How can genomic data inform biological invasions?

A thesis

submitted in partial fulfilment

of the requirements for the degree

of

Master of Science (Research) in Ecology and Biodiversity

at

The University of Waikato

by

Paige Matheson

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2022

Abstract

Rates of biological invasions are increasing, with global trade and climate change causing significant damage to biodiversity, human well-being, primary industries, and economies around the world. However, our ability to predict and prevent future invasions is limited by significant gaps in our mechanistic understanding of the invasion process. Advances in next generation sequencing technologies and bioinformatics make it possible to investigate potential genomic factors that drive invasion success with much higher resolution and accuracy than prior research based on a small number of genetic loci. My thesis argues for the value of population genomic data in invasion biology, first examining the uptake of genomics in invasion research and then providing a case study for using genomic data to understand invasion patterns of pink bollworm (*Pectinophora gossypiella*).

The first analysis (Chapter 2) compares the extent to which population *genetic* data versus population *genomic* data, including reference genomes, have been used or are publicly available to study globally invasive species from the International Union for Conservation of Nature (IUCN) "100 of the World's Worst Invasive Alien Species" (WAS) list. In this chapter, I demonstrate that 'invasion genomics' is still in its infancy with respect to research uptake: while 82% of species on the WAS list have been studied using some form of population genetic data, just 32% have been studied using population genomic data. Further, 55% of the WAS list species lack a reference genome, however 18% of these were sequenced in the last three years, indicating a growing investment in genomic resources that looks promising for future invasion genomics research.

The second analysis (Chapter 3) showcases population genomic data being used as a tool to inform a biological invasion. Pink bollworm is one of the most destructive global pests

of cotton, costing farmers millions of dollars each year in productivity losses and management efforts. A small population of pink bollworm is currently established in North West Australia, where it poses a significant threat to the expanding cotton industry there. In this chapter, I analysed genomic data in the form of single nucleotide polymorphisms (SNPs) – obtained through a reduced representation, genotyping-by-sequencing technique (DArTseq) – for global populations of pink bollworm to elucidate the population structure and connectivity patterns of the pest. My results show that pink bollworm populations in my dataset have low genetic diversity and strong differentiation between populations from different continents. Importantly, the high genetic differentiation between Australia and other continents reduces concerns about gene flow to North West Australia, particularly from populations in India and Pakistan that have evolved resistance to transgenic insecticidal cotton.

As species continue to move globally beyond their natural ranges, understanding how genome-driven processes facilitate invasion is critical. Genomic data can enhance our mechanistic understanding of the invasion process and inform proactive management of invasive species. Yet, despite progress in this space, there remain limitations to the breadth and depth of such studies that are highlighted throughout my thesis. These represent valuable research opportunities. With the cost of generating genomic data constantly decreasing and long-read sequencing bridging the gap for many taxon-specific challenges, genomic data is starting to address many previously intractable research questions and has the potential to improve overall biosecurity outcomes worldwide.

Acknowledgements

I am so incredibly fortunate to have such wonderful people in my life that have supported me in so many ways throughout my Master's Degree.

I express my deepest gratitude to my brilliant, kind, and intelligent supervisor, Dr Ang McGaughran. Ang - I absolutely adore you, not only as a scientist, but also as a human. The passion you have for your work and students is so inspiring and infectious, and I feel really honored to have worked with you over the past few years. I could add an entire chapter to my thesis about what I am grateful for, but I will try to narrow it down because it is already over 150 pages long! Thank you for always being so patient and kind, especially in the beginning when I was still learning the ropes. Thank you for genuinely caring about my success, and always providing me with immeasurable amounts of opportunities, support, advice, and encouragement. Thank you for making these few years so fun and positive that I never want to leave. You're the best.

Thank you to all of the friends I have made during my time at University. Particularly to the members of the Invasomics Lab group (and honorary members) - thank you! You are such an awesome group to work with. Thank you for letting me tag along to your field work, nerd-out about bugs/animals with you, relate with you about the struggles we are facing, and for just being really, really, cool people. You guys are so great and so smart, and I know you will all be super successful in your future studies/careers.

I also wish to thank Eli Parvizi, Bruce Tabashnik, Jeff Fabrick, Hamid Anees Siddiqui, and Tom Walsh for the time they took to provide careful and considered

feedback on Chapter 3. I am very grateful for how this chapter improved following your expert feedback.

One of my biggest thanks goes to my managers and co-workers. Rita - you have genuinely been one of my biggest supporters since the day I decided I wanted to go to university. Thank you for being so flexible and always finding a way to work around my university hours. Thank you for always having lollies in the drawer and coffee in your locker. Allen - I have learnt a lot about life from you, from the best shares to invest in to how to make the perfect sandwich, thank you. And to everybody else – thank you for the random conversations that were a welcome distraction and for always checking in on how I'm going – your support does not go unnoticed!

To my gorgeous friends Brooke and Kaitlyn - thank you for your unwavering hype and support. You both bring such a light to my life and inspire me in so many ways. Brooke you are one of my oldest and dearest friends - whenever I hear The Gambler I get wholesome flashbacks to Friday nights in Shenanigans. You are so smart, kind, and driven and I love you (and your beautiful son) to pieces. Kaitlyn - if someone told us in high school that our futures would entail you becoming a high school teacher and me doing a Masters degree in science - we both would have cracked up laughing. Your students are so lucky to have such a funny, relatable, and intelligent teacher, and I love and miss you so much! Thank you both for the continuous support over these years; you have honestly been integral to this endeavour and I could not have done it without you.

My family has been a fundamental part of this experience and I can't thank them enough for always encouraging me to pursue my interests and for being my number one supporters. Mum and Dad – you have both shaped me into the person I am today. Thank you for all of your sound advice, hugs, and comfort whenever the going gets tough - they have certainly gotten me through the past few years. Thank you for always pushing me

to be my best and always believing in me. Thank you for entertaining my (sometimes) hours-long spiels about science. Tayne - thanks for being such a cool, funny, and entertaining little brother. I love all of the random pictures I get sent of wildlife throughout the day from each of you because it “seems like something I would appreciate”. You’re right – I do. Tucker the dog - you can’t read this, but I appreciate you too. I love you all so much.

Lastly, to Tavis. I am so grateful to have had you by my side for the longest time, through the good times and the bad. I am so proud of how far you have come and I can’t imagine life with you. You’re the best.

Table of Contents

List of Figures

Appendices:

List of Tables

Chapter 1.

General Introduction

1.1 Introduction

Ecosystems have evolved in isolation for the bulk of Earth's geologic history – separated by natural physical barriers, such as oceans or mountain chains. This has resulted in the organisation of diverse populations of flora and fauna, representative of the environment in which they have evolved (Vitousek et al. 1997). We are currently experiencing a new geological era, faced with the consequences of planetary-scale environmental changes driven primarily by anthropogenic activities, i.e., the Anthropocene (Hautier et al. 2015). In real time, the diversity of life is being redistributed, with some species shifting to more tolerable habitats to meet demands imposed by land-use and climate change (Pecl et al. 2017), or being transported by humans – the latter of which is recognised as the primary cause of biological invasions (Hulme, 2008).

Biological invasions are defined as occurring when a species is introduced beyond its natural range and establishes a sustainable population in the 'invaded range', where it negatively impacts ecosystem functioning (Blackburn et al. 2011), biodiversity (Doherty et al. 2016), human health (Mazza et al. 2013), primary industries (Paini et al. 2016), and/or economies (Diagne et al. 2021). These adverse impacts are how most scientists discriminate 'invasive' species from other 'non-native' species, which may have relatively benign or even beneficial impacts – such as providing habitat, food or valuable ecosystem functions (Schlaepfer et al. 2011).

The total global reported cost of mitigating and managing invasive species impacts was at least $USD1.288 trillion between the years 1970-2017, with this figure expected to rise exponentially as changing climate and trade accelerate the pace of invasion (Diagne et al. 2021). While some invasives have very clear direct impacts, others act more ambiguously, with consequences that indirectly cascade through the ecosystem and lead to long term 'regime shifts', which are virtually impossible to reverse (Linders et al. 2019; Shackleton et al. 2018). For instance, the introduction of a single invasive

species can cause a decrease of up to 16.6% in species richness by introducing ecological pressures such as predation, competition, and/or parasitism (Blackburn et al. 2004; Butchart et al. 2010; Mollot et al. 2017; Tobin, 2018), and invasive mammals are directly attributed to approximately 58% of all bird, mammal, and reptile extinctions worldwide (Doherty et al. 2016). Invasive species are also capable of modifying fire regimes (García et al. 2015), nutrient cycling (Allison & Vitousek, 2004), and trophic webs (Wainright et al. 2021). As a result of these collective impacts, many components of human well-being are also jeopardised, with reduced food security and income generation, and increased exposure to vector-borne diseases (Shackleton et al. 2019). In Aotearoa New Zealand, invasive species compromise the integrity of natural spaces that are essential for hauora – the understanding that contact with nature improves physical, emotional, and spiritual well-being (see: https://natlib.govt.nz/blog/posts/nature-and-nurture-connecting-conservation-and-wellbeing).

Not only is it essential to research biological invasions to aid measures for reducing their collective impacts; they are also interesting natural experiments for studying rapid adaptive processes in the face of global change. In the next sections, I describe the invasion process, outline the challenges associated with invasive biology research, discuss how invasion biology can benefit from a marriage with population genomics, and introduce the main objectives of my thesis.

1.2 The invasion process

The invasion process can be viewed as a series of stages (transport, introduction, establishment, and spread), with the graduation of each stage contingent on overcoming barriers set forth by each preceding stage (Blackburn et al. 2011; Grarock et al. 2013; Colautti & MacIsaac, 2004) (Figure 1.1).

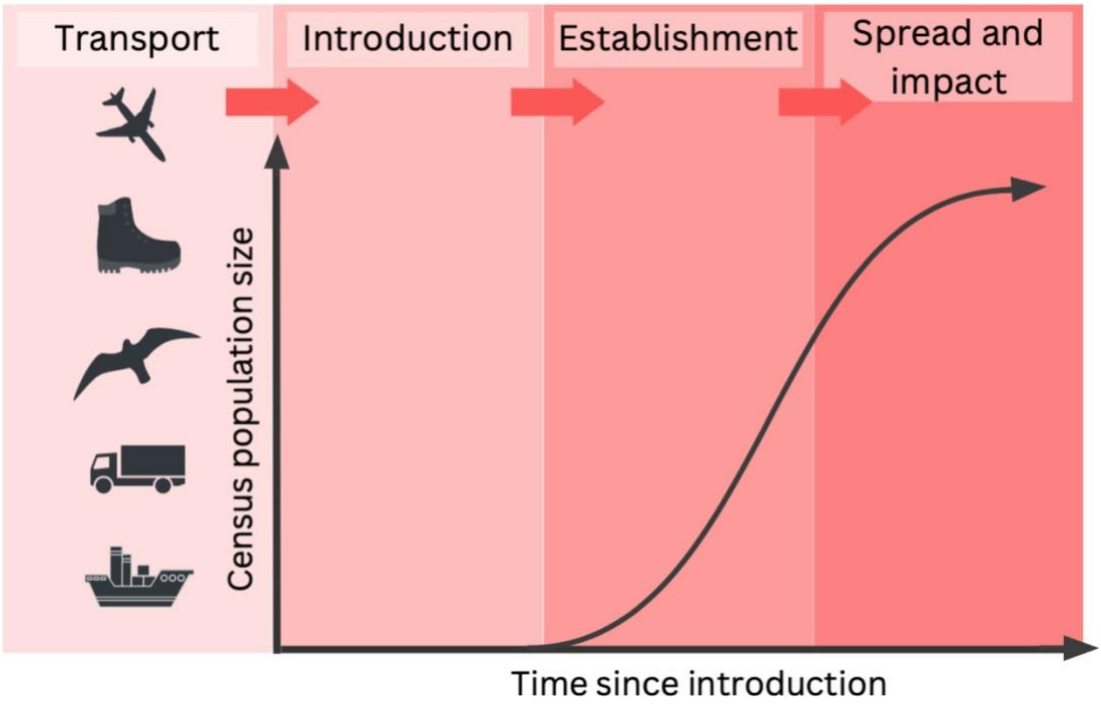


Figure 1.1 Each stage of the invasion process: transport (and icons of potential human-mediated invasion vectors), introduction, establishment, and spread/impact.

Invasion is initiated when individuals (eggs, seeds, larvae, mature adults, etc.) are transported to regions beyond their native range through various intentional and unintentional introduction pathways (Meyerson & Mooney, 2007). For example, species may be introduced deliberately by being imported as a commodity (e.g., livestock or forestry), for aesthetic purposes (e.g., ornamental plants), for biocontrol of pest species, and/or for game or sport. Meanwhile, accidental introductions may occur as a by-product of commodity transport (e.g., parasites and diseases), as a result of species stowaways on ships or aeroplanes, or via the opening of new dispersal pathways mediated by artificially engineered infrastructures (e.g., water canals, bridges; Ascensão, & Capinha, 2017; Hulme, 2008).

Globalisation has shaped the magnitude, frequency, geographic, and taxonomic trends of invasions (Grarock et al. 2013; Hulme, 2009). The quantity and diversity of

individuals being introduced to new regions every day is large, however very few individuals are likely to survive the journey, and even less will survive the physical and biotic agents they encounter upon arrival in a new habitat. In order for a species to establish a viable population from a small number of founders, they must overcome the consequent lower reproductive fitness and the lower adaptive potential that can often result from a lack of genetic diversity during invasion (e.g., due to genetic bottlenecks – rapid reductions in population size that limit genetic diversity, often associated with foundation of new populations from a small number of individuals with limited genetic diversity that establish in a new area; Sherpa et al. 2019).

Establishment has been shown to depend greatly on genetic, demographic, and environmental factors, and the ways in which these interact with one another (Fautley et al. 2012; Sherpa & Després, 2021). For example, introduced mammals that have successfully established populations in Australia were shown to have a larger climatically-suitable geographic range and traits associated with a faster population growth rate (e.g., shorter life span and more offspring per year), facilitating a greater frequency of introduction events with higher rates of individuals (i.e., propagule pressure; Blackburn et al. 2015; Forsyth et al. 2004). Successfully established populations are said to be 'naturalised' and may stay this way for many years, having little to no impact. In fact, some populations experience lag periods of tens to hundreds of years until conditions optimise, or species sufficiently adapt to the new environment, sparking population growth and range expansion (Coutts et al. 2018; Spear et al. 2021). In New Zealand, 91% of 142 invasive plant species exhibited such lag periods prior to population expansion, and 5% of these periods lasted over 40 years (Aikio et al. 2010). Once a population experiences rapid growth and range expansion, and causes ecological and economic damage, it is considered invasive (Blackburn et al. 2011).

1.3 Challenges of invasion biology research

The majority of invasion biology research has only been published within the last three decades, thus it is a relatively new and rapidly developing field (Guiaşu & Tindale, 2017), with some criticism and debate about how research outcomes can be effectively implemented in policy (Richardson & Ricciardi, 2013). Indeed, a lack of global datasets, our inability to predict invasions and quantify their impacts, and the argument that successful eradication regimes are never definitive continue to place limits on what we are capable of achieving from a biosecurity standpoint (Cassini, 2020). These uncertainties arise from the complex evolutionary and ecological factors that underpin each unique biological invasion – including interspecific relationships (e.g., trophic webs), demographic dynamics (e.g., founder events), co-evolutionary processes, abiotic effects, and anthropogenic influences (e.g., land-use change; Courchamp et al. 2017) – and the way these factors interact.

Invasive species are capable of modifying specific aspects of their physiology and behaviour in ways that may select for or otherwise promote invasion success (Lee, 2002; Sherpa & Després, 2021). This may include wider dispersal, rapid reproduction, faster growth rates, greater size, greater breadth of diet, advantages over competitive species, and/or increased aggression (Adrian-Kalchhauser et al. 2020; Pearce et al. 2017). Much of invasion biology research therefore tends to focus on identifying the presence or absence of the particular phenotypes known to facilitate invasion success. However, invasiveness has been shown to vary even at the population scale, and such idiosyncrasies make studying invasions challenging (Arim et al. 2006).

Understanding the physical factors that underlie invasive potential can certainly lead to excellent outcomes if they can be exploited in mitigation plans. For example, when a small, isolated population of the invasive marine algae *Caulerpa taxifolia* was discovered in the Mediterranean Sea in 1984, its invasive potential was largely unknown.

This resulted in limited management action, with the algae then proliferating rapidly to cover approximately 200 km of coastline by the year 2000 (Meinesz et al. 2001). However, when later discovered in California, early detection and new knowledge of the species' phenotypic biology prompted stringent biosecurity measures within 17 days of detection, and the algae was declared eradicated in California in 2006 (Williams & Grosholz, 2008). Thus, using phenotypic information to identify mitigative targets for future invaders based on preventing or limiting traits like dispersal or reproduction is an important method in the biosecurity toolbox, with prevention being much less costly than management or eradication (Diagne et al. 2021). Indeed, much of invasive species research and management is motivated by economic cost – somewhat ironically, considering invasions are themselves by-products of economic activity, with strong correlations existing between a country's Gross Domestic Product (GDP) and the richness of its invasive species (Hulme, 2009).

However, recent research is demonstrating the role of the genome in invasions by identifying clear links between genomic changes and invasive potential (e.g., Pearce et al. 2017). A greater incorporation of genomic factors into biosecurity and invasion biology research therefore has the potential to achieve more effective genome-informed outcomes in management, as well as broader understanding of rapid evolutionary processes. This argument forms the basis of my research and I elaborate on the value of leveraging genomics data to the study of biological invasions in the next section.

1.4 Population genomics as a tool for study invasions

The genetic composition of invasive populations is dynamic over time and space, with several genetic factors potentially facilitating invasion success by acting at various stages of the invasion process. In particular, starting diversity of native lineages, duration and intensity of foundation-induced bottleneck events, levels of population connectivity (e.g., propagule pressure and post-invasion gene flow), and adaptive/plastic processes (Lee,

2002; Sherpa & Després, 2021) may each contribute to invasion outcomes. Meanwhile, specific genomic features, such as transposable elements, genetic diversity, genome size, allelic frequencies, and/or genes under selection are starting to be identified as important mechanistic components of invasion success (e.g., Pearce et al. 2017; Olazcuaga et al. 2020).

Historically, studying genes at the ecological and evolutionary scale has involved genetic markers such as microsatellites, mitochondrial and nuclear DNA, allozymes, and/or a small number of single nucleotide polymorphisms (SNPs; Lee et al. 2002). Recent advances in high-throughput sequencing and bioinformatic pipelines are revolutionising the field of biology, with the ability to produce high-quality, whole-genome datasets from multiple populations (McCartney et al. 2019). Population genomics – the application of genomic technologies to study populations of individuals – has the potential to benefit invasion biology by revealing the intrinsic genetic factors that underpin successful invasion at much higher resolution and accuracy than the more traditional markers (Chen et al. 2021). For example, Parvizi et al. (2022) used genome-wide polymorphisms from native, invasive, and intercepted populations of the highly invasive brown marmorated stink bug, *Halyomorpha halys*, to identify incursion pathways. These authors discovered high levels of gene flow and admixture among global populations, with multiple incursions from genetically diverse source populations likely contributing to its success in new habitats (Parvizi et al. 2022). Thus, population genomic tools can clearly inform and guide invasive species management, providing a valuable tool for implementation in biosecurity frameworks (e.g., Sjoden et al. 2020).

1.5 Thesis structure

My thesis aims to explore the use of population genomics in invasion biology – particularly for improving understanding of the evolutionary mechanisms embedded in the genetic composition of species that might contribute to their invasion success.

Following this Introductory chapter, **Chapter 2** investigates the extent to which population genetic and genomics data have been used to study invasive species in the literature, including the extent to which genomic data and resources (i.e., reference genomes) exist for a sample of one hundred of the world's worst invasive species. **Chapter 3** provides a case study to demonstrate how population genomic data can be used to understand the invasion of a major agricultural pest moth – pink bollworm (*Pectinophora gossypiella*) – in Australia. **Chapter 4** provides an overall discussion of my research, placing findings into a broader context and considering future steps. Finally, **Appendix 1** reviews how climate change might favour the establishment of non-native arthropods and affect the adaptive responses of polar arthropods.

1.6 Literature cited

Adrian-Kalchhauser, I., Blomberg, A., Larsson, T., Musilova, Z., Peart, C. R., Pippel, M. et al. (2020). The round goby genome provides insights into mechanisms that may facilitate biological invasions. *BMC Biology*, *18*, 11. https://doi.org/10.1186/s12915-019-0731-8.

Aikio, S., Duncan, R. P., & Hulme, P. E. (2010). Lag-phases in alien plant invasions: Separating the facts from the artefacts. *Oikos*, *119*, 370-378. https://doi.org/10.1111/j.1600-0706.2009.17963.x.

Allison, S. D., & Vitousek, P. M. (2004). Rapid nutrient cycling in leaf litter from invasive plants in Hawai'i. *Oecologia*, *141*, 612-619. https://doi.org/10.1007/s00442-004-1679-z.

Arim, M., Abades, S. R., Neill, P. E., Lima, M., & Marquet, P. A. (2006). Spread dynamics of invasive species. *Proceedings of the National Academy of Sciences*, *103*, 374-378. https://doi.org/10.1073/pnas.0504272102.

Ascensão, F., & Capinha, C. (2017). Aliens on the move: transportation networks and non-native species. In L. Borda-de-Água, R. Barrientos, P. Beja, & H. M. Pereira (Eds.), *Railway Ecology* (pp. 65-80). Springer International Publishing. https://doi.org/10.1007/978-3-319-57496-7_5.

Blackburn, T. M., Cassey, P., Duncan, R. P., Evans, K. L., & Gaston, K. J. (2004). Avian extinction and mammalian introductions on oceanic islands. *Science*, *305*, 1955-1958. https://doi.org/10.1126/science.1101617.

Blackburn, T. M., Lockwood, J. L., & Cassey, P. (2016). The influence of numbers on invasion success. In *Invasion Genetics* (pp. 25-39). https://doi.org/10.1002/9781119072799.ch2.

Blackburn, T. M., Pyšek, P., Bacher, S., Carlton, J. T., Duncan, R. P., Jarošík, V. et al. (2011). A proposed unified framework for biological invasions. *Trends in Ecology and Evolution, 26*, 333-339. https://doi.org/10.1016/j.tree.2011.03.023.

Butchart, S. H. M., Walpole, M., Collen, B., van Strien, A., Scharlemann, J. P. W., Almond, R. et al. (2010). Global biodiversity: Indicators of recent declines. *Science, 328*, 1164-1168. https://doi.org/10.1126/science.1187512.

Cassini, M. H. (2020). A review of the critics of invasion biology. *Biological Reviews, 95*, 1467-1478. https://doi.org/10.1111/brv.12624.

Chen, M., Mei, Y., Chen, X., Chen, X., Xiao, D., He, K. et al. (2021). A chromosome-level assembly of the harlequin ladybird *Harmonia axyridis* as a genomic resource to study beetle and invasion biology. *Molecular Ecology Resources, 21*, 1318-1332. https://doi.org/10.1111/1755-0998.13342.

Colautti, R. I., & MacIsaac, H. J. (2004). A neutral terminology to define 'invasive' species. *Diversity and Distributions, 10*, 135-141. https://doi.org/10.1111/j.1366-9516.2004.00061.x.

Courchamp, F., Fournier, A., Bellard, C., Bertelsmeier, C., Bonnaud, E., Jeschke, J. M., & Russell, J. C. (2017). Invasion biology: Specific problems and possible solutions. *Trends in Ecology and Evolution, 32*, 13-22. https://doi.org/10.1016/j.tree.2016.11.001.

Coutts, S. R., Helmstedt, K. J., & Bennett, J. R. (2018). Invasion lags: The stories we tell ourselves and our inability to infer process from pattern. *Diversity and Distributions, 24*, 244-251. https://doi.org/10.1111/ddi.12669.

Diagne, C., Leroy, B., Vaissière, A.-C., Gozlan, R. E., Roiz, D., Jarić, I. et al. (2021). High and rising economic costs of biological invasions worldwide. *Nature, 592*, 571-576. https://doi.org/10.1038/s41586-021-03405-6.

Doherty, T. S., Glen, A. S., Nimmo, D. G., Ritchie, E. G., & Dickman, C. R. (2016). Invasive predators and global biodiversity loss. *Proceedings of the National Academy of Sciences*, *113*, 11261-11265. https://doi.org/10.1073/pnas.1602480113.

Fautley, R., Coulson, T., & Savolainen, V. (2012). A comparative analysis of the factors promoting deer invasion. *Biological Invasions*, *14*, 2271-2281. https://doi.org/10.1007/s10530-012-0228-7.

Forsyth, D. M., Duncan, R. P., Bomford, M., & Moore, G. (2004). Climatic suitability, life-history traits, introduction effort, and the establishment and spread of introduced mammals in Australia. *Conservation Biology*, *18*, 557-569. https://doi.org/10.1111/j.1523-1739.2004.00423.x.

García, R. A., Engler, M. L., Peña, E., Pollnac, F. W., & Pauchard, A. (2015). Fuel characteristics of the invasive shrub *Teline monspessulana*. *International Journal of Wildland Fire*, *24*, 372-379. https://doi.org/10.1071/WF13078.

Grarock, K., Lindenmayer, D. B., Wood, J. T., & Tidemann, C. R. (2013). Using invasion process theory to enhance the understanding and management of introduced species: A case study reconstructing the invasion sequence of the common myna (*Acridotheres tristis*). *Journal of Environmental Management*, *129*, 398-409. https://doi.org/10.1016/j.jenvman.2013.08.005.

Guiaşu, R. C., & Tindale, C. W. (2018). Logical fallacies and invasion biology. *Biology & Philosophy*, *33*, 34. https://doi.org/10.1007/s10539-018-9644-0.

Hautier, Y., Tilman, D., Isbell, F., Seabloom, E. W., Borer, E. T., & Reich, P. B. (2015). Anthropogenic environmental changes affect ecosystem stability via biodiversity. *Science*, *348*, 336-340. https://doi.org/doi:10.1126/science.aaa1788.

Hulme, M. (2008). The conquering of climate: Discourses of fear and their dissolution. *The Geographical Journal*, *174*, 5-16. https://doi.org/10.1111/j.1475-4959.2008.00266.x.

Hulme, P. E. (2009). Trade, transport and trouble: Managing invasive species pathways in an era of globalization. *Journal of Applied Ecology, 46*, 10-18. https://doi.org/10.1111/j.1365-2664.2008.01600.x.

Lee, C. E. (2002). Evolutionary genetics of invasive species. *Trends in Ecology and Evolution, 17*, 386-391. https://doi.org/10.1016/S0169-5347(02)02554-5.

Linders, T. E. W., Schaffner, U., Eschen, R., Abebe, A., Choge, S. K., Nigatu, L. et al. (2019). Direct and indirect effects of invasive species: Biodiversity loss is a major mechanism by which an invasive tree affects ecosystem functioning. *Journal of Ecology*, *107*, 2660-2672. https://doi.org/10.1111/1365-2745.13268.

Mazza, G., Tricarico, E., Genovesi, P., & Gherardi, F. (2014). Biological invaders are threats to human health: an overview. *Ethology Ecology & Evolution*, *26*, 112-129. https://doi.org/10.1080/03949370.2013.863225.

McCartney, M. A., Mallez, S., & Gohl, D. M. (2019). Genome projects in invasion biology. *Conservation Genetics*, *20*, 1201-1222. https://doi.org/10.1007/s10592-019-01224-x.

Meinesz, A., Belsher, T., Thibaut, T., Antolic, B., Mustapha, K. B., Boudouresque, C.-F. et al. (2001). The introduced green alga *Caulerpa taxifolia* continues to spread in the Mediterranean. *Biological Invasions, 3*, 201-210. https://doi.org/10.1023/A:1014549500678.

Meyerson, L. A., & Mooney, H. A. (2007). Invasive alien species in an era of globalization. *Frontiers in Ecology and the Environment*, *5*, 199-208. https://doi.org/10.1890/1540-9295(2007)5[199:IASIAE]2.0.CO;2.

Mollot, G., Pantel, J. H., & Romanuk, T. N. (2017). Chapter two - the effects of invasive species on the decline in species richness: a global meta-analysis. In D. A. Bohan, A. J. Dumbrell, & F. Massol (Eds.), *Advances in Ecological Research* (Vol. 56, pp. 61-83). Academic Press. https://doi.org/10.1016/bs.aecr.2016.10.002.

Mooney, H. A., & Cleland, E. E. (2001). The evolutionary impact of invasive species. *Proceedings of the National Academy of Sciences*, *98*, 5446-5451. https://doi.org/10.1073/pnas.091093398.

Moran, E. V., & Alexander, J. M. (2014). Evolutionary responses to global change: Lessons from invasive species. *Ecology Letters*, *17*, 637-649. https://doi.org/10.1111/ele.12262.

Olazcuaga, L., Loiseau, A., Parrinello, H., Paris, M., Fraimout, A., Guedot, C., et al. (2020). A whole-genome scan for association with invasion success in the fruit fly *Drosophila suzukii* using contrasts of allele frequencies corrected for population structure. *Molecular Biology and Evolution, 37*, 2369-2385.

Paini, D. R., Sheppard, A. W., Cook, D. C., De Barro, P. J., Worner, S. P., & Thomas, M. B. (2016). Global threat to agriculture from invasive species. *Proceedings of the National Academy of Sciences*, *113*, 7575-7579. https://doi.org/10.1073/pnas.1602205113.

Parvizi, E., Dhami, M. K., Yan, J., & McGaughran, A. (2022). Population genomic insights into invasion success in a polyphagous agricultural pest, *Halyomorpha halys*. *Molecular Ecology*, https://doi.org/10.1111/mec.16740.

Pearce, S. L., Clarke, D. F., East, P. D., Elfekih, S., Gordon, K., Jermiin, L. S., McGaughran, A. et al. (2017). Genomic innovations, transcriptional plasticity and gene loss underlying the evolution and divergence of two highly polyphagous and invasive *Helicoverpa* pest species. *BMC biology*, *15*, 1-30. https://doi.org/10.1186/s12915-017-0402-6.

Pecl, G. T., Araújo, M. B., Bell, J. D., Blanchard, J., Bonebrake, T. C., Chen, I.-C. et al. (2017). Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. *Science*, *355*, eaai9214. https://doi.org/doi:10.1126/science.aai9214.

Richardson, D. M., & Ricciardi, A. (2013). Misleading criticisms of invasion science: A field guide. *Diversity and Distributions*, *19*, 1461-1467. https://doi.org/10.1111/ddi.12150.

Schlaepfer, M. A., Sax, D. F., & Olden, J. D. (2011). The potential conservation value of non-native species. *Conservation Biology*, *25*, 428-437. https://doi.org/10.1111/j.1523-1739.2010.01646.x.

Shackleton, R. T., Biggs, R., Richardson, D. M., & Larson, B. M. H. (2018). Social-ecological drivers and impacts of invasion-related regime shifts: consequences for ecosystem services and human wellbeing. *Environmental Science & Policy*, *89*, 300-314. https://doi.org/10.1016/j.envsci.2018.08.005.

Shackleton, R. T., Shackleton, C. M., & Kull, C. A. (2019). The role of invasive alien species in shaping local livelihoods and human well-being: A review. *Journal of Environmental Management*, *229*, 145-157. https://doi.org/10.1016/j.jenvman.2018.05.007.

Sherpa, S., Blum, M. G. B., & Després, L. (2019). Cold adaptation in the Asian tiger mosquito's native range precedes its invasion success in temperate regions *Evolution*, *73*, 1793-1808. https://doi.org/10.1111/evo.13801.

Sherpa, S., & Després, L. (2021). The evolutionary dynamics of biological invasions: A multi-approach perspective. *Evolutionary Applications*, *14*, 1463-1484. https://doi.org/10.1111/eva.13215.

Sjodin, B. M. F., Irvine, R. L., Ford, A. T., Howald, G. R., & Russello, M. A. (2020). *Rattus* population genomics across the Haida Gwaii archipelago provides a framework for guiding invasive species. *Evolutionary Applications*, *13*, 889-904. https://doi.org/10.1111/eva.12907.

Spear, M. J., Walsh, J. R., Ricciardi, A., & Zanden, M. J. V. (2021). The invasion ecology of sleeper populations: prevalence, persistence, and abrupt shifts. *BioScience*, *71*, 357-369. https://doi.org/10.1093/biosci/biaa168.

Tobin, P. C. (2018). Managing invasive species. *F1000Res*, *7*. https://doi.org/10.12688/f1000research.15414.1.

Vitousek, P. M., D'Antonio, C. M., Loope, L. L., RejmÁNek, M., & Westbrooks, R. (1997). Introduced species: A significant component of human-caused global change. *New Zealand Journal of Ecology*, *21*, 1-16. http://www.jstor.org/stable/24054520.

Wainright, C. A., Muhlfeld, C. C., Elser, J. J., Bourret, S. L., & Devlin, S. P. (2021). Species invasion progressively disrupts the trophic structure of native food webs. *Proceedings of the National Academy of Sciences*, *118*, e2102179118. https://doi.org/10.1073/pnas.2102179118.

Williams, S. L., & Grosholz, E. D. (2008). The invasive species challenge in estuarine and coastal environments: marrying management and science. *Estuaries and Coasts*, *31*, 3-20. https://doi.org/10.1007/s12237-007-9031-6.

Chapter 2.

Genomic data is missing for many highly invasive species, restricting our preparedness for escalating incursion rates

Images obtained with Creative Commons license

Published as: Matheson, P. & McGaughran, A. (2022). Genomic data is missing for many highly invasive species, restricting our preparedness for escalating incursion rates. *Scientific Reports, 12, 13987*. https://doi.org/10.1038/s41598-022-17937-y.

Contributions: A.M conceived the research project, P.M obtained and analysed the data, and A.M and P.M co-wrote the manuscript.

2.1 Abstract

Biological invasions drive environmental change, potentially threatening native biodiversity, human health, and global economies. Population genomics is an increasingly popular tool in invasion biology, improving accuracy and providing new insights into the genetic factors that underpin invasion success compared to research based on a small number of genetic loci. We examine the extent to which population genomic resources, including reference genomes, have been used or are available for invasive species research. We find that 82% of species on the International Union for Conservation of Nature "100 Worst Invasive Alien Species" list have been studied using some form of population genetic data, but just 32% of these species have been studied using population genomic data. Further, 55% of the list's species lack a reference genome. With incursion rates escalating globally, understanding how genome-driven processes facilitate invasion is critical, but despite a promising trend of increasing uptake, "invasion genomics" is still in its infancy. We discuss how population genomic data can enhance our understanding of biological invasion and inform proactive detection and management of invasive species, and we call for more research that specifically targets this area.

