SCISPACE
formerly Typeset

Open access • Dissertation • DOI:10.14264/UQL.2018.552

Conserving evolutionary processes — Source link

Jeffrey Owen Hanson

Published on: 22 Jul 2018

Topics: Protected area, Conservation biology and Convention on Biological Diversity

Related papers:

- Rapid plant diversification: Planning for an
- Integrating Biosystematic Data into Conservation Planning: Perspectives from Southern Africa's Succulent Karoo
- Decision Theory in Conservation Biology Case Studies in Mathematical Conservation
- raptr: representative and adequate prioritization toolkit in R
- Planning for persistence: using the risk of extinction in multi-species approaches to conservation planning

Share this paper:

View more about this paper here: https://typeset.io/papers/conserving-evolutionary-processes-576u6d8o3s



THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

Conserving evolutionary processes

Jeffrey Owen Hanson

BSc (Hons)

A thesis submitted for the degree of Doctor of Philosophy at The University of Queensland in 2018

School of Biological Sciences

Abstract

One of the most fundamental aims in conservation is to ensure the long-term persistence of biodiversity. To achieve this goal, hundreds of thousands of protected areas have been set aside globally to buffer species from anthropogenic impacts and provide a platform for management actions. To be effective, protected areas must preserve existing patterns of biodiversity (e.g. species, ecosystems) and also the evolutionary processes that create new patterns of biodiversity. By ensuring that disruptions to evolutionary processes are minimized, protected areas can help maintain existing patterns of genetic diversity and facilitate adaptation to new threats. This is particularly important in a world where environmental change is rapidly accelerating. However, despite this, evolutionary processes are rarely considered when siting new protected areas or evaluating existing protected area systems.

The main goal of this thesis is to improve our understanding of how evolutionary processes can be incorporated into conservation planning to deliver more effective protected areas. To achieve this goal, I develop a novel decision support tool to target intra-specific variation in conservation prioritizations (**Chapter 2**). I then investigate potential surrogates for representing intra-specific genetic variation (**Chapter 3**) and maintaining gene flow in prioritizations (**Chapter 4**). Finally, I evaluate how well the existing protected area system is representing adaptive processes for nearly every vertebrate species on Earth (**Chapter 5**).

Building prioritizations that conserve evolutionary processes has been a long standing challenge in conservation planning. In **Chapter 2**, I develop a new decision support tool—the `raptr` `R` package—for generating multi-species prioritizations that minimize the overall cost of the solution whilst (i) securing a representative sample of the intra-specific variation for each species, (ii) providing an adequate amount of habitat for each species, and (iii) minimizing the level of overall fragmentation in the solution. By applying this method to simulated and case-study species, I show that conservation planners need to explicitly target intra-specific variation—otherwise they risk losing it. This chapter paves the way for using intra-specific variation to guide the selection of nature reserves. After developing this decision support tool, I use it to examine potential surrogates for conserving evolutionary processes.

One of the key reasons that evolutionary processes are not often used to guide reserve selection is that substantial resources and expertise are needed to obtain and analyze genetic data. In **Chapters 3 and 4**, I investigate the effectiveness of strategies for capturing evolutionary processes using freely available genetic data for multiple alpine plant species. In **Chapter 3**, I show that prioritizations capturing a representative sample of the climatic variation and geographic spread across species' distributions tend also to capture a large proportion of species' adaptive and neutral genetic variation. In **Chapter 4**, I found that conventional approaches for increasing connectivity may not actually result in prioritizations that maintain strong levels of gene flow. These findings illustrate how genetic data can be used to

guide conservation planning. Furthermore, they show that freely available data can, at least in some cases, be used to deliver effective protected area systems even when genetic data are not available. After demonstrating that environmental data can be used as a surrogate for conserving evolutionary processes, I then use environmental data to explore how well the existing global protected area system is conserving adaptive evolutionary processes.

In response to the biodiversity crisis, 195 governments have signed the Convention on Biological Diversity. These nations have pledged to protect at least 17 % of the Earth's land and improve the conservation status for species at risk of extinction by the year 2020 (Aichi Targets 11 and 12). One of the components for monitoring progress made towards conserving biodiversity is the protection of habitats that contain "key evolutionary processes" (Annex I). By conserving populations in a wide range of climatic conditions, protected area systems can maximize the range of existing and potential local adaptations within a species. In **Chapter 5**, I discover that the global protected area system poorly represents the climatic conditions found across species' geographic distributions. To begin to address this shortfall, I also identify priority areas for protected area expansion. This work provides the first global assessment for how well protected areas are fostering adaptive evolutionary process for the world's amphibians, birds, and mammals.

This thesis sheds new light on understanding how well evolutionary processes are conserved, and how conservation decisions can be made in a way that safeguards evolutionary processes. It combines ideas from conservation biology, decision science, and evolutionary biology. The discoveries made here will be relevant to a broad range of scientists working in conservation and genetics research, and also policy makers and planners engaged in protecting biodiversity. Careful utilization of the decision support tool (**Chapter 2**), genetic and surrogate data (**Chapters 3 and 4**), and priority areas (**Chapter 5**) outlined in this thesis could substantially increase the chances for the long-term persistence of biodiversity.

Declaration by author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, financial support and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my higher degree by research candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the policy and procedures of The University of Queensland, the thesis be made available for research and study in accordance with the Copyright Act 1968 unless a period of embargo has been approved by the Dean of the Graduate School.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis and have sought permission from co-authors for any jointly authored works included in the thesis.

Publications during candidature

Peer reviewed publications

- **Hanson JO**, Rhodes JR, Possingham HP & Fuller RA (2018) raptr: Representative and adequate prioritization toolkit in R. *Methods in Ecology & Evolution*, **9**, 320–330.

  Hanson JO, Rhodes JR, Riginos C & Fuller RA (2017) Environmental and geographic variables are effective surrogates for genetic variation in conservation planning. *Proceedings of the National Academy of Sciences of the United States of America*, **114**, 12755–12760.

- Dhanjal-Adams KL, **Hanson JO**, Murray NJ, Phinn SR, Wingate VR, Mustin K, Lee JR, Allan JR, Cappadonna JL, Studds CE, Clemens RS, Roelfsema CM & Fuller RA (2016). Distribution and protection of intertidal habitats in Australia. *Emu*, **116**, 208–214.

- Dudaniec RY, Worthington Wilmer J, **Hanson JO**, Warren M, Bell S & Rhodes JR (2016) Dealing with uncertainty in landscape genetic resistance models: A case of three co-occurring marsupials. *Molecular Ecology*, **25**, 470–486.

- **Hanson JO**, Salisbury SW, Campbell HA, Dwyer RG, Jardine TD & Franklin CE (2015) Feeding across the food web: The interaction between diet, movement and body size in estuarine crocodiles (*Crocodylus porosus*). *Austral Ecology*, **40**, 275–286.

- Auerbach NA, Wilson KA, Tulloch AI, Rhodes JR, **Hanson JO** & Possingham HP (2015). Effects of threat management interactions on conservation priorities. *Conservation Biology*, **29**, 1626–1635.

- Bunton JD, Ernst AT, **Hanson JO**, Beyer HL, Hammill E, Runge CA, Venter O, Possingham HP & Rhodes JR (2015) Integrated planning of linear infrastructure and conservation offsets. In Weber, T, McPhee, MJ & Andersson RS (eds) *MODSIM 2015, 21st International Congress on Modelling and Simulation*. Modelling and Simulation Society of Australia and New Zealand, December 2015, pp. 1427–1433.

- Rabeb D, Othman DS, Essilfie AT, Hansbro PM, **Hanson JO**, McEwan AG & Kappler U (2015) Maturation of molybdoenzymes and its influence on the pathogenesis of non-typeable *Haemophilus* influenzae. *Frontiers in Microbiology*, **6**, 01219.

- Runge CA, Watson JEM, Butchart HM, **Hanson JO**, Possingham HP & Fuller RA (2015) Protected areas and global conservation of migratory birds. *Science*, **350**, 1255–1258.

Conference presentations

- **Hanson JO**, Rhodes JR, Riginos C, Fuller RA (2017) Effective surrogates for genetic variation in conservation planning. Oral presentation to the 28th International Congress for Conservation Biology, Cartagena, Colombia.
- **Hanson JO**, Rhodes JR, Possingham HP, Fuller RA (2016) RAPR: Representative and Adequate Prioritizations in R. Oral presentation to Society for Conservation Biology 4th Oceania Congress, Brisbane, Australia.

Publications included in this thesis

- **Hanson JO**, Rhodes JR, Possingham HP, & Fuller RA (2017) `raptr`: Representative and adequate prioritization toolkit in `R`. *Methods in Ecology & Evolution*. In press, Available at `https://dx.doi.org/10.1111/2041-210X.12862`. Incorporated as **Chapter 2**.

Contributor	Statement of contribution
Jeffrey O. Hanson (Candidate)	Conception and design (60%) Analysis and interpretation (70%) Drafting and production (70%)
Jonathan R. Rhodes	Conception and design (20%) Analysis and interpretation (10%) Drafting and production (10%)
Hugh P. Possingham	Conception and design (0%) Analysis and interpretation (10%) Drafting and production (10%)
Richard A. Fuller	Conception and design (20%) Analysis and interpretation (10%) Drafting and production (10%)

- **Hanson JO**, Rhodes JR, Riginos C, & Fuller RA (2017) Environmental and geographic variables are effective surrogates for genetic variation in conservation planning. *Proceedings of the National Academy of Sciences of the United States of America*. In press. Available at `https://dx.doi.org/10.1073/pnas.1711009114`. Incorporated as **Chapter 3**.

Contributor	Statement of contribution
Jeffrey O. Hanson (Candidate)	Conception and design (70%) Analysis and interpretation (70%) Drafting and production (70%)
Jonathan R. Rhodes	Conception and design (10%) Analysis and interpretation (10%) Drafting and production (10%)
Cynthia Riginos	Conception and design (10%) Analysis and interpretation (10%) Drafting and production (10%)
Richard A. Fuller	Conception and design (10%) Analysis and interpretation (10%) Drafting and production (10%)

Contributions by others to the thesis

The majority of the thesis was developed, executed, and written by the candidate Jeffrey Hanson. Contributions made by other individuals are listed below.

Chapter 2: The conceptual framework of this chapter was developed by the candidate with substantial contributions by Jonathan Rhodes (supervisor) and Richard Fuller (supervisor). Hugh Possingham assisted with the mathematical concepts and definitions. This chapter was written by the candidate with contributions from all authors.

Chapter 3: The candidate and Richard Fuller developed the main idea for this chapter. Cynthia Riginos provided guidance on preparing and analyzing the genetic data. Jonathan Rhodes provided statistical advice. This chapter was written by the candidate with contributions from all authors.

Chapter 4: The candidate conceived the main idea for this chapter. Jonathan Rhodes and Richard Fuller advised on the methodology and interpreting the results. This chapter was written by the candidate with contributions from all authors.

Chapter 5: The main aim of this chapter was developed by the candidate and Richard Fuller. Carlo Rondinini, Gentile Ficetola, Graeme Buchanan, and Stuart Butchart provided species distribution maps. This chapter was written by the candidate with contributions from all authors.

Statement of parts of the thesis submitted to qualify for the award of another degree

None.

Research involving human or animal subjects

No animal or human participants were involved in this research.

Acknowledgements

I would not have been able to complete this thesis without the assistance, guidance, and support from many people. Among this crowd of colleagues, friends, and family, I would first like to thank my supervisors Richard Fuller and Jonathon Rhodes.

Rich, you are an inspiration. Many years ago, long before I started this PhD project, you first taught me that computational and statistical analyses can be used to address important and interesting questions. Later, your unwavering enthusiasm and encouragement helped me deliver a thesis that achieved this. You opened my eyes to the big questions and taught me how to communicate my findings.

Jonathan, you are an outstanding supervisor. You taught me how to plan my research. You kindly offered statistical advice and helped me frame my results in the context of all those who have come before me. I really value the skills you have taught me.

I would also like to express my sincerest gratitude to Hugh Possingham and James Watson. These extremely busy people made time to offer guidance throughout my PhD project. Hugh, you taught me how to express my ideas using mathematics. James, you helped me to connect the dots and find the gaps. You both helped make my research more relevant.

I am also extremely grateful to my collaborators. I greatly appreciate Cynthia Riginos for explaining concepts and sharing thoughts on my work. I am immensely thankful to Carlo Rondinini, Francesco Ficetola, Graeme Buchanan, Stuart Butchart for contributing data and reading through drafts. All of these people played a role in helping me deliver this thesis.

Zoë, you were unexpected. Your unending support means the world to me.

I also thank my family for their support, cheesecake, and lasagna.

I also thank the National eResearch Collaboration Tools and Resources project (NeCTAR) for providing computational resources.

Finally, my sincere thanks to all my friends in the Centre of Biodiversity and Conservation Science (CBCS) and the School of Biological Sciences at The University of Queensland. Over the last few years, we have shared many accomplishments, many commiserations, and many, many beers on the Goddard building rooftop.

Financial support

This research was supported by an Australian Government Training Program Scholarship.

Key words

Adaptation, alpine plants, connectivity, conservation, evolution, gene flow, integer programming, optimization, prioritization, reserve selection.

Australian and New Zealand Standard Research Classifications

ANZSRC code: 050202 Conservation and Biodiversity 35 %

ANZSRC code: 120504 Land Use and Environmental Planning 35 %

ANZSRC code: 060411 Population, Ecological and Evolutionary Genetics 30 %

Fields of Research (FoR) Classification

FoR code: 0502 Environmental Science and Management 35 %

FoR code: 1205 Urban and Regional Planning 35 %

FoR code: 0604 Genetics 30 %

Table of contents

Abstract ii
Declaration by author iv
Publications during candidature v
Publications included in this thesis vi
Contributions by others to the thesis viii
Statement of parts of the thesis submitted to qualify for the award of another degree viii
Research involving human or animal subjects viii
Acknowledgements ix
Financial support x
Key words x
Australian and New Zealand Standard Research Classifications x
Fields of Research (FoR) Classification x
Table of contents xi
List of figures xiv
List of tables xiv
Abbreviations xv

Chapter 1 Introduction **1**
1.1 General background 2
1.1.1 Systematic conservation planning 2
1.1.2 Evolutionary processes 4
1.1.3 Knowledge gaps 8
1.2 Thesis objectives and structure 9

Chapter 2 `raptr`: Representative and adequate prioritization toolkit in `R` **11**
2.1 Abstract 12
2.2 Introduction 13
2.3 Methods 14
2.3.1 Problem formulation 14
2.3.2 Optimization 20
2.4 Examples 20
2.4.1 Simulation study 20
2.4.1.1 Methods 20
2.4.1.2 Uniformly distributed species 21
2.4.1.3 Normally distributed species 23
2.4.1.4 Bimodally distributed species 23
2.4.2 Case study 1 24
2.4.2.1 Methods 24
2.4.2.2 Results 26
2.4.3 Case study 2 26

2.4.3.1 Methods 26
2.4.3.2 Results 28
2.5 Implications and future directions 29
2.6 Acknowledgements 31
2.7 Data accessibility 31

Chapter 3 Environmental and geographic variables are effective surrogates for genetic variation in conservation planning 33
3.1 Abstract 34
3.2 Introduction 35
3.3 Methods 36
3.3.1 Landscape data 37
3.3.2 Adaptive and neutral genetic data 38
3.3.3 Prioritization method 39
3.3.4 Computational and statistical analyses 39
3.4 Results 41
3.5 Discussion 44
3.6 Acknowledgements 47
3.7 Data accessibility 47

Chapter 4 Conservation planning needs genetic data to maintain gene flow 49
4.1 Abstract 50
4.2 Introduction 51
4.3 Methods 52
4.3.1 Habitat suitability maps 54
4.3.2 Resistance maps 55
4.3.3 Prioritizations 57
4.4 Results 58
4.5 Discussion 61
4.6 Acknowledgements 63

Chapter 5 Global conservation of realized niches 65
5.1 Abstract 66
5.2 Introduction 67
5.3 Methods 68
5.3.1 Data 68
5.3.2 Analysis 69
5.4 Results and Discussion 70
5.5 Conclusion 72
5.6 Acknowledgements 73

Chapter 6 General discussion 75
6.1 Overview 76
6.2 New methods 76

6.2.1 Research summary 76
6.2.2 Limitations 77
6.2.3 Future research directions 78
6.3 Surrogates for genetic data 79
6.3.1 Research summary 79
6.3.2 Limitations 79
6.3.3 Future research directions 80
6.4 Evaluating the conservation of adaptive processes 82
6.4.1 Research Summary 82
6.4.2 Limitations 82
6.4.3 Future research directions 83
6.5 Research perspectives 84
6.6 Concluding remarks 85

References **87**

Appendices **116**
Chapter 2 Appendices 116
Appendix A 116
Appendix B 119
Appendix C 120
Appendix D 122
Chapter 3 Appendices 123
Appendix E 123
Appendix F 124
Appendix G 128
Appendix H 129
Appendix I 138
Appendix J 139
Appendix K 153
Appendix L 154
Appendix M 156
Appendix N 157
Chapter 4 Appendices 158
Appendix O 158
Appendix P 159
Appendix Q 159
Appendix R 160
Appendix S 160
Appendix T 161
Appendix U 163
Appendix V 164

Appendix W 164
Appendix X 164
Chapter 5 Appendices 165
Appendix Y 165
Appendix Z 165

List of figures

2.1 Attribute space example 16
2.2 Distributions of three simulated species 21
2.3 Prioritizations for the simulation study 22
2.4 Climatic variation across Queensland, Australia 24
2.5 Distribution of the species used in the first case study 25
2.6 Prioritizations for the first case study 26
2.7 Data used for the second case study 27
2.8 Prioritizations for the second case study 29

3.1 Study area map 37
3.2 Relationship between amount of surrogate and genetic variation captured in randomly generated prioritizations 42
3.3 Proportion of genetic variation secured in different solutions 44

4.1 Study area 53
4.2 Habitat suitability maps 55
4.3 Landscape resistance maps 56
4.4 Ability for prioritizations to maintain gene flow 59

5.1 Representation of species' climatic partitions 71
5.2 Prioritizations to represent species' niches 72

List of tables

None.

Abbreviations

AFLP	Amplified Fragment Length Polymorphisms
ALA	Atlas of Living Australia
APA	Australian Postgraduate Award
AUC	Area under Receiver Operating Characteristic Curve
CBCS	Centre for Biodiversity and Conservation Science
CBD	Convention on Biological Diversity
DNA	Doxyribonucleic Acid
ESU	Evolutionarily Significant Unit
ED	Environmental Diversity
GBIF	Global Biodiversity Information Facility
GLM	Generalized Linear Model
GLMM	Generalized Linear Mixed-effects Model
GeOMe	Genomic Observatory Metadatabase
IBD	Isolation By Distance
IUCN	International Union for Conservation of Nature
MaxEnt	Maximum Entropy
MLPE	Maximum Likelihood Population Effects
NMDS	Non-metric Dimensional Scaling
PCA	Principal Components Analysis
RAPTR	Representative and Adequate Prioritization Toolkit in R
RRAP	Reliable formulation of the Representative and Adequate Prioritization problem
URAP	Unreliable formulation of the Representative and Adequate Prioritization problem

CHAPTER 1

Introduction

1.1 General background

Biodiversity is under siege (Alford & Richards, 1999; Thomas *et al.*, 2004b). Threatening processes—such as habitat destruction (Hansen *et al.*, 2013; Murray *et al.*, 2014; Allan *et al.*, 2017; Alroy, 2017), disease (Skerratt *et al.*, 2007), and invasive species (Clavero *et al.*, 2009; Doherty *et al.*, 2016)—are driving species to extinction. In fact, the present day extinction rate is so great that it far out-paces previous mass extinctions in the geological record (Wake & Vredenburg, 2008; Barnosky *et al.*, 2011). Over the next century, urban expansion (Seto *et al.*, 2012) and changes to the world's climate regimes (Crowley, 2000) are projected to present even more challenges (Thomas *et al.*, 2004a; Deutsch *et al.*, 2008). In response to this centuries-old trend, the overarching aim of conservation is to protect biodiversity and the processes that sustain it (Moritz, 2002). Today, nearly 15 % of the Earth's surface is covered by protected areas (H.M. *et al.*, 2015), though many of these reserves are sited in places that offer little value to biodiversity (Fuller *et al.*, 2010). More protected areas are needed to guarantee the long-term persistence of biodiversity (McCarthy *et al.*, 2012; Barnes, 2015)—but where should they be located? In this thesis, I present new ways to plan for the long-term persistence of biodiversity by safeguarding evolutionary processes.

