbriata* larvae are mostly restricted to cool streams in native forests, and fragmented stream corridors provide less optimal microclimate conditions (e.g., lower humidity) that may limit dispersal of adults outside of forest areas. We found the lowest Φ_{ST} values between sites within Stream C (Te Pahu), and individuals from this stream, which is covered by forest riparian

Table 4. Haplotype frequencies based on COI data for each of three study species: *Coloburiscus humeralis*, *Zelandobius confusus*, and *Hydropsyche fimbriata*.

Coloburiscus humeralis

Population	A1	A2	A3	B1	B2	B3	B4	C1	C2	C3	C4	n
H1	4	8	6	10	3	5	3	8	6	4	3	60
H2		1										1
H3			1						2		2	5
H4				1								1
H5										1		1
H6											1	1
H7											1	1
H8		1	1					1		3		6
H9							1					1
H10										1		1

Zelandobius confuses

Population	A1	A2	A3	B1	B2	B3	B4	C1	C2	C3	C4	n
H1	1	1	1	5	3	5		3	7	3		29
H2	8	7	7	4	1	3	4	5	3	6	1	49
H3					1							1
H4					1							1
H5					2			1		2		5
H6								1				1
H7	1	1	3	1	1	1	1	1		1	1	12
H8					1							1
H9					1							1
H10		1										1

Hydropsyche fimbriata

Population	A1	A2	A3	B1	B2	B3	B4	C1	C2	C3	C4	n
H1				1								1
H2		1		1		1			1			4
H3			1	5	1	3	5	4	1			20
H4											1	1
H5			1				1					2
H6							1					1
H7				2								2
H8		1	3		1	1	9					15
H9		1										1
H10	1											1
H11	1											1
H12	4	3	1	2		3	10					23
H13			1									1
H14	1	2			1		3	1		1		9
H15						1						1
H16		1	4			1	2	3	2	2	3	18
H17		1	1				1					3
H18							1	1	1			3

Note: Grey shaded haplotypes are private. *n*, number of individuals sharing each haplotype.

vegetation, tended to form a single genetic group in the PCA. In contrast, greater genetic differentiation was found within the other more fragmented streams, where riparian vegetation is mainly dominated by low growing pasture grasses

**Table 5.** Pairwise genetic distances (Φ_{ST}) between sampling sites for each of the three study species.

Coloburiscus humeralis

	A1(0)	A2(4)	A3(2)	B1(5)	B2(0)	B3(2)	B4(2)	C1(3)	C2(0)	C3(1)	C4(0)
A1	—	0.052*	NA	NA	NA	NA	NA	NA	NA	NA	NA
A2	0.000	—	NA	NA	NA	NA	NA	0.006	NA	NA	NA
A3	0.000	0.000	—	NA	NA	NA	NA	NA	NA	NA	NA
B1	0.000	0.006	0.018	—	NA	NA	NA	0.000	NA	NA	NA
B2	0.000	0.000	0.000	0.000	—	NA	NA	NA	NA	NA	NA
B3	0.000	0.000	0.000	0.000	0.000	—	NA	NA	NA	NA	NA
B4	0.000	0.007	0.079	0.189	0.000	0.063	—	NA	NA	NA	NA
C1	0.116	0.000	0.000	0.002	0.000	0.000	0.145	—	NA	NA	NA
C2	0.020	0.046	0.000	0.124	0.000	0.063	0.127	0.105	—	NA	NA
C3	0.000	0.003	0.000	0.082*	0.000	0.000	0.000	0.007	0.063	—	NA
C4	0.000	0.027	0.000	0.124*	0.000	0.023	0.059	0.097*	0.000	0.054	—

Zelandobius confuses

	A1(6)	A2(9)	A3(9)	B1(9)	B2(8)	B3(7)	B4(1)	C1(4)	C2(9)	C3(9)	C4(2)
A1	—	0.000	0.000	0.025	0.019	0.000	NA	0.017	0.007	0.007	NA
A2	0.000	—	0.000	0.024	0.016	0.000	NA	0.025	0.014	0.000	NA
A3	0.000	0.000	—	0.000	0.021	0.000	NA	0.000	0.000	0.000	NA
B1	0.111	0.018	0.005	—	0.000	0.000	NA	0.000	0.041*	0.000	NA
B2	0.319***	0.195	0.145*	0.074	—	NA	NA	NA	NA	NA	NA
B3	0.167	0.064	0.044	0.000	0.014	—	NA	0.000	0.004	0.000	NA
B4	0.000	0.000	0.000	0.053	0.216	0.104	—	NA	NA	NA	NA
C1	0.067	0.000	0.000	0.000	0.066	0.000	0.011	—	0.042*	0.000	NA
C2	0.277*	0.162	0.146	0.000	0.121	0.000	0.258*	0.000	—	0.011	NA
C3	0.041	0.000	0.000	0.000	0.090	0.000	0.000	0.000	0.213	—	NA
C4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	—