2.1 Introduction

Anthropogenic activities, such as global trade and transport over recent decades, have strongly shaped the geographic scope, frequency, and taxonomic trends of species movement beyond their natural ranges (Early et al. 2016). Many species introductions have historically had benign, relatively small, or even beneficial impacts – for example, providing habitat or food resources to native species or crucial ecosystem functions (Tobin, 2018). However, some introduced species have the potential to become 'invasive' – that is, expand demographically and spatially and impose negative consequences in their new environment (Kumschick & Richardson, 2013). Invasive species are pervasive drivers of global change (Early et al. 2016; Thompson et al. 2021), potentially altering ecosystem function (Chown et al. 2015), introducing new ecological pressures (Doherty et al. 2016), and leading to genotype loss (e.g., via hybridisation; Neill & Arim 2019) or blurred regional distinctiveness of native biota (Kumschick & Richardson, 2013), population decline, or extinction of indigenous species (Ficetola et al. 2008). The total global reported cost of invasions between 1970-2017 was estimated to be at least US$1.288 trillion (Diagne et al. 2021). However, the true economic effects of invasive species are difficult to quantify due to their indirect impacts (e.g., reduced plant cover, increased soil erosion, increased eutrophication; Neill & Arim, 2019).

As a leading driver of environmental change, increasing attention has been dedicated towards a unified understanding of the invasion process (e.g., Blackburn et al. 2011; Novoa et al. 2020). Despite this, there are significant gaps in our mechanistic conception of invasion, as well as our ability to quantify and forecast impacts caused by invasive species (Lenzner et al. 2019; Novoa et al. 2020). Yet, rates of biological incursion are increasing with international trade and climate change (Chown et al. 2015; Hulme 2021): 37% of all globally established invasive species in the last 200 years are estimated to have been introduced after 1970 (Seebens et al. 2017), and established

invasive numbers per continent are predicted to increase by 36% between 2005 and 2050 (Seebens et al. 2021).

Though pre-existing traits, such as broad physical tolerance and high rates of dispersal, reproduction, and growth (Diez et al. 2012; Bellard et al. 2013) are key parameters of successful invasion, certain intrinsic genetic features may make for more successful invaders (Suda et. al. 2015; Brandies et al. 2019). Emerging research is identifying links between invasive potential and genomic changes (e.g., Colautti & Lau, 2015; Wagner et al. 2017; Ochocki & Miller, 2017), with alterations to gene expression, gene interaction, or genomic architecture potentially leading to a greater diet breadth (e.g. Pearce et al. 2017), competitive advantage (e.g., Adrian-Kalchhauser et al. 2020), and/or adaptive response to environmental change (e.g., Suda et al. 2015). More generally, genetic, demographic, and environmental factors interact to determine invasion success, and understanding of genetic characteristics, such as pre-adaptation and population connectivity (see Box 1 in Sherpa & Després, 2021), is crucial for monitoring, managing, and mitigating the impact of invasive species.

Invasion ecology has benefited from a union with population genetic approaches (i.e., 'invasion genetics') for over 55 years (Baker & Stebbins 1965). This has resulted in broad understanding of the evolutionary processes associated with invasion, such as the general effects of bottlenecks and genetic drift on invasion success and the specific adaptive responses of some invasive species (Burgess et al. 2021). However, much invasive biology research still suffers from a lack of information around complex processes operating at the genomic level (Bock et al. 2015; Neinavaie et al. 2021). Moving from a 'genetic' (single or few loci) lens to a genome-wide ('genomic') one can improve analytical accuracy in some scenarios (North et al. 2021). For example, the ability of mitochondrial DNA (mtDNA) to track recent invasions can be limited as this marker accumulates variation over longer timescales – for invasive mammals in

particular, mtDNA can incorrectly identify an invasive populations' country of origin compared to higher resolution genome-wide markers (Browett et al. 2020). Such was the case for raccoons (*Procyon lotor*), which show low mtDNA variation in their invasive European range (Frantz et al. 2013), and brown rats (*Rattus norvegicus*) that have invaded New Zealand and show a European origin with mtDNA, but an admixed Asian and non-Asian ancestry using genome-wide markers (Puckett et al. 2016). In other contexts, genomic data can allow new questions to be addressed that are intractable with a small number of loci. For example, genome-wide scans in invasive populations of *Drosophila suzukii* and monkeyflower *(Mimulus guttatus*) have identified new genes that are associated with invasion routes and stress adaptation during invasion, respectively (Olazcuaga et al. 2020; Puzey & Vallejo-Marin, 2014). With respect to management, next generation sequencing technologies can facilitate proactive *community*-wide detection and identification, and ongoing monitoring programmes (e.g., in aquatic systems; Deiner et al. 2017); it can also provide targeted frameworks for eradication plans by revealing crucial information, such as dispersal patterns and population connectivity (e.g., Sjodin et al. 2019).

Recent advances in sequencing, and associated downstream analytical approaches, are thus cementing the link between genomics and invasion biology (North et al. 2021) in the new field of 'invasion genomics'. Population genomics in particular, involves the analysis of genomic patterns within and among populations to make evolutionary inferences (De Wit et al. 2015). Associated high-throughput sequencing of entire genomes or genome-wide SNPs (single nucleotide polymorphisms) for multiple individuals and populations of interest is facilitating research into population structure, demographic history, and selective processes (Rius et al. 2015; McCartney et al. 2019). In an invasive context, population genomics can be used to provide greater insights than genetic studies based on a small number of loci by accurately identifying source or high

risk populations, pinpointing genomic weaknesses, studying demo-genetic factors involved in the invasion process (e.g., genetic bottlenecks, founder effects), and examining particular 'invasive' genes and their roles in rapid evolution (Hohenlohe et al. 2010; Sherpa & Després, 2021). Meanwhile, complete genomic sequences, i.e., 'reference genomes', provide the basis for within- and between-species insights (such as the genomic architecture of important phenotypic traits; Brandies et al. 2019), and support the development of new technologies that may be applied to pest management (e.g., targeted SNP panels or gene drives; Zoonomia Consortium, 2020).

Minimising the impact of invasive species in the Anthropocene will require a strong emphasis on proaction and prevention, and genomic data can be leveraged to support this. Reviews by Rius et al. (2015) and McCartney et al. (2019) investigated the use of next-generation sequencing techniques to study invasive species, and documented the availability of genome assemblies for species from the International Union for Conservation of Nature (IUCN) "100 of the World's Worst Invasive Alien Species" list ('WAS List', hereafter), respectively. Here, we investigate the extent to which *population genomic* data has been used or is available to study globally invasive species from the WAS List, and provide an update on how many of these species currently have assembled reference genomes. We also analyse data in a *population genetic* context, determining the number of species that have not been analysed using any form of genetic marker and marking the shift between genetic and genomic studies. We begin by illustrating a promising trend of increasing uptake of genomic research for invasive species generally before showing that, despite this, the majority of such research for WAS List species has lacked a population genomics context and genomic resources are still entirely absent for many of these species. These discouraging gaps must be addressed if we are to prepare for escalating rates of biological invasion in the future.

2.2 Methods

IUCN "100 of the World's Worst Invasive Alien Species"

The International Union for Conservation of Nature (IUCN) is an organisation of governments, civil society organisations, and experts perhaps best known for publishing the 'Red List of Threatened Species', which provides a comprehensive index of the conservation status of species worldwide and their associated risk of extinction. The Invasive Species Specialist Group (ISSG) is a network of experts and policy makers organised under IUCN that aims to increase awareness of invasive species and their impact on the environment, as well as blueprint prevention, management, and/or eradication plans (Pagad et al. 2015). The Global Invasive Species Database (GISD) is a product of the ISSG, developed by Clout & Lowe (1996) to aid the early detection and management of invasive species in developing countries. The WAS List was first published in 2000 for both scientific and communication purposes (e.g., Bellard et al. 2013; Courchamp, 2013). Species on the list are chosen based on their impact on biodiversity and human activities, as well as their illustration of issues surrounding biological invasion and representation of a diverse selection of taxonomic groups, from microorganisms to plants and vertebrates (Luque et al. 2014).

Database searches

Web of Science and PubMed searches were performed (May 2022) to examine: the uptake of 'population genetics' and 'population genomics' analysis in an invasion biology context, and the degree to which population genetic, genomic, and/or reference genome resources exist for each of the WAS List species.

In the first search, the terms `("population genetic*" OR "next generation sequencing" OR "SNP*" OR "single nucleotide polymorphism*" OR allozyme* OR AFLP* OR microsatellite* OR mtDNA OR "mitochond* DNA" OR "nuclear`

DNA") AND ("invasive" OR "weed" OR "pest") AND ("animal*" OR "species" OR "organism*") were applied to titles, abstracts, and author keywords in the Web of Science and PubMed databases, yielding a total of 3,276 results. Publication years for each search were obtained using the Web of Science 'analyse results' tool. To identify differences between population genetic and population genomic trends though time, this was followed by a second search using the terms ("population genomic*" OR "next generation sequencing" OR "SNP*" OR "single nucleotide polymorphism*") AND ("invasive" OR "weed" OR "pest") AND ("animal*" OR "species" OR "organism*"), which returned 779 results.

In a separate search across both databases, keywords for each species associated with the WAS List were used to establish whether: (a) population genetic; and (b) population genomic data was available for inferring evolutionary patterns and processes. The keyword string used for each species and search was: (a) ("common name*" OR "species name") AND ("population genetic*" OR "next generation sequencing" OR "SNP*" OR "single nucleotide polymorphism*" OR allozyme* OR AFLP* OR microsatellite* OR mtDNA OR "mitochond* DNA" OR "nuclear DNA") and (b) ("common name*" OR "species name") AND ("population genom*" OR "next generation sequencing" OR "SNP*" OR "single nucleotide polymorphism*"); and titles, abstracts, and author keywords were searched in each case. For (a), this search yielded 0-535 results per species, and 4,399 articles were retrieved overall. For (b), the search yielded from 0-258 results per species, and 1,217 total articles were retrieved. The relevance of each document for these searches was determined based on a screening of the abstract, resulting in the removal of articles that did not contain: samples from wild individuals, samples from different populations, and for (b) samples that lacked a focus on genome-wide data. For (a) and (b), if at least one abstract contained data and terminology relevant to population genetics or genomics (e.g., population structure, gene flow/genetic drift, genetic diversity, phylogeography), then the species

was considered ‘positive’ for either data type and was further examined and scored for invasive context – in this case, each study was evaluated and scored for its *dominant* research focus: the history or route of incursion, the demography of the invading population, or the evolution of invasiveness. For both (a) and (b), metrics such as year of publishing, origin country of the research organisations affiliated with each author, and publication availability (i.e., open access status) were collected for each species using the Web of Science ‘analyse results’ tool.

NCBI searches

The National Centre for Biotechnology Information (NCBI) database was used to track whether each species on the WAS List had a publicly-available reference genome associated with it. Although there are likely other public repositories for genomic data, NCBI contains the largest bank of molecular biological and genetic data available and its genome database contains the most up to date sequence and mapping data for a range of organisms (Benson et al. 2010); as a result, we feel it best captures the most publicly accessible genome data available. In March, 2022 the scientific name of each of the 100 species was entered into the search bar of the NCBI website (https://www.ncbi.nlm.nih.gov/) with the database category set to ‘genome’. If the resulting search indicated that there was a reference genome, that species was recorded as ‘positive’ for this data type.

2.3 Results

Genomics of invasive species research is escalating

We searched across two academic databases to examine the number of published articles that target genomics, population genetics, and/or population genomics of invasive species. We found that publications utilising genetic markers (e.g., mtDNA, microsatellites, allozymes, amplified fragment length polymorphisms/AFLPs and/or

small numbers of SNPs) largely dominate invasive biology research (n = 3,128), despite an increasing focus towards population genomics over the last ~20 years (Fig. 2.1). In 2011, genomics-based research made up just 9% (15 out of 165) of population studies on invasive species conducted that year; ten years later this figure had increased to 31% (116 out of 378). In fact, 15% (116 out of 779) of articles targeting population genomics of invasive species were published in 2021 alone.

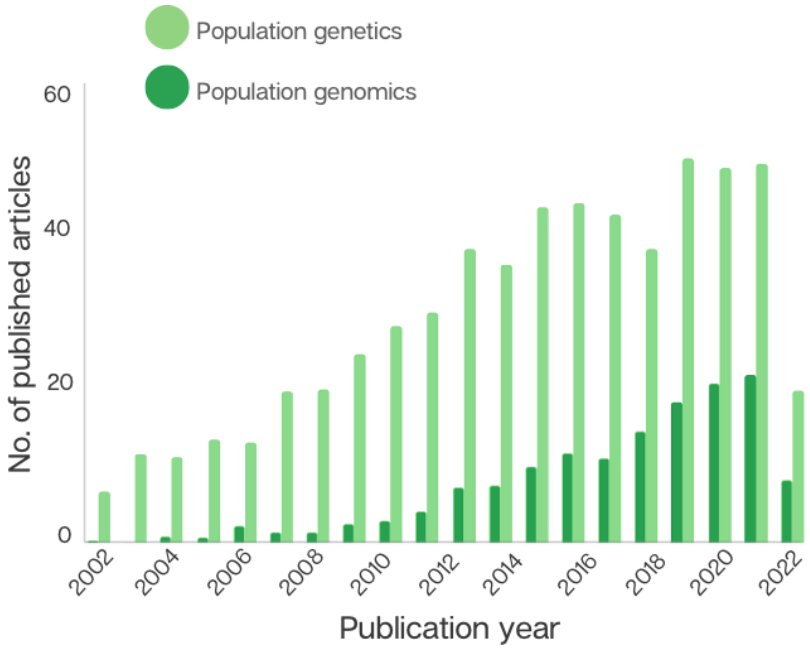


Figure 2.1 Number of published articles that apply "population genomics", "next generation sequencing", or "SNPs" (dark green) and "population genetics", "allozyme", "mitochondrial DNA", "nuclear DNA", or "microsatellites" (light green) in an invasive context over time.

Population genetic data is largely available for invasives, but limited in scope

As of May 2022, 82% of the WAS List species had been examined using some form of population genetic data (Fig. 2.2). Of 807 retrieved studies, at least one publication utilised genetic data in an invasive context for 74 of the examined species (90%). These invasion-focused population genetic studies dominantly targeted the history/routes of incursion (51%) and the demography of colonising populations (39%), while the

evolution of invasiveness has only rarely been examined using population genetic data (10%; Fig. 2.2).

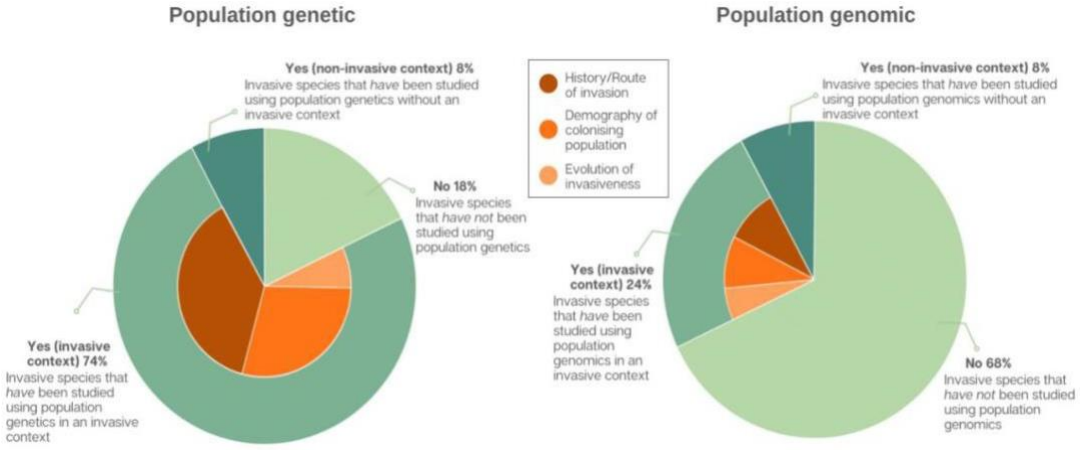


Figure 2.2 Proportion of invasive species from the WAS List (n = 100) for which researchers have utilised population genetic and population genomic data as a research tool.

Population genomic data for invasives is predominantly absent

Despite the encouraging pattern outlined above, our specific analysis of WAS List species found that only 32% of these species had at least one publication that utilised population genomic data as a research tool. Thus, roughly two-thirds of globally important, highly invasive species on the WAS List currently lack publicly-available population genomic data (Fig. 2.2). Of the 32% of species for which population genomic data is available, this data has been applied in an invasive context for the majority (75%), though this represents a total of just 24 of the 100 listed species. These population genomics-focused studies predominantly targeted the history/routes of incursion and the demography of colonising populations in similar proportions (~38%), while the evolution of invasiveness has received the least focus (24%; Fig 2.2).

The study of invasives is subject to a limited geographical distribution of resources

We extracted the origin country of research organisations affiliated with the authors of each publication to investigate the likely geographical distribution of invasive genomic resources (e.g., tools and funding). Of the 809 articles that used population genetic data to study invasive species, author country of origin records (n = 1,140) indicated that the top five countries that utilising this tool are high income countries: United States (n = 242), France (n = 118), Australia (n = 71), Germany (n = 70), and Spain (n = 67). While a smaller number of publications had author country of origins from low-income countries, there were never more than 10 (i.e., < 1% of the total) publications per country. The United States also dominated the author country of origin records for the 91 articles that included population genomic data as a research tool to study invasive species, making up 49 of the 239 total records. In the top ten author countries of origin, Australia (n = 12) was again the only country in the Southern Hemisphere represented, and no countries from Africa were present (Table A2.1B)

Meanwhile, the majority (86%) of the total articles that were identified as having a population genomic context were published within an open access framework. This contrasts with our population genetic analysis, where less than half (41%) of the articles were open access. There was no significant relationship between geography and a presence or lack of open access publishing for either population genetic or population genomic publications (Table A2.1).

Invasive species commonly lack reference genomes

We examined the National Centre for Biotechnology Information (NCBI) database and found that 45% of the WAS List species had a publicly-available reference genome (Fig. 3A). The WAS List is largely dominated by plant species (n = 37), followed by invertebrates (n = 26), and mammals (n = 14) (outer ring, Fig. 2.3B). However, mammals are disproportionately over-represented in terms of available reference genomes (78.6%).

Plants are conversely under-represented, with ~89% of the WAS List species lacking genomic resources. Meanwhile, two of the three birds from the list lack reference genomes entirely, as do half of the list's 26 invertebrate species (Fig. 2.3B).

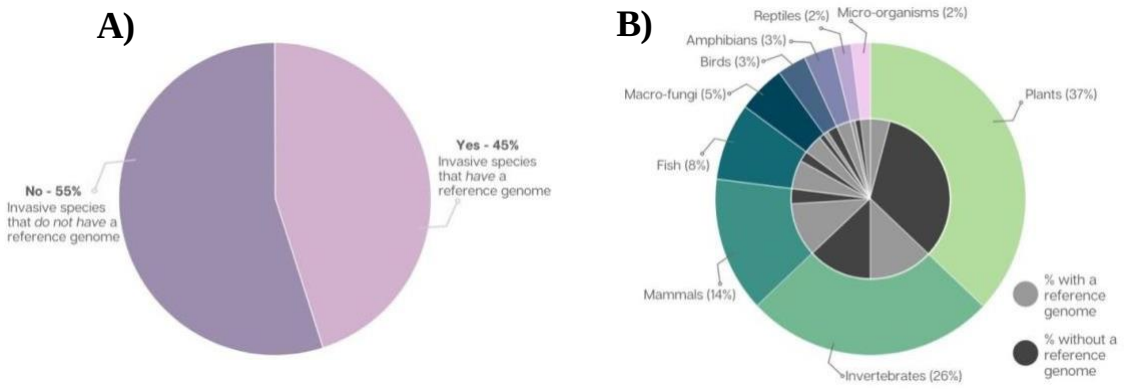


Figure 2.3 Proportion of invasive species from the WAS List (n = 100): (A) for which researchers have deposited a reference genome to NCBI; and (B) that correspond to the indicated taxonomic groups (outer circle), the associated proportion for which have (light grey) or lack (dark grey) reference genomes (inner circle).

2.4 Discussion

We investigated the extent to which population genetic and genomic data have been used to study globally invasive species from the WAS List. We found that genetic data, as opposed to genomic data, is used more widely as a tool to study population dynamics of invasive species, though this is mainly limited to elucidating invasion history, such as identifying routes of colonisation and source populations. Despite this, we found that publications relating generally to biological invasions and genomics (including population genomics) are gaining momentum in the literature, showing an increasing trajectory over the past ~20 years that aligns well with the reducing costs of next generation sequencing. However, only 32% of species on the WAS List have currently been studied in a population genomics context, and there is a large depauparacy (55%) of

reference genomes that are available for these species in the commonly-used NCBI genome repository.

Recent studies exemplify the value of population genomic resources as tools for informing, monitoring, and managing biological invasions (North et al. 2021). For example, whole genome scans of the predatory Northern snakehead fish, *Channa argus*, were used to identify the source population of invasions in parts of the United States for future prohibition of accidental and deliberate introductions (Resh et al. 2021), and whole genome resequencing data has led to more targeted management of glyphosate resistance in populations of the weed, *Amaranthus tuberculatus* (Kreiner et al. 2019). Despite this, Rius et al. (2015) found that just 33% of 117 published studies applying next-generation sequencing to invasive species between 2008 and 2015 had an invasive context. Similarly, we found here that only 32% of species on the WAS List have been studied in a population genomics context, though we note that the list of Rius et al. (2015) included only 13 species from the WAS List. Within the 32% of species on the WAS List that had been studied in a population genomics context in our study, an encouraging majority (n = 24; 75%) focused on invasive objectives, however the least targeted aspect of invasion biology in these studies was the evolution of invasiveness. Although genome sequencing has been used in only a handful of invasive studies and a small number of organisms to date (McCartney et al. 2019; North et al. 2021), there is accumulating evidence that genetic changes contribute to invasion success (Wagner et al. 2017; Welles & Dlugosch, 2019), so the underutilisation of population genomics for detecting the genomic architecture of invasion is disappointing. Comparing genomic divergence between invasive and native-range populations of the same species in particular holds great promise for elucidating and predicting the role of the genome in biological invasion (Rius et al. 2015) – an area that is clearly still yet to gain great traction.

Despite a variety of genome-generating initiatives (e.g. the Earth Biogenome Project: https://www.earthbiogenome.org/), under half (45%) of species on the WAS List currently have accessible reference genomes. However, this represents a sizable increase in the last three years, with McCartney et al. (2019) identifying 27/100 of species on the same WAS List as having reference genomes – a promising result on the surface that may indicate a rapidly growing investment in genomic resources. (This parallels the trend seen for the IUCN threatened species list, for which published genomes were available for 2.4% of the total 15,521 listed species as at January 2022 - an increase from 0.8% in 2018; Hogg et al. 2020). However, reference genomes may be assembled for invasive species in a nevertheless non-invasive context, e.g., the species may have high economic value, or high merit as a research model. Indeed, just 13 of the 27 species in McCartney et al. (2019) that had a reference genome had invasive status as an *a priori* rationale for genome assembly. In our case, species such as *Sus scrofa* (pig), *Oncorhynchus mykiss* (rainbow trout), and *Mus musculus* (field mouse) returned hundreds of documents in the Web of Science literature searches, however, very few of these were relevant or translatable to invasion.

The lack of widespread application of genomic resources to invasion (and other) biology that we detect here is undoubtedly driven by the associated costs of generating such data. Financial and computational burdens, together with the required time and expertise, continue to place limits on the breadth and depth of genomic studies (Brandies et al. 2019), despite progress in technology, analysis pipelines, and bioinformatics training. Fortunately, many important questions in invasion biology can be addressed with fewer genetic markers and our population genetic results indicate that, although genomic approaches are superior in some instances, individual markers, such as mtDNA, still have an important ongoing role to play in invasive species research. However, a lack of equity in this space may explain our finding that most authors of invasive species

research are predominantly based in high-income countries, such as the United States and countries in Europe, rather than in locations in Africa or the Southern Hemisphere (Appendix 2; Table A2.1) – irrespective of whether the data was population genetic or genomic in nature (though we noted a slight increase in representation of lower-income author countries of origin in the population genetic versus genomic records, this never exceeded 1% of the total publications for these countries).

Mammals make up just 14% of the WAS List, including familiar species such as red deer, domestic cats, and stoats. However, they constitute roughly a quarter of the species that have a reference genome. Plants show the converse pattern, making up 37% of the WAS List but having a reference genome for only four species. These findings do not reflect the relative impacts of each invasive group (e.g., Panda et al. 2018; Kumar Rai & Singh 2020); rather taxon-specific idiosyncrasies likely play a role for some groups. For example, ploidy in plants can increase the complexity and challenge of genomic analysis compared to other taxonomic groups (Schatz et al. 2012) and amphibians also have large and highly heterozygous genomes (Sun et al. 2020). Fortunately, recent technological advancements, especially relating to long-read sequencing, are making genomic research more accessible and accurate, particularly for organisms with large and complex genomes (Marks et al. 2021).

Of course, it is possible to learn a great deal about a species' evolutionary properties without the use of a reference genome (Díaz-Arce & Rodríguez-Ezpeleta, 2019). Such approaches are particularly useful for studying non-model species, but reduced genome complexity and potentially high degrees of missing data (Wickland et al. 2017; Díaz-Arce & Rodríguez-Ezpeleta, 2019) make them unsuitable for addressing certain study questions (e.g., genomic rearrangements; Worley et al. 2017), while access to a reference genome can make answering other questions more efficient (McCartney et al. 2019). The lack of reference genomes identified here limits the resolution of genomic studies

available for invasive species on the WAS List. However, several recent initiatives aim to sequence genomes of pests and/or pathogens (e.g., Ag100Pest Initiative: http://i5k.github.io/ag100pest; Plant Pathogen 'Omics Initiative: https://bioplatforms.com/projects/plant-pathogen-omics/) and we argue that more funding, effort, and expertise should be allocated to such projects, particularly for the taxa that we have identified as having received little research attention, such as plants.

The limited taxonomic scope of invasive species from the WAS List that have received population genomic attention to date likely represents a broader limiting of evolutionary understanding of invasive species that is required to predict and prevent future incursions. Generally, the incorporation of population genetic research into policy decisions is becoming more widely adopted – particularly in its use for identifying invasion routes and clarifying taxonomic uncertainties prior to management (Searle 2008; Le Roux & Wieczorek, 2009; Browett et al. 2020). However, incorporation of population genomic data into such policy has been minimal (a similarly slow translation of population genomics findings to applied wildlife conservation is also common; Hohenlohe et al. 2020) despite its clear advantage over genetic data in many scenarios, as outlined here. As invasions are predicted to increase in frequency and magnitude with climate change, the implications of this will affect pest management at a global scale and, although the highest number of invasive species are found in developed nations, their threat to developing nations, where there are less resources available for invasion management, is much higher (Early et al. 2016).

Population genomic data and methods are revolutionising the field of biology and have the potential to change the way we study invasive organisms and accelerate the pace at which we can ultimately apply genomic resources to a policy and management setting. However, while genomics and population genomics are gaining momentum in invasive species research, there is much to be done. First, reference genomes need to be assembled

and made publicly available for the vast proportion of invasive species that lack them, including those on the WAS List – we need to see more, targeted 'invasomics' reference genome initiatives. Second, more research should target population genomic analysis of invasive species, allowing for a greater understanding of the demo-genetic factors and intrinsic genetic mechanisms that lead to invasion success. This will aid in the development of proactive responses against invasive species that take a genome-informed approach to exploit specific species weaknesses to prevent their spread and limit their impact. Third, much of this research is cutting edge and, although 68% of species from the WAS List are yet to be studied with population genomic methods, over half of those that have been were published in the last five years. Further genomic uptake in this space should be maintained to ensure that genomic insights into invasive species continue at a pace that meets the escalating demands imposed by future climate change. In conjunction, the accessible nature of at least some of the population-based genomic data that has not currently been applied in a population genomic context could be retrospectively analysed with appropriate bioinformatic techniques to address invasive questions.

2.5 Literature cited

Adrian-Kalchhauser, I., Blomberg, A., Larsson, T., Musilova, Z., Peart, C. R., Pippel, M., et al. (2020). The round goby genome provides insights into mechanisms that may facilitate biological invasions. *BMC Biology*, *18*, 11. https://doi.org/10.1186/s12915-019-0731-8.

Baker, H. G., & Stebbins, G. L. (1965). The genetics of colonizing species. New York: Academic Press.

Bellard, C., Thuiller, W., Leroy, B., Genovesi, P., Bakkenes, M., & Courchamp, F. (2013). Will climate change promote future invasions? *Global Change Biology*, *19*, 3740-3748. https://doi.org/10.1111/gcb.12344.

Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2010). GenBank. *Nucleic Acids Research*, *38*, 46-51. https://doi.org/10.1093/nar/gkp1024.

Blackburn, T. M., Pyšek, P., Bacher, S., Carlton, J. T., Duncan, R. P., Jarošík, V., Wilson, J. R. U., & Richardson, D. M. (2011). A proposed unified framework for biological invasions. *Trends in Ecology and Evolution*, *26,* 333-339. https://doi.org/10.1016/j.tree.2011.03.023.

Bock, D. G., Caseys, C., Cousens, R. D., Hahn, M. A., Heredia, S. M., Hübner, S., et al. (2015). What we still don't know about invasion genetics. *Molecular Ecology, 24*, 2277–2297. https://doi.org/10.1111/mec.13032.

Brandies, P., Peel, E., Hogg, C. J., & Belov, K. (2019). The value of reference genomes in the conservation of threatened species. *Genes*, *10*, 846. https://doi.org/10.3390/genes10110846.

Browett, S. S., O'Meara, D. B., & McDevitt, A. D. (2020). Genetic tools in the management of invasive mammals: Recent trends and future perspectives. *Mammalian Review, 50,* 200–210. https://doi.org/10.1111/mam.12189.

Burgess, B. T., Irvine, R. L., Howald, G. R., & Russello, M. A. (2021). The promise of genetics and genomics for improving invasive mammal management on islands. *Frontiers in Ecology and Evolution, 9,* https://doi.org/10.3389/fevo.2021.704809.

Chown, S. L., Hodgins, K. A., Griffin, P. C., Oakeshott, J. G., Byrne, M., & Hoffmann, A. A. (2015). Biological invasions, climate change and genomics. *Evolutionary Applications*, *8*, 23-46. https://doi.org/10.1111/eva.12234.

Colautti, R. I., & Lau, J. A. (2015). Contemporary evolution during invasion: Evidence for differentiation, natural selection, and local adaptation. *Molecular Ecology*, *24*, 1999-2017. https://doi.org/10.1111/mec.13162.

Courchamp, F. (2013). Monster fern makes IUCN invader list. *Nature*, *498*, 37-37. https://doi.org/10.1038/498037a.

De Wit, P., Pespeni, M. H., & Palumbi, S. R. (2015). SNP genotyping and population genomics from expressed sequences – current advances and future possibilities. *Molecular Ecology*, *24*, 2310-2323. https://doi.org/10.1111/mec.13165.

Deiner, K., Bik, H. M., Elvira, M., Seymour, M., Lacoursiére-Roussel, A., Altermatt, F., et al. (2017) Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology, 26,* 5872–5895. https://doi.org/10.1111/mec.14350.

Diagne, C., Leroy, B., Vaissière, A. C., Gozlan, R. E., Roiz, D., Jarić, I., et al. (2021). High and rising economic costs of biological invasions worldwide. *Nature, 592,* 571-576. https://doi.org/10.1038/s41586-021-03405-6.

Díaz-Arce, N. & Rodríguez-Ezpeleta, N. (2019). Selecting RAD-Seq data analysis parameters for population genetics: The more the better? *Frontiers in Genetics*, *10,* 533. https://doi.org/10.3389/fgene.2019.00533.

Diez, J. M., D'Antonio, C. M., Dukes, J. S., Grosholz, E. D., Olden, J. D., Sorte, C. J., et al. (2012). Will extreme climatic events facilitate biological invasions? *Frontiers in Ecology and the Environment, 10*, 249-257. https://doi.org/10.1890/110137.

Doherty, T. S., Glen, A. S., Nimmo, D. G., Ritchie, E. G., & Dickman, C. R. (2016). Invasive predators and global biodiversity loss. *Proceedings of the National Academy of Sciences*, 113, 11261. https://doi.org/10.1073/pnas.1602480113.

Early, R., Bradley, B. A., Dukes, J. S., Lawler, J. J., Olden, J. D., Blumenthal, D. M., et al. (2016). Global threats from invasive alien species in the twenty-first century and national response capacities. *Nature Communications, 7*, 12485. https://doi.org/10.1038/ncomms12485.

Ficetola, G. F., Bonin, A., & Miaud, C. (2008). Population genetics reveals origin and number of founders in a biological invasion. *Molecular Ecology*, *17*, 773-782. https://doi.org/10.1111/j.1365-294X.2007.03622.x.

Frantz, A. C., Heddergott, M., Lang, J., Schulze, C., Ansorge, H., Runge, M., et al. (2013) Limited mitochondrial DNA diversity is indicative of a small number of founders of the German raccoon (*Procyon lotor*) population. *European Journal of Wildlife Research, 59*, 665–674. https://doi.org/10.1007/s10344-013-0719-6.

Hogg, C. J., Ottewell, K., Latch, P., Rossetto, M., Biggs, J., Gilbert, A., et al. (2022). Threatened species initiative: Empowering conservation action using genomic resources. *Proceedings of the National Academy of Sciences, 119*, e2115643118. https://doi.org/10.1073/pnas.2115643118.

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., & Cresko, W. A. (2010). Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics, 6*, e1000862. https://doi.org/10.1371/journal.pgen.1000862.

Hohenlohe, P. A., Funk, W. C. & Rajora, O. P. (2021) Population genomics for wildlife conservation and management. *Molecular Ecology, 30,* 62-82. https://doi.org/10.1111/mec.15720.

Hulme, P. E. (2021) Unwelcome exchange: International trade as a direct and indirect driver of biological invasions worldwide. *One Earth, 4*, 666-679. https://doi.org/10.016/j.oneear.2021.04.015.

Kreiner, J. M., Giacomini, D. A., Bemm, F., Waithaka, B., Regalado, J., Lanz, C., et al. (2019). Multiple modes of convergent adaptation in the spread of glyphosate-resistant *Amaranthus tuberculatus*. *Proceedings of the National Academy of Sciences*, *116*, 21076-21084. https://doi.org/10.1073/pnas.1900870116.

Kumar Rai, P., & Singh, J. S. (2020). Invasive alien plant species: Their impact on environment, ecosystem services and human health. *Ecological Indicators*, *111*, 106020. https://doi.org/10.1016/j.ecolind.2019.106020.

Kumschick, S., & Richardson, D. M. (2013). Species-based risk assessments for biological invasions: Advances and challenges. *Diversity and Distributions*, *19*, 1095-1105. https://doi.org/10.1111/ddi.12110.

Le Roux, J., & Wieczorek, A. M. (2009). Molecular systematics and population genetics of biological invasions: Towards a better understanding of invasive species management: Systematics and population genetics of biological invasions. *Annals of Applied Biology, 154,* 1–17. https://doi.org/10.1111/j.1744-7348.2008.00280.x

Lenzner, B., Latombe, G., Capinha, C., Bellard, C., Courchamp, F., Diagne, C., et al. (2020). What will the future bring for biological invasions on islands? An expert-based assessment. *Frontiers in Ecology and Evolution*, *8,* 280. https://doi.org/10.3389/fevo.2020.00280.

Luque, G. M., Bellard, C., Bertelsmeier, C., Bonnaud, E., Genovesi, P., Simberloff, D., et al. (2014). The 100th of the world's worst invasive alien species. *Biological Invasions, 16,* 981-985. https://doi.org/10.1007/s10530-013-0561-5.