1.1.1 Systematic conservation planning

Protected areas are one of the most tangible achievements of conservation (Watson *et al.*, 2014). These areas buffer species from anthropogenic threats and provide places for biodiversity processes to continue. Although strategic management of other land-use types can enhance biodiversity (Polasky *et al.*, 2005; Reyers *et al.*, 2012), protected areas are important for providing a foundation for direct management interventions—such as habitat restoration (e.g. Babu *et al.*, 2009), reintroductions (e.g. King *et al.*, 2012), and pest eradication (e.g. Innes *et al.*, 1999)—that are crucial for species' recovery. For instance, pest eradication in the New Zealand Mapara Wildlife Reserve has helped bring the North Island kōkako (*Callaeas cinerea wilsoni*) back from the brink of extinction (Innes *et al.*, 1999). Not only do conservation areas areas benefit biodiversity, they also play a role in human wellbeing (Bryce *et al.*, 2016). The world's first national park—Yellowstone National Park in The United States of America—was provided "for the benefit and enjoyment of the people" (US Congress 1872). Yet, despite the critical importance of these areas, recent assessments have found serious shortfalls in the global protected area network (H.M. *et al.*, 2015; Runge *et al.*, 2015; Kuempel *et al.*, 2016).

Modern systematic conservation planning was born from the profound realization that "[s]ince there are limited funds available to purchase land as nature reserves, an order of priority [must be] established" (Wright, 1977, p. 293). To establish such an order of priority, conservation scientists require (i) standardized biological, cultural, and cost data to compare can-

didate areas, and (ii) methods to prioritize these areas for protection (reviewed in Margules & Pressey, 2000). Early reserve selection methods used multiple criteria—such as species rarity or diversity (Ratcliffe, 1971; Bunce & Shaw, 1973; Margules & Usher, 1981)—to score candidate areas and rank them for prioritization (e.g. Helliwell, 1969; Tubbs & Blackwood, 1971; Gehlbach, 1975; Wright, 1977; Goldsmith, 1975). These efforts marked some of the earliest attempts at integrating cost-efficiency into conservation planning.

In 1983, James Kirkpatrick developed a new reserve selection method (Kirkpatrick, 1983). Kirkpatrick recognized that prioritizing areas "on the basis of a single application of a formula there is no guarantee that the priority area second or third on the list might not duplicate the species, communities or habitats that could successfully be preserved in the first priority area" (Kirkpatrick, 1983, p. 128). Based on this idea, Kirkpatrick developed an "iterative algorithm" aiming to identify a set of sites that in combination would yield the greatest conservation benefit. This fundamental breakthrough inspired the rapid development of ever more sophisticated algorithms (e.g. Pressey & Nicholls, 1989; Rebelo & Siegfried, 1990, 1992; Nicholls & Margules, 1993). In time, this idea that reserves should complement each other became formally recognized as the principle of "complementarity" (Vane-Wright *et al.*, 1991).

Towards the early 1990s, conservation scientists began to realize that they needed plans for protected area expansion that would fulfill well-defined objectives (Cocks & Baird, 1989). The issue of establishing an order of priority was now a secondary issue to be resolved—using measures such as "irreplaceability" (Pressey *et al.*, 1994)—after obtaining a candidate set of reserves (termed a "prioritization"). Reserve selection methods began to incorporate criteria pertaining to species' persistence, such as reserve size and shape (Nicholls & Margules, 1993; Lombard *et al.*, 1995), and also configuration (Bedward *et al.*, 1992). By formally recognizing reserve selection as an optimization problem, researchers working in applied mathematical research fields began to take an interest in conservation planning (Underhill, 1994). But in these fields, iterative algorithms—known as heuristics—had long been superseded by "exact algorithms" (e.g. branch and bound search processes; Land & Doig, 1960).

During the mid-1990s, the application of exact algorithms to reserve selection problems sparked furious debate (Underhill, 1994; Camm *et al.*, 1996; Williams & ReVelle, 1998; Pressey *et al.*, 1996; Csuti *et al.*, 1997; Pressey *et al.*, 1997). On the one hand, exact algorithm solvers identified the most cost-effective solution (Underhill, 1994; Camm *et al.*, 1996; Williams & ReVelle, 1998). Yet, on the other hand, exact algorithm solvers available at the time (e.g. LINDO; Schrage, 1984) could not solve large planning problems fast enough (Pressey *et al.*, 1996; Csuti *et al.*, 1997; Pressey *et al.*, 1997). There was little consensus between advocates of either camp. In 1996, Ian Ball found a middle-ground. Ball implemented a meta-heuristic algorithm—named "simulated annealing" (Kirkpatrick *et al.*, 1983)—to find near-optimal solutions to large reserve selection problems in a feasible amount of time (Ball,

2000; Ball *et al.*, 2009).

Conservation scientists have come to realize that careful thought is needed to ensure that habitat protection ensures species' long-term persistence (Cowling *et al.*, 1999). For instance, population viability analyses (PVAs) have been used to estimate the minimum population size needed to yield an acceptably low extinction probability (e.g. Reed *et al.*, 1988). Additionally, conservation planning can now plan for the spatial allocation of different management actions in differences places—as implemented in the `MarZone` software (Watts *et al.*, 2009)—leading to more cost-effective implementations (Wilson *et al.*, 2007; Auerbach *et al.*, 2015). Conservation plans can also now be designed to accommodate predictions about future threats such as climate change and land-use transformations (Carvalho *et al.*, 2011a).

Today, conservation planning for the establishment of protected areas makes extensive use of decision support tools, many of which are freely available, such as `Marxan` (Ball *et al.*, 2009), `raptr` (Hanson *et al.*, 2017b), `ResNet` (Sarkar *et al.*, 2002), and `Zonation` (Moilanen *et al.*, 2005). Following technological advances over the last few decades (Elith *et al.*, 2006; Foody, 2008), conservation scientists can now tackle threats to biodiversity at the global scale (Ceballos *et al.*, 2005) using data sets of unprecedented spatial and temporal coverage (e.g. Hansen *et al.*, 2013). We are also beginning to see a resurgence of exact algorithm solvers (Stigner *et al.*, 2016; Hanson *et al.*, 2017b; Law *et al.*, 2017). Modern exact algorithm solvers (e.g. `GUROBI`; Gurobi Optimization Inc., 2016) are now capable of solving bigger problems than either heuristics or meta-heuristics (Beyer *et al.*, 2016). They can also accommodate more complex reserve design criteria to maximize species' persistence (Billionnet, 2013) and foster ecological and evolutionary processes (e.g. Hanson *et al.*, 2017b).

1.1.2 Evolutionary processes

Human activities are threatening the evolutionary processes that generate new patterns of biodiversity (Moritz, 2002; Sgro *et al.*, 2011; Carvalho *et al.*, 2017). On the inter-specific level, evolutionarily distinct species—such as Tuatara (*Sphenodon punctatus*; Daugherty *et al.*, 1990)—represent a unique period of evolutionary history. Their extinction would mean an irrevocable loss in global biodiversity (Faith, 1992; Safi *et al.*, 2013). On the intra-specific level, species often comprise multiple genetic lineages—groups of populations that have been evolving independently for many years—and the loss of a single lineage can entail serious consequences (e.g. Daugherty *et al.*, 1990). The extinction of genetic lineages represents not only the loss of evolutionary history, but also the loss of genetic diversity which directly threatens the ability for an entire species to persist (reviewed in Frankham, 2005). Moreover, such lineages represent budding branches on the tree of life, and their extinction constrains future patterns of biodiversity (Frankel, 1974; Moritz, 1994). Broadly speaking, evolutionary processes can be classified as adaptive or neutral (reviewed in Holderegger

et al., 2006). To achieve the long-term persistence of biodiversity, conservation actions must preserve both types of evolutionary processes (Moritz, 2002). Below, I will discuss each of these evolutionary processes and strategies for conserving them.

In evolutionary terms, the ultimate measure of success for an organism is its ability to survive and reproduce (termed "fitness"; reviewed in Pough, 1989; Kawecki & Ebert, 2004). The fitness of an organism is governed by the phenotypical characters it possesses—termed "traits"—and the external forces that it experiences—termed "selection pressures" (reviewed in Kawecki & Ebert, 2004; Violle *et al.*, 2007). Traits pertain to the intrinsic qualities of an organism (Violle *et al.*, 2007). They may, for example, be morphological (e.g. beak shape; Boag & Grant, 1981), behavioral (e.g. risk aversion; van Oers *et al.*, 2004), or phenological (e.g. flowering time; Kuittinen *et al.*, 1997) in nature. Selection pressures, on the other hand, can originate from an organism's surrounding environment (e.g. soil salinity; Ahmad *et al.*, 2010), conspecifics (e.g. competition for mating opportunities; Klappert & Reinhold, 2003), and other species (e.g. predation; Møller, 1990). Both selection and traits play a role in the survival of an organism, and to a larger extent, the species as a whole.

Across a population, each individual will exhibit a set of traits (reviewed in Pough, 1989). Individuals that possess certain traits may exhibit greater reproductive output. For instance, in cold environments, larger female fruit flies (*Drosophila melanogaster*) live longer and produce more offspring (McCabe & Partridge, 1997). If such advantageous traits arise from heritable genetic variation, they may be passed on to an individual's offspring (Mitchell & Shaw, 1993; Pulido *et al.*, 2001; Leiss *et al.*, 2004). As a consequence, advantageous traits can become more prevalent over subsequent generations, and populations can become better adapted to their surroundings over time (Boag & Grant 1981; but see Kruuk *et al.* 2002). Although evolutionary time-scales often involve thousands of years, such adaptive processes have been observed over much shorter time scales (e.g. Grant & Grant, 1992; Hairston *et al.*, 2005; Franks *et al.*, 2007). Adaptive processes have been recorded in simulation studies (e.g. Endy *et al.*, 2000; Wilke *et al.*, 2001), laboratory experiments (e.g. Wichman *et al.*, 1999; Blount *et al.*, 2008; Sørensen *et al.*, 2017), and wild populations (e.g. Boag & Grant, 1981).

Populations can develop adaptations that are specific to their own local selection pressures (reviewed in Kawecki & Ebert, 2004; Kawecki, 2008). This process—termed "local adaptation"—can enable a species to thrive under a broad range of environmental conditions (e.g. McKay *et al.*, 2001; Gomez-Mestre *et al.*, 2003; Fournier-Level *et al.*, 2011). For example, plant populations have repeatedly shown the capacity to adapt to locally high concentrations of metal in soil (Wright *et al.*, 2006; Dechamps *et al.*, 2008; Turner *et al.*, 2010). In a transplant experiment, Mahieu *et al.* (2013) found that individuals from populations in soil with high metal concentrations experienced less mortality in areas contaminated from mining activities than individuals from populations growing in lower metal concentrations. This result suggests that restoration activities can be more effective by sourcing individuals

from populations that are pre-adapted to local sources of degradation (Mahieu *et al.*, 2013). In addition to enabling species to persist over a range of different environmental conditions in the present, local adaptations also enable species to become resilient against changing environmental conditions over time (Sgro *et al.*, 2011; Nicotra *et al.*, 2015; Breiner *et al.*, 2017).

Over the last few decades, there has been significant debate over the best conservation strategy to foster adaptive processes (reviewed in Moritz, 2002). One strategy has recommended the protection of populations with complementary local adaptations—identified using genetic data (e.g. McKay *et al.*, 2001)—to ensure that they are not lost (Crandall *et al.*, 2000). This strategy aims to fortify species against threats in the near future (e.g. climate change; Grady *et al.*, 2011). Building on this idea, Lesica & Allendorf (1995) proposed that conservation efforts should target populations found along range margins since they may contain key genetic innovations. On the other hand, rather than conserving existing adaptive genetic variation, a different strategy involves conserving populations in habitats that will give rise to new and complementary adaptive variants—since these populations are expected to develop adaptations (as shown in Losos *et al.*, 1997; Reznick *et al.*, 1997)—to ensure these evolutionary powerhouses are not lost (Moritz, 2002; Rouget *et al.*, 2003; Smith *et al.*, 2001). Today, it is generally accepted that protected area systems must be constructed with consideration to both strategies (Moritz, 2002). It is worth noting that the capacity for species to develop local adaptations is mediated by many different factors that also need to be considered during reserve selection (Kawecki & Ebert, 2004). Thus far we have discussed adaptive evolutionary processes, however, non-adaptive evolutionary processes—termed "neutral processes"—have also played a major role in shaping many of the species that we see today.

Neutral evolutionary processes are the result of random genetic changes (reviewed in Masel, 2011). Unlike adaptive processes, such neutral processes affect regions of an organism's genome that have a near-zero effect on fitness (e.g. Endo *et al.*, 1996). Across a population, different individuals may possess different genetic variants at these putatively "neutral" regions (Wang, 2005). Following reproduction, some of the neutral genetic variants may be passed on to the next generation of individuals (Wright, 1929). Since these genetic variants have little effect on fitness, the frequency of these genetic variants in the next generation is expected to follow a random sample of those in the current generation (plus a small amount of new variants created through mutation and recombination; Wright 1931, 1932; but see genetic hitch hiking in Hedrick 1982; Barton 2000). Over multiple generations, some neutral genetic variants may become more prevalent in the population whilst others may become lost (Wright, 1931; Kimura, 1954; Lande, 1976). Populations that cease to exchange genetic material—a break down in gene flow—may develop different frequencies of neutral genetic variants and, over time, become separate species (Savolainen *et al.*, 2006; Comes *et al.*, 2008; Sobel *et al.*, 2010). Such neutral evolutionary processes—termed "genetic drift"—

have been recorded in experimental (e.g. Buri, 1956), managed (e.g. Sugama *et al.*, 1988; Durand *et al.*, 1993), and wild populations (e.g. Hedgecock, 1994).

Protected areas need to safeguard the patterns of existing neutral genetic variation within species of conservation interest (Moritz, 1994). The rationale behind this is that the erosion of neutral genetic variation (i) represents a quantifiable loss in biodiversity (Frankel, 1974), (ii) increases the risk of inbreeding depression (Reed & Frankham, 2003), and (iii) reduces the raw material available for adaptation (Ellstrand & Elam, 1993). The main strategy for conserving intra-specific neutral genetic variation currently involves splitting species into sub-units that represent groups of populations that are evolving independently—termed evolutionarily significant units (ESUs; Moritz, 1994, 1999)—and protecting enough habitat for each unit (e.g. Vasconcelos *et al.*, 2012). For example, Ciofi *et al.* (1999) identified four distinct units for conservation in the Komodo dragon (*Varanus komodoensis*) using mitochondrial DNA. More recently, methodologies have been developed to directly maximize genetic diversity in prioritizations (Bonin *et al.*, 2007; Diniz-Filho *et al.*, 2012). In addition to preserving neutral genetic variation, reserves must also maintain gene flow among protected populations to foster evolutionary processes.

Conservation actions need to protect—and in some cases restore—gene flow across species' geographic distributions (Keyghobadi, 2007; Ellstrand, 2014). This is because gene flow plays a role in both adaptive and neutral evolutionary processes (reviewed in Slarkin, 1985; Kawecki, 2008). Although gene flow is generally thought to restrict adaptive processes (Tufto, 2001), it can facilitate adaptation in marginal populations. For instance, in North Wales, *Agrostis capillaris* plant populations in contaminated soil obtained zinc tolerance adaptations from migrants—possessing these adaptations through chance—that originated from genetically diverse populations (Al-Hiyaly *et al.*, 1993). Gene flow also plays a role in maintaining neutral genetic diversity (e.g. Hedrick, 1995). For instance, in southeastern California, neutral genetic diversity in desert bighorn sheep (*Ovis canadensis nelsoni*) plummeted following the construction of inter-state highways, canals, and urban areas that restricted gene flow between populations (Epps *et al.*, 2005). By siting protected areas in places that offer little impediment to gene flow (e.g. Creech *et al.*, 2014; Torrubia *et al.*, 2014; Krosby *et al.*, 2015), conservation actions can maintain gene flow and foster evolutionary processes.

Conservation planning must take evolutionary processes into account or risk failing to conserve biodiversity (Moritz, 2002). Here I have reviewed adaptation, genetic drift, and gene flow, and how the interaction between them affects species' persistence. Over the last few years, these ideas have only just started to penetrate systematic conservation planning (e.g. Bonin *et al.*, 2007; Diniz-Filho *et al.*, 2012; Carvalho *et al.*, 2017). By and large, evolutionary concepts and data are rarely used to inform conservation efforts (Hendry *et al.*, 2010). Given that the only way to achieve the long-term persistence of biodiversity—the stated goal of conservation—is to employ the knowledge gained in evolutionary biology, what is the

reason for this discrepancy?

1.1.3 Knowledge gaps

Conservation planning needs to account for evolutionary processes to achieve the long-term persistence of biodiversity (Moritz, 2002)—and yet most spatial prioritizations are not designed with this in mind. The main reasons for this stem from: (i) the lack of a decision support tool that can generate prioritizations to foster evolutionary processes, (ii) the lack of verified surrogate data that can be used to guide conservation planning when insufficient genetic data are available, and (iii) a general lack of knowledge surrounding how well existing protected areas are conserving evolutionary processes. These knowledge gaps need to be filled for conservation scientists, planners, and policy makers to critically consider the evolutionary consequences of conservation actions. Below, I discuss each of these issues in more detail.

One of the key issues in conservation planning is the development of prioritizations that conserve intra-specific variation. Conventionally, species have been split into evolutionarily significant units (ESUs; Moritz, 1994; Crandall *et al.*, 2000; Carvalho *et al.*, 2011b). In practice, however, this approach is difficult to apply because (i) this inevitably leads to a loss of information (Orians, 1993; Faith & Walker, 1996) and (ii) prioritizations are sensitive to the chosen number of units (Pressey & Logan, 1994). Both of these issues mean that a prioritization may conserve enough habitat for each unit but still fail to represent intra-specific variation (e.g. Moilanen *et al.*, 2013). Although other approaches have been proposed (e.g. Bonin *et al.*, 2007; Diniz-Filho *et al.*, 2012), they do not accommodate other criteria—such as acquisition cost, reserve fragmentation, and habitat targets—that are crucial for effective reserve design (Margules & Pressey, 2000). To fill this gap, in **Chapter 2**, I developed a decision support tool that can generate prioritizations targeting intra-specific variation while simultaneously accounting for cost, fragmentation, and habitat targets, without needing to split species into units. However, even when equipped with this tool, conservation planners still need genetic data to capture evolutionary processes—or do they?

Genetic data require substantial resources to obtain and analyze. Since conservation planning exercises often involve thousands of species, it is generally simply not feasible to obtain the genetic data required to map the evolutionary processes for each species of conservation interest. As a consequence, conservation planners desperately need cost-effective surrogates for genetic data to generate prioritizations that will capture spatial patterns of intra-specific genetic variation and gene flow across the landscape. Previous studies have investigated the potential use of environmental data and geographic distances as surrogates for capturing adaptive and neutral intra-specific genetic variation in prioritizations (Carvalho *et al.*, 2011b; Ponce-Reyes *et al.*, 2014). Additionally, based on the idea that areas with high quality habitat will offer little impediment to gene flow (reviewed in Zeller *et al.*, 2012),

other studies have explored the potential use of habitat suitability as a surrogate for gene flow (Wang *et al.*, 2008; Keeley *et al.*, 2016). However, the effectiveness of these surrogates in the context of guiding reserve selection remains largely untested. In **Chapters 3 and 4**, I examined the effectiveness of these surrogates using genetic data for multiple alpine plant species collected by the IntraBioDiv project (Taberlet *et al.*, 2012a,b). If broad-scale environmental data can be used to generate plans for protected areas that capture patterns of adaptive genetic variation, can such data also be used to understand how well existing protected areas are protecting the capacity for species to adapt to future threats?

As changes to the Earth's climate rapidly accelerate (Crowley, 2000), many species will need to adapt to new environmental conditions or face extinction (Hoffmann & Sgrò, 2011). As such, there has been increasing interest in identifying species that have a limited capacity to adapt to new conditions (Nicotra *et al.*, 2015; Breiner *et al.*, 2017), and identifying habitats that harbor evolutionary processes that will fortify species against such future threats (Bonin *et al.*, 2007; Carvalho *et al.*, 2011b; Sgro *et al.*, 2011). One strategy to maintain a species' adaptive capacity is to protect populations in a broad range of different habitats—a representative sample of its realized niche (Hutchinson, 1957)—which, in turn, will maximize the potential range of local adaptations within the species (Moritz, 1994, 2002). Based on this idea, in **Chapter 5**, I measured how well protected areas currently cover species' realized climatic niches globally. This chapter identifies species that need targeted conservation in specific habitats, and pinpoints habitats that will offer substantial benefits for many species.

1.2 Thesis objectives and structure

The main goal of this thesis is to understand how evolutionary processes can be incorporated into conservation planning to deliver more effective protected areas. I develop (**Chapter 2**) and evaluate (**Chapters 3 and 4**) methods for representing evolutionary processes in spatial prioritizations. I then provide insight into how well global protected areas may be representing species' adaptive evolutionary processes (**Chapter 5**). An outline of each chapter is presented below.

