SPATIOTEMPORAL EVOLUTIONARY DYNAMICS AND BIOLOGICAL RESPONSE OF MOUNTAIN PINE BEETLE (*Dendroctonus ponderosae* Hopkins) (COLEOPTERA: CURCULIONIDAE) IN WESTERN CANADA AFTER RANGE EXPANSION

by

Kirsten M. Thompson

B. Sc. Trinity Western University, 2011
M. Sc. University of Northern British Columbia, 2015

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
NATURAL RESOURCES AND ENVIRONMENTAL STUDIES

UNIVERSITY OF NORTHERN BRITISH COLUMBIA

September 2022

Abstract:

Global ecosystems are increasingly affected by climate change leading to alterations in expected disturbance regimes. Outbreaks of irruptive insect pests represent some of the most destructive disturbance events possible that occur in the forests of North America. The mountain pine beetle, *Dendroctonus ponderosae* Hopkins (Coleoptera: Cuculionidae) (MPB), an irruptive tree-killing pest of several species of pine, is responsible for the most damaging insect outbreak in recent history in western Canada, leading to a massive range expansion from established territory in British Columbia to pine stands in central and northern Alberta during the mid-2000s. The adaptive genetic structure of these new populations is relatively unknown a decade after range expansion and the selective response of the beetle to novel habitats is likewise understudied. There is increasing concern among forestry stakeholders that MPB may develop novel genetic traits or behaviours in response to their new habitat.

In this study, I sampled beetle DNA and used double-digest restriction fragment genotyping-by-sequencing to generate single nucleotide polymorphism (SNP) loci. I used these data to investigate the establishment of genetic structure throughout the newly expanded and historic MPB range. I identified two distinct genetic clusters, a southern cluster for beetles located south of Banff National Park and a northern for beetles located north of the park. The data presented here suggest that Jasper National Park and the surrounding region represents an area of admixture between the two genetic clusters, caused in part by the movement of beetles from both the north and the south. This area of admixture may have the potential to differentiate into a separate third genetic cluster.

Outlier analysis indicated that several landscape variables including mean annual

precipitation and relative humidity contributed to the selective pressures on MPB, while frost free period contributes to the genotypes of beetles from the expanded range in central and northern Alberta. We also found that novel colonized MPB sites caused by in-flights from BC to Alberta did not display changes in genetic structure from their source population. Semivoltine MPB new adults collected from a site within the admixed zone near Jasper National Park displayed accumulation of cryoprotectants in response to cooling autumn temperatures but are not capable of surviving below -30°C. This is the first time that new adults have been demonstrated to cold harden, but also indicates that beetles from the admixed zone are likely not displaying aberrant forms of dispersal, host selection, or cold tolerance. I propose that beetles within the expanded range are not experiencing extreme selection events that would lead to genetic changes that will cause notable colonization, gustation, or dispersal differences to that of their counterparts in the historical range of BC. Warming climate is likely to contribute to beetle survival throughout central and northern Alberta, however, MBP populations will likely reach an endemic state and display historically expected population cycling (endemic to epidemic to endemic) in the future. The homogenous nature of MPB genetic structure throughout the expanded range does not support the development of populations-specific control techniques.

Permission to Reproduce:

Chapter 4 is a slightly modified version of Thompson et al. (2020) and is reproduced with the permission of PLoSOne (permission granted under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) license).

Thompson, K.M., Huber, D.P.W., and Murray, B.W. 2020. Autumn shifts in cold tolerance metabolites in overwintering adult mountain pine beetles. PLoS One 15: e0227203. Public Library of Science. doi:10.1371/journal.pone.0227203.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank my supervisor Dr. Brent Murray for accepting me as a PhD student and encouraging my pivot from fungal DNA barcoding work to the truly vast world of population genomics and MPB. His deep knowledge of genetics was an asset to me in shaping the conception of this project. I am also very grateful to Dr. Felix Sperling for his incredible entomology expertise and approach to science in general. Both Dr. Murray and Dr. Sperling helped to guide my MPB work and provided excellent sounding boards for my ideas. Funding for this work was provided by NSERC via the TRIA network and the UNBC Research Project Award program.

I am also grateful to my committee members Dr. Dezene Huber and Dr. Chris Johnson. Dr. Huber was instrumental in the conception and execution of my metabolite work and hugely encouraging during the publishing process. Dr. Johnson has consistently answered all of my statistical questions and provided invaluable advice on writing and research in general.

I must also thank Kinton Sy (my field assistant), Caroline Whitehouse, Erica Samis, Conan Ma, the members of FIRG, and the members of the TRIA network, particularly Jasmine Janes and Phil Batista. Special thanks to the members of the Sperling lab, including Victor Shegelski and Erin O'Dell Campbell. The staff of the Lucerne campground and Jasper National Park were also instrumental in supporting my sampling efforts.

Thanks go to my excellent circle of friends, including Sara Ho and Elizabeth Kreiter who have been my stalwart companions and kindred spirits in the world of graduate studies (and now you are mentioned at the beginning of TWO archived scientific works). I am grateful to my family (Dad, Mom, Liam, Josiah) for their support even as they all have exited student life and are wondering why I'm not hurrying up the process. Thanks to Brenda and Da for putting up with my scholastic zoom calls and note-taking during the holidays. Thanks also to Cayla for a weekend of fine art wonders right when I needed it.

To my husband Neil, you are the first, best thing in my life and I am forever grateful to you for your support during the completion of this work in various rentals on dubious country internet connections. You have been my backwoods driver, GIS wizard, camp chef, and tree de-barker extraordinaire. I both thank and blame you for the existence of the two most northernly sites in this study.

1. Chapter 1: Introduction

1.1. General Overview

Disturbance regimes, the cumulative changes ecosystems face over space and time, are inherently linked to climate and respond dynamically as climate conditions change. Since the early 2000s, changes in the distribution, range, and phenology behaviour of multiple organisms have been reported across global ecosystems concurrent with detection of increasing global mean temperatures (www.ipcc.ch; Parmesan, 2006). While some species may experience detrimental impacts due to the alteration of their habitat by climate, mobile species that have a broad ecological niche can adapt and move into newly hospitable areas (Dormann, 2007; Wiens, 2016).

Range expansions attributed to climate change have already been observed within the last 30 years in several species of birds, mammals, invertebrates, and plants (Gibbs & Sheffield, 2009; Hallegraeff, 2010; Cullingham *et al.*, 2011; Allendorf *et al.*, 2013; Davey *et al.*, 2013; Moran & Alexander, 2014). These range expansions follow changes in temperature and precipitation and may disrupt coevolved systems by causing species to move into areas with naïve hosts that are less capable of defending themselves (Parmesan & Yohe, 2003; Raffa *et al.*, 2008; Sambaraju & Goodsman, 2021). Evolutionary changes in response to warmer climates have been detected in multiple taxa (Klopfstein *et al.*, 2006; Parmesan, 2006). Rapid genetic changes are likely at species range margins due to population alterations due to dispersal effects and selection pressures from new environmental conditions, however the long-term effects of such changes on success and fitness are unclear (Young *et al.*, 2017).

Insects have ectothermic physiology; thus, they are highly responsive to changes in

temperature due to their reliance on external heat sources. Some insect guilds also benefit from weakened host plant defenses caused by changes in climate regime, such as drought or extreme heat (Logan & Amman, 1991; Hofstetter & Gandhi, 2022). Bark beetles are key parts of many natural disturbance regimes of conifer forests in the northern hemisphere. Bark beetles often act as the most impactful disturbance agent in those ecosystems, causing many hectares of damage every year (Logan *et al.*, 1995; Raffa *et al.*, 2008; Hrinkevich, 2012). Though these beetles are a natural part of the renewal process of conifer forests, coevolved with their host trees, human intervention by fire exclusion and selective harvest has led to the development of unnatural forest conditions. These densely stocked stands can influence population dynamics by favoring the reproductive success of beetle populations. The increase in beetle populations coupled with warming climate has created a positive feedback cycle of increased size and severity of outbreaks across the landscape leading to detrimental ecosystem impacts and economic damage to the forest industry and forest-dependant communities (Logan & Powell, 2001; Six & Bracewell, 2015; Sambaraju & Goodsman, 2021).

Mountain pine beetle, *Dendroctonus ponderosae* Hopkins (Coleoptera: Curculionidae) (MPB)), is responsible for the most destructive insect outbreak in the history of western Canada (Aukema *et al.,* 2006; Janes *et al.*, 2014). This "hyperepidemic" resulted in a massive range expansion of MPB north and east of its known recent territory in western Canada (Sambaraju *et al.*, 2019). A combination of favourable stand conditions, stressed hosts from warmer, drier climate conditions, and milder winters led to dramatic synchronous population irruptions starting in multiple locations in British Columbia (BC) in the early 1990s (Safranyik & Wilson, 2007). By 1999–2003 populations had reached

epidemic levels, affecting over 7 million hectares of lodgepole pine (*Pinus contorta* var. *latifolia* (Engelm.) Critchfield) and ponderosa pine (*P. ponderosa* Douglas ex C.Lawson) forest. In 2003, prior to the peak of the outbreak in BC, researchers already referred to the scale of the damage as "unprecedented" (Aukema *et al.*, 2006). During the peak of the outbreak in 2006, atmospheric winds pushed billions of MPB over the northern part of the Rocky Mountains and extended the range of the species into novel forests in Alberta (Jackson *et al.*, 2008). As the outbreak declined in BC, leaving over 18 million hectares of pine forest affected, MPB continued to expand its range in Alberta and spread eastward, expanding toward the Alberta hybrid zone between jack pine (*Pinus banksiana*, Lamb) and lodgepole pine (Cullingham *et al.*, 2011; McKee *et al.*, 2015).

MPB infests several species in the genus *Pinus* in BC, primarily attacking lodgepole pine, ponderosa, western white (*P. monticola* Douglas ex D. Don), limber (*P. flexilis* E.James), and whitebark pines (*P. albicaulis* Engelm.) (Six & Bracewell, 2015). Outbreaks of MPB typically leave patches of dead trees, which in sub-epidemic conditions will lead to uneven-aged stand regeneration (Safranyik & Carroll, 2006). Before aggressive wildfire suppression started in the mid-20th century, it is theorized that forest fires helped to maintain heterogeneous age structures in lodgepole pine forests, reducing the probability of MPB outbreak synchrony by reducing host availability (Seidl *et al.*, 2016). Fire suppression in the interior of BC after European settlement led to an accumulation of vulnerable over-mature, decadent pine on the landscape, which was one of the factors that contributed to the severity of the most recent outbreak (Whitehead *et al.*, 2001; Bleiker, 2016).

MPB outbreaks have been recorded in BC forests since 1910, affecting up to

450,000ha annually through 1995 (Wood & Unger, 1996; Bleiker, 2016). Dendrochronological records show periodic MPB outbreaks of moderate to large scale across the central part of the province prior to European settlement (Hrinkevich & Lewis, 2011). Outbreaks ranging in size from 4,000 to 650,000ha have occurred asynchronously in regions of BC throughout much of the 20th century. An outbreak in Banff, Alberta affected 10,000ha in the 1940s (Powell, 1961), but no infestations east of the Rocky Mountains have been recorded north of Banff prior to the most recent infestation, though isolated outbreaks have occurred in the south of Alberta near the Crowsnest Pass and Cypress Hills in the 1980s and 1990s (Langor, 1989; Emond & Cerezke, 1991).

Historical MPB outbreaks would have followed a cyclic pattern of population expansion and collapse: endemic phase, incipient epidemic phase, epidemic phase, and a return to endemic phase (Safranyik & Carroll, 2006; Cooke & Carroll, 2017). Endemic populations persist in small patches of weakened or recently killed trees until favorable climatic conditions and host availability permit expansion to the epidemic (outbreak) phase. These populations are by nature small and isolated, with beetles competing for sub-cortical resources with other wood and bark boring insects. During the incipient epidemic to epidemic phase, mass attacks of healthier trees with stronger constitutive and induced defenses become possible as beetle numbers become sufficient to overwhelm those defenses (Safranyik & Wilson, 2007; Six & Bracewell, 2015). Population numbers drive this change in host preference, with trees with weakly concentrated monoterpenes in phloem resin preferred in times of low MPB density and trees with more concentrated resin and associated thicker phloem preferred at times of high MPB density (Boone *et al.*, 2011).

Tree colonization is initiated by the female beetle who locates a tree and bores a chamber under the bark. Males are attracted by the aggregation pheromone *trans*-verbenol released by the female and the male-female pair will then mate and construct a vertical gallery into the phloem of the tree. Male beetles then release the pheromones *exo*-brevicomin and frontalin (Borden *et al.*, 1987). The *trans*-verbenol released by the females will attract both males and more females, contributing to a mass attack of the host tree, where great numbers of beetles will attack together and bore into the phloem. Beetles will colonize the tree until enough pairs arrive to overwhelm the host trees' resin defenses, and then begin to release anti-aggregation pheromones, mostly verbenone from females and frontalin from males, to direct incoming conspecifics to other trees in the stand (Borden *et al.*, 1987).

Mated females lay their eggs (~60) into egg niches partitioned along the sides of the gallery under the bark (Safranyik & Carroll, 2006). As the female ingests more nutritious phloem and moves down the galleries, the eggs will increase in size. The action of excavating tunnels under the bark damages the tree's resin ducts, but also introduces the MPB symbiote complex, blue stain fungi (Ophiostomatales, Ascomycota). Though trees will release resin as part of both their terpenoid-based constitutive and induced response to the beetle (Raffa & Berryman, 1983), in the case of a mass attack, the damage to the resin ducts effectively girdles the trunk and the infection of the fungi species blocks water transport, both processes that rapidly kill the tree (Plattner, 2008, Chiu, 2018).

MPB has four larval instars, prior to pupation. A majority of the MPB life cycle is spent under the bark of the host tree in larval form. Phloem and cambium tissue and the symbiotic fungus that surrounds the egg gallery provides the food source for the larva as

they excavate out tunnels at right angles to the parental gallery. Larval development will continue until temperatures grow too cold in the winter. During the fall and early winter months, larva of all stages will accumulate cryoprotectants such as glycerol in response to freezing temperatures (Robert *et al.*, 2016; Fraser *et al.*, 2017). Fourth instar larvae will clear a chamber at the end of their gallery after winter in preparation for pupation, then pupate, and metamorphosize into teneral adults. Teneral adults will continue to feed on the tree and go through sclerotization, hardening and darkening in colour prior to emergence, thus completing the life cycle (Six & Bracewell, 2015).

Dispersal of MBP is mediated by stand conditions, pheromone cues, and wind. A majority of MPB are short-distance dispersers, moving below the canopy to relatively nearby trees. Beetles that detect pheromone plumes in flight will be attracted to the tree that is the source of that plume (Borden *et al.*, 1987). Flight distance is highly variable (0.3 to 30km) and depends on the stored energy a beetle has, and the scent cues it receives from pheromone plumes released by other attacking beetles (Evenden *et al.*, 2014). In some cases, MPB will fly upward and exit the canopy to be caught in updrafts and convective winds. This wind transport is estimated to occur in less than 3% of all dispersing beetles, but when MPB numbers are high, can lead to large-scale transfer of beetles across the landscape (Jackson *et al.*, 2008).

Climate plays a contributing role in the phases of development and emergence for MPB. MPB usually has a univoltine life cycle, taking place over the course of one year, with adults typically laying eggs in the late summer to early fall with larvae overwintering to pupate and emerge as new adults in the summer months. Emergence of adults and the development of the 4 larval instars is closely tied to ambient temperatures, with larval

development slowing as temperatures cool (Bentz & Mullins, 1999). Both eggs and pupae are freeze-intolerant and incapable of weathering the colder winter temperatures that larvae are resilient to (Reid, 1962; Amman & Cole, 1983). This is most likely due to their inability to generate cryoprotectants the way that MPB larvae do. Eggs can withstand a short exposure to -20°C, but also suffer high mortality when held at temperatures of 0°C and below (Bleiker *et al.*, 2017). In high-elevation areas, MPB has been observed to have a semivoltine life cycle, with new, un-emerged adults overwintering under the bark and emerging in their second summer due to the limited amount of summer development time (Amman, 1973). Though earlier warm summer temperatures have caused faster beetle emergences and multiple generations in a single year in other *Dendroctonus* species, MPB has not yet been documented with a bivoltine life cycle (Bentz & Powell, 2014).

Exposure to temperatures of –40°C will kill overwintering MPB larvae above the snow line, placing a natural check on population growth at higher latitudes and elevations (Safranyik & Wilson, 2007). Freezing-related die-off of larval broods in appreciable numbers was last recorded in the winters of 1984 and 1985 within known BC MPB habitat, with few to no winters in regions south of the city of Quesnel achieving brood killing temperatures (Wood & Unger, 1996; Dawson *et al.*, 2008). Warmer temperatures since then in the summers may have stressed host trees and shortened beetle development time, further facilitating population growth (Bleiker, 2016). Taken together, stand conditions, fire suppression, and climate change all contributed to the increase in MPB numbers in BC and the migration of MPB into Alberta.

MPB is now found at higher elevations and latitudes in both BC and Alberta than previously recorded (Aukema *et al.*, 2006; Cullingham *et al.*, 2011). Alberta's climate

within the expanded range is notably different than that of historical MPB habitat within BC, receiving influence from chinooks and arctic winds, meaning that beetles may face new selective climate pressures (Nkemdirim, 1986; Whitfield & Cannon, 2000). The large number of breeding individuals has also possibly introduced mutations into the gene pool (Allendorf *et al.*, 2013). Interbreeding between previously isolated populations is also likely, given the movement of the outbreak across BC prior to entering Alberta (Janes *et al.*, 2014). As MPB moves eastward in Alberta, it presents an opportunity to study molecular evolutionary dynamics after a range expansion.

1.2. Previous Research on MPB on Western North America:

In-depth genetic research into the population structure of the current MPB outbreak in western North America did not begin until near the outbreak peak in 2006, though there are some earlier studies prior to the "hyperepidemic" investigating host use and MPB speciation in North America (Lanier & Wood, 1968; Stock & Guenther, 1979; Langor & Spence, 1991). Most research conducted post-2000 has been neutral polymorphic genetic markers, suitable for genetic analysis of gene flow and population structure. Mock *et al.* (2007) used amplified fragment length polymorphisms (AFLP) markers and mitochondrial DNA (mtDNA) sequenced from seven locations within the USA and one location from the northern interior of BC to explore the populations structure of MPB within the western USA. Using 159 loci, they reported a pattern of isolation-by-distance that moved northward with host pine species expanding away from source populations near the southwestern USA. They also found preliminary indications that MPB does not demonstrate panmictic breeding across the USA, and that the single Canadian MPB population had lower genetic diversity patterns that mirrored lower genetic diversity found

in host lodgepole pine samples taken from regions north of known pine glacial refugia (Mock *et al.*, 2007).

Several other studies further built on the work of Mock *et al.* (2007), exploring genetic structure of MPB within the established range in the USA and Canada. Cullingham *et al.* (2012) used mtDNA, sampled from the eight sites from Mock *et al.* (2007) with an additional 26 western Canadian sites to explore the phylogeographic relationships of MPB in North America. Using 267 individuals, they reported considerable diversity in haplotypes with a strong isolation-by-distance pattern across North America. Both Cullingham *et al.* (2012) and another study by Janes *et al.* (2018) found that northern regions also had low MPB genetic diversity attributed to relatively recent post-glacial expansion (Cullingham *et al.*, 2012; Janes *et al.*, 2018). Bracewell *et al.* (2017) found that that reproductive isolation between population clusters in MPB sampled in the USA is driven by fixed neo-Y haplotypes, leading to reproductive incompatibility between population clusters (Bracewell *et al.*, 2017). Dowle *et al.* (2017) found that autosomal gene flow was reduced between populations due to this limitation caused by these fixed Y-haplogroups, with three distinct populations found within the USA. The two Canadian sites used in this study (one from BC and one from Alberta) belong to the same Y-haplogroup.

Samarasekera *et al.* (2012) took advantage of the increased numbers of MPB on the landscape during the mid-2000s and collected MPB samples from 49 locations in BC and Alberta, genotyping a total of 4607 beetles at 16 microsatellite loci, using specimens collected in 2005–2008. An analysis of these neutral markers delineated a northern and southern population cluster within BC, with more genetic homogeneity in northern populations and more structure in the more established southern BC populations. The

partitioning of the two genetic clusters did not correlate with any apparent landscape barrier but was theorized to related to climate factors and post-glacial expansion of the beetle, following the northward movement of pine forests. Isolation-by-distance patterns were found in southern populations but were absent in northern beetle populations. In addition, they found that the northern population, while influenced by the large outbreak located near Tweedsmuir Provincial Park - South, was produced by the merging of multiple population centers (Bartell, 2008; Samarasekera *et al.*, 2012).

I participated in a joint follow up study to Samarasekera *et al.*'s (2012) work that also used microsatellite loci to map MPB population structure and relatedness across western Canada and the USA (Boone *et al.*, 2022). Beetles were collected from 153 sites in the USA and Canada and 3858 individuals were genotyped at 16 microsatellite loci (collection years 2003–2012). Our data indicated that MPB may have moved from glacial refugia in Oregon northward over the past several thousand years. We also found that host use does not influence population structure across the landscape, a finding that is supported by a smaller allozyme work using beetles from southwest Alberta (Langor, 1989). Both microsatellite studies show the basic population structure of MPB in western Canada, which provides context for the origins of the currently active MPB infestations. Neither study found explicit barriers to MPB populations interbreeding, particularly within the northern and southern clusters contained within Canada (Boone *et al.*, 2022).

Janes *et al.* (2016) used single nucleotide polymorphisms (SNPs), a different neutral genetic marker system, to assess beetle population structure in BC and Alberta. Beetles (n=548) sampled from multiple years (2006–2010) were genotyped using Illumina GoldenGate® technology, generating a panel of 1536 SNPs. Two clusters were detected on

the landscape, with more diverse genetic structure present in the southern cluster, supporting the results of Samarasekera *et al.* (2012). Janes *et al.* also found evidence of genetic diversity at the hypothesized intersections of these clusters near the border of BC and Alberta in the region of Mount Robson Provincial Park and Valemont (Janes *et al.*, 2014). Outliers in this study were identified with loci linked to ion transport, actin contraction, and sterol association. Batista *et al.* (2016). also published research showing that both adaptive and neutral SNPs can be used to delineate population structure in MPB. In their study, 1115 individuals, collected in 2005–2011, were genotyped at 92 SNPs, generated using the Sequenom iPLEX Gold® assay (Batista *et al.*, 2016). The inclusion of adaptive markers in this dataset (36 of the 92) improved the resolution of genetic structure. A subset of these adaptive loci was identified as outliers linked to the cell cycle and to DNA/RNA processes. In both cases, these studies used SNP loci generation methods that are considered more "targeted" than those generated by genotype-by-sequencing methods that survey loci from throughout the entire genome.

Trevoy *et al.* (2018) piloted the use of double-digest restriction-site associated DNA sequencing (ddRADseq) genotype-by-sequencing on 175 MPB individuals in BC and Alberta, including 13 lab-bred crosses to simulate interbreeding between northern and southern MPB. They found that these crossed individuals shared some similarities with beetles originating from Jasper National Park, close to the region of genetic admixture identified in previous studies (Janes *et al.*, 2014). I was also a part of a collaborative study by Shegelski *et al.* (2021) which used a similar ddRADseq based study on 294 individuals from nine sites at 2872 genomic SNPs, finding little genetic structure in the five sites contained in central Alberta. Shegelski *et al.* (2021) used the genetic material from a small

subset of beetles used in this dissertation.

These studies (Janes *et al.*, 2014; Batista *et al.*, 2016; Trevoy *et al.*, 2018; Shegelski *et al.*, 2021) showed the versatility of SNPs as a genetic marker to use in the study of MPB and present possible gene flow patterns that will be informative for my study. While some of these studies have identified MPB outlier loci, they have not yet explored the genetic correlations of outlier loci with potential landscape drivers (e.g., precipitation, temperature, stand condition). The impact of these landscape factors on MPB genetic dynamics is critically important to document if we are to understand how MPB will evolve throughout its new range.

1.3. Study Objectives and Rationale:

As of 2022, MPB continues to infest parts of Alberta, though populations appear to be in a state of decline (Belanger, 2022). The descendants of these beetle populations within the expanded range represent the foundational genetic character of future outbreaks across the north of both BC and Alberta. A study of genetically adaptive signatures of MPB in Alberta a decade after range expansion will provide insight into how irruptive pest population structure changes in the face of long-range migration and colonization of new territory. This work will explore the possibility of rapid evolution and potential development of advantageous loci or behaviours, with particular focus on functional cold hardening behaviour when MPB is exposed to climate in its new range. There is also concern within the forestry-connected stakeholder community in Alberta that Jasper National Park may produce beetles that are notably different from others throughout the province and that the park may act as a source population of new beetles, intensifying the damage caused by the initial MPB colonization event (Weber, 2017; Gonzalès & Parrott,

2019; Cowley, 2022). An understanding of the current population structure and adaptive status of particular loci will aid in determining if targeted management recommendations are required and indicate if there are region-specific actions that should be taken.

My overall research objectives were 1) to identify predictors of selective pressure on MPB in Alberta and BC approximately 10 years after population expansion, 2) to look for strong selection events in advantageous MPB loci, 3) to fully describe the MPB population structure remaining on the landscape within the north of British Columbia and Alberta, particularly as it relates to spread from still-active centers 4) to describe the short-term genetic changes in site-level populations structure using current (2016) and historical MPB samples (2005 and 2007), and 5) to document the metabolomic profiles and cold hardening behaviour of a population of overwintering new MPB adults, sampled from an area of putative admixture, in response to ambient fall and winter temperatures.

1.4. Organization of Dissertation

I have divided my dissertation into an introduction, three data chapters, and a concluding summary chapter with three appendices for code and analysis outputs. Chapters 2 and 3 address objective 2 of the Populations Genomics-MPB section of the NSERC (Natural Sciences and Engineering Research Council) TRIA (Turning Risk into Action for the Mountain Pine Beetle Epidemic) network research grant proposal by exploring the changes in MPB population structure and the presence of outlier loci within the expanded range. Funding for this dissertation was provided in large part by the TRIA network grant, and this research is part of a large body of collaborative research on MPB produced by the grant. Chapter 4 of this work explores the overwintering processes of new adult mountain pine beetle. Chapters are structured as manuscripts prepared for publication with figures

and tables presented at the end of each chapter. References to other chapters within data chapters are intended to guide the reader and will be edited out in the event of publication. For brevity, all references are contained in one section at the end of this work.

2. Chapter 2: In-Flights of Outbreak Populations of Mountain Pine Beetle Alter the Local Genetic Structure of Established Populations a Decade After Range Expansion

Kirsten M. Thompson, Dezene P. W. Huber, Chris J. Johnson, Felix A. H. Sperling, Brent W. Murray

2.1. Abstract

Mountain pine beetles began to appear at epidemic levels in Alberta, Canada, in 2006, following six years of extensive outbreaks in neighboring British Columbia. We assessed the effect of genetic MPB in-flights from the peak of the outbreak on the genetic structure of established populations of MPB and the change over time in novel regions colonized by these inflights. We used five locations sampled during the peak of the outbreak (2005/2007) and re-sampled in 2016. We performed a ddRADseq protocol to generate a SNP dataset via single-end Illumina sequencing. We detected a northern and southern genetic cluster in both sampling sets (2005/2007 and 2016) and a demographic shift in cluster assignment after ~10 generations from south to north in two of the sites in the path of the northern outbreak. Fst values were significantly different between most sites in the same years and between the same sites at different years, with some exceptions for northern sites established by inflights. Overall, sites in the spreading path of the MPB outbreak have taken on the genetic structure of the contiguous northern outbreak except for an isolated site in Golden, BC, and in Mount Robson Provincial Park where populations are admixed between north and south. Our results suggest that range expansion during insect outbreaks can alter the genetic structure of established populations and lead to interbreeding between populations.

2.2. Introduction

Global ecosystems are experiencing dramatic changes in climate regime, leading to species from multiple taxa undergoing range shifts or range expansions (Parmesan, 2006; Clements & DiTommaso, 2012). In some cases, these climatic changes lead to detrimental effects and species loss due to habitat exclusion, but in other cases species benefit from climatic changes by experiencing population growth and increased movement (Ciosi *et al.*, 2008; Mona *et al.*, 2014). Model-based genetic studies suggest that range expansion causes established population structure to change over time, with changes in patch occupation, gene flow between populations, and allele surfing contributing to alterations in genetic character (Klopfstein *et al.*, 2006; Mayrand *et al.*, 2019). Populations on the leading edge of a range expansion may also lose structure as they carry only a subset of the genetic diversity of the species, while at the same time establishing the foundational population throughout the new habitat range (Ibrahim *et al.*, 1996).

The effects of range expansions and invasion often focus on phenotypic changes or population demographic shifts to document the changes a species undergoes after migration to a new area (Young *et al.*, 2017). Short-term genetic consequences of range expansions are addressed less frequently in the literature. Instead, many studies focus on historical genetic changes comparing geological time scales or movement away from glacial refugia (Hellberg *et al.*, 2001; Roberts & Hamann, 2015; Hagen *et al.*, 2015). The studies that do document genetic impacts over relatively short periods of time (several years to decades) focus on population structure changes within established species ranges, relying on archived samples and citizen collections with imprecise location data, as seen with studies of red deer (Nussey *et al.*, 2005) and bobcats (Carroll *et al.*, 2019). Other

genetic research on recently expanded species ranges utilises contemporary samples from the current population only to assess the establishment of current structure, as seen in the study of invasive crabs (Herborg *et al.*, 2007), sparrows (Liebl *et al.*, 2013), geckos (Short & Petren, 2011), and wasp spiders (Krehenwinkel *et al.*, 2016). Rarely are the short-term temporal ramifications of species range expansion documented among years across repeated sites, as was the case in a study of brown bears that demonstrated the rapid genetic changes possible during range alterations (Hagen *et al.*, 2015). In many cases, these genetic studies of range expansion focus on macrofauna that have relatively small numbers of offspring, or somewhat limited dispersal capabilities. Irruptive insects that produce high numbers of individuals that can disperse easily by air are not widely represented in the literature.

Mountain pine beetle, *Dendroctonus ponderosae* Hopkins (Coleoptera: Curculionidae) (MPB)) is one such irruptive insect. MPB is a tree killing sub-cortical pine bark beetle, responsible for one of the most widespread and damaging insect outbreaks in recent history (Raffa *et al.*, 2008; Six & Bracewell, 2015). Over 18.3 million hectares of pine forests were affected by MPB in the wake of this outbreak in British Columbia from 2000–2012 (McKee *et al.*, 2015). This native beetle is found in pine species in the southwestern USA to the Black Hills of South Dakota, most of British Columbia (BC), and recently expanded into Alberta. In western Canada, mountain pine beetle has been a common part of the natural disturbance regime of pine forests with recorded outbreaks of differing size recorded throughout the 20th century, and dendrochronologically reconstructed outbreaks dated prior to European settlement (Wood & Unger, 1996; Hrinkevich & Lewis, 2011).

Warmer temperatures linked to climate change prevented MPB winter brood die-off and increased breeding success in the early 1990s and into the 2000s (Safranyik & Linton, 1998; Carroll *et al.*, 2006). Fire suppression and lack of human interference within established pine stands in favor of focus on more lucrative softwood species created a contiguous food source for spot outbreaks to grow and coalesce together (Whitehead *et al.*, 2001). Historically, the mountainous terrain of the eastern portion of BC and the Rocky Mountains acted as range barriers to prevent excessive movement of MPB into Alberta, except in the areas around Banff National Park, the Crowsnest Pass, and Cypress Hills (Powell, 1961; Langor, 1989). MPB populations within the southern portion of BC were typically isolated from one another by the steep mountain terrain and patchiness of pine habitat interspersed with interior grasslands.

Mountain pine beetles began to appear in large numbers in Alberta in 2006 (Bartell, 2008), crossing over the Rocky Mountains assisted by strong convective winds (Jackson *et al.*, 2008). These initial inflights of beetles were concentrated near Canmore in the south and Grande Prairie in the north but have since spread eastward as far as Lac La Biche (Bartell, 2008; Cullingham *et al.*, 2011; Pokorny, 2021). A decade after the initial colonization of these regions during the peak of the outbreak, beetles are still present and continuing to attack surviving pine. Some areas appear to be experiencing epidemic-level attacks while others appear to be collapsing down to endemic levels.

Several genetic studies addressed the population structure of MPB within western Canada using a variety of marker systems. One of the first within BC used microsatellite markers to investigate the MPB outbreak throughout its original range and parts of the expanded range (Bartell, 2008). Populations in the north were found to be more

homogenous, while southern populations were more structured, though although the authors did not identify a landscape barrier leading to this division (Samarasekera *et al.*, 2012). A similar study using single nucleotide polymorphisms (SNPs) confirmed the differentiation between the northern and southern geographic clusters, also finding that the Robson Valley and nearby Jasper National Park had greater genetic diversity in an intermediate area between the north and south (Samarasekera *et al.*, 2012; Janes *et al.*, 2014; Batista *et al.*, 2016). These studies combined beetles from multiple years, but also were conducted during or within several years of the initial MPB range expansion in the mid-2000s. In all cases, the researchers did not have the opportunity to re-sample regions and compare MPB genetic structure between age cohorts.

It is unknown if the populations of MPB in regions that received a high level of migrant beetles, like Smithers or Canmore, retained their original endemic population structure 10 years after the initial in-flights, or if they have taken on the character of the northerly MPB population that expanded into the north of BC and into Alberta. It is also unknown how locations with novel colonization, like Grand Prairie, have changed since the first in-flights. In both cases, beetles will have completed approximately nine to eleven generations at each location, as MPB is known to re-attack relatively close (<3 km) to their original host tree if conditions are favourable (Evenden *et al.*, 1943) and the correct aggregation pheromone cues are received (Evenden *et al.*, 2014).

Our study took advantage of beetles archived by research teams during the first wave of expansion of the BC outbreak in 2005 and 2007 and compares those specimens to populations active in 2016. Our objective was to document the short-term temporal genetic structure changes that occur in the aftermath of an insect range expansion with a high

degree of immigration into novel territory and into areas with established populations. Temporal genetic sampling methods have been recommended for use in assessing genetic structure post-range expansion (Short & Petren, 2011; Hagen *et al.*, 2015), though their use in insects at this time has largely been restricted to agricultural pests responding to abiotic inputs (Pélissié *et al.*, 2018). We hypothesized that samples from the first time point would have greater genetic structure than those sampled ten years on, but that samples from established populations would retain, in part, their original structure due to dispersal limitations. We also anticipated that admixture would increase over time due to gene flow between sampling locations (Hagen *et al.*, 2015).

2.3. Methods

2.3.1. Sampling and Extraction of Genomic DNA

Field sampling was conducted in the summer of 2016. Adult and teneral (un-hardened) beetles were collected from trees in British Columbia and Alberta. Sampling was informed by aerial survey data provided by the BC Ministry of Forests, Lands and Natural Resource Operations (https://catalogue.data.gov.bc.ca/, 2022) and field survey data from Alberta Agriculture and Forestry (C. Whitehouse, personal communication, June 2016). Infested trees were identified in the field by observing yellowing foliage and the presence of pitch tubes accompanied by frass on the bole. Bark was peeled using a draw knife to expose pupal chambers. A minimum of 40 adult or teneral beetles were collected per site, with individuals selected from separate galleries to prevent comparison of full siblings. Samples collected from 2005 and 2007 used the cardinal direction collection method, where beetles were selected from four points around the bole of the tree to prevent gallery overlap. In both cases, a maximum of four per tree were retained when possible.

Beetles collected in 2016 were immediately placed in 95% ethanol upon removal from the tree in the field, transferred to a lab, and then stored at -20℃ prior to DNA extraction. All the 2005 and 2007 specimens from Alberta and BC used in this study followed the same method of collection and ethanol preservation as above. Samples from 2007 were collected as second instar larva and lyophilized for room-temperature, shelf-stable storage prior to DNA extraction.

Genomic DNA was isolated from all MPB specimens collected in 2016, and those collected in 2007 using Qiagen Dneasy Blood & Tissue spin column and 96-well plate kits (Germantown, Maryland, USA) following manufacturer's instructions, with the addition of an overnight tissue lysis step incubated at 56℃ and the use of an optional RNA removal step using 2µL of RNaseA with a 15-minute incubation. DNA product was eluted into DNase-free water for ease of later sequencing. DNA from the samples collected in 2005 was previously extracted using a phenol-chloroform method and stored at –80℃ (Samarasekera *et al.*, 2012). Final concentration was quantified using both a NanoDrop™ 1000 Spectrophotometer and a first-generation Qubit fluorometer with the Quant-iT dsDNA BR Assay Kit.