Hydropsyche fimbriata

|	A1(7)	A2(9)	A3(7)	B1(10)	B2(0)	B3(9)	B4(9)	C1(11)	C2(10)	C3(10)	C4(10)
A1	—	0.080**	0.327***	0.248***	NA	0.075	0.486***	0.440***	0.273***	0.320***	0.252***
A2	0.106	—	0.016	0.000	NA	0.000	0.143***	0.106***	0.087***	0.085**	0.000
A3	0.276***	0.000	—	0.000	NA	0.075	0.000	0.015	0.054*	0.000	0.050
B1	0.391**	0.168*	0.127	—	NA	0.000	0.086**	0.077***	0.092***	0.089**	0.015
B2	0.341*	0.000	0.000	0.070	—	NA	NA	NA	NA	NA	NA
B3	0.168*	0.000	0.000	0.005	0.000	—	0.295***	0.281***	0.220**	0.164**	0.000
B4	0.166*	0.000	0.008	0.081*	0.000	0.000	—	0.000	0.020	0.000	0.105*
C1	0.432***	0.099	0.000	0.156*	0.000	0.036	0.108*	—	0.013	0.000	0.079**
C2	0.394**	0.030	0.000	0.182*	0.000	0.000	0.068	0.000	—	0.000	0.042
C3	0.467*	0.118	0.033	0.494***	0.182	0.195	0.219***	0.063	0.000	—	0.047
C4	0.509**	0.200*	0.030	0.411***	0.195	0.182	0.232***	0.023	0.000	0.000	—

Notes: Results for SNP and COI data presented in the upper and lower diagonals, respectively. SNP data sample size for each site is provided in parentheses. Significant values are shown in bold (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$). NA indicates values that were not calculated (for populations with ≤2 individuals).

and weed species. Thus, four diverged genetic groups were observed in the PCA and these groups somewhat correlated with the four most frequent haplotypes in the COI data: one group consisted of individuals from all three streams; one group consisted predominately of individuals from Stream C; and two groups consisted exclusively or predominantly of individuals from Streams A and B only. Further landscape genetics analyses are needed to determine the relevance of riparian land cover and thus, habitat preference, in shaping the spatial genetic structure in *H. fimbriata* populations. In particular, environmental variation (e.g., via altitudal differences), habitat adaptation, and evolutionary history can result in distinct cryptic lineages, as revealed previously for other caddisflies species groups (Previsic et al. 2014; Saito et al. 2018), although our data do not indicate that the diverged groups are reproductively isolated. Local extinction–recolonisation dynamics can also increase genetic drift and differentiation among populations when migration is limited (Slatkin 1977; Hastings and Harrison 1994). For example, founder events occur in temporary streams if the population is founded from

Fig. 3. Visualisation of population structure and admixture of the study species from FastStructure software. Results are presented for *K*-values of 2 for *C. humeralis*, 2 and 3 for *Z. confusus*, and 5 for *H. fimbriata*, with colours in each panel representing the assignment of individuals to each genetic cluster. Single bars with >1 colour indicate admixture of that particular individual (i.e., sharing of genetic structure across more than one genetic group). Each sampling site is designated by its label (Streams A1, A2, etc.).
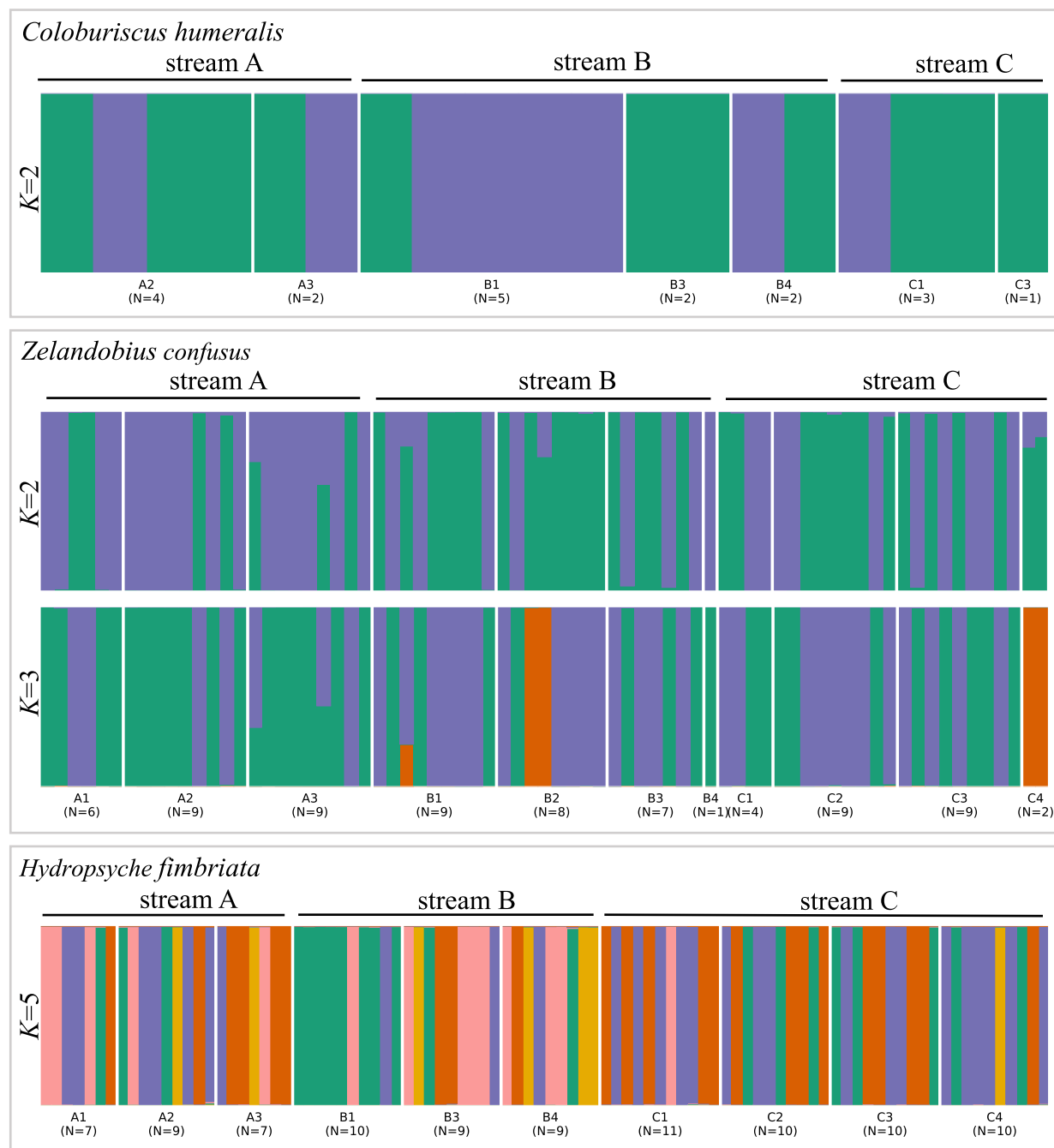
a few individuals of low genetic diversity from source populations and thus affect population structure, as previously reported in the caddisfly *Hydropsyche sitalai* (Múrria et al. 2010). In our study, the small geographic area shows differing vegetation cover but is not characterised by extreme environmental or habitat variation. Nevertheless, a comprehensive phylogenetic analysis, combined with ecology and distributional expansion knowledge, are necessary to further evaluate the