Chapter 1 has reviewed the major technical advances in conservation biology, and the key concepts that underpin evolutionary biology. This chapter uncovered gaps in the literature that prevent the widespread operationalization of evolutionary concepts in conservation planning. Finally, it sets the stage for the following chapters in this thesis.

Chapter 2 outlines a new method for representing intra-specific variation in conservation planning. This chapter provides a detailed mathematical formulation of the method. It also includes simulations and case-studies to explore the behavior of the method and shows how conservation planners can target selection pressures and existing patterns of genetic variation.

Chapter 3 explores the potential for environmental and geographic variables to serve as surrogates for adaptive and neutral genetic data during reserve selection. This chapter uses the method outlined in the previous chapter to build and assess plans for protected areas using surrogate and genetic data. This chapter provides insight into how freely available data could be used to conserve species' evolutionary processes and increase the chances for their long-term persistence.

Chapter 4 investigates strategies commonly used to increase connectivity in prioritizations. This chapter examines whether prioritizations generated using fragmentation penalties and constraints to ensure that species are represented in contiguous areas of suitable habitat can actually maintain gene flow. This chapter provides further insight into how genetic data can be used to guide the reserve selection process and ultimately plan for species' long-term persistence.

Chapter 5 assesses how well protected areas represent the range of climatic selection pressures that are currently shaping species' evolutionary trajectories. This chapter involves a global analysis comprising most of the world's amphibian, avian, and terrestrial mammalian species. This chapter identifies species that need targeted conservation interventions to maintain their capacity to adapt to new environmental conditions, and highlights places that yield substantial benefits for achieving this.

Chapter 6 integrates the key findings in this thesis. It contextualizes these discoveries amongst the broader scientific literature, considers the limitations of the approaches taken in this thesis, and discusses directions for future research.

CHAPTER 2

raptr: Representative and adequate prioritization toolkit in R

Protected area systems need to secure an adequate amount of habitat for each species. They also need to secure a representative sample of each species—for instance, in terms of their existing patterns of genetic variation—to maximize their long-term persistence. This chapter outlines a new reserve selection method for generating plans for protected areas that fulfill both of these objectives. Furthermore, it includes a simulation study and two case-studies to explore the behavior of the new method. This chapter shows that conservation planners must explicitly target representativeness to generate plans that will capture it.

A version of this chapter has been published in *Methods in Ecology and Evolution* as:

Hanson JO, Rhodes JR, Possingham HP & Fuller RA (2018) raptr: Representative and adequate prioritization toolkit in R. *Methods in Ecology & Evolution*, **9**, 320–330.

2.1 Abstract

An underlying aim in conservation planning is to maximize the long-term persistence of biodiversity. To fulfill this aim, the ecological and evolutionary processes that sustain biodiversity must be preserved. One way to conserve such processes at the feature level (e.g. species, ecosystem) is to preserve a sample of the feature (e.g. individuals, areas) that is representative of the intrinsic or extrinsic physical attributes that underpin the process of interest. For example, by conserving a sample of populations with local adaptations—physical attributes associated with adaptation—that is representative of the range of adaptations found in the species, protected areas can maintain adaptive processes by ensuring these adaptations are not lost. Despite this, current reserve selection methods overwhelmingly focus on securing an adequate amount of area or habitat for each feature. Little attention has been directed towards capturing a representative sample of the variation within each feature. To address this issue, we developed the raptr R package to help guide reserve selection. Users set "amount targets"—similar to conventional methods—to ensure that solutions secure a sufficient proportion of area or habitat for each feature. Additionally, users set "space targets" to secure a representative sample of variation in ecologically or evolutionarily relevant attributes (e.g. environmental or genetic variation). We demonstrate the functionality of this package using simulations and two case studies. In these studies, we generated solutions using amount targets—similar to conventional methods—and compared them with solutions generated using amount and space targets. Our results demonstrate that markedly different solutions emerge when targeting a representative sample of each feature. We show that using these targets is important for features that have multimodal distributions in the process-related attributes (e.g. species with multimodal niches). We also found that solutions could conserve a far more representative sample with only a slight increase in reserve system size. The raptr R package provides a toolkit for making prioritizations that secure an adequate and representative sample of variation within each feature. By using solutions that secure a representative sample of each feature, prioritizations may have a greater chance of achieving long-term biodiversity persistence.

2.2 Introduction

Perhaps the most fundamental aim of conservation is to maximize the long-term persistence of biodiversity (McNeely, 1994; Margules & Pressey, 2000). To achieve this, conservation actions must preserve biodiversity patterns (e.g. populations, species, ecosystems), but also crucially the processes that sustain them. One of the major tangible achievements of modern conservation has been the act of setting aside areas for preservation (McNeely, 1994; Sanderson *et al.*, 2015). Reserve networks buffer species from gross threatening processes and set the stage for enhanced management interventions (Gaston *et al.*, 2008). Since the resources available for conservation action are limited, protected area networks must be sited in places that satisfy conservation objectives for minimal cost (Margules & Pressey, 2000).

Today, the most widely used conservation planning tools focus on biodiversity patterns (`Zonation` and `Marxan`; Moilanen, 2007; Ball *et al.*, 2009). Decision makers can use these tools to obtain solutions that secure a proportion of the geographic range of each biological feature (populations, species, or ecosystems) of interest, by setting targets. One of the main strategies for incorporating ecologically and evolutionary processes into the these tools has involved identifying ecologically or evolutionarily significant areas as a pre-processing step (e.g. climate refugia, habitat corridors, or places with steep environmental gradients; Rouget *et al.*, 2003; Klein *et al.*, 2009a) and ensuring that they are conserved in solutions. However, a more effective approach is to conserve a representative sample of the internal or external physical attributes that underpin these processes for each feature to ensure that these processes are maintained for each feature (Carvalho *et al.*, 2011b). Current methods for achieving this second approach involve partitioning features into sub-groups based on an attribute variable that relates to a biodiversity process of interest (Klein *et al.*, 2009b; Beger *et al.*, 2014) or phylogenetic trees (Carvalho *et al.*, 2017). For instance, by dividing species' distributions into sub-groups according to habitat discontinuities and ensuring that each sub-group is conserved in the solution, conservation planners can obtain prioritizations that promote adaptive processes (Carvalho *et al.*, 2011b). However, using these methods is challenging because biodiversity often cannot be divided into operational groups without information loss (Orians, 1993; Faith & Walker, 1996), and because the number of groups can substantially alter solutions (Pressey & Logan, 1994). This limitation has been known for quite some time, and dates back to some of the earliest reserve selection methods (e.g. Kirkpatrick, 1983).

To overcome this limitation, Faith and Walker (1994) developed the "environmental diversity" (ED) reserve selection framework to conserve a representative sample of the environmental variation found across a study area. Drawing inspiration from the *p*-median problem (Faith & Walker, 1993), this method used Euclidean distances to express differences among candidate sites and maximize the representativeness of the solution (Faith & Walker, 1996; Faith,

2003; Faith *et al.*, 2004). Recent work has built on this method using more advanced optimization algorithms (Engelbrecht *et al.*, 2016). However, existing environmental diversity based methods are not designed to work with multiple features (Faith & Walker, 1996). As a consequence, these methods cannot identify cost-effective solutions that meet specific criteria (e.g. habitat targets) for multiple features (c.f. Marxan), and so cannot be applied to many real-world conservation problems (Margules & Pressey, 2000).

Conservation planners lack a decision support tool that lets them set explicit targets to obtain solutions that secure (i) an adequate amount of habitat for each feature and (ii) a representative sample of the variation in each feature. To begin to fill this gap, we unite the ideas underpinning environmental diversity and Marxan into new formulations of the reserve selection problem, and implement them in the raptr R package. This R package provides decision makers with the tools to generate prioritizations based on data that relate to biodiversity patterns and processes. Here, we aim to provide an in-depth understanding of this R package and explore its functionality.

2.3 Methods

2.3.1 Problem formulation

Biodiversity features are defined as the entities that the prioritization is required to preserve (e.g. species, ecosystems). Each biodiversity feature may have one or more "attributes" which vary across its geographic range as a result of ecological or evolutionary processes. These attributes can be intrinsic (e.g. genetic or phenotypic) or extrinsic (e.g. environmental conditions) to the feature. The overarching goal of the prioritization is to conserve these biodiversity processes, and by capturing the variation in these species-specific attributes, the prioritization can make some progress towards achieving this. Thus there should be a reasonable underlying hypothesis that relates the attributes to the biodiversity processes that need to be conserved.

The variation in the attributes of a feature can be described using an *n*-dimensional "attribute space". For example, a decision maker may require a prioritization that captures populations along climatic gradients. To achieve this, the decision maker might use a "climatic" attribute space with dimensions relating to mean annual temperature (°C) and precipitation (mm). Any given combination of temperature and precipitation may be conceived as a point in this climatic space. Although they exist as polygons, for simplicity, each planning unit may be thought to exist as a single point inside a given attribute space. By associating the planning units with climatic data, and calculating a descriptive statistic for each planning unit (e.g. mean), they can be mapped from geographic space to this climatic attribute space.

"Demand points" (Faith & Walker, 1993, 1996; Faith, 2003) are designated by the decision maker to indicate regions of the attribute space that they wish to conserve in the prioritization (see below for discussion on generating demand points for real-world data sets). For instance, by siting demand points throughout an attribute space, planners can obtain solutions that better capture existing variation. Alternatively, by siting demand points in specific regions of an attribute space, planners can obtain solutions that target specific samples of an attribute space. For a given set of demand points, the shorter the sum of the distances between the demand points and the planning units selected for prioritization: the better a solution is at securing the desired variation in the attribute space. Demand points are used heavily in the *p*-median and facility location problems (reviewed in Owen & Daskin, 1998). Additionally, different features and attribute spaces may require different demand points. For instance, values that might be valid in one attribute space (e.g. mean annual temperature -5°C) may be invalid for another attribute space (e.g. mean annual rainfall -5 mm). Additionally, there may be some regions that are desirable for some features and undesirable for others (e.g. conditions known to be outside the physiological tolerance of certain species).

To illustrate these concepts, consider the following example: we wish to develop a prioritization for a single species that has four populations. Since we can only afford to preserve three of the four populations, our objective is to conserve the most representative sample possible. To achieve this goal, we obtained annual rainfall (mm) and temperature (°C) data at the location of each population. We used this data to construct a two-dimensional climatic attribute space. Next, we generated demand points as equidistant points inside this space. By comparing the distribution of the demand points to the distribution of the populations in the attribute space, we can identify the optimal solution (Figure 2.1). We can see that a solution that prioritizes both populations *A* and *B* constitutes considerable redundancy by selecting populations that inhabit similar conditions. Instead, a more representative sample of the intra-specific variation could be captured by securing populations *A* (or *B*), *C*, and *D*. However, if the goal of the prioritization was to preserve populations living in warmer temperatures, then instead of siting demand points across the full range of conditions, we could site demand points in environmental conditions with temperatures over 30 °C (i.e. the top two rows of demand points in Figure 2.1). Given this new set of demand points, populations *A*, *B*, and *C* would be prioritized. Since demand points can be sited and weighted in any configuration, they provide a flexible means to guide the reserve selection process.

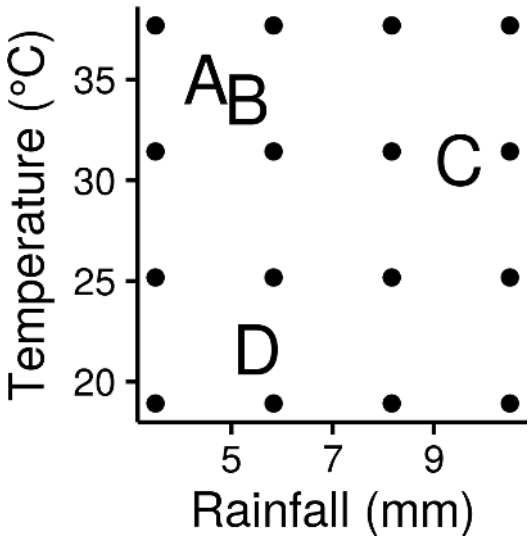


Figure 2.1 Attribute space example. This environmental attribute space has dimensions relating to annual temperature (°C) and rainfall (mm). Letters denote the environmental conditions associated with the geographic locations where four hypothetical populations are found. Points denote demand points. In this space, populations close to each other inhabit similar environmental conditions.

The raptr R package utilizes two novel formulations of the reserve selection problem to generate prioritizations. These formulations are based on ideas that underpin the Marxan, environmental diversity, and uncapacitated facility location problems (Owen & Daskin, 1998). Since they are based on the unreliable and reliable facility location problems (Cui *et al.*, 2010), the formulations are hereafter referred to as the "unreliable" and "reliable" formulations. The difference between the two formulations is that the reliable formulation explicitly accommodates uncertainty in the probability of features occupying the planning units when calculating how well a given solution samples a feature's attribute space. For brevity, we will define the simpler formulation—the unreliable formulation—below and define the more complex version—the reliable formulation—in Appendix A. All mathematical terms defined hereafter are described in Appendix B. For convenience, the cardinality of a given set will be denoted using the same symbol used to denote the set.

Define F to be the set of features one wishes to conserve (indexed by f). Let J be a set of planning units (indexed by j), and C_j denote the cost of preserving planning unit $j \in J$. Also, let A_j denote the area—or some other metric such as habitat quality or amount of habitat—associated with planning unit j. To assess the extent to which each feature is secured in a given prioritization, let q_{fj} denote the probability of feature f occupying planning unit j. The level of fragmentation associated with a prioritization is parameterized as the total exposed boundary length (as in Marxan). Let the shared edges between each planning unit $j \in J$ and $k \in J$ be e_{jk}.

Let S denote a set of attribute spaces (indexed by s). Each $j \in J$ is associated with spatially explicit data that serve as coordinates in each attribute space $s \in S$. Let I_{fsi} denote a set of demand points (indexed by i) for each feature $f \in F$ and each attribute space $s \in S$. Let

λ_{fsi} denote the weighting for each demand point $i \in I$, $f \in F$ and $s \in S$. Let d_{fsij} denote the distance between each demand point $i \in I$ and each planning unit $j \in J$ for each feature $f \in F$ and attribute space $s \in S$. To describe the inherent variation in the distribution of demand points for feature f and space s, let δ_{fsi} denote the distance between each demand point $i \in I$ and the centroid of the demand points in space s.

Demand points with greater weight λ_{fsi} are more important, and so solutions may need to select planning units closer to highly weighted demand points. Since increasing the number of demand points in the problem also increases the time required to solve it, conservation planners may find it more useful to increase the weighting of demand points in important regions of an attribute space rather than increasing the number of demand points in the region. Decision makers will need to choose an appropriate weighting for each demand point to ensure that the resulting solution reflects their overarching goals.

Targets are used to ensure that prioritizations sufficiently conserve each feature. Amount-based targets specify the minimum amount of habitat required for each feature to be adequately conserved (similar to those used in `Marxan`). Let T_f denote the area or amount of habitat that needs to be preserved for each feature $f \in F$. Space-based targets specify the minimum proportion of variation in the demand points that needs to be secured for each feature. These targets directly relate to a continuous and multi-dimensional attribute space and the set of demand points within it. Space-based targets are expressed as proportions—instead of a sum of weighted distances—by scaling the sum of weighted distances between the demand points and the selected planning units in a solution relative to the distances between demand points and the demand points' centroid. This scaling is conceptually similar to that used in calculating the R^2 statistic for *k*-means cluster analyses from the within sums of squares and total sums of squares (Greenacre & Primicerio, 2014, page 106). Let τ_{fs} denote the space-based targets for feature $f \in F$ and attribute space $s \in S$. The control variables for the unreliable formulation are B, T_s, and τ_{fs}.

$T_s =$ amount target for feature f eqn 2.1a

$\tau_{fs} =$ space target for feature f in attribute space s eqn 2.1b

$B =$ boundary length modifier (BLM) to penalize fragmented solutions eqn 2.1c

The decision variables are X_j and Y_{fsij}.

$$X_j = \begin{cases} 1, & \text{if planning unit } j \text{ is selected for conservation action} \\ 0, & \text{otherwise} \end{cases} \quad \text{eqn 2.2a}$$

$$Y_{fsij} = \begin{cases} 1, & \text{if demand point } i \text{ for feature } f \text{ in space } s \text{ is conserved by planning unit } j. \\ 0, & \text{otherwise} \end{cases} \quad \text{eqn 2.2b}$$

Each demand point $i \in I$ for feature $f \in F$ and space $s \in S$ is conserved by a single selected planning unit (i.e. a $j \in J$ where $X_j = 1$). The degree to which a demand point i is conserved by a planning unit j is determined by the distance between them (d_{fsij}). Generally—unless near zero space targets are used so that the problem is effectively unconstrained by the target—demand points are conserved by their closest selected planning units. In poorer quality solutions, demand points will be conserved by planning units that are further away from them. As a consequence, the sum of the weighted distances between all of the demand points and the planning units used to conserve them will be larger, and so, the solution will capture less of the variation described by the demand points.

The unreliable formulation of the representative and adequate prioritization problem (URAP) is defined as a multi-objective optimization problem.

$$
\begin{aligned}
&\text{(URAP)} \quad \text{Min} \sum_{j=0}^{J-1} \left(X_j C_j\right) + \sum_{j=0}^{J-1} \sum_{k=j}^{J-1} X_j \left(1 - X_k\right) \left(Be_{jk}\right) && && \text{eqn 2.3a}\\
&\quad \text{s.t.} \sum_{j=0}^{J-1} A_j q_{fj} \geq T_f && \forall 0 \leq f \leq F-1 && \text{eqn 2.3b}\\
&\quad 1 - \frac{\sum_{i=0}^{I-1} \sum_{j=0}^{J-1} \lambda_{fsi} {d_{fsij}}^2 Y_{fsij}}{\sum_{i=0}^{I-1} \lambda_{fsi} {\delta_{fsi}}^2} \geq \tau_{fs} && \forall 0 \leq f \leq F-1, && \text{eqn 2.3c}\\
&\qquad && 0 \leq s \leq S-1 &&\\
&\quad \sum_{j=0}^{J-1} Y_{fsij} = 1 && \forall 0 \leq f \leq F-1, && \text{eqn 2.3d}\\
&\qquad && 0 \leq s \leq S-1, &&\\
&\qquad && 0 \leq i \leq I-1 &&\\
&\quad Y_{fsij} \leq X_j && \forall 0 \leq f \leq F-1, && \text{eqn 2.3e}\\
&\qquad && 0 \leq s \leq S-1, &&\\
&\qquad && 0 \leq i \leq I-1, &&\\
&\qquad && 0 \leq j \leq J-1 &&\\
&\quad X_j, Y_{fsij} \in 0, 1 && \forall 0 \leq f \leq F-1, && \text{eqn 2.3f}\\
&\qquad && 0 \leq s \leq S-1, &&\\
&\qquad && 0 \leq i \leq I-1 &&
\end{aligned}
$$

The objective (eqn 2.3a) is to minimize the total cost and fragmentation of the solution. Constraints ensure that all amount-based targets are met (eqn 2.3b), and all space-based targets are met for each feature and each attribute space (eqn 2.3c). By treating targets as constraints, rather than including them in the objective function, all feasible solutions will fulfill the targets. For each feature and attribute space, the total weighted distance between the demand points and their closest selected planning units is calculated ($\sum_{i=0}^{I-1} \sum_{j=0}^{J-1} \lambda_{fsi} {d_{fsij}}^2 Y_{fsij}$). This total weighted distance is then scaled by the inherent variation in the demand points ($\sum_{i=0}^{I-1} \lambda_{fsi} {\delta_{fsi}}^2$). The resulting fraction is used to calculate a proportion conceptually similar to the R^2 statistic used in *k*-means cluster analysis. Constraints (eqn 2.3d) ensure that only one planning unit is assigned to each demand point. Constraints (eqn 2.3e) ensure that demand points are only assigned to selected planning units. Constraints (eqn 2.3f) ensure that the X and Y variables are binary.

2.3.2 Optimization

Although the reserve selection problems presented here are non-linear (see Appendix A for the reliable formulation), they can be linearized using methods described by Beyer *et al.* (2016) and Cui *et al.* (2010). The raptr R package provides functions to express conservation planning data as linearized versions of the optimization problems and solve them using the commercial Gurobi software suite (https://gurobi.com). Presently, academics can obtain a Gurobi license at no cost.

2.4 Examples

To showcase the behavior of the unreliable formulation, we conducted a simulation study and two case studies. To assess how long it would take to solve various sized problems, we also conducted a benchmark analysis (Appendix C). We completed the analyses using R (version 3.3.2; R Core Team, 2016) and raptr (version 0.0.3). We solved all optimization problems to within 10 % of optimality using Gurobi (version 7.0.2).