2.3.2. Sequencing and Data Cleaning

Double-digest restriction-site associated DNA sequencing (ddRADseq), a type of genotyping-by-sequencing, was used to generate the SNP markers for this project. Genomic library preparation and sequencing was completed by the Molecular Biology Services Unit (MBSU) of University of Alberta. DNA Library preparation followed the MBSU specifications for MPB sampling (Peterson *et al.*, 2012; Campbell *et al.*, 2017). Total purified DNA input per sample was 80ng. A non-methylation sensitive two-enzyme

system (*Pst* I and *Msp* I) was used to fragment whole genomic DNA. Samples were sequenced on a single lane of an Illumina NextSeq 500 at the MBSU to generate single-end 75 bp reads.

Initial sequence read demultiplexing and mapping was performed on Compute Canada's Graham cluster (Toronto, Ontario, Canada). Reads were demultiplexed and quality checked using the default pipeline in STACKS 2.0b (Rochette *et al.*, 2019). Cutadapt (version 1.9.1) was used to trim the *Pst* I restriction site on the 5'-end of each read (Martin, 2011). Reads were also trimmed to a length of 67 bp. Processed reads were mapped to the female draft MPB reference genome (Keeling *et al.*, 2013) using BWA-MEM (Burrows-Wheeler Aligner Maximal Exact Match – version 0.7.17) (Li & Durbin, 2009). Quality of the alignments was assessed using SAMtools (version 1.9) (Li *et al.*, 2009b).

Single nucleotide polymorphism variants were called by the ref_map.pl perl wrapper provided with STACKS 2.0b with a minimum minor allele frequency of 0.05. We also used the r80 principle where any retained locus must appear in at least 80% of all individuals (Paris *et al.*, 2017; Rochette *et al.*, 2019). VCF files generated by ref_map.pl were filtered again using VCFtools (version 0.1.14) for a minor allele frequency of 0.05, a minimum quality score of 30, and minimum read depth of 10 (Danecek *et al.*, 2011). To mitigate the confounding signal of sex-linked markers, the dataset was loaded into R (version 3.6.3) and SNP loci were filtered using the PCA based method in (Trevoy *et al.*, 2019) by removing loci associated with high PC 1 values (R Core Team, 2018). We filtered for linkage disequilibrium ($R^2 > 0.5$) using the R package dartR (Abdellaoui *et al.*, 2013; Gruber *et al.*, 2018) retaining only one random SNP from each linkage group to prevent

excessive clustering during population structure analysis. Individuals were rechecked for quality after linkage filtering and any remaining poor-quality individuals (more than 20% data missing) were removed.

2.3.3. Assessment of Population Structure

SNP data were first visualized using the adegenet package (version 2.1.2) in R (Jombart, 2008; Jombart *et al.*, 2010). A combination of PCA and DAPC was used to assess similarity between the two time periods and between the individual locations. PCA was performed without assigning populations to specific groups based on collection year. DAPC maximizes the variation between groups and uses a priori population assignment. Individuals were assigned to populations corresponding to their location of origin. Cross validation calculations were performed following the defaults contained in the DAPC vignette to assess the number of PCs to retain and the DF to retain for final visualization (Jombart *et al.*, 2010).

We used STRUCTURE (version 2.3.4) to explore population structure among all of our populations (Pritchard *et al.*, 2000; Falush *et al.*, 2003, 2007; Hubisz *et al.*, 2009). For this dataset, we used the admixture model and did not specify location information or group individuals into populations. We performed an initial STRUCTURE run with 100,000 Markov chain Monte Carlo generations and a burn-in period of 50,000. We then increased our STRUCTURE run to 1 million Markov chain Monte Carlo generations and a burn-in period of 250,000 and tested for values of K between 1 and 10 with 10 replicates of each K value. STRUCTURE results were visualized using CLUMPAK (Kopelman *et al.*, 2015) to create an average for each value of K. Optimal K was determined by comparing the results of the ΔK Evanno method (Evanno *et al.*, 2005) between likelihood

of K values and the mean estimated natural logarithm for the K value probability (Ln(PrK))(Pritchard *et al.*, 2000). Both ΔK and Ln(PrK) were compared using STRUCTUREHARVESTER (Earl & VonHoldt, 2012). A hierarchical AMOVA was conducted in Arlequin (version 3.5.2.2) using pairwise Fst comparisons on the 10 sites, among the two collection dates, among sites within the collection dates, and among individuals within populations, and within individuals, set at a 0.05 significance level with 1000 permutations (Excoffier & Lischer, 2010). Isolation by distance was tested using a Mantel test in dartR (version 1.1.11) with a Pearson's product-moment correlation and a Fst/1-Fst vs. the log distance in meters based on the Mercator projection (Gruber *et al.*, 2018). Summary statistics of the cleaned datasets were conducted using GenoDive (version 3.05) to assess observed heterozygosity (Ho), heterozygosity within populations (Hs), and inbreeding coefficients (*Gis*) (Meirmans, 2020).

2.4. Results

We retained 4,899 genomic SNPs and 175 individuals within the dataset after quality filtering. Pairwise Fst values ranged from 0 to ~0.14 and were largely significantly different population-to-population, regardless of collection year (Table 2.2 and Table 2.3). The exceptions to this were the two Grand Prairie timepoints (2007 and 2016) which were non-significant, the initial Canmore and Golden site sampling (2007 and 2005 respectively), and the initial 2007 Grand Prairie sampling and the 2016 Canmore sampling. Observed heterozygosity (Ho) was comparable on all sites (0.293 – 0.265) with no major trends connected to sampling date. The highest Ho was detected at the Golden site from 2007. Inbreeding coefficient values ranged from 0.093 to -0.036 and were highest at the

2016 collection at Canmore (Table 2.1).

Both ΔK (Figure 2.3) and Ln(PrK) metrics from our STRUCTURE analysis support an optimal K of 2 at both time points, indicating two detectable population clusters on the landscape (Figure 2.2). Canmore showed a southern affinity during the first sampling and a mixed northern and south affinity during the second sampling. Smithers at the time of the first sampling showed a more mixed population and in the 2016 sampling grouped completely north. In contrast, Robson moved from a blend of individuals assorting fully to south and north, to most individuals having probability of cluster membership divided evenly between north and south (this admixed signature is addressed more fully in Chapter 3). Golden remained southern in nature at both sampling times, though there was a low-level increase in probability of northern assignment. Finally, Grande Prairie maintained a northern population assignment at both sampling times. An additional cluster (K = 3) is not supported by the data and additional K values beyond 3 are also not supported. Clusters were also assessed using the find.clusters algorithm in adegenet which relies on Bayesian information criterion (BIC) to infer optimal populations structures. K = 2 was also supported in this analysis. Isolation by distance analysis of sites grouped by both year and taken together showed no pattern of geographic population differences by distance, $r^2 = 0.2189$ (Figure 2.5).

DAPC and PCA analysis echoed the results of the STRUCTURE analysis with association of populations with collection year, but more association with geographic clustering. In the case of both STRUCTURE and DAPC (80 PCs retained), sites that were in the expanding path of the BC outbreak took on a more northerly signature (ellipses overlapping), with Mount Robson Provincial Park taking an intermediate position and

Golden remaining distinct (Figure 2.4C). PCA investigation of both time points taken together showed roughly the same geographic clustering reflected in the STRUCTURE analysis, with PC 1 explaining 9.68% of the variation found in the SNP dataset and PC 2 explaining 4.5% of the remaining variation. PC 1 is likely linked to geographic location, while PC 2 is most likely linked to variation between individuals (Figure 2.4A) (Shegelski *et al.*, 2021). Analysis of molecular variance (AMOVA) of 4733 loci was used to test the differences between sites and to look for similarity based on year or location (Table 2.4). Our analysis found that there was considerable variation within populations and between individuals. Grouping populations by year explained the least amount of variation. All Fstatistics sampled (Fis, Fsc, Fct, Fit) were significant ($p < 0.0001$).

2.5. Discussion

Our study assessed the changes in population structure among MPB populations sampled during the height of the British Columbian outbreak of the early 2000s and those sampled approximately ten years after expansion and of the initial population to the north and east. We explored the possibility that beetles would retain their original population signature in the path of an expanding outbreak due to dispersal limitations based on their flight capacity and proximity to attractive viable host trees. We found instead that populations in the path of the outbreak (Canmore in the east to Alberta and Smithers in the north) transitioned from more southerly or mixed assignments to more northernly assignments and took on the characteristics of the large spreading BC outbreak. However, beetles in all collection years across all sites still had a distinguishable north-south population divide, with two detectable genetic clusters on the landscape that assorted geographically, even though there was no discernible isolation-by-distance detected. It is

likely that the smaller number of total individuals in this study and the greater number of sites that assort to the north obscured any pattern of isolation by distance. This echoes previous research on MPB in western Canada that also has shown a geographic north/south divide between populations (Samarasekera *et al.*, 2012; Janes *et al.*, 2014; Batista *et al.*, 2016; Shegelski *et al.*, 2021).

The 2016 samples from Mount Robson Provincial Park population had an intermediate split of populations assigned between north and south for all individuals. Mount Robson Provincial Park is located close to the area that was identified by Janes *et al.* (2014) as region of higher genetic admixture within BC and as an area of genetic mixing by Samarasekera *et al.* (2012). The original sampling of Robson in 2005 produced more individuals that assorted completely north and completely south (a mixed population) as opposed to these more admixed individuals collected in 2016 (Figure 2.2). This indicates that the Robson area in 2005 contained beetles from both landscape clusters that have likely since interbred to produce a blended population. It is important to note that STRUCTURE analysis did not identify a distinct population in this area as a separate cluster, nor did the find.clusters algorithm with this dataset.

AMOVA analysis of the SNP dataset indicated that while MPB individuals have high genetic variability among each other, both populations and collection-time based groupings produced small but significant differences. The recent expansion of the MPB outbreak, combined with the presence of two population clusters and mixed populations within the dataset likely reduced the size of the difference between the two time points, but it is important to recognize that there are changes in the geographic population structure between 2005/2007 and 2016. There are significant site-to-site pairwise Fst differences

over time in almost all sites sampled in 2005/2007 and 2016, likely reflecting a history of differentiation due geographic isolation with limited long-distance dispersal between populations prior to the large MPB outbreak in BC during the mid-2000s.

The mid-2000s outbreak moved both north and east, leaving southern BC outside of the path of most dispersing MPB in the north of the province. In addition to the Golden population's isolation in a high, remote mountain valley in the south of BC, the outbreak's movement away from the southern locations following prevailing winds likely contributed to the maintenance of the distinct southern structure detected at Golden (Figure 2.4 B and C). The two Grande Prairie MPB collections in the north of Alberta did not differ by pairwise Fst, likely due to the relatively recent establishment of the population and the fact that all MPB populations in the expanded range area come from the northern BC epidemic. Approximately ten generations in one location in a region without repeated population replenishment from other MPB sources is likely insufficient to cause a significant change in the original genetic structure.

In most of British Columbia, MPB populations have now returned to endemic levels due to lack of food resources and extremes of temperature in the expanded range in the north of the province. Alberta has seen similar declines due to very vigorous control efforts, and a series of non-ideal climate conditions including excessively cold winters in 2019 and 2020, and an unseasonably cool and wet summer in 2019 (M. Undershultz, personal communication, April 22, 2022). The removal of the host pine by several consecutive summers of intense wildfire in both provinces is also likely driving beetle numbers down, returning both areas to an endemic level of infestation (Tan *et al.*, 2019; Daniels *et al.*, 2020).

MPB populations in northern Alberta, specifically in the Grand Prairie area and eastward, are considered locally invasive as there has never been a recorded instance of MPB colonization prior to 2006 (Burke, 2016). This is despite the fact that the Canmore population in southern Alberta is connected to the well-established Banff populations that have been documented since the 1940s (Powell, 1961; Cooke & Carroll, 2017). MPB in Canmore has taken on the more northernly genetic signature suggesting that gene flow toward the park is significant enough to prevent new local adaptations from establishing at this time (Allendorf *et al.*, 2013). In addition, the presence of northernly beetles in Canmore may indicate that the northern population is beginning to destabilize existing genetic structure in the Banff National Park area.

Despite the shorter geographic distance between Robson, located between north and south clusters, and Canmore compared to Grand Prairie and Canmore, STRUCTURE grouped most MPB individuals in Canmore in the south with Grande Prairie in the north which may indicate either long-distance dispersal events from the colonizing MPB populations in Alberta or local transportation of infested wood into the area. Transportation of firewood is discouraged specifically to prevent movement of pest insects (Government of Alberta, 2008), but Banff National Park and the surrounding towns accommodate a large number of tourists with demands for camp wood that may be sourced from infested areas in the north of BC or Alberta (Cheng, 1980). There is also the continued possibility of long-distance dispersal throughout Alberta naturally, as prevailing winds in all seasons move west to east through the north of the province, and down towards Canmore in the summer (Government of Alberta, 2014).

Though MPB outbreaks within the expanded range in central Alberta are in a state

of flux, our study indicates that epidemic groups of beetles can establish a genetically homogenous population on a landscape even when separated by great distances. In the case of Alberta, the first colonizing wave of MPB from the northern BC outbreak has left a strong signature on the landscape. We also have found that there can be shifts in population assignment between MPB populations over short periods of time, meaning that care should be taken when combining sample cohorts from different years, particularly for irruptive species like MPB.

We found that regional genetic structure can be lost or altered in the face of epidemic level in-flights of beetles. Most of the beetles found in the northern parts of both BC and Alberta assort to one homogenous population. For this reason, the development of population-specific control methods should not be used over established methods of control as MPB populations on the landscape are likely not yet displaying meaningful behavioural differences site to site. Traditional MPB management methods tailored to regional differences in stand compositions and site characteristics including: spot eradication through fall and burn, sanitation cuts, controlled burns, and other forms of host denial (Fettig *et al.*, 2014) are all likely to be effective on current outbreaks.

The genetic characteristics of sites like Grande Prairie and Canmore are also important to document as they provide context for the character of MPB populations that are likely to remain in Alberta. To our knowledge, this is the only study to date that has compared the population genomics of an irruptive beetle pest by assessing genetic changes by time cohort. As repeated collections are now encouraged to track changes in genetic demography (Hagen *et al.*, 2015), our data support the separation of collection cohort to better understand how irruptive pest movements are influencing population structure.

2.6. Acknowledgements

We would like to thank the British Columbia Ministry of Environment and Climate Change Strategy (Park Use Permit, Authorization # 107171), Alberta Agriculture and Forestry, and Robson Provincial Park staff for providing guidance on potential site locations. This research was supported in part by funding provided by the Natural Science and Engineering Research Council of Canada (grant no. NET GP 434810-12) to the TRIA Network. Computational work was supported by the cluster resources provided by Compute Canada (www.computecanada.ca).

Table 2.1: MPB site collection date, location, observed heterozygosity (Ho), heterozygosity within populations (Hs), and inbreeding coefficient (Gis).

Population	Site Code	Latitude	Longitude	N	Year	Stage	Ho	Hs	Gis
Canmore_1	M11	50.93900	-115.14220	9	2007	Larva	0.271	0.293	0.077
Canmore_2	xS	51.06070	-115.27756	28	2016	Adult	0.266	0.293	0.093
Golden_1	M06	51.07400	-116.38200	13	2007	Larva	0.293	0.288	-0.018
Golden_2	xT	50.84039	-116.63792	17	2016	Adult	0.282	0.298	0.055
Grande Prairie_1	M01	54.79882	-119.79597	12	2007	Adult	0.273	0.275	0.01
Grande Prairie_2	xL	54.62643	-119.82640	20	2016	Adult	0.285	0.275	-0.036
Robson_1	YH	52.89490	-118.73480	21	2005	Adult	0.265	0.292	0.091
Robson_2	A	52.84985	-118.57265	14	2016	Adult	0.279	0.294	0.051
Smithers_1	TE	54.66740	-127.08870	24	2005	Adult	0.276	0.293	0.061
Smithers_2	xX	54.94973	-127.38175	17	2016	Adult	0.283	0.276	-0.024

Table 2.2: Significance of pairwise Fst values based on P-values generated in Arlequin (+ = $P < 0.05$).

	Robson_1	Grand Prairie_1	Golden_1	Canmore_1	Smithers_1	Smithers_2	Robson_2	Grand Prairie_2	Golden_2	Canmore_2
Robson_1										
Grand Prairie_1	+									
Golden_1	+	+								
Canmore_1	+	+	-							
Smithers_1	+	+	+	+						
Smithers_2	+	+	+	+	+					
Robson_2	+	+	+	+	+	+				
Grand Prairie_2	+	-	+	+	+	+	+			
Golden_2	+	+	+	+	+	+	+	+		
Canmore_2	+	-	+	+	+	+	+	+	+	

Table 2.3: Heatmap of Pairwise Fst values for MPB sites generated in Arlequin. Red indicates the highest levels of Fst while green indicates the lowest.

	Robson_1	Grand Prairie_1	Golden_1	Canmore_1	Smithers_1	Smithers_2	Robson_2	Grand Prairie_2	Golden_2	Canmore_2
Robson_1	0.00									
Grand Prairie_1	0.03	0.00								
Golden_1	0.04	0.11	0.00							
Canmore_1	0.04	0.11	0.00	0.00						
Smithers_1	0.02	0.04	0.03	0.03	0.00					
Smithers_2	0.06	0.01	0.13	0.14	0.06	0.00				
Robson_2	0.01	0.02	0.05	0.05	0.02	0.04	0.00			
Grand Prairie_2	0.03	0.00	0.11	0.11	0.04	0.01	0.02	0.00		
Golden_2	0.05	0.09	0.03	0.04	0.04	0.10	0.05	0.09	0.00	
Canmore_2	0.01	0.00	0.06	0.06	0.02	0.02	0.01	0.01	0.05	0.00

Table 2.4: Global AMOVA design for MPB individuals and results as a weighted average over loci (averaged over 4733 loci). An asterisk denotes significant values.

Source of variation	SS	Var. comp.	% var	Fstat
Among groups	2305.948	3.44104	0.48023	FCT :0.00480*
Among populations within groups	12859.198	27.21377	3.79797	FSC :0.03816*
Among individuals within populations	113213.075	29.94989	4.17983	FIS :0.04367*
Within individuals	110537.500	655.92895	91.54196	FIT :0.08458*
Total	238915.720	716.53365	---	

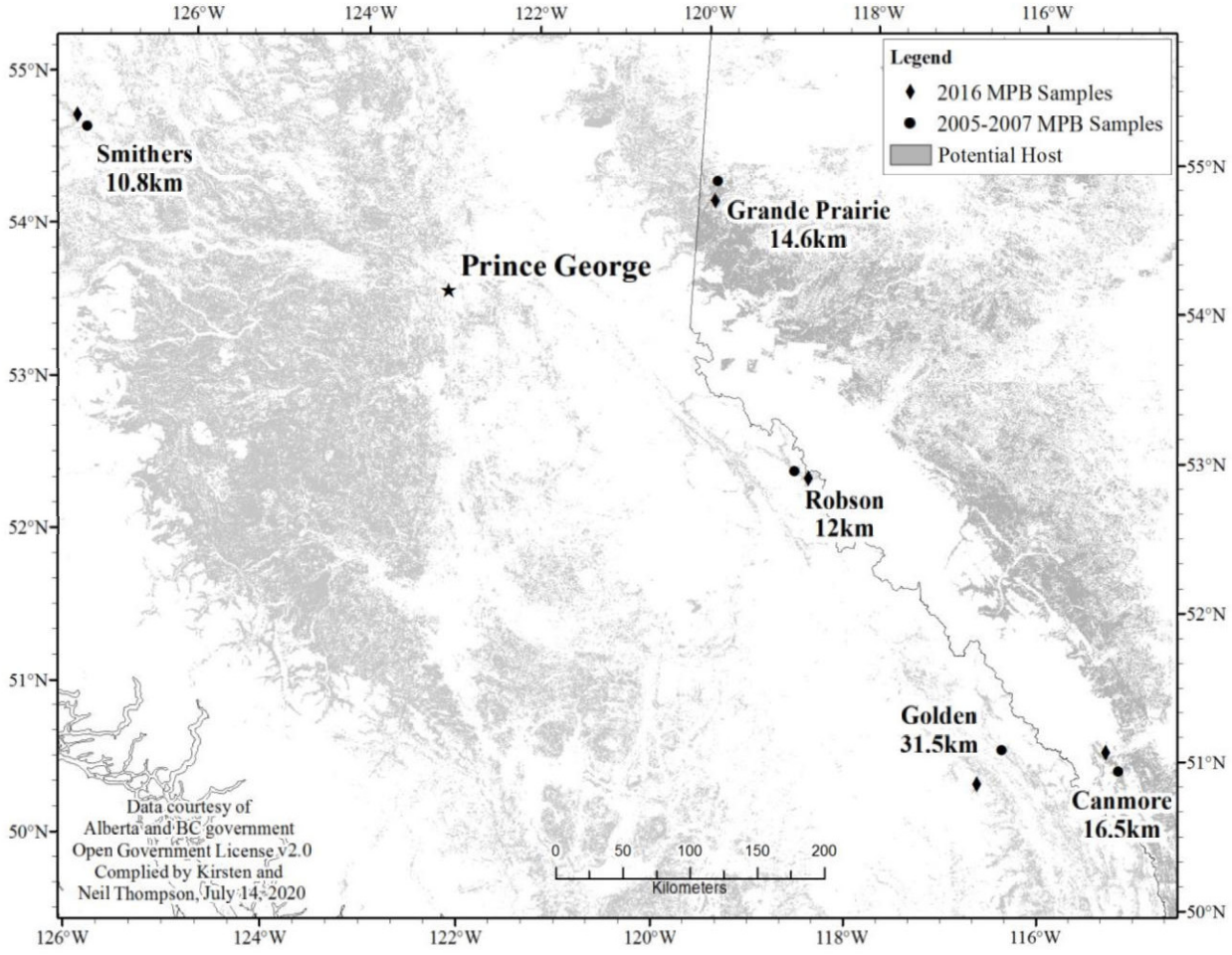

Figure 2.1: Map of MPB collection sites sampled in both 2005/2007 and 2016 with the distances between site locations.

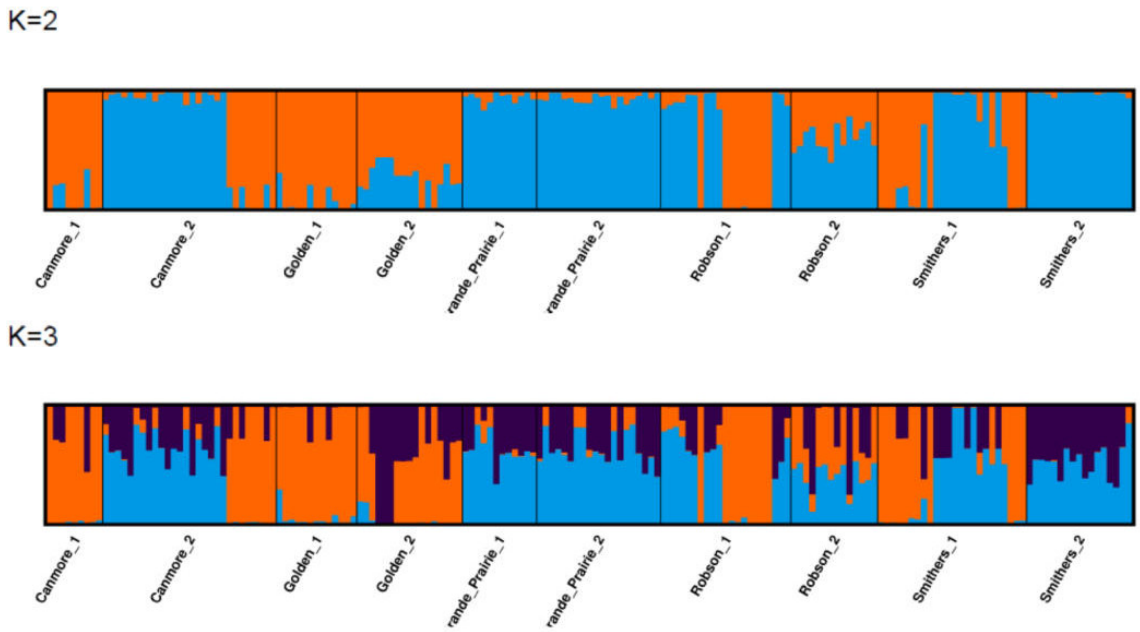

Figure 2.2: STRUCTURE plots with cluster membership per individual MPB (n=175) averaged by CLUMPAK. Sites visually divided by a black bar and, paired together by year, and analyzed at 4899 loci. Individuals are represented by a partitioned vertical bar with cluster membership colour coded by K. The proportion of colour within the bar represents the probability of the individual's assignment to each cluster. For K = 2, orange denotes a southern cluster while blue denotes a northern cluster.

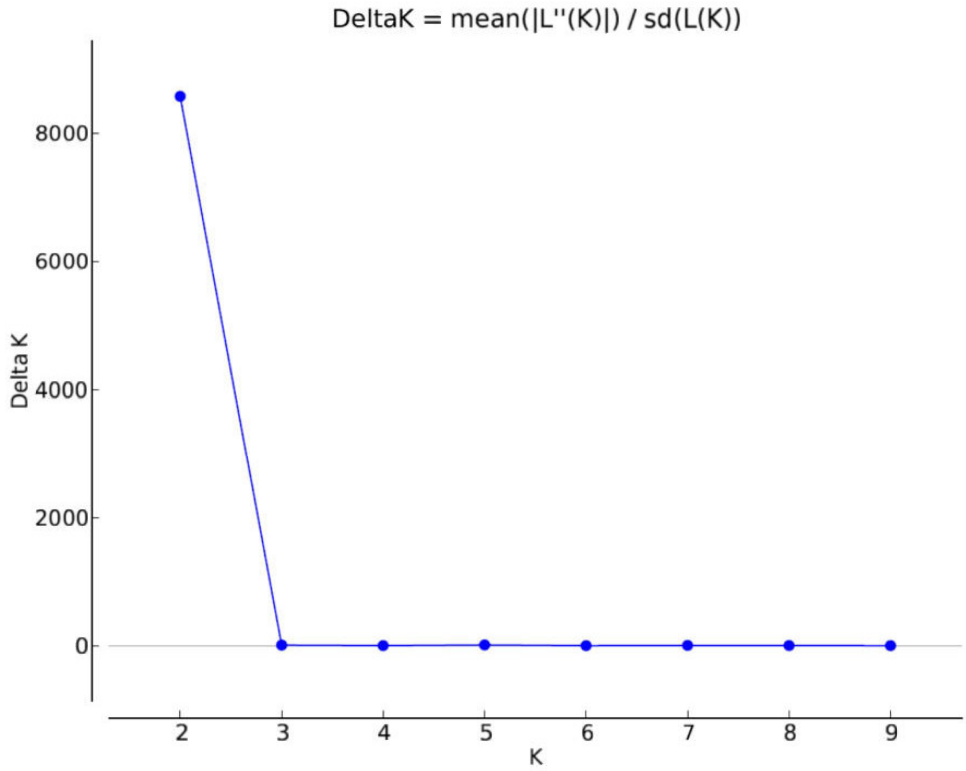

Figure 2.3: ΔK method plot with the optimal K of 2 for MPB, calculated in CLUMPAK using the Evanno method.

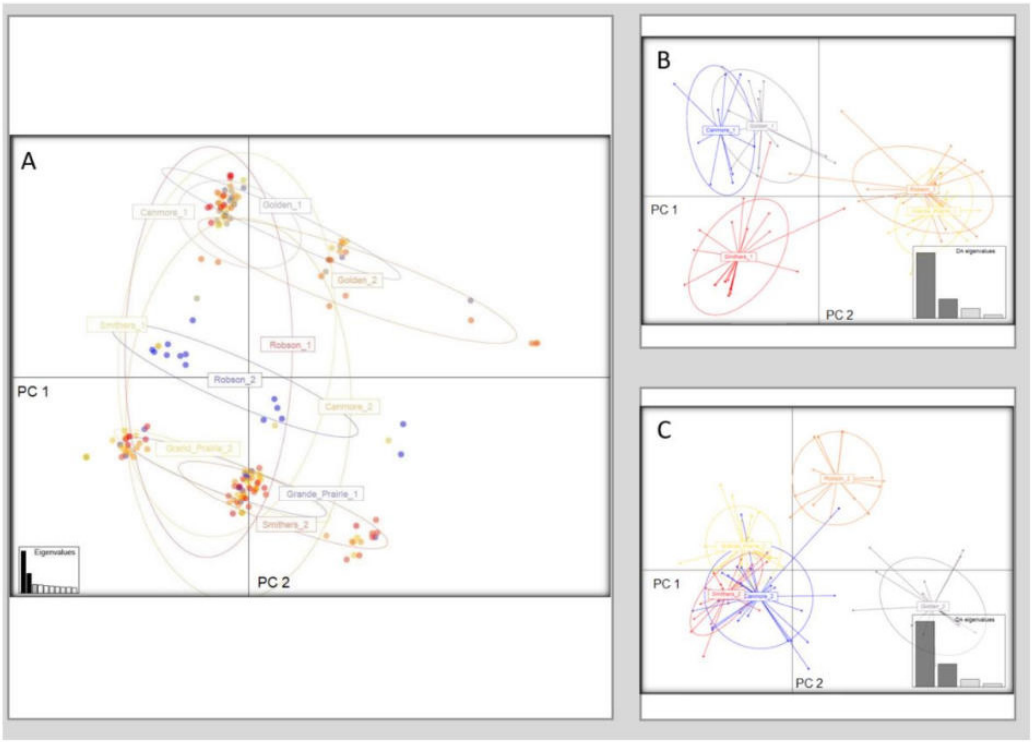

Figure 2.4: A) PCA of all MPB sites showing PC 1 and PC 2 and respective eigenvalues. B) DAPC scatterplot of MPB SNP genotypes displaying principal components 1 and 2 of years 2005/2007 C)DAPC scatterplot of MPB SNP genotypes displaying principal components 1 and 2 of year 2016.

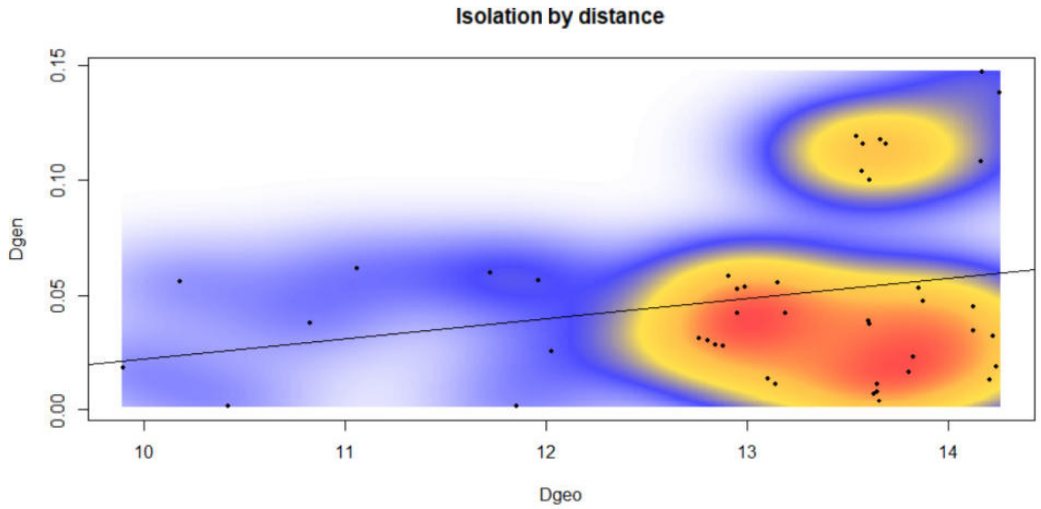

Figure 2.5: Isolation by distance for sites in 2005/2007 and 2016. The solid line through the points represents a Pearson-Moment correlation (non-significant) for the Mantel test (r2 = 0.2189, p = 0.102).

3. Chapter 3: Mountain Pine Beetle Populations Retain Characteristics of Source Populations a Decade After Range Expansion and Do Not Display Rapid Specialization of Genetic Traits

Kirsten M. Thompson, Dezene P. W. Huber, Chris J. Johnson, Felix A. H. Sperling, Brent W. Murray

3.1. Abstract

Mountain pine beetle, *Dendroctonus ponderosae* (Coleoptera: Curculionidae) a natural irruptive forest pest of pines in western North America, has recently undergone a massive range expansion in novel parts of western Canada, and as such has experienced high levels of gene flow across many thousands of hectares. In this study, we use SNPs (single nucleotide polymorphisms) to explore the population structure of this range expansion approximately ten years after the initial establishment in addition to identifying putative outlier loci and related environmental predictors. We used a ddRADseq protocol to generate SNP data via single-end Illumina to sequence whole genomic DNA. We identified 3612 SNPs from 1018 individuals across 55 sites in a ~ 1,000,000 km^2 area, sampled from 2016–2018.We found a clear north/south population split within the expanded range, with the possible emergence of a third cluster on the landscape near Jasper National Park. Fst values are significantly different between most sites, with blocks of low Fst between Jasper, the Southern cluster, and the Northern clusters. We used RDA analysis to assess the impact of landscape factors and found that mean annual precipitation, relative humidity, and frost free period contribute most to outlier SNPs across the sampling area. The remaining active MPB infestations in the north of British Columbia and Alberta all retained similar genetic characteristics and did not yet display evidence of rapid evolutionary change.

3.2. Introduction

Many mobile organisms have broad ecological niches and the capacity to expand over large geographic areas (Lal *et al.*, 2016; Maclauchlan *et al.*, 2018; Yadav *et al.*, 2019). This capacity for rapid dispersal and lack of landscape barriers often leads to states of high gene flow which can obscure local signatures of selection and mask weak selection events (Malet, 2001; Allendorf *et al.*, 2013). This phenomenon is particularly prevalent in organisms that have contiguous habitats, whether marine or terrestrial, as the low cost of movement creates areas of near-admixture conditions (Geffen *et al.*, 2004; Grummer *et al.*, 2019). However, even in areas of considerable gene flow, local selection processes still occur and can overcome the homogenizing effects of repeated mixing of alleles, leading to the establishment of locally distinct populations (Räsänen & Hendry, 2008). These populations show the influence of selective forces, such as climate, through retention of adaptive loci.

Climate change is causing rapid range expansions and habitat changes for a variety of species worldwide (Parmesan, 2006; Dawson *et al.*, 2008; Kirk *et al.*, 2013b). While more sensitive niche-bound organisms are often at a disadvantage as climate shifts (Damschen *et al.*, 2010), mobile pest insects have benefited from longer, warmer summers that decrease generation time and strain host defenses (Raffa *et al.*, 2008; Sgrò *et al.*, 2016). Neutral genetic markers including microsatellites and SNPs have been used to explore landscape variables and breeding barriers for pest insects (Janes *et al.*, 2014; Dowle *et al.*, 2017; Yadav *et al.*, 2019). Adaptive variation has also been studied in agricultural pest systems with human selection inputs, such as pesticides (Crossley *et al.*, 2017), though neutral-marker population structure remains the key focus of most studies

(Duan *et al.*, 2017; Zhang *et al.*, 2019). Despite widespread acknowledgement of the far reaching impacts that climate change is expected to have on the success and spread of insect species into new regions (Hofstetter & Gandhi, 2022), the interaction of climate factors on adaptive loci has not yet been thoroughly explored in pest systems (Yadav *et al.*, 2019).

Mountain pine beetle, *Dendroctonus ponderosae* Hopkins (Coleoptera: Curculionidae) (MPB), is an example of a highly mobile forest pest experiencing climate change-related success. MPB is a subcortical-feeding irruptive bark beetle native to western North America (Safranyik & Carroll, 2006). Prior to 2006, MPB was found in pine forests ranging from southwest USA and the Black Hills in South Dakota, to parts of the interior of British Columbia, and some highly isolated populations in Banff National Park and Cypress Hills Alberta (Mitton & Sturgeon, 1982; Emond & Cerezke, 1991; Janes *et al.*, 2016). Historically, MPB persisted in small, endemic populations attacking weak, stressed pine trees in low numbers and periodically progresses from endemic numbers to outbreaks of moderate to large size (Powell, 1961).

Low interference pine stand management practices, fire exclusion, and increasingly warm winters in the late part of the 20th century allowed for the development of massive populations of MPB that eventually coalesced into an epidemic-level outbreak of unprecedented scale, resulting in trillions of beetles moving across 18.3 million hectares of pine forests in British Columbia alone (Bentz *et al.*, 2010; Gillette *et al.*, 2014; McKee *et al.*, 2015). This allowed for further beetle movement into the north of British Columbia, and novel range expansion over the Rocky Mountains into central Alberta (Cullingham *et al.*, 2011). The range expansion of MPB led to the death of millions of hectares of pine and

serious economic and social effects for forest-dependant communities in British Columbia (Turner & Clifton, 2009; Fettig *et al.*, 2014; Howe *et al.*, 2022).