possibility of cryptic species as well as stochastic events in driving population divergence of *H. fimbriata*. Ploidy variation in the sequences could also explain the observed divergent groups, although to our knowledge, this has not been identified in aquatic insects, including for nuclear markers.

Overall, our combined COI and SNP data suggests limited population structure and non-limited dispersal at small spatial scales within and between neighbouring catchments. The

small differences we observed in patterns of genetic differentiation among the three species—suggesting higher population connectivity for *C. humeralis* and *Z. confusus* compared to *H. fimbriata*—highlight the potential influence of their dispersal abilities and (or) landscape features, such as the riparian land cover (forested vs. open pasture) in shaping connectivity among populations. However, we caution that smaller sample sizes and lower data quality for the SNP datasets may have limited their resolution in determining population genetic structure within our study area. We note that genetic differentiation within the stream channel was slightly more pronounced in the SNP than the COI dataset. However, the generally consistent results between markers suggest that either COI or SNPs can independently offer suitable estimates of population differentiation at small spatial scales. Used in combination, they could strengthen inferences of population structure for estimates of contemporary gene flow and elucidation of fine-scale dispersal patterns that might be otherwise confounded by the resolution power and specific characteristics of a given single marker. Extending this approach to a landscape genetics context with increasing geographic coverage will further enhance our knowledge of genetic structure and landscape influences on aquatic insect dispersal and population connectivity.

Acknowledgements

Article information

Copyright

Data availability
Mitochondrial DNA sequences: BOLD dataset DS-EPTNZNI (dx.doi.org/10.5883/DS-EPTNZNI) and GenBank under accession numbers OK502554–OK502876. Full SNP data are provided in VCF format in the Supplementary material (A1 = less conservative; and A2 = more conservative; results from which are presented in the manuscript).

Author information

Author ORCIDs
Vanessa de Araujo Barbosa https://orcid.org/0000-0002-8338-5377

Author notes
Ian Hogg served as an Associate Editor at the time of manuscript review and acceptance; peer review and editorial decisions regarding this manuscript were handled by Sarah Adamowicz.

Author contributions
Project design: VAB, EG, BS, IH; aquatic insect collection: VAB, EG, BS; aquatic insect morphological identification: VAB, BS; data analysis and interpretation: VAB, AM; manuscript preparation: VAB. All authors reviewed, edited, and approved the manuscript.

Competing interests
The authors declare that there are no competing interests.

Supplementary material

Supplementary data are available with the article at https://doi.org/10.1139/gen-2022-0021.

References

Alexander, L.C., Hawthorne, D.J., Palmer, M.A., and Lamp, W.O. 2011. Loss of genetic diversity in the North American mayfly *Ephemerella invaria* associated with deforestation of headwater streams. Freshw. Biol. **56**(7): 1456–1467. doi:10.1111/j.1365-2427.2010.02566.x.