Marks, R. A., Hotaling, S., Frandsen, P. B., & VanBuren, R. (2021). Representation and participation across 20 years of plant genome sequencing. *Nature Plants, 7,* 1571-1578. https://doi.org/10.1038/s41477-021-01031-8.

McCartney, M. A., Mallez, S., & Gohl, D. M. (2019). Genome projects in invasion biology. *Conservation Genetics, 20,* 1201-1222. https://doi.org/10.1007/s10592-019-01224-x.

Neill, P. E., & Arim, M. (2019). Human health link to invasive species. *Encyclopaedia of Environmental Health,* 570-578. https://doi.org/10.1016/B978-0-12-409548-9.11731-2.

Neinavaie, F., Ibrahim-Hashim, A., Kramer, A. M., Brown, J. S., & Richards, C. L. (2021) The genomic processes of biological invasions: From invasive species to cancer metastases and back again. *Frontiers in Ecology and Evolution, 9,* 681100. https://doi.org/10.3389/fevo.2021.681100.

North, H. L., McGaughran, A., & Jiggins, C. D. (2021). Insights into invasive species from whole-genome resequencing. *Molecular Ecology, 30,* 6289-6308. https://doi.org/10.1111/mec.15999.

Novoa, A., Richardson, D. M., Pyšek, P., Meyerson, L. A., Bacher, S., Canavan, S., et al. (2020) Invasion syndromes: A systematic approach for predicting biological invasions and facilitating effective management. *Biological Invasions, 22,* 1801–1820. https://doi.org/10.1007/s10530-020-02220-w.

Ochocki, B. M., & Miller, T. E. X. (2017). Rapid evolution of dispersal ability makes biological invasions faster and more variable. *Nature Communications, 8,* 14315. https://doi.org/10.1038/ncomms14315.

Olazcuaga, L., Loiseau, A., Parrinello, H., Paris, M., Fraimout, A., Guedot, C., et al. (2020). A whole-genome scan for association with invasion success in the fruit fly *Drosophila suzukii* using contrasts of allele frequencies corrected for population structure. *Molecular Biology and Evolution, 37*, 2369-2385. https://doi.org/10.1093/molbev/msaa098.

Pagad, S., Genovesi, P., Carnevali, L., Scalera, R., & Clout, M. (2015). IUCN SSC Invasive Species Specialist Group: Invasive alien species information management supporting practitioners, policy makers and decision takers. *Management of Biological Invasions, 6*, 127-135. https://doi.org/10.3391/mbi.2015.6.2.03.

Panda, R. M., Behera, M. D., & Roy, P. S. (2018). Assessing distributions of two invasive species of contrasting habits in future climate. *Journal of Environmental Management, 213*, 478-488. https://doi.org/10.1016/j.jenvman.2017.12.053.

Pearce, S. L., Clarke, D. F., East, P. D., Elfekih, S., Gordon, K. H. J., Jermiin, L. S., McGaughran, A., et al. (2017). Genomic innovations, transcriptional plasticity and gene loss underlying the evolution and divergence of two highly polyphagous and invasive *Helicoverpa* pest species. *BMC Biology*, *15*, 63. https://doi.org/10.1186/s12915-017-0402-6.

Puckett, E. E., Park, J., Combs, M., Blum, M. J., Bryant, J. E., Caccone, A. et al. (2016) Global population divergence and admixture of the brown rat (*Rattus norvegicus*). *Proceedings of the Royal Society B: Biological Sciences, 283*, 20161762. https://doi.org/10.1098/rspb.2016.1762.

Puzey, J., & Vallejo-Marín, M. (2014). Genomics of invasion: Diversity and selection in introduced populations of monkeyflowers (*Mimulus guttatus*). *Molecular Ecology, 23*, 4472-85. https://doi.org/10.1111/mec.12875.

Resh, C. A., Galaska, M. P., Benesh, K. C., Gardner, J. P., Wei, K. J., Yan, R. J. et al. (2021). Using genomics to link populations of an invasive species to its potential

sources. *Frontiers in Ecology and Evolution, 9.* https://doi.org/10.3389/fevo.2021.575599.

Rius, M., Bourne, S., Hornsby, H.G., Chapman, M.A. (2015) Applications of next-generation sequencing to the study of biological invasions. *Current Zoology, 614*(3), 88–504. https://doi.org/10.1093/czoolo/61.3.488.

Schatz, M. C., Witkowski, J., & McCombie, W. R. (2012). Current challenges in *de novo* plant genome sequencing and assembly. *Genome Biology*, *13*, 243. https://doi.org/10.1186/gb-2012-13-4-243.

Searle, J. B. (2008). The genetics of mammalian invasions: a review. *Wildlife Research, 35*, 185–192. https://doi.org/10.1071/WR07123.

Seebens, H., Bacher, S., Blackburn, T. M., Capinha, C., Dawson, W., Dullinger, S., et al. (2021). Projecting the continental accumulation of alien species through to 2050. *Global Change Biology, 27*, 970-982. https://doi.org/10.1111/gcb.15333.

Seebens, H., Blackburn, T. M., Dyer, E. E., Genovesi, P., Hulme, P. E., Jeschke, J. M., et al. (2017). No saturation in the accumulation of alien species worldwide. *Nature Communications*, *8*, 14435. https://doi.org/10.1038/ncomms14435.

Sherpa, S, Després, L. (2021). The evolutionary dynamics of biological invasions: A multi-approach perspective. *Evolutionary Applications*, *14*, 1463– 1484. https://doi.org/10.1111/eva.13215.

Sjodin, B. M. F., Irvine, R. L., Ford, A. T., Howald, G. R., & Russello, M. A. (2020). *Rattus* population genomics across the Haida Gwaii archipelago provides a framework for guiding invasive species management. *Evolutionary Applications, 13,* 889-904. https://doi.org/10.1111/eva.12907.

Suda, J., Meyerson, L. A., Leitch, I. J., & Pyšek, P. (2015). The hidden side of plant invasions: The role of genome size. *New Phytologist, 205*, 994-1007. https://doi.org/10.1111/nph.13107.

Sun, Y.-B., Zhang, Y., & Wang, K. (2020). Perspectives on studying molecular adaptations of amphibians in the genomic era. *Zoological Research*, *41*, 351-364. https://doi.org/10.24272/j.issn.2095-8137.2020.046.

Thompson, B. K., Olden, J. D., & Converse, S. J. (2021). Mechanistic invasive species management models and their application in conservation. *Conservation Science and Practice*, *3*, e533. https://doi.org/10.1111/csp2.533.

Tobin, P. C. (2018). Managing invasive species. *F1000Research*, *7*, 1686. https://doi.org/10.12688/f1000research.15414.1.

Wagner, N. K., Ochocki, B. M., Crawford, K. M., Compagnoni, A., & Miller, T. E. X. (2017). Genetic mixture of multiple source populations accelerates invasive range expansion. *Journal of Animal Ecology*, *86*, 21-34. https://doi.org/10.1111/1365-2656.12567.

Welles, S. R., & Dlugosch, K. M. (2019). Population genomics of colonization and invasion. In O. P. Rajora (Ed.), *Population genomics: Concepts, approaches and applications* (pp. 655-683). Springer International Publishing. https://doi.org/10.1007/13836_2018_22.

Wickland, D. P., Battu, G., Hudson, K. A., Diers, B. W. & Hudson, M. E. (2017). A comparison of genotyping-by-sequencing analysis methods on low-coverage crop datasets shows advantages of a new workflow, GB-eaSy. *BMC Bioinformatics,1*8, 586. https://doi.org/10.1186/s12859-017-2000-6.

Worley, K. C., Richards, S., & Rogers, J. (2017). The value of new genome references. *Experimental Cell Research*, *358*, 433-438. https://doi.org/10.1016/j.yexcr.2016.12.014.

Zoonomia Consortium. (2020) A comparative genomics multitool for scientific discovery and conservation. *Nature, 587*, 240–245. https://doi.org/10.1038/s41586-020-2876.

Chapter 3.

Genome-wide SNP data reveals minimal connectivity between Bt-resistant and non-resistant populations of pink bollworm (*Pectinophora gossypiella*)

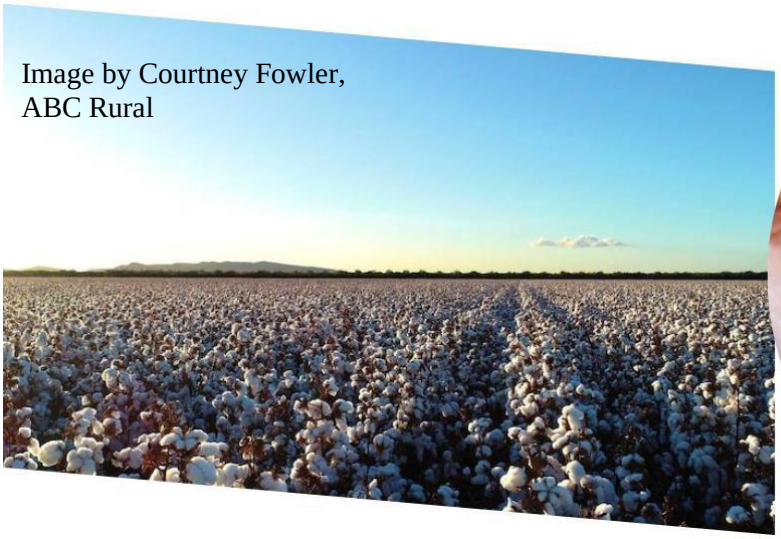


Image by Courtney Fowler, ABC Rural

Image by Peggy Greb, USDA

To be submitted as: Matheson, P., Parvizi, E., Fabrick, J. A., Siddiqui, H. A., Tabashnik, B. E., Walsh, T. W., & McGaughran, A. (2022). Genome-wide SNP data reveals minimal connectivity between Bt-resistant and non-resistant populations of pink bollworm (*Pectinophora gossypiella*). *Scientific Reports*, in prep.

Contributions: A.M and T.W wrote the original grant application and A.M designed the eventual analysis plan, with J.F, B.T, H.S, and T.W collecting or facilitating sample collection. P.M analysed the data with occassional assistance from A.M and E.P, and created all figures. PM wrote the first draft of the manuscript and all authors provided feedback, with PM leading subsequent revision of the final manuscript.

3.1 Abstract

The pink bollworm (*Pectinophora gossypiella*) is one of the world's most destructive pests of cotton. This invasive Lepidopteran occurs in nearly all cotton-growing countries. Its presence in the Ord Valley of North West Australia poses a potential threat to the expanding cotton industry there. To assess this threat and better understand gene flow in pink bollworm, we analysed genomic data from individuals collected in the field from North West Australia, India, and Pakistan, as well as from four laboratory colonies that originated in the United States. We identified single nucleotide polymorphisms (SNPs) using a reduced-representation, genotyping-by-sequencing technique (DArTseq). The final filtered dataset includes 6,355 SNPs and 88 individual genomes that clustered into five groups: Australia, India-Pakistan, and three groups from the United States. We found low genetic diversity within populations and high differentiation between populations from different continents. The high genetic differentiation between Australia and the other continents reduces concerns about gene flow to North West Australia, particularly from populations in India and Pakistan that have evolved resistance to transgenic insecticidal cotton. We attribute the observed population structure to pink bollworm's narrow host plant range and limited natural dispersal between continents.

3.2 Introduction

Invasive species represent a significant threat to agriculture, due to economic costs associated with management and reduced crop yields (Gippet et al. 2019). The pink bollworm (*Pectinophora gossypiella*) is a major pest of cotton (*Gossypium* spp.) that has colonised more than 100 countries worldwide (Henneberry & Naranjo, 1998; CABI Compendium, https://www.cabidigitallibrary.org/doi/10.1079/cabicompendium.39417). The origin of this invasive Lepidopteran pest is not known, but a leading candidate is India (Ingram, 1994; Liu et al. 2009), where it was first discovered damaging cotton in 1843 (Sridhar et al. 2017). It has also been hypothesised to have originated in Australia or Southeast Asia (Wang et al., 2011).

In Australia, pink bollworm was first reported on cotton in Queensland in 1924 (Holdaway, 1926). Currently, it is thought to occur primarily in Western Australia and the Northern Territory (Common, 1958; Holdaway, 1926; Naumann & Sands, 1984; Richards, 1964) which are over 1,000 km from Australia's primary cotton production areas of New South Wales and southern Queensland. However, the pink bollworm's presence in the Ord Valley in the northern part of Western Australia is a potential threat to the expanding cotton industry there, where over 1,000 hectares of cotton were planted in 2020 for the first time since 2011 (https://www.abc.net.au/news/rural/2020-08-31/ord-cotton-industry-hits-new-highs-in-was-far-north/12610730). Cotton genetically engineered to produce insecticidal proteins from the bacterium *Bacillus thuringiensis* (Bt) has been grown in Australia since 1996 and has proven to be effective against pink bollworm in the United States and China (Tabashnik et al. 2020; Wan et al. 2017). However, pink bollworm populations in India and Pakistan have evolved resistance to Bt cotton, with serious practical consequences (Tabashnik & Carrière, 2019). In principle, gene flow could introduce Bt-resistant pink bollworm into North West Australia. Thus, a

greater understanding of pink bollworm gene flow could be useful for assessing the threat to cotton in Australia as well for improving surveillance and management elsewhere.

Five previous studies of population structure in pink bollworm have analysed genetic variation based on DNA sequences from one to 13 loci (Liu et al. 2009; Liu et al. 2010; Wang et al. 2011; Sridhar et al. 2017; Naik et al. 2020). Two of these studies evaluated only the mitochondrial DNA (mtDNA) cytochrome *c* oxidase subunit I (COI) locus in samples from India (Sridhar et al. 2017; Naik et al. 2020), one focused primarily on China and used two mtDNA loci (Liu et al. 2010), or 13 microsatellites (Liu et al. 2009), and one was based on a *piggyBac*-like transposon insertion and its flanking sequences (Wang et al., 2011). Of these, only the latter included populations from Australia. The conclusions from these studies are mixed, including apparently conflicting results from mtDNA and microsatellite DNA analysis of the same set of populations (Liu et al., 2009, 2010; see Discussion). Population structure of pink bollworm is yet to be explored using genome-wide data that includes Australian samples and a large number of genetic markers. Such analysis may provide increased accuracy and resolution to address these discrepancies.

Here, we aimed to understand pink bollworm population structure by using a genotyping-by-sequencing technique (DArTseq; Kilian et al. 2012) to produce genome-wide data in the form of 6,355 single nucleotide polymorphisms (SNPs) identified from 88 individual pink bollworm from 11 populations in Australia, India, Pakistan, and the U.S. We use these data to assess mean relatedness of individuals (i.e., population structure), demographic history, and connectivity patterns between populations, with a particular focus on the Australian population. Our results show minimal gene flow between Australia and the other countries, which reduces concerns about the threat of natural introductions of Bt-resistant pink bollworm from India and Pakistan.

3.3 Methods

Pink bollworm samples

We analysed the genomes of 88 pink bollworm from Australia, India, Pakistan, and the U.S. (Table A3.1). Male moths were collected from near Kununurra, Western Australia (-15.65, 128.70) in March and April 2017 using pheromone traps (Agrisense Recharge Lures for *Pectinophora gossypiella*; Entosol (Australia) Pty Ltd, NSW Australia). Fourth instar larvae were collected from fields of Bt cotton producing Cry1Ac + Cry2Ab (Bt toxins) in 2010 from the states of Telangana, Maharashtra, Karnataka, and Andhra Pradesh in India (Mathew et al. 2018), and from fields of Bt cotton producing Cry1Ac in 2016-2017 from the provinces of Punjab and Sindh in Pakistan.

For pink bollworm from the U.S., we used individuals from four laboratory colonies in Arizona: APHIS-S_1, APHIS-S_2, Bt4R, and Bt4-R2. Both APHIS-S colonies were obtained from the United States Department of Agriculture (USDA) Animal and Plant Health Inspection Service (APHIS) laboratory in Phoenix, Arizona, where they have been maintained for over 40 years without exposure to Bt toxins or other insecticides (Liu et al. 2001, Bartlett 1995). Subsets of the APHIS-S colony were provided to USDA Agricultural Research Service Maricopa in 2006 (APHIS-S_2) and 2018 (APHIS-S_1). Bt4R and Bt4-R2 are laboratory-selected resistant colonies: Bt4R was derived from the Bt-susceptible Western Cotton Research Laboratory (WCRL) colony from 2007-2008 by laboratory selection for Cry1Ac resistance (Fabrick et al 2009). Bt4-R2 was obtained from Bt4R in 2010 and had 28-fold resistance to Cry1Ac and >10,000-fold resistance to Cry2Ab (Fabrick et al. 2015). Larvae from all four colonies were reared in the laboratory on wheat germ diet (Bartlett & Wolf 1985) at 26 °C and a photoperiod of 14 h light:10 h dark. Hereafter, we refer to each of these colonies as U.S. 'populations'.

DArTseq genotyping

We sent 88 pink bollworm samples to Diversity Arrays Technology (DArT Pty Ltd, Canberra, Australia) for DNA extraction, quantification, and genotyping (see Appendix 3; Table A3.1 for further sample details). DArTseq™ is a restriction enzyme-based, complexity reduction method that employs a next generation sequencing platform to generate genomic data in the form of SNPs, as detailed by Killian et al. (2012).

SNP filtering

We received genotyping outputs in the DArT '2 row' format, where alleles are scored in a binary fashion for each individual ('1' = presence of allele, '0' = absence of allele, and '-' = failure to score). This initial dataset contained 59,262 SNPs across 90 individuals. Using R version 4.1.2 (R core team, 2021), we converted this data into a genlight object using 'adegenet' version 2.1.7 (Jombart et al. 2008), and then used 'dartR' version 2.0.4 (Gruber et al. 2018) and 'radiator' version 1.2.2 (Gosselin et al. 2020) for data manipulation and filtering. We filtered SNPs by reproducibility (threshold: 0.98), call rate (threshold: 0.95), and minor allele frequency (threshold: 0.02). The final filtered dataset contained 88 individual genotypes, 6,355 SNP markers, and 1.41% missing data.

Heterozygosity

In R, we used 'hierfstat' version 0.5.11 (Goudet et al. 2005) to calculate observed (Ho) and expected (He) heterozygosity, as well as inbreeding coefficients (F_{IS}; a metric that ranges from -1 to 1, where values close to 0 meet the neutral expectation, values that approach 1 indicate a deficit of heterozygotes indicating inbreeding, and values approaching -1 indicate an excess of heterozygotes) for each population from the filtered dataset.

Population structure

We used R to analyse genetic variation between and within populations by performing an Analysis of Molecular Variance (AMOVA) with the 'poppr' package version 2.9.3

(Kamvar et al. 2014). We randomly permuted the AMOVA output 1,000 times to test if populations were significantly different using the function ‘randtest’ from the package ‘ade4’ version 1.7.19 (Dray et al. 2007). We then estimated pairwise genetic differentiation (F_{ST}) between populations using the function ‘genet.dist’ and method ‘WC84’ from the ‘hierfstat’ package, version 0.5.11 (Weir & Cockerham 1984; Weir & Goudet, 2017).

We conducted a principal component analysis based on Euclidean genetic distances, using the function ‘glPCA’ implemented in the ‘adegenet’ package in R. We performed two analyses, one with only the 11 populations of pink bollworm, and another that also included an ‘outgroup’ population of spotted pink bollworm (*Pectinophora scutigera*; see Fig. A3.1).

We calculated individual admixture coefficients by first converting SNP data into ‘STRUCTURE’ format using the ‘gl2faststructure’ function implemented in the R package ‘dartR’, and then to ‘.geno’ format using the ‘struct2geno’ function from ‘LEA’ version 3.6.0 (Frichot & Francios, 2015). We ran sparse non-negative matrix factorisation on individuals using the ‘sNMF’ function also implemented in ‘LEA’ and analysed K values of 1 to 10, with 100 replications for each K value. We identified the K value that best explained our results using the cross-entropy criterion (Fig. A3.2).

To examine relative migration levels and directions among populations, we first converted our SNP data to variant call format (VCF) using the R function ‘genomic_converter’ from ‘radiator’ version 1.2.2 (Gosselin, 2020). We then used the ‘vcf2genepop.pl’ script (available at: https://github.com/z0on/2bRAD_denovo/blob/master/vcf2genepop.pl) to convert the VCF file to genepop format, and the function ‘divMigrate’ from the R ‘diveRsity’ package version 1.9.90 (Keenan et al. 2018) to measure relative migration. The method of genetic distance was set to ‘Gst’ with a filter threshold of 0.5 and significance of values

were tested using 1,000 bootstraps and an alpha value of 0.05. DivMigrate identifies relative levels and directions of ancestral migration by defining a hypothetical pool of migrants for a pair of populations and then estimating genetic differentiation between each of those two populations and the hypothetical population. This directional differentiation is then used to estimate relative levels of migration between the two populations, with the larger and smaller values indicating which of the two populations is the most likely source and sink population, respectively (Sundqvist et al. 2016).

To examine relationships among samples in a phylogenetic context, we converted the same VCF from the migration analysis to phylip format using the command 'vcf2phylip' script (available at: https://github.com/edgardomortiz/vcf2phylip). A maximum likelihood (ML) phylogeny was constructed in IQtree version 1.2.1 (Nguyen et al. 2015) using the best-fit substitution model automatically selected by the software, with 10,000 bootstrap iterations to assess clade support. The resulting output was read into R using 'ape' version 5.6.2 (Paradis & Schliep, 2018) for data visualisation.

Finally, to further investigate genetic diversity and population structure of global pink bollworm populations, we obtained 31 mitochondrial COI sequences in FASTA format from GenBank (Table A3.2) for populations from the U.S., India, Australia, Pakistan, Israel, and Kenya. COI sequences were aligned using MEGA version 11 (Tamura et al. 2021). The nucleotide diversity of each population was then calculated using the 'nuc.div' function in R from the 'pegas' package version 1.1 (Paradis, 2010). COI sequences were converted to haplotypes using the same package, and the function 'haploNet' was used to construct a haplotype network.

Demographic modelling

To investigate the possible demographic history of pink bollworm, we used an approximate Bayesian computation approach in DIYABC version 1.0 (Collin et al. 2021) using a Random Forest algorithm. We grouped samples based on maximum likelihood

phylogeny clustering results in order to maintain population structure patterns similar to the dataset. We then formulated 12 invasion scenarios, all characterised by a genetic bottleneck occurring at the approximate time pink bollworm were discovered in each region according to historical records (Fig. A3.3; Table A3.3). Following the software guidelines, we simulated two thousand summary statistics per scenario and identified scenarios with higher classification votes (i.e., the number of times a scenario is chosen as best suited to the target dataset among the set of compared scenarios). Based on these results, we formulated a further 14 scenarios under the assumption that the India-Pakistan group diverged first in our tested set of populations. In this new set of scenarios, we considered the possibility of unsampled 'ghost' populations in addition to genetic bottlenecks (Fig. A3.4; Table A3.4). We simulated two thousand summary statistics per scenario, and identified the scenario with the highest classification vote and posterior probability (i.e., the scenario deemed most likely; Fig. A3.5; Table A3.5). The goodness of fit of the parameter-posterior combination for the chosen scenario was validated by checking the position of the observed dataset among the summary statistics on a PCA (Fig. A3.5).

Outlier analysis

To detect outlier SNPs, we compared the pooled India-Pakistan populations with either the Australian population or the different U.S. populations (considering Bt4R and Bt4-R2 as one population because of their close genetic relationship, as indicated by PCA) and used 'pcadapt' version 4.3.3 (Luu et al. 2017). According to Cattell's rule, we used a K value of 2 (Cattell, 1966), and set the distance statistic to Mahalanobis and the LD clumping r^2 threshold to 0.2 for a window of 200 SNPs, and we considered q-values less than 0.05 as candidates for 'outlier' status (Storey & Tibshirani, 2003; Duforet-Frebourg et al. 2016).

3.4 Results

Heterozygosity

Heterozygosity was low across all populations, with values of observed heterozygosity (Ho) that were consistently lower than expected heterozygosity (He) (t_{10} = 4.72; P < 0.001). As expected, mean Ho was lower in the four U.S. laboratory populations (0.09) than the seven wild India-Pakistan and Australian populations (0.19) (t_9 = 4.87; P < 0.001). Ho was not correlated with the number of field sites sampled per population (r_5 = 0.34; P = 0.23), nor the population sample size (r_5 = 0.12; P = 0.72). However, the mean excess in He relative to Ho as a proportion of He (i.e., [(He - Ho) / He] was significantly higher for the seven field populations (0.17) compared to the U.S. lab populations (0.06) (t_9 = 2.87; P = 0.018). Similarly, field populations displayed greater putative inbreeding coefficients (range = 0.068 – 0.280) than laboratory populations (range = 0.015 – 0.083) (t_9 = 2.63, P = 0.03) (Table 3.1).

Table 3.1 Location, number of field sites, and number of individuals included in the analysis, average observed heterozygosity (Ho), expected heterozygosity (He), and inbreeding coefficients (F_{IS}) for 11 populations of pink bollworm.

Population	Country	State or Province	n	No. field sites	Ho	He	F_{IS}
Aus-K	Australia	Western Australia	12	1	0.141	0.209	0.280
Ind-A	India	Andra Pradesh	10	3	0.197	0.231	0.117
Ind-K	India	Karnataka	4	1	0.194	0.233	0.110
Ind-M	India	Maharashtra	9	3	0.195	0.23	0.121
Ind-T	India	Telangana	5	2	0.198	0.232	0.101
Pak-P	Pakistan	Punjab	14	5	0.199	0.236	0.129
Pak-S	Pakistan	Sindh	4	1	0.207	0.235	0.068
APHIS-S_1	U.S.	Arizona	12	NA	0.160	0.178	0.083
APHIS-S_2	U.S.	Arizona	11	NA	0.060	0.060	0.000
Bt4R	U.S.	Arizona	4	NA	0.08	0.089	0.055
Bt4-R2	U.S.	Arizona	3	NA	0.072	0.075	-0.015

Population structure

AMOVA indicated that 25.6% of genetic variation was partitioned between populations, while 12.3% and 62.1% of variation was partitioned between and within individuals, respectively ($P < 0.001$ in all cases; Table 3.2).

Table 3.2 Analysis of molecular variance (AMOVA) to assess variation between populations, and between and within individuals, for 11 populations of pink bollworm.

Source	**df**	**SS**	**MS**	**Est.Var.**	**PV**	***P*-value**
Between populations	10	79,799.7	7.979.9	421.4	25.6	< 0.001
Between individuals	77	107,396.5	1,394.8	193.2	12.3	< 0.001
Within individuals	88	88,741.7	1,008.4	62.1	62.1	< 0.001
Total	175	275,937.9	1,576.8	1623.0	100.0	-

df: degrees of freedom, SS: sum of squares, MS: mean square, Est. Var: estimated variance, PV: percentage variance

Table 3.3 Pairwise FST values for the 11 populations of pink bollworm listed in Table 3.1.

	Aus-K	Ind-A	Ind-K	Ind-M	Ind-T	Pak-P	Pak-S	APHIS-S_1	APHIS-S_2	Bt4R
Ind-A	0.260	-	-	-	-	-	-	-	-	-
Ind-K	0.259	0	-	-	-	-	-	-	-	-
Ind-M	0.259	0.020	0.014	-	-	-	-	-	-	-
Ind-T	0.255	0.001	0.003	0.007	-	-	-	-	-	-
Pak-P	0.243	0.016	0.011	0.012	0.006	-	-	-	-	-
Pak-S	0.253	0.022	0.021	0.018	0.016	0.003	-	-	-	-
APHIS-S_1	0.347	0.204	0.224	0.206	0.212	0.185	0.211	-	-	-
APHIS-S_2	0.550	0.450	0.549	0.458	0.518	0.408	0.541	0.393	-	-
Bt4R	0.431	0.312	0.369	0.318	0.351	0.287	0.365	0.273	0.558	-
Bt4-R2	0.423	0.310	0.366	0.313	0.337	0.282	0.356	0.327	0.583	0.327

We observed clear genetic structuring of populations based on geography, with pairwise F_{ST} values greater than 0.185 for populations pertaining to Australia, U.S., and India-Pakistan. U.S. populations were strongly genetically differentiated (pairwise F_{ST} = 0.273 – 0.583; Table 3.3), whereas Indian and Pakistani populations were genetically similar to one another (pairwise $F_{ST} < 0.022$; Table 3.3). These findings were reinforced by the principal component analysis (PCA; Fig. 3.1) and maximum likelihood (ML) phylogeny (Fig. 3.2). The first two principal components in the PCA explained 26.6% of the total genetic variance in our dataset, and clearly demonstrated the distinction between individuals assigned to U.S. (blue shades), Australian (green), and India-Pakistan (red/orange/yellow shades) populations (Fig. 3.1), with a bootstrap confidence value of 100 supporting the divergence of the three main groups in the ML phylogeny (Fig. 3.2). The PCA assigned U.S. individuals to populations forming three main clusters: 1) APHIS-S_1, 2) APHIS-S_2, and 3) Bt4R and Bt4-R2 (Fig. 1). ML further distinguished this grouping, placing each of the four U.S. populations as genetically distinct with a bootstrap confidence value of 100 (Fig. 3.2). Consistent with pairwise F_{ST} values, individuals from India and Pakistan lacked population structure, forming one genetic aggregate in the PCA plot and showing limited divergence in the ML phylogeny (Fig. 3.2).

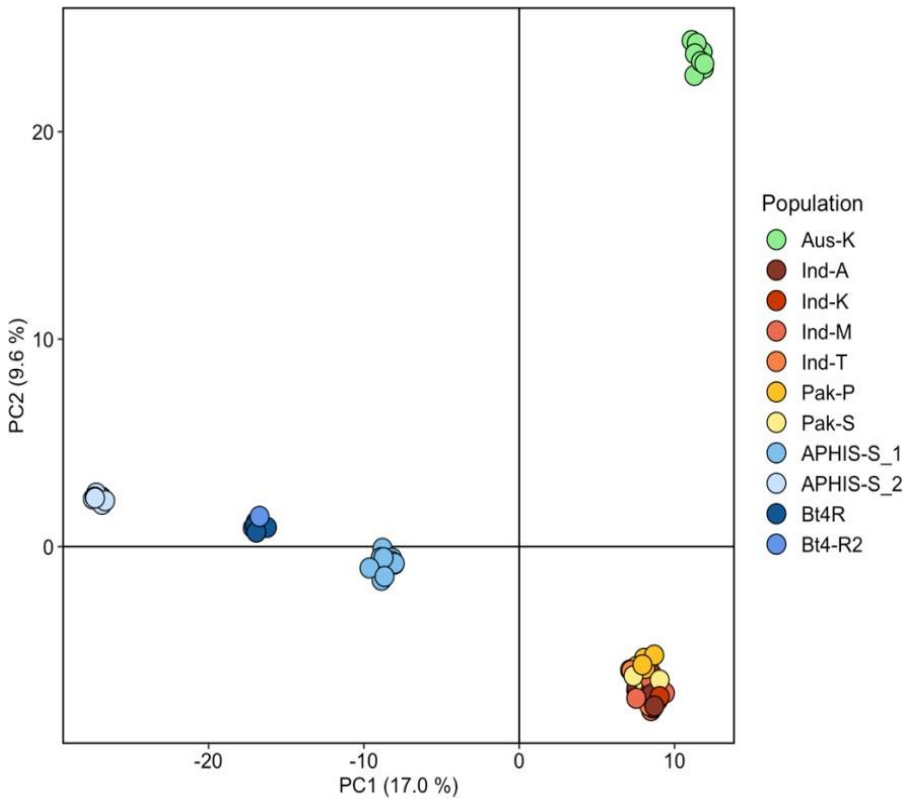


Figure 3.1 Principal component analysis (PCA) of the filtered dataset of 6,355 SNP loci for 11 populations of pink bollworm. Refer to Table 3.1 for population code details.

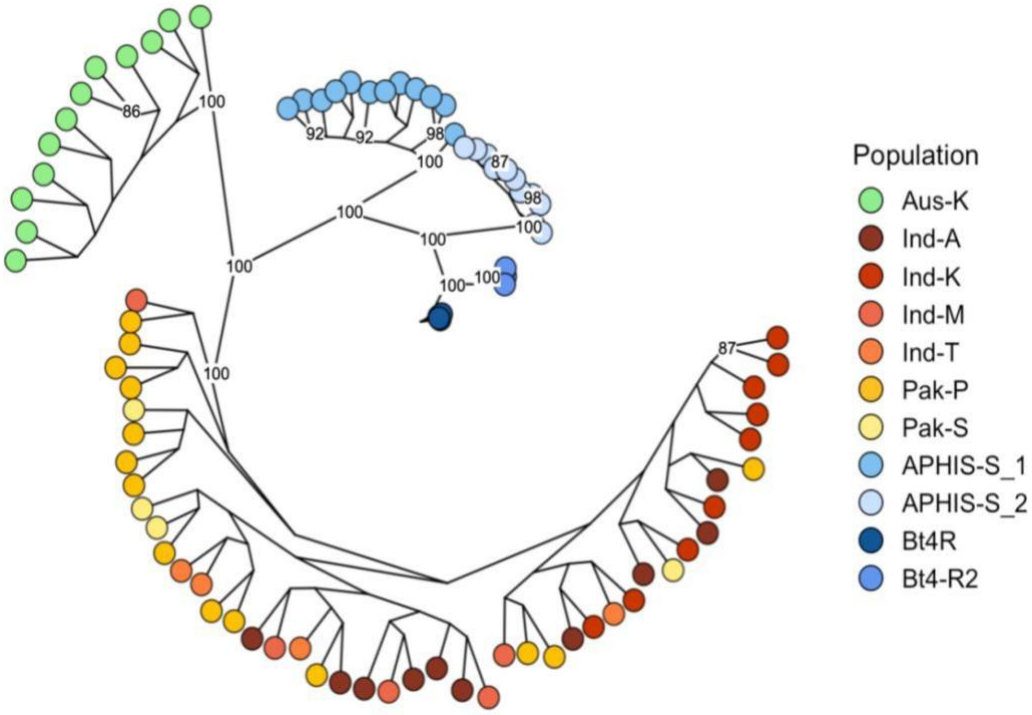


Figure 3.2 Maximum likelihood tree for individuals from 11 populations of pink bollworm with bootstrap confidence values > 85 attached to nodes. Refer to Table 3.1 for population code details.

Sparse Non-Negative Matrix Factorisation (sNMF) analysis of the filtered dataset (6,355 SNPs) indicated an optimal K value of five clusters (Fig. A3.2; see Fig. A3.6 for admixture proportions at K = 3, K = 4, and K = 6). These corresponded to Australia, India-Pakistan, and three groups for the U.S. populations: 1) APHIS-S_2 (light blue), 2) APHIS-S_1 (medium blue), and 3) Bt4R and Bt4-R2 (dark blue) (Fig. 3.3). We found limited admixture, with each individual corresponding primarily to one genetic group (Fig. 3.3A). The India-Pakistan populations showed the highest absolute admixture, with very small proportions (< ~5%) of shared ancestry with U.S. and Australian clusters (Fig. 3.3A). This was supported by our migration analysis, which showed highest *relative* directional gene flow (range = 0 - 1; see Methods) estimates from the U.S. to Pakistan-India (blue arrows; Fig. 3.3B), though we note that this is likely to be low in absolute terms (see Discussion). Meanwhile, Australia was highly isolated, sharing limited genetic ancestry or migration pathways with other populations (Fig. 3.3).