Insect pests with a broad climatic niche, like MPB, are often successful due to their ability to interact successfully with a variety of environmental factors. MPB already demonstrates an ability to attack and reproduce in a variety of pine species across broad ranging ecological conditions in western North America. Following epidemic conditions, multiple populations of MPB have demonstrated genetic mixing across large areas (Janes & Batista, 2016). For this reason, we would expect unprecedented dispersal across BC and into Alberta to lead to relatively diffuse population structure with the potential of beneficial subsets of loci showing signs of selection whether by genetic hitch-hiking, past selection events, or influence of the novel environment. Neutral population structure of MPB across portions of the expanded territory have been studied at several points during the recent outbreak (Samarasekera *et al.*, 2012; Janes *et al.*, 2014; Batista *et al.*, 2016), sampling heavily within the historic BC range and into parts of Alberta.

One such study of post-expansion outbreak MPB populations in Alberta found that beetles in some parts of northern Alberta clustered as one genetic population (Janes *et al.*, 2016), in addition to other studies that have found evidence of outlier loci linked to cell signalling and ion transport (Batista *et al.*, 2016). However, the entirety of the new MPB range border within Alberta and British Columbia has not been yet been sampled, nor have these previous studies specifically addressed how landscape and climatic factors may have influenced adaptive loci of these populations. Previous studies have also used relatively restrictive sets of markers, as opposed to genotyping-by-sequencing generated SNPs that represent the entire genome. Recent advances in the field of landscape genomics have

resulted in new methods to quantify genotype-environment relationships. Those methods, including redundancy analysis (RDA) have great utility for detecting weak signatures of selection in multi-locus datasets (Forester *et al.*, 2018). RDA is designed for use in populations with low structure and high movement, which makes it an ideal technique for studying MPB genomics. RDA also allows for the assessment of both neutral population structure and adaptive selection on the same dataset to delineate evolutionary response with the expanded range of MPB.

MPB is likely to establish a permanent presence in Alberta due to size of the new area colonized and the difficulty of detection and removal of beetles when populations regress to low levels. The genetic characteristics of remaining active MPB populations on the landscape will influence the behaviour of future MPB outbreaks within Alberta by providing the genetic foundation of future spread. For this reason, our objective is to thoroughly catalogue the dominant population structure of MPB on the landscape 10 years after the establishment of beetle populations in novel territory in central and northern Alberta and BC. We also seek to identify the influence of environmental variables from this novel range on the genetics of resident MPB populations, searching for patterns of adaptive genetic variation due to selection events. Our use of environmental climate and landscape data follows recent recommendations in environmental association analysis (Rellstab *et al.*, 2015; Hoban *et al.*, 2016) in order to provide a complete picture of how MPB is adjusting and interacting with novel habitats as it moves eastward. We hypothesize that MPB within the expanded range in central and northern Alberta retains the genetic signature of the original northern BC outbreak and that temperature metrics exert the greatest influence on MPB outliers due to evidence from research on cold tolerance genes

and MPB ectothermic physiology (Régnière & Bentz, 2007; Robert *et al.*, 2016).

3.3. Methods

A note to the reader: these methods for SNP generation are the same as those used in Chapter 2, though they have been rewritten using different text in anticipation of both chapters being submitted for publication.

3.3.1. Beetle Collection and Extraction of Genomic DNA

Live MPB were collected in the summer of 2016 (5) and again in 2017 (5). An additional site from 2018 was provided by V. Shegelski. (Figure 3.1). Aerial survey data provided by the BC Ministry of Forests, Lands and Natural Resource Operations (https://catalogue.data.gov.bc.ca/, 2022) and data from Alberta Agriculture and Forestry field crews was used to identify active areas of beetle infestation (C. Whitehouse, personal communication, June 2016). Sites were chosen on approximate axes that spread north and east in Alberta, and south and west in BC. Yellow foliage and the presence of pitch tubes with frass were used as an indicator to identify the trees with current MPB infestation. Bark was peeled from the bole of selected trees using a draw knife, and beetles were chosen from different galleries in order to prevent the collection of full siblings. In most cases, the boles of the trees were peeled completely to a height of 2 m to make an accurate assessment of the connection of galleries. Approximately 40 beetles were sampled at each site, with a maximum of four individuals per tree selected randomly for analysis. Both adult and teneral (freshly metamorphosized) beetles were selected from trees in Alberta and British Columbia. Collected live beetles were placed in 95% ethanol and then frozen at -20°C for long-term storage.

Whole genomic DNA was extracted from the collected beetles using the Qiagen

DNeasy Blood & Tissue spin column and 96 well plate kits (Germantown, Maryland, USA). Extractions were conducted using the default Qiagen protocol, with an overnight tissue lysis step incubated at 56℃ and an RNA lysis step using 2µL of RNaseA with a 15-minute incubation. To make sure that future amplification and sequencing did not require extra purification steps, DNA was eluted into DNase-free water instead of elution buffer. DNA concentration was measured twice, using both a first-generation Qubit fluorometer with the Quant-iT dsDNA BR Assay Kit and a NanoDrop™ 1000 Spectrophotometer.

3.3.2. DNA Sequencing and Data Set Construction

We used a double-digest restriction-site associated DNA sequencing (ddRADseq) protocol, a genotype-by-sequencing method, to generate whole genome SNP markers. Genomic libraries were constructed from the whole genomic DNA. These libraries were sequenced by the Molecular Biology Services Unit (MBSU) of University of Alberta (Peterson *et al.*, 2012; Campbell *et al.*, 2017). The weight of purified DNA per sample was 80ng per individual. We used a non-methylation sensitive two-enzyme system (*Pst* I and *Msp* I) to fragment whole genomic DNA for sequencing, run on a single lane of an Illumina NextSeq 500 at the MBSU, producing single-end 75 bp reads.

Sequences generated on Illumina machines must be computationally processed prior to analysis. Sequence reads were demultiplexed and mapped on Compute Canada's Graham cluster (Toronto, Ontario, Canada). We used the default STACKS 2.0b pipeline to perform initial demultiplexing and quality checking (Rochette *et al.*, 2019). We trimmed the *Pst* I restriction site from the 5'-end of reads using Cutadapt (version 1.9.1), trimming the reads to 67 bp in length (Martin, 2011). Using BWA-MEM (Burrows-Wheeler Aligner Maximal Exact Match version 0.7.17), these trimmed reads were then mapped onto the

female MPB reference genome and quality checked using SAMtools (version 1.9) (Li & Durbin, 2009; Li *et al.*, 2009a; Keeling *et al.*, 2013). We then used the ref_map.pl perl wrapper from STACKS 2.0b to call SNP variants and produce a raw VCF file, using a minimum minor allele frequency of 0.05, and the r80 principle, we only retained loci that appeared in 80% of all MPB samples (Paris *et al.*, 2017; Rochette *et al.*, 2019).

We used VCFtools (version 0.1.14) to perform additional filtering, with a minor allele frequency of 0.05, a minimum quality score of 30, and setting minimum read depth to 10 (Danecek *et al.*, 2011). MPB is known to have a strong signal from sex-linked markers, so to address this signal we loaded the VCF datafile into R (version 3.6.3) and filtered out sex-linked SNPs using the PCA method outlined in Trevoy *et al.* (2019), removing loci associated with high peaks in PC 1 distribution (R Core Team, 2018). We then filtered for linkage disequilibrium (R disequilibrium ($R^2 > 0.5$) using the R package dartR (version 1.1.11), retaining one random SNP from each linkage group in order to prevent excessive clustering from confounding population structure analysis (Abdellaoui *et al.*, 2013; Gruber *et al.*, 2018). We then rechecked individuals for poor quality data after linkage filtering and removed individuals with > 20% data missing.

3.3.3. Analysis of Population Structure

Preliminary assessment of range-wide population structure was conducted using the adegenet package (version 2.1.2) in R (Jombart, 2008; Jombart *et al.*, 2010) to perform both PCA and DAPC analysis. PCA and DAPC analysis were used to look for structural or geographic patterns within the SNP dataset. PCA was performed without *a priori* assignment to population clusters on the landscape. We also used the find.clusters algorithm in adegenet to assign individuals to putative populations on the landscape for

later analysis. DAPC was used to further explore populations structure as it tends to show greater separation between closely related groups with weak differentiation (Jombart *et al.*, 2010). The cross-validation process recommended in the DAPC vignette were used during the visualization process to retain the optimal number of PCs and DF (Jombart *et al.*, 2010).

For our analysis, we used the admixture model without using location information in STRUCTURE (version 2.3.4) to group individuals into populations (Pritchard *et al.*, 2000; Falush *et al.*, 2003, 2007; Hubisz *et al.*, 2009). Our primary analysis of this dataset used 100,000 Markov chain Monte Carlo generations and a burn-in period of 50 000 for 10 replicates of K values set between 1 and 10. We used this first run to limit the value of K to between 1 and 5 and then increased our STRUCTURE run to 1 million Markov chain Monte Carlo generations and a burn-in period of 250, 000, also with ten replicates. We used CLUMPAK (Kopelman *et al.*, 2015) to visualize our results by creating mean values for each K. We determined optimal K by comparing the results of the ΔK Evanno method (Evanno *et al.*, 2005) between likelihood of K values and the mean estimated natural logarithm for the K value probability (Ln(PrK)) (Pritchard *et al.*, 2000). Both ΔK and Ln(PrK) were compared using the online software package STRUCTUREHARVESTER (Earl & VonHoldt, 2012). Summary statistics and a hierarchical AMOVA was conducted in GenoDive (version 3.05) to assess observed heterozygosity (Ho), heterozygosity within populations (Hs), and inbreeding coefficients (Gis) (Meirmans, 2020).

We used Arlequin (version 3.5.2.2) to compare pairwise Fst values generated for the 55 sites, set at a 0.05 significance level with 1000 permutations (Excoffier & Lischer, 2010). Isolation by distance was tested using a Mantel test in adegenet (version 2.1.2) in

conjunction with package ade4 (version 1.7-15) with a Monte-Carlo test with 999 permutations on the reduced dataset used for RDA (see section 3.3.4 for dataset modification) (Gruber *et al.*, 2018).

3.3.4. Detection of Outlier SNPs

We used two different methods to detect outlier SNPs within the full range-wide dataset: OUTFLANK for loci identification and redundancy analysis (RDA) to identify outlier SNPs associated with landscape variables. We used OUTFLANK (version 0.2), which uses a likelihood method to calculate Fst outliers from an inferred neutral Fst distribution based on a trimmed distribution of Fst. We implemented OUTFLANK using the default settings with a q-value threshold of 0.05, which equates to a false discover rate (FDR) of 5% (Whitlock & Lotterhos, 2015). Putative outlier SNPs were isolated with their flanking sequence and annotated using NCBI BLAST (Johnson *et al.*, 2008). SNPs were compared to the MPB transcriptome using the blastx mode. Annotations were inspected for query coverage and retained based on E values of less than 0.0001.

Redundancy analysis (RDA) is a type of genotype and environmental association (GEA) method that can be used to identify selection for putative loci in datasets with many variables and weak selection signatures. RDA uses multivariate linear regression to produce fitted values that are then used to create a PCA matrix from which significant environmental variables are identified. This approach is particularly well suited to species with relatively low population structure, or those that are highly mobile (Rellstab *et al.*, 2015; Forester *et al.*, 2018).

For our RDA analysis we used the package vegan (version 2.5-6) in R with the default settings (Oksanen *et al.*, 2020). Environmental variables were collected for each

site based on GPS locations, extracted from ClimateNA, using the default settings, including temperature, precipitation, and humidity data (Wang *et al.*, 2016a). Correlation between the environmental predictors was assessed using the R package psych (version 2.0.12) and highly correlated variables were removed from the dataset (Revelle, 2010). Variables were initially selected based on their possible influence on insect or host biology, including precipitation, temperature, and radiation metrics (Ojeda Alayon *et al.*, 2017; Yadav *et al.*, 2019; Sambaraju & Goodsman, 2021).

Elevation data were collected using Google Earth and an estimation of vegetation cover was produced using Forest Elevation (Ht) Mean 2015 data calculated as the percentage of cells in a 500m radius plot with greater than zero vegetation height (https://opendata.nfis.org/mapserver/nfis-change_eng.html, 2015). Cluster identification per site was derived from the results of the find.clusters algorithm previously mentioned. We collapsed our site data down to allele frequencies within demes by taking the most common allele identity (mode) per locus from each site. This was also done to reduce the computational intensity of significance tests on the constrained axes. A subset consisting of beetles from Jasper National Park was also analyzed by RDA.

3.4. Results

After quality filtering the range-wide dataset, we obtained 3612 SNPs from 1018 individual beetles. STRUCTURE analysis of this dataset using both ΔK and Ln(PrK) methods supported two genetic clusters (K = 2) as optimal across the entirety of British Columbia and Alberta. Sites from the southern portion of BC were grouped as one population, while all other sites north and east of Jasper National Park assorted into a northern cluster. This pattern was subtly different at the Canmore site (xS – seen in Ch 2)

which had many individuals assort completely to the northern cluster, one of the two Smithers sites (xY) and one of the three Grand Prairie sites (xM) that had a mix of northern and southern assorting individuals. Additional population clusters (K = 3 and greater) were not supported, however, individuals from the Jasper National Park area had approximately equal probability of membership within the north and south clusters and did not cleanly assort to either genetic cluster (Figure 3.2).

In addition to STRUCTURE, we also used the find.clusters algorithm in adegenet, along with DAPC and PCA to explore population structure. In this case, BIC analysis (Bayesian Information Criterion) indicated that optimal clustering showed three groups (K = 3) with beetles from the Jasper area assorting as a third cluster on the landscape (refer to Figure 3.6 for cluster identity from adegenet). The third cluster represents the populations from Jasper National Park that STRUCTURE calls as individuals that assort equally between north and south. PCA analysis also shows with three clusters instead of two on the landscape, though it is clear that there is some degree of genetic variation between individuals from the same sites (Figure 3.4A). PC 1 explains 7.31% of the variation within the dataset, with PC 2 explaining a further 3.04% of variation. PC 1 represents geographic coordinates, while PC 2 shows more variation and relates to individual genotype differences. For DAPC (with 400 PCs retained and 54 DF), we see the separation of three groups (northern, Jasper region, and southern) that corresponds with the three clusters identified by find.clusters (Figure 3.4B). There was no broad pattern of isolation-by-distance on the landscape despite the presence of disjunct populations ($r^2 = 0.07$, $p > 0.2$).

Analysis of molecular variance (AMOVA) of the 3612 SNP loci tested the genetic similarity within populations, among populations, and among clusters on the landscape

(Table 3.2). The greatest amount of variation was explained by within populations and not among populations on the landscape (93.3% and 1.2% respectively). Pairwise Fst values for all sites varied between 0 to ~ 0.12 with the highest pairwise values found between the southern sites and the sites assorting to the northern cluster, excepting Canmore, one Smithers site, and one Grand Prairie site (Figure 3.5). Observed heterozygosity (Ho) and heterozygosity within populations (Hs) had a very small range among all sites, with Ho ranging from 0.269–0.317 and Hs ranging from 0.278–0.302. Inbreeding coefficients (Gis) were spread between -0.065 to 0.061 with the population from Canmore showing the highest value and a population from the far north of Alberta having the lowest (Table 3.1).

OUTFLANK identified 25 SNPs under putative selection (Whitlock & Lotterhos, 2015). Of those 25 loci, nine had a positive blastx hit with a putative protein or gene product. Most of the hits related to calcium interactions, energy metabolism, and cellular detoxification (Table 3.3). RDA analysis showed environmental correlations that corresponded positively to the genotypes contained within specific genetic clusters. Southern sites were more positively linked to mean annual precipitation (MAP) and relatively humidity (RH) compared to northern sites, which were more positively connected to frost free period (FFP). Of the original 278 landscape variables, eight were retained in addition to latitude, longitude (see Appendix 4 for a full list). Sites in the Jasper area were somewhat positively linked to the percent of vegetation cover present relative to the other vectors (Figure 3.7). Outlier analysis of the RDA SNP plots identified 102 outlier loci in the range-wide dataset (Figure 3.8 and Table 3.4). Most outlier loci were linked with mean annual precipitation (MAP) and relative humidity (RH) with latitude contributing to a relatively low number of outliers (n = 11). Outlier analysis of a SNP subset from Jasper

National Park produced no significant outlier loci.

3.5. Discussion

Our study details the genetic structure of the newly expanded range of MPB in Western Canada across both provinces while identifying SNPs under selection throughout the remaining active portions of the outbreak using a bilateral approach pairing of both neutral and adaptive analysis techniques (Batista *et al.*, 2016; Yadav *et al.*, 2019). DAPC, PCA, and STRUCTURE analysis reveal the persistence of a strong north and south divide across the landscape, as reported in previous studies from both Canada and the USA (Samarasekera *et al.*, 2012; Janes *et al.*, 2014; Batista *et al.*, 2016). Based on our sampling of the landscape, this divide occurs north of Canmore, in close proximity to the location of the north-south delimitation identified by Samarasekera *et al*. (2012).

The northern population, as described by STRUCTURE, represents the greatest proportion of active infestations on the landscape in Alberta. However, our research also found that the Jasper-area population represents an area of admixture, contained in a third putative cluster identified by find.clusters in adegenet, DAPC, and PCA, a finding that was first suggested by previous work from Trevoy *et al.* (2018) and Shegelski *et al.* (2021). A possible explanation for the third cluster in our analysis is that beetles in this area may have originated from a degree of mixing between northern and southern populations within Jasper National Park, caused by movement of beetles from the west of BC migrating towards Alberta concurrently with spread northward from the southern clusters. This would align with our results from chapter 2, where we found that mixed groups of MPB (north and south) interbred and moved to admixed populations over time. Studies of the MPB

neo-Y haplogroup identify populations in British Columbia and Alberta as originating from the same haplogroup, suggesting that there are no genetic barriers to interbreeding in this area (Dowle *et al.*, 2017).

We found that some sites in the north contain individuals that assort with the southern populations (one Smithers site and one Grande Prairie site). In the case of Smithers, there is the possibility that the population originated from long-time local MPB resident populations with a relatively developed genetic signature (Aukema *et al.*, 2006), as opposed to the population being the sole result of immigration from the spreading northern outbreak. The Smithers site is located on the leeward of a mountain and may not have received wind-driven beetle inflights. However, in both cases, the sampling areas are proximal to highways with recreational sites that attract large amounts of tourists traveling through southern British Columbia north towards Alaska in the summer (Hardy & Gretzel, 2008). The transportation of camp wood is a very real concern for assisting beetle movement and may be a possible source of southern-assorting beetles in the north (Batista *et al.*, 2016). The reverse may also be true for the Canmore site, which has a more northerly signal despite its proximity to Banff. Canmore and Banff National Park likewise attract very large numbers of tourists from the northern parts of British Columbia and Alberta each year (Cheng, 1980; Draper, 2000).

In our study, we found nine outlier SNPs in or near gene coding regions with a positive blastx hit, with functions linking to putative calcium ion use, energy metabolism, cell signaling, and photo response. Our two calcium channel-related hits were an agrin-like protein which is involved in the development of neural synapses and neuromuscular junctions (Zong *et al.*, 2012) which may connect to dispersal ability and a mucolipin-3-like

protein (TRPML) which is a part of endosome function and immunity (Himmel & Cox, 2020). We also found a TRPL translocation defect protein which relates to how cells respond to light and may also involve Ca^{2+} conductance (Franchini *et al.*, 2014). This connects to previous research on photosensitivity response demonstrated at several life-stages in MPB (Wertman *et al.,* 2018). Ion transportation outlier SNPs were also previously identified from smaller scale studies of MPB SNPs in addition to functions linked to actin contraction, sterol association and RNA and DNA processing proteins (Bonnett *et al.*, 2012; Janes *et al.*, 2014; Batista *et al.*, 2016). Across the landscape, MPB has moved into regions with harsher, more challenging climates, meaning that beetles with faster response to stimuli, whether relating to photoperiod or other environmental factors, may lead to greater fitness and selection signature on the SNPs.

We found two energy pathway-related outliers: glutamate dehydrogenase, used in energy metabolism and activated by ADP (Bond & Sang, 1968) and a structural maintenance of chromosomes protein 4, an ATPase used in DNA repair and epigenetic silencing (Harvey *et al.*, 2002). In addition, a tetratricopeptide repeat protein (TPR) was noted, which is involved in a variety of processes including transcription regulation and the cell cycle (Kreppel & Hart, 1999). We also found a neurotrimin-like protein which may be related to learning and information exchange (Alleman *et al.*, 2019) and a TLR4 interactor with leucine rich repeats, also involved in cell signaling (Chaturvedi & Pierce, 2009).

Zinc carboxypeptidase, an enzyme involved in the resistance to evolved plant defensive protease inhibitors which can inhibit digestive enzymes (Bayé *et al.*, 2005), was also an outlier. Zinc carboxypeptidase was also found to be downregulated in new adult beetles preparing to overwinter as part of a collaborative study using beetles from Chapter

4 of this work (Penfold *et al.* in progress). These particular outlier SNPs are likely related to the harsher environment faced by MPB in its northern range. Particularly with the connection to epigenetic gene silencing outliers which may be expected as a result of an induced response to climatic factors (Sgrò *et al.*, 2016). As a phytophagous insect, MPB also benefits from an ability to prevent damage from the digestion of toxic terpenoid compounds and other constitutive defenses (Raffa *et al.*, 2017) so zinc carboxypeptidase may provide increased ability maintain digestion function. This would potentially increase beetle success and longevity, particularly in endemic conditions as non-tree killing strip attacks become more common and beetle colonization by mass attack decreases.

RDA analysis demonstrates that MPB genetic clusters on the landscape have differing genetic responses to landscape variables. Individual genotypes in southern sites are more strongly correlated with increasing mean annual precipitation (MAP) and relative humidity (RH) than all other sites. Sites in the south are more isolated and more likely to experience greater and consistent amounts of rainfall than parts of Alberta close to the Rockies (Harris & Brown, 1978; Meidinger & Pojar, 1991; Samarasekera *et al.*, 2012). Genotypes in the northern sites comparatively have the weakest correlation with both precipitation metrics. This may be due to the harsher, drier conditions in the north reducing the influence of moisture on beetle populations. A longer frost free period (FFP), however, has the strongest positive correlation with genotypes of northern populations, while being negatively associated all other clusters. While northern British Columbia and Alberta are known to experience comparatively much harsher, colder climate than the southern portions of both provinces (Nkemdirim, 1986), the northern populations experience a longer FFP than any of the mountainous areas sampled and may be experiencing selection

pressure on their genotypes towards the center of the expanded range as expected during range expansions (Vucetich & Waite, 2003; Klopfstein *et al.*, 2006).

Though we did not identify specific SNPs linked to cold tolerance, some of the ATP and mitochondrial function-linked SNPs may play a role in glycerol and cryoprotectant production. Elevation and percent land cover have the strongest positive correlation with the Jasper cluster and have a weaker influence on all other populations. Forests in Jasper National Park can vary greatly due to their elevational position within the park and the soil and site conditions overlaying the bedrock parent material. Jasper National Park also has different forest management practices than crown lands managed by forestry tenure holders, which may contribute to a higher degree of vegetation cover than other sites sampled (Rhemtulla *et al.*, 2002).

Over the entire dataset, MAP exerted the strongest correlation on outlier SNPs, with RH having a secondary influence. These two climatic factors, while interlinked, are not directly related to each other as RH can depend on a variety of factors including topography and other landscape metrics. It is well known that insects are responsive to various climatic and environmental effects (Logan & Amman, 1991; Sandrock *et al.*, 2011; Goodsman *et al.*, 2018; Bentz *et al.*, 2022), though the long- and short-term genetics effects are not well studied (Sgrò *et al.*, 2016; Fricke *et al.*, 2022) particularly in insect pests (Kirk *et al.*, 2013a). MAP has far-reaching effects on the health of host trees as water is a key determinant of plant growth and fitness. Drought and lower RH can deplete carbohydrate reserves and contribute to weakness within trees (Allen *et al.*, 2010; Erbilgin *et al.*, 2021), making them potentially more attractive hosts to MPB. Trees with adequate access to moisture will be more capable of mounting a successful defense against endemic

level beetle attacks, though epidemic level mass attacks are usually capable of overwhelming even healthy trees (Shore *et al.*, 2006).

Mean annual temperature (MAT) contributed to only one outlier SNP, which is in contrast to other studies of invertebrates where temperature has been a leading influence on outlier SNPs (Sandrock *et al.*, 2011; Xuereb *et al.*, 2018; Yadav *et al.*, 2019). While not a direct measure of temperature, FFP had a greater influence on MPB outlier SNPs, which may function somewhat analogously to mean annual temperature as both climatic conditions can influence MPB development time (Bentz & Powell, 2014). It is quite possible FFP will contribute to an upper limit of successful spread of MPB under both epidemic and endemic conditions where the summer periods in the far north of Alberta and British Columbia would be too short for MPB to successfully complete a full generation. This has already been partially demonstrated during the peak of the outbreak where beetles that dispersed into the Yukon in the mid-2000s were unable to establish and maintain populations there. However, as climate conditions continue to warm, this barrier to northern expansion is likely to increase in latitude.

Our study demonstrates that the vast majority of beetles in Alberta are the likely progeny of original northern source beetles from British Columbia. These in-flight-caused populations in Alberta did not display fine-scale structure due to their origin from the same source, relatively high gene flow, and recent colonization. This genetic homogeneity in the northern cluster on the landscape, indicated by very low Fst values, makes it likely that beetle populations on the whole will display similar behaviours across the newly established range. While one analysis (BIC-based clustering) suggests that the Jasper National Park populations represented a discernable genetic cluster, beetles from Jasper did

not make up a majority of the expanded range of the beetle in Alberta, though they have been detected ~80 km from the park in the town of Hinton (Shegelski *et al.*, 2021). It is possible that as time progresses, beetles within this admixed zone will differentiate into a uniquely discernible population, but at this time there is no evidence of an increase in selective processes within the Jasper National Park population.

We did not find evidence that indicates extremes of local adaptation at this time. While beetle populations were originally from the more contiguous landscape in central British Columbia, Alberta has a naturally more fragmented distribution of pine, with large areas of prairie that separate stands, and therefore, groups of beetles. This naturally patchy landscape may contribute to the establishment of a metapopulation dynamic for MPB within Alberta, with smaller scale outbreaks that face local expansion and resurgence over time (Hanski, 1998). Stand characteristics and more importantly, regional climatic factors, as demonstrated by our data, are likely to be the main drivers of genetic and behavioural changes.

In this study, we used landscape genomic techniques to explore the influence of environmental metrics on the genomic processes of MPB as it expands in novel habitat. We did this by taking advantage of differential filtering on the same SNP dataset to use both neutral and adaptive analysis methods to separate landscape-level structure and multi-locus selection. We found that most remaining viable MPB populations in western Canada are of the same or similar character, produced by the initial in-flight from northern BC. We did not find evidence that Jasper National Park acts as a major source of beetle spread throughout Alberta. Our study indicated that some climatic factors contributed to selection pressure within MPB populations, particularly those related to precipitation and frost free

period. It is likely that continued warming of the climate within western Canada, particularly Alberta, will contribute to the maintenance of successful endemic populations of MPB across the landscape and the need for vigilance and implementation of forest practices that maximize the health of pine to provide resistant, resilient hosts.

3.6. Acknowledgments

We are deeply grateful to the British Columbia Ministry of Environment and Climate Change Strategy (Park Use Permit, Authorization # 107171) and Alberta Agriculture and Forestry for the MPB survey data they provided. We would like to thank the Weyerhaeuser, Graymont, and Sunoco staff who provided access to various gated areas during sample collection. Research funding was partially provided by the Natural Science and Engineering Research Council of Canada (grant no. NET GP 434810-12) to the TRIA Network. Computational resources used in analysis of SNP data and job code troubleshooting was provided by Compute Canada and their support team (www.computecanada.ca).

Table 3.1: Site collection dates, with observed heterozygosity (Ho), heterozygosity within populations (Hs), and inbreeding coefficient (Gis).

Site Code	N	Latitude	Longitude	Year	Ho	Hs	Gis
xV	19	49.07629	-116.89520	2016	0.273	0.28	0.024
xW	18	49.24843	-117.95980	2016	0.281	0.286	0.019
xU	19	49.27647	-115.87400	2016	0.269	0.278	0.031
xT	18	50.84039	-116.63790	2016	0.285	0.286	0.002
xS	28	51.06070	-115.27760	2016	0.279	0.297	0.061
H	15	52.70070	-117.90750	2016	0.293	0.297	0.013
JC	17	52.70070	-117.90746	2017	0.301	0.3	-0.002
F	14	52.73098	-118.04700	2016	0.317	0.301	-0.052
G	20	52.81334	-118.05170	2016	0.29	0.297	0.021
JB	19	52.81334	-118.05170	2017	0.31	0.302	-0.026
A	14	52.84985	-118.57260	2016	0.291	0.296	0.015
JA	20	52.84985	-118.57265	2017	0.293	0.298	0.017
E	14	52.85660	-118.13200	2016	0.292	0.299	0.021
D	14	52.86663	-118.25450	2016	0.292	0.297	0.017
C	16	52.89184	-118.48460	2016	0.29	0.296	0.021
I	19	52.89498	-117.87590	2016	0.294	0.298	0.013
K	9	52.90983	-118.09890	2016	0.296	0.297	0.006
J	17	52.92076	-117.99660	2016	0.304	0.297	-0.022
JD	18	52.92076	-117.99663	2017	0.293	0.3	0.024
N	25	52.98470	-117.36470	2016	0.299	0.299	0.002
JE	39	53.12952	-117.77269	2017	0.306	0.301	-0.017
L	30	53.12952	-117.77270	2016	0.296	0.298	0.005
M	17	53.34372	-117.58260	2016	0.293	0.296	0.011
xA	13	53.39775	-115.85040	2016	0.282	0.281	-0.005
xR	20	53.49925	-117.33850	2016	0.301	0.289	-0.04
xC	18	53.64446	-117.03800	2016	0.282	0.284	0.007
xB	19	53.83089	-116.56860	2016	0.296	0.286	-0.033
Z	16	54.10527	-116.07160	2016	0.281	0.285	0.015
xQ	18	54.29064	-118.49020	2016	0.285	0.284	-0.001
Y	19	54.37585	-115.69060	2016	0.279	0.28	0.006
xN	20	54.50274	-117.52660	2016	0.282	0.281	-0.001
xP	19	54.50808	-118.69190	2016	0.282	0.282	0.001
xG	20	54.56863	-119.43100	2016	0.278	0.282	0.015
xH	17	54.59904	-119.27610	2016	0.296	0.284	-0.043
xF	20	54.60057	-119.06580	2016	0.281	0.283	0.009

xO	17	54.60632	-118.22430	2016	0.281	0.285	0.015
xJ	18	54.61417	-119.91350	2016	0.283	0.283	0.001
xL	20	54.62643	-119.82640	2016	0.299	0.286	-0.044
xK	19	54.64256	-119.87110	2016	0.283	0.285	0.008
xE	17	54.65595	-119.00670	2016	0.28	0.282	0.007
xI	17	54.66468	-119.99570	2016	0.278	0.282	0.014
xM	27	54.66766	-119.80470	2016	0.287	0.299	0.039
LLB	24	54.68386	-112.03320	2018	0.281	0.285	0.014
xY	17	54.72940	-127.21700	2016	0.28	0.297	0.056
O	14	54.80227	-120.06490	2016	0.296	0.287	-0.031
P	17	54.85801	-120.13560	2016	0.282	0.282	-0.003
Q	13	54.87391	-120.24830	2016	0.281	0.283	0.007
xX	17	54.94973	-127.38180	2016	0.299	0.284	-0.054
S	13	55.61294	-114.32640	2016	0.294	0.283	-0.037
T	20	55.63810	-113.93650	2016	0.28	0.28	0.001
U	20	56.58712	-118.43490	2016	0.275	0.281	0.021
R	9	56.64878	-119.02720	2016	0.28	0.28	0.002
V	33	57.31179	-117.54710	2016	0.282	0.284	0.008
X	14	58.23651	-118.92790	2016	0.305	0.286	-0.065
W	14	58.54397	-119.44140	2016	0.284	0.283	-0.003

Table 3.2: Analysis of Molecular Variance on best clustering according to BIC (n = 1018 individuals at 3618 loci.

Source of Variation	Nested in	SSD	d.f.	MS	%var	Rho	Rho-value
Within Populations	--	1107067	963	1149.603	0.933	st	0.067
Among Populations	Clusters	72196.67	51	1415.621	0.012	sc	0.012
Among Clusters	--	48199.69	3	16066.56	0.055	ct	0.055
Total (SST)	--	1227464	1017	1206.946	--	--	--

Table 3.3: Identified outlier SNPs from OUTFLANK with predicted proteins from the MPB annotated transcriptome and gene ontology. Only the 9 loci with a positive BLAST hit are shown.

Locus Name	Gene Description	Uniprot GO Term
240556:31:+	agrin-like	calcium ion binding
168565:4:+	glutamate dehydrogenase, mitochondrial-like	oxidoreductase activity
111602:27:-	mucolipin-3-like	calcium channel activity
218165:43:-	neurotrimin-like isoform X2	cell adhesion
164187:9:+	structural maintenance of chromosomes protein 4	ATP binding
36757:21:-	tetratricopeptide repeat protein 17	actin filament polymerization
131840:55:-	TLR4 interactor with leucine rich repeats	lipopolysaccharide binding
148082:61:-	TRPL translocation defect protein 14	cellular response to light stimulus
82249:8:-	zinc carboxypeptidase	zinc ion binding

Table 3.4: RDA Environmental predictor variables with numbers of associated SNPs for MPB SNP dataset (north, south, and Jasper) (n=102 outlier SNPs)

Explanatory Variable	**Number of Associated Loci**
MAP (Mean Annual Precipitation)	50
RH (Relative Humidity)	34
Lat (Latitude)	11
FFP (Frost Free Period)	6
MAT (Mean Annual Temperature)	1

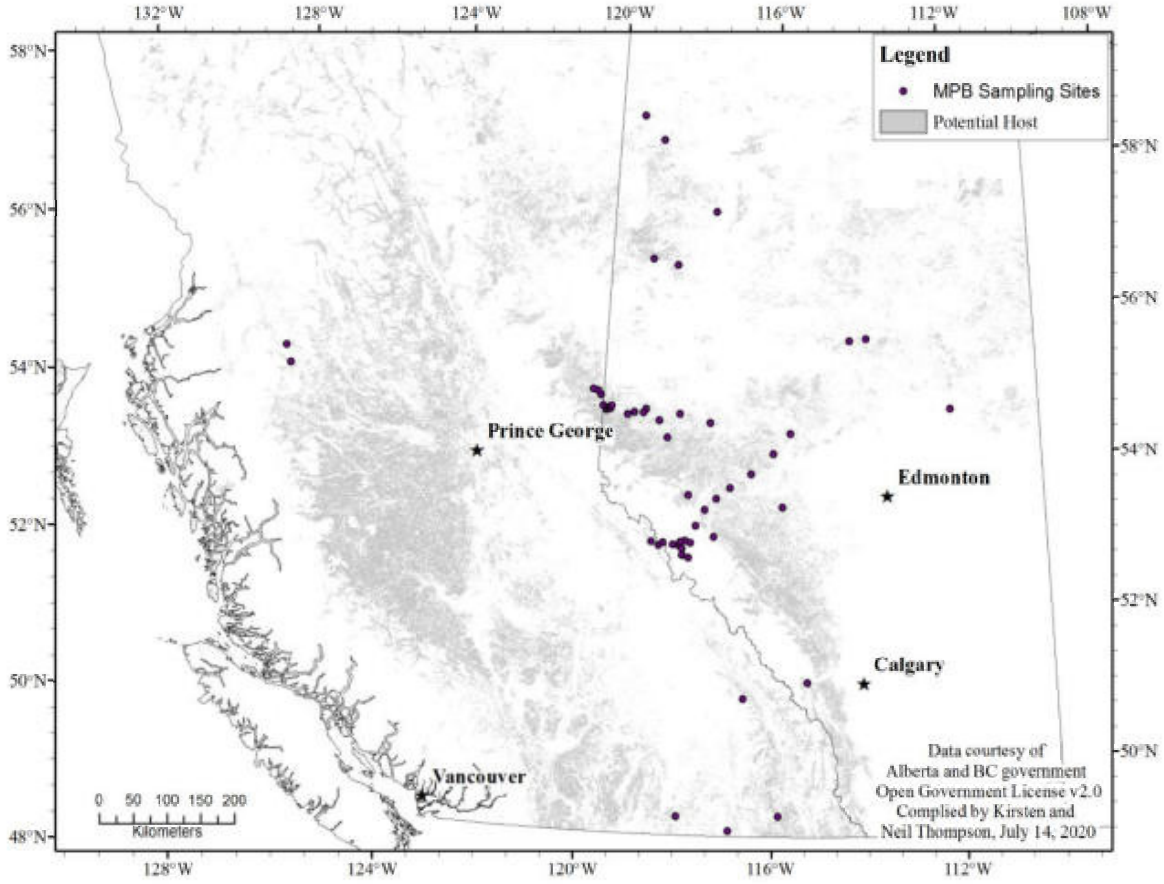

Figure 3.1: Location of MPB study sites in British Columbia and Alberta (n=55). Potential host pine is shaded in grey. Distances between outermost sites are approximately 900 km between east and west sites and 1000 km between north and south sites.