Alp, M., Keller, I., Westram, A.M., and Robinson, C.T. 2012. How river structure and biological traits influence gene flow: a population genetic study of two stream invertebrates with differing dispersal abilities. Freshw. Biol. **57**(5): 969–981. doi:10.1111/j.1365-2427.2012.02758.x.

Andrews, S. 2010. FastQC: a quality control tool for high throughput sequence data.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L. Lewis, Z.A., et al. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS ONE, **3**(10): e3376. doi:10.1371/journal.pone.0003376. PMID: 18852878.

Bohonak, A.J. 1999. Effect of insect-mediated dispersal on the genetic structure of postglacial water mite populations. Heredity, **82**: 451–461. doi:10.1038/sj.hdy.6885050. PMID: 10383664.

Briers, R.A., Cariss, H.M., and Gee, J.H.R. 2002. Dispersal of adult stoneflies (Plecoptera) from upland streams draining catchments with contrasting land-use. Arch. Hydrobiol. **4**: 627–644. doi:10.1127/archiv-hydrobiol/155/2002/627.

Briers, R.A., Gee, J.H.R., Cariss, H.M., and Geoghegan, R. 2004. Interpopulation dispersal by adult stoneflies detected by stable isotope enrichment. Freshw. Biol. **49**(4): 425–431. doi:10.1111/j.1365-2427.2004.01198.x.

Brittain, J. 2003. Biology of Mayflies. Annu. Rev. Entomol. **27**: 119–147. doi:10.1146/annurev.en.27.010182.001003.

Coates, B., Bayles, D., Wanner, K., Robertson, H., Hellmich, R., and Sappington, T. 2011. The application and performance of single nucleotide polymorphism markers for population genetic analyses of lepidoptera. Front. Genet. **2**:38. doi10.3389/fgene.2011.00038. PMID: 22303334.

Collier, K.J., and Smith, B.J. 1997. Dispersal of adult caddisflies (Trichoptera) into forests alongside three New Zealand streams. Hydrobiologia, 361: 53–65. doi:10.1023/A:1003133208818.

Collier, K.J., and Smith, B.J. 2000. Interactions of adult stoneflies (Plecoptera) with riparian zones I. Effects of air temperature and humidity on longevity. Aquat. Insects, 22(4): 275–284. doi:10.1076/0165-0424(200010)22:4;1-Y;FT275.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E. DePristo, M. A., et al. 2011. The variant call format and VCFtools. Bioinformatics (Oxford, England), 27(15): 2156–2158. doi:10.1093/bioinformatics/btr330. PMID: 21653522.

Davey, J.W., and Blaxter, M.L. 2010. RADSeq: next-generation population genetics. Briefings in. Funct. Genomics, 9(5–6): 416–423. doi:10.1093/bfgp/elq031. PMID: 21266344.

De Cario, R., Kura, A., Suraci, S., Magi, A., Volta, A. Marcucci R., et al. 2020. Sanger validation of high-throughput sequencing in genetic diagnosis: still the best practice? Front Genet. 11: 592588. doi:10.3389/fgene.2020.592588.

Devlin-Durante, M.K., and Baums, I.B. 2017. Genome-wide survey of single-nucleotide polymorphisms reveals fine-scale population structure and signs of selection in the threatened Caribbean elkhorn coral, *Acropora palmata*. PeerJ, 5: e4077. doi:10.7717/peerj.4077. PMID: 29181279.

Downes, B.J., and Lancaster, J. 2010. Does dispersal control population densities in advection-dominated systems? A fresh look at critical assumptions and a direct test. J. Anim. Ecol. 79(1): 235–248. Available from. doi:10.1111/j.1365-2656.2009.01620.x. PMID: 19840172.

Driscoll, D.A., Banks, S.C., Barton, P.S., Ikin, K., Lentini, P. Lindenmayer, D.B., et al. 2014. The trajectory of dispersal research in conservation biology. Systematic Review. PLoS ONE, 9(4): e95053. doi:10.1371/journal.pone.0095053.

Dussex, N., Chuah, A., and Waters, J.M. 2016. Genome-wide SNPs reveal fine-scale differentiation among wingless alpine stonefly populations and introgression between winged and wingless forms. Evolution, 70(1): 38–47. doi:10.1111/evo.12826. PMID: 26614689.

Eaton, D., and Overcast, I. 2020. Ipyrad: interactive assembly and analysis of RADseq datasets. Bioinformatics (Oxford, England), 36(8): 2592–2594. doi:10.1093/bioinformatics/btz966.

Ekblom, R., and Galindo, J. 2011. Applications of next generation sequencing in molecular ecology of non-model organisms. Heredity, 107(1): 1–15. doi:10.1038/hdy.2010.152. PMID: 21139633.

Elbrecht, V., Feld, C.K., Gies, M., Hering, D., Sondermann, M., Tollrian, R., and Leese, F. 2014. Genetic diversity and dispersal potential of the stonefly *Dinocras cephalotes* in a central European low mountain range. Freshw. Sci. 33(1): 181–191. doi:10.1086/674536.