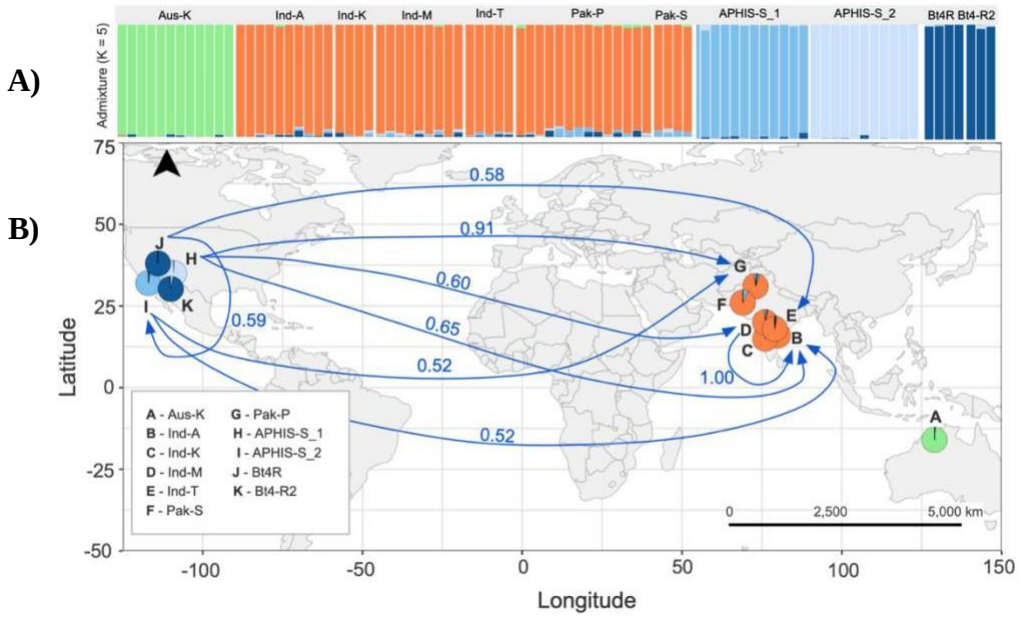


Figure 3.3 Admixture proportions for 11 populations of pink bollworm with approximate relative gene flow between countries under a filter threshold of 0.5 (blue arrows). Bar plots in A) represent admixture coefficients at K = 5 for each individual. In B), pie charts of admixture coefficients are plotted on a map and labelled A-K for each population of pink bollworm based on the provided key. Refer to Table 3.1 for population code details.

We identified thirteen haplotypes from 31 COI sequences downloaded from GenBank that included individuals from the U.S., India, Australia, Israel, Kenya, and Pakistan (see Methods). The haplotype network was dominated by a central haplotype (A), which was observed in 16 individuals across all six populations (n = 3, 2, 1, 4, 3, and 3, for Australia, India, Pakistan, Kenya, Israel, and the U.S., respectively; Fig. 3.4). Consistent with our genetic diversity findings (Table 3.1), we found that India had the highest diversity, with ten haplotypes (C-L) differing from the central haplotype by a single substitution, forming a star-shaped phylogeny (with a common central haplotype surrounded by less-common haplotypes that are one, or a few, mutational steps different). The only other frequent haplotype (B) was found in four individuals from the Pakistan population (Fig. 3.4). India had the highest nucleotide diversity (0.002), followed by Israel (0.0007), while populations from Australia, America, Kenya, and Pakistan all had nucleotide diversity values of zero due to the presence of only a single haplotype in those populations.

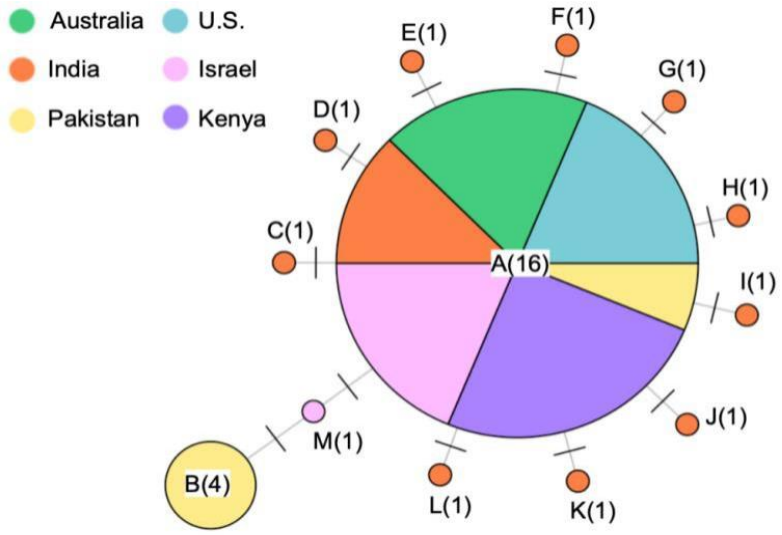


Figure 3.4 Haplotype network based on 31 pink bollworm COI sequences from Australia, India, Pakistan, Israel, Kenya, and the U.S. Circles represent haplotypes and the numbers in parentheses indicate the frequency of each haplotype. Lines between haplotypes represent mutations and dashes along lines represent mutational steps between haplotypes.

Demographic modelling

Our second round of demographic modelling was based on scenarios that included unsampled 'ghost' populations to account for the fact that our dataset may not include the entire available gene pool (Cao et al. 2016). The most likely scenario among the 14 tested suggested that pink bollworm populations in our dataset likely originated from an unsampled population, after which India-Pakistan diverged first (consistent with first records in the literature; Naranjo et al. 2002), Australia diverged second, and the U.S. diverged last (Fig 3.5; 439 classification votes; P = 0.518; Table A3.5).

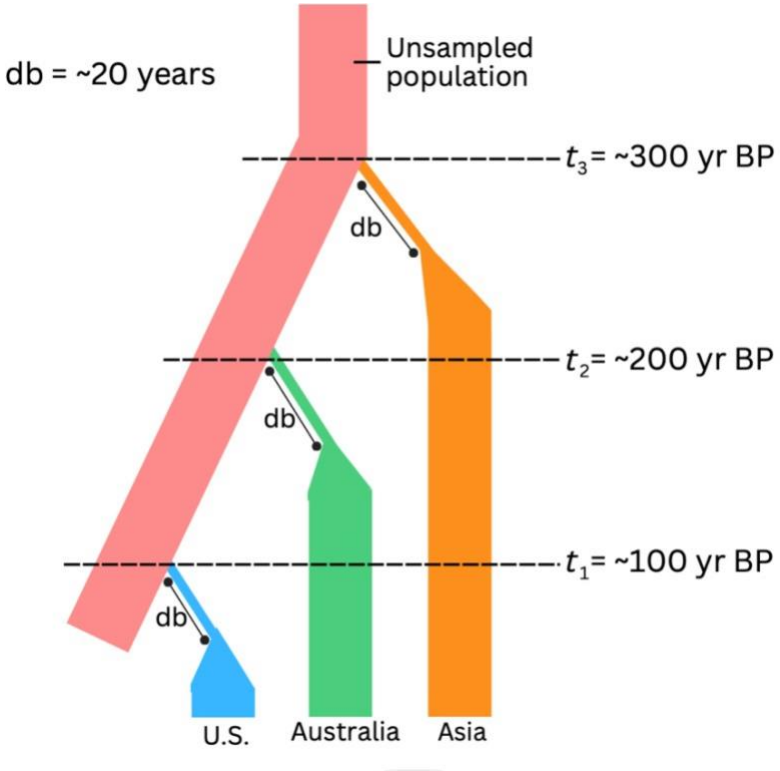


Figure 3.5 Model and estimated parameters from demographic modelling analysis of pink bollworm populations. t_1, t_2, t_3: approximate timing of pink bollworm detection in U.S., Australia, and India-Pakistan, respectively; db: bottleneck periods at the time of each colonisation event (see Table A3.4 and Appendix 3 for further details).

Signatures of selection

We performed four separate outlier analyses on three U.S. colonies (APHIS-S_1, APHIS-S_2, and Bt4R + Bt4-R2 combined) and one Australian population (see Methods) to identify any shared loci showing putative evidence of selection (i.e., 'outlier SNPs'). We used India-Pakistan as the comparison population for these analyses, treating it as the native range population to identify shared outliers in the other populations that may be related to invasion success.

The most outliers were detected in APHIS-S_1 (n = 305), followed by APHIS-S_2 (n = 200), Aus-K (n = 66), and Bt4R + Bt4-R2 combined (n = 36). All but Bt4R + Bt4-R2/Aus-K pairwise population comparisons shared some of the same outlier SNPs (range 3-21), however, no three-way or four-way comparisons identified any shared outlier loci (Fig. 3.6). The density and linkage among these outlier SNPs, and their functional importance, is not clear from the current analysis (see Discussion).

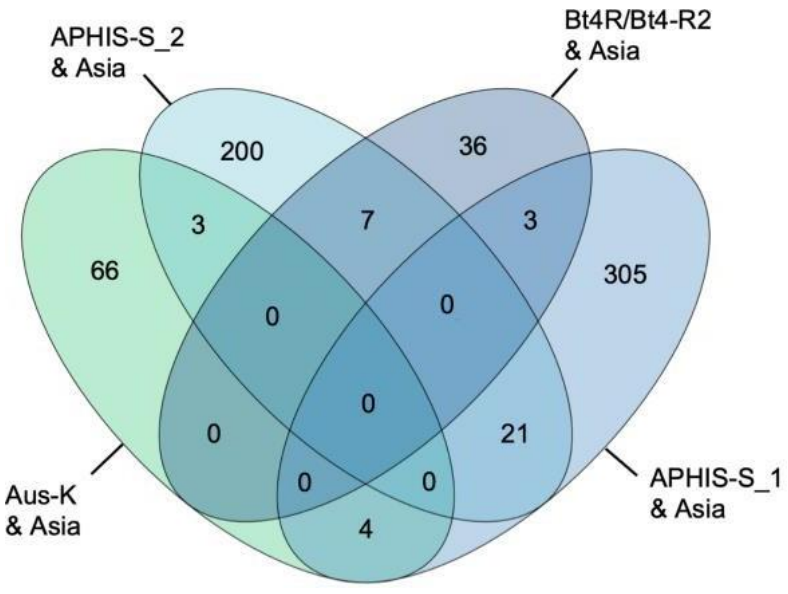


Figure 3.6 Venn diagram showing numbers of shared outlier SNPs for putatively "invasive" pink bollworm populations ('APHIS-S_1', 'APHIS-S_2', and 'Bt4R + Bt4-R2' from the U.S., and 'Aus-K' from Australia) *versus* combined "native" samples from India-Pakistan (i.e., Asia).

3.5 Discussion

We used population genomic analyses to assess population structure, demographic history, and connectivity of pink bollworm, with particular interest in a population from Kununurra in North West Australia, to evaluate its risk to the expanding cotton industry there. Overall, our results showed minimal gene flow between Australia and other countries in the dataset, which reduces concern about the threat of introduction there of Bt-resistant insects from India and Pakistan.

Conflicting results were obtained in five previous studies using genetic markers to investigate the population structure and genetic diversity of pink bollworm populations. For example, using microsatellites (Liu et al. 2009), relatively high genetic diversity and clear population structure were found among pink bollworm from China, Pakistan, and the U.S. However, low genetic variation and weak differentiation was later found for the same populations analysing mtDNA (Liu et al. 2010). Among the other studies, high haplotype diversity within populations from Australia, China, India, Israel, Mexico, and the U.S. was found by analysing transposable element sequence data (Wang et al., 2011), while low diversity was found in two studies that analysed mtDNA sequences from India (Sridhar et al. 2016; Naik et al. 2020). Such discrepancies may be due to the different types of genetic markers used. Namely, the different marker regions could have variable mutation rates, mtDNA is more prone to genetic drift due to its maternal inheritance (Xu et al. 2012; Song et al. 2014) and each of these studies used only a limited number of genetic loci. Here, we used a comprehensive approach to analyse genome-wide SNP data, which is generally expected to increase resolution and accuracy for understanding genetic processes (North et al. 2021).

Our results showed that genetic diversity (heterozygosity at SNP markers, and nucleotide diversity at mtDNA COI) was low among all populations. However, while there was significantly lower heterozygosity in the laboratory populations (U.S.)

compared to the field populations (Australia and India-Pakistan), putative inbreeding was higher in field populations. This was a surprising result, given that laboratory environments would generally be expected to cause a reduction in heterozygosity due to their small population sizes and increased likelihood of mating between related individuals (Briscoe et al. 1992; Ross et al. 2019). Estimates of diversity using SNP data may be influenced by sample size, rare alleles, missing data, and population structure (Schmidt et al. 2021). However, we found that Ho was not correlated with sample size nor with the number of field sites sampled, and the amount of missing data in our SNP matrix was minimal (< 1.5%). Consistent with the microsatellite-based findings of Liu et al. (2009), we found strong genetic differentiation between populations of pink bollworm from different continents, and this may explain the F_{IS} patterns. More specifically, differential allele frequencies in isolated populations may cause many variant sites to be polymorphic in only one or two populations and monomorphic in others, resulting in lower estimates of heterozygosity overall (Schmidt et al. 2021).

Previous studies also showed strong genetic differentiation, limited gene flow, and low heterozygosity (although with variable consistency; see above) in wild pink bollworm populations (Liu et al. 2009, Liu et al. 2010, Wang et al., 2011, Sridhar et al. 2016; Naik et al. 2020). Such results were attributed to potential genetic bottlenecks caused by larval mortality from Bt-cotton, post-Pleistocene range expansion with limited founders, limited flight activity and/or narrow host-specificity (Liu et al. 2009, Liu et al. 2010, Wang et al., 2011, Sridhar et al. 2016; Naik et al. 2020). Here, we expected to observe patterns of directional gene flow that would reveal the colonisation history of pink bollworm, most likely radiating from the supposed native range in India-Pakistan (Ingram, 1994; Liu et al. 2009). However, the divMigrate gene flow analysis suggested that the main migration direction was from the U.S. toward India and Pakistan. Although evidence does exist showing pink bollworm moths have been caught at altitudes up to

1,000 m and that long-distance dispersal occurs in the U.S. (Tabashnik et al. 1999, McDonald & Loftin 1935, Bariola et al. 1973 and references therein), long distance, inter-continental dispersal (from Asia to Australia and/or the U.S.) is not expected to be within the realm of natural dispersal capabilities for this moth. Together with our observations of extremely low admixture between pink bollworm populations from Australia, U.S. and India-Pakistan, it is likely that these ambiguous gene flow findings are due to the fact that overall dispersal is highly limited, and that any realised gene flow would most likely be due to prior human-assisted movement rather than recent introductions.

A further aim of our study was to determine the origin of pink bollworm, with previous studies suggesting India as the leading source (Ingram, 1994; Liu et al. 2009). We found that observed heterozygosity was highest in India-Pakistan, while nucleotide and haplotype diversity based on COI sequences was highest in India, supporting an India-Pakistan origin. The most likely scenario in our demographic modelling also indicated that the India-Pakistan populations diverged first. However, the modeling also suggested that unsampled populations play a role in the invasion history of pink bollworm. However, genomic data from populations covering a wider geographical scope is necessary for future studies aiming to reconstruct the evolutionary history of pink bollworm and determine its origin with greater confidence.

The high genetic differentiation observed between Australia and other continents indicates that the potential for gene flow to North West Australia from Bt-resistant populations in India and Pakistan is likely to be low. Furthermore, we found little to no evidence to suggest overlap in terms of shared outlier SNPs among populations, suggesting a lack of shared adaptive responses among the populations analysed - at least for the SNP sites in our dataset. Future studies may investigate such responses in more detail using the newly available reference genome for pink bollworm (Stahlke et al. 2022). For now, the invasion hypothesis that offers the most parsimonious explanation for the

Australian population is that pink bollworm was introduced via transport of cotton seed with a small number of founders, followed by a long period of isolation from the ancestral population. Through strong genetic drift, divergent selection (possibly caused by different ways pink bollworm have been managed; see Tabashnik & Carrière, 2019), and a lack of gene flow due to limited natural dispersal frequency, these factors likely shaped the population genomic dynamics reported here. Additional research is needed on pink bollworm populations from a wider geographic scope to gain a greater insight into the origin and global dynamics of this cosmopolitan pest.

3.6 Literature cited

Bariola, L. A., Keller, J. C., Turley, D. L., & Farris, J. R. (1973). Migration and population studies of the pink bollworm 1 in the Arid West 23. *Environmental Entomology*, *2*, 205-208. https://doi.org/10.1093/ee/2.2.205.

Bartlett, A. C. (1995). Resistance of the pink bollworm to Bt transgenic cotton. Beltwide Cotton Conferences (USA), In *Beltwide Cotton Conferences (USA)*.

Bartlett, A. C. & Wolf, W. W. (1985). *Pectinophora gossypiella*. Handbook of Insect Rearing. Vol. 2 (eds. Singh, P. & Moore, R. F.) pp. 415-430. Elsevier Scientific Publishing Co., Amsterdam.

Briscoe, D. A., Malpica, J. M., Robertson, A., Smith, G. J., Frankham, R., Banks, R. G., & Barker, J. S. F. (1992). Rapid loss of genetic variation in large captive populations of *Drosophila* flies: implications for the genetic management of captive populations. *Biology*, *6*, 416-425. https://doi.org/10.1046/j.1523-1739.1992.06030416.x.

Cao, L. J., Wei, S. J., Hoffmann, A. A., Wen, J.-B., & Chen, M. (2016). Rapid genetic structuring of populations of the invasive fall webworm in relation to spatial expansion and control campaigns. *Diversity and Distributions*, *22*, 1276-1287. https://doi.org/10.1111/ddi.12486.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*, 245-276. https://doi.org/10.1207/s15327906mbr0102_10.

Collin, F.-D., Durif, G., Raynal, L., Lombaert, E., Gautier, M., Vitalis, R. et al. (2021). Extending approximate Bayesian computation with supervised machine learning to infer demographic history from genetic polymorphisms using DIYABC Random Forest. *Molecular Ecology Resources*, *21*, 2598-2613. https://doi.org/10.1111/1755-0998.13413.

Common, I. (1958). A revision of the pink bollworms of cotton (*Pectinophora busck* (Lepidoptera: Gelechiidae) and related genera in Australia. *Australian Journal of Zoology*, *6*, 268-306. https://doi.org/10.1071/ZO9580268.

Dray, S., & Dufour, A.-B. (2007). The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software*, *22*, 1 - 20. https://doi.org/10.18637/jss.v022.i04.

Duforet-Frebourg, N., Luu, K., Laval, G., Bazin, E., & Blum, M. G. B. (2016). Detecting genomic signatures of natural selection with principal component analysis: Application to the 1000 genomes data. *Molecular Biology and Evolution*, *33*, 1082-1093. https://doi.org/10.1093/molbev/msv334.

Fabrick, J. A., & Tabashnik, B. E. (2012). Similar genetic basis of resistance to Bt toxin Cry1Ac in boll-selected and diet-selected strains of pink bollworm. *PLOS ONE*, *7*, e35658. https://doi.org/10.1371/journal.pone.0035658.

Fabrick, J. A., Unnithan, G. C., Yelich, A. J., DeGain, B., Masson, L., Zhang, J. et al. (2015). Multi-toxin resistance enables pink bollworm survival on pyramided Bt cotton. *Scientific Reports*, *5*, 16554. https://doi.org/10.1038/srep16554.

Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, *6*, 925-929. https://doi.org/10.1111/2041-210X.12382.

Gippet, J. M. W., Liebhold, A. M., Fenn-Moltu, G., & Bertelsmeier, C. (2019). Human-mediated dispersal in insects. *Current Opinion in Insect Science*, *35*, 96-102. https://doi.org/10.1016/j.cois.2019.07.005.

Goudet, J. (2005). hierfstat, a package for r to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, *5*, 184-186. https://doi.org/10.1111/j.1471-8286.2004.00828.x.

Gosselin, T. (2020). radiator: RADseq data exploration, manipulation and visualization using R. R package version 1.1.9. https://thierrygosselin.github.io/radiator/. Accessed November 3 2022.

Gruber, B., Unmack, P. J., Berry, O. F., & Georges, A. (2018). dartr: An R package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Molecular Ecology Resources*, *18*, 691-699. https://doi.org/10.1111/1755-0998.12745.

Henneberry, T., & Naranjo, S. (1998). Integrated management approaches for pink bollworm in the southwestern United States. *Integrated Pest Management Reviews*, *3*, 31-52. https://doi.org/10.1023/A:1009673622862.

Holdaway, F. G. (1926). The pink bollworm of Queensland. *Bulletin of Entomological Research*, *17*, 67-83. https://doi.org/10.1017/S000748530001909X.

Ingram, W. (1994). *Pectinophora* (Lepidoptera: Gelechiidae), insect pests of cotton. In: CAB International Publisher, Wallingford, United Kingdom.

Jombart, T. (2008). adegenet: An R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*, 1403-1405. https://doi.org/10.1093/bioinformatics/btn129.

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, *2*, e281. https://doi.org/10.7717/peerj.281.

Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., & Prodöhl, P. A. (2013). diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, *4*, 782-788. https://doi.org/10.1111/2041-210X.12067.

Kilian, A., Wenzl, P., Huttner, E., Carling, J., Xia, L., Blois, H. et al. (2012). Diversity arrays technology: A generic genome profiling technology on open platforms.

Methods in Molecular Biology, *888*, 67-89. https://doi.org/10.1007/978-1-61779-870-2_5.

Liu, Y., Hou, M., & Wu, K. (2010). Genetic variation of mitochondrial DNA in Chinese populations of *Pectinophora gossypiella* (Lepidoptera: Gelechiidae). *Environmental Entomology*, *39*, 1344-1351. https://doi.org/10.1603/EN09355.

Liu, Y.-B., Tabashnik, B. E., Dennehy, T. J., Patin, A. L., Sims, M. A., Meyer, S. K. et al. (2001). Effects of Bt cotton and Cry1Ac toxin on survival and development of pink bollworm (Lepidoptera: Gelechiidae). *Journal of Economic Entomology*, *94*, 1237-1242. https://doi.org/10.1603/0022-0493-94.5.1237.

Liu, Y.-D., Wu, K.-M., & Guo, Y.-Y. (2009). Population structure and introduction history of the pink bollworm, *Pectinophora gossypiella*, in China. *Entomologia Experimentalis et Applicata*, *130*, 160-172. https://doi.org/10.1111/j.1570-7458.2008.00800.x.

Luu, K., Bazin, E., & Blum, M. G. B. (2017). pcadapt: An R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, *17*, 67-77. https://doi.org/10.1111/1755-0998.12592.

McDonald, R. E., & Loftin, U. C. (1935). Dispersal of the pink bollworm by flight or wind carriage of the moths. *Journal of Economic Entomology*, *28*, 745-755. https://doi.org/10.1093/jee/28.5.745.

Mathew, L. G., Ponnuraj, J., Mallappa, B., Chowdary, L. R., Zhang, J., Tay, W. et al. (2018). ABC transporter mis-splicing associated with resistance to Bt toxin Cry2Ab in laboratory- and field-selected pink bollworm. *Scientific Reports*, *8*, 13531. https://doi.org/10.1038/s41598-018-31840-5.

Naik, V. C. B., Pusadkar, P. P., Waghmare, S. T., K. P, R., Kranthi, S., Kumbhare, S. et al. (2020). Evidence for population expansion of cotton pink bollworm

Pectinophora gossypiella (Saunders) (Lepidoptera: Gelechiidae) in India. *Scientific Reports*, *10*, 4740. https://doi.org/10.1038/s41598-020-61389-1.

Naranjo, S. E., Butler, G. D., & Henneberry, T. J. (2002). A bibliography of the pink bollworm, *Pectinophora gossypiella* (Saunders). *US Department of Agriculture, Agricultural Research Service Maryland*. Available at: https://handle.nal.usda.gov/10113/353.

Naumann, I. D., & Sands, D. P. A. (1984). Two Australian *Elasmus spp.* (Hymenoptera: Elasmidae), parasitoids of *Pectinophora gossypiella* (Saunders) (Lepidoptera: Gelechiidae): Their taxonomy and biology. *Australian Journal of Entomology*, *23*, 25-32. https://doi.org/10.1111/j.1440-6055.1984.tb01901.x.

Nguyen, L. T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, *32*, 268-274. https://doi.org/10.1093/molbev/msu300.

North, H. L., McGaughran, A., & Jiggins, C. D. (2021). Insights into invasive species from whole-genome resequencing. *Molecular Ecology*, *30*, 6289-6308. https://doi.org/10.1111/mec.15999.

Paradis, E. (2010). pegas: An R package for population genetics with an integrated–modular approach. *Bioinformatics*, *26*, 419-420. https://doi.org/10.1093/bioinformatics/btp696.

Paradis, E., & Schliep, K. (2018). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*, 526-528. https://doi.org/10.1093/bioinformatics/bty633.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ross, P. A., Endersby-Harshman, N. M., & Hoffmann, A. A. (2019). A comprehensive assessment of inbreeding and laboratory adaptation in *Aedes aegypti* mosquitoes. *Evolutionary Applications*, *12*, 572-586. https://doi.org/10.1111/eva.12740.

Richards, K. (1964). Insect pests of cotton in the Ord River irrigation area. *Journal of the Department of Agriculture, Western Australia, Series 4*, *5*(2), 79-86. Available at: https://library.dpird.wa.gov.au/journal_agriculture4/vol5/iss2/4.

Schmidt, T. L., Jasper, M.-E., Weeks, A. R., & Hoffmann, A. A. (2021). Unbiased population heterozygosity estimates from genome-wide sequence data. *Methods in Ecology and Evolution*, *12*, 1888-1898. https://doi.org/10.1111/2041-210X.13659.

Song, W. H., Ballard, J. W., Yi, Y. J., & Sutovsky, P. (2014). Regulation of mitochondrial genome inheritance by autophagy and ubiquitin-proteasome system: implications for health, fitness, and fertility. *Biomed Research International*, *2014*, 981867. https://doi.org/10.1155/2014/981867.

Sridhar, J., Chinna Babu Naik, V., Ghodke, A., Kranthi, S., Kranthi, K. R., Singh, B. P. et al. (2017). Population genetic structure of cotton pink bollworm, *Pectinophora gossypiella* (Saunders) (Lepidoptera: Gelechiidae) using mitochondrial cytochrome oxidase I (COI) gene sequences from India. *Mitochondrial DNA Part A*, *28*, 941-948. https://doi.org/10.1080/24701394.2016.1214727.

Stahlke, A. R., Change, J. Chaudalayandi, S., Heu, C. C., Geib, S. M., Childers, A. K. et al. (2022). Chromosome-scale genome assembly of the pink bollworm, *Pectinophora gossypiella*, a global pest of cotton. *bioRxiv*, https://doi.org/10.1101/2022.10.07.511331.

Storey, J. D., & Tibshirani, R. (2003). Statistical significance for genome wide studies. *Proceedings of the National Academy of Sciences*, *100*, 9440-9445. https://doi.org/10.1073/pnas.1530509100.

Sundqvist, L., Keenan, K., Zackrisson, M., Prodöhl, P., & Kleinhans, D. (2016). Directional genetic differentiation and relative migration. *Ecology and Evolution*, *6*, 3461-3475. https://doi.org/10.1002/ece3.2096.

Tabashnik, B. E., Liesner, L. R., Ellsworth, P. C., & Carrière, Y. (2020). Transgenic cotton and sterile insect releases synergize eradication of pink bollworm a century after it invaded the United States. *Proceedings of the National Academy of Sciences*, *118*, e2019115118. https://doi.org/10.1073/pnas.2019115118.

Tabashnik, B. E., & Carrière, Y. (2019). Global patterns of resistance to Bt crops highlighting pink bollworm in the United States, China, and India. *Journal of Economic Entomology*, *112*, 2513-2523. https://doi.org/10.1093/jee/toz173.

Tabashnik, B. E., Patin, A. L., Dennehy, T. J., Liu, Y.-B., Miller, E., & Staten, R. T. (1999). Dispersal of pink bollworm (Lepidoptera: Gelechiidae) males in transgenic cotton that produces a *Bacillus thuringiensis* toxin. *Journal of Economic Entomology*, *92*, 772-780. https://doi.org/10.1093/jee/92.4.772.

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution*, *38*, 3022-3027. https://doi.org/10.1093/molbev/msab120.

Wan, P., Xu, D., Cong, S., Jiang, Y., Huang, Y., Wang, J., et al. (2017). Hybridizing transgenic Bt cotton with non-Bt cotton counters resistance in pink bollworm. *Proceedings of the National Academy of Sciences*, *114*, 5413-5418. https://doi.org/10.1073/pnas.1700396114.

Wang, J., Simmons, G. S., Miller, T. A., Tabashnik, B. E., & Park, Y. (2011). Global variation in the piggyBac-like element of pink bollworm, *Pectinophora gossypiella*. *Journal of Asia-Pacific Entomology*, *14*, 131-135. https://doi.org/10.1016/j.aspen.2010.09.006.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-Statistics for the analysis of population structure. *Evolution, 38*, 1358-1370. https://doi.org/10.2307/2408641.

Weir, B. S., & Goudet, J. (2017). A unified characterization of population structure and relatedness. *Genetics, 206,* 2085-2103. https://doi.org/10.1534/genetics.116.198424.

Xu, S., Schaack, S., Seyfert, A., Choi, E., Lynch, M., & Cristescu, M. E. (2012). High mutation rates in the mitochondrial genomes of *Daphnia pulex*. *Molecular Biology and Evolution*, *29*, 763-769. https://doi.org/10.1093/molbev/msr243.

Chapter 4.

Thesis discussion

4.1 General overview

My thesis aimed to demonstrate the use of population genomics as a tool to address specific gaps in invasion biology and enhance our mechanistic understanding of the invasion process. In **Chapter 2**, I provided specific examples of how genomic data can elucidate the genetic factors that underpin biological invasions, yet my research showed that such data is missing for the majority of highly invasive species, limiting our ability to prepare for, and ultimately prevent, invasions. **Chapter 3** highlighted the value of genomic data for informing biological invasion management by demonstrating that resistance to Bt in Australia is not an immediate priority concern for the management of pink bollworm , though further studies with more genomic data from global populations will aid in evaluatung its invasive potential moving forward. In **Appendix 1**, I highlighted the fact that even ecosystems as harsh and isolated as the Arctic and Antarctic are not immune to invasive species, nor will they be in the future. In this chapter, I conclude my thesis by identifying gaps and limitations in the current implementation of genomics to invasion biology research.

4.2 Considerations and caveats

In 2017, Courchamp et al. identified 24 problems in theoretical invasion biology categorised into four groups: understanding, alerting, supporting, and implementing (Courchamp et al. 2017). The underlying theme of each of the problems in these categories consistently draws back to a lack of knowledge, consensus, and communication among scientists and stakeholders, preventing the provision of effective solutions for invasive species management. Part of this problem is that invasions are complex evolutionary events, determined by a series of interactions between intrinsic (i.e., genetic) and extrinsic (i.e., environmental) factors (Courchamp et al. 2017), creating

'invasive' signatures in the genome that were, until recently, very challenging to pinpoint.

Research based on genetic markers has paved the way for tracing invasive signatures (Rollins et al. 2006) and holds great value as an affordable and, depending on the research question, effective tool for understanding biological invasions (e.g., Qin et al. 2022; Goryacheva et al. 2022; Hashem et al. 2022). This is reinforced by my finding that population *genetic* based research still largely dominates the invasion genetics literature (**Chapter 2**; Matheson & McGaughran, 2022). However, moving to a whole genome lens can improve the accuracy and resolution of a study, and address some questions that are intractable using research based on single or small numbers of genes (Chen et al. 2021; North et al. 2021). Many studies have proven the value of genomic data in various biology-adjacent disciplines – it is transforming healthcare (Josephs et al. 2019), aquaculture, agriculture (Bernatchez et al. 2017), and conservation (Supple & Shapiro 2018). Yet, invasion biology has received slower uptake – in my research, I discovered just 90 publications that used population genomics to investigate invasive species specific to the WAS list, and these were limited to only 26 species (**Chapter 2**; Appendix 2). This finding forms the basis of my future research recommendations – to sequence more genomic data (see below).

Open-access data sharing has undoubtedly driven the advancement of genomics, fostering a productive and collaborative approach to research that is more transparent, reproducible, and accountable. However, enforcing globally unrestricted access to genomic data reflects just one world view, potentially undermining the traditional knowledge, rights, and interests of indigenous communities who rightfully wish to retain control over data generated from their land, species, and water (Hudson et al. 2019; McCartney et al. 2022). Thus, open access policies must adapt to a multi-paradigmatic approach that considers the expectations of, and prioritises partnership with, indigenous

communities (McCartney et al. 2022). This is relevant for researchers studying invasion biology, where there may be some crossover between species that have the potential to be invasive, but provide cultural or social value. A valuable resource for recognising indigenous rights in data derived from genetic resources are the 'biocultural' labels and notices (see: https://localcontexts.org/labels/biocultural-labels/). These tools enhance the indigenous control of indigenous data by outlining the provenance of collections and community expectations and consents, and by connecting data to people and environments (Anderson & Hudson, 2020). Although I recommend increased sequencing of genomic data below, an important caveat is that this needs to be led by the prioritisation of data sovereignty and indigenous rights to ensure the best overall outcomes.

4.3 Recommendations

Sequence more genomic data from more populations

Species respond uniquely to environmental change (Silva et al. 2020) and invasive species are no exception: findings from one invasive system are not always translatable to another, nor necessarily are findings from different populations of the same species (Colautti & MacIsaac, 2004). Population genomics seeks to address this gap, but the only way to use genomic data to its full capacity in a biosecurity context is to *sequence more genomic data*.

A major finding from **Chapter 2** was that certain taxonomic groups were over-represented with respect to genomic resources compared to others. Notably, I found that 11/14 of mammals on the WAS list had a chromosome-level reference genome. In comparison, reference genomes existed for just 13/26 of insects (three to chromosome level) and 3/37 plants (one to chromosome level). This finding is most likely due to species-specific idiosyncrasies in the genome (e.g., polyploidy and highly repetitive content in plants), and I make the point in **Chapter 2** that long-read nanopore sequencing

is helping to bridge this gap. However, taxonomic biases and societal preferences in ecological research moving forward should be addressed.

Investing substantial resources toward the management of a few species prevents our ability to reach local and global conclusions, and implement genomic-informed biosecurity plans (Troudet et al. 2017). In New Zealand, this is exemplified by the Predator Free 2050 campaign. Established in 2016, Predator Free 2050 aims to eradicate all possums, rats, and stoats from mainland New Zealand by the year 2050 through traditional management techniques, such as 1080 pesticide and trapping. So far, $75.6, $9.6, and $11.7 million New Zealand dollars have been invested toward landscape projects, research projects, and tool and guideline development, respectively (see: https://pf2050.co.nz/app/uploads/2022/11/PF2050-Limited-Annual-Report-2022.pdf). This is a substantial investment for a highly ambitious project that, if successful, would see the eradication of just three invasive species. In New Zealand, and elsewhere, funding for invasive species management could be allocated much more efficiently – particularly toward proactive technologies that might help to prevent invasions in the first place. Ongoing advances in sequencing technologies and the reduction of sequencing costs make genomic-based tools a feasible and affordable tool worthy of investment for monitoring, management, and mitigation of invasive species (North et al. 2021).