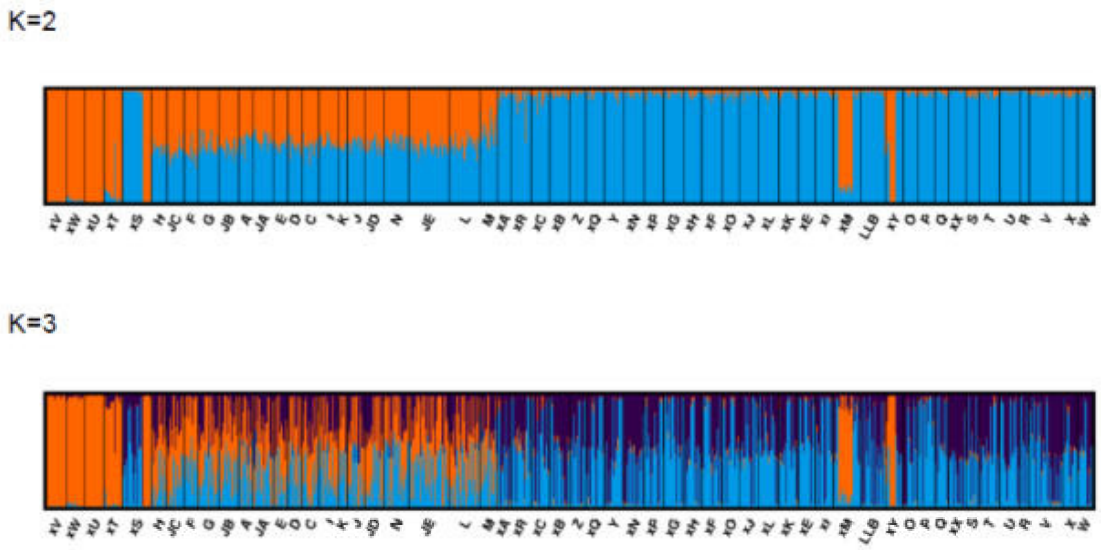

Figure 3.2: STRUCTURE plots with cluster membership per individual (n = 1018 individuals at 3612 loci) as averaged by CLUMPAK. Sites are arranged by increasing latitude and separated by black bars (n=55). Each cluster (K) is denoted by a different colour. The probability of cluster membership is shown by the proportion of each colour with the vertical bar. For K =2, orange sites assort to the southern cluster, while blue denotes the northern cluster.

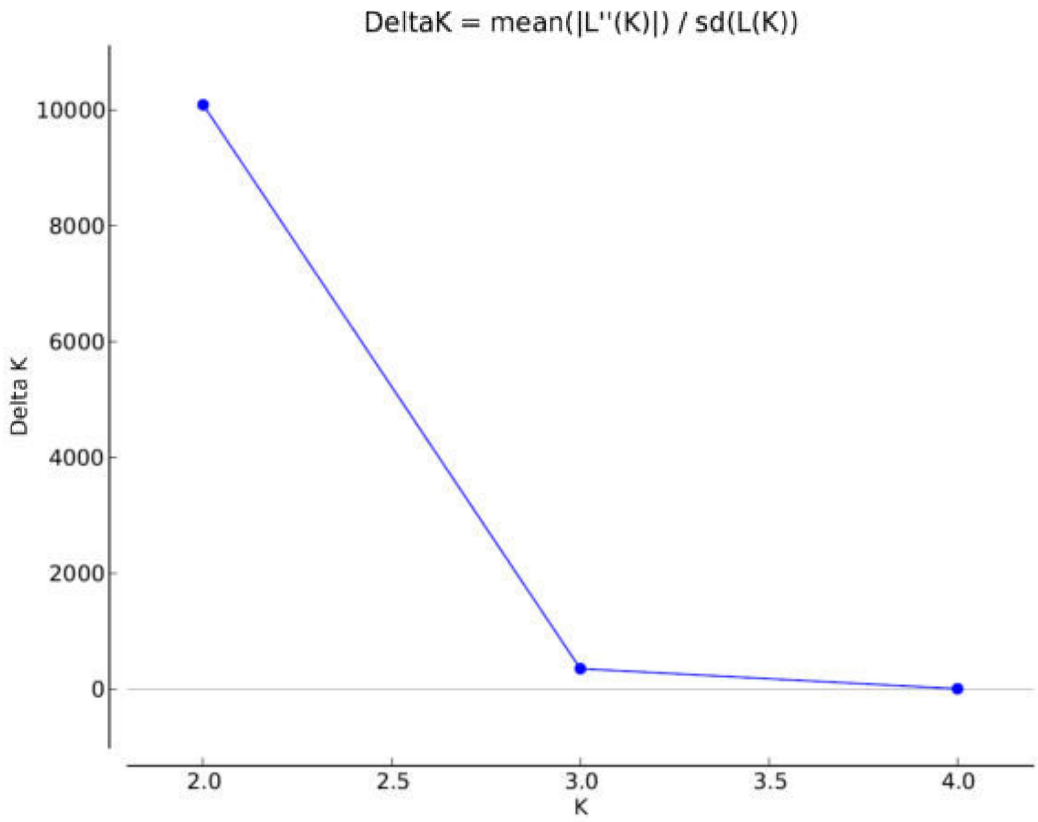

Figure 3.3: ΔK method plot, calculated in CLUMPAK using the Evanno method. Optimal K for this dataset is 2, shown by the elevated datapoint at K=2.

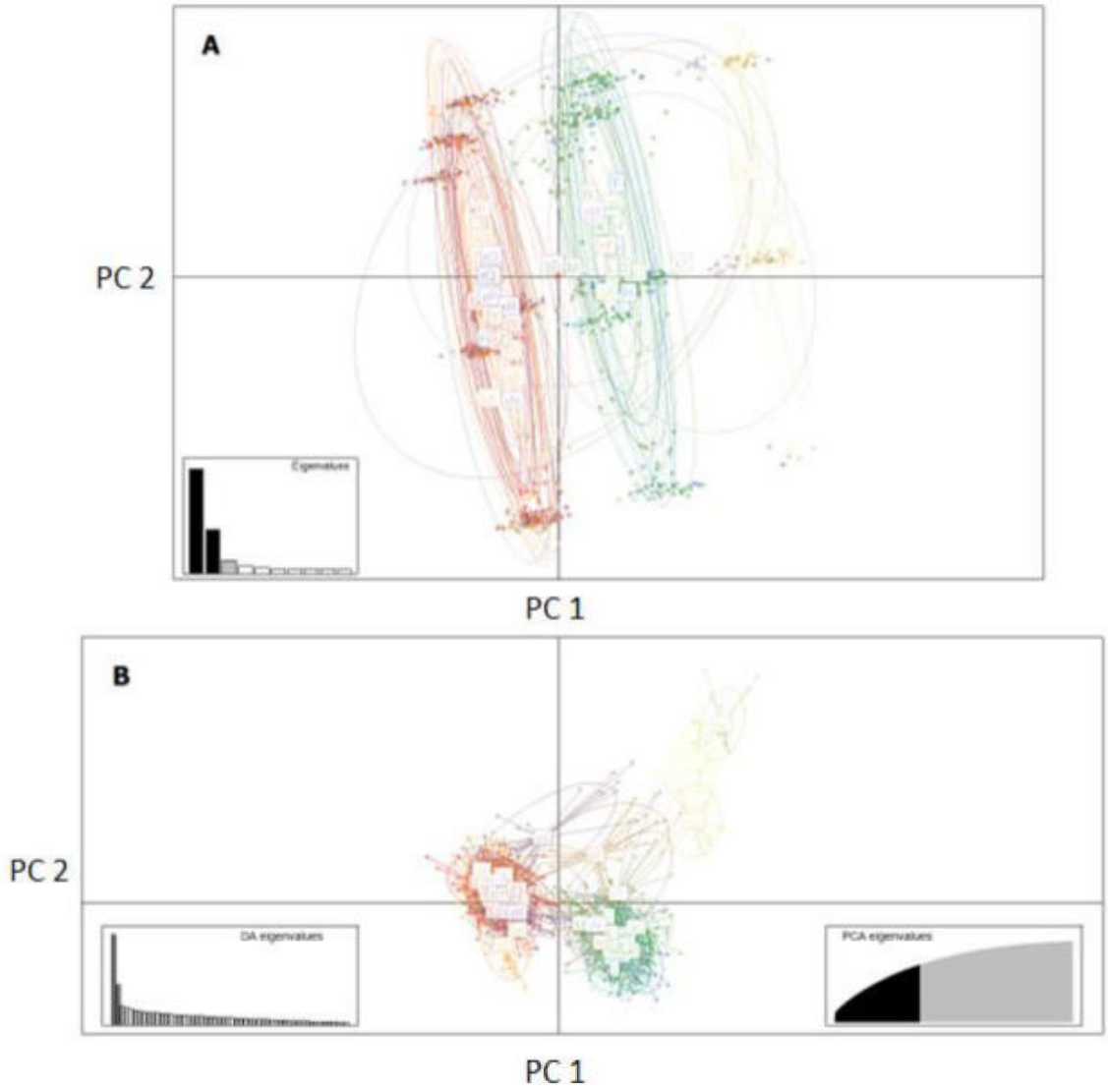

Figure 3.4: A)PCA of all MPB sites(n=55) with PC 1 and PC 2 and eigenvalues B)DAPC scatterplot of MPB SNP genotypes displaying principal components 1 and 2 of all sites (n=1018 individuals, and n=3612 loci). Orange tone denotes sites in the northern cluster, Green tone denotes sites in the Jasper cluster, and Yellow tone denotes sites in the southern cluster.

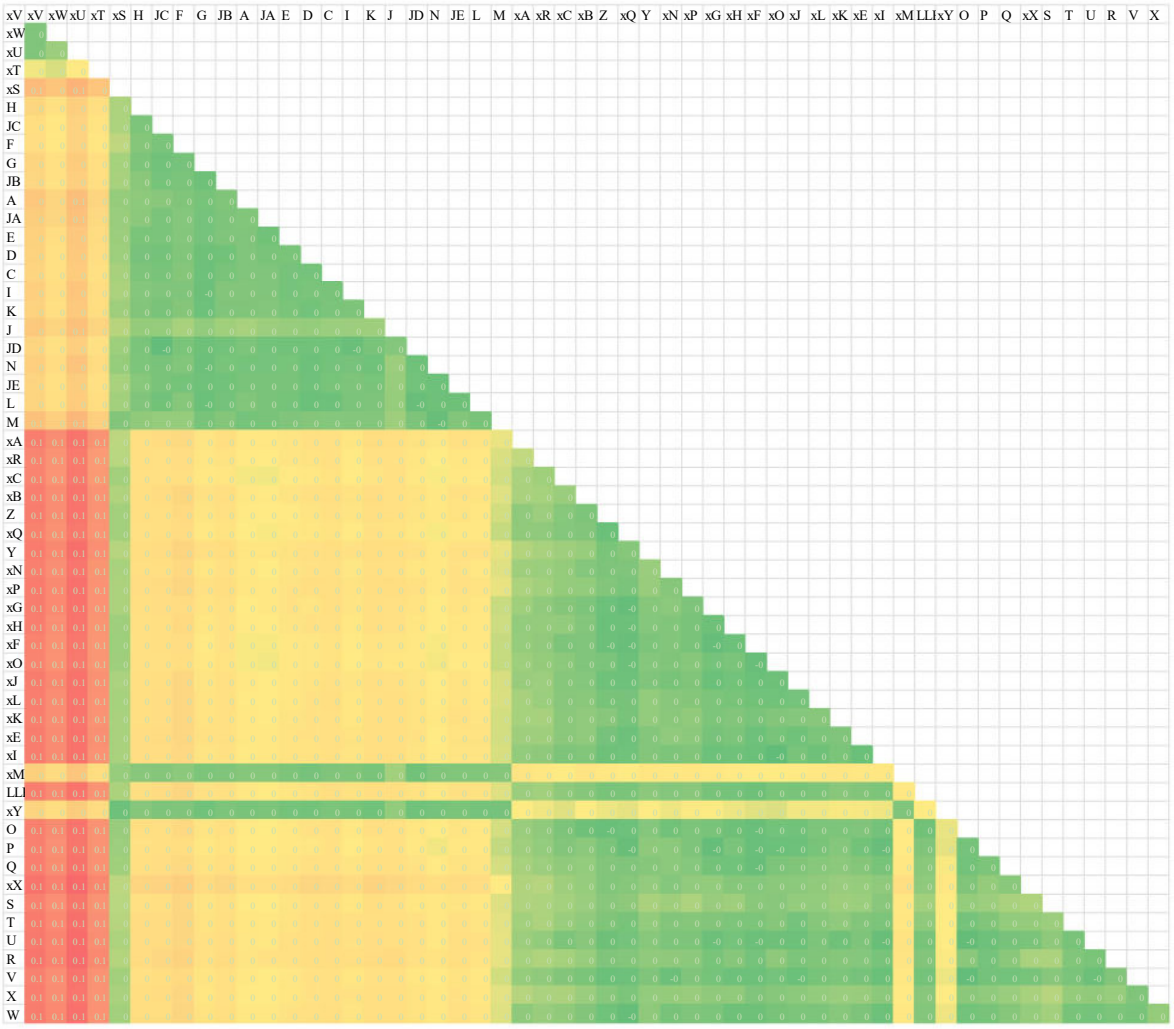

Figure 3.5: Heatmap of pairwise Fst values between MPB sites. Red indicates the highest levels of Fst while green indicates the lowest.

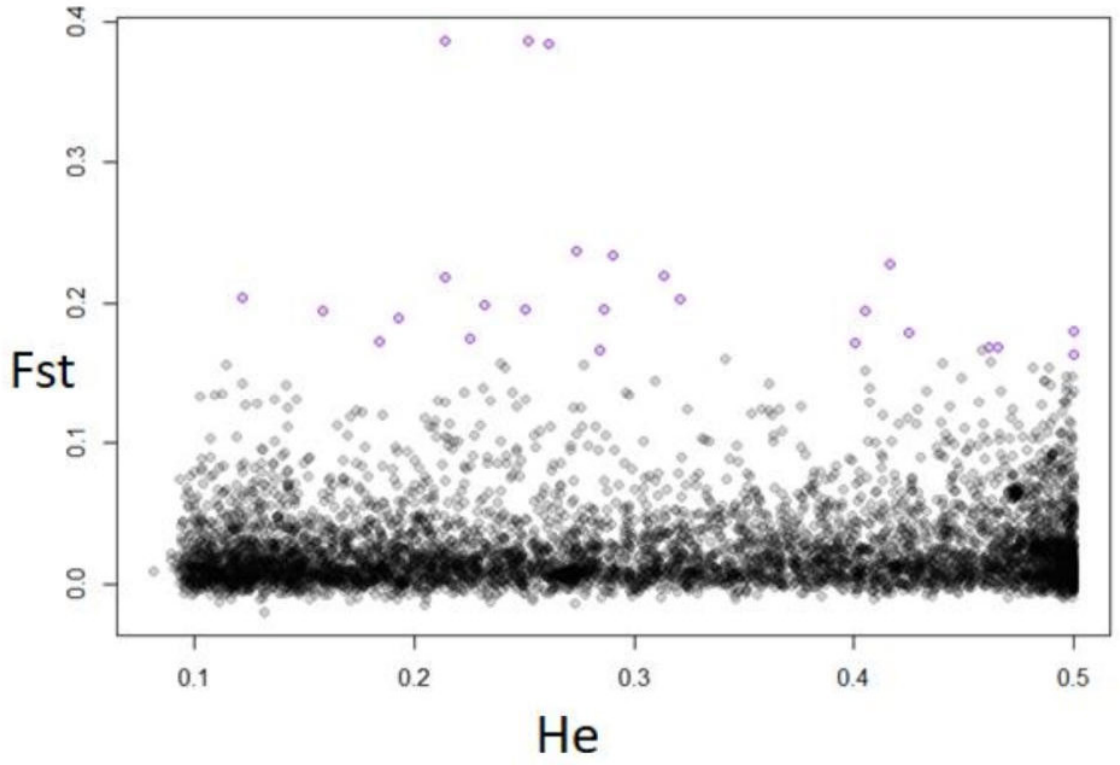

Figure 3.6: OUTFLANK MPB SNP plot showing Fst and He values. Outlier loci are marked in purple (n=25).

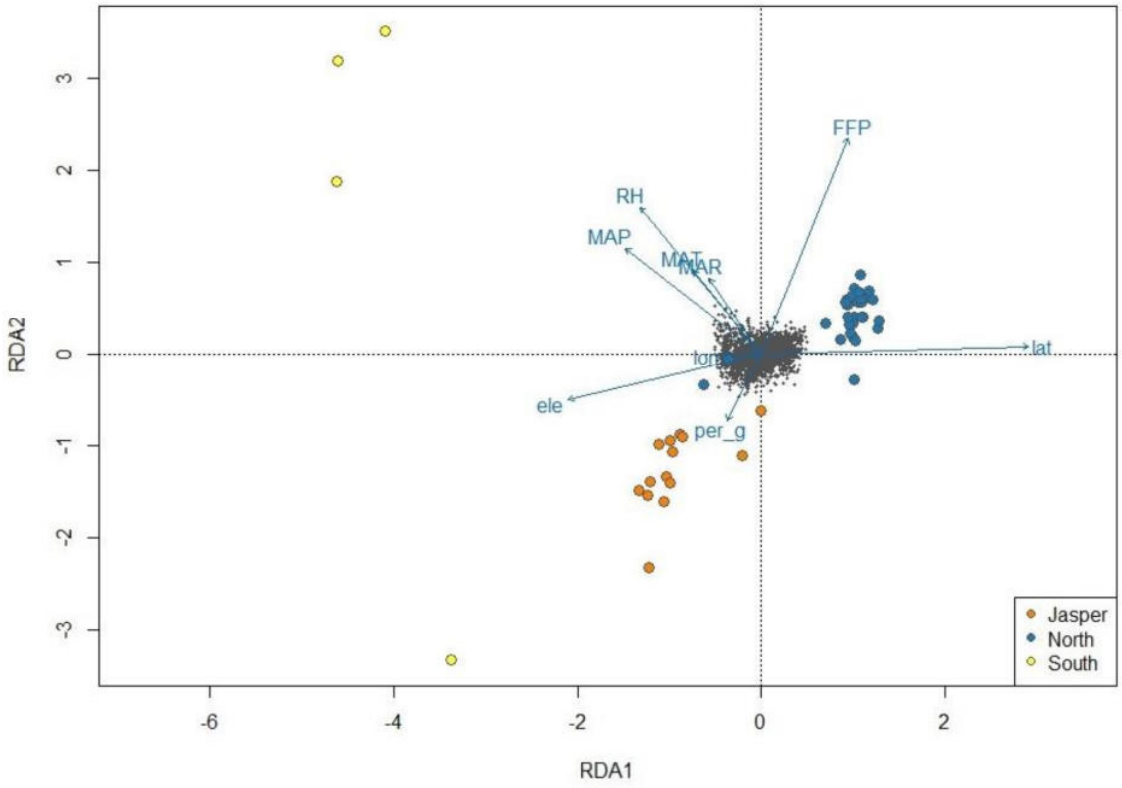

Figure 3.7: RDA triplot showing MPB SNPs in grey, and MPB sites (n=50) as filled circles. Cluster membership wass determined by find.clusters in adegenet. Blue vectors are environmental predictors. Arrangement in ordination space shows relationships with axes that are linearized mixtures of predictors with symmetrical scaling.

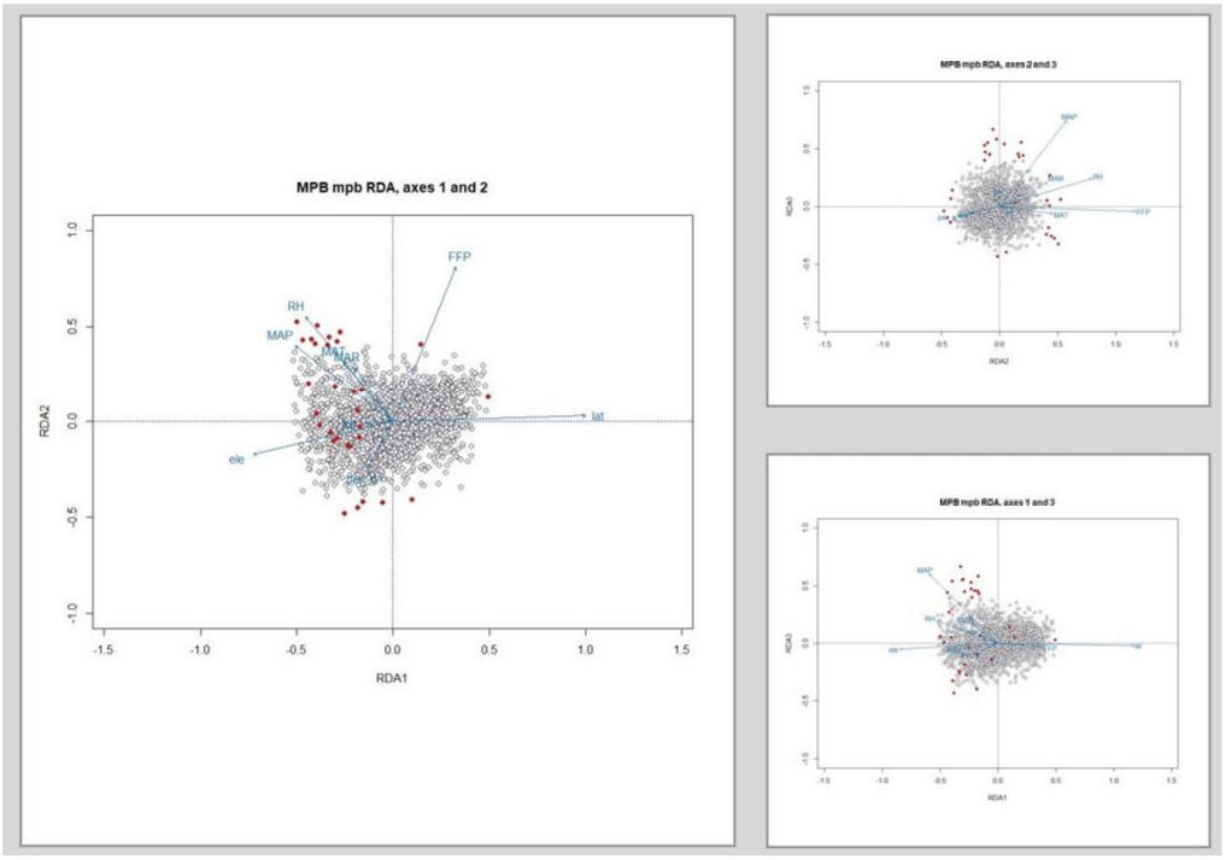

Figure 3.8: RDA triplot showing neutral MPB SNPs in grey and outlier SNPs in red (n=3612 loci total, 102 as outliers) for the range-wide dataset. Blue vectors are environmental predictors. Proximity to a vector indicates positive correlation between the predictor and the SNP.

4. Chapter 4: Autumn Shifts in Cold Tolerance Metabolites in Overwintering Adult Mountain Pine Beetles (Published)

Kirsten M. Thompson, Dezene P. W. Huber, Brent W. Murray

This chapter of my dissertation is based in full on the previously published article listed below: Thompson, K.M., Huber, D.P.W., and Murray, B.W. 2020. Autumn shifts in cold tolerance metabolites in overwintering adult mountain pine beetles. PLoS One 15: e0227203. Public Library of Science. doi:10.1371/journal.pone.0227203.

4.1. Abstract

The mountain pine beetle, *Dendroctonus ponderosae* (Coleoptera: Curculionidae) is a major forest pest of pines in western North America. Beetles typically undergo a one-year life cycle with larval cold hardening in preparation for overwintering. Two-year life cycle beetles have been observed but not closely studied. This study tracks cold hardening and preparation for overwintering by adult mountain pine beetles in their natal galleries. Adults were collected *in situ* between September and December 2016 for a total of nine time points during 91 days. Concentrations of 41 metabolites in these pooled samples were assessed using quantitative nuclear magnetic resonance (NMR). Levels of glycerol and proline increased significantly with lowering temperature during the autumn. Newly eclosed mountain pine beetles appear to prepare for winter by generating the same cold-tolerance compounds found in other insect larvae including mountain pine beetle, but high on-site mortality suggested that two-year life cycle adults have a less efficacious acclimation process. This is the first documentation of cold acclimation metabolite production in overwintering new adult beetles and is evidence of physiological plasticity that would allow evolution by natural selection of alternate life cycles (shortened or lengthened) under a changing climate or during expansion into new geoclimatic areas.

4.2. Introduction

The mountain pine beetle, *Dendroctonus ponderosae* (Coleoptera: Curculionidae), is an irruptive forest insect native to western North America (Bracewell *et al.*, 2017). In the past 20 years, mountain pine beetle outbreaks of unusual size killed much of the mature pine in British Columbia (Hrinkevich & Lewis, 2011) and expanded beyond their historical range over the Rocky Mountains to Alberta and into the north of British Columbia (Cullingham *et al.*, 2011; Janes *et al.*, 2014). Cold winters – greater than two weeks at –40°C – are thought to have limited the scope of previous outbreaks by killing off mountain pine beetle brood by freezing (Safranyik & Linton, 1998; Safranyik & Carroll, 2006). Climate change has led to a warming trend in the past 30 years, reducing the length and frequency of reaching and sustaining this temperature threshold (Carroll *et al.*, 2006). This trend is particularly the case in the early- and late-winter season when overwintering insects would normally be the most vulnerable due to a lack of overwintering metabolites, therefore increasing mountain pine beetle winter survival rates (Bentz *et al.*, 2010; Goodsman *et al.*, 2018).

Most mountain pine beetles have a univoltine lifecycle, meaning they mature over the course of a single year (Bentz & Powell, 2014; Six & Bracewell, 2015). Life cycle development is dictated by climatic factors; adults lay eggs in the late summer to early fall, allowing for partial larval development prior to freezing winter temperatures. Mountain pine beetles are freeze-intolerant and will experience mortality if water within their soft tissue crystalizes (Bentz & Mullins, 1999). For this reason, mountain pine beetles generate cryoprotectants in the autumn, especially glycerol, to reduce their super cooling point and protect against the formation of ice crystals (Régnière & Bentz, 2007; Robert *et al.*, 2016;

Bleiker *et al.*, 2017; Fraser *et al.*, 2017). Early instar larvae (instars 1 and 2) are marginally less cold tolerant than late-instars (instars 3 and 4), with the first three instars generating similar proportions of glycerol in relation to bodyweight, and the final instar producing slightly more (Logan *et al.*, 1995; Safranyik & Carroll, 2006). Larvae also void their guts in preparation for cooler temperatures in order to reduce the number of internal ice nucleation surfaces (Keeling *et al.*, 2013). Pupation and maturation follow in the spring and beetles fly in the summer to find new hosts and repeat the life cycle.

Mountain pine beetle phenology is known to be most responsive to temperature changes, with hormonal regulation and photo period playing little part in cold acclimation (Régnière & Bentz, 2007). Photoperiod, historically thought to have little effect on cold hardening, may be involved to some degree as new evidence suggests that mountain pine beetle larvae respond negatively to light, even when located in a sub-cortical environment (Wertman *et al.*, 2018). Reliance on temperature occasionally results in phenological delays causing an extension of the life cycle beyond one year. Such larvae overwinter, pupate later in the summer, and overwinter a second time as new adults (Amman, 1973). Larval cold hardening and cryoprotectant generation in insects is mainly understood (Storey & Storey, 1983), with mountain pine beetle-specific studies of cryoprotectant production (Régnière & Bentz, 2007), RNA transcript generation (Robert *et al.*, 2016; Fraser *et al.*, 2017), and proteomic data (Bonnett *et al.*, 2012) suggesting some of the mechanisms and pathways of larvae winter survival. While overwintering new adults are occasionally recorded in the literature (DeLeon *et al.*, 1934; Amman, 1973), and have been observed as an early-May flight during the height of the outbreak in central BC (DPWH, personal communication), they represent a less common life strategy and the precise

mechanisms of overwintering are unknown. In this study we recorded the production of metabolites in newly eclosed adult mountain pines beetle *in situ* from late-fall to the coldest day of the winter in order to quantify their cold acclimation process.

4.3. Methods

4.3.1. New Adult Beetle Collection and Climate Metrics

Beetle collection began in late-summer (17 Sept 2016) at the Lucerne Campground (Robson Provincial Park, British Columbia, 52°50'59.48"N, 118°34'21.84"W, 1125m (Fig 1)). The infested stand consisted mostly of lodgepole pine (*Pinus contorta)* with evidence of recent mountain pine beetle attack dating from several previous years, including the most recent summer. Sampling continued weekly (25 Sept, 2 Oct, 11 Oct) and then bi-weekly (23 Oct, 6 Nov, 20 Nov, 3 Dec, 16 Dec) for nine total collection days during a 91-day period spanning almost the entire autumn of 2016. Brood galleries were exposed using a draw knife to remove bark from the tree. A minimum of 40 new adult beetles were collected from separate galleries in five randomly selected trees during each collection event (eight beetles per tree). New adults were identified based on their proximity to larval galleries and their relative size. For the first eight sampling days, beetles were confirmed to be living (by observing movement) prior to collection. Temperatures below –30°C on the final day of sampling precluded this step as beetles were too cold for movement. Ice crystals were observed within sampled galleries on the final sampling day. Beetles were placed in either 2 mL or 1.5 mL dry snap-capped microcentrifuge tubes and flash frozen in the field using liquid nitrogen and then transported to the lab and stored at –80°C until metabolite processing. Three HOBO U23 Pro v2 data loggers (HOBOware, Onset Computer Corporation) were placed at breast height on well separated trees throughout the

field site to track temperature. Climate loggers were moved from their original positions after the eighth sampling day and relocated to other trees near to their original positions within the sampling area due to sanitation logging removing the original trees used for placement.

4.3.2. Metabolite Extraction

Mountain pine beetles were quickly thawed on ice. Approximately 1g of beetle tissue was transferred to a mortar and pestle with 3 mL of chloroform:methanol (1:2 v/v). The beetles were then ground for 3 min and the extract was transferred to a 4 dram vial. The mortar was rinsed with 2 mL of a chloroform:methanol (1:2 v/v) mixture for the complete recovery of the extract. The entire extract was filtered using vacuum filtration. The residue was transferred into a 15 mL sterile screw-capped plastic centrifuge tube and 5 mL chloroform:methanol (1:2 v/v) mixture was added to the residue and was shaken at 250 rpm for 30 min at ambient temperature on a shaker. This extract was filtered again using vacuum filtration, combined with the first filtrate in a 13.5 mL Teflon lined screw-capped glass vial and the combined filtrate was transferred to a 50 mL sterile screw-capped plastic centrifuge tube. To this filtrate one quarter of the total volume of the filtrate 0.88% KCl was added. The tube was vortexed for 1.5 min and placed aside for 10 min for phase separation of an upper aqueous layer and lower organic layer. The tube was then centrifuged for 30 minutes at 3000 rpm. The upper aqueous layer (water-soluble metabolites) was transferred into a 15 mL sterile screw-capped plastic centrifuge tube and 2.5 mL HPLC water was added to the water-soluble metabolite extract and flash frozen in liquid nitrogen. This sterile screw-capped plastic centrifuge tube was lyophilized with frozen water-soluble metabolites for 24 h and the resultant freeze-dried powder of was

divided into 15 mg aliquots for NMR analysis.

A single 15 mg aliquot of the lyophilized water-soluble extract from pine beetles was taken in 1.5 mL snap-capped microcentrifuge tube. To this powder, 570 µL of water was added. The sample was sonicated for 15 min in a bath sonicator. To this sample, 60 µL of reconstitution buffer (585 mM phosphate buffer with 11.67 mM DSS) and 70 µL of D2O were added. The solution was vortexed for 1 min and centrifuged at 10,000 rpm for 15 min at ambient temperature. The clear supernatant was transferred into an NMR tube for NMR analysis.

All 1H-NMR spectra were collected on a 700 MHz Avance III (Bruker) spectrometer equipped with a 5 mm HCN Z-gradient pulsed-field gradient (PFG) cryoprobe. 1H-NMR spectra were acquired at 25°C using the first transient of the NOESY pre-saturation pulse sequence (noesy1dpr), chosen for its high degree of quantitative accuracy. All FID's (free induction decays) were zero-filled to 250 K data points. The singlet produced by the DSS methyl groups was used as an internal standard for chemical shift referencing (set to 0 ppm) and for quantification all 1H-NMR spectra were processed and analyzed using the online Bayesil software package. Bayesil allows for qualitative and quantitative analysis of an NMR spectrum by automatically and semi-automatically fitting spectral signatures from an internal database to the spectrum. Specifically, the spectral fitting for metabolites was completed using the standard serum metabolite library. Typically, all visible peaks were assigned. Most of the visible peaks are annotated with a compound name. This fitting procedure provides absolute concentration accuracy of 90% or better. Each spectrum was further inspected by an NMR spectroscopist to minimize compound misidentification and mis-quantification.

4.3.3. Exploration of Metabolic Data

A one-way ANOVA was used to determine significance between measured metabolites, with a post-hoc Tukey's honestly significant different test used for pair-wise comparison of metabolite mean concentration between timepoints [adjusted p-value (FDR) = 0.05] following statistically significant ($p < 0.05$) ANOVA results. Both measures were conducted using MetaboAnalystR (Chong & Xia, 2018). Concentration values and temperature measurements were visualized in Microsoft Excel (2016). Correlation of glycerol, proline, and trehalose to temperature values was performed in Rstudio (R version 3.4.4) using Pearson's product-moment correlation and negatively transformed temperature values.

4.4. Results

Temperature at the study site initially decreased towards freezing but warmed and remained at or above freezing during the day during October and November. Temperatures dropped rapidly during early December and remained below –20 ℃ for approximately two weeks, reaching the coldest on-site temperature of the winter (–30.1℃) on the final sampling day, 16 December 2016 (Fig 2). A spring collection was attempted when the site was again accessible following snow melt on 26 May 2017, but no living new adults could be found in a search of the area, including thorough investigation of trees previously sampled (number of trees checked on site > 50).

We detected 41 metabolites in overwintering adults (full list provided in S1 file). A one-way ANOVA with a post-hoc Tukey's HSD test for pairwise comparison showed 27 of these metabolites differed significantly at one or more of the time points taken during the study (Fig 3, but see also S2 file). Of these 27 significant measures, three metabolites –

glycerol, trehalose, and proline – became highly elevated and are likely biologically significant in addition to being statistically significant (Fig 4). Glycerol concentrations were highest in beetles at the end of the sampling period, reaching 468.91 µg/mg of body weight (SE ± 49.74). Trehalose levels increased into the month of October but decreased in December, reaching a peak of 180.16 µg/mg (SE ± 13.67) on 6 Nov. Proline levels increased until the middle of the sampling period and remained stable for the final four sampling days, reaching a final concentration of 46.02 µg/mg (SE ± 2.93) on the final sampling day in mid-December. Glycerol ($p = 0.002$) and proline ($p= 0.016$) both had a significant strong positive correlation to decreasing temperatures while trehalose had a non-significant correlation (Fig 5).

4.5. Discussion

We found that new adult mountain pine beetles form their own metabolic antifreeze compounds in response to autumn temperature cues. The most responsive metabolites to increasing cold in adults were glycerol, trehalose, and proline, for all of which there is also previous transcriptomic and proteomic evidence of biosynthesis during larval cold hardening (Bonnett *et al.*, 2012; Robert *et al.*, 2016; Fraser *et al.*, 2017). Mountain pine beetle larvae typically survive winters if ambient temperatures do not fall below approximately –40°C, though larvae insulated below the snow line may survive these cooler conditions (Safranyik & Carroll, 2006; Bleiker *et al.*, 2017). Studies that have referenced the presence of newly eclosed adult mountain pine beetles and that tracked emergence rates have indicated that new adults have higher winter mortality rates compared to their larval counterparts, but did not investigate potential mechanisms for this reduced success (DeLeon *et al.*, 1934; Amman, 1973). We observed, but did not quantify,

high mortality at our site during our site which we postulate is linked to several factors including site temperature regime, differing physiology compared to larvae, and below-bark conditions.