2.4.1 Simulation study

2.4.1.1 Methods

We simulated a hypothetical study area with square planning units arranged in a 10×10 grid (Figure 2.2), and three species inhabiting this study area. Firstly, we simulated a hypergeneralist species (hereafter referred to as the "uniformly distributed species"). It occupied all planning units with equal probability (Figure 2.2a; eqn 2.4a). Secondly, we simulated a species with simple habitat requirements (hereafter referred to as the "normally distributed species"; Figure 2.2b). This species was most likely to be found in planning units nearest to the center of the study area. It was simulated using the density function of a multivariate normal distribution (denoted by $\mathcal{N}$; eqn 2.4b). Thirdly, we simulated a species with two distinct populations (hereafter referred to as the "bimodally distributed species"; Figure 2.2c). It was simulated using the maximum density of two multivariate normal distributions (eqn 2.4c). For a given species, planning unit occupancy was calculated using the (X, Y) coordinates of the units' centroids and the relevant equation. We used a geographic attribute space to provide an intuitive visualization of the solutions. Demand points were set as the planning units' centroids and were weighted according to the units' probability of occupancy.

$$P\left(\text{uniformly distributed species}|\left(X,Y\right)\right)=0.5 \quad \text{eqn 2.4a}$$

$$P\left(\text{normally distributed species}|\left(X,Y\right)\right)=\frac{\mathcal{N}\left(\left[\begin{smallmatrix}x\\y\end{smallmatrix}\right]\left[\begin{smallmatrix}0.0\\0.0\end{smallmatrix}\right]\left[\begin{smallmatrix}12.58 & 0\\0 & 12.5\end{smallmatrix}\right]\right)}{2} \quad \text{eqn 2.4b}$$

$$P\left(\text{bimodally distributed species}|\left(X,Y\right)\right)=\text{Max}\begin{Bmatrix}\mathcal{N}\left(\left[\begin{smallmatrix}x\\y\end{smallmatrix}\right]\left[\begin{smallmatrix}-3.75\\-3.75\end{smallmatrix}\right]\left[\begin{smallmatrix}10 & 0\\0 & 10\end{smallmatrix}\right]\right),\\ \frac{\mathcal{N}\left(\left[\begin{smallmatrix}x\\y\end{smallmatrix}\right]\left[\begin{smallmatrix}3.75\\3.75\end{smallmatrix}\right]\left[\begin{smallmatrix}8 & 0\\0 & 8\end{smallmatrix}\right]\right)}{2}\end{Bmatrix} \quad \text{eqn 2.4c}$$

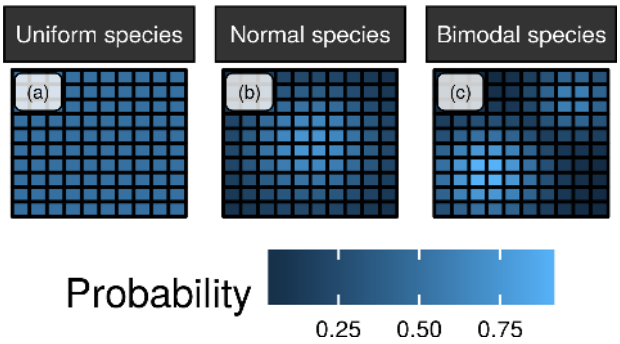


Figure 2.2 Distributions of three simulated species. Squares denote planning units. Colors indicate probability of occupancy.

We generated four solutions for each species. First, to portray solutions generated using conventional methods (e.g. `Marxan`), we generated solutions using 20 % amount targets. Second, to show how the addition of space targets can affect solutions, we generated solutions using 20 % amount targets and 90 % space targets to capture the geographic spread of each species. Third, to portray solutions generated using conventional planning methods that penalize for fragmentation, we generated solutions using 20 % amount targets and a boundary length modifier of 2.5. Finally, to illustrate the combined effects of using amount and space targets as well as fragmentation penalties, we generated solutions using 20 % amount targets, 90 % space targets, and boundary length modifiers of 2.5.

2.4.1.2 Uniformly distributed species

The solution generated for the uniformly distributed species using an amount target prioritized 20 planning units (Figure 2.3a). For this scenario, all solutions containing 20 units are optimal because this species has an equal chance of occurring in any given planning unit. Although it may seem odd that the solution prioritized planning units along the southern end of the study area, this behavior occurred because there were many optimal solutions and the solver—without any other criteria to compare solutions—returned a solution that reflected the order that data were encoded in the optimization problem. As we can see in later examples, this behavior only manifests when there are many optimal solutions. It is unlikely that conservation planners would encounter this behavior when using real-world data, and if they did, it would suggest that they need to obtain more data to inform the reserve selection

process.

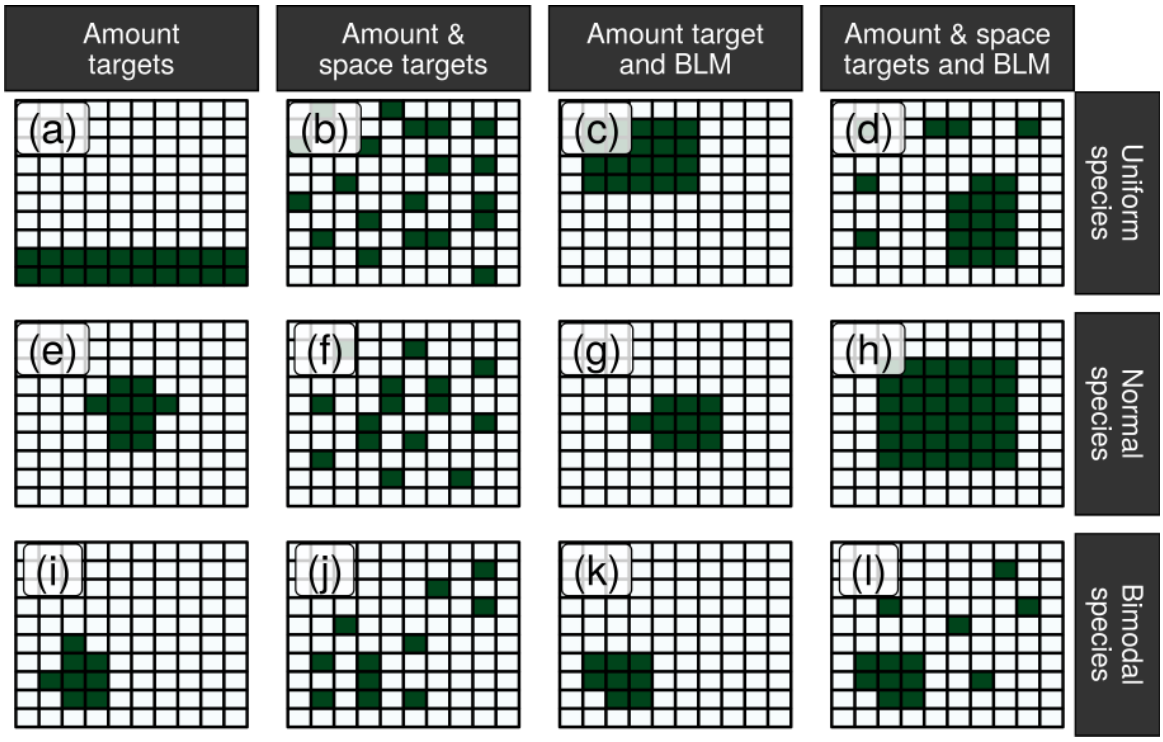


Figure 2.3 Prioritizations for the simulation study. Each panel shows a prioritization generated for a single species using a set of parameters. Squares denote planning units. Dark green planning units were selected for protection. Each row of panels show prioritizations generated for a different species. Each column of panels corresponds to a different set of parameters used to generate the prioritization.

The solution generated using just amount targets only captured a small proportion of the variation in the geographic attribute space (-23.64 % sampled). The coverage was so poor that this "proportion" was negative because the solution captured a less representative sample than a solution containing one planning unit in the center of the species' distribution. In other words, the total distance between the demand points and the 20 selected planning units was larger than the total distance between the demand points and a single planning unit in the center of the study area. Similarly, R^2 statistics can be negative when they are describing models that perform worse than a null model. This solution performed worse than the solution generated using amount targets and a boundary length modifier to reduce fragmentation (47.27 % sampled; Figure 2.3c) because the latter solution sited more planning units closer to center of the study area, reducing the total distance between the selected planning units and the demand points.

By explicitly targeting a representative sample of the species' geographic spread, we obtained a solution that captured the geographic spread of the uniform species (93.39 % sampled; Figure 2.3b). Although this solution secured an adequate amount of habitat and a representative sample of the species' geographic spread, this solution was highly fragmented. However, by penalizing fragmentation using a boundary length modifier, we were able to

obtain a well connected solution that met all of the objectives (90 % sampled; Figure 2.3d). This solution prioritized a similar number of planning units as the previous solutions even though it is far superior. As we can see, under the simplest of circumstances, reserve selection methods may not yield solutions that secure a representative sample of features unless constraints are used to guarantee this property.

2.4.1.3 Normally distributed species

The solution generated for the normally distributed species using just an amount target prioritized planning units in the species' core distribution (Figure 2.3e). Although this strategy may seem cost-effective, simply conserving the places where this species is most likely to be found is a poor strategy for securing a representative sample of the species' geographic spread (59.09 % sampled) because it did not protect any peripheral habitat. By using amount and space targets, we obtained a solution that conserved an adequate amount of habitat and also secured a representative sample of the species' geographic spread (90.02 % sampled; Figure 2.3f). Similar to the solutions for the uniformly distributed species, this solution was highly fragmented and we were able to obtain a better connected solution by specifying fragmentation penalties (90.88 % sampled; Figure 2.3h). However, unlike the solutions for the uniformly distributed species, the solution generated using an amount target, a space target, and fragmentation penalties required more planning units than the other solutions. These results suggest that solutions may need to select more planning units to meet additional conservation objectives.

2.4.1.4 Bimodally distributed species

The solution generated for the bimodally distributed species using just an amount target conserved individuals belonging to one of the two populations (Figure 2.3i). As a consequence, this solution did not secure a representative sample of the species' geographic spread (12 % sampled). The addition of fragmentation penalties exacerbated this issue, and resulted in a solution that sampled even less of the species' geographic spread (8.38 % sampled; Figure 2.3k). However, this issue was resolved by using a space target to obtain a solution that secured both populations (90.23 % sampled; Figure 2.3l). This finding suggests that species with large intra-specific variation could benefit the most from prioritizations generated using space-based targets.

2.4.2 Case study 1

2.4.2.1 Methods

We investigated how space-based targets can be used in a multi-species planning context to generate a prioritization that sufficiently preserves the species' realized niches. By preserving the populations in suitable habitats with different environmental conditions, conservation planners can preserve the species' adaptive landscape and foster resilience to environmental change (Moritz, 2002). We selected Queensland, Australia as the study area, and used an equal area coordinate system for geospatial analyses (Australian Albers GDA94; EPSG:3577). We used a 50 $\times$ 50 km grid within the state boundary as planning units for this case study. We obtained data for 19 bioclimatic variables across the region (at $30''$ resolution from http://worldclim.org; Hijmans *et al.*, 2005) and subjected them to a principal components analysis (using ArcMap 10.3.1). We used scores from the first two principal components to characterize the environmental variation across the study area (explaining 99.5 % of the total variation; Figure 2.4).

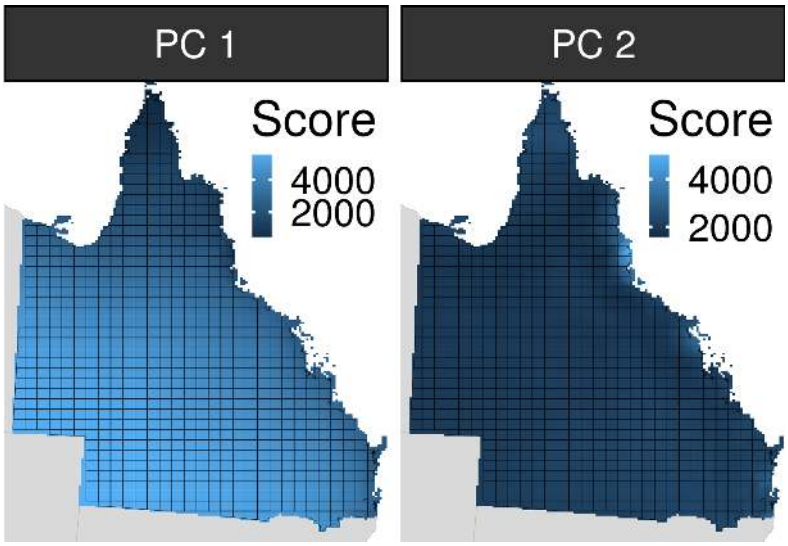


Figure 2.4 Two main gradients of climatic variation across Queensland, Australia. Polygons denote planning units. The map is rendered in an equal-area coordinate system (EPSG:3577).

We used four bird species in this case study: blue-winged kookaburra (*Dacelo leachii*), brown-backed honeyeater (*Ramsayornis modestus*), brown falcon (*Falco berigora*), and pale-headed rosella (*Platycercus adscitus*). These species span a range of different habitat requirements. We mapped the extent of occurrence for each species (Figure 2.5). To do this, we obtained occurrence records from the Atlas of Living Australia across the whole of Australia (using the ALA4R R package; Raymond *et al.*, 2015), spatially thinned the data to omit points within 10 km of each other to ameliorate the effects of sampling bias (using a modified version of the spThin R package; https://github.com/jeffreyhanson/spThin; Aiello-Lammens *et al.*, 2015), and fit 85 % minimum convex polygons (using the adehabitatHR R package; Calenge, 2006). We used this method because it is entirely reproducible using

freely available data.

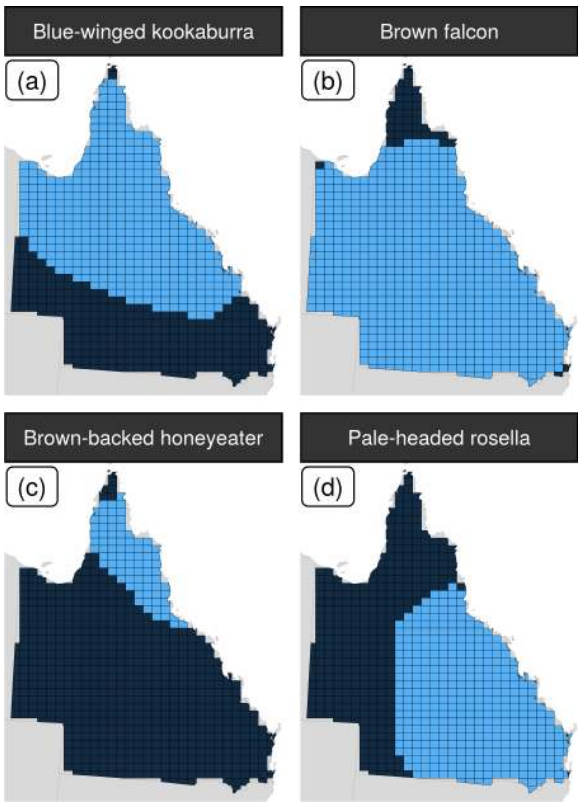


Figure 2.5 Distribution of the species used in the first case study. See Figure 2.4 caption for conventions. Planning units occupied by a given species are shown in light blue.

We generated 500 demand points for each species (Appendix D). To ensure that the demand points reflected the core parts of the species' realized niches, we used the following method: for each species we generated random points inside the species' geographic range and at each point extracted the principal component values at that location. We then fitted hyperbox kernels to the distribution of principal component values to characterize the realized niche of each species using a manually chosen bandwidth of 0.2 and a 0.5 quantile to map the core parts of the species' niches (implemented in the `hypervolume R` package; Blonder *et al.*, 2014). We then generated uniformly distributed points inside the species' distribution in environmental space, and extracted the kernel density at their locations. These uniformly distributed points and associated density estimations were used as demand point coordinates and weights (respectively).

We generated two multi-species prioritizations. The first solution was generated using 20 % amount targets for each species. The second solution was generated using the same amount targets with additional 75 % space targets to capture a representative sample of the realized niche for each species.

2.4.2.2 Results

Generally, the solution generated using just amount targets preserved a representative sample of each the four bird species' niches (Figure 2.6a; left column Appendix D). This solution secured a large proportion of the realized niche of blue-winged kookaburra (90.18 %), brown falcon (92 %), and the pale-headed rosella (84.76 %), but only a small proportion for the brown-backed honeyeater (29.09 %). On the other hand, the solution generated using space targets secured a large proportion of the realized niche for all four of the species (Figure 2.6b; right column Appendix D). This result demonstrates that although conventional methods may yield solutions that conserve a representative sample of the variation in some features, only through the use of explicit targets can planners obtain cost-effective solutions that secure a representative sample of the variation in all features.

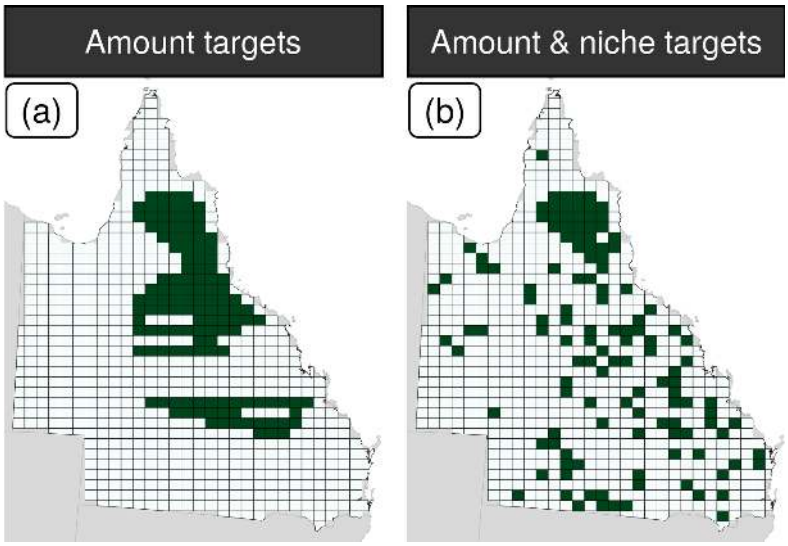


Figure 2.6 Prioritizations for the first case study. Polygons denote planning units. Dark green planning units were selected for protection. Panel (a) shows the solution generated when using 20 % amount targets. Panel (b) shows the solution when using 20 % amount targets and 75 % space targets.

2.4.3 Case study 2

2.4.3.1 Methods

Here we used space-based targets to generate a prioritization securing a representative sample of a species' intra-specific genetic variation. We used species occurrence and genetic data collected by the international IntraBioDiv project in the European Alps (Taberlet *et al.*, 2012a,b). Although this dataset contains multiple species, we used data for the betony-leaved rampion (*Phyteuma betonicifolium*) as a simple case study, because it exhibits significant inter-population genetic structure (Meirmans *et al.*, 2011). Members of the IntraBioDiv project collected data using a $20'$ longitude by $21'$ latitude grid (approx. 22.3 km $\times$ 25 km;

Figure 2.7a). They visited every second grid cell, and if the species was detected in a cell, samples were collected from three individuals. Samples were genotyped using amplified fragment length polymorphisms (Vos *et al.*, 1995), and used to construct matrices denoting the presence/absence of polymorphisms at loci. In total, 131 individuals were genotyped at 138 markers.

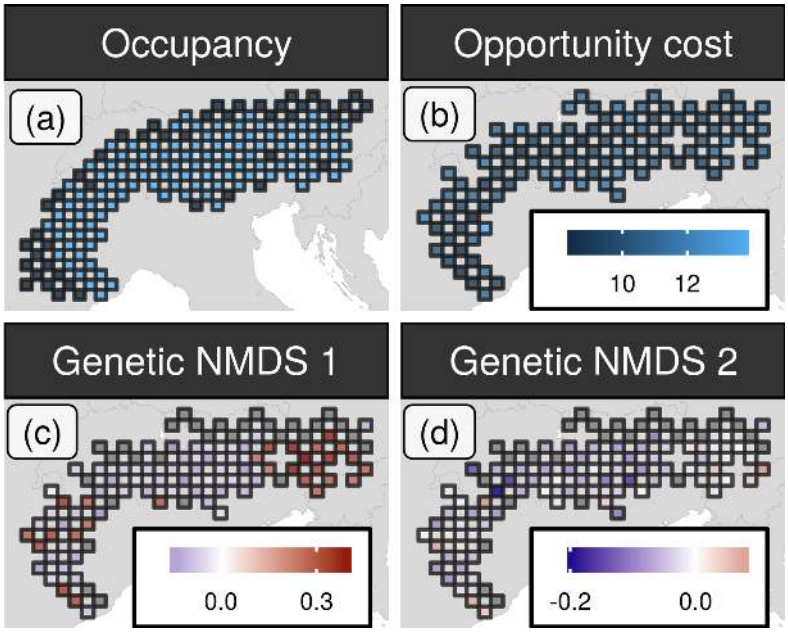


Figure 2.7 Data used for the second case study. Squares denote planning units. Panel (a) shows all grid cells surveyed by the IntraBioDiv project. Grid cells occupied by the betony-leaved rampion are shown in bright blue. The subsequent panels only show occupied grid cells. Panel (b) shows the acquisition cost of each planning unit (estimated as the total human population density). Panels (c–d) show the spatial distribution of the ordinations describing genetic variation. These values describe the typical genetic characteristics of individuals in each planning unit. Planning units with similar values/colors contain individuals with similar loci polymorphisms. Note that data were not collected in every grid cell, and the planning units are therefore arranged in a checkerboard pattern.