Emerson, K.J., Merz, C.R., Catchen, J.M., Hohenlohe, P.A., Cresko, W.A., Bradshaw, W.E., and Holzapfel, C.M. 2010. Resolving postglacial phylogeography using high-throughput sequencing. Proc. Natl. Acad. Sci. U.S.A. 107(37): 16196. doi:10.1073/pnas.1006538107.

Engelhardt, C.H.M., Haase, P., and Pauls, S. U. 2011. From the Western Alps across Central Europe: postglacial recolonisation of the tufa stream specialist *Rhyacophila pubescens* (Insecta, Trichoptera). Front. Zool. 8:10. doi:10.1186/1742-9994-8-10. PMID: 21569621.

Excoffier, L., Smouse, P.E., and Quattro, J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics, 131(2): 479–491. doi:10.1093/genetics/131.2.479. PMID: 1644282.

Geraci, C.J., Zhou, X., Morse, J. C., and Kjer, K.M. 2010. Defining the genus *Hydropsyche* (Trichoptera: Hydropsychidae) based on DNA and morphological evidence. J. North Am. Benthol. Soc. 29(3): 918–933. doi:10.1899/09-031.1.

Goudet, J. 2005. Hierfstat, a package for R to compute and test hierarchical *F*-statistics. Mol. Ecol. Notes, 5(1): 184–186. doi:10.1111/j.1471-8286.2004.00828.x.

Harding, J., and Winterbourn, M. 1993. Life history and production of *Coloburiscus humeralis* (Ephemeroptera: Oligoneuriidae) in two South Island high-country streams, New Zealand. New Zeal. J. Mar. Freshw. Res. 27: 445–451. doi:10.1080/00288330.1993.9516586.

Hastings, A., and Harrison, S. 1994. Metapopulation dynamics and genetics. Annu. Rev. Ecol. Syst. 25: 167–188. doi:10.1146/annurev.es.25.110194.001123.

Head, S.R., Komori, H.K., LaMere, S.A., Whisenant, T., Van Nieuwerburgh, F., Salomon, D.R., and Ordoukhanian, P. 2014. Library construction for next-generation sequencing: overviews and challenges. Biotechniques, 56(2): 61. doi:10.2144/000114133. PMID: 24502796.

Heather, J.M., and Chain, B. 2016. The sequence of sequencers: the history of sequencing DNA. Genomics, 107(1): 1–8. doi:10.1016/j.ygeno.2015.11.003. PMID: 26554401.

Hebert, P.D.N., Penton, E.H., Burns, J.M., Janzen, D. H., and Hallwachs, W. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proc. Natl. Acad. Sci. U.S.A. 101(41): 14812–14817. doi:10.1073/pnas.0406166101. PMID: 15465915.

Hebert, P.D.N., Ratnasingham, S., and deWaard, J.R. 2003. Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species. Proc. Biol. Sci. 270(1): S96–S99. PMID: 12952648.

Hogg, I.D., Willmann-Huerner, P., and Stevens, M.I. 2002. Population genetic structures of two New Zealand stream insects: *Archichauliodes diversus* (Megaloptera) and *Coloburiscus humeralis* (Ephemeroptera). New Zeal. J. Mar. Freshw. Res. 36(3): 491–501. doi:10.1080/00288330.2002.9517104.

Hotaling, S., Giersch, J.J., Finn, D.S., Tronstad, L.M., Jordan, S. Serpa, L.E., et al. 2019. Congruent population genetic structure but differing depths of divergence for three alpine stoneflies with similar ecology and geographic distributions. Freshw. Biol. 64: 335–347.

Hotaling, S., Kelley, J.L., and Frandsen, P.B. 2020. Aquatic insects are dramatically underrepresented in genomic research. Insects, 11(9): 601. doi:10.3390/insects11090601.

Hughes, J.M. 2007. Constraints on recovery: using molecular methods to study connectivity of aquatic biota in rivers and streams. Freshw. Biol. 52(4): 616–631. doi:10.1111/j.1365-2427.2006.01722.x.

Hughes, J.M., Schmidt, D.J., and Finn, D.S. 2009. Genes in streams: using DNA to understand the movement of freshwater fauna and their riverine habitat. BioScience, 59(7): 10. doi:10.1525/bio.2009.59.7.8.

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(4): 998–1002. doi:10.1111/j.1471-8286.2006.01428.x.

Kamvar, Z.N., Tabima, J.F., and Grünwald, N.J. 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ, 2: e281–e281. doi:10.7717/peerj.281. PMID: 24688859.

Kear, D. 1960. Sheet 4-Hamilton. Geological map of New Zealand 1:250000. Department of Scientific and Industrial Research, Wellington.

Knapp, I., Puritz, J., Bird, C., Whitney, J., Sudek, M., Forsman, Z., and Toonen, R. 2016. ezRAD: an accessible next-generation RAD sequencing protocol suitable for non-model organisms_v3.2. protocols.io.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A. McWilliam, H., et al. 2007. Clustal W and Clustal X version 2.0. Bioinformatics, 23(21): 2947–2948. doi:10.1093/bioinformatics/btm404. PMID: 17846036.