Glycerol is a known cryoprotectant in many insects, including other *Dendroctonus* spp., and has been documented in mountain pine beetle larvae (Miller & Werner, 1987; Bentz & Mullins, 1999; Régnière & Bentz, 2007; Wang *et al.*, 2016b). It is a relatively inert compound that can be maintained at high concentration without interfering with other cellular processes or enzymatic reactions (Leather *et al.*, 1993). It is also nontoxic, so insects experience few fitness trade-offs when generating this compound, and it can be converted into glycogen when temperatures begin to warm (Leather *et al.*, 1993; Storey & Storey, 2012; Fraser *et al.*, 2017). Our previous studies have shown that larvae increase their capacity to generate glycerol in correlation to temperature similar to what we have observed here with adults (Bonnett *et al.*, 2012; Robert *et al.*, 2016; Fraser *et al.*, 2017) (Fig 4). Recent cold acclimation metabolite work has shown that mountain pine beetle larvae produce a concentration of glycerol an order of magnitude more per mg of tissue compared to the new adults profiled in our study (Batista pers. comm., visiting scholar, UNBC, batista@unbc.ca). This lower concentration of glycerol may have reduced the new adult beetles' ability to supercool and thereby increased mortality rates.

Trehalose is a major sugar constituent of insect haemolymph that acts as a mobile energy source for cellular respiration (Leather *et al.*, 1993; Thompson, 2003). We observed increasing levels of trehalose in the mid- to late-fall, but no continual increase as temperatures grew cooler (Fig 4). In European populations of *Ips typographus* (Coleoptera: Curculionidae), trehalose undergoes a similar increase through October in response to cold

temperatures (Koštál *et al.*, 2011). Trehalose can also act as a cryoprotectant, stabilizing proteins at cold temperatures and keeping cellular membranes intact (Feng *et al.*, 2016). Increasing the durability of cellular membranes would reduce cellular damage in the event of changing osmotic pressure or ice crystal formation. Trehalose has also be linked to changing dietary cues for insects (Thompson, 2003), meaning changing trehalose levels in the blood might lead to a reduction or cessation of feeding behaviour as temperatures cool, helping with voiding of the gut prior to onset of winter.

Proline is known to be a cryoprotectant in both plant and yeast cells (Pemberton *et al.*, 2012) but is not well-documented as a cryoprotectant in insects. The increase of proline levels in response to temperature within the new adults suggests a connection to cold acclimation in mountain pine beetles (Fig 4). In the red flat bark beetle, *Cujucus clavipes* (Coleoptera: Cucujidae), proline and alanine are thought to work together with trehalose to slow the freezing process (Carrasco *et al.*, 2012). Alanine was detected in the new adults, though it did not correlate with temperature over the duration of the study (S1 file). It is possible that new mountain pine beetle adults use proline to decrease their supercooling point, though further study would be needed to confirm this possibility. Proline is also metabolized along with carbohydrates during flight (Gäde & Auerswald, 2002; Teulier *et al.*, 2016). Proline generation may serve a dual purpose where beetles use the amino acid as a cryoprotectant in the winter and then metabolize remaining proline as a flight fuel for dispersal.

On-site phenology drivers and additional metabolic demands on new adults are likely contributing factors to the level of mortality observed in the field. While field temperatures initially dropped at a steady rate from September to early October, they

rewarmed between mid-October and November (Fig 3). The extended period of warmer weather may have confused the cues that normally trigger the beetles to produce cryoprotectants. New adults feed on fungal associates after eclosing but prior to emergence from under the bark (Safranyik & Carroll, 2006); based on our observations of bark and phloem conditions, adequate food resources were available to beetles on the study site. Adults beetles do, however, have different metabolic demands compared to larvae. They must develop fat reserves to support flight and also maintain gonadal tissue for reproduction (Gäde & Auerswald, 2002; Six & Bracewell, 2015). It should be noted that adult females partition their metabolic resources to at least some extent – for instance, they do not generate vitellogenin until they come into contact with a host tree following dispersal flight (Pitt *et al.*, 2014). Larvae have not yet developed these tissues beyond imaginal disks, and may thus have more resources to allocate to cold hardiness.

As ice crystals were observed within galleries on the coldest sampling day, this may have been a contributing factor to the increased mortality observed on the sampling site. Direct contact of ice crystals on the surface of an insect's exoskeleton creates a surface where point nucleation of ice can occur (Elnitsky *et al.*, 2008; Bleiker *et al.*, 2017). New adults are melanized with hardened carapaces and have more surface area compared to larva. Having undergone pupation, new adults also have thin legs that are liable to freeze faster due to their exposure. It is probable that beetles with elevated levels of cryoprotectants still experienced internal ice crystal formation due to the external ice contact. In addition, it is unknown if new adults are capable of voiding their guts as larvae do in preparation for freezing temperatures (Keeling *et al.*, 2013). If they are not capable, new adults would have more internal surfaces for ice crystal formation due to retained

food, likely making them yet more susceptible to freezing.

We found overwintering, newly eclosed mountain pine beetle adults produce three known antifreeze metabolites. Previous research suggests that these are the same three major metabolites produced by larvae during cold acclimation (Bonnett *et al.*, 2012; Robert *et al.*, 2016; Fraser *et al.*, 2017), but it is likely newly eclosed adults produce less of each cryoprotectant. This is the first time that a metabolic mechanism for new adult survival has been documented. While the new adults in our study experienced high mortality, larvae at nearby sites in Jasper National Park were found to have survived the winter of 2016-2017. New adult mountain pine beetles have been most commonly described at high elevation, due to a comparatively late start to spring and early start to cooling autumn temperatures (Amman, 1973; Bentz *et al.*, 2016). In areas where winters begin earlier in the year, thus reducing the cold acclimation period, and have temperatures reaching below –30 ℃ for several days, it is likely that new adult beetles that experience an extended life cycle will not be able successfully overwinter. Our observation of high mortality further supports the original field records of lower overwintering success in new adults (DeLeon *et al.*, 1934; Amman, 1973), and suggests both metabolic and physical drivers. This work provides new parameters for modeling the spread of mountain pine beetles in their expanding geographic range and is evidence of physiological plasticity in this insect. In both novel, colder regions like the Boreal Forest and warmer ecosystems which experience more developmental degree days, we may see the effect of natural selection amplifying varied life cycle lengths (longer or shorter) in mountain pine beetle.

4.6. Acknowledgments

We are grateful to the British Columbia Ministry of Environment and Climate

Change Strategy (Park Use Permit, Authorization #107171) and Robson Provincial Park staff for providing on-site assistance. Thanks also to Dr. Lisa Poirier for providing the data logger used in this study. Thanks to my husband Neil Thompson, for calling me back from Alberta so we could sample on the coldest day of the year, and to Daisy, the bravest little canine field assistant ever.

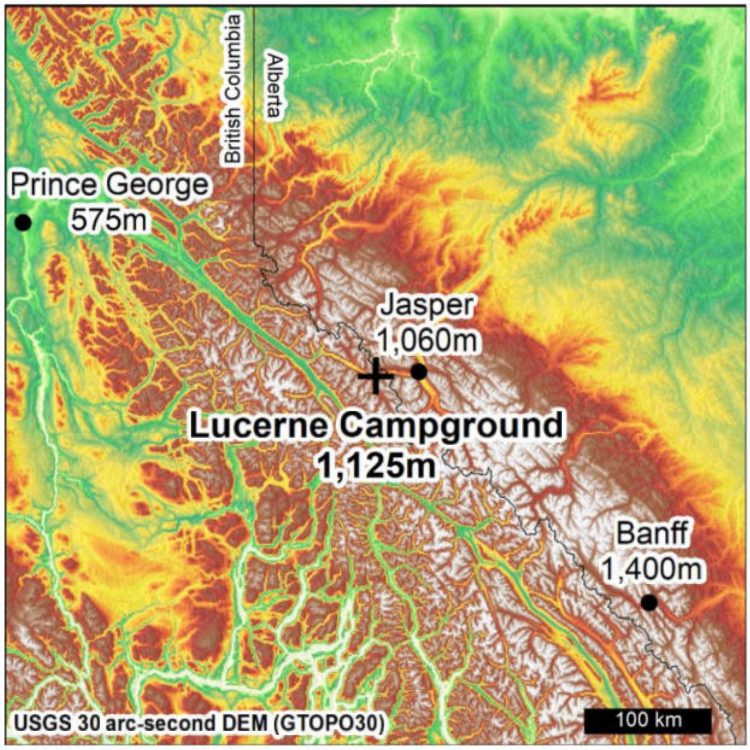

Figure 4.1: Location of the study site (cross) in relation to the closest major urban centers (dots). The sample site itself is located proximal to the Continental Divide of the Americas (latitude 52°50'59.48"N, longitude 118°34'21.84"W).

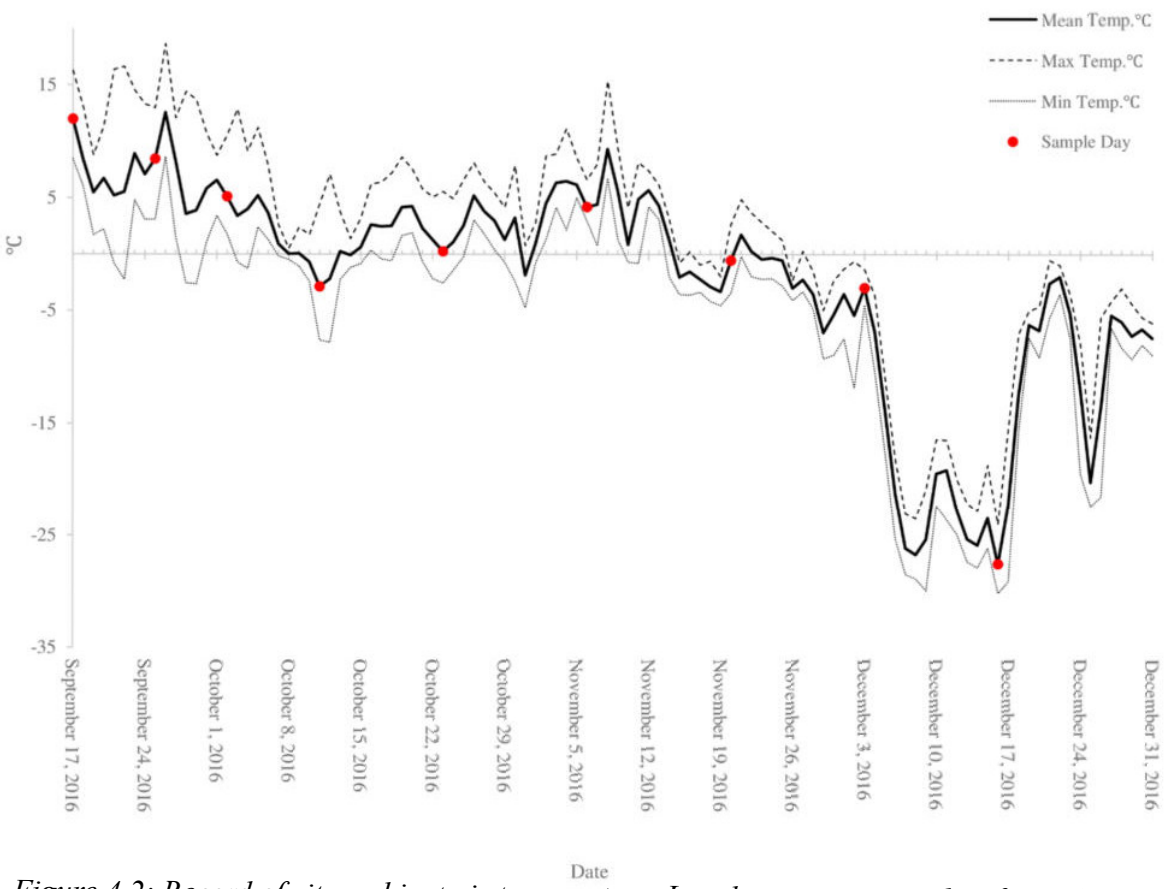

Figure 4.2: Record of site ambient air temperature. Local temperatures taken from climate data loggers at Lucerne Campground in Robson Provincial Park from September to December 2016.

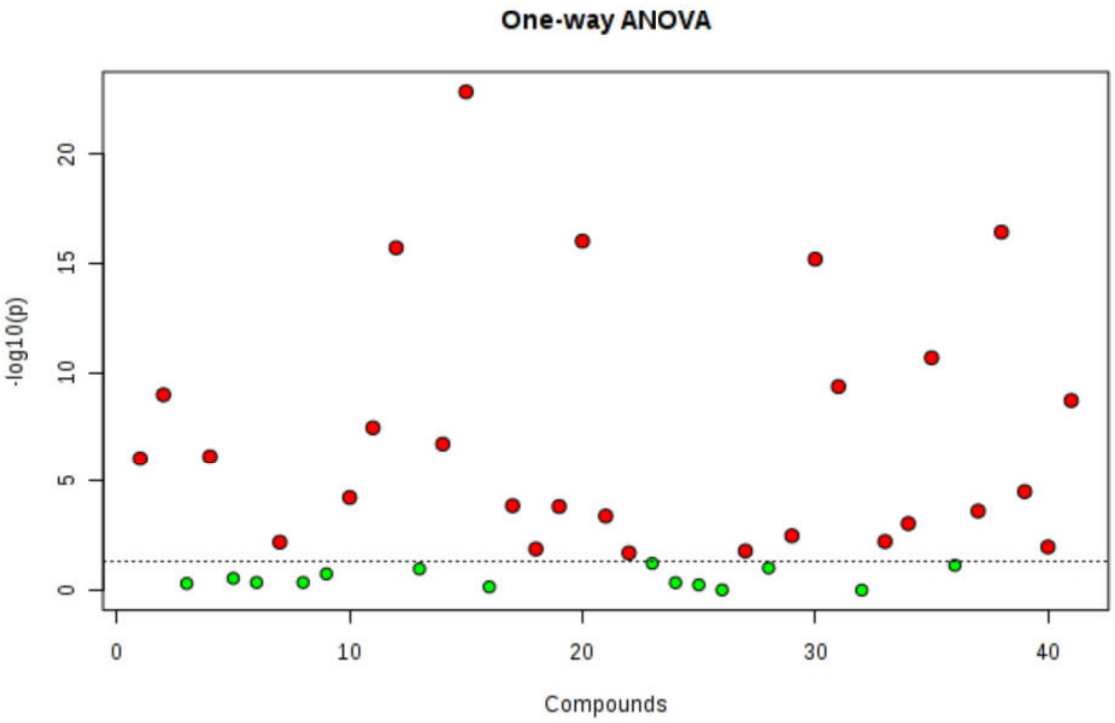

Figure 4.3: One-way ANOVAs for all metabolites sampled. All points marked in red exhibited significant differences between time points during the study while all points marked in green did not vary significantly during the study (FDR=0.05, n=8 samples per time point).

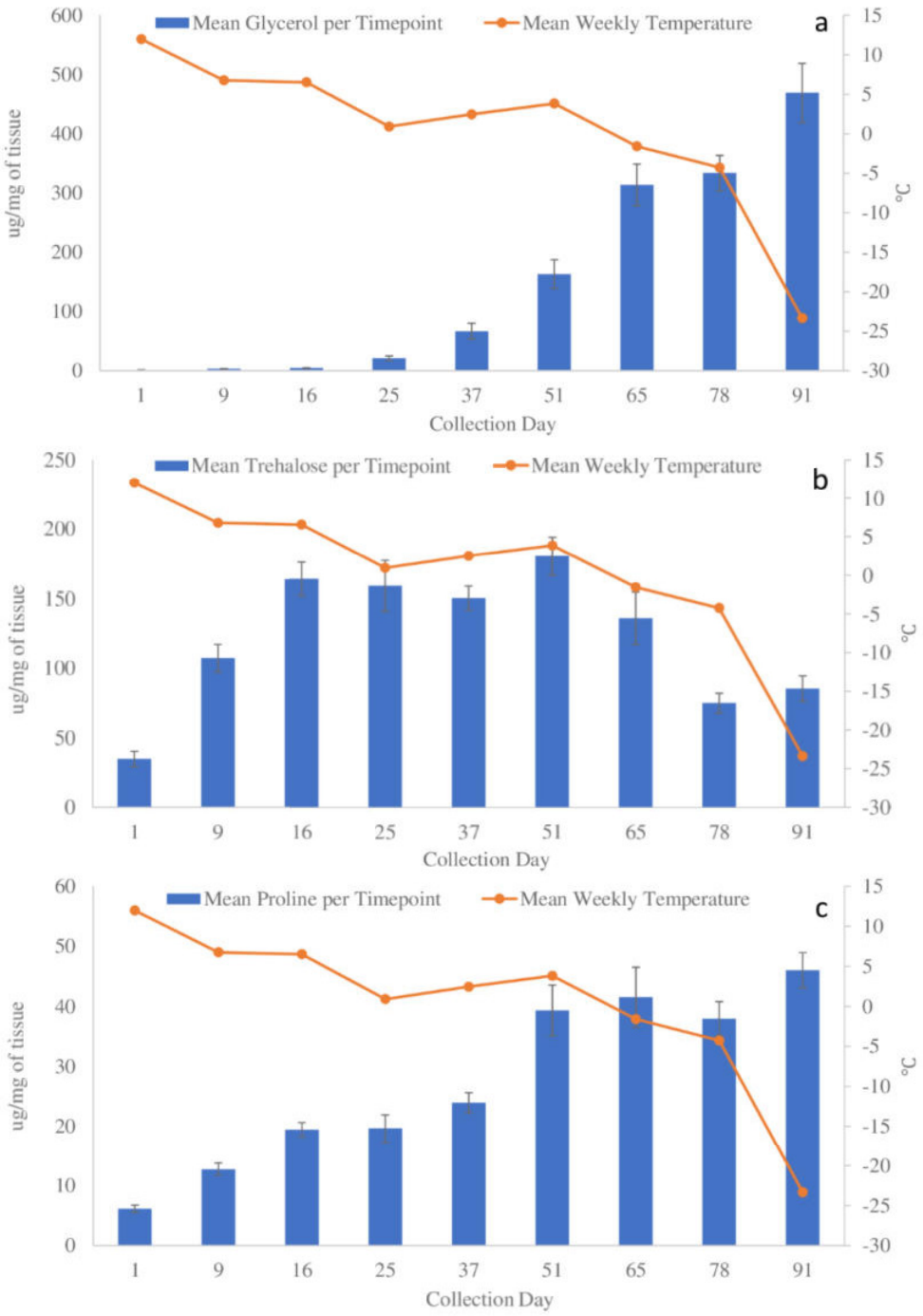

Figure 4.4: Mean glycerol (a), trehalose (b), and proline (c) concentrations of new adult beetles (n=8 samples per time point) in relation to ambient site temperature. Error bars show standard error of the mean.

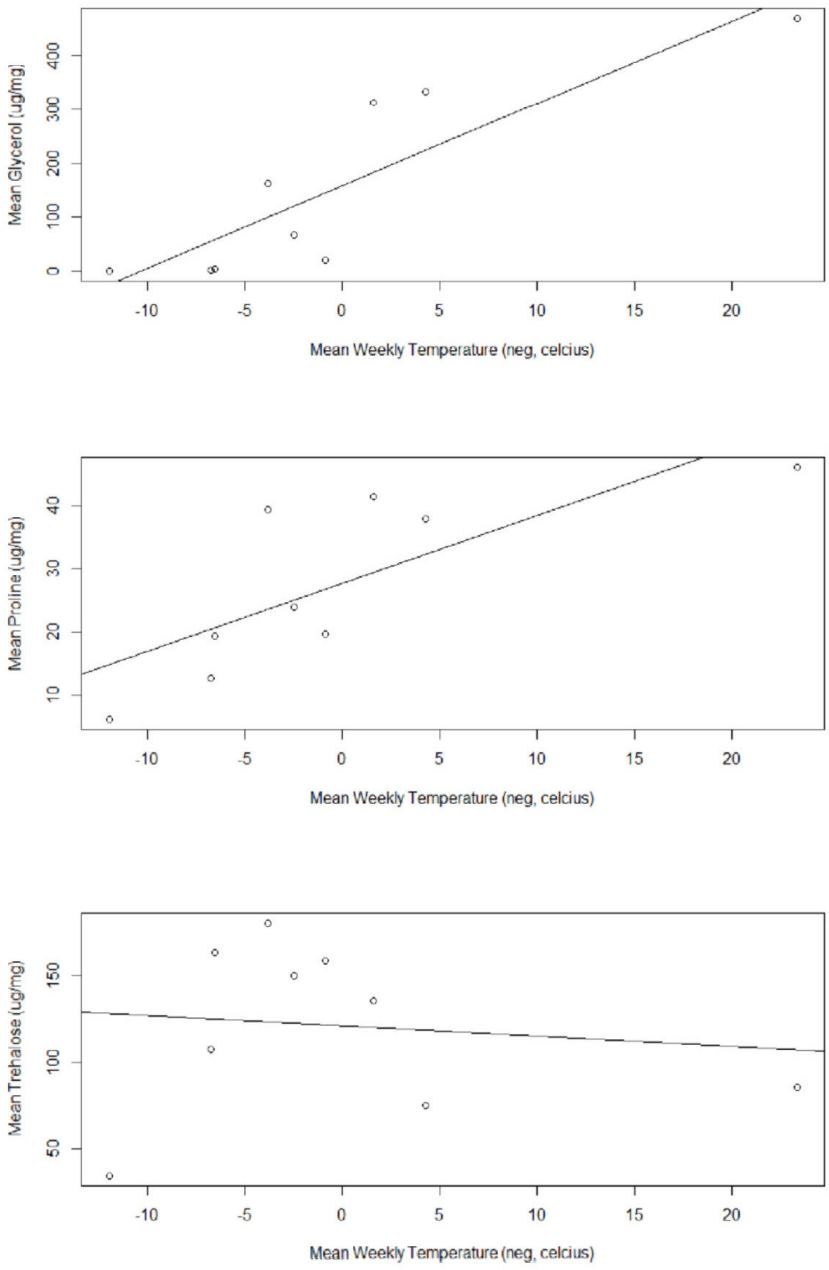

Figure 4.5: Pearson's product-moment correlation between mean metabolite concentration and negatively transformed temperature (°C) over time (n=8 samples per timepoint, µg/mg of tissue). Mean glycerol concentration (r = 0.873, p < 0.01, R2 = 0.7615) and mean proline concentration (r = 0.767, p < 0.05, R2 = 0.588) were both significant while mean trehalose concentration (r = 0.125, p > 0.05, R2 = 0.0157) was non-significant.

5. Concluding Synthesis

5.1. Dissertation Synopsis

In this dissertation, I explored the genetic structure and genetic properties of the destructive sub-cortical insect pest, the mountain pine beetle, throughout its recently expanded range within the north of British Columbia and Alberta. This outbreak slowly increased in size during the late 1990s to peak within BC in the mid-2000s and sent billions of beetles into the neighbouring province of Alberta. My work takes advantage of samples collected from the peak of the outbreak within BC and from the results of presumed in-flight populations in Alberta to establish the historical population structure when beetle numbers were at their highest. I also completed the most extensive field collection of MPB to date within the expanded range in Alberta to develop a comprehensive genetic structure profile of the new range of the species and explore the influence of environmental factors on outlier SNPs.

My work demonstrated that a majority of the novel colonized regions of central and northern Alberta, excluding Jasper National Park, retain the characteristics of their originating northern BC source population. I also demonstrated that mixed populations of southern and northern beetles on the landscape are capable of interbreeding, with none of the presumed sterility issues identified in other MPB populations clusters (Bracewell *et al.*, 2017), and then transition to more admixed state population over time, as seen in the Jasper National Park region. Also, I have found that new adult MPB from this admixed zone experiencing a semivoltine lifecycle complete a similar cold hardening process to that documented in their larval counterparts in BC. In this synthesis, I will review these results

and discuss the implications of this research for management of MPB in western Canada.

Mountain pine beetle is one of the first forest pests to experience a large-scale climate change-related range expansion in North America. Due to the life cycle of the beetles and their ability to grow and reproduce in the absence of brood-killing winters, the population grew to an extent beyond the control of normal ecological constraints (parasites, predators, tree constitutive defenses) could not mitigate their growth on the landscape. The outbreak spread from multiple locations across the province of British Columbia flowing northward and east to Alberta, limited in some areas only by prevailing winds and high-elevation mountains. Though MPB has been a consistent part of BC's pine ecosystems prior to European settlement, the movement of beetles northward across the province and the sheer number of beetles on the landscape in the mid-2000s led to serious concern about causes of this outbreak and the ramifications for the future of pine forests in BC and Alberta.

For **Chapters 2** and **3** of this work, I collected beetles from as many active MPB infestation sites as possible within BC, Jasper National Park, and the rest of Alberta during 2016. Collection sites were identified from a combination of aerial survey records from both BC and Alberta, coupled with extensive travel throughout both provinces to visually identify active areas of infestation, particularly in areas with reduced MPB activity in the west, north, and south of BC. When a subset of these beetles from 2016 were compared to counterpart sites from 2005/2007 several distinct patterns of genetic change emerged. Population structure was maintained in the southern parts of BC known to have a history of MPB outbreaks prior the early 2000s, particularly in regions that were isolated from the main northward spread of the MPB outbreak. I also found that novel sites of colonization

retained the character of their source population and did not differentiate into a separate population group.

By and large, it appears that MPB within the north of BC and the north of Alberta are all contained within the same population cluster and the passing of the MPB outbreak not only established new population sites on the landscape, but also removed almost all population structure in the north, excluding the Jasper National Park and Mount Robson Provincial Park region that now has a clear admixed structure. During sample collection within Alberta, I established transects that approximated linear paths eastward and northward toward the edges of the known MPB infestations. Across these gradients, we found that beetles from the Jasper National Park area (including Mount Robson Provincial Park but excluding the Banff National Park area) showed an admixed population of individuals with mixed populations assignments for all individuals. Moving away from Jasper National Park into Alberta, we found that sampled MPB populations immediately switched to a northern assignment, matching the character of their original inflights. This pattern was repeated along those rough transects through Alberta with individuals assorting to the northern population cluster moving both east and north.

While it is quite likely that Jasper National Park will act as an ongoing source of MPB for locations proximal to the park, it is my assessment that a majority of the beetles from regions in central and northern Alberta originated from the first major inflight of beetles from northern BC identified by Jackson *et al.* (2008). Jasper itself has not contributed meaningfully to the spread of MPB across Alberta. What is clear from my data is that Jasper represents a region of genetic admixture, and evidence from **Chapter 2** indicates that this admixture has likely been caused by the region receiving beetles from

both southern and northern populations. It is possible that these admixed beetles will develop into a distinct third genetic cluster over time, but I did not find evidence that they have genetic traits that indicate they will behave differently than other MPB. As populations have severely declined in northern and southern BC, continued replenishment of pure individuals into the park is not likely, and as the Jasper population likewise regresses to an endemic state, we may see a refining of the cluster signature due to the expected bottleneck effects from low population size until the eventual occurrence of the next outbreak.

One of the major concerns with a range expansion is the possibility of development of new adaptive mutations that change the behaviour or reproductive success of the species. There remain questions about MPB's ability to adapt and succeed in the north of Alberta (Trevoy *et al.*, 2018; Cullingham *et al.*, 2019; Pokorny, 2021), particularly as parts of the lodgepole forests within the province are exposed to much colder temperatures than that of historic MPB territory in BC. While many of the colonized parts of Alberta do have winter temperatures that are considered sufficient to cause extensive brood mortality above the snow line (i.e.,< –40°C), there is concern that MPB may develop greater cold tolerance or other traits related to landscape variables in the new territory that will amplify the ability of the beetle to cause damage. MPB is inexorably linked with climatic conditions, therefore these new environmental factors are likely to lead to selection events throughout their genome. There is also the possibility of randomly arising mutations gaining widespread distribution through expanding populations outside of selection processes. This process is similar to a founder effect with mutant alleles "surfing" the expanding edge of the population (Klopfstein *et al.*, 2006). For this reason, I explored the presence of MPB

SNP outliers throughout the new edges of the range in **Chapter 3.**

I found that, while there are outlier loci present within the SNP dataset, they are linked to a diverse set of potential genes with a variety of functions and not the product of one unique environmental factor. Many of the outlier loci are linked to putative ion channel proteins that relate to neuron function or energy metabolism. Though none of these genes are directly linked to metabolic pathways related to cold tolerance, it is possible that they represent an indirect response to the harsher and colder climate encountered within the north of both BC and Alberta. Neuro-muscular linked outliers may also have connections to dispersal ability within these novel stands. The expanded population of beetles has only been a resident of northern Alberta for a little over a decade and it is likely that the outliers I detected represent a combination of response to conditions experienced by beetles within northern BC in the early 2000s in addition to those currently experienced within Alberta.

While Alberta does experience colder temperatures in the north of the province, with winter temperatures often reaching below –40 °C, climate in the north of Canada is expected to continue to warm at a faster rate than lower latitudes (www.ipcc.ch; Bush & Lemmen, 2019). MPB populations are likely to experience harsher conditions in Alberta that start to become slightly milder in some regions as the climate warms. This means that the selective pressure from colder temperatures may decrease over time. We do already see the positive correlation of a longer as opposed to shorter frost free period on the genotypes within the Alberta range expansion. This means that colder regions sampled with shorter frost free periods are not causing outlier loci within MPB populations. I also found evidence that MPB SNP loci are being influenced by precipitation and humidity factors in all sampled regions, both elements of the environment that will continue to change as the

climate changes. Certain areas of BC and Alberta will experience more precipitation dependant on the time of year, but what is likely to occur in Alberta is an increase in droughty conditions, which will lead to stressed trees that are less capable of mounting a strong resin defense against the beetle (Erbilgin *et al.*, 2021; https://droughtmonitor.unl.edu, 2022). My data did not support the hypothesis that rapid selection events are occurring within Albertan beetle populations, nor did my results indicate that MPB in the expanded range have genetic reasons to behave dramatically differently from MPB studied elsewhere. For this reason, genetically tailored management techniques are likely not needed at this time.

During the course of collecting beetles in BC and Alberta in 2016 I encountered several high-elevation sites that displayed delayed development compared to other locations. As these sites seemed likely to produce the semivoltine MPB life cycle identified by Amman (1972), one location was chosen for further investigation in **Chapter 4** to track the metabolite profiles of un-emerged new adults as they prepared for overwintering. We found that these new adults produced the same cryoprotectants as MPB larvae, though they do experience high mortality at cold temperatures (< -30°C). While these results indicate that new adults will not become part of the life cycle of the following year in very cold locations, they also indicate that MPB will produce viable semivoltine beetles in areas with milder temperatures than the area sampled. As the climate continues to warm, we will likely see changes within MPB life cycles at higher elevation sites in Alberta, resulting in summer flights that blend adults from both univoltine and semivoltine life cycles.

5.2. Future Directions

This work represents the most recent comprehensive sampling of MPB within Alberta and BC. I am also the first to track the short-term changes in MPB genetic structure over time. What is clear, based on both my own data, that of other researchers, and anecdotal field observations, is that MPB is likely to become an endemic part of the lodgepole pine forests of Alberta, beyond the historic regions of Banff and Cypress Hills (MacCormick, 2020; Pokorny, 2021). Across novel territory MPB is still colonizing predominantly lodgepole stands that are a preferred host and considered part of a coevolved system, even if that particular host may be considered evolutionarily naïve. I expect MPB through the expanded range to follow the same population dynamics experienced by other species of irruptive *Dendroctonus*, moving from epidemic numbers and behaviours to endemic numbers and behaviours. One site sampled near the Whitecourt area (Site Y) in Alberta in 2016 already showed clear evidence of endemic-style behaviour, with beetles on this site observed to be preferentially attacking severely weakened trees infected by the root pathogen *Armillaria ostoyae* (Family: Physalacriaceae. Order: Agaricales) (Figure 5.1).

Figure 5.1: Photo of an affected lodgepole pine from the Whitecourt site, 3m from the sampled tree

MPB populations will continue to decline in Alberta as part of a natural progression toward an endemic state, a population decline that was recently accelerated by overall cooler winter temperatures in 2019 and 2020 and a cool and wet summer in 2019 (Belanger, 2022). While MPB numbers will decrease, complete eradication is not a realistic possibility. Resident MPB in Alberta are likely to persist in small numbers, attacking weakened trees at such a low level they will be difficult to detect by aerial surveys, and may only be observable by extremely thorough land-based surveys.

Excluding Jasper National Park, it is clear that MPB in the north of Alberta represent one genetic population at the time of sampling. As beetles regress to low numbers, we may see the establishment of new population structure on the Alberta landscape, by the development of the admixed population within Jasper National Park, but

also in response to regions of warmer and cooler temperatures across the new range. This change is likely to take many decades, if not longer. Continual monitoring of all known active populations should be undertaken to watch for changes in behaviour over time or increased survival in colder regions, similar to previous monitoring in the southwestern USA for voltinism changes. If possible, resampling of MPB from the same areas from this dissertation in 10 to 20 years, similar to what I accomplished in **Chapter 2**, would be very informative in tracking these changes. There are many hundreds of excess georeferenced ethanol-preserved beetles from this dissertation that will remain viable for DNA extraction in the future. Samples from this collection could also be lyophilized for dry, long-term storage in order to produce a collection of beetles that could be re-tested against resident beetles in the event of new outbreaks in 50 to 100 years.

There is strong evidence to suggest that MPB will struggle to establish a niche within the pure jack pine of boreal forest due to competition from other beetles and wood borers (Raffa *et al.* 2015, Pokorny, 2021). However, the hybrid zone between jack and lodgepole pine may be an area where MPB would have greater ability to establish a niche. It would also be very informative to compare and contrast future beetles from the purer lodgepole pine areas of Alberta and those from the admixed zone with the extra beetles from this work that remain in storage, again after a period of 10 to 20 years. Specific site variables could also be co-collected including soil pH, moisture, nutrient content, stand characteristics, and needle clippings to quantify host pine genetics, and used within a redundancy analysis or other ecological association analysis similar to **Chapter 3** to further explore the landscape predictors of selection within the beetle populations.

6. References

Abdellaoui, A., Hottenga, J.J., Knijff, P. De, Nivard, M.G., Xiao, X., Scheet, P., Brooks, A., Ehli, E.A., Hu, Y., Davies, G.E., Hudziak, J.J., Sullivan, P.F., Van Beijsterveldt, T., Willemsen, G., De Geus, E.J., Penninx, B.W.J.H., and Boomsma, D.I. 2013. Population structure, migration, and diversifying selection in the Netherlands. Eur. J. Hum. Genet. **21**: 1277–1285. Nature Publishing Group. doi:10.1038/ejhg.2013.48.

Alleman, A., Stoldt, M., Feldmeyer, B., and Foitzik, S. 2019. Tandem-running and scouting behaviour are characterized by up-regulation of learning and memory formation genes within the ant brain. Mol. Ecol. **28**: 2342–2359. John Wiley & Sons, Ltd. doi:10.1111/mec.15079.

Allen, C.D., Macalady, A.K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., Kitzberger, T., Rigling, A., Breshears, D.D., Hogg, E.H., Gonzalez, P., Fensham, R., Zhang, Z., Castro, J., Demidova, N., Lim, J.H., Allard, G., Running, S.W., Semerci, A., and Cobb, N. 2010. A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. For. Ecol. Manage. **259**: 660–684. Elsevier. doi:10.1016/j.foreco.2009.09.001.

Allendorf, F.W., Luikart, G.H., and Aitken, S.N. 2013. Conservation and the Genetics of Populations. 2nd edition. Wiley-Blackwell, West Sussex, UK.

Amman, G.D. 1973. Population Changes of the Mountain Pine Beetle in Relation to Elevation. Environ. Entomol. **2**: 541–548. doi:10.1093/ee/2.4.541.

Amman, G.D., and Cole, W.E. 1983. Mountain pine beetle dynamics in lodgepole pine forests. Part II: Populations dynamics. USDA Forest Service. Gen. Tech. Rep. INT-145.

Aukema, B.H., Carroll, A.L., Zhu, J., Raffa, K.F., Sickley, T.A., and Taylor, S.W. 2006. Landscape level analysis of mountain pine beetle in British Columbia, Canada: spatiotemporal development and spatial synchrony within the present outbreak. Ecography (Cop.). **29**: 427–441. doi:10.1111/j.2006.0906-7590.04445.x.

Bartell, N. V 2008. A Microsatellite Analysis of the Western Canadian Mountain Pine Beetle (*Dendroctonus ponderosae*) Epidemic: Phylogeography and Long Distance Dispersal Patterns. University of Northern British Columbia.

Batista, P.D., Janes, J.K., Boone, C.K., Murray, B.W., and Sperling, F.A.H. 2016. Adaptive and neutral markers both show continent-wide population structure of mountain pine beetle (*Dendroctonus ponderosae*). Ecol. Evol. **6**: 6292–6300. doi:10.1002/ece3.2367.

Bayé, A., Comellas-Bigler, M., Rodríguez De La Vega, M., Maskos, K., Bode, W., Aviles, F.X., Jongsma, M.A., Beekwilder, J., and Vendrell, J. 2005. Structural basis of the resistance of an insect carboxypeptides to plant protease inhibitors. Proc. Natl. Acad.

Sci. U. S. A. **102**: 16602–16607. doi:10.1073/pnas.0505489102.

Belanger, D. 2022. AFRED MPB Survey and Control Program Update for 2021-22. Bugs Dis. **33**: 1–14.