We used non-metric multi-dimensional scaling (NMDS; using Gower distances to accommodate sparsity implemented using the `cluster R` package; Gower, 1971; Maechler *et al.*, 2015) to ordinate the presence (or absence) of locus-specific alleles within individuals into two continuous variables (implemented in the `vegan R` package; Oksanen *et al.*, 2015). These continuous variables described the main axes of genetic variation within the species. We averaged together values corresponding to samples collected in the same grid cell, and used them to create a genetic attribute space (Figure 2.7c and Figure 2.7d). To assess spatial auto-correlation, we calculated Moran's I auto-correlation index for each NMDS axis using inverse great circle distances between the grid cells' centroids as weights (using the `ape R` package; Paradis *et al.*, 2004).

The grid cells were used as planning units for generating prioritizations, noting that data were not collected in every planning unit, and so the planning units were arranged in a checkerboard pattern. The spatially averaged ordinations were used to describe the typical

genetic characteristics of individuals in each planning unit. To ensure that solutions did not prioritize particularly costly areas, we obtained population density data (1 km resolution from the Global Rural-Urban Mapping Project; GRUMP V1; CIESEN, Columbia University *et al.*, 2011) and estimated the total population density inside each grid cell. We used these values to denote opportunity cost (Figure 2.7b).

We generated two solutions. The first solution was generated using a 10 % amount-based target. The second solution was generated using the same amount-based target and an additional 95 % space target to conserve genetic variation.

2.4.3.2 Results

A two-dimensional genetic attribute space was able to capture most of the variation between individuals (stress = 0.17; Figure 2.7c and Figure 2.7d). Planning units located near each other contained individuals with similar genetic characteristics (Moran's I: NMDS axis 1, $I = 0.4$, $P < 0.001$; Moran's I: NMDS axis 2, $I = 0.32$, $P < 0.001$). In terms of the average genetic characteristics of individuals in the planning units, they tended to cluster into two main groups, with evidence of within-group structure inside the larger group (Fig 8c). This analysis supports previous work by Alvarez *et al.* (2009) who also found evidence of genetic structure within this species.

The solution generated using just the amount target failed to preserve a representative portion of the species' genetic variation (15.52 % sampled; Figure 2.8 8a and Figure 2.8c). This solution only conserved individuals in one of the two main genetic groups. In fact, this solution is just a collection of the cheapest planning units occupied by the species needed to fulfill the target. Alternatively, the solution generated using both amount and space targets conserved individuals in each of the two main genetic groups (95.2 % sampled; Figure 2.8b and Figure 2.8d). Although the solutions only differ by a single planning unit, by securing individuals in both main genetic groups the second solution was able to secure a much more representative sample of the species' genetic variation for only a minor increase in cost (99.66 compared to 99.41 total cost).

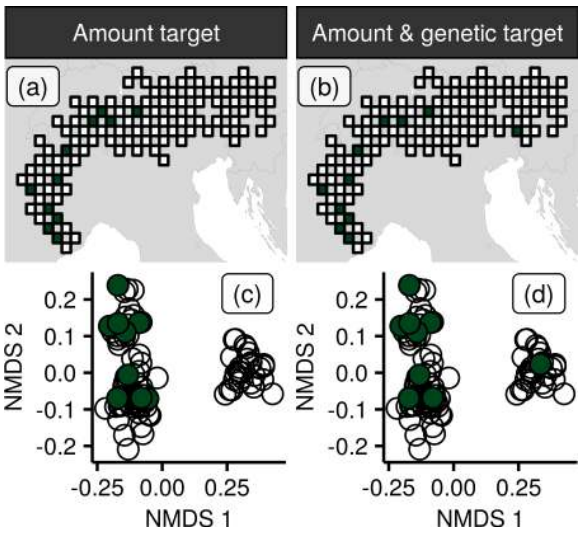


Figure 2.8 Prioritizations for the second case study. Panels (a–b) show prioritizations generated using different parameters. Polygons denote planning units. Dark green planning units were selected for protection. Panel (a) shows the planning units selected when using 10 % amount targets. Panel (b) shows the planning units selected when using 10 % amount targets and 95 % space targets. Panels (c–d) show the solutions in the genetic space. Each point corresponds to a planning unit. The coordinates of the points denote the typical genetic characteristics of individuals sampled in that planning unit (based on an NMDS of the binary loci data). Planning units associated with points that are closer together contain individuals with more similar genetic characteristics than planning units that are further apart.

2.5 Implications and future directions

The `raptr R` package provides a unified approach to reserve selection. Conservation planners can use this `R` package to generate prioritizations that secure intra- and inter-specific biodiversity patterns. Both the simulations and case studies show that explicitly targeting a representative sample of the variation within each feature in a conservation planning exercise can substantially alter prioritizations. Additionally, we found prioritizations often needed to secure more habitat to conserve a representative sample of the variation within each feature. Although it may not be practical to conserve substantially more habitat than the amount specified using amount-based targets, we did find that even small increases in reserve system size could yield large gains.

One of the key advantages of this package is that it is general enough to incorporate almost any spatially explicit variable. This package can accommodate intrinsic or extrinsic variation in the feature(s). For example, adaptation processes could be secured using environmental variation (e.g. Carvalho *et al.*, 2011b), and trophic processes could be conserved by capturing the overlapping distributions of predator and prey species (e.g. Rayfield *et al.*, 2009; Chernomor *et al.*, 2015). As is often the case with multi-objective conservation planning problems, trade-offs between different goals may be unavoidable. For example, maximizing geographic spread of sites will almost always have an impact on connectivity. Additionally, space targets can be used to secure a representative sample of features for reasons that

are not related to biodiversity processes, such as obtaining a geographically representative set of reserves so there is equitable access to parks by people (e.g. Moilanen *et al.*, 2013). As long as the variation can be described using Euclidean distances—which could be achieved through transformation or dimension reduction—the R package can be used to obtain a representative sample.

The degree to which a prioritization truly secures a representative sample of a feature depends on (i) the attribute space(s), (ii) the distribution of demand points chosen by the conservation planner, and (iii) the space target used. The optimal solution will not effectively capture the decision maker's objectives if an inappropriate set of spatial variables or demand points are used to construct the attribute space(s). Demand points should be distributed across the full range of variation in an attribute space to obtain solutions that secure a representative sample (see the "`make.DemandPoints`" function for statistical routines to achieve this). As with conventional amount-based targets, planners will need to ensure that space targets are high enough to fulfill conservation objectives (e.g. to secure genetic diversity as mandated in the Convention on Biological Diversity). The most appropriate space target will depend on the species of conservation interest (e.g. threatened species may need more genetic variation to be conserved; Reed & Frankham, 2003), and how conserving more variation relates to the species' persistence (e.g. the relationship between inbreeding and extinction has a threshold; Frankham, 1995). We encourage conservation planners to consider carefully the relevant biological information to identify suitable targets.

Attribute data need to be spatially comprehensive to map planning units to attribute spaces. However, many real-world data sets are patchy. For example, due to constrained resources, genetic data may only be available for some planning units. To use such patchy data, planners could omit the planning units that are missing data, or estimate the missing data using spatially explicit models (e.g. kriging or generalized dissimilarity models; Oliver & Webster, 1990; Ferrier *et al.*, 2007). This approach has been successfully applied to a range of biological data sets (e.g. Thomassen *et al.*, 2010). We caution that inaccuracy and uncertainty in data can negatively impact prioritizations (Wilson *et al.*, 2005), and recommend that planners ensure that data are sufficiently reliable or use the reliable formulation (described in Appendix A) to accommodate uncertainty into the reserve selection process.

To maximize the long-term persistence of biodiversity, decision makers need to identify prioritizations that preserve existing patterns of biodiversity and the processes that support them. It is becoming increasingly apparent that simply protecting a large area of land will fail to achieve this (Barnes, 2015). Here, we developed the `raptr R` package to provide conservation planners with the tools to deliver cost-effective prioritizations that secure an adequate amount of a representative sample of biodiversity features. By exploring the functionality of this software package, we found that explicitly targeting a representative sample of the variation in each feature can result in substantially different solutions.

2.6 Acknowledgements

JOH is supported by an Australian Government Research Training Program (RTP) Scholarship. RAF and HPP have Australian Research Council Fellowships. This work was supported by the Centre for Biodiversity and Conservation Science (CBCS). We are also grateful to Dan Rosauer and an anonymous reviewer for their suggestions that greatly improved this work.

2.7 Data accessibility

The raptr R package can be downloaded from The Comprehensive R Archive Network (https://CRAN.R-project.org/package=raptr). To permit replication and validation of this work, all data, code, and results are available in an online code repository (https://github.com/jeffreyhanson/raptr-manuscript) and an online digital data repository (DOI: 10.5281/zenodo.847016).

CHAPTER 3

Environmental and geographic variables are effective surrogates for genetic variation in conservation planning

Protected areas must maintain evolutionary processes to facilitate the long-term persistence of biodiversity. The previous chapter showed that conservation planners should aim to incorporate information on species' patterns of genetic variation to ensure that it will be conserved. However, the genetic data required to map the spatial distribution of evolutionary processes are often prohibitively expensive to obtain and analyze. This chapter shows that freely available environmental and geographic variables can be used to generate plans for protected areas that represent intra-specific patterns of adaptive and neutral genetic variation. Thus future efforts should aim to use these surrogates to generate more effective protected areas when genetic data are not available.

A version of this chapter has been published in the *Proceedings of the National Academy of Sciences of the United States of America* as:

Hanson JO, Rhodes JR, Riginos C & Fuller RA (2017) Environmental and geographic variables are effective surrogates for genetic variation in conservation planning. *Proceedings of the National Academy of Sciences of the United States of America*, **114**, 12755–12760.

3.1 Abstract

Protected areas buffer species from anthropogenic threats and provide places for the processes that generate and maintain biodiversity to continue. However, genetic variation, the raw material for evolution, is difficult to capture in conservation planning, not least because genetic data require considerable resources to obtain and analyze. Here we show that freely available environmental and geographic distance variables can be highly effective surrogates in conservation planning for representing adaptive and neutral intra-specific genetic variation. We obtained occurrence and genetic data from the IntraBioDiv project for 27 plant species collected over the European Alps using a gridded sampling scheme. For each species, we identified loci that were potentially under selection using outlier loci methods, and mapped their main gradients of adaptive and neutral genetic variation across the grid cells. We then used the cells as planning units to prioritize protected area acquisitions. First, we verified that the spatial patterns of environmental and geographic variation were correlated, respectively, with adaptive and neutral genetic variation. Second, we showed that these surrogates can predict the proportion of genetic variation secured in randomly generated solutions. Finally, we discovered that solutions based only on surrogate information secured substantial amounts of adaptive and neutral genetic variation. Our work paves the way for widespread integration of surrogates for genetic variation into conservation planning.

3.2 Introduction

Protected areas spearhead conservation efforts (Watson *et al.*, 2014). They buffer species from anthropogenic impacts, providing places for them to persist. Since the resources available for conservation are limited, protected areas need to be sited in places that fulfill conservation objectives for minimal cost (Margules & Pressey, 2000). To achieve this, conservation planning exercises often generate plans for entire networks of protected areas (prioritizations) to preserve species and broad-scale biodiversity processes (Cowling & Pressey, 2001). After determining which species to protect (e.g. using evolutionary distinctiveness; Isaac *et al.*, 2007), conservation planners need to ensure that prioritizations represent the intra-specific genetic variation found within species to secure their long-term persistence (Crandall *et al.*, 2000; Moritz, 2002; Hendry *et al.*, 2010). As a consequence, there has been increasing interest in designing prioritizations that fulfill this objective (Diniz & Telles, 2006; Potts *et al.*, 2013; Carvalho *et al.*, 2017; Hanson *et al.*, 2017b; Nielsen *et al.*, 2017).

Although the strength of natural selection is a continuous force, genetic variation is often classified as adaptive or neutral (reviewed in Schoville *et al.*, 2012). Adaptive genetic variation is associated with loci that significantly affect fitness. Typically, such "adaptive" loci are detected due to anomalous patterns among individuals (e.g. Foll & Gaggiotti, 2008; Duforet-Frebourg *et al.*, 2014), though phenotypic-genetic associations are more likely to identify functionally important genomic regions for individual species (e.g. Steiner *et al.*, 2007). By representing the full range of adaptive variation, including adaptations of which we are currently unaware, protected areas can enhance a species' capacity to persist in a range of different conditions (Crandall *et al.*, 2000; Sgro *et al.*, 2011). In contrast, neutral genetic variation is associated with loci that do not significantly affect fitness, but instead reflect the evolutionary history of different populations. By representing the full range of neutral variation, protected areas can safeguard against the adverse effects of low genetic diversity (Moritz, 1994). Thus optimally sited protected area networks would represent patterns of both adaptive and neutral genetic variation (Moritz, 2002).

Recently, the use of surrogate data has been proposed to generate prioritizations that capture intra-specific genetic variation (Carvalho *et al.*, 2011b; Potts *et al.*, 2013; Ponce-Reyes *et al.*, 2014)—without needing to utilize genetic data directly. Since adaptation is ultimately driven by selection pressures, environmental variables have been proposed as surrogates for adaptive genetic variation (e.g. Cowling & Pressey, 2001; Carvalho *et al.*, 2011b). Specifically, by conserving individuals in a broad range of environmental conditions, we might expect to capture individuals with a diverse set of local adaptations, and overall capture a large proportion of the adaptive genetic variation present in the species. However, the effectiveness of this approach remains unverified. Neutral genetic variation arises from a reduction in gene flow between populations. Over the last few years, the field of phylogeography has predominantly focused on describing neutral genetic diversity, and so the landscape factors

that affect neutral variation are relatively well understood (reviewed in Avise, 2009). Potential surrogates for neutral genetic variation have been based on variables that predict the level of connectivity between different areas. For instance, according the isolation by distance model (Wright, 1943), populations located further apart are predicted to experience less gene flow, and in turn, share less genetic material at neutral loci. Based on this idea, spreading conservation priorities evenly across the geographic distribution of an island network has been found to capture more neutral genetic variation (Ponce-Reyes *et al.*, 2014). However this remains untested in spatially contiguous systems where connectivity is complicated by many additional factors (e.g. land use; Dudaniec *et al.*, 2016)—the typical situation in most conservation planning exercises.

Here we determine whether environmental and geographic surrogates capture adaptive and neutral genetic variation respectively in the context of conservation planning. We use distribution and genomic data for 27 alpine plant species in the European Alps that was obtained by the IntraBioDiv project (Taberlet *et al.*, 2012a,b). The data were collected following a gridded sampling scheme which we adopted as planning units for this study. We show that for most species the spatial patterns of adaptive and neutral genetic variation correlated with environmental (obtained from `http://worldclim.org`; Hijmans *et al.*, 2005) and geographic variables. We also show that for most species the level of association is strong enough to be operational for conservation planning. Finally, we demonstrate that using these surrogates results in prioritizations that secure a substantial proportion of intra-specific genetic variation.

3.3 Methods

We used data for 27 alpine plant species in the European Alps collected by the IntraBioDiv project (Figure 3.1a; Taberlet *et al.*, 2012a,b). This data set has been used to explore patterns of adaptive (e.g. Bothwell *et al.*, 2013) and neutral genetic variation (e.g. Alvarez *et al.*, 2009), and the potential for species richness as a surrogate for genetic diversity (Taberlet *et al.*, 2012a). Data were collected using a $20'$ longitude by $12'$ latitude grid (approx. 20 $\times$ 22.5 km; Appendix E). Project surveyors visited every second grid cell, and for each species, they collected samples from three individuals if any individuals were found. They genotyped samples using AFLPs, and constructed matrices denoting the presence of polymorphisms at loci for each species (mean 130.7 $\pm$ 54.9 SD markers genotyped per species; for more information see Alvarez *et al.*, 2009). Thus the data set contains information describing the genomic properties of individuals for each species over a geographic grid. We used these data because they comprised comparable genetic information for a range of species with different evolutionary and life histories collected using a standardized sampling scheme. All spatial and statistical analyses were conducted in `R` (version 3.3.2; R Core Team, 2016).

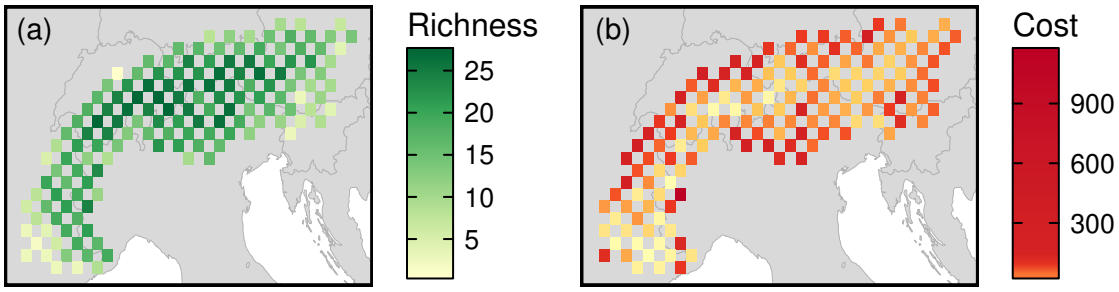


Figure 3.1 Map of the study area showing (a) the pattern of species richness, and (b) the distribution of acquisition cost depicted using a quantile-based color ramp. Squares denote planning units.

3.3.1 Landscape data

We adopted the sampling grid used to collect data as planning units to develop prioritizations. Of the total 388 cells in the grid, we used the 149 cells that contained samples for subsequent analysis. We calculated the total human population density inside each planning unit (1 km^2 resolution from the Global Rural-Urban Mapping Project; CIESEN, Columbia University *et al.*, 2011) and used this to represent acquisition cost (Figure 3.1b).

We created environmental and geographic surrogate variables for each species (Appendix F). To describe the geographic location of each planning unit (Appendix F), we projected the grid into an equidistant coordinate system (Europe Equidistant Conic; `ESRI:102031`), calculated the centroid of each grid cell, and extracted their two-dimensional coordinates. To describe the environmental characteristics of each planning unit (Appendix F), we obtained 19 climatic layers ($30''$ resolution; obtained from `http://worldclim.org`; Hijmans *et al.*, 2005), projected them and the planning units into an equal-area coordinate system (Europe Lambert Conformal Conic; `ESRI:102014`), and computed planning unit averages for each climatic variable. To reduce dimensionality, for each species, we subjected the climatic values associated with the planning units they were found in to a principal components analysis (PCA). We used the first three principal components to characterize climatic variation found across the species' geographic distributions inside the study area (mean 90.26 % variation described $\pm$ 0.99 SD of the total climatic variation; Appendix G). Thus we constructed a two-dimensional geographic space as a potential surrogate for neutral genetic variation, and a three-dimensional environmental space as a potential surrogate for adaptive genetic variation for each species. In these spaces, each planning unit was associated with a single point. Planning units with more comparable environmental conditions, or higher spatial proximity in the real world, were associated with points that were closer together in these environmental or geographic spaces.

3.3.2 Adaptive and neutral genetic data

To investigate the effectiveness of our surrogates, we first needed to identify which of the sampled loci were adaptive (Appendix F). Following recommended practices, we used two outlier detection methods to achieve this (an individual- and a population-level method; de Villemereuil *et al.*, 2014). The basic premise underpinning such methods is that neutral loci are expected to exhibit a certain level of variation, and loci that deviate from this expectation are likely to be under selection (Schoville *et al.*, 2012). The advantage of these methods—in contrast with environmental association analyses—is that they do not use environmental data, which would have introduced an element of circularity into our analysis. Loci identified by both outlier detection methods were treated as adaptive and the remainder were treated as neutral. To minimize false-positives, we omitted loci from both methods where the global frequency of the minor allele was less than 10 % and treated them as neutral.

The first outlier detection method involved fitting multinomial-Dirichlet models implemented in `BayeScan` (version 2.1; Foll & Gaggiotti, 2008). We adopted a similar methodology to Bothwell *et al.* (2013) and applied it to each of the species separately. Following their methodology, we initially grouped conspecifics into genetic lineages to further minimize false-positives (Appendix H), by fitting admixture and correlated alleles models implemented in `Structure` (version 2.3.4; 20 replicates per species; 5,000 admixture burn-in iterations; 300,000 burn-in iterations; 400,000 total iterations; Falush *et al.*, 2007) using the number of lineages previously determined by Alvarez *et al.* (2009) and combining replicate runs using `ClumPP` (version 1.1.2; greedy algorithm based on the G' statistic; 1,000 iterations; Jakobsson & Rosenberg, 2007). We then ran `BayeScan` for each species using these lineages (1:1 prior odds; 4 replicates per species; 20 pilot runs; 100,000 burn-in iterations; 110,000 total iterations thinned by 10 iterations) using a suitable false discovery rate ($q \leq 0.1$; Benjamini & Hochberg, 1995). We omitted individuals if their population membership was uncertain (maximum membership probability < 0.75).