My findings from **Chapter 3** were certainly limited by the lack of geographically widespread genomic data available in my dataset. I was able to obtain data from wild populations in the presumed invasive range (Australia) and the likely native range (India/Pakistan), as well as from laboratory populations from America, but I lacked genomic data from any other countries, despite the fact that pink bollworm persists throughout tropical America, Africa, Australasia, Asia, and Mexico (Holdaway, 1926). For this reason, I was unable to define a source population for the Australian population, and the evolutionary history I constructed was restricted to the available populations, with

the demographic modelling indicating that unsampled populations likely played an important role in pink bollworm evolutionary history. Genomic data from a larger geographic range would have improved the resolution of **Chapter 3**, and allowed me to more confidently draw conclusions about its invasive potential.

Assemble more reference genomes to explore adaptive evolution

Genomic data can be analysed to understand whether and how invasions are facilitated by adaptive evolution by identifying particular genes under selection and shedding light on the importance of pre- and post-establishment adaptations in the invasion process (McCartney et al. 2019). Such findings are critical for predicting the conditions under which invasiveness is suppressed or enhanced (Olazcuaga et al. 2020), yet in **Chapter 2**, I found that 'evolution of invasiveness' was identified as the research area that garnered the least amount of attention in invasion biology. This is likely because the majority of species on the WAS List lacked an annotated, chromosome-quality reference genome, preventing the identification of specific genes under selection.

There are currently a range of global initiatives to address this gap – for example, the Earth Biogenome project aims to characterise and catalogue reference genomes for all eukaryotic life (https://www.earthbiogenome.org). However, in order for these initiatives to create resources that produce accurate and detailed results in downstream analyses, they need to be high quality and annotated (i.e., a large percentage of sequences assembled into chromosomes and low numbers of errors and gaps; Yu et al. 2021). Furthermore, knowing the metadata (particularly date and location of sampling) for reference genome samples is fundamentally important, especially given the degree of genomic differentiation that can evolve over time between native and invasive range populations (Vaughan et al. in prep.) – new guidelines for lodging sequence data to online repositories alongside standardised metadata that ensures reproducible research have recently been released (Leipzig et al. 2021). Investing in 'pan-genome' approaches (i.e.,

sequencing a collective set of whole genomes from multiple individuals that represent the diversity of a given species, as opposed to a single non-representative reference genome; Eizenga et al. 2020) and standardised assembly and validation pipelines based on optimised procedures might bridge this gap and provide higher reliability of the data being shared in repositories (e.g., 'A reference standard for genome assembly' see: https://www.nature.com/articles/nbt.4318).

Of course, it is possible to study invasive species without a reference genome – in **Chapter 3**, I used reduced-representation DArTseq data. However, due to the lack of a reference genome at the time of analysis, I was limited in my ability to make inferences about pink bollworm's adaptive potential. While I was able to detect the number of outlier SNPs putatively under selection, I was unable to classify them (e.g., identify potential functions). Emerging research highlights the potential of classifying 'invasive' genes under selection to predict invasion success (see Olazcuaga et al. 2020; Parvizi et al. 2022; Puzey & Vallejo-Marin, 2014) – providing new data that may be of direct relevance to applied management settings. Since the completion of my pink bollworm analysis for this thesis, an annotated reference genome has become available, opening the door for a range of future analyses that can more comprehensively investigate the underlying factors that drive evolution of resistance to Bt (Stahlke et al. 2022).

Empower citizen scientists to aid management efforts

Encouraging citizens to contribute to the collection of scientific data has several clear benefits, including improved public understanding and literacy of scientific issues (Novoa et al. 2017), reduced cost and/or labour of data collection, and access to data from large geographical areas (Crall et al. 2009). For invasive species management – where early detection and subsequent rapid action is key for mitigation – citizen science can potentially speed up biosecurity responses. In fact, the Ministry for Primary Industries (MPI) has many resources accessible to the public for this reason – including how to

identify species of high invasive risk, their threat to New Zealand, and what to do if one is found (e.g., https://www.mpi.govt.nz/biosecurity/major-pest-and-disease-threats/brown-marmorated-stink-bug-threat-to-nz-and-identification/). Furthermore, large-scale, community-led management projects such as Capital Kiwi, which aims to use traps to eradicate mustelids from 23,000 ha of central Wellington, continue to play an important role in the monitoring and control of invasive species in New Zealand (https://www.capitalkiwi.co.nz/the-project/). Moving forward, emerging sequencing technologies, such as environmental DNA (eDNA; the identification of genetic material derived from environmental samples, such as soil and water) have potential to complement traditional monitoring technologies and detect new incursions more rapidly. Collecting eDNA is not only effective and affordable, but also a standardised and simple procedure, and is already providing great value for citizen scientists (Novoa et al. 2017; https://www.wilderlab.co.nz/blog/thames-high-school-students-find-possible-fish-migration-barrier-with-edna) and volunteer groups (Agersnap et al. 2022). It would be great to see this trend continue, and potentially act as a gateway toward improving literacy and reducing hesitancy among stakeholders with concerns about genomic technologies and their application to invasion biology.

4.4 Concluding remarks

The issues examined in my thesis collectively illustrate that population genomics is a promising tool for understanding the invasion process, informing future invasions, and potentially leading to more effective biosecurity outcomes. I found that invasion genomics researchers are expanding the amount of data being generated for invasive species, but the field is overall still in its infancy. Population genomic analyses were highly beneficial for elucidating information about pink bollworm in Australia, but would have benefited from a greater geographic scale of population genomic data, as well as access to a (then unavailable) high-quality reference genome. An increased investment

towards genomic data generation for invasive species would have addressed each of these issues and, moving forward, genome sequencing initiatives that carefully consider data sovereignty and metadata standards will lead to better research capabilities, collaboration, and equity among researchers and local communities. Furthermore, applying new genomic tools to the long-standing challenge of invasive species management promises more effective armaments for the current biosecurity toolbox and the delivery of better outcomes that benefit the biodiversity of ecosystems alongside the well-being/hauora of their caretakers.

4.5 Literature cited

Agersnap, S., Sigsgaard, E. E., Jensen, M. R., Avila, M. D. P., Carl, H., Møller, P. R. et al. (2022). A national scale "BioBlitz" using citizen science and eDNA metabarcoding for monitoring coastal marine fish. *Frontiers in Marine Science, 9*. https://doi.org/10.3389/fmars.2022.824100.

Anderson, J. & Hudson, M. (2020). The Biocultural Labels Initiative: Supporting Indigenous rights in data derived from genetic resources. *Biodiversity Information Science and Standards*, *4*. https://doi.org/10.3897/biss.4.59230.

Bernatchez, L., Wellenreuther, M., Araneda, C., Ashton, D. T., Barth, J. M. I., Beacham, T. D. et al. (2017). Harnessing the power of genomics to secure the future of seafood. *Trends in Ecology and Evolution, 32,* 665-680. https://doi.org/10.1016/j.tree.2017.06.010.

Chen, T., Chen, X., Zhang, S., Zhu, J., Tang, B., Wang, A. et al. (2021). The genome sequence archive family: Toward explosive data growth and diverse data types. *Genomics, Proteomics and Bioinformatics, 19,* 578-583. https://doi.org/10.1016/j.gpb.2021.08.001.

Colautti, R. I., & MacIsaac, H. J. (2004). A neutral terminology to define 'invasive' species *Diversity and Distributions*, *10*, 135-141. https://doi.org/10.1111/j.1366-9516.2004.00061.x.

Courchamp, F., Fournier, A., Bellard, C., Bertelsmeier, C., Bonnaud, E., Jeschke, J. M. et al. (2017). Invasion biology: Specific problems and possible solutions. *Trends in Ecology and Evolution, 32,* 13-22. https://doi.org/10.1016/j.tree.2016.11.001.

Crall, A. W., Newman, G. J., Jarnevich, C. S., Stohlgren, T. J., Waller, D. M., & Graham, J. (2010). Improving and integrating data on invasive species collected by citizen scientists. *Biological Invasions, 12,* 3419-3428. https://doi.org/10.1007/s10530-010-9740-9.

Eizenga, J. M., Novak, A. M., Sibbesen, J. A., Heumos, S., Ghaffaari, A., Hickey, G. et al. (2020). Pangenome graphs. *Annual Reviews of Genomics and Human Genetics. 21*, 139-162. https://doi.org/10.1146/annurev-genom-120219-080406.

Goryacheva, I. I., Schepetov, D. M., Blekhman, A. V., & Zakharov, I. A. (2022). On the genetic structure of *Harmonia axyridis* (Coleoptera, Coccinellidae) populations in native and invasive ranges, view from a position of molecular genetics. *Russian Journal of Genetics*, *58*, 1118-1128. https://doi.org/10.1134/S1022795422090095.

Hashem, M. H., Alshaya, D. S., Jalal, A. S., Abdelsalam, N. R., Abd El-Azeem, R. M., Khaled, A. E. et al. (2022). Genetic divergence and phylogenetic relationship of the rabbitfish *Siganus rivulatus* inferred from microsatellite and mitochondrial markers. *Journal of King Saud University - Science*, *34*, 101943. https://doi.org/10.1016/j.jksus.2022.101943.

Holdaway, F. G. (1926). The pink bollworm of Queensland. *Bulletin of Entomological Research*, *17*, 67-83.

Hudson, M., Garrison, N. A., Sterling, R., Caron, N. R., Fox, K., Yracheta, J. et al. (2020). Rights, interests and expectations: Indigenous perspectives on unrestricted access to genomic data. *Nature Reviews Genetics*, *21*, 377-384. https://doi.org/10.1038/s41576-020-0228-x.

Josephs, K. S., Berner, A., George, A., Scott, R. H., Firth, H. V., & Tatton-Brown, K. (2019). Genomics: The power, potential and pitfalls of the new technologies and how they are transforming healthcare. *Clinical Medicine Journal*, *19*, 269-272. https://doi.org/10.7861/clinmedicine.19-4-269.

Leipzig, J., Nüst, D., Hoyt, C. T., Ram, K., & Greenberg, J. (2021). The role of metadata in reproducible computational research. *Patterns, 2,* 100322. https://doi.org/10.1016/j.patter.2021.100322.

Matheson, P., & McGaughran, A. (2022). Genomic data is missing for many highly invasive species, restricting our preparedness for escalating incursion rates. *Scientific Reports*, *12*, 13987. https://doi.org/10.1038/s41598-022-17937-y.

Mc Cartney, A. M., Anderson, J., Liggins, L., Hudson, M. L., Anderson, M. Z., TeAika, B. et al. (2022). Balancing openness with Indigenous data sovereignty: An opportunity to leave no one behind in the journey to sequence all of life. *Proceedings of the National Academy of Sciences*, *119*, e2115860119. https://doi.org/10.1073/pnas.2115860119.

McCartney, M. A., Mallez, S., & Gohl, D. M. (2019). Genome projects in invasion biology. *Conservation Genetics*, *20*, 1201-1222. https://doi.org/10.1007/s10592-019-01224-x.

North, H. L., McGaughran, A., & Jiggins, C. D. (2021). Insights into invasive species from whole-genome resequencing. *Molecular ecology*, *30*, 6289-6308. https://doi.org/https://doi.org/10.1111/mec.15999.

Novoa, A., Dehnen-Schmutz, K., Fried, J., & Vimercati, G. (2017). Does public awareness increase support for invasive species management? Promising evidence across taxa and landscape types. *Biological Invasions*, 19, 3691-3705. https://doi.org/10.1007/s10530-017-1592-0.

Olazcuaga, L., Loiseau, A., Parrinello, H., Paris, M., Fraimout, A., Guedot, C. et al. (2020). A whole-genome scan for association with invasion success in the fruit fly *Drosophila suzukii* using contrasts of allele frequencies corrected for population structure. *Molecular Biology and Evolution*, *37*, 2369-2385. https://doi.org/10.1093/molbev/msaa098.

Parvizi, E., Dhami, M. K., Yan, J., & McGaughran, A. (2022). Population genomic insights into invasion success in a polyphagous agricultural pest, *Halyomorpha halys*. *Molecular ecology*, *0*, 1-14. https://doi.org/10.1111/mec.16740.

Puzey, J., & Vallejo-Marín, M. (2014). Genomics of invasion: diversity and selection in introduced populations of monkeyflowers (Mimulus guttatus). *Molecular Ecology, 23*, 4472-4485. https://doi.org/10.1111/mec.12875.

Qin, T., Lee, C., Li, S., Cavalcante, R. G., Orchard, P., Yao, H. et al. (2022). Comprehensive enhancer-target gene assignments improve gene set level interpretation of genome-wide regulatory data. *Genome Biology, 23*, 105. https://doi.org/10.1186/s13059-022-02668-0.

Rollins, L., Woolnough, A., & Sherwin, W. (2006). Population genetic tools for pest management: A review. *Wildlife Research, 33*, 251-261. https://doi.org/10.1071/WR05106.

Silva, A. P., Mukherjee, S., Ramakrishnan, U., Fernandes, C., & Björklund, M. (2020). Closely related species show species-specific environmental responses and different spatial conservation needs: *Prionailurus* cats in the Indian subcontinent. *Scientific Reports, 10*, 18705. https://doi.org/10.1038/s41598-020-74684-8.

Stahlke, A. R., Chang, J., Chudalayandi, S., Heu, C. C., Geib, S. M., Scheffler, B. E. et al. (2022). Chromosome-scale genome assembly of the pink bollworm, *Pectinophora gossypiella*, a global pest of cotton. *bioRxiv*, https://doi.org/10.1101/2022.10.07.511331.

Supple, M. A., & Shapiro, B. (2018). Conservation of biodiversity in the genomics era. *Genome Biology, 19*, 131. https://doi.org/10.1186/s13059-018-1520-3.

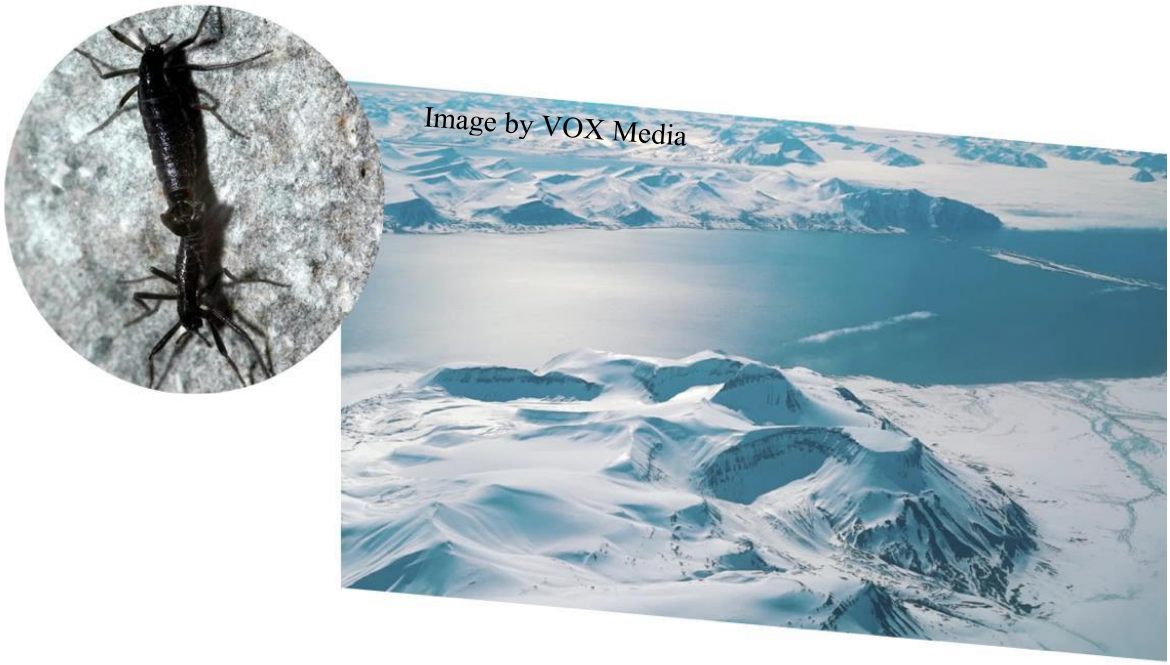
Troudet, J., Grandcolas, P., Blin, A., Vignes-Lebbe, R., & Legendre, F. (2017). Taxonomic bias in biodiversity data and societal preferences. *Scientific Reports, 7*, 9132. https://doi.org/10.1038/s41598-017-09084-6.

Yu, J., Hulse-Kemp, A. M., Babiker, E., & Staton, M. (2021). High-quality reference genome and annotation aids understanding of berry development for evergreen

blueberry (*Vaccinium darrowii*). *Horticulture Research*, *8*, 228. https://doi.org/10.1038/s41438-021-00641-9.

Appendix 1.

How might climate change affect adaptive responses of polar arthropods?



Submitted as: Matheson, P. & McGaughran, A. (2022). How might climate change affect adaptive responses of polar arthropods? *Diversity, submitted on 10 December 2022, and currently under review.*

Contributions: A.M and P.M conceptualised the manuscript, P.M wrote the first draft and created the figures, and A.M and P.M together finalised the text.

A1.1 Abstract

Climate change is expected to impact the global distribution and diversity of arthropods, with warmer temperatures forcing species to relocate, acclimate, adapt, or go extinct. The Arctic and Antarctic regions are extremely sensitive to climate change, and have displayed profound and variable changes over recent decades, including decreases in sea ice extent, greening of tundra, and changes to hydrological and biogeochemical cycles. It is unclear how polar-adapted arthropods will respond to such changes, though many are expected to be at great risk of extinction. Here, we review the adaptive mechanisms that allow polar arthropods to persist in extreme environments and discuss how the effects of climate change at the poles will likely favour non-native species, or those with the ability to rapidly evolve and/or acclimate. We find that physiological, behavioural, plastic, and genetic data are limited in scope for polar arthropods, and research on adaptive responses to change is scarce. This restricts our ability to predict how they may respond to a warming climate. We call for a greater investment in research that specifically targets the ecology and evolution of these taxa, including genomic and transcriptomic approaches that can evaluate the potential for plastic and evolved environmental responses.

A1.2 Introduction

Climate change is having increasingly detrimental impacts on the natural environment. In particular, potential warming beyond 1.5°C is predicted to increase extreme weather events (e.g., floods, droughts, cyclones), impact human livelihoods by limiting potable water/food availability and damaging infrastructure, and threaten many species with extinction due to physically intolerable conditions, changing the entire landscape of ecosystems (Pörtner et al. 2022).

Situated at opposite ends of the Earth, the polar regions (Arctic and Antarctic) are some of the most harsh, remote, and inhospitable habitats on the planet (Figure A1.1). Where the Arctic is an ocean covered by perennial sea ice and surrounded by land, Antarctica is a continent covered by thick ice and surrounded by the Southern Ocean. Both are considered 'barometers' of global health for the role they play in providing planetary-scale balance and circulation – regulating energy exchange between climatic and oceanic systems, driving atmospheric and weather processes, acting as sinks for carbon dioxide, and providing thermal density gradients that drive thermohaline circulation (Monteiro et al. 2022; Shadwick et al. 2021). Observed and projected climatic changes in the Anthropocene are comparable to some of the largest environmental changes of the past 65 million years, though human activities such as fossil fuel combustion and land-use change have meant that this is occurring much more rapidly than previously observed in the geological record (Diffenbaugh & Field, 2013). Despite their isolation from civilisation, the Arctic and Antarctic polar regions have responded rapidly to human induced climate change, with small changes in temperature having large impacts on species diversity and distribution (Høye, 2020; Hughes et al. 2021; Koenigk et al. 2020).

Arthropods – insects, arachnids, myriapods, and crustaceans – comprise the bulk diversity of species on Earth, with some taxa that have survived all five mass extinction

events in evolutionary history (e.g., horseshoe crabs; Rudkin & Young, 2009). As such, they are fundamental members of many ecosystems, playing crucial roles in food webs, pollination, decomposition, nutrient cycling, and pest control (Høye & Culler, 2018). Polar regions are no exception, with some biota that have been present for millions of years throughout glacial cycles and since the breakup of Gondwana, and others that have arrived since the last glacial maximum (< 30 kya; Convey et al. 2008; Hughes et al. 2010). Resident polar arthropods must endure environmental stresses beyond the physiological limits of many other species, such as extremely low temperatures (routinely < -15°C), intense heat stress due to solar radiation that exceeds air temperature (> 20°C) during summer months, dry and windy conditions, and 24 hours of darkness during winter (Michaud et al. 2008; Turner et al. 2021; Vanstreels et al. 2021). As a result, these taxa have evolved unique physiological adaptations to tolerate their local environment.

Nevertheless, arthropods are highly sensitive to environmental changes owing to their diverse life-histories, typically short generation times, large population sizes, and often high rates of adaptability (Høye & Culler, 2018). Therefore, they are excellent research models for elucidating the extrinsic and intrinsic mechanisms that promote survival in a rapidly changing world. Here, we review the literature to ask how climate change might affect arthropods in polar regions. We first review current Arctic and Antarctic arthropod diversity, and some of the most important adaptations that have allowed these species to persist in extreme polar environments. Next, we evaluate how climate change might affect arthropod diversity and distribution in the polar regions, with comments on biological invasions and ecological responses of native arthropods. We consider examples from the high Arctic, low Arctic, and sub-Arctic in our definition of the Arctic region; and both continental Antarctica and nearby sub-Antarctic maritime islands in our definition of the Antarctic region (Figure A1.1).

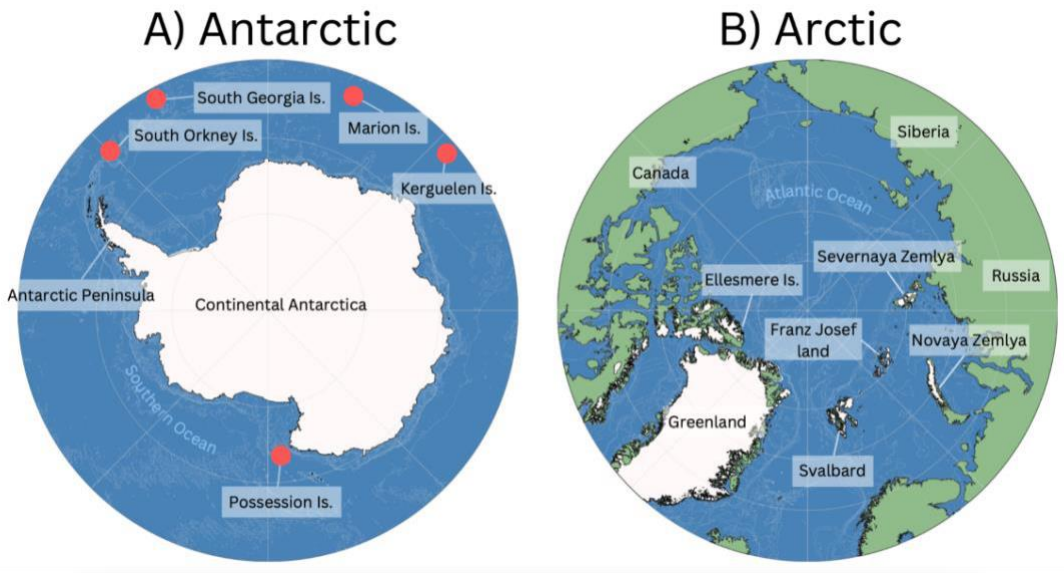


Figure A1.1 Map of A) Antarctic and B) Arctic regions identifying key locations. Red circles in A) represent relative locations of the main sub-Antarctic Islands.

A1.3 Polar arthropods are abundant and diverse, and exhibit a range of adaptations to the cold

Abundance and diversity

The biodiversity of polar arthropods was once thought to be scarce owing to an overly simplistic understanding of their ecology and roles in the ecosystem. For instance, past Arctic food web models collated individual species into related groups (e.g., considered all beetle species one entity), neglecting to account for the wide range of feeding specialisations and adaptations that exist at the species level (Hodkinson et al. 2003). In reality, arthropods are the most diverse and abundant phylum found near the poles, accounting for approximately 90% of all known species there, though many are small and have only basic body forms (Giribet & Edgecombe, 2019; Hodkinson et al. 2003).

In Antarctica, most terrestrial arthropods are restricted to the less harsh sub-Antarctic islands. For example, at least 41 endemic species of insects and spiders occupy Possession and Kerguelen Islands alone (Maurice & Philippe, 2021). However, some species inhabit mainland Antarctica in high densities – including soil microarthropods

(springtails, mites), most of which are endemic to the continent (Sinclair et al. 2006), and two species of chironomid midges (endemic *Belgica antarctica* and native *Parochlus steineii*). The majority of Arctic arthropod diversity is similarly distributed in less harsh environments of the sub-Arctic and low-Arctic regions, and predominantly consists of springtails, lice, chironomid midges, crane flies, aphids, beetles, moths, and wasps. In the high-Arctic, the same orders exist, though with much fewer species (Hodkinson et al. 2003; Høye & Forchhammer, 2008).

Adaptations to the cold

All polar-dwelling arthropods are exposed to cold temperatures, often below the freezing point of their body fluids. Resident species, at least to some degree, have a variety of adaptations that allow them to tolerate these local conditions, including physiological, plastic, behavioural, and genetic (Figure A1.2). For example, some microarthropods (e.g., ticks) have developed parasitic relationships with warm-blooded hosts (e.g., seals or birds), while others buffer against immediate environmental temperatures by living in a microclimate (e.g., soil or water), or migrating seasonally (Vanstreels et al. 2020). By far the greatest focus of research to date has been on physiological adaptations of polar arthropods; we outline some of these below.

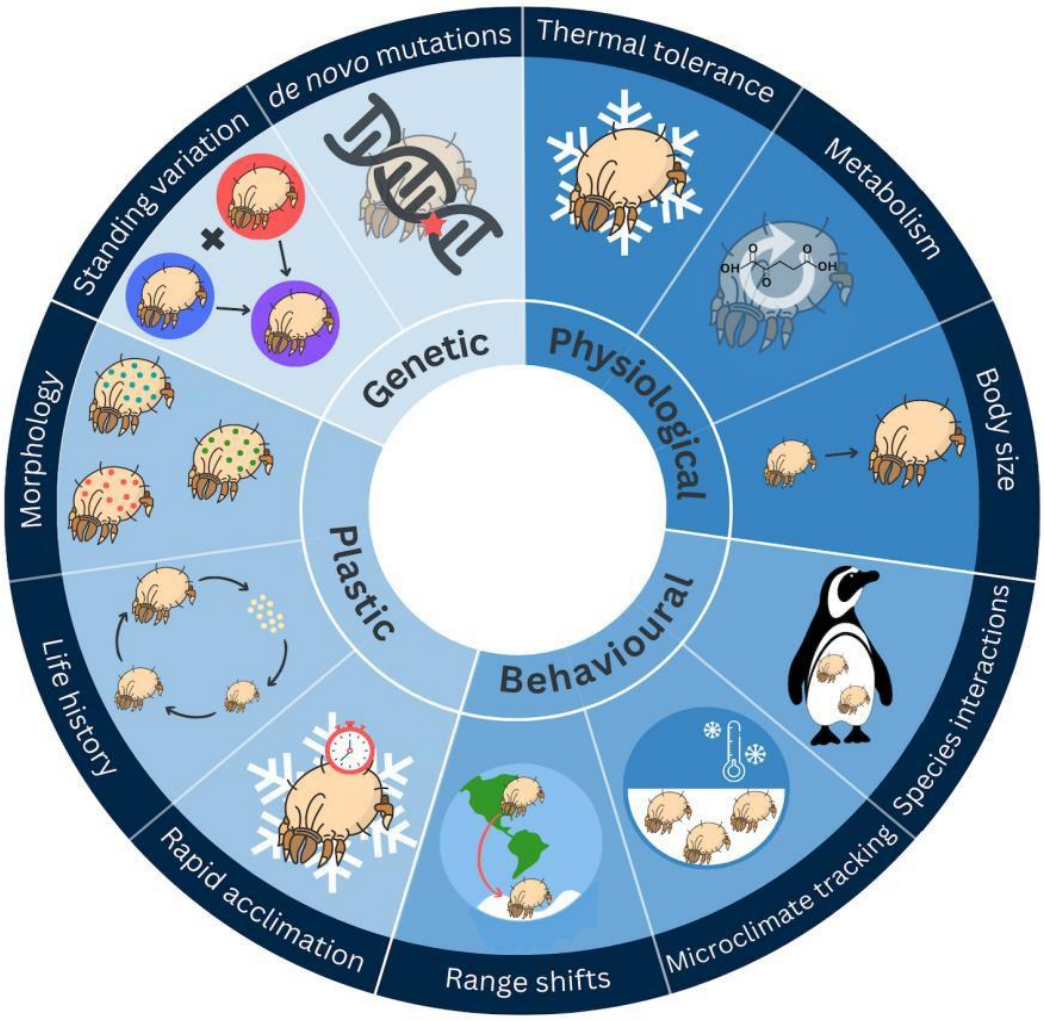


Figure A1.2 The potential physiological, behavioural, plastic, and genetic responses associated with surviving life at the poles. For simplicity, we depict discrete categories, though responses may fall into > 1 category.

Thermal tolerance

Thermal tolerance is the ability of an organism to survive short- or long-term exposure to temperature extremes. The ectothermic nature of arthropods means that temperature plays a major role in the effectiveness of their chemical and biological functions, therefore, thermal tolerance is a critically important trait for survival and success at the poles (Sinclair et al. 2015). Polar arthropods can be classified as either freeze-avoidant or freeze-tolerant depending on their strategy to survive overwintering, although the factors that determine which species employ which strategy are largely unclear.

Freeze-avoidance

Most polar arthropods are freeze-avoidant – they survive subzero temperatures in a supercooled state by lowering the point at which their body fluids freeze, referred to as the super-cooling point (SCP; Rozsypal & Košťál, 2018; Sinclair et al. 2006). This is achieved through the accumulation of antifreeze and/or heat shock proteins (Duman et al. 2004), or by the removal of ice-nucleating particles, such as food and microbes in the digestive tract. Bimodal SCP distributions are commonly described, with less cold-hardy species being more active and higher-foraging, while more cold-hardy species are thought to be non-feeding and therefore lacking in ice nucleators from food in the gut (Sinclair et al. 2006; Worland et al. 2005). However, this is not the case for all species. For example, the Antarctic springtail *Gomphiocephalus hodgsoni* prevents inoculative freezing of body fluids to temperatures as low as -35.4°C through the accumulation of thermal hysteresis proteins and increased haemolymph osmolality (Sinclair et al. 2006; Sinclair et al. 2001; Zettel, 1984).

Freeze-tolerance

Freeze-tolerant arthropods can withstand the formation of ice between their body tissues. They prevent intra-tissue damage by producing ice-nucleating and heat-shock proteins (Rinehart et al. 2007), and/or accumulating cryoprotectants (polyols and sugars) that control the rate of freezing and protect the membrane bilayer (Storey & Storey, 2012). Freezing is often initiated slowly at relatively high subzero temperatures in order to promote the growth of small, site-specific ice crystals in the extracellular spaces, as opposed to the rapid formation of ice that can cause injury (Duman et al. 2004). For example, an Arctic species of stonefly, *Nemoura arctica*, cools to around -1.5°C before freezing and can then survive temperatures as low as -15°C for over two weeks by increasing hemolymph glycerol concentrations and antifreeze proteins (Walters et al. 2009).

Cryoprotective dehydration

Cryoprotective dehydration (CPD) is a form of freeze-tolerance that occurs when an organism loses extreme amounts of body fluids to reach a vapour pressure equilibrium with surrounding ice. It has only been observed in a handful of arthropod species, including several Arctic springtails such as *Megaphorura arctica*, *Hypogastrura viatica*, and *Folsomia quadrioculata* (Clark et al. 2009; Sørenson et al. 2011), the Alaskan beetle *Cucujus clavipes puniceus* (Sfromo et al. 2010) and larvae of the Antarctic midge *Belgica antarctica*. In *B. antarctica*, ~40% of the body fluid is lost and high concentrations of solutes, such as osmolytes, accumulate until vapour pressure equilibrium with the surrounding ice is reached. This prevents internal freezing, allowing survival over 7-8 months of continuous subzero temperatures in a matrix of soil and ice (Elnitsky et al. 2008).

Rapid cold hardening

Polar environments provide increasingly variable microhabitat temperatures and unpredictable freeze-thaw periods in any season throughout the year. In response, some species have the ability to plastically enable cold tolerance via 'rapid cold hardening' (RCH) – a fast response to cold that can last a matter of minutes or hours (Teets et al. 2020) – providing an ecological advantage over species that must acclimate or adapt over much longer timescales. The mechanisms that underpin RCH are relatively unclear: it was thought that since RCH has similar ecological outcomes to seasonal freeze-tolerance or avoidance (i.e., higher survival rate, increased fitness at low temperatures), it is achieved through similar mechanisms albeit at shorter timescales. However, recent work suggests that RCH is rather mechanistically distinct to seasonal cold-hardening. For instance, cryoprotectant/antifreeze protein synthesis and up-regulated stress related genes (i.e., heat shock proteins) are important factors associated with seasonal cold hardening, but either are not involved in, or act ambiguously during, RCH. Factors that are important

to RCH include calcium signalling, inhibition of apoptotic cell death, membrane restructuring, and adjustments to ion transport mechanisms (Teets et al. 2013).

RCH has been identified in some native polar arthropods and microarthropods (e.g., Antarctic mites, *Alakaozetes antarcticus* and *Halozete belgicae*; Worland & Convey, 2001) and is hypothesised to reduce the developmental cost of employing long-term overwintering strategies, such as freeze tolerance. However, others argue that RCH offers little advantage to some polar arthropods (such as Arctic springtail *Hypogastrura tullbergi*) that migrate through soil profiles during cold snaps, as this often offers sufficient protection from short-term low temperatures (Hawes et al. 2006).

Metabolism

While metabolism is mainly driven by the availability of food and oxygen, environmental temperatures have profound effects on metabolism, with higher metabolic rates across all phyla generally occurring in warmer climates and vice versa (Clarke & Fraser, 2004; Schulte, 2015). However, some research proposes that ectotherms from cold habitats have elevated metabolisms compared to those from warm habitats, owing to short, cool growing seasons that select for an increase in growth rates and development times – referred to as 'metabolic cold adaptation' (Addo-Bediako et al. 2002; Ayres & Scriber; Terblanche et al. 2009; Williams et al. 2016).