Bentz, B.J., Duncan, J.P., and Powell, J.A. 2016. Elevational shifts in thermal suitability for mountain pine beetle population growth in a changing climate. Forestry **89**: 271–283. doi:10.1093/forestry/cpv054.

Bentz, B.J., Hansen, E.M., Davenport, M., and Soderberg, D. 2022. Complexities in predicting mountain pine beetle and spruce beetle response to climate change. Pages 31–54 *in* K.J.K. Gandhi and R.W. Hofstetter, eds. Bark Beetle Management, Ecology, and Climate Change. Academic Press. doi:10.1016/b978-0-12-822145-7.00013-1.

Bentz, B.J., and Mullins, D.E. 1999. Ecology of mountain pine beetle (Coleoptera: Scolytidae) cold hardening in the intermountain west. Environ. Entomol. **28**: 577–587. doi:10.1093/ee/28.4.577.

Bentz, B.J., and Powell, J.A. 2014. Mountain pine beetle seasonal timing and constraints to bivoltinism: A comment on Mitton and Ferrenberg, "Mountain pine beetle develops an unprecedented summer generation in response to climate warming." Am. Nat. **184**: 787–796. doi:10.1086/678405.

Bentz, B.J., Rgnire, J., Fettig, C.J., Hansen, E.M., Hayes, J.L., Hicke, J.A., Kelsey, R.G., Negron, J.F., and Seybold, S.J. 2010. Climate change and bark beetles of the western United States and Canada: Direct and indirect effects. Bioscience **60**: 602–613. doi:10.1525/bio.2010.60.8.6.

Bleiker, K. 2016. Mountain pine beetle (factsheet). [Online] Available: http://www.nrcan.gc.ca/forests/fire-insects-disturbances/top-insects/13397 [2021 Apr. 28].

Bleiker, K.P., Smith, G.D., and Humble, L.M. 2017. Cold tolerance of mountain pine beetle (coleoptera: Curculionidae) eggs from the historic and expanded ranges. Environ. Entomol. **46**: 1165–1170. doi:10.1093/ee/nvx127.

Bond, P.A., and Sang, J.H. 1968. Glutamate dehydrogenase of *Drosophila* larvae. J. Insect Physiol. **14**: 341–359. Pergamon. doi:10.1016/0022-1910(68)90076-0.

Bonnett, T.R., Robert, J.A., Pitt, C., Fraser, J.D., Keeling, C.I., Bohlmann, J., and Huber, D.P.W.W. 2012. Global and comparative proteomic profiling of overwintering and developing mountain pine beetle, Dendroctonus ponderosae (Coleoptera: Curculionidae), larvae. Insect Biochem. Mol. Biol. **42**: 890–901. Elsevier Ltd. doi:10.1016/j.ibmb.2012.08.003.

Boone, C.K., Aukema, B.H., Bohlmann, J., Carroll, A.L., and Raffa, K.F. 2011. Efficacy of tree defense physiology varies with bark beetle population density: A basis for positive feedback in eruptive species. Can. J. For. Res. **41**: 1174–1188. Canadian Science Publishing. doi:10.1139/x11-041.

Boone, C.K., Thompson, K.M., Henry, P., and Murray, B.W. 2022. Host use does not

drive genetic structure of mountain pine beetles in western North America. bioRxiv: 2022.06.28.498011. Cold Spring Harbor Laboratory. doi:10.1101/2022.06.28.498011.

Borden, J.H., Ryker, L.C., Chong, L.J., Pierce Jr., H.D., Johnston, B.D., and Oehlschlager, A.C. 1987. Response of the mountain pine beetle, *Dendroctonus ponderosae* Hopkins (Coleoptera: Scolytidae), to five semiochemicals in British Columbia lodgepole pine forests. Can. J. For. Res. **17**: 118–128. NRC Research Press Ottawa, Canada. doi:10.1139/x87-023.

Bracewell, R.R., Bentz, B.J., Sullivan, B.T., and Good, J.M. 2017. Rapid neo-sex chromosome evolution and incipient speciation in a major forest pest. Nat. Commun. **8**: 1593. Springer US. doi:10.1038/s41467-017-01761-4.

Burke, J.L. 2016. Consequences of climate-induced range expansion of a native invasive herbivore in western Canada. University of British Columbia.

Bush, E., and Lemmen, D.. (eds.) 2019. Canada's Changing Climate Report. Environment and Climate Change Canada, Government of Canada. Government of Canada, Ottawa, ON. [Online] Available: www.ChangingClimate.ca/CCCR2019.

Campbell, E.O., Davis, C.S., Dupuis, J.R., Muirhead, K., and Sperling, F.A.H. 2017. Cross-platform compatibility of de novo-aligned SNPs in a nonmodel butterfly genus. Mol. Ecol. Resour. **17**: 1–10. doi:10.1111/1755-0998.12695.

Carrasco, M.A., Buechler, S.A., Arnold, R.J., Sformo, T., Barnes, B.M., and Duman, J.G. 2012. Investigating the deep supercooling ability of an Alaskan beetle, Cucujus clavipes puniceus, via high throughput proteomics. J. Proteomics **75**: 1220–1234. doi:10.1016/j.jprot.2011.10.034.

Carroll, a, Régnière, J., Logan, J. a, Taylor, S.W., and Powell, J. a 2006. Impacts of climate change on range expansion by the mountain pine beetle. [Online] Available: http://cfs.nrcan.gc.ca/pubwarehouse/pdfs/26601.pdf.

Carroll, R.P., Litvaitis, M.K., Clements, S.J., Stevens, C.L., and Litvaitis, J.A. 2019. History matters: contemporary versus historic population structure of bobcats in the New England region, USA. Conserv. Genet. **20**: 743–757. Springer Netherlands. doi:10.1007/s10592-019-01170-8.

Chaturvedi, A., and Pierce, S.K. 2009. How location governs toll-like receptor signaling. Traffic **10**: 621–628. John Wiley & Sons, Ltd. doi:10.1111/j.1600-0854.2009.00899.x.

Cheng, J.R. 1980. Tourism: how much is too much? Lessons for Canmore from Banff. Can. Geogr. / Le Géographe Can. **24**: 72–80. doi:10.1111/j.1541-0064.1980.tb00977.x.

Chiu, C.C. 2018. Detoxification of Pine Terpenoids by Mountain Pine Beetle Cytochromes P450. University of British Columbia.

Chong, J., and Xia, J. 2018. MetaboAnalystR: An R package for flexible and reproducible analysis of metabolomics data. Bioinformatics **34**: 4313–4314.

doi:10.1093/bioinformatics/bty528.

Ciosi, M., Miller, N.J., Kim, K.S., Giordano, R., Estoup, A., and Guillemaud, T. 2008. Invasion of Europe by the western corn rootworm, *Diabrotica virgifera virgifera*: Multiple transatlantic introductions with various reductions of genetic diversity. Mol. Ecol. **17**: 3614–3627. doi:10.1111/j.1365-294X.2008.03866.x.

Clements, D.R., and DiTommaso, A. 2012. Predicting weed invasion in Canada under climate change: Evaluating evolutionary potential. Can. J. Plant Sci. **92**: 1013–1020. doi:10.4141/CJPS2011-280.

Cooke, B.J., and Carroll, A.L. 2017. Predicting the risk of mountain pine beetle spread to eastern pine forests: Considering uncertainty in uncertain times. For. Ecol. Manage. **396**: 11–25. doi:10.1016/j.foreco.2017.04.008.

Cowley, P. 2022.June 13. Mountain pine beetles continue trek east. Red Deer Advocate. Red Deer, Atla. [Online] Available: https://www.reddeeradvocate.com/news/mountain-pine-beetles-continue-trek-east/ [2022 Jul. 6].

Crossley, M.S., Chen, Y.H., Groves, R.L., and Schoville, S.D. 2017. Landscape genomics of Colorado potato beetle provides evidence of polygenic adaptation to insecticides. Mol. Ecol. **26**: 6284–6300. John Wiley & Sons, Ltd. doi:10.1111/mec.14339.

Cullingham, C.I., Cooke, J.E.K., Dang, S., Davis, C.S., Cooke, B.J., and Coltman, D.W. 2011. Mountain pine beetle host-range expansion threatens the boreal forest. Mol. Ecol. **20**: 2157–2171. doi:10.1111/j.1365-294X.2011.05086.x.

Cullingham, C.I., Janes, J.K., Hamelin, R.C., James, P.M.A., Murray, B.W., and Sperling, F.A.H. 2019. The contribution of genetics and genomics to understanding the ecology of the mountain pine beetle system. Can. J. For. Res. **49**: 721–730. doi:10.1139/cjfr-2018-0303.

Cullingham, C.I., Roe, A.D., Sperling, F.A.H., and Coltman, D.W. 2012. Phylogeographic insights into an irruptive pest outbreak. Ecol. Evol. **2**: 908–919. doi:10.1002/ece3.102.

Damschen, E.I., Harrison, S., and Grace, J.B. 2010. Climate change effects on an endemic-rich edaphic flora: Resurveying Robert H. Whittaker's Siskiyou sites (Oregon, USA). Ecology **91**: 3609–3619. doi:10.1890/09-1057.1.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., and Durbin, R. 2011. The variant call format and VCFtools. Bioinformatics **27**: 2156–2158. doi:10.1093/bioinformatics/btr330.

Daniels, L., Gray, R., and Burton, P. 2020. 2017 megafires in British Columbia: urgent need to adapt and improve resilience to wildfire. Pages 51–62 *in* Proceedings of the Fire Continuum-Preparing for the future of wildland fire. [Online] Available: https://www.fs.usda.gov/treesearch/pubs/62325 [2022 Jul. 6].

Davey, C.M., Devictor, V., Jonzén, N., Lindström, Å., and Smith, H.G. 2013. Impact of

climate change on communities: Revealing species' contribution. J. Anim. Ecol. **82**: 551–561. John Wiley & Sons, Ltd. doi:10.1111/1365-2656.12035.

Dawson, R.J., Werner, A.T., and Murdock, T.Q. 2008. Preliminary analysis of climate change in the Cariboo-Chilcotin area of British Columbia. Pacific Clim. Impacts Consortium, Univ. Victoria, Victoria BC: 1–49. [Online] Available: https://pacificclimate.org/sites/default/files/publications/Werner.ClimateChangeCaribooChilcotin.Sep2008.pdf.

DeLeon, D., Bedard, W.D., and Terrell, T.T. 1934. Recent discoveries concerning the biology of the mountain pine beetle and their effect on control in western white pine stands. J. For. **32**: 430–436. doi:10.1093/jof/32.4.430.

Dormann, C.F. 2007. Promising the future? Global change projections of species distributions. Basic Appl. Ecol. **8**: 387–397. doi:10.1016/j.baae.2006.11.001.

Dowle, E.J., Bracewell, R.R., Pfrender, M.E., Mock, K.E., Bentz, B.J., and Ragland, G.J. 2017. Reproductive isolation and environmental adaptation shape the phylogeography of mountain pine beetle (*Dendroctonus ponderosae*). Mol. Ecol. **26**: 6071–6084. doi:10.1111/mec.14342.

Draper, D. 2000. Toward sustainable mountain communities: Balancing tourism development and environmental protection in Banff and Banff National Park, Canada. Ambio **29**: 408–415. Royal Swedish Academy of Sciences. doi:10.1579/0044-7447-29.7.408.

Duan, X., Peng, X., Qiao, X., and Chen, M. 2017. Life cycle and population genetics of bird cherry-oat aphids *Rhopalosiphum padi* in China: an important pest on wheat crops. J. Pest Sci. (2004). **90**: 103–116. Springer Berlin Heidelberg. doi:10.1007/s10340-016-0752-9.

Earl, D.A., and VonHoldt, B.M. 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. **4**: 359–361. Springer Verlag. doi:10.1007/s12686-011-9548-7.

Elnitsky, M.A., Hayward, S.A.L., Rinehart, J.P., Denlinger, D.L., and Lee, R.E. 2008. Cryoprotective dehydration and the resistance to inoculative freezing in the Antarctic midge, *Belgica antarctica*. J. Exp. Biol. **211**: 524–530. doi:10.1242/jeb.011874.

Emond, F.J., and Cerezke, H.F. 1991. Forest insect and disease conditions in Alberta, Saskatchewan, Manitoba and the Northwest Territories in 1989 and predictions for 1990. Northern Forestry Centre Information Report NOR-X-318.

Erbilgin, N., Zanganeh, L., Klutsch, J.G., Chen, S., Zhao, S., Ishangulyyeva, G., Burr, S.J., Gaylord, M., Hofstetter, R., Keefover-Ring, K., Raffa, K.F., and Kolb, T. 2021. Combined drought and bark beetle attacks deplete non-structural carbohydrates and promote death of mature pine trees. Plant Cell Environ. **44**: 3636–3651. John Wiley & Sons, Ltd. doi:10.1111/pce.14197.

Evanno, G., Regnaut, S., and Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. **14**: 2611–2620.

Evenden, J.C., Bedard, W.D., and Struble, G.R. 1943. The mountain pine beetle, an important enemy of western pines. Circular No. 664. United States Department of Agriculture, Washington, D. C. [Online] Available: https://hdl.handle.net/2027/uiug.30112019274122.

Evenden, M.L., Whitehouse, C.M., and Sykes, J. 2014. Factors Influencing Flight Capacity of the Mountain Pine Beetle (Coleoptera: Curculionidae: Scolytinae). Environ. Entomol. **43**: 187–196. doi:10.1603/EN13244.

Excoffier, L., and Lischer, H.E.L. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. **10**: 564–567. Mol Ecol Resour. doi:10.1111/j.1755-0998.2010.02847.x.

Falush, D., Stephens, M., and Pritchard, J.K. 2003. Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. Genetics **164**: 1567–1587. [Online] Available: http://pritch.

Falush, D., Stephens, M., and Pritchard, J.K. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol. Ecol. Notes **7**: 574–578. doi:10.1111/j.1471-8286.2007.01758.x.

Feng, Y., Xu, L., Li, W., Xu, Z., Cao, M., Wang, J., Tao, J., and Zong, S. 2016. Seasonal changes in supercooling capacity and major cryoprotectants of overwintering Asian longhorned beetle (*Anoplophora glabripennis*) larvae. Agric. For. Entomol. **18**: 302–312. doi:10.1111/afe.12162.

Fettig, C.J., Gibson, K.E., Munson, A.S., and Negrón, J.F. 2014. Cultural practices for prevention and mitigation of mountain pine beetle infestations. For. Sci. **60**: 450–463. doi:10.5849/forsci.13-032.

Forester, B.R., Lasky, J.R., Wagner, H.H., and Urban, D.L. 2018. Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. Mol. Ecol. **27**: 2215–2233. John Wiley & Sons, Ltd. doi:10.1111/mec.14584.

Franchini, P., Fruciano, C., Spreitzer, M.L., Jones, J.C., Elmer, K.R., Henning, F., and Meyer, A. 2014. Genomic architecture of ecologically divergent body shape in a pair of sympatric crater lake cichlid fishes. Mol. Ecol. **23**: 1828–1845. doi:10.1111/mec.12590.

Fraser, J.D., Bonnett, T.R., Keeling, C.I., and Huber, D.P.W. 2017. Seasonal shifts in accumulation of glycerol biosynthetic gene transcripts in mountain pine beetle, *Dendroctonus ponderosae* Hopkins (Coleoptera: Curculionidae), larvae. PeerJ **2017**: e3284. doi:10.7717/peerj.3284.

Fricke, U., Steffan-Dewenter, I., Zhang, J., Tobisch, C., Rojas-Botero, S., Benjamin, C.S., Englmeier, J., Ganuza, C., Haensel, M., Riebl, R., Uhler, J., Uphus, L., Ewald, J.,

Kollmann, J., and Redlich, S. 2022. Landscape diversity and local temperature, but not climate, affect arthropod predation among habitat types. PLoS One **17**: e0264881. Public Library of Science. doi:10.1371/journal.pone.0264881.

Gäde, G., and Auerswald, L. 2002. Beetles' choice - Proline for energy output: Control by AKHs. Comp. Biochem. Physiol. - B Biochem. Mol. Biol. **132**: 117–129. doi:10.1016/S1096-4959(01)00541-3.

Gayathri Samarasekera, G.D.N., Bartell, N. V., Lindgren, B.S., Cooke, J.E.K., Davis, C.S., James, P.M.A., Coltman, D.W., Mock, K.E., and Murray, B.W. 2012. Spatial genetic structure of the mountain pine beetle (*Dendroctonus ponderosae*) outbreak in western Canada: Historical patterns and contemporary dispersal. Mol. Ecol. **21**: 2931–2948. doi:10.1111/j.1365-294X.2012.05587.x.

Geffen, E., Anderson, M.J., and Wayne, R.K. 2004. Climate and habitat barriers to dispersal in the highly mobile grey wolf. Mol. Ecol. **13**: 2481–2490. John Wiley & Sons, Ltd. doi:10.1111/j.1365-294X.2004.02244.x.

Gibbs, J., and Sheffield, C.S. 2009. Rapid range expansion of the wool-carder bee, *Anthidium manicatum* (linnaeus) (Hymenoptera: Megachilidae), in North America. J. Kansas Entomol. Soc. **82**: 21–29. doi:10.2317/JKES805.27.1.

Gillette, N.E., Wood, D.L., Hines, S.J., Runyon, J.B., and Negrón, J.F. 2014. The once and future forest: Consequences of mountain pine beetle treatment decisions. For. Sci. **60**: 527–538. doi:10.5849/forsci.13-040.

Gonzalès, R., and Parrott, L. 2019. How stakeholders structure their collaborations to anticipate and tackle the threat of mountain pine beetle in the Jasper–Hinton (Alberta, Canada) area. Can. J. For. Res. **49**: 480–490. NRC Research Press. doi:10.1139/cjfr-2018-0314.

Goodsman, D.W., Grosklos, G., Aukema, B.H., Whitehouse, C., Bleiker, K.P., McDowell, N.G., Middleton, R.S., and Xu, C. 2018. The effect of warmer winters on the demography of an outbreak insect is hidden by intraspecific competition. Glob. Chang. Biol. **24**: 3620–3628. doi:10.1111/gcb.14284.

Government of Alberta 2008. Keep Your Forest Green - Do Not Cut and Transport Mountain Pine Beetle-Infested Pine Trees for Firewood. Government of Alberta. [Online] Available: https://www1.agriculture.alberta.ca/$Department/deptdocs.nsf/ba3468a2a8681f69872569d60073fde1/2a342f030a941e2c87257f6f00764426/$FILE/KeepYourForestsGreen.pdf.

Government of Alberta 2014. Shelterbelts for Livestock Farms in Alberta Overview. Alberta Ag and Forestry, Edmonton, Alberta. doi:Agdex 400/092-2.

Gruber, B., Unmack, P., Berry, O., and Georges, A. 2018. dartr: An r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. Mol. Ecol. Resour. **18**: 691–699. doi:10.1111/1755-0998.12745.

Grummer, J.A., Beheregaray, L.B., Bernatchez, L., Hand, B.K., Luikart, G., Narum, S.R., and Taylor, E.B. 2019. Aquatic Landscape Genomics and Environmental Effects on Genetic Variation. Trends Ecol. Evol. **34**: 641–654. Elsevier Ltd. doi:10.1016/j.tree.2019.02.013.

Hagen, S.B., Kopatz, A., Aspi, J., Kojola, I., and Geir Eiken, H. 2015. Evidence of rapid change in genetic structure and diversity during range expansion in a recovering large terrestrial carnivore. Proc. R. Soc. B Biol. Sci. **282**: 20150092. The Royal Society. doi:10.1098/rspb.2015.0092.

Hallegraeff, G.M. 2010. Ocean climate change, phytoplankton community responses, and harmful algal blooms: A formidable predictive challenge. J. Phycol. **46**: 220–235. John Wiley & Sons, Ltd. doi:10.1111/j.1529-8817.2010.00815.x.

Hanski, I. 1998. Metapopulation dynamics. Nature **396**: 41–49. Nature Publishing Group. doi:10.1038/23876.

Hardy, A., and Gretzel, U. 2008. It's All About Me: Understanding Recreational Vehivle Usage (Caravanning) on the Alaska Highway. Pages 1–11 *in* Cauthe.

Harris, S.A., and Brown, R.J.E. 1978. Plateau Mountain: A Case Study of Alpine Permafrost in the Canadian Rocky Mountains. Pages 385–391 *in* Proceedings 3rd International Conference on Permafrost. Edmonton, Alberta.

Harvey, S.H., Krien, M.J.E., and O'Connell, M.J. 2002. Structural maintenance of chromosomes (SMC) proteins, a family of conserved ATPases. Genome Biol. **3**: 1–5. BioMed Central. doi:10.1186/GB-2002-3-2-REVIEWS3003.

Hellberg, M.E., Balch, D.P., and Roy, K. 2001. Climate-driven range expansion and morphological evolution in a marine gastropod. Science (80-.). **292**: 1707–1710. American Association for the Advancement of Science. doi:10.1126/science.1060102.

Herborg, L.M., Weetman, D., Van Oosterhout, C., and Hänfling, B. 2007. Genetic population structure and contemporary dispersal patterns of a recent European invader, the Chinese mitten crab, *Eriocheir sinensis*. Mol. Ecol. **16**: 231–242. John Wiley & Sons, Ltd. doi:10.1111/j.1365-294X.2006.03133.x.

Himmel, N.J., and Cox, D.N. 2020. Transient receptor potential channels: current perspectives on evolution. Proc. R. Soc. B Biol. Sci. **287**. The Royal Society. doi:10.1098/rspb.2020.1309.

Hoban, S., Kelley, J.L., Lotterhos, K.E., Antolin, M.F., Bradburd, G., Lowry, D.B., Poss, M.L., Reed, L.K., Storfer, A., and Whitlock, M.C. 2016. Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. Am. Nat. **188**: 379–397. doi:10.1086/688018.

Hofstetter, R.W., and Gandhi, K.J.K. 2022. Interactions among climate, disturbance, and bark beetles affect forest landscapes of the future. Pages 395–404 *in* Bark Beetle Management, Ecology, and Climate Change. doi:10.1016/b978-0-12-822145-7.00003-9.

Howe, M., Raffa, K.F., Aukema, B.H., Gratton, C., and Carroll, A.L. 2022. Numbers matter: how irruptive bark beetles initiate transition to self-sustaining behaviour during landscape-altering outbreaks. Oecologia **198**: 681–698. Springer. doi:10.1007/s00442-022-05129-4.

Hrinkevich, K., and Lewis, K.J. 2011. Northern range limit mountain pine beetle outbreak dynamics in mixed sub-boreal pine forests of British Columbia. Ecosphere **2**: 1–16. doi:10.1890/ES11-00150.1.

Hrinkevich, K.H. 2012. Northern range limit mountain pine beetle (*Dendroctonus ponderosae*) outbreak dynamics and climate interactions in mixed sub-boreal pine forests of British Columbia. University of Northern British Columbia.

https://catalogue.data.gov.bc.ca/ 2022. The B.C. Geographic Warehouse. [Online] Available: https://catalogue.data.gov.bc.ca/ [2022 Jul. 3].

https://droughtmonitor.unl.edu 2022. Alberta | North American Drought Monitor. [Online] Available: https://droughtmonitor.unl.edu [2022 Jun. 27].

https://opendata.nfis.org/mapserver/nfis-change_eng.html 2015. NFIS Canadian Council of Forest Ministers. [Online] Available: https://opendata.nfis.org/mapserver/nfis-change_eng.html [2021 May 7].

Hubisz, M.J., Falush, D., Stephens, M., and Pritchard, J.K. 2009. Inferring weak population structure with the assistance of sample group information. Mol. Ecol. Resour. **9**: 1322–1332. doi:10.1111/j.1755-0998.2009.02591.x.

Ibrahim, K.M., Nichols, R.A., and Hewitt, G.M. 1996. Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. Heredity (Edinb). **77**: 282–291. doi:10.1038/sj.hdy.6880320.

Jackson, P.L., Straussfogel, D., Lindgren, B.S., Mitchell, S., and Murphy, B. 2008. Radar observation and aerial capture of mountain pine beetle, *Dendroctonus ponderosae* Hopk. (Coleoptera: Scolytidae) in flight above the forest canopy. Can. J. For. Res. **38**: 2313–2327. doi:10.1139/X08-066.

Janes, J.K., and Batista, P.D. 2016. The Role of Population Genetic Structure in Understanding and Managing Pine Beetles. Pages 75–100 *in* Advances in Insect Physiology. doi:10.1016/bs.aiip.2016.01.001.

Janes, J.K., Li, Y., Keeling, C.I., Yuen, M.M.S., Boone, C.K., Cooke, J.E.K., Bohlmann, J., Huber, D.P.W., Murray, B.W., Coltman, D.W., and Sperling, F.A.H. 2014. How the mountain pine beetle (*Dendroctonus ponderosae*) breached the canadian rocky mountains. Mol. Biol. Evol. **31**: 1803–1815. doi:10.1093/molbev/msu135.

Janes, J.K., Roe, A.D., Rice, A. V., Gorrell, J.C., Coltman, D.W., Langor, D.W., and Sperling, F.A.H. 2016. Polygamy and an absence of fine-scale structure in *Dendroctonus ponderosae* (Hopk.) (Coleoptera: Curculionidae) confirmed using molecular markers. Heredity (Edinb). **116**: 68–74. Nature Publishing Group. doi:10.1038/hdy.2015.71.

Janes, J.K., Worth, J.R.P., Batista, P.D., and Sperling, F.A.H. 2018. Inferring Ancestry and Divergence Events in a Forest Pest Using Low-Density Single-Nucleotide Polymorphisms. Insect Syst. Divers. **2**. Oxford Academic. doi:10.1093/isd/ixy019.

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., and Madden, T.L. 2008. NCBI BLAST: a better web interface. Nucleic Acids Res. **36**: W5–W9. Oxford Academic. doi:10.1093/NAR/GKN201.

Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics **24**: 1403–1405. Oxford Academic. doi:10.1093/bioinformatics/btn129.

Jombart, T., Devillard, S., and Balloux, F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. **11**: 94. BioMed Central Ltd. doi:10.1186/1471-2156-11-94.

Keeling, C.I., Yuen, M.M.S., Liao, N.Y., Docking, T.R., Chan, K.S.S.K., Taylor, G.A., Palmquist, D.L., Jackman, S.D., Nguyen, A., Li, M., Henderson, H., Janes, J.K., Zhao, Y., Pandoh, P., Moore, R., Sperling, F.A.H., Huber, D.P.W., Birol, I., Jones, S.J.M., and Bohlmann, J. 2013. Draft genome of the mountain pine beetle, *Dendroctonus ponderosae* Hopkins, a major forest pest. Genome Biol. **14**: R27. doi:10.1186/gb-2013-14-3-r27.

Kirk, H., Dorn, S., and Mazzi, D. 2013a. Molecular genetics and genomics generate new insights into invertebrate pest invasions. Evol. Appl. **6**: 842–856. doi:10.1111/eva.12071.

Kirk, H., Dorn, S., Mazzi, D., Dominique Mazzi, C., and Zurich, E. 2013b. Molecular genetics and genomics generate new insights into invertebrate pest invasions. Evol. Appl. **6**: 842–856. John Wiley & Sons, Ltd. doi:10.1111/eva.12071.

Klopfstein, S., Currat, M., and Excoffier, L. 2006. The fate of mutations surfing on the wave of a range expansion. Mol. Biol. Evol. **23**: 482–490. doi:10.1093/molbev/msj057.

Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A., and Mayrose, I. 2015. Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. Mol. Ecol. Resour. **15**: 1179–1191. Blackwell Publishing Ltd. doi:10.1111/1755-0998.12387.

Koštál, V., Doležal, P., Rozsypal, J., Moravcová, M., Zahradníčková, H., and Šimek, P. 2011. Physiological and biochemical analysis of overwintering and cold tolerance in two Central European populations of the spruce bark beetle, *Ips typographus*. J. Insect Physiol. **57**: 1136–1146. doi:10.1016/j.jinsphys.2011.03.011.

Krehenwinkel, H., Rödder, D., Năpăruş-Aljančič, M., and Kuntner, M. 2016. Rapid genetic and ecological differentiation during the northern range expansion of the venomous yellow sac spider *Cheiracanthium punctorium* in Europe. Evol. Appl. **9**: 1229–1240. John Wiley & Sons, Ltd. doi:10.1111/eva.12392.

Kreppel, L.K., and Hart, G.W. 1999. Regulation of a cytosolic and nuclear O-GlcNAc transferase. Role of the tetratricopeptide repeats. J. Biol. Chem. **274**: 32015–32022. Elsevier. doi:10.1074/jbc.274.45.32015.

Lal, M.M., Southgate, P.C., Jerry, D.R., and Zenger, K.R. 2016. Fishing for divergence in a sea of connectivity: The utility of ddRADseq genotyping in a marine invertebrate, the black-lip pearl oyster *Pinctada margaritifera*. Mar. Genomics **25**: 57–68. Elsevier B.V. doi:10.1016/j.margen.2015.10.010.

Langor, D., and Spence, J. 1991. Host effects on allozyme and morphological variation of the mountain pine beetle, *Dendroctonus ponderosae* Hopkins (Coleoptera: Scolytidae). Can. Entomol. **123**: 395–410. doi:10.4039/Ent123395-2.

Langor, D.W. 1989. Host effects on the phenology, development, and mortality of field populations of the mountain pine beetle, *Dendroctonus ponderosae* Hopkins (Coleoptera: Scolytidae). Can. Entomol. **121**: 149–157. Cambridge University Press. doi:10.4039/Ent121149-2.

Lanier, G.N., and Wood, D.L. 1968. Controlled Mating, Karyology, Morphology, and Sex-Ratio in the *Dendroctonus ponderosae* Complex. Ann. Entomol. Soc. Am. **61**: 517–526. [Online] Available: http://cfs.nrcan.gc.ca/bookstore_pdfs/30216.pdf [2015 Jun. 11].

Leather, S.R., Walters, K.F.A., and Bale, J.S. 1993. The ecology of insect overwintering. The ecology of insect overwintering. Cambridge University Press, Cambridge. doi:10.2307/3495521.

Li, H., and Durbin, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics **25**: 1754–1760. Oxford Academic. doi:10.1093/bioinformatics/btp324.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. 2009a. The Sequence Alignment/Map format and SAMtools. Bioinformatics **25**: 2078–2079. Oxford Academic. doi:10.1093/bioinformatics/btp352.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Subgroup, 1000 Genome Project Data Processing 2009b. The Sequence Alignment/Map format and SAMtools. Bioinformatics **25**: 2078–2079. Oxford Academic. doi:10.1093/bioinformatics/btp352.

Liebl, A.L., Schrey, A.W., Richards, C.L., and Martin, L.B. 2013. Patterns of DNA methylation throughout a range expansion of an introduced songbird. Pages 351–358 *in* Integrative and Comparative Biology. Oxford Academic. doi:10.1093/icb/ict007.

Logan, J.A., and Amman, G.D. 1991. Temperature-dependent development of the mountain pine beetle (coleoptera: Scolytidae) and simulation of its phenology. Can. Entomol. **123**: 1083–1094. doi:10.4039/Ent1231083-5.

Logan, J.A., Bolstad, P. V, Bentz, B.J., and Perkins, D.L. 1995. Assessing the effects of

changing climate on mountain pine beetle dynamics. Pages 92–105 *in* Proceedings: Interior West Global Change Workshop, 25-27 April 1995, Fort Collins, CO. [Online] Available: http://www.usu.edu/beetle/documents2/1995Logan etal_Assessing the Effects of Changing Climate.pdf.

Logan, J.A., and Powell, J.A. 2001. Ghost Forests, Global Warming, and the Mountain Pine Beetle (Coleoptera: Scolytidae). Am. Entomol. **47**: 160–172.

MacCormick, J. 2020. Spread Management Action Collaborative. Bugs Dis. **31**: 1–10. Government of Alberta.

Maclauchlan, L.E., Daniels, L.D., Hodge, J.C., and Brooks, J.E. 2018. Characterization of western Spruce budworm outbreak regions in the British Columbia interior. Can. J. For. Res. **48**: 783–802. doi:10.1139/cjfr-2017-0278.

Malet, J. 2001. Gene flow in insect movement: mechanisms and consequences. Pages 337–360 *in* I. Woiwood, D. Reynolds, and C. Thomas, eds. Proceedings of the Royal Entomological Society's 20th Symposium, London, UK, September 1999. doi:10.1079/9780851994567.0337.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal **17**: 10. doi:10.14806/ej.17.1.200.

Mayrand, P., Filotas, É., Wittische, J., and James, P.M.A. 2019. The role of dispersal, selection, and timing of sampling on the false discovery rate of loci under selection during geographic range expansion. Genome **62**: 715–727. NLM (Medline). doi:10.1139/gen-2019-0004.

McKee, F.R., Huber, D.P.W., Lindgren, B.S., Hodgkinson, R.S., and Aukema, B.H. 2015. Effect of natal and colonised host species on female host acceptance and male joining behaviour of the mountain pine beetle (Coleoptera: Curculionidae) using pine and spruce. Can. Entomol. **147**: 39–45. doi:10.4039/tce.2014.22.

Meidinger, D., and Pojar, J. 1991. ECOSYSTEMS OF BRITISH COLUMBIA Special Report Series 6. *Edited By*D.V. Meidinger and J. Pojar B.C. Ministry of Forests. [Online] Available: papers://9ecb3f19-9827-472a-b222-ab202fcd4bac/Paper/p24169.

Meirmans, P.G. 2020. genodive version 3.0: Easy-to-use software for the analysis of genetic data of diploids and polyploids. Mol. Ecol. Resour. **20**: 1126–1131. John Wiley & Sons, Ltd. doi:10.1111/1755-0998.13145.

Miller, L.K., and Werner, R.A. 1987. Cold-hardiness of adult and larval spruce beetles *Dendroctonus rufipennis* (Kirby) in interior Alaska. Can. J. Zool. **65**: 2927–2930. doi:10.1139/z87-444.

Mitton, J.B., and Sturgeon, K.B. 1982. Bark beetles in North American conifers: a system for the study of evolutionary biology. University of Texas Press. [Online] Available: https://utpress.utexas.edu/books/mitbar [2018 Jun. 13].

Mock, K.E., Bentz, B.J., O'Neill, E.M., Chong, J.P., Orwin, J., and Pfrender, M.E. 2007. Landscape-scale genetic variation in a forest outbreak species, the mountain pine

beetle (*Dendroctonus ponderosae*). Mol. Ecol. **16**: 553–568. doi:10.1111/j.1365-294X.2006.03158.x.

Mona, S., Ray, N., Arenas, M., and Excoffier, L. 2014. Genetic consequences of habitat fragmentation during a range expansion. Heredity (Edinb). **112**: 291–299. Nature Publishing Group. doi:10.1038/hdy.2013.105.

Moran, E. V., and Alexander, J.M. 2014. Evolutionary responses to global change: lessons from invasive species. Ecol. Lett. **17**: 637–649. John Wiley & Sons, Ltd. doi:10.1111/ELE.12262.

Nkemdirim, L.C. 1986. Chinooks in Southern Alberta: Some distinguishing nocturnal features. J. Climatol. **6**: 593–603. John Wiley & Sons, Ltd. doi:10.1002/joc.3370060603.

Nussey, D.H., Coltman, D.W., Coulson, T., Kruuk, L.E.B., Donald, A., Morris, S.J., Clutton-Brock, T.H., and Pemberton, J. 2005. Rapidly declining fine-scale spatial genetic structure in female red deer. Mol. Ecol. **14**: 3395–3405. John Wiley & Sons, Ltd. doi:10.1111/j.1365-294X.2005.02692.x.

Ojeda Alayon, D.I., Tsui, C.K.M., Feau, N., Capron, A., Dhillon, B., Zhang, Y., Massoumi Alamouti, S., Boone, C.K., Carroll, A.L., Cooke, J.E.K., Roe, A.D., Sperling, F.A.H., and Hamelin, R.C. 2017. Genetic and genomic evidence of niche partitioning and adaptive radiation in mountain pine beetle fungal symbionts. Mol. Ecol. **26**: 2077–2091. doi:10.1111/mec.14074.

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R., Simpson, G., Solymos, P., Stevens, M., Szoecs, E., and Wagner, H. 2020. vegan: Community Ecology Package. R package version 2.5-6. 2019. Community Ecol. **8**: 732–740. [Online] Available: https://cran.r-project.org/package=vegan.

Paris, J.R., Stevens, J.R., and Catchen, J.M. 2017. Lost in parameter space: a road map for stacks. Methods Ecol. Evol. **8**: 1360–1373. doi:10.1111/2041-210X.12775.

Parmesan, C. 2006. Ecological and evolutionary responses to recent climate change. Annu. Rev. Ecol. Evol. Syst. **37**: 637–669. doi:10.1146/annurev.ecolsys.37.091305.110100.

Parmesan, C., and Yohe, G. 2003. A globally coherent fingerprint of climate change impacts across natural systems. Nature **421**: 37–42. Nature Publishing Group. doi:10.1038/nature01286.