The second outlier detection method involved fitting PCAs to identify outlier loci (implemented in the `pcadapt R` package; Luu *et al.*, 2016). To enable comparisons between the two outlier detection methods, we used the same individuals in this analysis as in the `BayeScan` analysis. For each of the 27 species, we first imputed missing data by replacing missing values with the average frequency among conspecifics, then ran a PCA over the loci matrix, and extracted the minimum number of components needed to secure the overarching population level variation among the loci (10 % of the variation in loci). We then computed *q*-values using Mahalanobis distances, and used the same false discovery rate used in the `BayeScan` analysis to identify loci under selection (using the `qvalue R` package; Storey *et al.*, 2015).

After classifying loci as adaptive or neutral, we mapped the main gradients of the adaptive

(if detected) and neutral genetic variation for each species. We discarded the population groupings, and partitioned species' adaptive and neutral loci into separate matrices. We applied NMDS (implemented in the `vegan R` package; Oksanen *et al.*, 2015) using Gower distances (via the `cluster R` package; Maechler *et al.*, 2015) to derive continuous variables that described the main gradients of adaptive and neutral genetic variation separately for each species (Appendix I). To ensure that the ordinations described a sufficient amount of the genetic variation, we ran successive scaling analyses with increasing dimensionality until a sufficient stress value was obtained (maximum stress value ≤ 0.25; 100 random starts for each analysis). Since each grid cell had multiple samples per species, we averaged the ordinated values for conspecifics in the same grid cell.

Thus we constructed an adaptive (if detected) and neutral genetic space for each species. For a given species, each planning unit in which it occurred was associated with a multi-dimensional point in the species' adaptive genetic space (if adaptive genetic variation was detected) and another multi-dimensional point in the species' neutral genetic space. Planning units that were closer together in these genetic spaces were occupied by individuals with more comparable AFLP data. By spreading out conservation effort across these genetic spaces, and in turn selecting planning units occupied by individuals with increasingly different polymorphisms, prioritizations can secure more genetic variation.

3.3.3 Prioritization method

We used the `raptr R` package to generate solutions (Hanson *et al.*, 2017b). This toolkit can identify the cheapest set of planning units required to preserve both a target proportion of the species' geographic range (using amount-based targets) and a target proportion of intra-specific variation (using space-based targets). We used the environmental and geographic surrogate spaces, and the adaptive (if detected) and neutral genetic spaces to generate and evaluate solutions. Solutions associated with negative values—because they secured very little genetic variation—were replaced with zeros to facilitate statistical analysis. We solved all reserve selection problems to within 10 % of optimality (using `Gurobi`; version 7.0.2).

3.3.4 Computational and statistical analyses

Our first aim was to determine if the environmental and geographic variables correlate with the spatial patterns of adaptive and neutral genetic variation. To achieve this, for each species, we created dissimilarity matrices using Euclidean distances and the data in the surrogate and genetic spaces. These matrices showed differences between the planning units occupied by the species in terms of the units' (i) geographic position, (ii) environmental characteristics, and the (iii) adaptive (if detected) and (iv) neutral genetic characteristics of the individuals inside them.

We fitted maximum likelihood population-effects (MLPE) models using maximum likelihood to investigate correlations between the dissimilarity matrices (using the `lme4 R` package; Bates *et al.*, 2015). These models use random effects to accommodate the structure of dissimilarity matrices. For each species, we fitted a MLPE model to the species' dissimilarity matrices based on geographic position and neutral genetic variation. If adaptive loci were detected, we also fitted a MLPE model to the species' dissimilarity matrices describing environmental and adaptive genetic variation. All data variables were *z*-transformed to improve convergence. To test if the surrogates explained the genetic variation, we compared each model to its null model using a χ^2 test, and applied Bonferroni corrections.

Our second aim was to determine if the environmental and geographic variables were effective surrogates for adaptive and neutral genetic variation. For each species, we generated 20,000 prioritizations by randomly selecting different combinations of planning units that the species occupied. For half of these random prioritizations, we calculated the proportion of geographic variation and neutral genetic variation they secured, and for the other half, we calculated the proportion of environmental variation and adaptive genetic variation they secured. These calculations were performed using the `raptr R` package.

We fitted two full generalized linear models (GLM) with logit link functions. The first model was fitted to the proportion of adaptive genetic variation secured in a prioritization using the proportion of environmental variation also secured in the prioritization. The second model was fitted to the proportion of neutral genetic variation secured in a prioritization using the proportion of geographic variation also secured in the prioritization. Additionally, both models contained a variable indicating the species for which the prioritizations were generated, and an interaction term. They were subjected to backwards stepwise term deletion routines to assess term significance. *Post-hoc* analyses were conducted to assess trends for each species using Bonferroni corrections (using the `multcomp R` package; Hothorn *et al.*, 2008). To assess the performance of the surrogates, we refit these models separately for each species and computed the Cragg and Uhler's pseudo-R^2 value (using the `pscl R` package; Jackman, 2015).

Our third aim was to determine if surrogate-based targets improved representation of genetic variation in prioritizations. As previously mentioned, we generated prioritizations using different combinations of targets and conservation planning scenarios. We used 20 % amount-based targets for each species in all prioritizations to secure an adequate proportion of the species' distributions. Based on the results from our previous analysis, we used 97.5 % surrogate-based targets and 90 % genetic-based targets. We generated a single solution for each target/scenario combination, except for the "single-species (equal costs) with amount-based targets" combination for which we generated 1,000 replicates because it had many optimal solutions. We computed the proportion of adaptive (if present) and neutral genetic variation secured for each species in the prioritizations.

We fitted generalized linear mixed-effects models (GLMMs) with logit link functions to evaluate the prioritizations (using the `lme4 R` package). We fitted a full model to the proportion of genetic variation secured in a given prioritization. This model contained categorical variables indicating the targets (amount only, amount and surrogate targets, or amount and genetic targets), the planning scenario (single-species, multi-species, or multi-species with cost), the type of genetic variation measured (adaptive or neutral), and all interactions between them. Data for same species were accommodated using a random intercept term. The full model was subjected to a backwards stepwise term deletion routine to assess significance. A *post-hoc* analysis was conducted using Tukey contrasts with Bonferroni corrections (using the `multcomp R` package).

3.4 Results

We detected putatively adaptive genetic variation in ten of the 27 plant species. We used two outlier detection methods—that solely used genetic data and did not use environmental data—to identify loci under selection. These methods returned reasonably consistent results (mean 87.92 % loci per species assigned the same classification $\pm$ 8.75 SD). Of the species that were associated with adaptive genetic variation, only a small proportion of loci were classified as being adaptive (mean 3.01 % loci per species $\pm$ 1.95 SD). After identifying the loci showing strong signals of selection, we used non-metric multi-dimensional scaling (NMDS) analyses to identify the main gradients of adaptive (if detected) and neutral variation for each species. Generally, only a small number of continuous dimensions were needed to sufficiently describe their patterns of adaptive ($K = 2$ for all species; Appendix I) and neutral genetic variation ($2 \leq K \leq 4$ for all species; Appendix I). The resulting ordinations were used to construct an adaptive (if detected) and neutral genetic space for each species. The spatial distribution of these genetic spaces (Appendix J) generally showed substantial spatial autocorrelation (e.g. *Cerastium uniflorum* and *Dryas octopetala*, Appendix J), suggesting that the proposed surrogate variables have the potential to be effective for conservation planning.

We verified that the spatial patterns in environmental variation correlated with the broad-scale patterns in adaptive genetic variation for each species. We also verified that there was a correlation between neutral genetic variation and variation in geographic position. For each species, we constructed dissimilarity matrices expressing differences between the planning units where individuals were detected based on the units' (i) environmental characteristics, (ii) geographic position, and also the (iii) adaptive (if adaptive loci detected), and (iv) neutral genetic characteristics of the individuals found inside them. The spatial patterns of environmental variation were significantly correlated with the patterns of adaptive genetic variation for eight of the ten species associated with adaptive variation ($P < 0.05$; mean 0.09 marginal R^2 $\pm$ 0.06 SD for significant models; Appendix K). Similarly, geographic distance between planning units was also correlated significantly with the spatial patterns of neutral genetic

variation among planning units for 26 of the 27 species ($P < 0.05$; mean 0.21 marginal R^2 $\pm$ 0.16 SD for significant models; Appendix K). Thus, for most species, planning units that contained different environmental conditions contained individuals with different adaptive genetic characteristics, and planning units that were further apart tended to contain individuals with different neutral genetic characteristics. After verifying these correlations, our next step was to determine whether they were strong enough to be useful for conservation planning.

For each species, we generated a suite of 10,000 random prioritizations, and calculated the proportion of environmental variation and adaptive (if present) genetic variation that each prioritization captured. We then repeated this process, and calculated the proportion of variation in geographic position and neutral genetic variation they sampled. The environmental and geographic variables were moderately effective predictors for the genetic variation represented by the randomly generated prioritizations (Figure 3.2, Appendix L). The relationship between the proportion of genetic and surrogate variation secured in the prioritizations varied among the species (species $\times$ proportion of environmental variation interaction term: ${\chi^2}_9 = 6717.74; P = 0.019$; species $\times$ proportion of geographic variation interaction term: ${\chi^2}_{26} = 4.588808 \times 10^4$; $P < 0.001$). *Post-hoc* analyses showed that these relationships were positive for all species (environmental: minimum $Z = 40.71$, maximum $P < 0.001$; geographic: minimum $Z = 23.23$, maximum $P < 0.001$; Appendix M). After establishing that the proportion of surrogate variation secured in a prioritization also predicted the proportion of genetic variation it secured, our final step was to determine if prioritizations generated using targets to represent variation in the surrogate variables were more effective at representing intra-specific genetic variation.

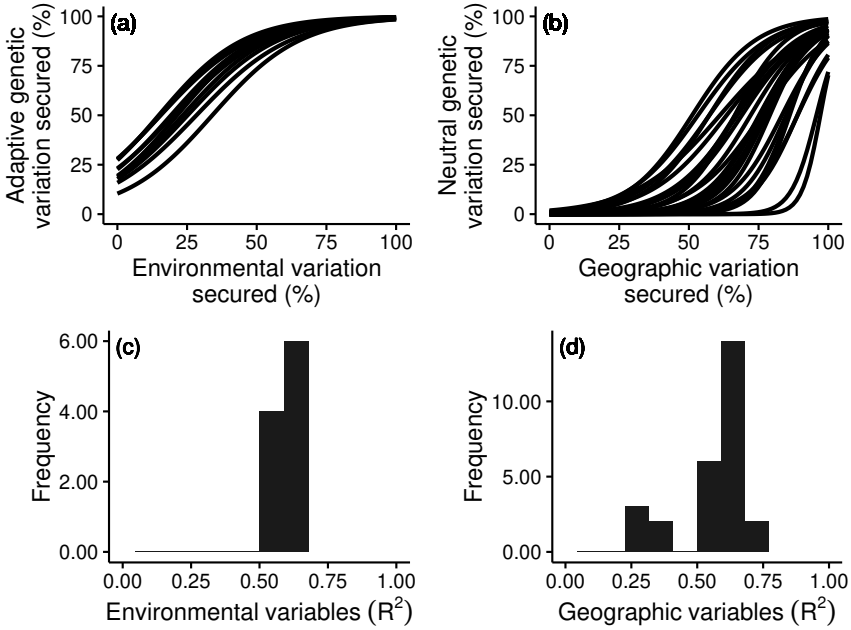


Figure 3.2 The proportion of adaptive (a; 10 species) and neutral (b; 27 species) genetic variation represented in a suite of randomly generated prioritizations for different species as a function of the surrogate variation they also secured (lines). Panels (c) and (d) show the distribution of R^2 values for the 37 models shown in panels (a) and (b).

To determine whether environmental and geographic targets could improve the effectiveness of prioritizations in representing genetic variation, we generated prioritizations using (i) "amount targets" reflecting the traditional approach of representing a certain proportion of each species' geographic distribution, (ii) "amount and surrogate targets" where we targeted representation of the environmental and geographic surrogates as well as a certain proportion of each species' geographic distribution, and (iii) "amount and genetic targets" where we targeted representation of the directly measured genetic variation as well as a certain proportion of each species' geographic distribution. To determine if these patterns were robust, for each of the three combinations of the targets, we generated prioritizations under four scenarios: (i) single-species with equal costs, (ii) single species with acquisition costs, (iii) multi-species with equal costs, and (iv) multi-species with acquisition costs (Figure 3.3, Appendix N).

The proportion of genetic variation secured in a given prioritization was not found to depend on any interaction between the planning scenario or targets used to generate the prioritization, or the type of genetic variation measured (${\chi^2}_{17} = 2.97$; $P > 0.99$). The proportion of genetic variation secured in a prioritization did depend, however, on the targets used to generate the prioritization (${\chi^2}_2 = 67.24$; $P < 0.001$). Most notably, prioritizations based on the surrogate variables represented significantly more genetic variation than amount-based prioritizations (93.25 % $\pm$ 13.81 SD overall genetic variation secured versus 78.39 % $\pm$ 23.54 SD; $Z_1 = 4.58$; $P < 0.001$), and were not distinguishable from prioritizations based on measured genetic data (94.68 % $\pm$ 3.69 SD; $Z_1 = 0.55$; $P > 0.99$). These results indicate that environmental and geographic surrogates in this case far outperform traditional amount-based conservation planning, and perform almost as well as a conservation plan based directly on genetic data.

Overall, the prioritizations tended to secure a greater proportion of adaptive genetic variation than neutral genetic variation (adaptive: mean 95.9 % $\pm$ 8.23 secured across all prioritizations, neutral: mean 72.08 % $\pm$ 24.06; ${\chi^2}_1 = 1381.57$; $P < 0.001$). Thus, regardless of the targets used to generate a prioritization, or the planning context under which the prioritization was generated, prioritizations tend to secure more adaptive than neutral genetic variation.

The average proportion of genetic variation secured by prioritizations varied under different scenarios ($\chi^2{}_3 = 33.97$; $P < 0.001$). Specifically, prioritizations generated under single-species scenario with equal costs (mean 78.42 % genetic variation secured $\pm$ 23.54) secured less genetic variation than those under the single species with acquisition costs (mean 88.74 % genetic variation secured $\pm$ 15.54; $Z_1 = 2.88$; $P = 0.024$) or the multi-species with acquisition cost scenarios (mean 91.91 % genetic variation secured $\pm$ 12.31; $Z_1 = 3.67$; $P = 0.001$). These results suggest the proportion of genetic variation secured in a prioritization may be sensitive to acquisition cost.

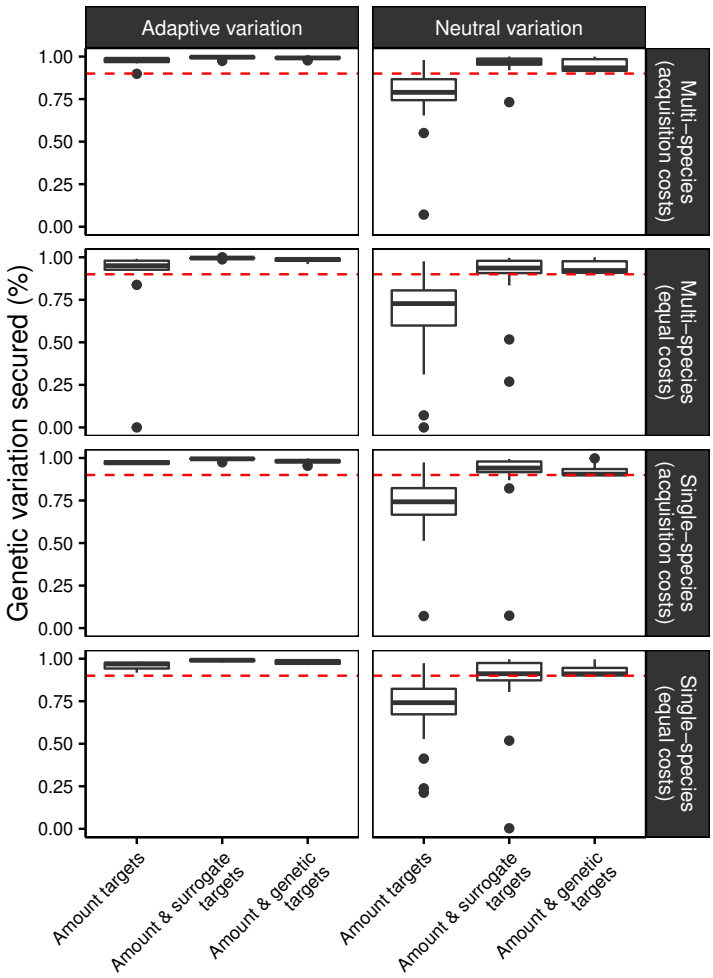


Figure 3.3 The proportion of adaptive and neutral genetic variation secured for each species in solutions based on different targets and scenarios. Box plots show the median, and the 25th and 75th percentiles. Whiskers show the data at 1.5 times the inter-quantile range. Points show data outside the whiskers. Red lines show the target proportion of genetic variation.

3.5 Discussion

We have shown that broad-scale environmental and geographic variables can be effective as surrogates for adaptive and neutral genetic variation in conservation planning. Our study is the first to investigate this using field measurements of genetic variation for a suite of plant species across an eco-region (Alvarez *et al.*, 2009). For most species, the proportion of

surrogate variation captured in a prioritization predicted the proportion of genetic variation that was also captured in the prioritization. Moreover, for most species, prioritizations generated using geographic targets secured a greater proportion of neutral genetic variation than traditional conservation planning methods. These environmental and geographic surrogates were based on freely available data sets and, with the exception of a few places with poor climatic data (Hijmans *et al.*, 2005), could be applied to any study region across the world.

Our results demonstrate that environmental and geographic variables are effective surrogates for most species considered here (Figure 3.2). Despite this, geographic distance was a surprisingly poor surrogate for neutral genetic variation for a few species (e.g. *Gentiana nivalis* R^2 = 0.3; *Gypsophila repens* R^2 = 0.25; *Luzula alpinopilosa* R^2 = 0.28). One explanation for this result is that geographic distance will be a poor surrogate for neutral genetic variation where spatial genetic structure is complicated by additional factors (Fortuna *et al.*, 2009; Potts *et al.*, 2013; Dudaniec *et al.*, 2016). For example, some species can maintain relatively high connectivity between distant populations (e.g. wind dispersing plants; Meirmans *et al.*, 2011), and in some places gene flow can be disrupted by landscape features (e.g. anthropogenically modified land; Dudaniec *et al.*, 2016). Overall, such species were the exception in our analysis, and both surrogates performed moderately well for most species.

Our results show that environmental and geographic surrogates can be used to capture genetic variation in prioritizations, the next question is: what percentage of surrogate variation should planners target to preserve genetic variation? To secure at least 90 % of the species' neutral genetic variation, prioritizations needed to sample 94.45 % $\pm$ 6.64 SD of the variation in geographic position among the planning units occupied by each species. Additionally, to secure at least 90 % of each species' adaptive genetic variation, prioritizations needed to sample 57.01 % $\pm$ 6.67 SD of the environmental variation in planning units occupied each the species. For all species, however, it was possible to generate solutions that secured a large proportion of the surrogate variation (> 90 %) and only a small proportion of genetic variation (< 20 %; Appendix L). Planners may avoid such outcomes by using both amount- and surrogate-based targets. Depending on the study area and the species of conservation interest, higher targets may be needed to increase the likelihood that prioritizations will secure a large proportion of the intra-specific genetic variation for all of the species in the planning exercise.

Our results suggest that conservation planners can secure a representative sample of intra-specific adaptive genetic variation using conventional reserve selection methods without needing surrogate or genetic data (Figure 3.3). This finding was evident in both the single- and multi-species prioritizations. One explanation for the single-species prioritizations is that because the spatial patterns of adaptive variation tended to cluster into a few main groups (e.g. *Carex sempervirens* and *Cerastium uniflorum*; Appendix J), a random selection of planning units across a species' geographic distribution would have a fairly high chance of capturing individuals that belonged to several of the main genetic groups. This mechanism

also explains why adaptive variation accumulated much more quickly than neutral genetic variation in the randomly generated prioritizations (Figure 3.2). On the other hand, one potential explanation for the multi-species prioritization is that because it was generated using a comprehensive set of species—each with their own habitat preferences—the solution was forced to secure each species in a range of different suitable habitats and, in turn, capture a representative sample of the species' adaptive genetic variation. If future studies on other taxa and biomes verify that conventional reserve selection methods do indeed conserve intra-specific adaptive genetic variation, this finding could have large implications for conservation planning.