Lean, J., Hammer, M., Unmack, P., Adams, M., and Beheregray, L.B. 2017. Landscape genetics informs mesohabitat preference and conservation priorities for a surrogate indicator species in a highly fragmented river system. Heredity, 118: 374–384. doi:10.1038/hdy.2016.111. PMID: 27876805.

Malmqvist, B. 2000. How does wing length relate to distribution patterns of stoneflies (Plecoptera) and mayflies (Ephemeroptera)? Biol. Conserv. 93: 271–276. doi:10.1016/S0006-3207(99)00139-1.

Mardis, E. 2017. DNA sequencing technologies: 2006–2016. Nat. Protoc. 12: 213–218. doi:10.1038/nprot.2016.182. PMID: 28055035.

McCulloch, G.A., Foster, B.J., Dutoit, L., Ingram, T., Hay, E. Veale, A.J., et al. 2019. Ecological gradients drive insect wing loss and speciation: the role of the alpine treeline. Mol. Ecol. 28: 3141–3150. PMID: 31038802.

McGaughran, A., Terauds, A., Convey, P., and Fraser, C.I. 2019. Genome-wide SNP data reveal improved evidence for Antarctic glacial refugia and dispersal of terrestrial invertebrates. Mol. Ecol. 28(22): 4941–4957. doi:10.1111/mec.15269. PMID: 31596994.

McLellan, I.D. 1993. Antarctoperlinae (Insecta: Plecoptera). Fauna New Zeal. 27: 271–276

McLeod, E.O., Pittari, A., Brenna, M., and Briggs, R. 2020. Geology of the Pirongia Volcano, Waikato: 1:30,000 Geological Map, Geoscience Society of New Zealand, New Zealand.

Monaghan, M.T., and Sartori, M. 2009. Genetic contributions to the study of taxonomy, ecology, and evolution of mayflies (Ephemeroptera): re-

view and future perspectives. Aquat. Insects, **31**: 19–39. doi:10.1080/01650420902734145. PMID:26973367.

Monaghan, M.T., Spaak, P., Robinson, C.T., and Ward, J.V. 2002. Population genetic structure of 3 alpine stream insects: influences of gene flow, demographics, and habitat fragmentation. J. North Am. Benthol. Soc. **21**: 114–131. doi:10.2307/1468304.

Morin, P. A., Luikart, G., and Wayne, R.K. 2004. SNPs in ecology, evolution and conservation. Trends Ecol. Evol. **19**: 208–216.

Múrria, C., Bonada, N., Ribera, C., and Prat, N. 2010. Homage to the Virgin of ecology, or why an aquatic insect unadapted to desiccation may maintain populations in very small, temporary Mediterranean streams. Hydrobiologia, **653**: 179–190. doi:10.1007/s10750-010-0353-1.

Múrria, C., Sáinz-Bariáin, M., Vogler, A.P., Viza, A., González, M., and Zamora-Muñoz, C. 2020. Vulnerability to climate change for two endemic high-elevation, low-dispersive *Annitella* species (Trichoptera) in Sierra Nevada, the southernmost high mountain in Europe. Insect Conserv. Diver. **13**: 283–295. doi:10.1111/icad.12387.

Nelson, C. 1994. Insects of the Great Basin and Colorado Plateau. *In* Natural History of the Colorado Plateau and Great Basin. *Edited by* K.T. Harper, L.L. St Clair, K.H. Thorne,, and W.M. Hess. University of Colorado Press, Niwot. pp. 211–237.

Paradis, E., and Schliep, K. 2018. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics (Oxford, England), **35**(3): 526–528. doi:10.1093/bioinformatics/bty633.

Pazmiño, D.A., Maes, G.E., Simpfendorfer, C.A., Salinas-de-León, P., and van Herwerden, L. 2017. Genome-wide SNPs reveal low effective population size within confined management units of the highly vagile Galapagos shark (*Carcharhinus galapagensis*). Conserv. Genet. **18**: 1151–1163. doi:10.1007/s10592-017-0967-1.

Petersen, I., Masters, Z., Hildrew, A.G., and Ormerod, S.J. 2004. Dispersal of adult aquatic insects in catchments of differing land use. J. Appl. Ecol. **41**: 934–950. doi:10.1111/j.0021-8901.2004.00942.x.

Phillipsen, I.C., Kirk, E.H., Bogan, M.T., Mims, M.C., Olden, J.D., and Lytle, D.A. 2015. Dispersal ability and habitat requirements determine landscape-level genetic patterns in desert aquatic insects. Mol. Ecol. **24**(1): 54–69. doi:10.1111/mec.13003. PMID:25402260.

Polato, N., Gray, M., Gill, B., Becker, C., Casner, K. Flecker, A., et al. 2017. Genetic diversity and gene flow decline with elevation in montane mayflies. Heredity, **119**: 107–116. doi:10.1038/hdy.2017.23. PMID:28489073.