Metabolic cold adaptation does not appear to be a phenomenon characterising all polar arthropods and warrants further research at the species-scale (Lardies et al. 2004). Nevertheless, metabolism produces the energy required for movement, growth, healing, feeding, digestion, and reproduction; thus the ability to metabolise efficiently in cold environments may represent an ecologically advantageous trait for polar arthropods (Nespolo et al. 2003). For example, during chill coma, arthropods lose ionic and osmotic homeostasis as the cellular membrane structure is altered and metabolic processes are down-regulated. As temperature increases again, whole-organism recovery involves

restoring neuromuscular movement, repairing damage, and re-establishing ionic and osmotic gradients to restore homeostasis, all of which is energetically expensive. Thus, faster metabolism (and therefore recovery time) may allow for additional opportunities for foraging, dispersal, and reproduction, while reducing susceptibility to predation (MacMillan & Sinclair, 2011; Williams et al. 2016). Indeed, arthropods from high-latitude environments may have a quicker chill coma recovery time than those from temperate regions (David et al. 2003). Meanwhile, metabolomic studies have revealed abundant variation in core metabolic enzymes between polar species in response to thermal stress (Marden, 2013; Michaud et al. 2008). For example, *B. antarctica* enhances thermotolerance by increasing concentrations of metabolites, such as *a*-ketoglutarate (a Krebs cycle intermediate), to aid the synthesis of cryoprotectants (Michaud et al. 2008). In contrast, the Arctic seed bug, *Nysius groenlandicus*, actively alters metabolite levels on a daily basis, with higher concentrations of sugars found in individuals that were caught at the lowest daily field temperatures (Noer et al. 2022).

Body size

Body size is integral to an organisms' ecological success – larger individuals tend to produce more offspring and live longer, and may be better at competing and avoiding predators than smaller individuals (Horne et al. 2017). Body size is influenced by many factors, including metabolic energy, resource availability, and selection (both sexual and predation-driven; Klok & Harrison, 2013; Verberk et al. 2021).

The temperature size rule (TSR) describes how individuals maintained at low temperatures tend to grow slower, but attain a larger body size upon maturity (Aguilar-Alberola & Mesquita-Joanes, 2014; Bowden et al. 2015). Similarly, Bergmann's rule describes how larger individuals are found in higher latitudes and colder environments (i.e., higher altitudes; Bowden et al. 2015). Body size distribution patterns have been found to be largely consistent with TSR and Bergmann's rule for many endothermic

species, such as birds and mammals (Scriven et al. 2016). Some ectothermic species follow these trends to a degree, though much less consistently, while other species follow a contradictory trend where body size decreases in response to cold temperatures ('converse Bergmann's rule'; Shelomi, 2012. For example, wing length in two high-Arctic butterfly species (*Boloria chariclea* and *Colias hecla*) decreased significantly between the years 1996 and 2013 in response to warmer Arctic temperatures, supporting TSR (Bowden et al. 2015), whereas small, simple insects characterised by rapid development (such as aphids and small flies) are dominant in both polar regions, in counter to Bergmann's rule (Danks, 2004). Variability in body size is likely to be idiosyncratic, depending largely on a number of species-specific and abiotic (e.g., oxygen) variables (Horne et al. 2017).

A1.4 Climate change will bring new challenges for polar arthropods: how might they cope?

Challenges of climate change

Anthropogenic activities, such as fossil fuel combustion, deforestation, land-use change, and pollution have variable and profound impacts on the global climate and are already driving changes in species' range extent (Thuiller et al. 2008). The poles are no exception.

The Arctic warms at roughly twice the rate as the rest of the globe due to complex interactions between the cryosphere, atmosphere, and ocean, referred to as 'Arctic Amplification' (Cohen et al. 2020; Koenigk et al. 2020). Consequently, Arctic sea ice has rapidly declined in thickness and extent over the last four decades (Maksym, 2019; Meier et al. 2014), particularly the Greenland Ice Sheet, which is melting faster than any climate models have projected (Wunderling et al. 2020). Future climate projections suggest that this trend will continue well into the 21st century, with potential warming by an additional

1.5-4°C predicted for scenarios that include completely ice-free summers (Jahn, 2018; Koenigk et al. 2020; Overland & Wang, 2013).

Climate change in Antarctica has not been uniform. Despite the fact that the Antarctic Peninsula, Scotia Arc, and Magellanic sub-Antarctic were some of the most rapidly warming areas on Earth for the second half of the 20th century, Antarctic sea ice extent across the entire continent has increased slightly since the 1970s (Contador et al. 2020; Convey & Peck, 2019). This paradox has been attributed to the ozone hole that occurs over Antarctica during the austral summer, which causes stratospheric radiative cooling from ozone depletion and strengthened wind speeds (Bandoro et al. 2014; Convey & Peck, 2019; Turner et al. 2009). When the ozone layer repairs, as it is predicted to do within the next century under a 'business as usual' scenario (i.e., continued increase of greenhouse gas emissions; Solomon et al. 2016, Antarctic surface temperatures are expected to increase by ~4°C by the year 2100 (Hughes et al. 2021).

Even small changes in temperature can cause irreversible changes to polar landscapes. For example, tundra greening in the low Arctic due to warming creates opportunities for plants in nearby sub-Arctic regions to expand their range toward higher latitudes, as seen in *Hieracium aurantiacum* (orange hawkweed), *Elodea nuttallii* (western waterweed), and *Bromus tectorum* (cheatgrass) – all of which are capable of dramatically changing the local ecosystem through changes in biogeochemical cycles, fire regimes, and hydrological processes (Lassuy & Lewis, 2013).

Potential responses to a changing climate

The ability for a species to adapt to climate change depends on its life-history traits and genetic architecture, both of which will affect the speed at which responses may occur (i.e., rapid/plastic versus slower/genetic). Despite this, most methods for evaluating responses to climate change rely largely on correlative species distribution models and limited ecological community data, neglecting to consider evolutionary potential (i.e.,

adaptive response) or species-specific idiosyncrasies (Bush et al. 2016; Kellermann & van Heerwaarden, 2019). Yet species in polar regions will almost certainly respond differentially to climate change, even those that occupy similar niches and/or are closely related. For example, larvae of the endemic Antarctic midge *B. antarctica* are freeze tolerant to -14°C, but die within a week at 10°C, and only survive a couple of hours at 30°C (Lee et al. 2006). By comparison, the native Antarctic midge *P. steinenii* has a wide thermal tolerance across all life stages, remaining active between temperatures ranging from -5°C to 28.6°C (Contador et al. 2020), while the crane fly *Trichocera maculipennis* – a recent Antarctic coloniser – has a lower thermal tolerance limit similar to *P. steineii* but a higher limit of 30.1°C (Pertierra et al. 2021). Though all three species will experience the same environmental changes, *P. steinenii* and *T. maculipennis* are better set to benefit from a warmer climate and may rapidly colonise new habitats as they become available.

Studies on species life histories have found that, while warming may benefit Arctic herbivores (de Sassi & Tylianakis, 2012), it will likely negatively impact detritivores, such as Collembola (Hodkinson et al. 1998) as well as certain above-ground and soil-dwelling Arctic arthropods (Hodkinson et al. 2003). Thus, in polar regions as elsewhere, a major consequence of ongoing climate change will be the redistribution of arthropod diversity: a) the reduction of native species diversity through extinction; b) the movement of species as they track changes in temperature to meet optimal conditions determined by their physiology (including the poleward shift of endemic species and the arrival of species inadvertently introduced to polar regions with increased survival and the potential to become invasive); and/or c) the response of native species to additional environmental and ecological pressures through changes in behaviour, physiology, plasticity, and/or genetics (Giribet & Edgecombe, 2019; Hodkinson et al. 2003; Figure A1.2). However, the intricacies of such responses are largely uncertain due to a lack of

relevant knowledge at the species level. And, as noted above, arthropods are likely to respond differentially to climate change: though many will respond negatively (e.g., heat stress, desiccation, phenological mismatches), some will benefit from warmer temperatures through alleviated cold thermal stress, lengthened active seasons that promote growth and reproduction, and increased habitat availability provided by new ice-free environments (Block et al. 1994; Schmidt et al. 2016; Wallingford et al. 2020). Overall, climate change may drive a variety of adaptive changes in species at the poles (Høye, 2020), as outlined in the following section.

Potential responses to a changing climate

Biological invasions occur when species are transported to regions beyond their native range by humans and there negatively impact native biodiversity, health, and economies. The invasion process can be viewed as a series of stages (transport, introduction, establishment, and spread), with the graduation of each stage dependent on overcoming specific barriers (e.g., founder effects, abiotic, and ecological pressures; Blackburn et al. 2011).

The quantity and diversity of individuals being introduced to the Arctic and Antarctic every day is high. For example, 1,376 individual alien invertebrates spanning 98 families were introduced to Antarctica during the summer field season in 2012-2013 (Houghton et al. 2016). However, very few individuals survive the journey, and even less survive the extreme physical conditions they encounter when they arrive at the poles (Frenot et al. 2005; Houghton et al. 2016). Increased patterns of human use in these areas (e.g., for research and tourism, shipping, fishing, etc.) will affect the rate of alien species introductions (Figure A1.3), while changes in climate can be expected to simultaneously select for, or otherwise promote, their survival upon arrival by creating conditions that are tolerable, if not favourable to invasive species (Alsos et al. 2015; Baird et al. 2020;

Duffy et al. 2017; Hughes & Worland, 2010; Lassuy & Lewis, 2013; Stachowicz et al. 2002; Williamson & Fitter, 1996).

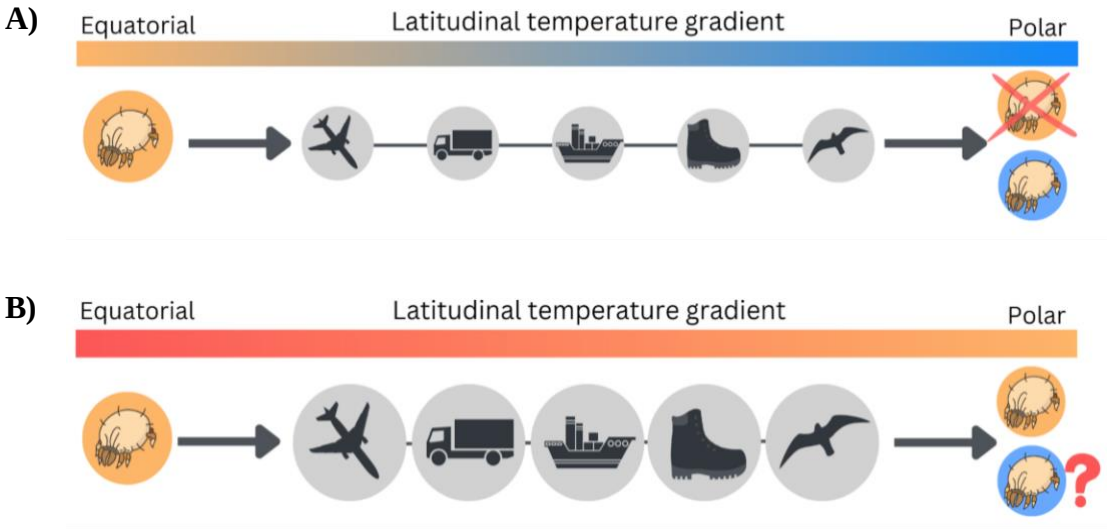


Figure A1.3 Depiction of possible arthropod distribution changes in response to warming temperatures associated with climate change. Blue circles represent native populations that prefer cold temperatures (i.e., polar regions) and orange circles represent non-native populations that prefer warm temperatures (i.e., non-polar regions). Grey circles represent different means of mediated dispersal, with size representing relative transport events to the poles (i.e., larger grey circles indicate higher propagule pressure). A) Under a normal latitudinal temperature gradient with fewer introduction events, non-native species are unlikely to establish populations upon arrival to the poles due to mismatches between local environment (e.g., air temperature) and physiology; B) Warmer temperatures and increased transport events may facilitate the establishment of non-native species at the poles by creating conditions within their optimal range, or facilitating rapid evolution in the new environment.

All 14 of the known non-native species established in Antarctica are found on the Antarctic Peninsula – the region where most warming has occurred since the 1950s – and almost exclusively in the direct vicinity of research centres and visitor sites (Hughes et al. 2020). Less data exists for invasive arthropods in the Arctic. However, a thorough inventory of species in the Svalbard region of the European high Arctic, includes ten established alien arthropods (e.g., the springtail *Hypogastrura assimilis*), seven vagrants

(e.g., the moth *Plutella xylostella*), and many other observed invasive species (e.g., the mite *Thanatus formicinus*), all of which were likely transported through imported soils (Coulson, 2015).

Of course, one of the greatest invasive risks at both poles is the local assisted dispersal of established or even native species to new areas, as such species are already tolerating polar conditions in their local habitats (Everatt et al. 2012). For example, the chironomid midge *Eretmoptera murphyi* was introduced to Signy Island (South Orkney Islands) from its native range on South Georgia Island and successfully established a population there due to pre-existing thermal adaptations (Worland, 2010); it now significantly impacts local litter turnover and availability of nutrients (Bartlett et al. 2021; Hughes et al. 2013). With at least 25 distinct biogeographic regions within Antarctica and the surrounding maritime sub-Antarctic (Chown et al. 2016), and movement of personnel and cargo between these regions occuring at high frequency, the risk of human-assisted movement of species between local populations is high – especially for environments that are becoming increasingly fragmented (Contador et al. 2020).

Plastic and genetic responses to a changing climate

Phenotypic plasticity

Phenotypic plasticity refers to the rapid expression of different phenotypes from the same genotype and is often initiated by an environmental change. For example, winter diapause in arthropods is triggered by shorter days and temperature responses from an earlier life stage or previous generation (Zhao et al. 2021). Plasticity may affect phenotypic outcomes – driving changes in the number of instars, or types and quantities of eggs or pupae produced – and also intrinsic processes, such as metabolism, gene expression, and protein/carbohydrate composition (Rodrigues & Beldade, 2020). Though plasticity can be maladaptive (e.g., seasonal mismatches between species and their resources, as

observed in pollinating insects; Buckley et al. 2017), the capacity and speed at which a species is able to adaptively respond by altering its behaviour, morphology, or physiology is an important component of its ability to cope with, and acclimate to, the effects of climate change (Bahrndorff et al. 2021; Rodrigues & Beldade, 2020; Sgrò et al. 2016).

Recent studies have suggested that invasive species may be more capable of responding plastically, suggesting that plasticity could be important for invasion success (Little et al. 2020). For example, the globally invasive European green crab (*Carcinus maenas*) demonstrates short-term acclimatory plasticity and high eurythermal tolerance compared to other crustaceans (18-31 days at either 5 or 25°C), which has likely facilitated its ability to colonise a wide range of thermally unique environments (Tepolt & Somero, 2014). RCH (see above) has been identified in many non-polar arthropods, particularly notoriously invasive species of Hemiptera (e.g., *Sitobion avenae*; Powell & Bale, 2004), Diptera (e.g., *Culicoides variipenni*; Nunamaker, 1993), Lepidoptera (e.g., *P. xylostella*; Park & Kim, 2014), and Orthoptera (e.g., *Locusta migratoria*; Findsen et al. 2013), highlighting their risk of potentially establishing populations in polar environments. The ability to assume ideal phenotypes for environmental variance over short timescales will likely play an important role in determining species responses to rapid climate change (Simons, 2011), however, limited research is devoted to understanding phenotypic plasticity of arthropod taxa. This includes research that utilises next generation sequencing technologies, such as transcriptomics (the study of RNA transcripts that are expressed in the cell) to examine changes in gene expression under specific environmental conditions (González-Aravena et al. 2021). As a result, the ability of polar arthropods to respond plastically to increasing temperatures, or compete with invading species that may also be more plastic, is unclear.

Genetic responses

Increasing temperature can create *de novo* genetic diversity by causing spontaneous replication errors and/or DNA damage (Chu et al. 2018), and can affect standing genetic variation by shifting allele frequencies in genes or linked genomic regions associated with climate response (De La Torre et al. 2019; McGaughran et al. 2021). *De novo* variation tends to operate on a longer timescale, as new mutations have to occur and then sweep through a population (Barrick & Lenski, 2013). Conversely, standing genetic variation can result in more rapid adaptation due to the co-option of existing variation towards new solutions (Chaturvedi et al. 2021).

Although the majority of mutations provide no additional fitness benefit (i.e., are most often deleterious or neutral), some give rise to variation within populations that provides an advantage to survival, with research identifying a clear link between temperature and mutation rate (Waldvogel & Pfenninger, 2021). The tendency for species density to be greater in lower as opposed to higher latitudes – i.e., the latitudinal biodiversity gradient – may be partly underlain by differential rates of mutation among species. For example, temperate and tropical arthropods have shorter generation times and larger population sizes and are therefore predicted to have a greater supply and faster rate of mutations for evolution to act upon, thus driving adaptation (Chu et al. 2020). In contrast, polar arthropods are expected to experience slower rates of mutation owing to the fact that many species must overwinter in a state of stupor and are only active for the few months of austral summer (thus have longer generation times; Chu et al. 2018; 2020). Therefore, *de novo* variation in polar arthropods may be insufficient to meet the demands of increased temperatures (Berteaux et al. 2004), including those driven by competition with invasives that have higher mutation rates.

Adaptive responses that build from standing variation have not been identified for polar arthropods to date, however temperate species show that existing genetic variation

across the genome in response to human impacts can drive changes in diapause timing (Troczka et al. 2012), contribute to insecticide resistance (Foster et al. 2021), and/or lead to phenotypic changes such as melanism or wing-loss (Foster et al. 2022; McColloch et al. 2022).

Genetic responses to climate warming (whether building from *de novo* or standing variation) for polar taxa have yet to gain great research traction more generally. This is an area that would clearly benefit from more focus, particularly because, even if such responses are unable to occur rapidly for endemic taxa, they may well promote rapid adaptation of newly invading species (McCulloch et al. 2022).

A1.5 Future steps

Published physiological studies on terrestrial Arctic and Antarctic arthropods are limited to a small number of species – undoubtedly driven by the various challenges associated with working at the poles, including financial, time, and resource burdens. For those taxa that have been studied, physiology is the target of the majority of research (e.g., freeze-tolerance, metabolism), with a view towards understanding how these unique adaptations promote survival in the cold.

However, physiology is dynamic across populations in time and space, as are other factors that shape evolutionary potential, such as genetic diversity, gene expression, population size, population structure, and mutation rates (Bahrndorff et al. 2021). Recent population genomic technologies have shown great potential for elucidating demographic and adaptive processes and, when combined with ecological niche modelling, can be used to reconstruct evolutionary and ecological histories and predict future trajectories (Krehenwinkel et al. 2015), aiding evaluation of the risk of localised biological invasions. Moreover, transcriptomic studies can shed light on the capacity of species to respond plastically to climate change (Snoeck et al. 2018). However, little research in these fields

targets arthropods, and even less employs a polar lens. Thus, evolutionary potential, the fitness implications and influence of plasticity, and their interacting effects on the other potential responses of polar species to climate change are key areas in which future research would provide exceptionally valuable insights.

A1.6 Literature cited

Addo-Bediako, A., Chown, S. L., & Gaston, K. J. (2002). Metabolic cold adaptation in insects: A large-scale perspective. *Functional Ecology*, *16*, 332-338. https://doi.org/10.1046/j.1365-2435.2002.00634.x.

Aguilar-Alberola, J. A., & Mesquita-Joanes, F. (2014). Breaking the temperature-size rule: Thermal effects on growth, development and fecundity of a crustacean from temporary waters. *Journal of Thermal Biology*, *42*, 15-24. https://doi.org/10.1016/j.jtherbio.2014.02.016.

Alsos, I. G., Ware, C., & Elven, R. (2015). Past Arctic aliens have passed away, current ones may stay. *Biological Invasions*, *17*, 3113-3123. https://doi.org/10.1007/s10530-015-0937-9.

Ayres, M. P., & Scriber, J. M. (1994). Local adaptation to regional climates in *Papilio canadensis* (Lepidoptera: Papilionidae). *Ecological Monographs*, *64*, 465-482. https://doi.org/10.2307/2937146.

Bahrndorff, S., Lauritzen, J. M. S., Sørensen, M. H., Noer, N. K., & Kristensen, T. N. (2021). Responses of terrestrial polar arthropods to high and increasing temperatures. *Journal of Experimental Biology*, *224*, 1630 -1037. https://doi.org/10.1242/jeb.230797.

Baird, H. P., Moon, K. L., Janion-Scheepers, C., & Chown, S. L. (2020). Springtail phylogeography highlights biosecurity risks of repeated invasions and intraregional transfers among remote islands. *Evolutionary Applications*, *13*, 960-973. https://doi.org/10.1111/eva.12913.

Bandoro, J., Solomon, S., Donohoe, A., Thompson, D. W. J., & Santer, B. D. (2014). Influences of the Antarctic ozone hole on Southern hemispheric summer climate change. *Journal of Climate*, *27*, 6245-6264. https://doi.org/10.1175/JCLI-D-13-00698.1.

Barrick, J. E., & Lenski, R. E. (2013). Genome dynamics during experimental evolution. *Nature Reviews Genetics, 14*, 827-839. https://doi.org/10.1038/nrg3564.

Bartlett, J. C., Convey, P., Hughes, K. A., Thorpe, S. E., & Hayward, S. A. L. (2021). Ocean currents as a potential dispersal pathway for Antarctica's most persistent non-native terrestrial insect. *Polar Biology, 44*, 209-216. https://doi.org/10.1007/s00300-020-02792-2.

Berteaux, D., Réale, D., McAdam, A. G., & Boutin, S. (2004). Keeping pace with fast climate change: can Arctic life count on evolution? *Integrative and Comparative Biology, 44*, 140-151. https://doi.org/10.1093/icb/44.2.140.

Blackburn, T. M., Pyšek, P., Bacher, S., Carlton, J. T., Duncan, R. P., Jarošík, V. et al. (2011). A proposed unified framework for biological invasions. *Trends in Ecology and Evolution, 26*, 333-339. https://doi.org/10.1016/j.tree.2011.03.023.

Block, W., Webb, N. R., Coulson, S., Hodkinson, I. D., & Worland, M. R. (1994). Thermal adaptation in the Arctic collembolan *Onychiurus arcticus* (Tullberg). *Journal of Insect Physiology*, *40*, 715-722. https://doi.org/10.1016/0022-1910(94)90099-X.

Bowden, J. J., Eskildsen, A., Hansen, R. R., Olsen, K., Kurle, C. M., & Høye, T. T. (2015). High-Arctic butterflies become smaller with rising temperatures. *Biology Letters, 11*, 20150574. https://doi.org/10.1098/rsbl.2015.0574.

Buckley, L. B., Arakaki, A. J., Cannistra, A. F., Kharouba, H. M., & Kingsolver, J. G. (2017). Insect development, thermal plasticity and fitness implications in changing, seasonal environments. *Integrative and Comparative Biology, 57*, 988-998. https://doi.org/10.1093/icb/icx032.

Bush, A., Mokany, K., Catullo, R., Hoffmann, A., Kellermann, V., Sgrò, C. et al. (2016). Incorporating evolutionary adaptation in species distribution modelling reduces

projected vulnerability to climate change. *Ecology Letters*, *19*, 1468-1478. https://doi.org/https://doi.org/10.1111/ele.12696.

Chaturvedi, A., Zhou, J., Raeymaekers, J. A. M., Czypionka, T., Orsini, L., Jackson, C. E. et al. (2021). Extensive standing genetic variation from a small number of founders enables rapid adaptation in *Daphnia*. *Nature Communications*, *12*, 4306. https://doi.org/10.1038/s41467-021-24581-z.

Chown, S. L., & Convey, P. (2016). Antarctic entomology. *Annual Review of Entomology*, *61*, 119-137. https://doi.org/10.1146/annurev-ento-010715-023537.

Chu, X.-L., Zhang, B.-W., Zhang, Q.-G., Zhu, B.-R., Lin, K., & Zhang, D.-Y. (2018). Temperature responses of mutation rate and mutational spectrum in an *Escherichia coli* strain and the correlation with metabolic rate. *BMC Evolutionary Biology*, *18*, 126. https://doi.org/10.1186/s12862-018-1252-8.

Chu, X.-L., Zhang, D.-Y., Buckling, A., & Zhang, Q.-G. (2020). Warmer temperatures enhance beneficial mutation effects. *Journal of Evolutionary Biology*, *33*, 1020-1027. https://doi.org/https://doi.org/10.1111/jeb.13642.

Clark, M. S., Thorne, M. A. S., Purać, J., Burns, G., Hillyard, G., Popović, Ž. D., Grubor-Lajšić, G., & Worland, M. R. (2009). Surviving the cold: Molecular analyses of insect cryoprotective dehydration in the Arctic springtail *Megaphorura arctica* (Tullberg). *BMC Genomics*, *10*, 328. https://doi.org/10.1186/1471-2164-10-328.

Clarke, A., & Fraser, K. P. P. (2004). Why does metabolism scale with temperature? *Functional Ecology*, *18*, 243-251. https://doi.org/https://doi.org/10.1111/j.0269-8463.2004.00841.x.

Cohen, J., Zhang, X., Francis, J., Jung, T., Kwok, R., Overland, J. et al. (2020). Divergent consensuses on Arctic amplification influence on midlatitude severe winter weather. *Nature Climate Change*, *10*, 20-29. https://doi.org/10.1038/s41558-019-0662-y.

Contador, T., Gañan, M., Bizama, G., Fuentes-Jaque, G., Morales, L., Rendoll, J. et al. (2020). Assessing distribution shifts and ecophysiological characteristics of the only Antarctic winged midge under climate change scenarios. *Scientific Reports*, *10*, 9087. https://doi.org/10.1038/s41598-020-65571-3.

Convey, P., Gibson, J. A., Hillenbrand, C. D., Hodgson, D. A., Pugh, P. J., Smellie, J. L. et al. (2008). Antarctic terrestrial life--challenging the history of the frozen continent? *Biological Reviews of the Cambridge Philosophical Society*, *83*, 103-117. https://doi.org/10.1111/j.1469-185X.2008.00034.x.

Convey, P., & Peck, L. S. Antarctic environmental change and biological responses. *Science Advances*, *5*, eaaz0888. https://doi.org/10.1126/sciadv.aaz0888.

Coulson, S. J. (2015). The alien terrestrial invertebrate fauna of the High Arctic archipelago of Svalbard: Potential implications for the native flora and fauna. *Polar Research*, *34*, 27364. https://doi.org/10.3402/polar.v34.27364.

Danks, H. V. (2004). Seasonal adaptations in Arctic insects. *Integrative and Comparative Biology*, *44*, 85-94. https://doi.org/10.1093/icb/44.2.85.

David, J. R., Gibert, P., Moreteau, B., Gilchrist, G. W., & Huey, R. B. (2003). The fly that came in from the cold: geographic variation of recovery time from low-temperature exposure in *Drosophila subobscura*. *Functional Ecology*, *17*, 425-430. https://doi.org/10.1046/j.1365-2435.2003.00750.x.

De La Torre, A. R., Wilhite, B., & Neale, D. B. (2019). Environmental genome-wide association reveals climate adaptation is shaped by subtle to moderate allele frequency shifts in loblolly pine. *Genome Biology and Evolution*, *11*, 2976-2989. https://doi.org/10.1093/gbe/evz220.

de Sassi, C., & Tylianakis, J. M. (2012). Climate change disproportionately increases herbivore over plant or parasitoid biomass. *PLoS ONE*, *7*, e40557. https://doi.org/10.1371/journal.pone.0040557.

Diffenbaugh, N. S., & Field, C. B. (2013). Changes in ecologically critical terrestrial climate conditions. *Science, 341*, 486-492. https://doi.org/10.1126/science.1237123.

Doellman, M. M., Egan, S. P., Ragland, G. J., Meyers, P. J., Hood, G. R., Powell, T. H. Q. et al. (2019). Standing geographic variation in eclosion time and the genomics of host race formation in *Rhagoletis pomonella* fruit flies. *Ecology and Evolution, 9,* 393-409. https://doi.org/10.1002/ece3.4758.

Duffy, G. A., Coetzee, B. W. T., Latombe, G., Akerman, A. H., McGeoch, M. A., & Chown, S. L. (2017). Barriers to globally invasive species are weakening across the Antarctic. *Diversity and Distributions, 23*, 982-996. https://doi.org/10.1111/ddi.12593.

Duman, J. G., Bennett, V., Sformo, T., Hochstrasser, R., & Barnes, B. M. (2004). Antifreeze proteins in Alaskan insects and spiders. *Journal of Insect Physiology, 50,* 259-266. https://doi.org/10.1016/j.jinsphys.2003.12.003.

Elnitsky, M. A., Hayward, S. A., Rinehart, J. P., Denlinger, D. L., & Lee, R. E., Jr. (2008). Cryoprotective dehydration and the resistance to inoculative freezing in the Antarctic midge, *Belgica antarctica*. *Journal of Experimental Biology, 211*, 524-530. https://doi.org/10.1242/jeb.011874.

Everatt, M. J., Worland, M. R., Bale, J. S., Convey, P., & Hayward, S. A. L. (2012). Pre-adapted to the maritime Antarctic? – Rapid cold hardening of the midge, *Eretmoptera murphyi*. *Journal of Insect Physiology, 58,* 1104-1111. https://doi.org/10.1016/j.jinsphys.2012.05.009.

Findsen, A., Andersen, J. L., Calderon, S., & Overgaard, J. (2013). Rapid cold hardening improves recovery of ion homeostasis and chill coma recovery time in the migratory locust, *Locusta migratoria*. *Journal of Experimental Biology, 216*, 1630-1637. https://doi.org/10.1242/jeb.081141.

Foster, B. J., McCulloch, G. A., Foster, Y., Kroos, G. C., & Waters, J. M. (2022). Ebony underpins Batesian mimicry in an insect melanic polymorphism. *bioRxiv*, https://doi.org/10.1101/2022.06.13.495778.

Foster, B. J., McCulloch, G. A., Vogel, M. F. S., Ingram, T., & Waters, J. M. (2021). Anthropogenic evolution in an insect wing polymorphism following widespread deforestation. *Biology Letters*, *17*, 20210069. https://doi.org/10.1098/rsbl.2021.0069.

Frenot, Y., Chown, S. L., Whinam, J., Selkirk, P. M., Convey, P., Skotnicki, M. et al. (2005). Biological invasions in the Antarctic: extent, impacts and implications. *Biological Reviews*, *80*, 45-72. https://doi.org/https://doi.org/10.1017/S1464793104006542.

Giribet, G., & Edgecombe, G. D. (2019). The phylogeny and evolutionary history of arthropods. *Current Biology*, *29*, R592-r602. https://doi.org/10.1016/j.cub.2019.04.057.

González-Aravena, M., Rondon, R., Font, A., Cárdenas, C. A., Toullec, J.-Y., Corre, E. et al. (2021). Low transcriptomic plasticity of Antarctic giant isopod *Glyptonotus antarcticus* juveniles exposed to acute thermal stress. *Frontiers in Marine Science*, *8*. https://doi.org/10.3389/fmars.2021.761866.

Hawes, T. C., Couldridge, C. E., Bale, J. S., Worland, M. R., & Convey, P. (2006). Habitat temperature and the temporal scaling of cold hardening in the high Arctic collembolan, *Hypogastrura tullbergi* (Schäffer). *Ecological Entomology*, *31*, 450-459. https://doi.org/10.1111/j.1365-2311.2006.00796.x.

Hodkinson, I. D., Coulson, S. J., & Webb, N. R. (2003). Community assembly along proglacial chronosequences in the high Arctic: vegetation and soil development in north-west Svalbard. *Journal of Ecology*, *91*, 651-663. https://doi.org/10.1046/j.1365-2745.2003.00786.x

Hodkinson, I. D., Webb, N. R., Bale, J. S., Block, W., Coulson, S. J., & Strathdee, A. T. (1998). Global change and Arctic ecosystems: Conclusions and predictions from experiments with terrestrial invertebrates on Spitsbergen. *Arctic and Alpine Research*, *30*, 306-313. https://doi.org/10.1080/00040851.1998.12002904.

Horne, C. R., Hirst, A. G., & Atkinson, D. (2017). Seasonal body size reductions with warming covary with major body size gradients in arthropod species. *Proceedings of the Royal Society B: Biological Sciences*, *284*, 20170238. https://doi.org/10.1098/rspb.2017.0238.

Houghton, M., McQuillan, P. B., Bergstrom, D. M., Frost, L., van den Hoff, J., & Shaw, J. (2016). Pathways of alien invertebrate transfer to the Antarctic region. *Polar Biology*, *39*, 23-33. https://doi.org/10.1007/s00300-014-1599-2.

Høye, T. T. (2020). Arthropods and climate change - Arctic challenges and opportunities. *Current Opinion in Insect Science*, *41*, 40-45. https://doi.org/10.1016/j.cois.2020.06.002.

Høye, T. T., & Culler, L. E. (2018). Tundra arthropods provide key insights into ecological responses to environmental change. *Polar Biology*, *41*, 1523-1529. https://doi.org/10.1007/s00300-018-2370-x.

Høye, T. T., & Forchhammer, M. C. (2008). Phenology of high-Arctic arthropods: Effects of climate on spatial, seasonal, and inter-annual variation. In *Advances in ecological research* (Vol. 40, pp. 299-324). Academic Press. https://doi.org/10.1016/S0065-2504(07)00013-X.

Hughes, K. A., Convey, P., & Turner, J. (2021). Developing resilience to climate change impacts in Antarctica: An evaluation of Antarctic Treaty System protected area policy. *Environmental Science & Policy*, *124*, 12-22. https://doi.org/10.1016/j.envsci.2021.05.023.

Hughes, K. A., Pescott, O. L., Peyton, J., Adriaens, T., Cottier-Cook, E. J., Key, G. et al. (2020). Invasive non-native species likely to threaten biodiversity and ecosystems in the Antarctic Peninsula region. *Global Change Biology*, *26*, 2702-2716. https://doi.org/10.1111/gcb.14938.

Hughes, K. A., & Worland, M. R. (2010). Spatial distribution, habitat preference and colonization status of two alien terrestrial invertebrate species in Antarctica. *Antarctic Science*, *22*, 221-231. https://doi.org/10.1017/S0954102009990770.

Hughes, K. A., Worland, M. R., Thorne, M. A. S., & Convey, P. (2013). The non-native chironomid *Eretmoptera murphyi* in Antarctica: Erosion of the barriers to invasion. *Biological Invasions*, *15*, 269-281. https://doi.org/10.1007/s10530-012-0282-1.

Jahn, A. (2018). Reduced probability of ice-free summers for 1.5°C compared to 2°C warming. *Nature Climate Change*, *8*, 409-413. https://doi.org/10.1038/s41558-018-0127-8.

Kellermann, V., & van Heerwaarden, B. (2019). Terrestrial insects and climate change: Adaptive responses in key traits. *Physiological Entomology*, *44*, 99-115. https://doi.org/10.1111/phen.12282.

Klok, C. J., & Harrison, J. F. (2013). The temperature size rule in arthropods: Independent of macro-environmental variables but size dependent. *Integrative and Comparative Biology*, *53*, 557-570. https://doi.org/10.1093/icb/ict075.

Koenigk, T., Key, J., & Vihma, T. (2020). Climate change in the Arctic. *Physics and chemistry of the Arctic atmosphere*, 673-705. https://doi.org/10.1007/978-3-030-33566-3_11.

Krehenwinkel, H., Rödder, D., & Tautz, D. (2015). Eco-genomic analysis of the poleward range expansion of the wasp spider *Argiope bruennichi* shows rapid adaptation

and genomic admixture. *Global Change Biology, 21,* 4320-4332. https://doi.org/10.1111/gcb.13042.

Lardies, M., Bacigalupe, L., & Bozinovic, F. (2004). Testing the metabolic cold adaptation hypothesis: An intraspecific latitudinal comparison in the common woodlouse. *Evolutionary Ecology Research*, *6*.

Lassuy, D. R., & Lewis, P. N. (2013). Invasive species: Human-induced. *Arctic biodiversity assessment. Status and trends in Arctic biodiversity*, 559-565.