There are several limitations associated with our analysis. Firstly, the size of the planning units we used (approx. 20 $\times$ 22.5 km) is larger than typically used in regional conservation planning. We used this resolution because the genetic data were collected at this scale. Whilst we could have interpolated the genetic data to a finer resolution, this would have introduced additional spatial auto-correlation and biased our analysis. Secondly, we used geographic distances as surrogates for neutral genetic variation. Although distances that incorporate data on dispersal cost may perform better (e.g. topography), such distances often require species-specific scaling (e.g. Dudaniec *et al.*, 2016) and so cannot easily be utilized in multi-species planning exercises. Thirdly, we used amplified fragment length polymorphism data (AFLP; Vos *et al.*, 1995) to describe genetic variation. Whilst next-generation sequencing provides higher resolution genetic information (Schoville *et al.*, 2012), we know of no suitable multi-species genomic data set and our methodology would still have utilized only the main gradients of the genetic variation to generate prioritizations in a feasible period of time. Furthermore, even with modern population genomic approaches, a survey can at best hope to identify markers that are linked to functionally adaptive variants. Fourthly, to our knowledge, none of the species we investigated here are at serious risk of extinction (Bornand *et al.*, 2016; IUCN, 2017). Since anthropogenic processes can alter patterns of species' genetic variation (e.g. Dudaniec *et al.*, 2016), future work should aim to establish if these surrogates are effective for imperiled species. Finally, conventional conservation planning exercises are sensitive to the quality and completeness of the underlying distribution data (Wilson *et al.*, 2005), and it is likely that the effectiveness of these surrogates is also sensitive to data quality. For instance, prioritizations generated using distribution data that omitted populations in specific habitats may fail to represent certain adaptations. Recent advances in distribution modeling have provided a wealth of data for conservation planners (Antoine *et al.*, 2013), and so they should endeavor to obtain high quality data where possible. Nevertheless, our findings pave the way for the widespread integration of evolutionary processes at the intra-specific level into reserve selection.

Genetic data provide a broad range of insights into a species' persistence and are fundamental to manging certain kinds of conservation problems (Hendry *et al.*, 2010). However, conservation interventions often need to be implemented urgently before the highest quality

data are available (Ceballos *et al.*, 2017). We found broad-scale environmental and geographic variables to be effective surrogates for representing adaptive and neutral genetic variation for most species in our study system. We call for further studies to examine these surrogates in other taxa and biomes. In cases where genetic data are not available, careful use of such surrogates in conservation planning could vastly improve the chances of long-term biodiversity persistence for relatively little additional cost.

3.6 Acknowledgements

JOH is supported by an Australian Government Research Training Program (RTP) Scholarship. RAF has an Australian Research Council Future Fellowship.

3.7 Data accessibility

To permit replication and validation of this work, all data, code, and results are stored in an online code repository (`https://github.com/jeffreyhanson/genetic-surrogates`) and an online data repository (DOI: `10.5281/zenodo.843625`).

CHAPTER 4

Conservation planning needs genetic data to maintain gene flow

Protected areas must maintain gene flow to facilitate species' long-term survival. Since genetic data are rarely available at the scale required for protected area planning, conventional approaches for maintaining gene flow involve spatially clustering reserves and representing species in contiguous areas containing suitable habitat. This chapter shows that these approaches may not be effective. This provides further insights into how genetic data can be used to guide conservation planning.

4.1 Abstract

To achieve the long-term persistence of biodiversity, protected area systems must conserve species in places with suitable habitat that are sited in a configuration that will facilitate gene flow. Since genetic data require considerable resources to obtain, many proxy methods have been developed with the aim to generate plans for protected area systems ("prioritizations") that will facilitate gene flow without needing genetic data. However, conventional methods—such as minimizing fragmentation among priority areas or representing species in contiguous areas—remain largely untested. Here we investigated the ability for prioritizations to maintain gene flow when they are generated using conventional methods for promoting connectivity. We used occurrence and genetic data obtained from the Global Biodiversity Information Facility and the IntraBioDiv project to create maps of habitat suitability and resistance to gene flow for eight alpine plant species. Next, we generated multi-species prioritizations that secured 10 % of the suitable habitat for each species and attempted to maintain gene flow by (i) penalizing fragmentation, (ii) representing species in contiguous areas of suitable habitat, and (iii) representing species in contiguous areas with minimal resistance to gene flow as modeled from genetic data. We found that prioritizations generated using fragmentation penalties failed to represent three of the eight species in areas that would facilitate moderate levels of gene flow. Similarly, the prioritizations that represented species in contiguous areas of suitable habitat were generally unable to sustain moderate levels of gene flow for two species. Although the prioritizations that represented species in contiguous areas of habitat with low resistance to gene flow were able to maintain high levels of gene flow, they selected over three times the amount of land in other prioritizations—suggesting that failing to account for gene flow when setting priorities may underestimate the scale of conservation action required. Our results suggest that conservation planners should make greater use of genetic data when designing protected area systems. By creating protected areas that foster evolutionary processes, conservation efforts stand a far greater chance at enhancing species' survival and, ultimately, the long-term persistence of biodiversity.

4.2 Introduction

We are witnessing a mass extinction (Ceballos *et al.*, 2017) and protected areas are urgently needed to buffer species from anthropogenic threats and provide places for biodiversity processes to continue (Watson *et al.*, 2014). Since it is rarely feasible to conserve all of the remaining habitat for a species, protected area systems must encompass a cost-effective subset of places that will enable species to persist indefinitely (Margules & Pressey, 2000). Gene flow plays an important role in species' persistence (Ellstrand, 2014), by helping populations adapt to new conditions (Hoffmann & Sgrò, 2011) and maintaining genetic diversity (O'Grady *et al.*, 2006). As a consequence, there has been increasing interest in developing plans for protected area systems (termed "prioritizations") and wildlife corridors to foster gene flow (Creech *et al.*, 2014; Hoffmann *et al.*, 2015; Dilkina *et al.*, 2017).

Over the last decade, the field of landscape genetics has made great strides towards understanding and mapping the spatial patterns of gene flow (Manel & Holderegger, 2013). State of the art methods use genetic data and optimization algorithms to transform maps showing the distribution of landscape features (e.g. roads or environmental conditions Dudaniec *et al.*, 2016) into resistance maps (Peterman, 2018; but see Cushman & Landguth, 2010). Resistance maps show how different areas hinder gene flow for a particular species (McRae *et al.*, 2008; Dudaniec *et al.*, 2016) and can be used to predict the effects of habitat modification on gene flow (Ruiz-Lopez *et al.*, 2016). Although resistance maps are often used to identify wildlife corridors (e.g. Dilkina *et al.*, 2017), they are rarely used to guide the acquisition of protected area systems (Keller *et al.*, 2015). This is in part, because genetic data remain prohibitively expensive to obtain at the scale required for planning protected area systems.

Many decision support tools for conservation planning aim to identify solutions that promote connectivity and enhance gene flow using proxies (Fuller & Sarkar, 2006; Moilanen & Wintle, 2007; Ball *et al.*, 2009). Conventional approaches for increasing connectivity have involved minimizing fragmentation among selected areas (Nicholls & Margules, 1993; Ball *et al.*, 2009; Hanson *et al.*, 2017b). More advanced approaches—based on the idea that areas containing more suitable habitat also facilitate greater levels of gene flow (Wang *et al.*, 2008; Mateo-Sánchez *et al.*, 2015; Keeley *et al.*, 2016)—represent species in contiguous areas that contain highly suitable habitat (Moilanen & Wintle, 2007; Krosby *et al.*, 2015; Wang & Önal, 2016). However, despite their widespread usage, the ability for these proxy-based approaches to identify prioritizations that actually maintain gene flow among protected populations remains poorly tested.

Using occurrence and genetic data for eight plant species in the European Alps (Taberlet *et al.*, 2012b,a), we generated multi-species prioritizations that were constrained to conserve an adequate amount of habitat for each species in (i) areas that exhibited varying levels of

spatial fragmentation, (ii) contiguous areas that contained habitat above a specific quality, and (iii) contiguous areas that facilitate specific levels of gene flow using resistance maps. We found very little evidence to suggest that prioritizations based on spatial measures of fragmentation or habitat suitability alone could maintain high levels of gene flow. This result means that, counter to previous studies which have found potential surrogates for genetic data (Ponce-Reyes *et al.*, 2014; Hanson *et al.*, 2017a), protected area systems may be more likely to enhance species' persistence they are designed using genetic data.

4.3 Methods

We used data collected by the IntraBioDiv project for eight plant species in the European Alps (*Androsace obtusifolia*, *Arabis alpina*, *Campanula barbata*, *Hornungia alpina*, *Phyteuma betonicifolium*, *Phyteuma hemisphaericum*, *Rhododendron ferrugineum*, and *Saxifraga stellaris*; Taberlet *et al.*, 2012b,a). Briefly, members of the IntraBioDiv project collected genetic samples for 27 species using a $20'$ longitude by $12'$ latitude grid (*c.* 22.3 $\times$ 25 km; Figure 4.1). They visited every second cell in this grid, and collected samples from three individuals from each of the 27 species present in the cell (Figure 4.1; elevation data shown was obtained from Jarvis *et al.*, 2008). Samples were genotyped using amplified fragment length polymorphisms (AFLP; Vos *et al.*, 1995), and were used to construct matrices denoting the presence/absence of polymorphisms at different loci (mean 151.58 $\pm$ 47.8 SD markers genotyped per species; for more information see Alvarez *et al.*, 2009). We used this data set because it contained comparable genetic data for a diverse group of species (Meirmans *et al.*, 2011). Although this data set contains information for 27 plant species, we used only eight species for our analysis. This is because we discarded (i) species that primarily used wind for pollination or seed dispersal (based on trait data in Thiel-Egenter *et al.*, 2009; Meirmans *et al.*, 2011) because contemporary landscape genetics methods are ill-suited for mapping asymmetric gene flow (e.g. Peterman, 2018), (ii) species with insufficient occurrence data to accurately model their distributions (i.e. species with less than 100 occurrence records remaining after data cleaning; *Cerastium uniflorum* and *Loiseleuria procumbens*), and (iii) species for which we were unable generate resistance maps that adequately explained their genetic variation ($R^2\beta$ statistic less than 0.35; *Gentiana nivalis* and *Gypsophila repens*; see below for more information). We performed all computational analyses using the `R` statistical computing environment (version 3.4.4; R Core Team, 2016).

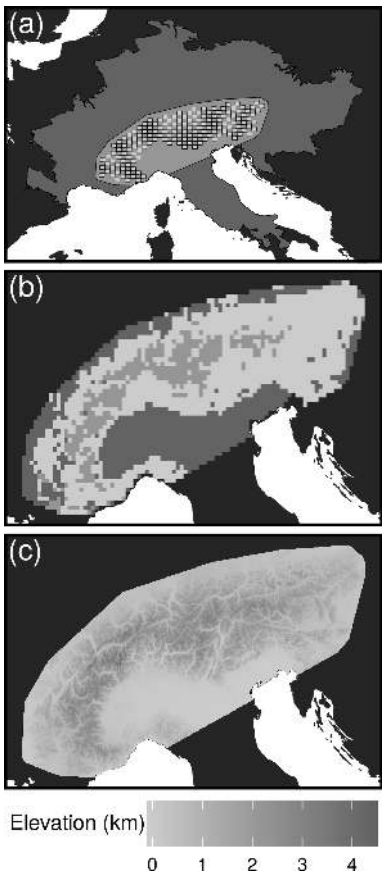


Figure 4.1 Panel (a) shows the spatial extent used for modeling species' distributions (dark gray), the spatial extent used to represent the study area (gray), and the grid cells used by the IntraBioDiv project to collect genetic samples (light gray). Panel (b) shows the distribution of forest (light gray), human modified land (dark gray), and other land cover types within the study area (gray). Panel (c) shows the elevation of different places within study area (note that elevational data was not used for model fitting and is shown here to help characterize the study area). All maps are rendered in an equal-area coordinate system (`EPSG:3035`).

We defined the study area for this project using the minimum convex hull that covered the sampling grid used by the IntraBioDiv project buffered by 20 km (*c.* 26.8 million ha in size; Figure 4.1). We obtained data on five predictor variables to make spatial predictions of habitat suitability and landscape resistance to gene flow (Appendix O). We obtained maps describing soil alkalinity (1 km resolution; Land Resources Management Unit *et al.*, 2010) and available growing conditions (i.e number of days in the year with temperatures over 5 °C; $2.5'$ resolution; Title & Bemmels, 2017) because they are likely important habitat requirements for these alpine plant species (Alvarez *et al.*, 2009; Meineri *et al.*, 2012). To account for spatial and seasonal variation in precipitation rates, we also obtained climatic layers describing annual precipitation (`BIO12`) and precipitation seasonality (`BIO15`; $30''$ resolution; available at `http://worldclim.org`; Hijmans *et al.*, 2005). Finally, to account for contemporary land cover, we obtained a land cover map (`MCD12Q1`; 500 m resolution corresponding to 2015; NASA LP DAAC, 2015) and reclassified it into three categorical classes representing forest (1–6), human-modified land cover (12–14), and other cover classes (6–11, 15–16; Figure 4.1b). All data were reprojected to an equal-area coordinate system (Lambert Azimuthal

Equal Area Europe; EPSG:3035), and then aggregated and resampled to 10 km resolution for subsequent analysis (using the sp and raster R packages; Pebesma & Bivand, 2005; Hijmans, 2016). We then used these maps for generating habitat suitability maps and resistance maps. We used the same predictor maps for generating habitat suitability maps as resistance maps to ensure that differences between them did not arise due to differences in data quality.

4.3.1 Habitat suitability maps

We fitted environmental niche models to generate the habitat suitability maps for each species. To increase the accuracy of predictions inside the study area, we used a much larger extent than the study area to train the environmental niche models (*c.* 108.9 million ha in size; Figure 4.1). Specifically, the modeling extent was defined by merging the study area with overlapping and neighboring World Wildlife Fund ecoregions (Olson *et al.*, 2001). Although the data set collected by the IntraBioDiv project contains high precision locality data, we elected not to use this data set for generating habitat maps so that our maps would be of comparable quality to those commonly used in conservation planning, and more independent of the genetic data. Thus we obtained 46,848 occurrence records for the eight study species from the Global Biodiversity Information Facility (using the rgbif R package; Chamberlain, 2017).

We cleaned the occurrence records prior to model fitting. Records were removed if they were duplicates, low in precision (i.e. coordinates were rounded to the nearest integer), not collected through direct human observations of wild individuals (e.g. data from reports, fossil records, and captive specimens were omitted), originally entered with invalid or suspicious coordinates, or entered with coordinates and tagged with a country that did not match. This resulted in a data set containing on average 490 records per species $\pm$ 355.75 SD (Appendix P). Additionally, to limit the effects of sampling bias on model performance, we extracted the total number of plant records in the Global Biodiversity Information Facility database over a 50 $\times$ 50 km grid, excluding records associated with the previously discussed issues, and used them to generate a map showing the spatial patterns of sampling bias (Appendix Q).

We fitted an environmental niche model for each of the eight alpine plant species (Figure 4.2). For each species, MaxEnt models (version 3.4.0; Phillips *et al.*, 2017) were fitted using its presence records and 10,000 background records randomly generated from the sampling bias map. The models were trained using ten-fold cross-validation and fitted using the five predictor variables. They were tuned using different feature class combinations ('L', 'LQ', 'H', 'LQH', 'LQHP', and 'LQHPT'; see Phillips *et al.*, 2017 for explanation) and regularization penalties (0.1, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, and 4; using the ENMeval R package; Muscarella *et al.*, 2014). The best performing model for each species was determined using

the average area under the receiver operating characteristic curve (AUC) statistic on the test data. In other words, we selected models that had the greatest predictive ability and did not over-fit the data used to train them. The models were then thresholded by maximizing the sum of the sensitivity and specificity statistics (Appendix R; Liu *et al.*, 2005). The thresholded habitat maps sufficiently described the species' geographic distributions (mean 0.89 AUC; ranging from 0.76 to 0.94; Appendix P).

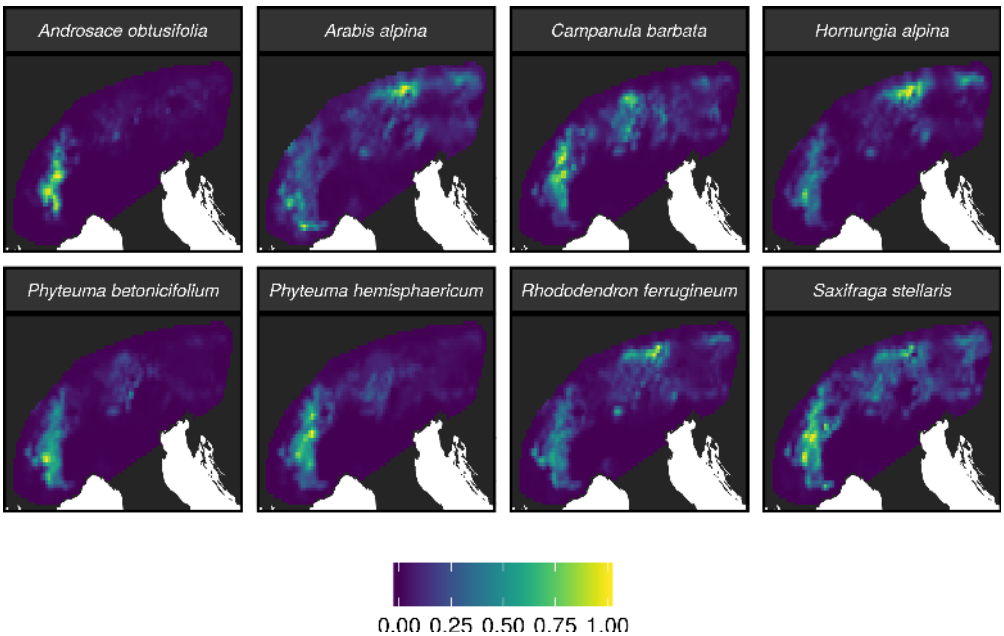


Figure 4.2 Distribution of habitat suitability for each species. Squares correspond to planning units. For visualization purposes, suitability data for each species have been rescaled to lay between zero and one.

4.3.2 Resistance maps

We generated resistance maps for each of the eight species using location and genetic data collected by the IntraBioDiv project. Since the genetic data set contained up to three replicate samples per sampling grid (*c.* 22.3 $\times$ 25 km in size) and because they were often collected with high spatial proximity (less than 1 km apart), we grouped together the samples collected in each sampling grid and calculated pairwise Nei's distances (using the `poppr R` package; Nei, 1978; Kamvar *et al.*, 2014). Thus we obtained measures of genetic dissimilarity that described how different the individuals in each sampling grid were compared to those in other sampling grid for each species. Although such genetic dissimilarity metrics can provide biased estimates of gene flow (e.g. due to sampling error, selection, and mutation; Whitlock & Mccauley, 1999), they are routinely used to model gene flow in landscape genetics (Mateo-Sánchez *et al.*, 2015; Ruiz-Lopez *et al.*, 2016). After calculating the pairwise genetic distances, they were used to produce maps describing the spatial patterns of gene flow for each species.

We used the `ResistanceGA R` package to generate resistance maps for each species (Figure 4.3; Peterman, 2018). Broadly speaking, this methodology involves using genetic algorithms to identify resistance maps that best describe the genetic distances calculated for different sampling locations using predictor maps (see Peterman, 2018 for more information). Following best practices (Ruiz-Lopez *et al.*, 2016), we trained the genetic algorithms using a suitable population size (i.e. 250), terminated them when they showed no improvement after 50 consecutive iterations, and employed a local search routine (using default parameters) to fine-tune the results. Commute distances were used to model dispersal across the resistance maps (via the `gdistance R` package; van Etten, 2017), because they are computationally efficient and can accommodate multiple pathways (similar to resistance distances calculated in `Circuitscape`; McRae *et al.*, 2008). The genetic algorithms evaluated resistance maps according to the log-likelihood of maximum likelihood population effects (MLPE; Clarke *et al.*, 2002) models fit to the genetic and resistance data.