Pélissié, B., Crossley, M.S., Cohen, Z.P., and Schoville, S.D. 2018. Rapid evolution in insect pests: the importance of space and time in population genomics studies. Curr. Opin. Insect Sci. **26**: 8–16. Elsevier. doi:10.1016/J.COIS.2017.12.008.

Pemberton, T.A., Still, B.R., Christensen, E.M., Singh, H., Srivastava, D., and Tanner, J.J. 2012. Proline: Mother Natures cryoprotectant applied to protein crystallography. Acta Crystallogr. Sect. D Biol. Crystallogr. **68**: 1010–1018. doi:10.1107/S0907444912019580.

Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., and 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. Public Library of Science. doi:10.1371/journal.pone.0037135.

Pitt, C., Robert, J.A., Bonnett, T.R., Keeling, C.I., Bohlmann, J., and Huber, D.P.W. 2014. Proteomics indicators of the rapidly shifting physiology from whole mountain pine beetle, *Dendroctonus ponderosae* (coleoptera: Curculionidae), adults during early host colonization. PLoS One **9**: e110673. doi:10.1371/journal.pone.0110673.

Plattner, A. B. L. 2008. Pathogenicity and taxonomy of fungi associated with the mountain pine beetle in British Columbia. University of British Columbia.

Pokorny, S.W. 2021. Novel biological interactions influence the persistence potential of invasive mountain pine beetle (*Dendroctonus ponderosae*) in the Canadian Boreal Forest. University of British Columbia.

Powell, J.M. 1961. The mountain pine beetle, Dendroctonus monticolae Hopk. in Western Canada. Canada Department of Forestry, Forest Entomology and Pathology Laboratory,. Edmonton, Alberta.

Pritchard, J.K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genetics **155**: 945–959. Genetics Society of America. [Online] Available: /pmc/articles/PMC1461096/?report=abstract [2021 Jan. 24].

R Core Team 2018. R: A language and environment for statistical computing. Foundation for Statistical Computing, Vienna, Austria.

Raffa, K.F., Aukema, B.H., Bentz, B.J., Carroll, A.L., Hicke, J.A., Turner, M.G., and Romme, W.H. 2008. Cross-scale drivers of natural disturbances prone to anthropogenic amplification: The dynamics of bark beetle eruptions. Bioscience **58**: 501–517. doi:10.1641/B580607.

Raffa KF, Aukema BH, Bentz BJ, Carroll AL, Hicke JA & Kolb TE. 2015. Responses of tree-killing bark beetles to a changing climate. In Bjorkman C & Niemela P. Climate Change and Insect Pests, CABI, Wallingfored England. Pp. 173-201.

Raffa, K.F., and Berryman, A.A. 1983. The Role of Host Plant Resistance in the Colonization Behaviour and Ecology of Bark Beetles (Coleoptera: Scolytidae). Ecol. Monogr. **53**: 27–49. John Wiley & Sons, Ltd. doi:10.2307/1942586.

Raffa, K.F., Mason, C.J., Bonello, P., Cook, S., Erbilgin, N., Keefover-Ring, K., Klutsch, J.G., Villari, C., and Townsend, P.A. 2017. Defence syndromes in lodgepole – whitebark pine ecosystems relate to degree of historical exposure to mountain pine beetles. Plant Cell Environ. **40**: 1791–1806. doi:10.1111/pce.12985.

Räsänen, K., and Hendry, A.P. 2008.June. Disentangling interactions between adaptive divergence and gene flow when ecology drives diversification. doi:10.1111/j.1461-0248.2008.01176.x.

Régnière, J., and Bentz, B. 2007. Modeling cold tolerance in the mountain pine beetle,

Dendroctonus ponderosae. J. Insect Physiol. **53**: 559–572. doi:10.1016/j.jinsphys.2007.02.007.

Reid, R.W. 1962. Biology of the Mountain Pine Beetle, *Dendroctonus monticolae* Hopkins, in the East Kootenay Region of British Columbia I. Life Cycle, Brood Development, and Flight Periods1. Can. Entomol. **94**: 531–538. doi:10.4039/Ent94531-5.

Rellstab, C., Gugerli, F., Eckert, A.J., Hancock, A.M., and Holderegger, R. 2015. A practical guide to environmental association analysis in landscape genomics. Mol. Ecol. **24**: 4348–4370. John Wiley & Sons, Ltd. doi:10.1111/MEC.13322.

Revelle, W. 2010. psych: Procedures for personality and psychological research [Computer software manual]. [Online] Available: http://www.personality-project.org/r/psych-manual.pdf [2022 May 19].

Rhemtulla, J.M., Hall, R.J., Higgs, E.S., and Macdonald, S.E. 2002. Eighty years of change: Vegetation in the montane ecoregion of Jasper National Park, Alberta, Canada. Can. J. For. Res. **32**: 2010–2021. NRC Research Press Ottawa, Canada. doi:10.1139/x02-112.

Robert, J.A., Bonnett, T., Pitt, C., Spooner, L.J., Fraser, J., Yuen, M.M.S., Keeling, C.I., Bohlmann, J., and Huber, D.P.W. 2016. Gene expression analysis of overwintering mountain pine beetle larvae suggests multiple systems involved in overwintering stress, cold hardiness, and preparation for spring development. PeerJ **2016**: e2109. doi:10.7717/peerj.2109.

Roberts, D.R., and Hamann, A. 2015. Glacial refugia and modern genetic diversity of 22 western North American tree species. Proc. R. Soc. B Biol. Sci. **282**: 0142903. doi:10.1098/rspb.2014.2903.

Rochette, N.C., Rivera-Colón, A.G., and Catchen, J.M. 2019. Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. Mol. Ecol. **28**: 4737–4754. Blackwell Publishing Ltd. doi:10.1111/mec.15253.

Safranyik, L., and Carroll, A.L. 2006. The biology and epidemiology of the mountain pine beetle in lodgepole pine forests. Pages 3–66 *in* The Mountain Pine Beetle – A Synthesis of Biology, Management, and Impacts in Lodgepole Pine. Canadian Forest Service. [Online] Available: http://www.ncbi.nlm.nih.gov/pubmed/22233385.

Safranyik, L., and Linton, D.A. 1998. Mortality of mountain pine beetle larvae, *Dendroctonus ponderosae* (Coleoptera: Scolytidae) in logs of lodgepole pine (*Pinus contorta var. latifolia*) at constant low temperatures. J. Entomol. Soc. Br. Columbia **95**: 81–87. [Online] Available: http://journal.entsocbc.ca/index.php/journal/article/view/476%5Cnpapers2://publication/uuid/E0C16AA1-9D5B-4B97-AE85-5F6621569B4C.

Safranyik, L., and Wilson, B. (eds.) 2007. The mountain pine beetle: a synthesis of biology, management and impacts on lodgepole pine. Canadian Forest Service. [Online] Available: http://www.cabdirect.org/abstracts/20073257193.html [2015 Jan.

5].

Sambaraju, K.R., Carroll, A.L., and Aukema, B.H. 2019. Multiyear weather anomalies associated with range shifts by the mountain pine beetle preceding large epidemics. For. Ecol. Manage. **438**: 86–95. Elsevier. doi:10.1016/J.FORECO.2019.02.011.

Sambaraju, K.R., and Goodsman, D.W. 2021. Mountain pine beetle: an example of a climate-driven eruptive insect impacting conifer forest ecosystems. CAB Rev. **16**: 1–18. doi:10.1079/PAVSNNR202116018.

Sandrock, C., Razmjou, J., and Vorburger, C. 2011. Climate effects on life cycle variation and population genetic architecture of the black bean aphid, Aphis fabae. Mol. Ecol. **20**: 4165–4181. John Wiley & Sons, Ltd. doi:10.1111/J.1365-294X.2011.05242.X.

Seidl, R., Donato, D.C., Raffa, K.F., and Turner, M.G. 2016a. Spatial variability in tree regeneration after wildfire delays and dampens future bark beetle outbreaks. Proc. Natl. Acad. Sci. U. S. A. **113**: 13075–13080. doi:10.1073/pnas.1615263113.

Sgrò, C.M., Terblanche, J.S., and Hoffmann, A.A. 2016. What Can Plasticity Contribute to Insect Responses to Climate Change? Annu. Rev. Entomol. **61**: 433–451. doi:10.1146/annurev-ento-010715-023859.

Shegelski, V.A., Campbell, E.O., Thompson, K.M., Whitehouse, C.M., and Sperling, F.A.H. 2021. Source and spread dynamics of mountain pine beetle in central Alberta, Canada. Can. Entomol. **153**: 314–326. doi:10.4039/tce.2020.83.

Shore, T.L., Safranyik, L., Hawkes, B.C., and Taylor, S.W. 2006. Effects of the mountain pine beetle on lodgepole pine stand structure and dynamics. *Edited By*L. Safranyik and B. Wilson The mountain pine beetle: A synthesis of biology, management, and impacts on lodgepole pine. Victoria, BC. [Online] Available: http://www.cfs.nrcan.gc.ca/pubwarehouse/pdfs/26116.pdf#page=106.

Short, K.H., and Petren, K. 2011. Fine-scale genetic structure arises during range expansion of an invasive gecko. PLoS One **6**: e26258. Public Library of Science. doi:10.1371/journal.pone.0026258.

Six, D.L., and Bracewell, R.R. 2015. Dendroctonus. Pages 305–350 *in* F.E. Vega and R.W. Hofstetter, eds. Bark Beetles: Biology and Ecology of Native and Invasive Species, 1st edition. Academic Press, London.

Stock, M.W., and Guenther, J.D. 1979. Isozyme Variation Among Mountain Pine Beetle (*Dendroctonus ponderosae*) Populations in the Pacific Northwest. Environ. Entomol. **8**: 889–893. Oxford University Press. doi:10.1093/ee/8.5.889.

Storey, J.M., and Storey, K.B. 1983. Regulation of cryoprotectant metabolism in the overwintering gall fly larva, Eurosta solidaginis: Temperature control of glycerol and sorbitol levels. J. Comp. Physiol. ■ B **149**: 495–502. doi:10.1007/BF00690008.

Storey, K., and Storey, J. 2012. Insect cold hardiness: metabolic, gene, and protein adaptation. Can. J. Zool. **90**: 456–475. doi:10.1139/Z2012-011.

Tan, X., Chen, S., Gan, T.Y., Liu, B., and Chen, X. 2019. Dynamic and thermodynamic changes conducive to the increased occurrence of extreme spring fire weather over western Canada under possible anthropogenic climate change. Agric. For. Meteorol. **265**: 269–279. Elsevier. doi:10.1016/j.agrformet.2018.11.026.

Teulier, L., Weber, J.M., Crevier, J., and Darveau, C.A. 2016. Proline as a fuel for insect flight: Enhancing carbohydrate oxidation in hymenopterans. Proc. R. Soc. B Biol. Sci. **283**: 20160333. doi:10.1098/rspb.2016.0333.

Thompson, S.N. 2003. Trehalose - The Insect "Blood" Sugar. Adv. In Insect Phys. **31**: 205–285. doi:10.1016/S0065-2806(03)31004-5.

Trevoy, S.A.L., Janes, J.K., Muirhead, K., and Sperling, F.A.H. 2019. Repurposing population genetics data to discern genomic architecture: A case study of linkage cohort detection in mountain pine beetle (*Dendroctonus ponderosae*). Ecol. Evol. **9**: 1147–1159. doi:10.1002/ece3.4803.

Trevoy, S.A.L., Janes, J.K., and Sperling, F.A.H. 2018. Where did mountain pine beetle populations in Jasper Park come from? Tracking beetles with genetics. For. Chron. **94**: 20–24. doi:10.5558/tfc2018-004.

Turner, N.J., and Clifton, H. 2009. "It's so different today": Climate change and indigenous lifeways in British Columbia, Canada. Glob. Environ. Chang. **19**: 180–190. doi:10.1016/j.gloenvcha.2009.01.005.

Vucetich, J.A., and Waite, T.A. 2003. Spatial patterns of demography and genetic processes across the species' range: Null hypotheses for landscape conservation genetics. Conserv. Genet. **4**: 639–645. doi:10.1023/A:1025671831349.

Wang, J., Zhang, R.R., Gao, G.Q., Ma, M.Y., and Chen, H. 2016a. Cold tolerance and silencing of three cold-tolerance genes of overwintering Chinese white pine larvae. Sci. Rep. **6**: 34698. doi:10.1038/srep34698.

Wang, T., Hamann, A., Spittlehouse, D., and Carroll, C. 2016b. Locally downscaled and spatially customizable climate data for historical and future periods for North America. PLoS One **11**: e0156720. Public Library of Science. doi:10.1371/journal.pone.0156720.

Weber, B. 2017.November 17. Mountain pine beetles from Jasper National Park moving east into commercial forest. Glob. News. [Online] Available: https://globalnews.ca/news/3866541/mountain-pine-beetle-jasper-national-park-hinton/ [2022 Jul. 6].

Wertman, D.L., Bleiker, K.P., and Perlman, S.J. 2018. The light at the end of the tunnel: Photosensitivity in larvae of the mountain pine beetle (Coleoptera: Curculionidae: Scolytinae). Can. Entomol. **150**: 622–631. doi:10.4039/tce.2018.38.

Whitehead, R., Martin, P., and Powelson, A. 2001. Reducing stand and landscape susceptibility to mountain pine beetle. Natural Resources Canada, Victoria, BC. [Online] Available: http://www.cfs.nrcan.gc.ca/publications/?id=18316.

Whitfield, P.H., and Cannon, A.J. 2000. Recent Variations in Climate and Hydrology in Canada. Can. Water Resour. J. **25**: 19–65. doi:10.4296/cwrj2501019.

Whitlock, M.C., and Lotterhos, K.E. 2015. Reliable detection of loci responsible for local adaptation: Inference of a null model through trimming the distribution of FST. Am. Nat. **186**: S24–S36. University of Chicago Press. doi:10.1086/682949.

Wiens, J.J. 2016. Climate-Related Local Extinctions Are Already Widespread among Plant and Animal Species. PLoS Biol. **14**: e2001104. Public Library of Science. doi:10.1371/journal.pbio.2001104.

Wood, C.S., and Unger, L. 1996. Mountain Pine Beetle-A history of outbreaks in pine forests in British Columbia 1910 to 1995. Natural Resources Canada, Victoria, BC. doi:10.1016/0378-1127(95)90011-X.

www.ipcc.ch (n.d.). The Intergovernmental Panel on Climate Change. [Online] Available: www.ipcc.ch [2022 Jan. 7].

Xuereb, A., Kimber, C.M., Curtis, J.M.R., Bernatchez, L., and Fortin, M.J. 2018. Putatively adaptive genetic variation in the giant California sea cucumber (*Parastichopus californicus*) as revealed by environmental association analysis of restriction-site associated DNA sequencing data. Mol. Ecol. **27**: 5035–5048. John Wiley & Sons, Ltd. doi:10.1111/mec.14942.

Yadav, S., Stow, A.J., and Dudaniec, R.Y. 2019. Detection of environmental and morphological adaptation despite high landscape genetic connectivity in a pest grasshopper (*Phaulacridium vittatum*). Mol. Ecol. **00**: 1–18. doi:10.1111/mec.15146.

Young, S.L., Clements, D.R., and DiTommaso, A. 2017. Climate Dynamics, Invader Fitness, and Ecosystem Resistance in an Invasion-Factor Framework. Invasive Plant Sci. Manag. **10**: 215–231. doi:10.1017/inp.2017.28.

Zhang, L., Wang, F., Qiao, L., Dietrich, C.H., Matsumura, M., and Qin, D. 2019. Population structure and genetic differentiation of tea green leafhopper, *Empoasca (Matsumurasca) onukii*, in China based on microsatellite markers. Sci. Rep. **9**: 1–10. doi:10.1038/s41598-018-37881-0.

Zong, Y., Zhang, B., Gu, S., Lee, K., Zhou, J., Yao, G., Figueiredo, D., Perry, K., Mei, L., and Jin, R. 2012. Structural basis of agrin-LRP4-MuSK signaling. Genes Dev. **26**: 247–258. Cold Spring Harbor Laboratory Press. doi:10.1101/gad.180885.111.

“There is nothing like looking, if you want to find something. You certainly usually find something, if you look, but it is not always quite the something you were after.”

J. R. R. Tolkien, *The Hobbit*

“And the first lesson of all was the basic trust that he could learn.”

Frank Herbert, *Dune*

7. Appendix 1: Code Used for SNP Generation and Data Analysis

7.1. DD-RAD Workflow

############## WORKFLOW FOR MPB DDRAD DATA ###############

DATA PROCESSING AND GENOTYPING IN STACKS

####

Step one: unzip and concatenate the 4 files for each N7XX index (usually labelled as L001-L004):

gzip -d * ## do this for each of the 4 .fastq files for each index, then:

cat * > N7XX_lanes1-4.fastq ## can also specify files individually separated by a space if in a common directory.

####

Step two: Demultiplex the data using process_radtags in Stacks, which is installed on Compute Canada's Cedar (and Graham) cluster:

8 bp barcodes and the PstI cut site are both on the 5'-end, need to remove the barcodes (see below), making total

read length 67 bp

command for running process_radtags:

module load nixpkgs/16.09 gcc/5.4.0 stacks/2.0b ## loads stacks on Cedar

process_radtags -f /home/kikit18/scratch/time_seq/N701_lanes1-4.fastq \

-b /home/kikit18/scratch/barcodes_time/mpb_barcodes_name701.txt --renz_1 pstI \

--inline_null -t 67 -w 0.15 -s 20 -c -r -D --filter_illumina -E phred33 \

-o /home/kikit18/scratch/demulti_time

####

Step three: Search for remnant Illumina adaptor sequences and remove affected reads in cutadapt using loop:

make sure to remove PstI site on 5'-end (-u 5), as there is sometimes sequencing error in the cut site, causing false SNPs (-u 5 is marked in the code below)

module load python

```
virtualenv ENV

source ENV/bin/activate

pip install --upgrade cutadapt


for fname in /home/kikit18/scratch/demulti_2/*.fq;


do


/scratch/kikit18/ENV/bin/cutadapt -u 5 -a
ACCGAGATCGGAAGAGCACACGTCTGAACTCCAGTCACNNNNNNNNATC -m 62
-o "${fname%.fq}.fastq" "$fname" &


done


deactivate



## make sure to remove PstI site on 5'-end (-u 5), as there is sometimes sequencing
error in the cut site, causing false SNPs
```

####

Step four: Align cleaned reads to MPB reference genome using bwa, genbank accession number: :

First, create reference index called "MPB_male" using male MPB genome:

bwa index -p MPB_male -a bwtsw /home/kikit18/scratch/MPB_genome/GCA_000355655.1_DendPond_male_1.0_genomic.fa

Then, use mem to align reads to the reference and output .sam files (must do this out of the genome directory):

bwa mem pigeon 696W2.fastq > 696W2.sam

or you can use a loop to align all files to reference with a single command:

updated June 2019, this now works as a for loop. Trying out the letter tag to tell between male/female (eg: filem for male and filef for female)

```
module load bwa

for fname in /home/kikit18/scratch/demulti_time/ROCK/*.fastq;

do

bwa mem /home/kikit18/project/kikit18/male_genome/MPB_male $fname >
"${fname%.*}"m.sam;

done
```

Step five: Quality check the bwa alignment using samtools or Picard (both on Cedar):

```
module load samtools/1.9

## to see a list of options type:
```

samtools help

to see how many reads aligned in each file (optional, but should check in case you need to tweak bwa):

samtools flagstat sample_name.sam

this should output something like this:

473319 + 0 in total (QC-passed reads + QC-failed reads)

0 + 0 secondary

121 + 0 supplementary

0 + 0 duplicates

431987 + 0 mapped (91.27% : N/A)

0 + 0 paired in sequencing

0 + 0 read1

0 + 0 read2

0 + 0 properly paired (N/A : N/A)

0 + 0 with itself and mate mapped

0 + 0 singletons (N/A : N/A)

```
## 0 + 0 with mate mapped to a different chr

## 0 + 0 with mate mapped to a different chr (mapQ>=5)


##KMT created for loop for this


module load samtools/1.9


for fname in /home/kikit18/scratch/demulti_time/ROCK/ROCK_sam/*.sam;


do


samtools flagstat $fname;


echo $fname ;


done


#### usually around 90% of reads are mapping to the reference genome, this is good.
```

```
##MPB maps very badly, so the threshold can be set lower (80%)


## convert samfiles to bam format and sort:

## code from Victor did not rename files properly. This version with basename corrects the problem, so files read xxx.bam instead of xxx.sam.bam

##MPB maps very badly, so the threshold can be set lower (80%)


module load samtools/1.9


for fname in ./*.sam;


do samtools view -bS $fname |

samtools sort > /home/kikit18/scratch/sorted_time/`basename $fname .sam`.bam;


done




###### Step 6: run the ref_map.pl script in Stacks:
```

```
module load nixpkgs/16.09 gcc/5.4.0 stacks/2.0b


ref_map.pl --samples /home/kikit18/scratch/sorted \
--popmap /home/kikit18/project/kikit18/popmaps/popmap_all_sam.txt \
-o /home/kikit18/scratch/refmap_redo \
-X "gstacks:--max-clipped 0" \
-X "populations:-p 1 -r 0.8 --min_maf 0.05 --write_random_snp --fstats --hwe --
vcf"\



### test different values of maf and -r to see how many SNPs we can get, then use
vcftools to filter by
### KMT NOTES: -r is usually set at .8 because it is more stringent.Could
potentially go down 0.5 if SNP numbers are poor (r80 means 80% of the population has
that SNP)
### GQ (genotype quality). Use a threshold of 30 (phred score of 30 = 0.01%
chance of error in genotype call)


module load vcftools/0.1.14
```

vcftools --vcf ./allmale_raw.vcf --minGQ 30 --recode --out ./allmale_gq30

NOTE that after using the --minGQ flag, genotypes that fail this filter

are not removed from the vcf file, but the genotype changes to ./. and

so isn't included in the final dataset - it's treated as missing

Remove individuals that sequenced poorly as well as duplicate seqs:

vcftools --vcf pigeon_gq30.recode.vcf --remove-indv 696W2 --remove-indv 907W2 --remove-indv 438W2 --remove-indv 1946V --remove-indv 1126W --remove-indv 14W --remove-indv 560W2 --remove-indv 852W --remove-indv 2E --remove-indv 10E --remove-indv 14S2 --remove-indv 1S --remove-indv 275KE --remove-indv 300KE --remove-indv XXXXX --remove-indv 2McDV --recode --out ./pigeon_gq30_good

after filtering by GQ and removing individuals, filter again for missing data to remove loci with poor coverage:

vcftools --vcf ./pigeon_gq30.recode.vcf --max-missing 0.95 --recode --out ./pigeon_gq30_max5

After this you should remove loci out of HWE and/or check for linkage, then data is good for popgen analyses

Victor has provided some code in dartR that removes linkage. dartR also can check for loci out of HWE and produce a list for removal in vcftools

7.2. VCFtools Codes

vcftools --vcf ./aug_f.vcf --minGQ 30 --recode --out ./aug_f_gq30

vcftools --vcf "time_raw.vcf" --max-missing 0.8 --recode --recode-INFO-all --out time_mm80

vcftools --vcf "time_mm80_maf05.recode.vcf" --maf 0.05 --recode --recode-INFO-all --out time_mm80_maf05

vcftools --vcf "time_mm80_maf05.recode.vcf" --minDP 10 --recode --recode-INFO-all --out time_mm80_maf05_dp10

vcftools --vcf time_mm80_maf05_dp10.recode.vcf --hardy --max-missing 0.8 --out output_mm80HWE

vcftools --vcf time_mm80_maf05_dp10.recode.vcf --site-mean-depth --out time_depth_site

vcftools --vcf time_mm80_maf05_dp10.recode.vcf --site-quality --out time_qual

vcftools --vcf time_mm80_maf05_dp10.recode.vcf --het --out time_het

vcftools --vcf time_mm80_maf05_dp10.recode.vcf --relatedness --out time_relate

vcftools --vcf time_mm80_maf05_dp10.recode.vcf --depth --out time_ind_depth

must remove individuals one at a time...

vcftools --vcf "aug_f_gq30_mm80_maf05_minDP10.recode.vcf" --remove-indv K1Abf --recode --recode-INFO-all --out aug_f_nd

##Remove loci by position

vcftools --vcf aug_f_gq30_mm95_maf05_minDP10_ND.recode.vcf --exclude-positions all_95_sex_snps_chrom_pos.txt --recode --recode-INFO-all --out aug_f_sexed

##Thin loci that are too close to each other (eg: 10 000 bp)

vcftools --vcf aug_f_sexed_LD.recode.vcf --thin 10000 --recode --recode-INFO-all

--out aug_f_sexed_LD_verythinned

7.3. STRUCTURE Code for GRAHAM cluster

##Basic commands for running structure in command line on Graham

##Navigate to your directory with your .str file and your mainparams and extraparams file (easy to make in the structure GUI)

##Convert the file in unix (removes DOS line breaks)

dos2unix time2005_pop_num.str

##This commmand runs structure locally (when you are in the folder with the params files and your input file)

module load structure

structure -K 1 -o k1run.txt

##This more detailed command will run structure from anywhere as long as you have the params files and the input files

structure -m /home/kikit18/scratch/time_structure/mainparams -e /home/kikit18/scratch/time_structure/extraparams -i

/home/kikit18/scratch/time_structure/time2005_pop_num.str -K 1 -o

/home/kikit18/scratch/time_structure/k1run2

7.4. PGDSpider Code for GRAHAM Cluster

```
##Code to run PGDSpider on Graham


##Load Java
module load nixpkgs/16.09
module load java/1.7.0_80


## Loads PGDSpider GUI on screen


java -Xmx1024m -Xms512m -jar
/home/kikit18/project/kikit18/pgdspider/src/PGDSpider_2.1.1.3/PGDSpider2.jar
```

7.5. Use of R for File Conversions

```
##Convert genlight to structure


gl2structure(time2016LD.HWE.gl, outfile = "time2016LD.HWE.str")


##Export pop info
write.csv(as.data.frame(range_clean.gl$pop),file="range_clean_pops.csv")
```

```
####Working with vcf files####

##set your working directory
setwd("C:/Users/kirst/Documents/Incoming Popgen/")

## call packages vcfR and dartR
library(vcfR)
library(dartR)
library(adegenet)

## read in the vcf file
all_vcf <-read.vcfR("aug_f_gq30_mm80_maf05_minDP10.recode.vcf")

## view file
all_vcf

## create a genelight file from the vcf file
all_gl <- vcfR2genlight(all_vcf)

## create a genind file from the vcf file. Both genelight and genind files
## do not have population data as VCF has no space for this
## genelight files allow you to retain the names of the individuals you are working
with
```

```
all_gi <- vcfR2genind(all_vcf)

## export list of individuals for generation of populations
write.table(all_gl@ind.names, "all.txt")

## add population data for each individual. This seems to work well n < 300 indivduals
pop(all_gl) <- as.factor(c("E", "E", "E", "E", "E", "E", "E", "E", "E", "E", "E", "E",
"E", "E", "E", "E", "E", "E", "F", "F", "F", "F", "F", "F", "F", "F", "F", "F", "F", "F", "F",
"F", "F", "F", "F", "F", "G", "G", "G", "G", "G", "G", "G", "G", "G", "G", "G", "G",
"G", "G", "G", "G", "G", "G", "G", "H", "H", "H", "H", "H", "H", "H", "H", "H", "H", "H",
"H", "H", "H", "H", "H", "I", "I", "I", "I", "I", "I", "I", "I", "I", "I", "I", "I", "I", "I", "I", "I",
"I", "I", "I", "I", "I", "J", "J", "J", "J", "J", "J", "J", "J", "J", "J", "J", "J", "J", "J", "J", "J",
"J", "J", "J", "J", "JA", "JA", "JA", "JA", "JA", "JA", "JA", "JA", "JA", "JA", "JA",
"JA", "JA", "JA", "JA", "JA", "JA", "JA", "JB", "JB", "JB", "JB", "JB", "JB", "JB",
"JB", "JB", "JB", "JB", "JB", "JB", "JB", "JB", "JB", "JB", "JB", "JB", "JC", "JC",
"JC", "JC", "JC", "JC", "JC", "JC", "JC", "JC", "JC", "JC", "JC", "JC", "JC", "JC",
"JD", "JD", "JD", "JD", "JD", "JD", "JD", "JD", "JD", "JD", "JD", "JD", "JD", "JD",
"JD", "JD", "JD", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE",
"JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE",
"JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "K", "K", "K", "K",
"K", "K", "K", "K", "K", "K", "K", "K", "K", "K", "K", "K", "K", "K", "K"))
```

```
## Create seperate lists using c() as there are numeric limits (over 499?)
part1 <- c("HV", "HV", "HV", "HV", "HV", "HV", "HV", "HV", "HV", "HV",
"HV", "HV", "HV", "HV", "HV", "HV", "HV", "HV", "HV", "HV", "HV", "HV",
"HV", "HV", "A", "A", "A", "A", "A", "A", "A", "A", "A", "A", "A", "A", "A", "A",
"A", "A", "A", "B", "B", "B", "B", "B", "B", "B", "B", "B", "B", "B", "B", "B",
"B", "B", "B", "B", "B", "B", "B", "B", "B", "B", "B", "B", "B", "B", "B", "C", "C",
"C", "C", "C", "C", "C", "C", "C", "C", "C", "C", "C", "C", "C", "C", "C", "C", "C",
"C", "C", "D", "D", "D", "D", "D", "D", "D", "D", "D", "D", "D", "D", "D", "D", "D",
"D", "D", "E", "E", "E", "E", "E", "E", "E", "E", "E", "E", "E", "E", "E", "E", "E",
"E", "E", "F", "F", "F", "F", "F", "F", "F", "F", "F", "F", "F", "F", "F", "F", "F",
"F", "G", "G", "G", "G", "G", "G", "G", "G", "G", "G", "G", "G", "G", "G", "G", "G",
"G", "G", "H", "H", "H", "H", "H", "H", "H", "H", "H", "H", "H", "H", "H", "H", "H",
"H", "I", "I", "I", "I", "I", "I", "I", "I", "I", "I", "I", "I", "I", "I", "I", "I",
"J", "J", "J", "J", "J", "J", "J", "J", "J", "J", "J", "J", "J", "J", "J", "J", "J",
"JA", "JA", "JA", "JA", "JA", "JA", "JA", "JA", "JA", "JA", "JA", "JA", "JA", "JA",
"JA", "JA", "JA", "JA", "JB", "JB", "JB", "JB", "JB", "JB", "JB", "JB", "JB", "JB",
"JB", "JB", "JB", "JB", "JB", "JB", "JB", "JB", "JC", "JC", "JC", "JC", "JC", "JC",
"JC", "JC", "JC", "JC", "JC", "JC", "JC", "JC", "JC", "JD", "JD", "JD", "JD",
"JD", "JD", "JD", "JD", "JD", "JD", "JD", "JD", "JD", "JD", "JE",
"JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE",
"JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE",
"JE", "JE", "JE", "JE", "JE", "JE", "JE", "JE", "K", "K", "K", "K", "K", "K", "K", "K",
"K", "K", "K", "K", "K", "K", "K", "K", "K", "L", "L", "L", "L", "L", "L", "L",
```

```
"L", "L", "L", "L", "L", "L", "L", "L", "L", "L", "L", "L", "L", "L", "L", "L", "L", "L", "L",
"L", "L", "L", "L", "L", "L", "L", "M", "M", "M", "M", "M", "M", "M", "M", "M", "M",
"M", "M", "M", "M", "M", "M", "M", "M", "M", "M", "M", "M", "M", "M", "M", "M",
"M", "M")

    ## Merge the lists using c()
    test1 <- c(part1, part2, part3)

    ##Check that it worked!
    head(test1)
    tail(test1)

    ##Check what populations you have
    popNames(all_gl)

    ##Remove pops as needed
    all80_sexed.gl <- gl.drop.pop(all80_sexed_gl, pop.list=c("B","xD"))

    ##add in gps coords if you have them
    jasper_sexed_gl@other$latlong <- jas_gps

    ##make an object with a list of individuals to remove
    jasperRM <-c("F3Dbf", "F3Dcf", "I1Dbf", "I1Dcf", "J3Cbf", "J3Ccf", "K1Abf",
```

"K1Acf")

use dartR to remove dups

library(dartR)

jasper_gl_nd <- gl.drop.ind(jasper_gl, jasperRM, recalc = TRUE, mono.rm = TRUE, v = 2)

##alternate: export the list of dupes and remove them from the vcf file, then start over

##make an object with a list of individuals to remove

##a good cutoff is 20% or more missing data per ind

ind20pMD <-c("F3Dbf", "F3Dcf", "I1Dbf", "I1Dcf", "J3Cbf", "J3Ccf", "K1Abf", "K1Acf")

##Remove the poor quality ind

range_clean <- gl.drop.ind(range_HWE.gl, range20pMD, recalc = TRUE, mono.rm = TRUE, v = 2)

##Convert edited genelight to genind (if desired) and save the file!

all_gi2 <- gl2gi(all_gl_nd, v = 1)

save(all_gi2, file="all_gi2.rdata")

7.6. Basis Statistics in R

```
## Basic stats, works, looks at per locus, not pop
library(dartR)
all80stats <- gl.basic.stats(all80_sexed_gl_BxD, digits = 4)
save(all80stats, file ="all80stats.RData")

## Diversity metrics, works
all80diversity <- gl.diversity(all80_sexed_gl_BxD, spectrumplot = TRUE,
confiplot = FALSE,
probar = TRUE, table = "DH")
##save outputs
save(all80stats, file ="all80stats.RData")

#Calculates HWE by pop.Works.
library(parallel)
all80HWE_pop <- gl.hwe.pop(all80_sexed_gi_BxD, pvalue = 0.05, plot = TRUE)
##Loci that deviate in the majority of populations can be identified via colSums on
the resulting matrix
HWElist1 <- colSums(all80HWE_pop)
write.table(HWElist1, "HWElist1.txt", sep="\t")
```

```
##Calculate Ho per loci

all80_Ho <- gl.Ho(all80_sexed_gl_BxD)

write.table(all80_Ho, "all80_Ho.txt", sep="\t")


##Calculate He per loci

time_He <- gl.Hs(time_sexed_gl)

write.table(time_He, "time_He.txt", sep="\t")


##Heat map test

d <- dist(as.matrix((time_noNAs)))

time_heatmap <- gl.dist.heatmap(d)

time_heatmap_col <- gl.dist.heatmap(d, ncolors=10, rank=TRUE)

save(time_heatmap_col, file = "time_heatmap_col.RData")



##Produce a table wiht HWE values for the whole dataset

timeLD_HWE <- gl.report.hwe(timeLD_gl)

write.table(timeLD_HWE, "timeLD_HWE.txt", sep="\t")


##Filter for HWE (Bonferroni doesn't filter)

time_HWE_list <- gl.filter.hwe(time_sexed_gl, 0.05, bon=TRUE)

time_HWE_nonBON <- gl.filter.hwe(time_sexed_gl, 0.05, bon=FALSE)

time_HWE_nonBON_gi <- gl2gi(time_HWE_nonBON, v = 1)
```

```
save(time_HWE_nonBON_gi, file = "time_HWE_nonBON_gi.RData")


##Save structure formatted file

library(dartR)

gl2structure(time2016LD.HWE.gl, outfile = "time2016LD.HWE.str")
```

7.7. Export PCA Loadings

```
####How to export PCA loadings####


##set your working directory

setwd("C:/Users/kirst/Documents/Incoming Popgen/")


##Create a new project to go with your directory

##if you need to save mid-work, use:

save.image("~/Incoming Popgen/Filtered Set June 2020/PCA_for_sexing.RData")


##Remove NA values, method 1 (tab method preferred)

jasper_noNAs <- tab(jasper_gi2, NA.method="mean")


##Run PCA on the noNAs (nf saves 8 axis comp)

jasper_PCA <- dudi.pca(jasper_noNAs, cent=TRUE, scale=FALSE,
scannf=FALSE, nf=8)


## Export the loadings ($c1) to text file
```

```
write.table(jasper_PCA$c1, "jasper.loading.txt", sep="\t")

write.csv(as.data.frame(time.05.pca$c1, file ="time2005.loadings.csv"))


##Export the coordinates ($li) to graph it out (look for a plateau)

write.table(jasper_PCA$li, "jasper.coords.txt", sep="\t")


##Export the eigenvectors ($eig)

write.table(jasper_PCA$eig, "jasper.eig.txt", sep="\t")


##Export scaffold, position, loci info

write.table(time_gl@chromosome, "time.chrom.txt", sep = "\t")

write.table(time_gl@position, "time.pos.txt", sep = "\t")

write.table(time_gl@loc.names, "time.locnames.txt", sep = "\t")

write.csv(as.data.frame(time_gl$chromosome),file="dartR.SNPChrInfo.csv")

write.csv(as.data.frame(time_gl$position),file="dartR.SNPPosInfo.csv")


##View a barplot of the eigenvalues

barplot(pca_all$eig[1:30],main="PCA eigenvalues", col=heat.colors(50))


##View a scatter graph of the PCA values

s.class(pca_all$li, pop(all_gl_nd), grid=FALSE)


#Add eigen plot to PCA
```

```
add.scatter.eig(pca_all$eig[1:10], 3,1,2)

##Print a coloured version of the PCA scatter (colors must = number of pops)
s.class(pca_all$li, pop(all_gl_nd), col=funky(55), grid=FALSE)

##Sort the loadings large to small in excel
##Make a list of loci to remove from the loadings plot (Trevoy et al. 2020)
list <- c("100049687|12-A/G","100050106|50-G/A")

##Alternate: Use the loci list and vcftools to remove the sex-linked loci (see text code)

##Modify the PCA graph
##Points only
s.class(time.16.pca$li, pop(time_clean.2016.gl), col=funky(5), grid=FALSE, cstar = 0, cellipse = 0)
#Stars and points
s.class(time.16.pca$li, pop(time_clean.2016.gl), col=funky(5), grid=FALSE, cstar = 1, cellipse = 0)
```