Lee, R. E., Jr., Elnitsky, M. A., Rinehart, J. P., Hayward, S. A., Sandro, L. H., & Denlinger, D. L. (2006). Rapid cold-hardening increases the freezing tolerance of the Antarctic midge *Belgica antarctica*. *Journal of Experimental Biology*, *209*, 399-406. https://doi.org/10.1242/jeb.02001.

Little, C. M., Chapman, T. W., & Hillier, N. K. (2020). Plasticity is key to success of *Drosophila suzukii* (Diptera: Drosophilidae) invasion. *Journal of Insect Science, 20,* 5. https://doi.org/10.1093/jisesa/ieaa034.

MacMillan, H. A., & Sinclair, B. J. (2011). Mechanisms underlying insect chill-coma. *Journal of Insect Physiology, 57,* 12-20. https://doi.org/10.1016/j.jinsphys.2010.10.004.

Maksym, T. (2019). Arctic and Antarctic sea ice change: contrasts, commonalities, and causes. *Annual Review of Marine Science, 11,* 187-213. https://doi.org/10.1146/annurev-marine-010816-060610.

Marden, J. H. (2013). Nature's inordinate fondness for metabolic enzymes: Why metabolic enzyme loci are so frequently targets of selection. *Molecular Ecology*, *22*, 5743-5764. https://doi.org/10.1111/mec.12534.

Maurice, H., & Philippe, V. (2021). Terrestrial macro-arthropods of the sub-Antarctic islands of Possession (Crozet Archipelago) and Kerguelen: Inventory of native

and non-native species. *Zoosystema*, *43*, 549-561. https://doi.org/10.5252/zoosystema2021v43a22.

McCulloch, G. A., & Waters, J. M. (2022). Rapid adaptation in a fast-changing world: Emerging insights from insect genomics. *Global Change Biology*, *00*, https://doi.org/10.1111/gcb.16512.

McGaughran, A., Laver, R., & Fraser, C. (2021). Evolutionary responses to warming. *Trends in Ecology and Evolution*, *36*, 591-600. https://doi.org/10.1016/j.tree.2021.02.014.

Meier, W. N., Hovelsrud, G. K., van Oort, B. E. H., Key, J. R., Kovacs, K. M., Michel, C. et al. (2014). Arctic sea ice in transformation: A review of recent observed changes and impacts on biology and human activity. *Reviews of Geophysics*, *52*, 185-217. https://doi.org/10.1002/2013RG000431.

Monteiro, M. R., Marshall, A. J., Hawes, I., Lee, C. K., McDonald, I. R., & Cary, S. C. (2022). Geochemically defined space-for-time transects successfully capture microbial dynamics along lacustrine chronosequences in a polar desert. *Frontiers in Microbiology*, *12*. https://doi.org/10.3389/fmicb.2021.783767.

Nespolo, R. F., Lardies, M. A., & Bozinovic, F. (2003). Intrapopulational variation in the standard metabolic rate of insects: Repeatability, thermal dependence and sensitivity (Q10) of oxygen consumption in a cricket. *Journal of Experimental Biology*, *206*, 4309-4315. https://doi.org/10.1242/jeb.00687.

Noer, N. K., Sørensen, M. H., Colinet, H., Renault, D., Bahrndorff, S., & Kristensen, T. N. (2022). Rapid adjustments in thermal tolerance and the metabolome to daily environmental changes – A field study on the Arctic seed bug *Nysius groenlandicus*. *Frontiers in Physiology*, *13*. https://doi.org/10.3389/fphys.2022.818485.

Nunamaker, R. A. (1993). Rapid cold-hardening in *Culicoides variipennis sonorensis* (Diptera: Ceratopogonidae). *Journal of Medical Entomology*, *30*, 913-917. https://doi.org/10.1093/jmedent/30.5.913.

Overland, J. E., & Wang, M. (2013). When will the summer Arctic be nearly sea ice free? *Geophysical Research Letters*, *40*, 2097-2101. https://doi.org/10.1002/grl.50316.

Park, Y., & Kim, Y. (2014). A specific glycerol kinase induces rapid cold hardening of the diamondback moth, *Plutella xylostella*. *Journal of Insect Physiology*, *67*, 56-63. https://doi.org/10.1016/j.jinsphys.2014.06.010.

Pertierra, L. R., Escribano-Álvarez, P., & Olalla-Tárraga, M. Á. (2021). Cold tolerance is similar but heat tolerance is higher in the alien insect *Trichocera maculipennis* than in the native *Parochlus steinenii* in Antarctica. *Polar Biology*, *44*, 1203-1208. https://doi.org/10.1007/s00300-021-02865-w.

Pörtner, H.-O., Roberts, D. C., Adams, H., Adler, C., Aldunce, P., Ali, E. et al. (2022). Climate change 2022: Impacts, adaptation and vulnerability. *IPCC Sixth Assessment Report*.

Powell, S. J., & Bale, J. S. (2004). Cold shock injury and ecological costs of rapid cold hardening in the grain aphid *Sitobion avenae* (Hemiptera: Aphididae). *Journal of Insect Physiology*, *50*, 277-284. https://doi.org/10.1016/j.jinsphys.2004.01.003.

Rinehart, J. P., Li, A., Yocum, G. D., Robich, R. M., Hayward, S. A. L., & Denlinger, D. L. (2007). Up-regulation of heat shock proteins is essential for cold survival during insect diapause. *Proceedings of the National Academy of Sciences*, *104*, 11130-11137. https://doi.org/doi:10.1073/pnas.0703538104.

Robert Michaud, M., Benoit, J. B., Lopez-Martinez, G., Elnitsky, M. A., Lee, R. E., & Denlinger, D. L. (2008). Metabolomics reveals unique and shared metabolic changes in response to heat shock, freezing and desiccation in the Antarctic

midge, *Belgica antarctica*. *Journal of Insect Physiology*, *54*, 645-655. https://doi.org/10.1016/j.jinsphys.2008.01.003.

Rodrigues, Y. K., & Beldade, P. (2020). Thermal plasticity in insects' response to climate change and to multifactorial environments. *Frontiers in Ecology and Evolution*, *8*. https://doi.org/10.3389/fevo.2020.00271.

Rozsypal, J., & Košťál, V. (2018). Supercooling and freezing as eco-physiological alternatives rather than mutually exclusive strategies: A case study in *Pyrrhocoris apterus*. *Journal of Insect Physiology*, *111*, 53-62. https://doi.org/10.1016/j.jinsphys.2018.10.006.

Rudkin, D., & Young, G. (2009). Horseshoe crabs–an ancient ancestry revealed. In *Biology and conservation of horseshoe crabs* (pp. 25-44). Springer.

Schmidt, N. M., Mosbacher, J. B., Nielsen, P. S., Rasmussen, C., Høye, T. T., & Roslin, T. (2016). An ecological function in crisis? The temporal overlap between plant flowering and pollinator function shrinks as the Arctic warms. *Ecography*, *39*, 1250-1252. https://doi.org/10.1111/ecog.02261.

Schulte, P. M. (2015). The effects of temperature on aerobic metabolism: Towards a mechanistic understanding of the responses of ectotherms to a changing environment. *Journal of Experimental Biology*, *218*, 1856-1866. https://doi.org/10.1242/jeb.118851.

Scriven, J. J., Whitehorn, P. R., Goulson, D., & Tinsley, M. C. (2016). Bergmann's body size rule operates in facultatively endothermic insects: evidence from a complex of cryptic bumblebee species. *PLoS One*, *11*, e0163307. https://doi.org/10.1371/journal.pone.0163307.

Sformo, T., Walters, K., Jeannet, K., Wowk, B., Fahy, G. M., Barnes, B. M. et al. (2010). Deep supercooling, vitrification and limited survival to -100°C in the Alaskan

beetle *Cucujus clavipes puniceus* (Coleoptera: Cucujidae) larvae. *Journal of Experimental Biology*, *213*, 502-509. https://doi.org/10.1242/jeb.035758.

Sgrò, C. M., Terblanche, J. S., & Hoffmann, A. A. (2016). What can plasticity contribute to insect responses to climate change? *Annual Review of Entomology*, *61*, 433-451. https://doi.org/10.1146/annurev-ento-010715-023859.

Shadwick, E. H., De Meo, O. A., Schroeter, S., Arroyo, M. C., Martinson, D. G., & Ducklow, H. (2021). Sea ice suppression of CO2 outgassing in the West Antarctic Peninsula: Implications for the evolving Southern Ocean carbon sink. *Geophysical Research Letters*, *48*, e2020GL091835. https://doi.org/10.1029/2020GL091835.

Shelomi, M. (2012). Where are we now? Bergmann's rule *sensu lato* in insects. *The American Naturalist*, *180*, 511-519. https://doi.org/10.1086/667595.

Simons, A. M. (2011). Modes of response to environmental change and the elusive empirical evidence for bet hedging. *Proceedings of the Royal Society B: Biological Sciences*, *278*, 1601-1609. https://doi.org/10.1098/rspb.2011.0176.

Sinclair, B. J., Coello Alvarado, L. E., & Ferguson, L. V. (2015). An invitation to measure insect cold tolerance: Methods, approaches, and workflow. *Journal of Thermal Biology*, *53*, 180-197. https://doi.org/10.1016/j.jtherbio.2015.11.003.

Sinclair, B. J., & Sjursen, H. (2001). Cold tolerance of the Antarctic springtail *Gomphiocephalus hodgsoni* (Collembola, Hypogastruridae). *Antarctic Science*, *13*, 271-279. https://doi.org/10.1017/S0954102001000384.

Sinclair, B. J., Terblanche, J. S., Scott, M. B., Blatch, G. L., Jaco Klok, C., & Chown, S. L. (2006). Environmental physiology of three species of Collembola at Cape Hallett, North Victoria Land, Antarctica. *Journal of Insect Physiology*, *52*, 29-50. https://doi.org/10.1016/j.jinsphys.2005.09.002.

Snoeck, S., Wybouw, N., Van Leeuwen, T., & Dermauw, W. (2018). Transcriptomic plasticity in the arthropod generalist *Tetranychus urticae* upon long-term acclimation to different host plants. *G3 Genes|Genomes|Genetics*, *8*, 3865-3879. https://doi.org/10.1534/g3.118.200585.

Solomon, S., Ivy, D. J., Kinnison, D., Mills, M. J., Neely, R. R., & Schmidt, A. (2016). Emergence of healing in the Antarctic ozone layer. *Science*, *353*, 269-274. https://doi.org/10.1126/science.aae0061.

Sørensen, J. G., & Holmstrup, M. (2011). Cryoprotective dehydration is widespread in Arctic springtails. *Journal of Insect Physiology*, *57*, 1147-1153. https://doi.org/10.1016/j.jinsphys.2011.03.001.

Stachowicz, J. J., Terwin, J. R., Whitlatch, R. B., & Osman, R. W. (2002). Linking climate change and biological invasions: Ocean warming facilitates nonindigenous species invasions. *Proceedings of the National Academy of Sciences*, *99*, 15497-15500. https://doi.org/10.1073/pnas.242437499.

Storey, K. B., & Storey, J. M. (2012). Insect cold hardiness: Metabolic, gene, and protein adaptation. *Canadian Journal of Zoology*, *90*, 456-475. https://doi.org/10.1139/z2012-011.

Teets, N. M., & Denlinger, D. L. (2013). Physiological mechanisms of seasonal and rapid cold-hardening in insects. *Physiological Entomology*, *38*, 105-116. https://doi.org/10.1111/phen.12019.

Teets, N. M., Gantz, J. D., & Kawarasaki, Y. (2020). Rapid cold hardening: Ecological relevance, physiological mechanisms and new perspectives. *Journal of Experimental Biology*, *223*. https://doi.org/10.1242/jeb.203448.

Tepolt, C. K., & Somero, G. N. (2014). Master of all trades: Thermal acclimation and adaptation of cardiac function in a broadly distributed marine invasive species,

the European green crab, *Carcinus maenas*. *Journal of Experimental Biology*, *217*, 1129-1138. https://doi.org/10.1242/jeb.093849.

Terblanche, J. S., Clusella-Trullas, S., Deere, Jacques A., Van Vuuren, Bettine J. et al. (2009). Directional evolution of the slope of the metabolic rate–temperature relationship is correlated with climate. *Physiological and Biochemical Zoology*, *82*, 495-503. https://doi.org/10.1086/605361.

Thuiller, W., Albert, C., Araújo, M. B., Berry, P. M., Cabeza, M., Guisan, A. et al. (2008). Predicting global change impacts on plant species' distributions: Future challenges. *Perspectives in Plant Ecology, Evolution and Systematics*, *9*, 137-152. https://doi.org/10.1016/j.ppees.2007.09.004.

Troczka, B., Zimmer, C. T., Elias, J., Schorn, C., Bass, C., Davies, T. G. E. et al. (2012). Resistance to diamide insecticides in diamondback moth, *Plutella xylostella* (Lepidoptera: Plutellidae) is associated with a mutation in the membrane-spanning domain of the ryanodine receptor. *Insect Biochemistry and Molecular Biology*, *42*, 873-880. https://doi.org/10.1016/j.ibmb.2012.09.001.

Turner, J., Comiso, J. C., Marshall, G. J., Lachlan-Cope, T. A., Bracegirdle, T., Maksym, T. et al. (2009). Non-annular atmospheric circulation change induced by stratospheric ozone depletion and its role in the recent increase of Antarctic sea ice extent. *Geophysical Research Letters, 36.* https://doi.org/10.1029/2009GL037524.

Turner, J., Lu, H., King, J., Marshall, G. J., Phillips, T., Bannister, D. et al. (2021). Extreme temperatures in the Antarctic. *Journal of Climate*, *34*, 2653-2668. https://doi.org/10.1175/JCLI-D-20-0538.1.

Vanstreels, R. E. T., Palma, R. L., & Mironov, S. V. (2020). Arthropod parasites of Antarctic and Subantarctic birds and pinnipeds: A review of host-parasite

associations. *International Journal for Parasitology: Parasites and Wildlife*, *12*, 275-290. https://doi.org/10.1016/j.ijppaw.2020.03.007.

Verberk, W. C. E. P., Atkinson, D., Hoefnagel, K. N., Hirst, A. G., Horne, C. R., & Siepel, H. (2021). Shrinking body sizes in response to warming: Explanations for the temperature–size rule with special emphasis on the role of oxygen. *Biological Reviews*, *96*, 247-268. https://doi.org/10.1111/brv.12653.

Waldvogel, A. M., & Pfenninger, M. (2021). Temperature dependence of spontaneous mutation rates. *Genome Research*, *31*, 1582-1589. https://doi.org/10.1101/gr.275168.120.

Wallingford, P. D., Morelli, T. L., Allen, J. M., Beaury, E. M., Blumenthal, D. M., Bradley, B. A. et al. (2020). Adjusting the lens of invasion biology to focus on the impacts of climate-driven range shifts. *Nature Climate Change*, *10*, 398-405. https://doi.org/10.1038/s41558-020-0768-2.

Walters, K. R., Jr., Sformo, T., Barnes, B. M., & Duman, J. G. (2009). Freeze tolerance in an Arctic Alaska stonefly. *Journal of Experimental Biology*, *212*, 305-312. https://doi.org/10.1242/jeb.020701.

Williams, C. M., Szejner-Sigal, A., Morgan, T. J., Edison, A. S., Allison, D. B., & Hahn, D. A. (2016). Adaptation to low temperature exposure increases metabolic rates independently of growth rates. *Integrative and Comparative Biology*, *56*, 62-72. https://doi.org/10.1093/icb/icw009.

Williamson, M., & Fitter, A. (1996). The varying success of invaders. *Ecology*, *77*, 1661-1666. https://doi.org/10.2307/2265769.

Worland, M. R. (2005). Factors that influence freezing in the sub-Antarctic springtail *Tullbergia antarctica*. *Journal of Insect Physiology*, *51*, 881-894. https://doi.org/10.1016/j.jinsphys.2005.04.004.

Worland, M. R. (2010). *Eretmoptera murphyi*: Pre-adapted to survive a colder climate. *Physiological Entomology*, *35*, 140-147. https://doi.org/10.1111/j.1365-3032.2010.00722.x.

Worland, M. R., & Convey, P. (2001). Rapid cold hardening in Antarctic microarthropods. *Functional Ecology*, *15*, 515-524. https://doi.org/10.1046/j.0269-8463.2001.00547.x.

Wunderling, N., Willeit, M., Donges, J. F., & Winkelmann, R. (2020). Global warming due to loss of large ice masses and Arctic summer sea ice. *Nature Communications*, *11*, 5177. https://doi.org/10.1038/s41467-020-18934-3.

Zettel, J. (1984). The significance of temperature and barometric pressure changes for the snow surface activity of *Isotoma hiemalis* (Collembola). *Experientia*, *40*, 1369-1372. https://doi.org/10.1007/BF01951894.

Zhao, C., Guo, Y., Liu, Z., Xia, Y., Li, Y., Song, Z. et al. (2021). Temperature and photoperiodic response of diapause induction in *Anastatus japonicus*, an egg parasitoid of stink bugs. *Insects*, *12*. https://doi.org/10.3390/insects12

.

Appendix 2

Supplementary Information: Chapter 2

Table A2.1 Open access information based on each author's institution location for publications included in the analysis relating to A) population genetics, and B) population genomics.

A) Population genetics			
Not Open Access (n=473)		**Open Access (n=334)**	
Country	**Count**	**Country**	**Count**
USA	131	Usa	111
France	61	France	57
Australia	44	England	28
Germany	44	Germany	27
Spain	40	Spain	27
Canada	39	Australia	26
England	28	Italy	22
Italy	26	Japan	22
Japan	26	Portugal	20
Peoples R China	23	Peoples R China	17
Russia	20	Brazil	15
New Zealand	17	Canada	14
Norway	15	Norway	13
Portugal	13	South Africa	13
Sweden	13	Sweden	12
South Korea	12	Poland	11
Greece	11	Scotland	11
Switzerland	11	Switzerland	11
Poland	9	Russia	10
Slovenia	9	Denmark	9
Austria	8	South Korea	8
Denmark	8	New Zealand	7
Serbia	8	Austria	6
Turkey	8	Finland	6
Czech Republic	7	Greece	6
Finland	7	Indonesia	6
South Africa	7	Cameroon	5
Belgium	6	Croatia	5
India	6	Ireland	5

A) Population genomics			
Not Open Access (n=39)		**Open Access (n=205)**	
Country	**Count**	**Country**	**Count**
USA	10	Usa	39
France	4	Italy	12
Australia	3	Peoples R China	12
Germany	3	Australia	9
Greece	2	France	9
Netherlands	2	Spain	9
Portugal	2	England	8
Czech Republic	1	Netherlands	7
Denmark	1	Switzerland	7
Finland	1	Canada	6
Hungary	1	Germany	6
Israel	1	Sweden	6
Slovenia	1	Brazil	5
Sweden	1	Denmark	5
Ukraine	1	Croatia	4
		Mexico	4
		Colombia	3
		Finland	3
		Iran	3
		Romania	3
		Russia	3
		Scotland	3
		Argentina	2
		Czech Republic	2
		Ecuador	2
		Estonia	2
		Hungary	2
		Ireland	2
		Israel	2

Iran	6	Malaysia	5
Ireland	6	Netherlands	5
Argentina	5	Serbia	5
Brazil	5	Taiwan	5
Bulgaria	5	Belgium	4
Croatia	5	Chile	4
Hungary	5	Slovenia	4
Scotland	5	Argentina	3
Taiwan	5	Bulgaria	3
Thailand	5	Colombia	3
Madagascar	4	Czech Republic	3
Mexico	4	Hungary	3
Morocco	4	Mexico	3
Tunisia	4	Morocco	3
Chile	3	North Ireland	3
Estonia	3	Tanzania	3
Indonesia	3	Uganda	3
Israel	3	Wales	3
Malaysia	3	Belarus	2
Romania	3	Burkina Faso	2
Slovakia	3	Cent Afr Republ	2
Bangladesh	2	Ecuador	2
Myanmar	2	Gabon	2
Singapore	2	Iran	2
Sri Lanka	2	Israel	2
Uganda	2	Kenya	2
Armenia	1	Luxembourg	2
Azerbaijan	1	Montenegro	2
Belarus	1	Myanmar	2
Benin	1	Niger	2
Brunei	1	Romania	2
Burundi	1	Turkey	2
Cuba	1	Vietnam	2
Gabon	1	Albania	1
Georgia	1	Cyprus	1
Ghana	1	Egypt	1
Grenada	1	Iceland	1
Iraq	1	India	1
Kenya	1	Kazakhstan	1
Lithuania	1	Latvia	1
Luxembourg	1	Lithuania	1
Macedonia	1	Macedonia	1

Norway	2
Slovakia	2
Slovenia	2
Armenia	1
Belgium	1
Bolivia	1
Costa Rica	1
Cuba	1
Greece	1
Guatemala	1
Kazakhstan	1
Luxembourg	1
Papua N Guinea	1
Paraguay	1
Peru	1
Poland	1
Portugal	1
Serbia	1
Tanzania	1
Ukraine	1
Uruguay	1
Vietnam	1

Monaco	1	Madagascar	1
Mozambique	1	Mongolia	1
Netherlands	1	North Macedonia	1
New Caledonia	1	Oman	1
North Ireland	1	Papua N Guinea	1
North Macedonia	1	Saudi Arabia	1
Rwanda	1	Senegal	1
St Kitts Nevi	1	Serbia	1
Tanzania	1	Singapore	1
Trinidad Tobago	1	Slovakia	1
Ukraine	1	Solomon Islands	1
Uruguay	1	South Sudan	1
Uzbekistan	1	Thailand	1
Venezuela	1	Togo	1
Vietnam	1	Tunisia	1
Wales	1	Ukraine	1
Zimbabwe	1	Zambia	1

Table A2.2 Year of publication for each publication included in the analysis that related to A) population genetics, and B) population genomics.

A) Population genetics	
Year	**Count**
2014	51
2016	42
2010	41
2012	41
2009	40
2011	40
2015	39
2007	38
2013	37
2020	35
2018	34
2019	34
2008	32
2006	31
2017	31
2021	29
2003	24
2004	24
2005	22
1999	18
2000	18
2001	18
2002	16
1996	13
1998	11
2022	11
1995	10
1997	7
1991	5
1993	5
1994	5
1992	4
1989	1

B) Population genomics	
Year	**Count**
2019	18
2021	16
2018	13
2020	13
2013	9
2016	6
2017	5
2012	4
2015	4
2022	2
2011	1

Table A2.3 A table of each species on the "IUCN 100 of the World's Worst Alien Invasive Species" list, whether they had population *genetic* data, whether it was in an invasive context, the total number of documents retrieved, and of those, how many were relevant to invasion biology.

Species name	Common name	Taxonomic group	Does population genetic data exist?	Has it been used in an invasive species context?	Total no. documents	Total no. relevant documents
Lithobates catesbeianus	Bullfrog	Amphibian	Yes	Yes	6	3
Rhinella marina	Cane toad	Amphibian	Yes	Yes	13	4
Eleutherodactylus coqui	Caribbean tree frog	Amphibian	Yes	Yes	4	3
Acridotheres tristis	Common myna	Bird	Yes	No	4	1
Pycnonotus cafer	Red-vented bulbul	Bird	Yes	No	0	0
Sturnus vulgaris	European starling	Bird	Yes	Yes	20	8
Clarias batrachus	Walking catfish	Fish	Yes	No	27	4
Cyprinus carpio	Common carp	Fish	Yes	Yes	229	16
Gambusia affinis	Gambusia	Fish	Yes	Yes	65	19
Lates niloticus	Nile perch	Fish	Yes	Yes	12	4
Micropterus salmoides	Largemouth bass	Fish	Yes	Yes	56	6
Oncorhynchus mykiss	Rainbow trout	Fish	Yes	Yes	535	66
Oreochromis mossambicus	Mozambique tilapia	Fish	Yes	Yes	23	6
Salmo trutta	Brown trout	Fish	Yes	Yes	419	104
Achatina fulica	Giant African snail	Invertebrate	Yes	No	7	0
Aedes albopictus	Tiger mosquito	Invertebrate	Yes	Yes	112	29
Anopheles quadrimaculatus	Common malaria mosquito	Invertebrate	Yes	Yes	7	1
Anoplolepis gracilipes	Yellow crazy ant	Invertebrate	Yes	Yes	13	7
Anoplophora glabripennis	Asian long-horned beetle	Invertebrate	Yes	Yes	11	6
Asterias amurensis	Flatbottom seastar	Invertebrate	Yes	Yes	9	2

Bemisia tabaci	Cotton whitefly	Invertebrate	Yes	Yes	124	43
Carcinus maenas	European shore crab	Invertebrate	Yes	Yes	32	9
Cercopagis pengoi	Fishhook waterflea	Invertebrate	Yes	Yes	4	2
Cinara cupressi	Cypress aphid	Invertebrate	No	No	0	0
Coptotermes formosanus	Formosan subterranean termite	Invertebrate	Yes	Yes	29	2
Dreissena polymorpha	Zebra mussel	Invertebrate	Yes	Yes	35	7
Eriocheir sinensis	Chinese freshwater edible crab	Invertebrate	Yes	Yes	33	3
Euglandina rosea	Cannibal snail	Invertebrate	No	No	1	0
Linepithema humile	Argentine ant	Invertebrate	Yes	Yes	37	4
Lymantria dispar	Spongy moth	Invertebrate	Yes	Yes	45	8
Mnemiopsis leidyi	Warty comb jelly	Invertebrate	Yes	Yes	12	6
Mytilus galloprovincialis	Blue mussel	Invertebrate	Yes	Yes	122	38
Pheidole megacephala	Big-headed ant	Invertebrate	Yes	Yes	3	1
Platydemus manokwari	Snail-eating flatworm	Invertebrate	No	No	0	0
Pomacea canaliculata	Apple snail	Invertebrate	Yes	Yes	20	4
Potamocorbula amurensis	Asian clam	Invertebrate	Yes	Yes	9	3
Solenopsis invicta	Red imported fire ant	Invertebrate	Yes	Yes	46	12
Trogoderma granarium	Khapra beetle	Invertebrate	No	No	5	0
Vespula vulgaris	Common wasp	Invertebrate	Yes	Yes	8	1
Wasmannia auropunctata	Little fire ant	Invertebrate	Yes	Yes	17	8
Aphanomyces astaci	Crayfish plague	Macro-fungi	Yes	Yes	29	5
Batrachochytrium dendrobatidis	Chytrid frog fungi	Macro-fungi	Yes	Yes	28	6
Cryphonectria parasitica	Chestnut blight	Macro-fungi	Yes	Yes	45	14
Ophiostoma ulmi sensu lato	Dutch elm disease	Macro-fungi	No	No	15	0
Phytophthora cinnamomi	Cinnamon fungus	Macro-fungi	Yes	Yes	30	5
Capra hircus	Domestic goat	Mammal	Yes	No	131	6

Cervus elaphus	Red deer	Mammal	Yes	Yes	259	74
Felis catus	Domestic cat	Mammal	Yes	Yes	196	9
Herpestes javanicus	Small indian mongoose	Mammal	Yes	Yes	10	4
Macaca fascicularis	Crab-eating macaque	Mammal	Yes	Yes	80	17
Mus musculus	Field mouse	Mammal	Yes	Yes	329	52
Mustela erminea	Stoat	Mammal	Yes	Yes	15	3
Myocastor coypus	Copyu	Mammal	Yes	Yes	10	5
Oryctolagus cuniculus	European rabbit	Mammal	Yes	Yes	85	12
Rattus rattus	Black rat	Mammal	Yes	Yes	75	20
Sciurus carolinensis	Gray squirrel	Mammal	Yes	Yes	14	4
Sus scrofa	Wild boar	Mammal	Yes	Yes	475	27
Trichosurus vulpecula	Brushtail possum	Mammal	Yes	Yes	22	6
Vulpes vulpes	European red fox	Mammal	Yes	Yes	113	31
Banana bunchy top virus	Banana bunchy top virus	Micro-organism	Yes	Yes	1	1
Plasmodium relictum	Avian malaria	Micro-organism	Yes	Yes	40	3
Acacia Mearnsii	Black wattle	Plant	Yes	Yes	8	2
Ardisia elliptica	Shoebutton ardisia	Plant	No	No	0	0
Arundo donax	Giant reed	Plant	Yes	Yes	4	4
Caulerpa taxifolia	Killer alga	Plant	Yes	No	5	2
Cecropia peltata	Trumpet tree	Plant	No	No	0	0
Chromolaena odorata	Bitterbush	Plant	Yes	Yes	4	1
Cinchona pubescens	Red cinchona	Plant	No	No	0	0
Clidemia hirta	Clidemia hirta	Plant	Yes	Yes	1	1
Eichhornia crassipes	Floating water hyacinth	Plant	Yes	Yes	9	1
Euphorbia esula	Leafy spurge	Plant	No	No	4	0
Hedychium gardnerianum	Kahili ginger	Plant	No	No	0	0

Hiptage benghalensis	Hiptage benghalensis	Plant	No	No	0	0
Imperata cylindrica	Cogon grass	Plant	Yes	Yes	6	2
Lantana camara	West Indian latana	Plant	Yes	Yes	7	3
Leucaena leucocephala	River tamarind	Plant	No	No	3	0
Ligustrum robustum	Common privet	Plant	No	No	1	0
Lythrum salicaria	Purple loosetrife	Plant	Yes	Yes	1	1
Melaleuca quinquenervia	Broad leaved paperbark tree	Plant	Yes	Yes	2	1
Miconia calvescens	Velvet tree	Plant	Yes	Yes	3	2
Mikania micrantha	american rope	Plant	Yes	Yes	13	7
Mimosa pigra	Bashful plant	Plant	Yes	Yes	1	1
Morella faya	Fire tree	Plant	Yes	No	1	1
Opuntia stricta	Araluen pear	Plant	No	No	0	0
Pinus pinaster	Maritime pine	Plant	Yes	Yes	88	15
Polygonum cuspidatum	Donkey rhubarb	Plant	Yes	No	5	1
Prosopis glandulosa	Honey mesquite	Plant	Yes	Yes	1	1
Psidium cattleianum	Strawberry guava	Plant	Yes	Yes	1	1
Pueraria montana	Kudzu	Plant	Yes	Yes	8	1
Rubus ellipticus	Asian wild raspberry	Plant	No	No	3	0
Salvinia molesta	African payal	Plant	No	No	1	0
Schinus terebinthifolius	Baie rose	Plant	Yes	Yes	5	4
Spartina anglica	Common cord grass	Plant	Yes	Yes	5	1
Spathodea campanulata	African tulip tree	Plant	No	No	1	0
Sphagneticola trilobata	Bay biscayne creeping oxeye	Plant	No	No	2	0
Tamarix ramosissima	Salt cedar	Plant	Yes	Yes	6	3
Ulex europaeus	Gorse	Plant	Yes	Yes	5	1
Undaria pinnatifida	Wakame	Plant	Yes	Yes	26	4

Boiga irregularis	Brown treesnake	Reptile	Yes	Yes	6	1
Trachemys scripta elegans	Red-eared slider	Reptile	Yes	Yes	7	1
				Total	4399	807

Table A2.4 A table of each species on the "IUCN 100 of the World's Worst Alien Invasive Species" list, whether they had population *genomic* data, whether it was in an invasive context, the total number of documents retrieved, and of those, how many were relevant to invasion biology.