Previsic, A., Schnitzler, J., Kucinic, M., Graf, W., Ibrahimi, H., Kerovec, M., and Pauls, S.U. 2014. Microscale vicariance and diversification of Western Balkan caddisflies linked to karstification. Freshw. Sci. **33**: 250–262. doi:10.1086/674430. PMID:27042385.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R. Bender, D., et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. **81**(3): 559–575. doi:10.1086/519795. PMID:17701901.

Rader, R.B., Unmack, P.J., Christensen, W.F., and Jiang, X.B. 2019. Connectivity of two species with contrasting dispersal abilities: a test of the isolated tributary hypothesis. Freshw. Sci. **38**: 142–155. doi:10.1086/701671.

Raj, A., Stephens, M., and Pritchard, J.K. 2014. fastSTRUCTURE: variational inference of population structure in large SNP data sets. Genetics, **197**(2): 573. doi:10.1534/genetics.114.164350. PMID:24700103.

Ratnasingham, S., and Hebert, P.D.N. 2007. BOLD: the barcode of life data system. Mol. Ecol. Notes, **7**(3): 355–364. doi:10.1111/j.1471-8286.2007.01678.x. Available from http://www.barcodinglife.org. PMID:18784790.

Saastamoinen, M., Bocedi, G., Cote, J., Legrand, D., Guillaume, F. Wheat, C.W., et al. 2018. Genetics of dispersal. Biol. Rev. Camb. Philos. Soc. **93**(1): 574–599. doi:10.1111/brv.12356. PMID:28776950.

Salokannel, J., Min Lee, K., Rinne, A., and Mutanen, M. 2021. ddRAD sequencing sheds light on low interspecific and high intraspecific mtDNA divergences in two groups of caddisflies. Insect Syst. Div. **5**(5): 4. doi:10.1093/isd/ixab013.

Saito, R., Kato, S., Kuranishi, R.B., Nozaki, T., Fujino, T., and Tojo, K. 2018. Phylogeographic analyses of the *Stenopsyche* caddisflies (Trichoptera: Stenopsychidae) of the Asian Region. Freshw. Sci. **37**: 562–572. doi:10.1086/699364.

Sanger, F., Nicklen, S., and Coulson, A.R. 1977. DNA sequencing with chain-terminating inhibitors. ProcNatl. Acad. Sci. U.S.A. **74**(12): 5463–5467. doi:10.1073/pnas.74.12.5463.

Schröder, O., Cavanaugh, K.K., Schneider, J.V., Schell, T., Bonada, N., Seifert, L., and Pauls, S.U. 2021. Genetic data support local persistence in multiple glacial refugia in the montane net-winged midge *Liponeura cinerascens cinerascens* (Diptera, Blephariceridae). Freshw. Biol. **66**: 859–868. doi:10.1111/fwb.13682.

Sd, S., Roessli, D., and Excoffier, L. 2000. ARLEQUIN: a software for population genetics data analysis. Genetics and Biometry Lab, Department of Anthropology, University of Geneva.

Slatkin, M. 1977. Gene flow and genetic drift in a species subject to frequent local extinctions. Theor. Popul. Biol. **12**: 253–262. doi:10.1016/0040-5809(77)90045-4. PMID:601717.

Slatkin, M. 1985. Gene flow in natural populations. Annu. Rev. Ecol. Syst. **16**(1): 393–430. doi:10.1146/annurev.es.16.110185.002141.

Smith, L.M., Sanders, J.Z., Kaiser, R.J., Hughes, P., Dodd, C. Connell, C.R., et al. 1986. Fluorescence detection in automated DNA sequence analysis. Nature, **321**(6071): 674–679. doi:10.1038/321674a0. PMID:3713851.

Smith, B.J., and Collier, K.J., 2000. Interactions of adult stoneflies (Plecoptera) with riparian zones II. Diet. Aqua. Insects, **22**(4): 285–296. doi:10.1076/0165-0424(200010)22:4;1-Y;FT285.

Smith, P.J., and Collier, K.J. 2001. Allozyme diversity and population genetic structure of the caddisfly *Orthopsyche fimbriata* and the mayfly *Acanthophlebia cruentata* in New Zealand streams. Freshw. Biol. **46**(6): 795–805. doi:10.1046/j.1365-2427.2001.00719.x.

Smith, P.J., and Smith, B.J. 2009. Small-scale population-genetic differentiation in the New Zealand caddisfly *Orthopsyche fimbriata* and the crayfish *Paranephrops planifrons*. New Zeal. J. Mar. Freshw. Res. **43**(3): 723–734. doi:10.1080/00288330909510037.

Smith, P.J., McVeagh, S.M., and Collier, K.J. 2006. Population-genetic structure in the New Zealand caddisfly *Orthopsyche fimbriata* revealed with mitochondrial DNA. New Zeal. J. Mar. Freshw. Res. **40**(1): 141–148. doi:10.1080/00288330.2006.9517408.

Sproul, J.S., Houston, D.D., Davis, N., Barrington, E., Oh, S.Y., Evans, R.P., and Shiozawa, D.K. 2014. Comparative phylogeography of codistributed aquatic insects in western North America: insights into dispersal and regional patterns of genetic structure. Freshw. Biol. **59**(10): 2051–2063. doi:10.1111/fwb.12406.