7.8. Generate DAPC Manually in R

```
adegenetServer("DAPC")
```

```
setwd("C:/Users/kirst/Documents/Incoming Popgen/")


library(adegenet)
library(lattice)
MPBSNPs <-read.genepop("allsnp_R_test.gen")
grp <- find.clusters(MPBSNPs)
dev.copy(pdf,"F_50_find.clusters_results.pdf", width=7, height=7)
dev.off()
MPBSNPs_NoNas <- tab(MPBSNPs, NA.method="mean")
xval <- xvalDapc(MPBSNPs_NoNas, pop(MPBSNPs), n.pca.max=300,
n.da=NULL, result="overall", n.pca=NULL, n.rep=30, xval.plot=FALSE)
xval
boxplot(xval$'Cross-Validation Results'$success~xval$'Cross-Validation
Results'$n.pca, xlab="Number of PCA components", ylab="Classificationsuccess",
main="DAPC - cross-validation")
dev.copy(pdf,"F_50_xvalBoxplot.pdf", width=7, height=7)
dev.off()
dapc <- dapc(MPBSNPs)
scatter(dapc)
scatter(dapc2005, cex = 1, legend = TRUE, col=c("red", "blue", "green", "purple",
"pink"),clabel = FALSE, posi.leg = "topleft", cleg = 1)
```

7.9. IBD in R and Patch for Large Datasets

```
##Simple IBD in dartR


##Load coordinate data into R

time_clean_all_coords <- read.table("time_clean_coords.txt")


##To create example gps table from dartR

write.table(testset.gl@other$latlong, "testset_gps.txt")


##Should be in the format below (5 decimal places)


##lat    lon

##"1"   52.84985        -118.57265

##"2"   52.84985        -118.57265

##"3"   52.84985        -118.57265

##"4"   52.84985        -118.57265



##Add coordinate data in decimal degrees to genlight

time_clean.gl@other$latlong <- time_clean_all_coords


##View newly added data

time_clean.gl@other$latlong
```

```
##Run IBD check
##Standard analysis performed on the genlight object. Mantel test and plot will be Fst/1-Fst versus log(distance)
##Coordinates transformed to Mercator (google) projection to calculate distances in meters.


gl <-gl.ibd(time_clean.gl)


###PATCH
gl.ibd_fix <- function(gl = NULL, Dgen = NULL, Dgeo = NULL, projected = FALSE,
                       permutations = 999, plot = TRUE)
{
 if (!(requireNamespace("dismo", quietly = TRUE))) {
  stop("Package dismo needed for this function to work. Please install it.")
 }
 else {
  if (!is.null(Dgen) & !is.null(Dgeo))
   cat("Analysis performed on provided genetic and Euclidean distance matrices.")
  if (class(gl) == "genlight") {
   cat("Standard analysis performed on the genlight object. Mantel test and plot
```

```
will be Fst/1-Fst versus log(distance)\n")
        if (nrow(gl@other$latlong) != nInd(gl))
          stop("Cannot find coordinates for each individual in slot @other$latlong")
        if (sum(match(names(gl@other$latlong), "long"), na.rm = T) == 1)
          gl@other$latlong$lon <- gl@other$latlong$long
        if (!projected) {
          xy <- dismo::Mercator(gl@other$latlong[, c("lon", "lat")])
          cat("Coordinates transformed to Mercator (google) projection to calculate
distances in meters.\n")
        }
        else {
          xy = gl@other$latlong[, c("lon", "lat")]
          cat("Coordinates not transformed. Distances calculated on the provided
coordinates.")
        }
        pop.xy <- apply(xy, 2, function(a) tapply(a, pop(gl),
                                                  mean, na.rm = T))
        Dgeo <- dist(pop.xy)
        Dgeo <- log(Dgeo)
        Dgen <- as.dist(StAMPP::stamppFst(gl, nboots = 1))
        Dgen <- Dgen/(1 - Dgen)
        ordering <- levels(pop(gl))
        Dgen <- as.dist(as.matrix(Dgen)[ordering, ordering])
```

```
  Dgeo <- as.dist(as.matrix(Dgeo)[ordering, ordering])
 }
 miss = FALSE
 if (sum(is.na(Dgen)) > 0 | sum(is.infinite(Dgen)) > 0 ) {
  miss = TRUE
  cat("There are missing values in the genetic distance matrix. No kernel distance
plot is possible.\n")
 }
 if (sum(is.na(Dgeo)) > 0 | sum(is.infinite(Dgeo)) > 0 ) {
  miss = TRUE
  cat("There are missing values in the geographic distance matrix. No kernel
distance plot is possible.\n")
 }
 manteltest <- vegan::mantel(Dgen, Dgeo, na.rm = TRUE, permutations = 999)
 print(manteltest)
 if (plot) {
  if (!miss) {
   dens <- MASS::kde2d(Dgeo, Dgen, n = 300)
   myPal <- colorRampPalette(c("white", "blue",
                               "gold", "orange", "red"))
   plot(Dgeo, Dgen, pch = 20, cex = 0.8)
   image(dens, col = transp(myPal(300), 0.7), add = TRUE)
   points(Dgeo, Dgen, pch = 20, cex = 0.8)
```

```
      abline(lm(Dgen ~ Dgeo))
      title("Isolation by distance")
    }
    else {
      plot(Dgeo, Dgen)
      abline(lm(Dgen ~ Dgeo))
      title("Isolation by distance")
    }
  }
  out <- list(Dgen = Dgen, Dgeo = Dgeo, mantel = manteltest)
  return(out)
 }
}
##Import loci data from RDA file

gl.test <- read.loci("MPB_GEN_MODE_num.csv", header = TRUE, allele.sep =
",", loci.sep = "," , col.pop = 2, row.names = 1)
range.gi <- loci2genind(gl.test)
range.gl <- gi2gl(range.gi)
range.gl@other$latlong <- range_cor
ibd <- gl.ibd(range.gl)

##More simple method
```

```
range <- genind2genpop(range.gi)

range.gp <- range

Dgen <- dist.genpop(range.gp,method = 2)

Dgeo <- dist(range.gl@other$latlong)

library(ade4)

ibd <-mantel.randtest(Dgen,Dgeo)

ibd

plot(ibd)


> mpb <-plot((log(Dgeo)), Dgen)

> abline(lm(Dgen ~ log(Dgeo)))


##Monte-Carlo test

##Call: mantel.randtest(m1 = Dgen, m2 = Dgeo)


##Observation: 0.07352354


##Based on 999 replicates

##Simulated p-value: 0.213

##Alternative hypothesis: greater

#

##Std.Obs Expectation    Variance

##0.750276236 0.002470174 0.008968648
```

7.10. Basic Workflow for OUTFLANK

```
##Basic Outflank use

library(dartR)

library(OutFLANK)

all_outflank <- gl.outflank(all_gi, plot = TRUE, LeftTrimFraction = 0.05,
RightTrimFraction = 0.05, Hmin = 0.1, qthreshold = 0.05)

save(time_outflank, file="time_outflank_results.RData")

write.table(all_outflank, "all_outflank_results.txt")


##Print out the list of outliers

all_outflank$outflank$results$OutlierFlag


##Print out the list of loci names

all_outflank$outflank$results$LocusName


##Check for high Fst loci with low He

plot(my_fst$He, my_fst$FST)


##Plot outliers

my_out <- all_outflank$outflank$results$OutlierFlag==TRUE

plot(all_outflank$outflank$results$He, all_outflank$outflank$results$FST, pch=19,
col=rgb(0,0,0,0.1))

points(all_outflank$outflank$results$He[my_out],
```

all_outflank$outflank$results$FST[my_out], col="purple")

7.11. Workflow for RDA Analysis

```
library(psych)

library(vegan)

library(lfmm)

library(qvalue)


##Reminder to save data of various kinds

write.table(bandicoot.gl@other$latlong, "bandicoot_test.txt")

save(range_clean.gl, file = "range_clean_gps.gl.rdata")



mpb.env <-as.data.frame(read.csv("MPB_ENV_Cut2_Site.csv", row.names = 1))

mpb.env <-as.data.frame(read.csv("MPB_ENV_SOUTH.csv"))

##Convert ind names to characters (or site names)


mpb.env$site <- as.character(mpb.env$site)

mpb.env$individual <- as.character(mpb.env$individual)


##Reduce if needed

pred <- subset(mpb.env, select=c(lat , lon , ele , MAT , MAP , FFP , MAR , RH,
per_g))
```

```
##Check import of ENV data
str(mpb.env)


##Import SNP info
mpb.gen <-as.data.frame(read.csv("MPB_GEN_MODE_cut.csv", row.names = 1))

##Check SNPS
dim(mpb.gen)

##Look for NAs
sum(is.na(mpb.gen))

##Remove NAs
mpb.gen.imp <- apply(mpb.gen, 2, function(x) replace(x, is.na(x),
as.numeric(names(which.max(table(x))))))

sum(is.na(mpb.gen.imp))

##Check row names
identical(rownames(mpb.gen.imp), mpb.env[,1])

##Look for correlated factors and remove those with high correlations (number
```

```
must match number of colunms)
        pairs.panels(mpb.env[,2:9], scale=T)

        ##Run RDA
        mpb.rda <- rda(mpb.gen.imp ~  lat + lon + ele + MAT + MAP + FFP + MAR + RH
+ per_g, data = mpb.env, scale = T)

        mpb.rda


        ##Check Rsquared and RSadjust. Constrained ordination explains % of variation.
        RsquareAdj(mpb.rda)

        ##Check Eigenvalues of contrained axes
        summary(mpb.rda)$concont

        ##Visualize Eigenvalues
        screeplot(mpb.rda)

        ##Check sig using F-stat for full model and constrained (R and Radjust)
        signif.full <- anova.cca(mpb.rda, parallel=getOption("mc.cores"))
        signif.full
        signif.axis <- anova.cca(mpb.rda, by="axis", parallel=getOption("mc.cores"))
```

```
signif.axis
write.table(signif.full, "signif.full.mpb.site.txt")
write.table(signif.axis, "signif.axis.mpb.site.txt")

##Look for multicollinearity with variables (consider removing those above 10 - good with currrent dataset)
vif.cca(mpb.rda)

##Plot the RDA 1-2
plot(mpb.rda, scaling=3)

##Plot the RDA 1-3
plot(mpb.rda, choices = c(1, 3), scaling=3)

##Plot the RDA 2-3
plot(mpb.rda, choices = c(2, 3), scaling=3)

##Colour the plots RDA 1-2
cluster <- mpb.env$cluster
bg <- c("#ff7f00","#1f78b4","#ffff33")
plot(mpb.rda, type="n", scaling=3)
points(mpb.rda, display="species", pch=20, cex=0.7, col="gray32", scaling=3) # the SNPs
```

```
points(mpb.rda, display="sites", pch=21, cex=1.3, col="gray32", scaling=3,
bg=bg[cluster])
text(mpb.rda, scaling=3, display="bp", col="#0868ac", cex=1)                    #
the predictors
legend("bottomright", legend=levels(cluster), bty="n", col="gray32", pch=21,
cex=1, pt.bg=bg)

##Colour the plots RDA 1-3
cluster <- mpb.env$cluster
bg <- c("#ff7f00","#1f78b4","#ffff33")
plot(mpb.rda, type="n", scaling=3, choices=c(1,3))
points(mpb.rda, display="species", pch=20, cex=0.7, col="gray32", scaling=3,
choices=c(1,3))
points(mpb.rda, display="sites", pch=21, cex=1.3, col="gray32", scaling=3,
bg=bg[cluster])
text(mpb.rda, scaling=3, display="bp", col="#0868ac", cex=1, choices=c(1,3))
legend("topleft", legend=levels(cluster), bty="n", col="gray32", pch=21, cex=1,
pt.bg=bg)

##Colour the plots RDA 2-3 (edit, can set graph limits, this is an example, play
with this)
cluster <- mpb.env$cluster
bg <- c("#ff7f00","#1f78b4","#ffff33")
```

```
plot(mpb.rda, type="n", scaling=3, xlim=c(-8,9), ylim=c(-8,8), choices=c(2,3))
points(mpb.rda, display="species", pch=20, cex=0.7, col="gray32", scaling=3,
choices=c(2,3))
points(mpb.rda, display="sites", pch=21, cex=1.3, col="gray32", scaling=3,
bg=bg[cluster])
text(mpb.rda, scaling=3, display="bp", col="#0868ac", cex=1, choices=c(2,3))
legend("topleft", legend=levels(cluster), bty="n", col="gray32", pch=21, cex=1,
pt.bg=bg)

##Search for Candidate SNPs
load.rda <- scores(mpb.rda, choices=c(1:3), display="species")  # Species scores
for the first three constrained axes

##Histogram of points
hist(load.rda[,1], main="Loadings on RDA1")
hist(load.rda[,2], main="Loadings on RDA2")
hist(load.rda[,3], main="Loadings on RDA3")

##Outlier SNP Script
outliers <- function(x,z){
  lims <- mean(x) + c(-1, 1) * z * sd(x) ## f.nd loadings +/- z SD from mean
loading
  x[x < lims[1] | x > lims[2]]          # locus names in these tails
```

```
}

cand1 <- outliers(load.rda[,1], 3)
cand2 <- outliers(load.rda[,2], 3)
cand3 <- outliers(load.rda[,3], 3)

##Count candidate loci
ncand <- length(cand1) + length(cand2) + length(cand3)
ncand

#Check for duplicates
mpb.rda.cand <- c(names(cand1), names(cand2), names(cand3)) ## the names of the candidates
length(mpb.rda.cand[duplicated(mpb.rda.cand)]) ## duplicate detections (detected on multiple RDA axes)
mpb.rda.cand <- mpb.rda.cand[!duplicated(mpb.rda.cand)] ## unique candidates
write.table(mpb.rda.cand, "mpb.rda.cand.txt")

##Colours for graphs
bgcol  <- ifelse(colnames(mpb.gen.imp) %in% mpb.rda.cand, 'gray32', '#00000000')
snpcol <- ifelse(colnames(mpb.gen.imp) %in% mpb.rda.cand, 'red', '#00000000')
```

```
##Plot SNP Outliers Axis 1-2  1-3 and 2-3
plot(mpb.rda, type="n", scaling=3, xlim=c(-1,1), ylim=c(-1,1), main="MPB mpb RDA, axes 1 and 2")
points(mpb.rda, display="species", pch=21, cex=1, col="gray32", bg='#f1eef6', scaling=3)
points(mpb.rda, display="species", pch=21, cex=1, col=bgcol, bg=snpcol, scaling=3)
text(mpb.rda, scaling=3, display="bp", col="#0868ac", cex=1)

plot(mpb.rda, type="n", scaling=3, xlim=c(-1,1), ylim=c(-1,1),choices=c(1,3), main="MPB mpb RDA, axes 1 and 3")
points(mpb.rda, display="species", pch=21, cex=1, col="gray32", bg='#f1eef6', scaling=3,choices=c(1,3))
points(mpb.rda, display="species", pch=21, cex=1, col=bgcol, bg=snpcol, scaling=3, choices=c(1,3))
text(mpb.rda, scaling=3, display="bp", col="#0868ac", cex=1, choices=c(1,3))

plot(mpb.rda, type="n", scaling=3, xlim=c(-1,1), ylim=c(-1,1),choices=c(2,3), main="MPB mpb RDA, axes 2 and 3")
points(mpb.rda, display="species", pch=21, cex=1, col="gray32", bg='#f1eef6', scaling=3,choices=c(2,3))
points(mpb.rda, display="species", pch=21, cex=1, col=bgcol, bg=snpcol, scaling=3, choices=c(2,3))
```

```
text(mpb.rda, scaling=3, display="bp", col="#0868ac", cex=1, choices=c(2,3))


##Separate out SNP loci into ENV predictors (must make the pred file for function to work)
cand1 <- cbind.data.frame(rep(1,times=length(cand1)), names(cand1), unname(cand1))
cand2 <- cbind.data.frame(rep(2,times=length(cand2)), names(cand2), unname(cand2))
cand3 <- cbind.data.frame(rep(3,times=length(cand3)), names(cand3), unname(cand3))
colnames(cand1) <- colnames(cand2) <- colnames(cand3) <- c("axis","snp","loading")
cand <- rbind(cand1, cand2, cand3)
cand$snp <- as.character(cand$snp)
foo <- matrix(nrow=(ncand), ncol=9)  # 9 columns for 9 predictors
colnames(foo) <- c("lat" , "lon" , "ele" , "MAT" , "MAP" , "FFP" , "MAR" , "RH", "per_g")


for (i in 1:length(cand$snp)) {
nam <- cand[i,2]
snp.gen <- mpb.gen.imp[,nam]
foo[i,] <- apply(pred,2,function(x) cor(x,snp.gen))
}
```

```
cand <- cbind.data.frame(cand,foo)

head(cand)

##Look for duplicated SNPs

length(cand$snp[duplicated(cand$snp)])
foo <- cbind(cand$axis, duplicated(cand$snp))

table(foo[foo[,1]==1,2]) # no duplicates on axis 1
table(foo[foo[,1]==2,2])
table(foo[foo[,1]==3,2])
cand <- cand[!duplicated(cand$snp),]

##Numbers here depend on your cand file X:X means where your variables start
and end
for (i in 1:length(cand$snp)) {
bar <- cand[i,]
cand[i,13] <- names(which.max(abs(bar[4:12]))) # gives the variable
cand[i,14] <- max(abs(bar[4:12]))             # gives the correlation
}
```

```
colnames(cand)[13] <- "predictor"

colnames(cand)[14] <- "correlation"

table(cand$predictor)



##Plot outliers for ENV (Does not work yet)

sel <- cand$snp

env <- cand$predictor


env[env=="lat"] <- '#1f78b4'

env[env=="lon"] <- '#a6cee3'

env[env=="ele"] <- '#6a3d9a'

env[env=="MAT"] <- '#e31a1c'

env[env=="MAP"] <- '#33a02c'

env[env=="FFP"] <- '#ffff33'

env[env=="MAR"] <- '#fb9a99'

env[env=="RH"] <- '#b2df8a'

env[env=="per_g"] <- '#b642f5'


# color by predictor:

col.pred <- rownames(mpb.rda$CCA$v)


for (i in 1:length(sel)) {          # color code candidate SNPs
```

```
foo <- match(sel[i],col.pred)
col.pred[foo] <- env[i]
}

col.pred[grep("chr",col.pred)] <- '#f1eef6' # non-candidate SNPs
empty <- col.pred
empty[grep("#f1eef6",empty)] <- rgb(0,1,0, alpha=0) # transparent
empty.outline <- ifelse(empty=="#00FF0000","#00FF0000","gray32")
bg <-
c('#1f78b4','#a6cee3','#6a3d9a','#e31a1c','#33a02c','#ffff33','#fb9a99','#b2df8a', '#b642f5')

# axes 1 & 2
plot(mpb.rda, type="n", scaling=3, xlim=c(-1,1), ylim=c(-1,1))
points(mpb.rda, display="species", pch=21, cex=1, col="gray32", bg=col.pred,
scaling=3)
points(mpb.rda, display="species", pch=21, cex=1, col=empty.outline, bg=empty,
scaling=3)
text(mpb.rda, scaling=3, display="bp", col="#0868ac", cex=1)
legend("bottomright", legend=c("lat" , "lon" , "ele" , "MAT" , "MAP" , "FFP" ,
"MAR" , "RH", "per_g"), bty="n", col="gray32", pch=21, cex=1, pt.bg=bg)

## axes 1 & 3
plot(mpb.rda, type="n", scaling=3, xlim=c(-1,1), ylim=c(-1,1), choices=c(1,3))
```

points(mpb.rda, display="species", pch=21, cex=1, col="gray32", bg=col.pred, scaling=3, choices=c(1,3))

points(mpb.rda, display="species", pch=21, cex=1, col=empty.outline, bg=empty, scaling=3, choices=c(1,3))

text(mpb.rda, scaling=3, display="bp", col="#0868ac", cex=1, choices=c(1,3))

legend("bottomright", legend=c("lat" , "lon" , "ele" , "MAT" , "MAP" , "FFP" , "MAR" , "RH", "per_g"), bty="n", col="gray32", pch=21, cex=1, pt.bg=bg)

8. Appendix 2: Time Study Arlequin Output

AMOVA-NO GROUPS

==========

== Comparisons of pairs of population samples

==========

List of labels for population samples used below:

Label Population name

----- ----------

1: pop_YH

2: pop_M01

3: pop_M06

4: pop_M11

5: pop_TE

6: pop_xX

7: pop_A

8: pop_xL

9: pop_xT

10: pop_xS

Population pairwise FSTs

Distance method: Pairwise difference

	1	2	3	4	5	6	7	8	9	10
1	0.00000									
2	0.03019	0.00000								
3	0.04386	0.10791	0.00000							
4	0.03910	0.10789	0.00193	0.00000						

5 0.01685 0.03812 0.03153 0.03014 0.00000

6 0.05717 0.00796 0.13006 0.13827 0.06254 0.00000

7 0.01419 0.02072 0.04954 0.04849 0.01689 0.04394 0.00000

8 0.03260 -0.00113 0.10653 0.10732 0.03843 0.01040 0.02255 0.00000

9 0.05452 0.09243 0.02621 0.03912 0.04428 0.10127 0.04645 0.09413 0.00000

10 0.01306 0.00438 0.05653 0.05781 0.01515 0.02033 0.00974 0.00701 0.05467 0.00000

FST P values

Number of permutations : 1023

1 2 3 4 5 6 7 8 9 10

1 *

2 0.00684+-0.0023 *

3 0.00098+-0.0010 0.00000+-0.0000 *

4 0.00000+-0.0000 0.00000+-0.0000 0.30078+-0.0150 *

5 0.01172+-0.0030 0.00000+-0.0000 0.01270+-0.0039 0.00488+-0.0020 *

6 0.00000+-0.0000 0.02832+-0.0051 0.00000+-0.0000 0.00000+-0.0000 0.00000+-0.0000 *

7 0.03809+-0.0066 0.00195+-0.0014 0.00000+-0.0000 0.00000+-0.0000 0.01270+-0.0031 0.00000+-0.0000 *

8 0.00000+-0.0000 0.36426+-0.0166 0.00000+-0.0000 0.00000+-0.0000 0.00000+-0.0000 0.00391+-0.0019 0.00000+-0.0000 *

9 0.00000+-0.0000 0.00000+-0.0000 0.03125+-0.0055 0.00000+-0.0000 0.00000+-0.0000 0.00000+-0.0000 0.00000+-0.0000 0.00000+-0.0000 *

10 0.02441+-0.0047 0.18066+-0.0117 0.00098+-0.0010 0.00000+-0.0000 0.00879+-0.0025 0.00098+-0.0010 0.03516+-0.0049 0.03613+-0.0057 0.00000+-0.0000 *

Matrix of significant Fst P values

Significance Level=0.0500

Number of permutations : 1023

	1	2	3	4	5	6	7	8	9	10
1		+	+	+	+	+	+	+	+	+
2	+		+	+	+	+	+	-	+	-
3	+	+		-	+	+	+	+	+	+
4	+	+	-		+	+	+	+	+	+
5	+	+	+	+		+	+	+	+	+
6	+	+	+	+	+		+	+	+	+
7	+	+	+	+	+	+		+	+	+
8	+	-	+	+	+	+	+		+	+
9	+	+	+	+	+	+	+	+		+
10	+	-	+	+	+	+	+	+	+	

9. Appendix 3: Range Study Genodive Output

– Summary of indices of genetic diversity

Statistic	Value	Std.Dev.	c.i.2.5%	c.i.97.5%	Description
Num	2.000	0.000	2.000	2.000	Number of alleles
Eff_num	1.438	0.005	1.429	1.447	Effective number of alleles

Ho	0.289	0.002	0.284	0.293	Observed Heterozygosity
Hs	0.289	0.002	0.285	0.294	Heterozygosity Within Populations
Ht	0.295	0.002	0.291	0.300	Total Heterozygosity
H't	0.296	0.002	0.291	0.300	Corrected total Heterozygosity
Gis	0.002	0.001	0.000	0.003	Inbreeding coefficient

Standard deviations of F-statistics were obtained through jackknifing over loci.

95% confidence intervals of F-statistics were obtained through bootstrapping over loci.

– Indices of genetic diversity per population

Population	Num	Eff_num	Ho	Hs	Gis
Pop001	1.889	1.450	0.273	0.280	0.024
Pop002	1.902	1.459	0.281	0.286	0.019
Pop003	1.876	1.447	0.269	0.278	0.031
Pop004	1.895	1.459	0.285	0.286	0.002
Pop005	1.992	1.474	0.279	0.297	0.061

Pop006	1.950	1.470	0.293	0.297	0.013
Pop007	1.971	1.475	0.301	0.300	-0.002
Pop008	1.938	1.476	0.317	0.301	-0.052
Pop009	1.979	1.470	0.290	0.297	0.021
Pop010	1.977	1.481	0.310	0.302	-0.026
Pop011	1.939	1.465	0.291	0.296	0.015
Pop012	1.978	1.472	0.293	0.298	0.017
Pop013	1.943	1.469	0.292	0.299	0.021
Pop014	1.935	1.466	0.292	0.297	0.017
Pop015	1.946	1.467	0.290	0.296	0.021
Pop016	1.975	1.472	0.294	0.298	0.013
Pop017	1.895	1.461	0.296	0.297	0.006
Pop018	1.953	1.471	0.304	0.297	-0.022
Pop019	1.975	1.476	0.293	0.300	0.024
Pop020	1.988	1.476	0.299	0.299	0.002
Pop021	1.999	1.483	0.306	0.301	-0.017
Pop022	1.991	1.475	0.296	0.298	0.005
Pop023	1.967	1.467	0.293	0.296	0.011
Pop024	1.884	1.445	0.282	0.281	-0.005

Pop025	1.954	1.463	0.301	0.289	-0.040
Pop026	1.942	1.450	0.282	0.284	0.007
Pop027	1.940	1.457	0.296	0.286	-0.033
Pop028	1.930	1.450	0.281	0.285	0.015
Pop029	1.945	1.451	0.285	0.284	-0.001
Pop030	1.942	1.444	0.279	0.280	0.006
Pop031	1.953	1.446	0.282	0.281	-0.001
Pop032	1.944	1.449	0.282	0.282	0.001
Pop033	1.954	1.448	0.278	0.282	0.015
Pop034	1.938	1.449	0.296	0.284	-0.043
Pop035	1.960	1.449	0.281	0.283	0.009
Pop036	1.952	1.452	0.281	0.285	0.015
Pop037	1.945	1.450	0.283	0.283	0.001
Pop038	1.950	1.457	0.299	0.286	-0.044
Pop039	1.946	1.453	0.283	0.285	0.008
Pop040	1.934	1.446	0.280	0.282	0.007
Pop041	1.940	1.447	0.278	0.282	0.014
Pop042	1.987	1.478	0.287	0.299	0.039
Pop043	1.964	1.455	0.281	0.285	0.014

Pop044	1.967	1.468	0.280	0.297	0.056
Pop045	1.929	1.453	0.296	0.287	-0.031
Pop046	1.927	1.445	0.282	0.282	-0.003
Pop047	1.906	1.444	0.281	0.283	0.007
Pop048	1.921	1.453	0.299	0.284	-0.054
Pop049	1.900	1.445	0.294	0.283	-0.037
Pop050	1.947	1.444	0.280	0.280	0.001
Pop051	1.948	1.446	0.275	0.281	0.021
Pop052	1.846	1.435	0.280	0.280	0.002
Pop053	1.979	1.455	0.282	0.284	0.008
Pop054	1.914	1.454	0.305	0.286	-0.065
Pop055	1.915	1.447	0.284	0.283	-0.003

Principal Components.

File: range_clean_gps.str

1018 of 1018 individuals included, 3612 of 3612 loci included

– Summary statistics for all 54 retained PCA-axes (out of 7224)

Axis	%Variance	Cumulative	G'st(Nei)	P-value
1	36.401	36.401	0.008	0.001
2	2.846	39.246	0.001	1.000
3	2.191	41.437	0.000	1.000
4	2.077	43.515	0.000	1.000
5	1.880	45.395	0.000	0.999
6	1.841	47.236	0.000	0.999
7	1.790	49.026	0.000	1.000
8	1.706	50.732	0.000	1.000
9	1.669	52.401	0.000	1.000
10	1.609	54.010	0.000	1.000
11	1.571	55.581	0.000	1.000
12	1.537	57.119	0.000	1.000

13	1.461	58.580	0.000	1.000
14	1.448	60.028	0.000	1.000
15	1.418	61.446	0.000	1.000
16	1.371	62.817	0.000	1.000
17	1.358	64.175	0.000	1.000
18	1.323	65.499	0.000	1.000
19	1.258	66.757	0.000	1.000
20	1.246	68.002	0.000	1.000
21	1.218	69.221	0.000	1.000
22	1.195	70.415	0.000	1.000
23	1.181	71.597	0.000	1.000
24	1.157	72.754	0.000	1.000
25	1.148	73.902	0.000	1.000
26	1.127	75.029	0.000	1.000
27	1.112	76.141	0.000	1.000
28	1.105	77.246	0.000	1.000
29	1.102	78.348	0.000	1.000
30	1.071	79.419	0.000	1.000
31	1.052	80.471	0.000	1.000

32	1.035	81.507	0.000	1.000
33	1.034	82.541	0.000	1.000
34	1.026	83.566	0.000	1.000
35	1.002	84.568	0.000	1.000
36	0.988	85.556	0.000	1.000
37	0.966	86.522	0.000	1.000
38	0.949	87.471	0.000	1.000
39	0.938	88.410	0.000	1.000
40	0.935	89.345	0.000	1.000
41	0.913	90.258	0.000	1.000
42	0.909	91.167	0.000	1.000
43	0.878	92.045	0.000	1.000
44	0.859	92.904	0.000	1.000
45	0.849	93.753	0.000	0.999
46	0.821	94.574	0.000	1.000
47	0.810	95.384	0.000	1.000
48	0.771	96.155	0.000	1.000
49	0.758	96.913	0.000	1.000
50	0.751	97.664	0.000	1.000

51 0.655 98.319 0.000 1.000

52 0.638 98.957 0.000 1.000

53 0.532 99.489 0.000 1.000

54 0.511 100.000 0.000 1.000

Overall value of G'st(Nei): 0.022

GenoDive 3.05, 2022-02-09 19:08:10 +0000

K-Means Clustering: Meirmans, 2012.

File: range_clean_gps.str

1018 of 1018 individuals included, 3612 of 3612 loci included

Clustering of 55 populations using among clusters sums of squares from an Analysis of Molecular Variance.

Simulated annealing using 50000 steps and 20 random starts

– Clustering statistics from k = 2 to k = 6

k	SSD(T)	SSD(AC)	SSD(WC)	r-squared	pseudo-F	BIC	Rho
2*	1227463.698	30964.870	1196498.828	0.257	18.351	635.082	0.047
3	1227463.698	40788.576	1186675.123	0.339	13.322	632.690	0.054
4&	1227463.698	48199.688	1179264.011	0.400	11.349	631.323	0.055
5	1227463.698	52807.134	1174656.564	0.439	9.766	631.703	0.054
6	1227463.698	55346.418	1172117.280	0.460	8.338	633.604	0.053

* Best clustering according to Calinski & Harabasz' pseudo-F: k = 2

& Best clustering according to Bayesian Information Criterion: k = 4

Best BIC clustering has been stored as population groups.

– Analysis of Molecular Variance on best clustering (according to BIC):

Source of Variation	Nested in	SSD	d.f.	MS	%var	Rho	Rho-value
Within Populations	--	1107067.338	963	1149.603	0.933	st	0.067
Among Populations	Clusters	72196.673	51	1415.621	0.012	sc	0.012
Among Clusters	--	48199.688	3	16066.563	0.055	ct	0.055
Total (SST)	--	1227463.698	1017	1206.946	--	--	--

---

10. Apppendix 4: Starting List of ClimateNA variables

Lat
long
elev
period

MAT
MWMT
MCMT
TD
MAP
MSP
AHM
SHM
DD_0
DD5
DD_18
DD18
NFFD
bFFP
eFFP
FFP
PAS
EMT
EXT
Eref
CMD
MAR
RH
CMI
DD1040
Tmax_wt
Tmax_sp
Tmax_sm
Tmax_at
Tmin_wt
Tmin_sp
Tmin_sm
Tmin_at
Tave_wt
Tave_sp
Tave_sm
Tave_at
PPT_wt
PPT_sp
PPT_sm

PPT_at
Rad_wt
Rad_sp
Rad_sm
Rad_at
DD_0_wt
DD_0_sp
DD_0_sm
DD_0_at
DD5_wt
DD5_sp
DD5_sm
DD5_at
DD_18_wt
DD_18_sp
DD_18_sm
DD_18_at
DD18_wt
DD18_sp
DD18_sm
DD18_at
NFFD_wt
NFFD_sp
NFFD_sm
NFFD_at
PAS_wt
PAS_sp
PAS_sm
PAS_at
Eref_wt
Eref_sp
Eref_sm
Eref_at
CMD_wt
CMD_sp
CMD_sm
CMD_at
RH_wt
RH_sp
RH_sm

RH_at
CMI_wt
CMI_sp
CMI_sm
CMI_at
Tmax_01
Tmax_02
Tmax_03
Tmax_04
Tmax_05
Tmax_06
Tmax_07
Tmax_08
Tmax_09
Tmax_10
Tmax_11
Tmax_12
Tmin_01
Tmin_02
Tmin_03
Tmin_04
Tmin_05
Tmin_06
Tmin_07
Tmin_08
Tmin_09
Tmin_10
Tmin_11
Tmin_12
Tave_01
Tave_02
Tave_03
Tave_04
Tave_05
Tave_06
Tave_07
Tave_08
Tave_09
Tave_10
Tave_11

Tave_12
Prec_01
Prec_02
Prec_03
Prec_04
Prec_05
Prec_06
Prec_07
Prec_08
Prec_09
Prec_10
Prec_11
Prec_12
Rad_01
Rad_02
Rad_03
Rad_04
Rad_05
Rad_06
Rad_07
Rad_08
Rad_09
Rad_10
Rad_11
Rad_12
DD_0_01
DD_0_02
DD_0_03
DD_0_04
DD_0_05
DD_0_06
DD_0_07
DD_0_08
DD_0_09
DD_0_10
DD_0_11
DD_0_12
DD5_01
DD5_02
DD5_03

DD5_04
DD5_05
DD5_06
DD5_07
DD5_08
DD5_09
DD5_10
DD5_11
DD5_12
DD_18_01
DD_18_02
DD_18_03
DD_18_04
DD_18_05
DD_18_06
DD_18_07
DD_18_08
DD_18_09
DD_18_10
DD_18_11
DD_18_12
DD18_01
DD18_02
DD18_03
DD18_04
DD18_05
DD18_06
DD18_07
DD18_08
DD18_09
DD18_10
DD18_11
DD18_12
NFFD_01
NFFD_02
NFFD_03
NFFD_04
NFFD_05
NFFD_06
NFFD_07

NFFD_08
NFFD_09
NFFD_10
NFFD_11
NFFD_12
PAS_01
PAS_02
PAS_03
PAS_04
PAS_05
PAS_06
PAS_07
PAS_08
PAS_09
PAS_10
PAS_11
PAS_12
Eref_01
Eref_02
Eref_03
Eref_04
Eref_05
Eref_06
Eref_07
Eref_08
Eref_09
Eref_10
Eref_11
Eref_12
CMD_01
CMD_02
CMD_03
CMD_04
CMD_05
CMD_06
CMD_07
CMD_08
CMD_09
CMD_10
CMD_11

CMD_12
RH_01
RH_02
RH_03
RH_04
RH_05
RH_06
RH_07
RH_08
RH_09
RH_10
RH_11
RH_12
CMI_01
CMI_02
CMI_03
CMI_04
CMI_05
CMI_06
CMI_07
CMI_08
CMI_09
CMI_10
CMI_11
CMI_12
COUNT
AREA
MIN
MAX
RANGE
MEAN
STD
SUM
PERCENT