Species name	Common name	Taxonomic group	Does population genomic data exist?	Has it been used in an invasive species context?	Total no. documents	Total no. relevant docs.
Lithobates catesbeianus	Bullfrog	Amphibian	No	No	5	0
Rhinella marina	Cane toad	Amphibian	Yes	Yes	2	1
Eleutherodactylus coqui	Caribbean tree frog	Amphibian	No	No	0	0
Acridotheres tristis	Common myna	Bird	Yes	Yes	2	1
Pycnonotus cafer	Red-vented bulbul	Bird	No	No	0	0
Sturnus vulgaris	European starling	Bird	Yes	Yes	2	2
Clarias batrachus	Walking catfish	Fish	No	No	8	0
Cyprinus carpio	Common carp	Fish	Yes	No	81	3
Gambusia affinis	Gambusia	Fish	Yes	Yes	7	1
Lates niloticus	Nile perch	Fish	No	No	1	0
Micropterus salmoides	Largemouth bass	Fish	Yes	No	15	0
Oncorhynchus mykiss	Rainbow trout	Fish	Yes	Yes	182	9
Oreochromis mossambicus	Mozambique tilapia	Fish	No	No	5	0
Salmo trutta	Brown trout	Fish	Yes	Yes	44	6
Achatina fulica	Giant African snail	Invertebrate	No	No	1	0
Aedes albopictus	Tiger mosquito	Invertebrate	Yes	Yes	35	9
Anopheles quadrimaculatus	Common malaria mosquito	Invertebrate	No	No	1	0
Anoplolepis gracilipes	Yellow crazy ant	Invertebrate	No	No	3	0
Anoplophora glabripennis	Asian long-horned beetle	Invertebrate	No	No	0	0
Asterias amurensis	Flatbottom seastar	Invertebrate	No	No	0	0

Bemisia tabaci	Cotton whitefly	Invertebrate	Yes	Yes	31	3
Carcinus maenas	European shore crab	Invertebrate	Yes	Yes	9	5
Cercopagis pengoi	Fishhook waterflea	Invertebrate	No	No	0	0
Cinara cupressi	Cypress aphid	Invertebrate	No	No	0	0
Coptotermes formosanus	Formosan subterranean termite	Invertebrate	Yes	Yes	2	1
Dreissena polymorpha	Zebra mussel	Invertebrate	No	No	4	0
Eriocheir sinensis	Chinese freshwater edible crab	Invertebrate	No	No	12	0
Euglandina rosea	Cannibal snail	Invertebrate	No	No	1	0
Linepithema humile	Argentine ant	Invertebrate	No	No	1	0
Lymantria dispar	Spongy moth	Invertebrate	Yes	Yes	13	2
Mnemiopsis leidyi	Warty comb jelly	Invertebrate	Yes	Yes	6	1
Mytilus galloprovincialis	Blue mussel	Invertebrate	Yes	Yes	24	3
Pheidole megacephala	Big-headed ant	Invertebrate	No	No	0	0
Platydemus manokwari	Snail-eating flatworm	Invertebrate	No	No	0	0
Pomacea canaliculata	Apple snail	Invertebrate	No	No	3	0
Potamocorbula amurensis	Asian clam	Invertebrate	No	No	2	0
Solenopsis invicta	Red imported fire ant	Invertebrate	Yes	Yes	9	2
Trogoderma granarium	Khapra beetle	Invertebrate	No	No	2	0
Vespula vulgaris	Common wasp	Invertebrate	No	No	1	0
Wasmannia auropunctata	Little fire ant	Invertebrate	No	No	1	0
Aphanomyces astaci	Crayfish plague	Macro-fungi	No	No	0	0
Batrachochytrium dendrobatidis	Chytrid frog fungi	Macro-fungi	Yes	Yes	6	1
Cryphonectria parasitica	Chestnut blight	Macro-fungi	Yes	Yes	3	2

Ophiostoma ulmi sensu lato	Dutch elm disease	Macro-fungi	No	No	2	0
Phytophthora cinnamomi	Cinnamon fungus	Macro-fungi	No	No	9	0
Capra hircus	Domestic goat	Mammal	Yes	No	55	4
Cervus elaphus	Red deer	Mammal	Yes	No	52	2
Felis catus	Domestic cat	Mammal	Yes	No	65	2
Herpestes javanicus	Small indian mongoose	Mammal	No	No	1	0
Macaca fascicularis	Crab-eating macaque	Mammal	Yes	No	25	1
Mus musculus	Field mouse	Mammal	Yes	Yes	87	4
Mustela erminea	Stoat	Mammal	No	No	0	0
Myocastor coypus	Copyu	Mammal	No	No	0	0
Oryctolagus cuniculus	European rabbit	Mammal	Yes	Yes	27	1
Rattus rattus	Black rat	Mammal	Yes	Yes	14	1
Sciurus carolinensis	Gray squirrel	Mammal	No	No	2	0
Sus scrofa	Wild boar	Mammal	Yes	Yes	258	13
Trichosurus vulpecula	Brushtail possum	Mammal	No	No	2	0
Vulpes vulpes	European red fox	Mammal	Yes	No	26	3
Banana bunchy top virus	Banana bunchy top virus	Micro-organism	No	No	0	0
Plasmodium relictum	Avian malaria	Micro-organism	No	No	11	0
Acacia Mearnsii	Black wattle	Plant	No	No	0	0
Ardisia elliptica	Shoebutton ardisia	Plant	No	No	0	0
Arundo donax	Giant reed	Plant	No	No	0	0
Caulerpa taxifolia	Killer alga	Plant	No	No	0	0
Cecropia peltata	Trumpet tree	Plant	No	No	0	0
Chromolaena odorata	Bitterbush	Plant	No	No	1	0
Cinchona pubescens	Red cinchona	Plant	No	No	0	0
Clidemia hirta	Clidemia hirta	Plant	No	No	0	0

Eichhornia crassipes	Floating water hyacinth	Plant	No	No	3	0
Euphorbia esula	Leafy spurge	Plant	No	No	0	0
Hedychium gardnerianum	Kahili ginger	Plant	No	No	0	0
Hiptage benghalensis	Hiptage benghalensis	Plant	No	No	0	0
Imperata cylindrica	Cogon grass	Plant	Yes	Yes	1	1
Lantana camara	West Indian latana	Plant	No	No	2	0
Leucaena leucocephala	River tamarind	Plant	No	No	2	0
Ligustrum robustum	Common privet	Plant	No	No	1	0
Lythrum salicaria	Purple loosetrife	Plant	No	No	0	0
Melaleuca quinquenervia	Broad leaved paperbark tree	Plant	No	No	0	0
Miconia calvescens	Velvet tree	Plant	No	No	0	0
Mikania micrantha	American rope	Plant	Yes	Yes	3	1
Mimosa pigra	Bashful plant	Plant	No	No	0	0
Morella faya	Fire tree	Plant	No	No	0	0
Opuntia stricta	Araluen pear	Plant	No	No	0	0
Pinus pinaster	Maritime pine	Plant	Yes	No	30	2
Polygonum cuspidatum	Donkey rhubarb	Plant	No	No	0	0
Prosopis glandulosa	Honey mesquite	Plant	No	No	0	0
Psidium cattleianum	Strawberry guava	Plant	No	No	0	0
Pueraria montana	Kudzu	Plant	No	No	2	0
Rubus ellipticus	Asian wild raspberry	Plant	No	No	0	0
Salvinia molesta	African payal	Plant	No	No	0	0
Schinus terebinthifolius	Baie rose	Plant	No	No	0	0
Spartina anglica	Common cord grass	Plant	No	No	0	0
Spathodea campanulata	African tulip tree	Plant	No	No	1	0
Sphagneticola trilobata	Bay biscayne creeping oxeye	Plant	No	No	0	0

Tamarix ramosissima	Salt cedar	Plant	No	No	0	0
Ulex europaeus	Gorse	Plant	No	No	0	0
Undaria pinnatifida	Wakame	Plant	Yes	Yes	6	2
Boiga irregularis	Brown treesnake	Reptile	No	No	2	0
Trachemys scripta elegans	Red-eared slider	Reptile	Yes	Yes	3	2
				Total	**1217**	**91**

Table A2.5 Each species on the "IUCN 100 of the World's Worst Invasive Alien Species" list, whether they have a reference genome. If they do, assembly quality information is given.

Species name	Is there a reference genome?	Strain or isolate (year), assembly accession	Release year	Assembly level	Total sequence length	Number of scaffolds	Scaffold N50	Scaffold L50	Number of contigs	Contig N50	Contig L50
Lithobates catesbeianus	Yes	Bruno; GCA_002284835.2	2017	Scaffold	6250.35	1,544,635	5,415	256,279	-	-	-
Rhinella marina	Yes	GCA_900303285.1	2018	Contig	2,551,760,146	-	-	-	31,391	167,498	3,374
Eleutherodactylus coqui	Yes	HN-11 Male; GCA_019857665.1	2021	Chromosome	2,789,347,782	105,233	109,467,076	8	480,031	10,856	54,726
Acridotheres tristis	No										
Pycnonotus cafer	No										
Sturnus vulgaris	Yes	715; GCA_001447265.1	2015	Scaffold	1,036,755,994	2,361	3,416,706	89	22,666	151,865	1,828
Clarias batrachus	Yes	FL201406; GCA_003987875.1	2018	Scaffold	821,750,104	10,042	361,123	611	78,047	24,893	8,096
Cyprinus carpio	Yes	SPL01; GCA_018340385.1	2021	Chromosome	1,680,134,903	6,701	29,545,497	24	19,838	1,558,716	229
Gambusia affinis	Yes	GCF_019740435.1	2021	Chromosome	680,166,006	37	29,761,488	11	218	12,906,370	17
Lates niloticus	No										
Micropterus salmoides	Yes	LMB-PRFRI; GCA_014851395.1	2020	Scaffold	963,613,914	4,753	36,481,429	12	7,662	1,227,323	202
Oncorhynchus mykiss	Yes	GCA_013265735.3	2020	Chromosome	2,341,688,614	939	39,165,350	22	1,229	15,579,713	41
Oreochromis mossambicus	No										
Salmo trutta	Yes	GCA_901001165.2	2021	Chromosome	2,371,880,186	1,441	52,209,666	18	5,378	1,703,178	294
Achatina fulica	No										
Aedes albopictus	Yes	FPA; GCA_006496715.1	2019	Scaffold	2,538,387,871	2,197	55,702,539	13	5,556	1,184,735	434
Anopheles quadrimaculatus	No										
Anoplolepis gracilipes	No										

Anoplophora glabripennis	Yes	ALB-LARVAE; GCA_000390285.2	2017	Scaffold	706,968,555	9,867	678,234	269	26,749	80,490	2,230
Asterias amurensis	No										
Bemisia tabaci	Yes	MEAM1;GCA_001854935.1	2016	Scaffold	615,017,152	19,751	3,232,964	56	31,571	84,501	2,037
Carcinus maenas	No										
Cercopagis pengoi	No										
Cinara cupressi	No										
Coptotermes formosanus	No										
Dreissena polymorpha	Yes	Duluth1; GCA_020536995.1	2021	Chromosome	1,798,009,535	194	117,515,028	6	2,862	1,111,654	444
Eriocheir sinensis	Yes	nwpu-v1-2020;GCA_013436485.1	2020	Chromosome	1,272,135,116	4,311	17,608,299	30	6,666	3,161,423	96
Euglandina rosea	No										
Linepithema humile	Yes	GCA_000217595.1	2011	Scaffold	219,500,750	3,030	1,402,257	40	18,227	35,858	1,697
Lymantria dispar	Yes	GCA_016802235.1	2021	Contig	998,430,749	-	-	-	4,622	661,876	305
Mnemiopsis leidyi	Yes	GCA_000226015.1	2011	Scaffold	155,865,547	5,100	187,314	242	24,927	11,914	3,659
Mytilus galloprovincialis	Yes	GCA_900618805.1	2020	Scaffold	1,282,208,009	10,577	207,642	1,904	22,883	77,157	4,913
Pheidole megacephala	No										
Platydemus manokwari	No										
Pomacea canaliculata	Yes	SZHN2017;GCA_003073045.1	2018	Chromosome	440,159,624	24	31,531,291	6	746	1,072,857	121
Potamocorbula amurensis	No										
Solenopsis invicta	Yes	M01_SB;GCA_016802725.1	2021	Chromosome	378,101,515	219	26,227,205	7	317	9,421,447	13
Trogoderma granarium	No										
Vespula vulgaris	Yes	GCA_905475345.1	2021	Chromosome	188,204,803	28	8,749,684	8	50	8,516,319	10

Wasmannia auropunctata	Yes	WASHAW1;GCA_000956235.1	2015	Scaffold	324,120,201	77,788	1,175,369	55	103,610	37,912	1,666
Aphanomyces astaci	Yes	Strain: APO3;GCA_000520075.1	2014	Scaffold	75,844,385	835	657,536	31	4,659	36,439	421
Batrachochytrium dendrobatidis	Yes	Strain: JEL423; GCA_000149865.1	2006	Scaffold	23,897,668	70	1,707,251	5	351	221,037	32
Cryphonectria parasitica	Yes	Strain: ES15; GCA_018104285.1	2021	Chromosome	43,305,322	18	4,077,748	5	19	4,077,748	5
Ophiostoma ulmi sensu lato	No										
Phytophthora cinnamomi	Yes	Strain: GKB4; GCA_018691715.1	2021	Scaffold	109,702,272	133	1,187,988	30	205	765,543	39
Capra hircus	Yes	GCA_001704415.1	2016	Chromosome	2,922,813,246	29,907	87,277,232	13	30,399	26,244,591	32
Cervus elaphus	Yes	GCA_910594005.1	2021	Chromosome	2,886,603,524	144	83,473,711	13	184	68,741,660	17
Felis catus	Yes	GCA_018350175.1	2021	Chromosome	2,425,747,038	71	148,491,486	7	110	90,731,473	10
Herpestes javanicus	No										
Macaca fascicularis	Yes	CE1976F; GCA_012559485.3	2021	Chromosome	2,800,899,072	401	149,803,873	8	1,006	26,251,577	31
Mus musculus	Yes	Strain: C57BL/6J; GCA_000001635.9	2020	Chromosome	2,728,222,451	102	106,145,001	11	306	59,462,871	15
Mustela erminea	Yes	mMusErm1; GCA_009829155.1	2020	Chromosome	2,445,217,270	94	130,149,454	8	293	36,329,944	19
Myocastor coypus	No										
Oryctolagus cuniculus	Yes	GCA_000003625.1	2009	Chromosome	2,737,462,810	3,318	35,972,871	22	84,024	64,648	12,076
Rattus rattus	Yes	New Zealand; GCA_011064425.1	2020	Chromosome	2,382,787,281	2,173	158,824,752	6	7,847	1,635,336	371
Sciurus carolinensis	Yes	GCA_902686445.2	2020	Chromosome	2,815,413,801	753	148,229,995	8	2,577	13,975,867	52

Sus scrofa	Yes	TJ Tabasco; GCA_000003025.6	2017	Chromosome	2,501,912,388	706	88,231,837	9	1,118	48,231,277	15
Trichosurus vulpecula	Yes	mTriVul1; GCA_011100635.1	2020	Chromosome	3,359,347,707	212	442,560,073	4	2,007	4,314,688	218
Vulpes vulpes	No										
Banana bunchy top virus	Yes	GCA_000847305.1	1993	Complete genome	6,396						
Plasmodium relictum	Yes	Strain: SGS1; GCA_900005765.1	2016	Chromosome	22,607,330	514	1,287,098	6	724	583,861	12
Acacia Mearnsii	No										
Ardisia elliptica	No										
Arundo donax	No										
Caulerpa taxifolia	No										
Cecropia peltata	No										
Chromolaena odorata	No										
Cinchona pubescens	No										
Clidemia hirta	No										
Eichhornia crassipes	No										
Euphorbia esula	Yes	GCA_002919075.1	2018	Scaffold	1,124,886,465	1,633,094	1,035	278,878	2,242,201	605	496,565
Hedychium gardnerianum	No										
Hiptage benghalensis	No										
Imperata cylindrica	No										
Lantana camara	No										
Leucaena leucocephala	No										
Ligustrum robustum	No										
Lythrum salicaria	No										
Melaleuca quinquenervia	No										

Miconia calvescens	No										
Mikania micrantha	Yes	NLD-2019; GCA_009363875.1	2019	Chromosome	1,790,643,622	2,815	86,674,320	9	4,414	1,353,263	250
Mimosa pigra	No										
Morella faya	No										
Opuntia stricta	No										
Pinus pinaster	No										
Polygonum cuspidatum	No										
Prosopis glandulosa	No										
Psidium cattleianum	No										
Pueraria montana	Yes	ZG-11; GCA_019096045.1	2021	Scaffold	1,386,382,709	281	130,661,842	5	5,042	600,355	560
Rubus ellipticus	No										
Salvinia molesta	No										
Schinus terebinthifolius	No										
Spartina anglica	No										
Spathodea campanulata	No										
Sphagneticola trilobata	No										
Tamarix ramosissima	No										
Ulex europaeus	No										
Undaria pinnatifida	Yes	A029; GCA_012845835.1	2020	Scaffold	511,280,173	114	16,510,065	12	618	1,670,562	93
Boiga irregularis	No										
Trachemys scripta elegans	Yes	TJP31775; GCA_013100865.1	2020	Chromosome	2,126,199,303	138	140,411,086	5	17,980	204,575	2,897

Appendix 3.

Supplementary Information: Chapter 3

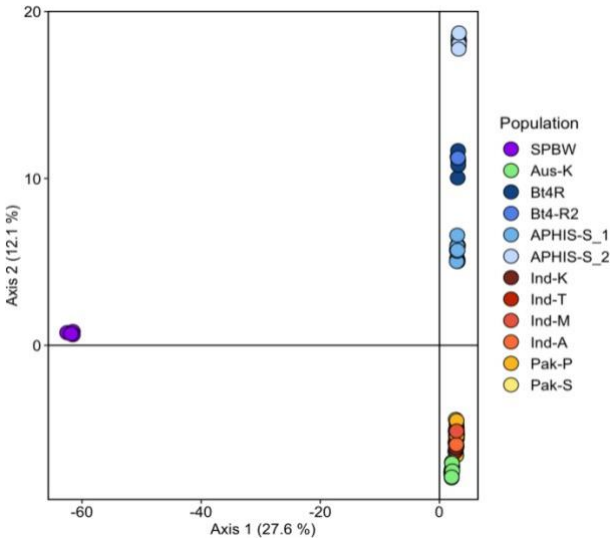


Figure A3.1 Principal component analysis (PCA) of 11 populations of pink bollworm and one population of spotted pink bollworm (*Pectinophora scutigera*; SPBW). Refer to Table 3.1 for population code details.

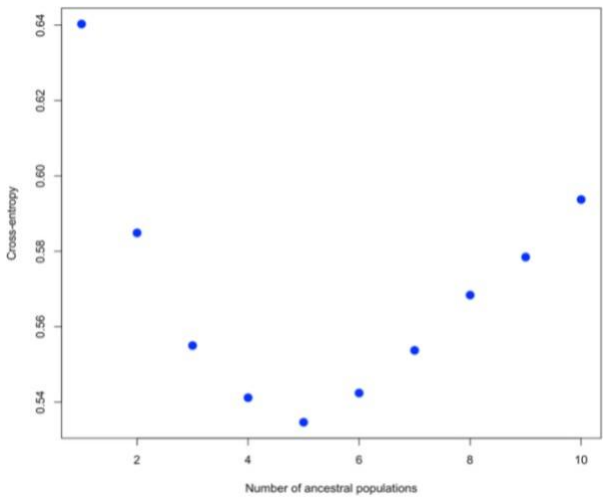


Figure A3.2 Cross-entropy graph indicating an optimal number of five ancestral populations for K in the admixture (sNMF) analysis.

Figure A3.3 The first round of demographic scenarios for demographic modelling (DIYABC analysis). Parameters are defined in Table A3.3; 'pop 1' = America, 'pop 2' = Australia, and 'pop 3' = Asia.

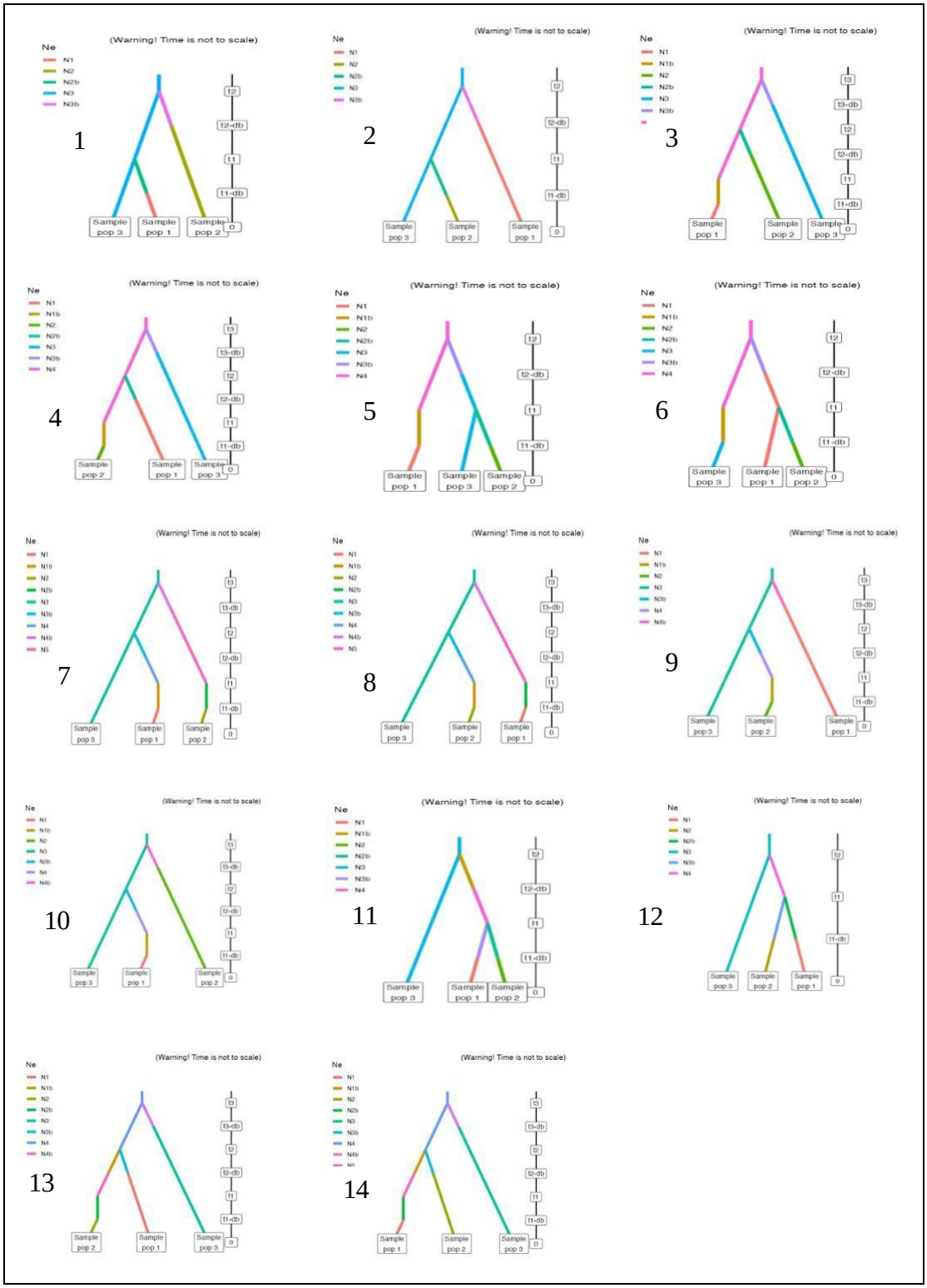


Figure A3.4 Second round of potential demographic scenarios for demographic modelling (DIYABC analysis). Parameters are defined in Table A3.4; 'pop 1' = America, 'pop 2' = Australia, and 'pop 3' = Asia.

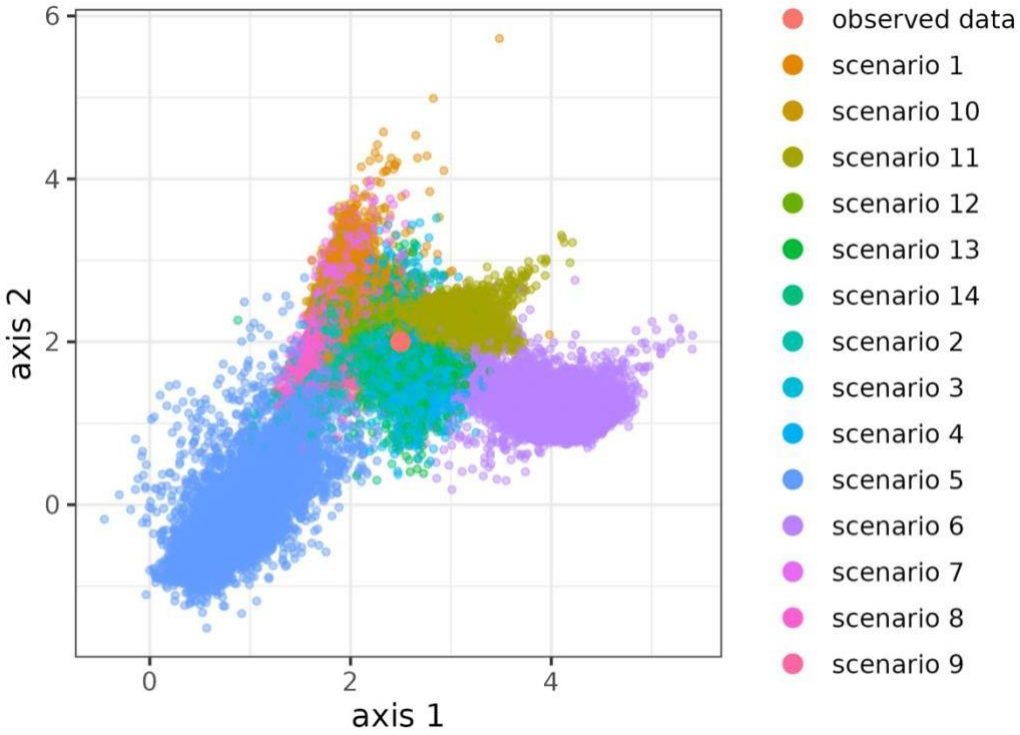


Figure A3.5 Evaluating the fit between the observed dataset and the simulated datasets for each scenario in the demographic modelling (DIYABC) analysis.

Figure A3.6 Admixture proportions for 11 populations of pink bollworm. Bar plots represent admixture proportions for each individual at K = 3, K=4, and K=6 as indicated. Refer to Table 3.1 for population code details.

Table A3.1 Sample information for each pink bollworm individual in the analysis.

Individual	Population	Country	State or Province	City or Lab	Stage
KNX11-4	Aus-K	Australia	Western Australia	Kununurra	Moth (male)
KNX11-5	Aus-K	Australia	Western Australia	Kununurra	Moth (male)
RT1-4	Aus-K	Australia	Western Australia	Kununurra	Moth (male)
RT1-5	Aus-K	Australia	Western Australia	Kununurra	Moth (male)
RT2-1	Aus-K	Australia	Western Australia	Kununurra	Moth (male)
RT2-11	Aus-K	Australia	Western Australia	Kununurra	Moth (male)
RT2-17	Aus-K	Australia	Western Australia	Kununurra	Moth (male)
RT2-3	Aus-K	Australia	Western Australia	Kununurra	Moth (male)
RT2-4	Aus-K	Australia	Western Australia	Kununurra	Moth (male)
RT2-5	Aus-K	Australia	Western Australia	Kununurra	Moth (male)
RT3-2	Aus-K	Australia	Western Australia	Kununurra	Moth (male)
RT3-6	Aus-K	Australia	Western Australia	Kununurra	Moth (male)
ATP-1	Ind-A	India	Andhra Pradesh	Anantapur	Larva
ATP-2	Ind-A	India	Andhra Pradesh	Anantapur	Larva
ATP-3	Ind-A	India	Andhra Pradesh	Anantapur	Larva
ATP-4	Ind-A	India	Andhra Pradesh	Anantapur	Larva
NDL-1	Ind-A	India	Andhra Pradesh	Nandyal	Larva
NDL-2	Ind-A	India	Andhra Pradesh	Nandyal	Larva
P1	Ind-A	India	Andhra Pradesh	Prakasam	Larva
P2	Ind-A	India	Andhra Pradesh	Prakasam	Larva
P3	Ind-A	India	Andhra Pradesh	Prakasam	Larva
P4	Ind-A	India	Andhra Pradesh	Prakasam	Larva
RCR-1	Ind-K	India	Karnataka	Raichur	Larva
RCR-2	Ind-K	India	Karnataka	Raichur	Larva
RCR-3	Ind-K	India	Karnataka	Raichur	Larva

RCR-4	Ind-K	India	Karnataka	Raichur	Larva
MA1	Ind-M	India	Maharashtra	Akola	Larva
MA2	Ind-M	India	Maharashtra	Akola	Larva
MA3	Ind-M	India	Maharashtra	Akola	Larva
MA4	Ind-M	India	Maharashtra	Akola	Larva
MP-1	Ind-M	India	Maharashtra	Parbhani	Larva
W1	Ind-M	India	Maharashtra	Warda	Larva
W2	Ind-M	India	Maharashtra	Warda	Larva
W3	Ind-M	India	Maharashtra	Warda	Larva
W4	Ind-M	India	Maharashtra	Warda	Larva
KNL-2	Ind-T	India	Telangana	Karnool	Larva
KNL-3	Ind-T	India	Telangana	Karnool	Larva
WGL-1	Ind-T	India	Telangana	Warangal	Larva
WGL-2	Ind-T	India	Telangana	Warangal	Larva
WGL-4	Ind-T	India	Telangana	Warangal	Larva
Pk1.1	Pak-P	Pakistan	Punjab	Faisalabad	Larva
Pk1.2	Pak-P	Pakistan	Punjab	Faisalabad	Larva
Pk1.3	Pak-P	Pakistan	Punjab	Faisalabad	Larva
Pk1.4	Pak-P	Pakistan	Punjab	Faisalabad	Larva
Pk10.1	Pak-P	Pakistan	Punjab	Bahawalpur	Larva
Pk10.2	Pak-P	Pakistan	Punjab	Bahawalpur	Larva
Pk2.1	Pak-P	Pakistan	Punjab	Toba Tek Singh	Larva
Pk3.1	Pak-P	Pakistan	Punjab	Toba Tek Singh	Larva
Pk4.1	Pak-P	Pakistan	Punjab	Multan	Larva
Pk4.2	Pak-P	Pakistan	Punjab	Multan	Larva
Pk5.1	Pak-P	Pakistan	Punjab	Multan	Larva
Pk5.2	Pak-P	Pakistan	Punjab	Multan	Larva

Pk7.1	Pak-P	Pakistan	Pubjab	Vehari	Larva
Pk8.1	Pak-P	Pakistan	Pubjab	Vehari	Larva
Pk18.1	Pak-S	Pakistan	Sindh	Hyderabad	Larva
Pk18.2	Pak-S	Pakistan	Sindh	Hyderabad	Larva
Pk18.3	Pak-S	Pakistan	Sindh	Hyderabad	Larva
Pk18.4	Pak-S	Pakistan	Sindh	Hyderabad	Larva
APHIS-1	APHIS-S_1	U.S.	Arizona	Lab	4th instar larva
APHIS-18	APHIS-S_1	U.S.	Arizona	Lab	4th instar larva
APHIS-19	APHIS-S_1	U.S.	Arizona	Lab	4th instar larva
APHIS-2	APHIS-S_1	U.S.	Arizona	Lab	4th instar larva
APHIS-20	APHIS-S_1	U.S.	Arizona	Lab	4th instar larva
APHIS-21	APHIS-S_1	U.S.	Arizona	Lab	4th instar larva
APHIS-22	APHIS-S_1	U.S.	Arizona	Lab	4th instar larva
APHIS-23	APHIS-S_1	U.S.	Arizona	Lab	4th instar larva
APHIS-3	APHIS-S_1	U.S.	Arizona	Lab	4th instar larva
APHIS-4	APHIS-S_1	U.S.	Arizona	Lab	4th instar larva
APHIS-5	APHIS-S_1	U.S.	Arizona	Lab	4th instar larva
APHIS-6	APHIS-S_1	U.S.	Arizona	Lab	4th instar larva
APHIS-41	APHIS-S_2	U.S.	Arizona	Lab	4th instar larva
APHIS-42	APHIS-S_2	U.S.	Arizona	Lab	4th instar larva
APHIS-43	APHIS-S_2	U.S.	Arizona	Lab	4th instar larva
APHIS-44	APHIS-S_2	U.S.	Arizona	Lab	4th instar larva
APHIS-45	APHIS-S_2	U.S.	Arizona	Lab	4th instar larva
APHIS-46	APHIS-S_2	U.S.	Arizona	Lab	4th instar larva
APHIS-53	APHIS-S_2	U.S.	Arizona	Lab	4th instar larva
APHIS-54	APHIS-S_2	U.S.	Arizona	Lab	4th instar larva
APHIS-55	APHIS-S_2	U.S.	Arizona	Lab	4th instar larva

APHIS-56	APHIS-S_2	U.S.	Arizona	Lab	4th instar larva
APHIS-57	APHIS-S_2	U.S.	Arizona	Lab	4th instar larva
Bt4R-6	Bt4R	U.S	Arizona	Lab	4th instar larva
Bt4R-7	Bt4R	U.S	Arizona	Lab	4th instar larva
Bt4R-8	Bt4R	U.S	Arizona	Lab	4th instar larva
Bt4R-9	Bt4R	U.S	Arizona	Lab	4th instar larva
Bt4R2-6	Bt4-R2	U.S	Arizona	Lab	4th instar larva
Bt4R2-7	Bt4-R2	U.S	Arizona	Lab	4th instar larva
Bt4R2-8	Bt4-R2	U.S	Arizona	Lab	4th instar larva

Table A3.2 Sample information for COI sequences downloaded from NCBI database (https://www.ncbi.nlm.nih.gov/).

Sample	Accession number	Genotype	Location information	Author
US_1	JF815079	A	University of Arizona, U.S.	Moore & Hughes, 2011
US_2	JF815080	A	University of Arizona, U.S.	Moore & Hughes, 2011
US_3	JF815081	A	University of Arizona, U.S.	Moore & Hughes, 2011
Aus_1	KF387796	A	Western Australia	McKeown et al. 2013
Aus_2	KF391287	A	Kununurra, Western Australia	Hebert et al. 2013
Aus_3	KF394480	A	Gulon Point, Northern Territory, Australia	Hebert et al. 2013
India_1	KM289071	A	Shankarnagar, Maharashtra, India	Sridhar et al. 2014
India_2	KM289072	A	Shankarnagar, Maharashtra, India	Sridhar et al. 2014
India_3	KM289073	C	Shankarnagar, Maharashtra, India	Sridhar et al. 2014
India_4	KM289074	D	Shankarnagar, Maharashtra, India	Sridhar et al. 2014
India_5	KM289075	E	Shankarnagar, Maharashtra, India	Sridhar et al. 2014
India_6	KM289076	F	Shankarnagar, Maharashtra, India	Sridhar et al. 2014
India_7	KM289077	G	Shankarnagar, Maharashtra, India	Sridhar et al. 2014
India_8	KM289078	H	Shankarnagar, Maharashtra, India	Sridhar et al. 2014
India_9	KM289079	I	Shankarnagar, Maharashtra, India	Sridhar et al. 2014
India_10	KM289080	J	Shankarnagar, Maharashtra, India	Sridhar et al. 2014
India_11	KM289081	K	Shankarnagar, Maharashtra, India	Sridhar et al. 2014
India_12	KM289082	L	Shankarnagar, Maharashtra, India	Sridhar et al. 2014
Pakistan_1	KX860287	A	Faisalabad, Punjab, Pakistan	Ashfaq et al. 2017
Pakistan_2	KX861047	B	Faisalabad, Punjab, Pakistan	Ashfaq et al. 2017
Pakistan_3	KX862450	B	Faisalabad, Punjab, Pakistan	Ashfaq et al. 2017
Pakistan_4	KX862806	B	Faisalabad, Punjab, Pakistan	Ashfaq et al. 2017
Pakistan_5	KX863147	B	Faisalabad, Punjab, Pakistan	Ashfaq et al. 2017
Kenya_1	MF121861	A	Siaya, Rarieda, Obaga, Kenya	Kinyanjui et al. 2017

Kenya_2	MF121862	A	Siaya, Rarieda, Obaga, Kenya	Kinyanjui et al. 2017
Kenya_3	MF121863	A	Siaya, Rarieda, Obaga, Kenya	Kinyanjui et al. 2017
Kenya_4	MF121864	A	Siaya, Rarieda, Obaga, Kenya	Kinyanjui et al. 2017
Israel_1	JF815075	M	Israel	Moore & Hughes, 2011
Israel_2	JF815076	A	Israel	Moore & Hughes, 2011
Israel_3	JF815078	A	Israel	Moore & Hughes, 2011
Israel_4	JF815077	A	Israel	Moore & Hughes, 2011

Table A3.3 Parameters used for demographic modelling (DIYABC analysis), their meaning, and chosen prior distribution value for the first set of scenarios in Fig. A3.3

Parameter	Parameter meaning	Uniform prior
N3b	Effective population size of bottleneck population	2-100
N2b	Effective population size of bottleneck population	2-100
N1	Effective population size for population 1	10-100,000
N2	Effective population size for population 2	10-100,000
N3	Effective population size for population 3	10-100,000
t1	Time in generations of first event	1,200 – 1,800
t2	Time in generations of second event	1,800 – 4,000
NA	Effective population size of the last common ancestral	10 – 1,000,000
db	Length of time of bottleneck event in generations	20-250

Table A3.4 Parameters used for demographic modelling (DIYABC analysis), their meaning, and chosen prior distribution value for the second set of scenarios in Fig. A3.4

Parameter	Parameter meaning	Uniform prior distribution (min-max)
N3b	Effective population size of bottleneck	10-100
N2b	Effective population size of bottleneck	10-100
N3b	Effective population size of bottleneck	10-100
N4b	Effective population size of bottleneck	10-100
N1	Effective population size for population 1	10-100,000
N2	Effective population size for population 2	10-100,000
N3	Effective population size for population 3	10-100,000
N4	Effective population size for ghost population	10-100,000
N5	Effective population size for ghost population	10-100,000
t1	Time in generations of first event	1,000-2,000
t2	Time in generations of second event	2,000-3,000
t3	Time in generations of third event	3,000-4,000
db	Length of time of bottleneck event in	20-250

Table A3.5 Scenario number and its associated number of classification votes (i.e., number of times the scenario was chosen as best suited to the target dataset among the set of compared scenarios) for demographic modelling (DIYABC analysis) of the second set of scenarios. Scenario numbers are given in Figure A3.4. The scenario with the highest number of classification votes and posterior probability (P) is highlighted in yellow.

Scenario	Votes	
1	132	
2	21	
3	439	P = 0.518
4	365	
5	95	
6	10	
7	60	
8	20	
9	8	
10	117	
11	23	
12	84	
13	276	
14	350	