Storfer, A., Patton, A., and Fraik, A.K. 2018. Navigating the interface between landscape genetics and landscape genomics. Front. Genet. **9**:68. doi:10.3389/fgene.2018.00068. PMID:29593776.

Tan, M.P., Wong, L.L., Razali, S.A., Afiqah-Aleng, N., Mohd Nor, S.A. Sung, Y.Y., et al. 2019. Applications of next-generation sequencing technologies and computational tools in molecular evolution and aquatic animals conservation studies: a short review. Evol. Bioinform. Online, **15**: 1176934319892284. doi:10.1177/1176934319892284. PMID:31839703.

Teske, P.R., Golla, T.R., Sandoval-Castillo, J., Emami-Khoyi, A., van der Lingen, C.D. von der Heyden, S., et al. 2018. Mitochondrial DNA is unsuitable to test for isolation by distance. Sci. Rep. **8**: 8448. doi:10.1038/s41598-018-25138-9. PMID:29855482.

Trense, D, Schmidt, TL, Yang, Q, Chung, J, Hoffmann, AA, and Fischer, K. 2021. Anthropogenic and natural barriers affect genetic connectivity in an Alpine butterfly. Mol. Ecol. **30**: 114–130. doi:10.1111/mec.15707.

Venables, S.K., Marshall, A.D., Armstrong, A.J., Tomkins, J.L., and Kennington, W.J. 2021. Genome-wide SNPs detect no evidence of genetic population structure for reef manta rays (*Mobula alfredi*) in Southern Mozambique. Heredity, **126**(2): 308–319. doi:10.1038/s41437-020-00373-x. PMID:33005043.

Vendrami, D.L.J., Telesca, L., Weigand, H., Weiss, M., Fawcett, K. Lehman, K., et al. 2017. RAD sequencing resolves fine-scale population structure in a benthic invertebrate: implications for understanding phenotypic plasticity. R. Soc. Open Sci. **4**(2): 160548. doi:10.1098/rsos.160548. PMID:28386419.

Wallace, F.M. 2013. Using population genetics to assess the dispersal patterns of the New Zealand mayfly *Coloburiscus humeralis* in a landscape context. Master of Science Dissertation, University of Canterbury, Christchurch, New Zealand.

Wang, P., Chen, B., Zheng, J., Cheng, W., Zhang, H. Wang, J., et al. 2020. Fine-scale population genetic structure and parapatric cryptic species of kuruma shrimp (*Marsupenaeus japonicus*), along the Northwestern

Pacific Coast of China. Front. Genet. **11**: 118. doi:10.3389/fgene.2020.00118. PMID: 32161618.

Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., and Barton, G.J. 2009. Jalview version 2—a multiple sequence alignment editor and analysis workbench. Bioinformatics (Oxford, England), **25**(9): 1189–1191. doi:10.1093/bioinformatics/btp033. PMID: 19151095.

Weir, B. S., and Cockerham, C. C. 1984. Estimating *F*-statistics for the analysis of population structure. Evolution, **38**(6): 1358–1370. doi:10.1111/j.1558-5646.1984.tb05657.x. PMID: 28563791.

Weston, M., Clarke, K., Maguire, G., and Sumner, J. 2020. Morphological and molecular evidence of population divergence in a widespread shorebird across its Southern Mainland Australian distribution. Conserv. Genet. **21**: 757–770. doi:10.1007/s10592-020-01286-2. PMID: 32607099.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. François, R., et al. 2019. Welcome to the Tidyverse. J. Open Source Softw. **4**(43).

Wilcock, H.R., Bruford, M.W., Nichols, R. A., and Hildrew, A.G. 2007. Landscape, habitat characteristics and the genetic population structure of two caddisflies. Freshw. Biol. **52**(10): 1907–1929. doi:10.1111/j.1365-2427.2007.01818.x.

Wilson, J.J. 2012. DNA barcodes for insects. Methods Mol. Biol. **858**: 17–46. doi:10.1007/978-1-61779-591-6_3. PMID: 22684951.

Winterbourn, M., Gregson, K.L.D., and Dolphin, C.H. 2000. Guide to the aquatic insects of New Zealand. Bull. Entomol. Soc. New Zeal. 13.

Winterbourn, M.J. 2005. Dispersal, feeding and parasitism of adult stoneflies (Plecoptera) at a New Zealand forest stream. Aqua. Insects, **27**(3): 155–166. doi:10.1080/01650420500062840.

Winterbourn, M.J., Chadderton, W., Entrekin, S., Tank, J., and Harding, J. 2007. Distribution and dispersal of adult stream insects in a heterogeneous montane environment. Archiv. Hydrobiol. **168**: 127–135. doi:10.1127/1863-9135/2007/0168-0127.

Zhang, J., Chiodini, R., Badr, A., and Zhang, G. 2011. The impact of next-generation sequencing on genomics. J. Genet. Genom. **38**(3): 95–109. doi:10.1016/j.jgg.2011.02.003. PMID: 21477781.