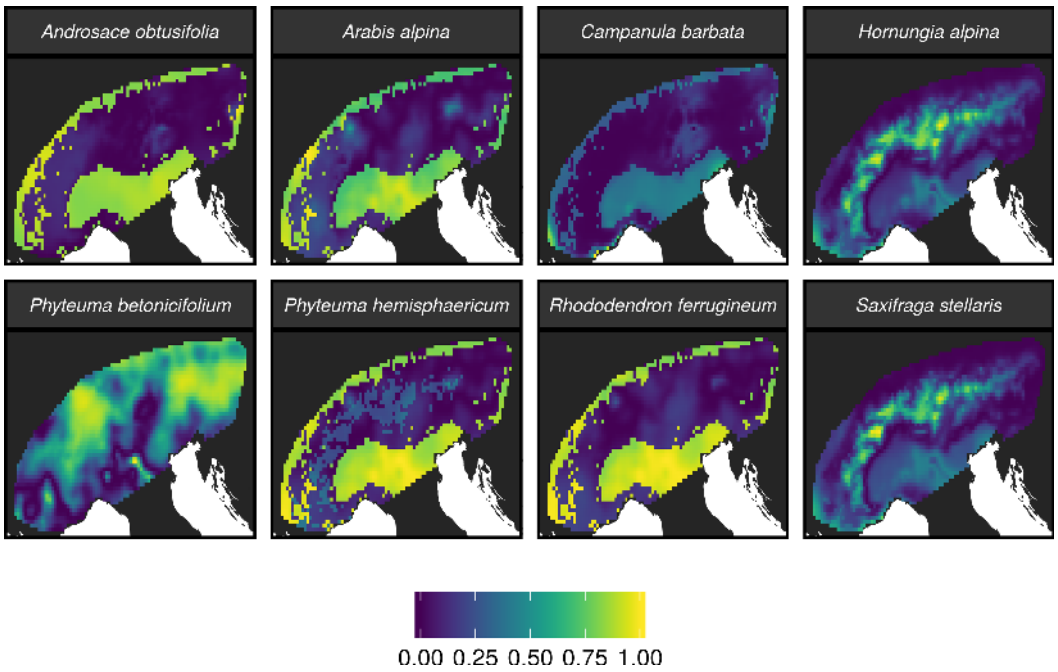


Figure 4.3 Resistance to gene flow for each species. Higher values denote areas that offer a greater barrier to gene flow. Squares correspond to planning units. For visualization purposes, resistance data have been rescaled to between zero and one.

We assessed the quality of the resistance maps to ensure that they adequately described the spatial patterns of gene flow for each species. To achieve this, the MLPE models were refitted using restricted maximum likelihood and $R^2\beta$ statistics were then calculated for each model (Edwards *et al.*, 2008; using the `r2glmm R` package; Jaeger, 2016). These statistics demonstrated that the resistance maps for the eight species explained an adequate proportion of the genetic differences among sampling grids (mean 0.66 $R^2\beta \pm$ 0.1 SD; Appendix S). We further verified that these statistics were robust to uncertainty by performing a bootstrap analysis (following Dudaniec *et al.*, 2016), comprising of 10,000 replicates with each replicate containing 80 % of the sampling grid cells where each species was sampled

(Appendix S). As mentioned earlier, the analysis originally included two additional species but they were omitted from the analysis because their resistance models were unable to explain an adequate proportion of their genetic variation ($R^2\beta \leq 0.35$; Keller *et al.*, 2015.

4.3.3 Prioritizations

We generated 2,649 planning units inside the study area using a 10 $\times$ 10 km grid. The planning units were generated to have the same spatial extent as the grid cells in the habitat and resistance maps. We then overlaid the planning units with these maps and extracted the data. Thus each planning unit was associated with the following data for each species: a continuous value for resistance to gene flow, a continuous value for habitat suitability, and a binary value indicating the presence or absence of suitable habitat. Using the planning units and their corresponding data, we then generated nine multi-species prioritizations to examine the performance of different approaches for maintaining gene flow. These prioritizations all shared the same primary objective—minimize the number of selected planning units in the solution—and contained constraints to secure 10 % of each species' distribution of suitable habitat (using the binary habitat suitability data). They only differed in their approach taken to promote gene flow.

The first three prioritizations were generated with additional penalties to punish fragmented solutions. To assess how the magnitude of the penalty influences performance, each prioritization was generated using a different penalty (low: 0, medium: 0.00000000004, and high: 0.0001). Note that these penalties were formulated using the same mathematical definitions used in the `Marxan` decision support tool (Ball *et al.*, 2009; Beyer *et al.*, 2016), and so the resulting prioritizations may exemplify conventional conservation prioritizations.

The next three prioritizations were generated using additional constraints to ensure that each species was represented in a spatially contiguous area (see Appendix T for mathematical formulation that builds on Önal & Briers, 2006; Cerdeira *et al.*, 2010; Beyer *et al.*, 2016). To explore how different habitat thresholds influenced performance, each prioritization was generated to represent species in a spatially contiguous unit of habitat above specified percentile threshold of habitat suitability. These thresholds were expressed along a continuum where the 0th percentile corresponded to all of the suitable habitat for each species (as defined using the binary suitability data) and the 100th percentile corresponded to the highest possible threshold of the continuous suitability data which resulted in a contiguous area containing 10 % of the suitable habitat for each species (defined using the binary suitability data). Given this scaling procedure, the thresholds were set as the 0th (low), 75th (medium), and 100th (high) percentiles of the continuous habitat suitability values for each species (see Appendix U for maps). This scaling procedure was used to standardize comparisons between species that have different levels of spatial auto-correlation in their habitat suitability.

The final three prioritizations were generated using additional constraints to represent each species in a spatially contiguous area containing resistance values below a specified threshold (using the same constraints detailed in Appendix T). For convenience, we expressed the resistance values as conductance values (calculated as the reciprocal value; McRae *et al.*, 2008), such that planning units with higher conductance values have a greater capacity to facilitate gene flow. Similar to the scaling procedure described for the continuous suitability data, the conductance percentile thresholds were expressed along a continuum where the 0th percentile corresponded to all of the conductance data inside areas with suitable habitat (defined using the binary suitability data) and the 100th percentile corresponded to the highest possible threshold for the continuous conductance data which resulted in a contiguous area containing 10 % of the suitable habitat for each species (defined using the binary suitability data). Given this scaling procedure, the thresholds were set as the 80th (low), 90th (medium), and 100th (high) percentiles of the continuous conductance data for each species (see Appendix V for maps).

All prioritizations were solved to within 15 % of optimality using `Gurobi` (version 7.5.1) and the `prioritizr R` package (Hanson *et al.*, 2017c). As a consequence, the level of fragmentation in species' distributions of suitable habitat did not affect the quality of the solutions. After generating the prioritizations, we then assessed their capacity to represent species in a configuration that would maintain various levels of gene flow by calculating the maximum amount of contiguous habitat for each species in each prioritization given various conductance thresholds (using the `igraph R` package; Csardi & Nepusz, 2006).

4.4 Results

We found that none of the prioritizations generated using fragmentation penalties sited priority areas in a configuration that would maintain moderate levels of gene flow for most species (i.e. minimum level of gene flow required greater than the 50th percentile; Figure 4.4). The prioritization generated without any fragmentation penalties was extremely fragmented (Figure 4.4a), and so none of the species were represented in a contiguous section of land containing 10 % of their suitable habitat. As a consequence, this prioritization was not able to maintain low levels of gene flow for any species (minimum level of gene flow required greater than the 20th percentile). Next, the prioritization generated with medium fragmentation penalties contained two reserves (Figure 4.4d). Although this prioritization represented one species in a contiguous section of land containing 10 % of its suitable habitat, this prioritization was not able to maintain even a moderate level of gene flow for this species. Moving on, the prioritization generated with high fragmentation penalties contained a single reserve (Figure 4.4g). As a consequence, this prioritization represented all species in a contiguous section of land that contained 10 % of their suitable habitat. It also sited priority areas in a configuration that would maintain moderate levels of gene flow for five species. These

results provide evidence that, for some species, fragmentation penalties can help prioritizations represent species in areas that will maintain gene flow. However, these results also suggest that simply representing species in spatially clustered areas is not likely to facilitate strong levels of gene flow because spatially clustered areas may not necessarily contain habitat that can facilitate strong levels of gene flow.

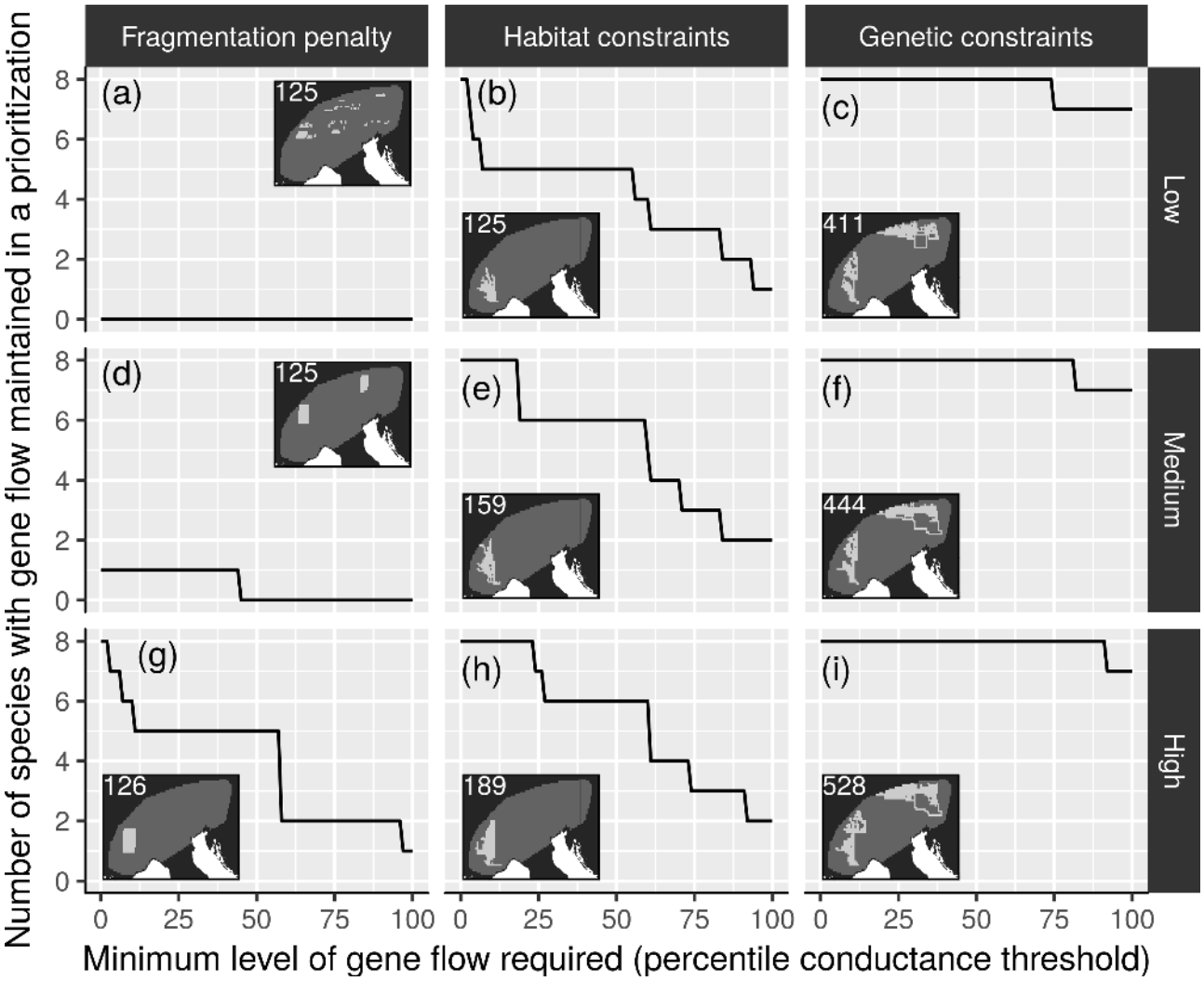


Figure 4.4 The ability for different prioritizations to represent species in areas that maintain gene flow. Each panel corresponds to a different prioritization and shows the number of species that are represented in areas that maintain gene flow as a function of the minimum level required to be maintained. Map insets depict the planning units (dark gray), the prioritizations (shown in in light gray), and the number of selected planning units is shown in the top left-hand corner of the inset. The left column of panels corresponds to prioritizations generated using low, medium, and high fragmentation penalties. The middle column of panels corresponds to prioritizations that represent species in contiguous areas containing low, medium, and high habitat suitability. The right column of panels corresponds to prioritizations generated using genetic data. These prioritizations represent species in contiguous areas that facilitate low, medium, and high levels of gene flow. Most of the prioritizations were able maintain low levels of gene flow for most of the studied species. However, only the prioritizations generated using genetic data were able to maintain strong levels of gene flow for the all of the studied species.

None of the prioritizations generated to represent species in a contiguous area of suitable habitat were able to maintain moderate levels of gene flow for all eight of the studied species (Figure 4.4). The prioritization generated to represent species in contiguous areas containing marginally suitable habitat (Figure 4.4b) was able to maintain moderate levels of gene flow for five species (minimum level of gene flow required greater than 50th percentile), and high levels of gene flow for three species (minimum level of gene flow required greater than 80th percentile). Similarly, the prioritizations generated to represent species in contiguous areas containing moderately suitable habitat (Figure 4.4e) and extremely suitable habitat (Figure 4.4h) were only able to maintain high levels of gene flow for three species (minimum level of gene flow required greater than 80th percentile). These results show that, for the species studied here, representing populations in contiguous areas of suitable habitat is not an effective strategy for maintaining high levels of gene flow—regardless of the suitability of the habitat located in these contiguous areas. To gain further insight, we investigated the relationships between habitat suitability and landscape resistance to gene flow.

We detected significant negative correlations between the spatial patterns of habitat suitability and resistance to gene flow for seven out of the eight species (one-sided Spearman's rank correlation tests adjusted using Bonferroni corrections; Appendix W). Thus for most species, areas that contained larger amounts of suitable habitat also tended to offer a smaller impediment to gene flow. This result was expected because we selected these species based on the fact that they exhibit dispersal mechanisms which would predispose them towards showing this relationship. Since the prioritizations generated to represent species in contiguous areas of suitable habitat were unable to maintain strong levels of gene flow, these results suggest that the correlation between habitat suitability and resistance to gene flow—despite being statistically significant—may not necessarily be strong enough to operationalize for conservation planning.

We confirmed that generating prioritizations to represent species in contiguous areas that offer low resistance to gene flow actually resulted in prioritizations that were able to maintain strong levels of gene flow. All three of the prioritizations generated to represent species in areas with pre-specified levels of resistance to gene flow (Figure 4.4c, Figure 4.4f, and Figure 4.4i) were able to maintain moderate levels of gene flow for all eight of the studied species (minimum level of gene flow required greater than 50th percentile). Additionally, the prioritizations generated to represent species in areas that contained medium and high conductance to gene flow (Figure 4.4f and Figure 4.4i) were able to maintain high levels of gene flow for all eight species (minimum level of gene flow required greater than 80th percentile). Furthermore, all of the prioritizations generated using gene flow contiguity constraints contained far more planning units than the prioritizations generated using the fragmentation penalties or habitat contiguity constraints. These findings suggest that failing to account for gene flow during reserve selection may underestimate the scale of conservation actions required to achieve species' persistence.

4.5 Discussion

We aimed to understand how effective commonly used approaches for increasing connectivity in prioritizations are at facilitating gene flow between protected populations. By evaluating the performance of prioritizations generated using fragmentation penalties and habitat contiguity constraints using measurements derived from genetic data, we found that these approaches can produce prioritizations that do not maintain strong levels of gene flow. This occurred despite the fact that areas which contained more suitable habitat tended to offer lower resistance to gene flow—a critical assumption that underpins many approaches for promoting connectivity (Moilanen & Wintle, 2007; Saura & Pascual-Hortal, 2007; Krosby *et al.*, 2015). Our results add to the growing body of evidence that direct integration of genetic data into conservation planning can markedly improve performance (Frankham, 2010; Hendry *et al.*, 2010; Taberlet *et al.*, 2012a). We encourage future studies to examine the return on investment for using genetic data with real-world cost data to help conservation planners better understand the benefits of genetic data.

We found that prioritizations needed to be generated directly using genetic data to maintain strong levels of gene flow for all eight of the study species. These results mean that commonly used proxy approaches for increasing genetic connectivity, such as penalizing spatial fragmentation (Ball *et al.*, 2009; Hanson *et al.*, 2017b) or representing species in contiguous areas of suitable habitat (Saura & Pascual-Hortal, 2007; Krosby *et al.*, 2015) may not be as effective as previously thought. Furthermore, these results contrast with earlier work which has identified cost-effective surrogates for genetic variation (Ponce-Reyes *et al.*, 2014; Hanson *et al.*, 2017a). One explanation for these results is that small changes to a prioritization can have a much larger influence on its ability to maintain gene flow compared to its ability to represent genetic variation. For instance, consider a prioritization comprising a set of planning units arranged in a chain. If we remove a single planning unit from the middle of the chain, this would mean that the prioritization is only able to facilitate gene flow among a much smaller subset of areas compared to the original prioritization. However, removing a single planning unit might not have such a negative impact on the genetic diversity represented in the prioritization if the remaining planning units still capture multiple genetic types. This result suggests that conservation planners will need to begin explicitly accounting for gene flow using genetic data, especially where protected area systems are being designed to help populations adapt to changing conditions (Sgro *et al.*, 2011).

Habitat suitability was significantly negatively correlated with landscape resistance to gene flow for seven out of the eight studied plant species (Appendix W). Although these relationships were not strong enough to operationalize for conservation planning, the fact that we found evidence to support their existence at least partially validates one of the key assumptions used to enhance functional connectivity in prioritizations (e.g. Moilanen & Wintle, 2007; Krosby *et al.*, 2015). Furthermore, these results support previous studies which have

detected similar relationships in mammalian species (Wang *et al.*, 2008; Mateo-Sánchez *et al.*, 2015; Keeley *et al.*, 2016). Initially, we expected all species to exhibit this relationship, but we were unable to find evidence to support its existence for *Phyteuma betonicifolium* (Appendix W). One potential explanation for this result is that *Phyteuma betonicifolium* might be pollinated by a range of species with different habitat preferences (as observed in other alpine species; Arroyo *et al.*, 1982), and so populations separated by less favorable habitat may still experience high rates of gene flow. Another explanation could be that the genetic data for this species correspond to historical patterns of gene flow which do not reflect present-day patterns of gene flow. However, this is less likely because we were able to produce a landscape resistance map using contemporary environmental data that explained a large proportion of its genetic variation between different sampling locations (0.71 $R^2\beta$).

The landscape features had different impacts on landscape resistance to gene flow for different species (Appendix X). For four out of the eight species, soil alkalinity had the greatest impact on determining how much a given area impeded gene flow. Previous work has found that soil substrate conditions have played a large role in shaping the recolonization patterns for these species following the Last Glacial Maximum (Alvarez *et al.*, 2009), and our results suggest that these conditions are still playing a role in their evolution today. For the other four species, available growing conditions and precipitation seasonality had the greatest impact on resistance to gene flow. These results support previous work which suggests that climatic conditions can have a large influence on gene flow in alpine plant species (Hirao & Kudo, 2004).

This study is associated with several limitations. Firstly, our prioritizations involved coarser-scaled planning units than commonly used in conservation planning (Margules & Pressey, 2000). Although we could have interpolated the resistance maps to a finer spatial resolution, this process would have introduced artificial spatial auto-correlation and biased our analysis. Secondly, we used eight alpine plant species, and though this constitutes many more species than commonly examined in landscape genetics studies (Manel & Holderegger, 2013), future work is needed to test our hypotheses in other taxa and biomes. Thirdly, the prioritization approaches for representing species in contiguous areas required thresholding continuous suitability and resistance maps. Although this means that practitioners will require a greater understanding of their study species to employ these approaches, similar to representation targets, such thresholds provide a transparent and effective method for ensuring that solutions exhibit a minimum level of performance (Carwardine *et al.*, 2009; Di Marco *et al.*, 2016). Fourthly, the prioritization approaches used to represent species in contiguous areas will currently not scale to scenarios involving hundreds of species or hundreds of thousands of planning units. Despite this, we were able to solve all nine prioritizations in less than a day, which suggests that recent advances in exact algorithm solvers mean that conservation planners can begin incorporating complex connectivity requirements into reserve selection (Beyer *et al.*, 2016). It would be a considerable advance for conserva-

tion planning to begin using genetic data to guide the selection of priority areas (Keller *et al.*, 2015).

Current approaches for increasing connectivity in prioritizations often involve spatially aggregating priority areas, or representing species in contiguous areas containing suitable habitat. Here we found that such approaches are not necessarily reliable for securing species in areas that facilitate high levels of gene flow. Although future studies are still needed to verify our findings in taxa and biomes, in the meantime, we urge conservation planners to carefully consider using genetic data when designing prioritizations to ensure that priority areas can maintain strong levels of gene flow among protected populations.

4.6 Acknowledgements

JOH was supported by an Australian Government Research Training Program (RTP) Scholarship. RAF had an Australian Research Council Future Fellowship. This work was supported by a research allocation from the National eResearch Collaboration Tools and Resources (NeCTAR) project.

CHAPTER 5

Global conservation of realized niches

Protected areas need to foster species' capacity to adapt to changing conditions—especially in the face of land use change and climate change—for them to survive in the long term. Building on ideas in Chapters 2 and 3, this chapter investigates how well the realized niches for most of the world's amphibians, birds, and mammals are represented by existing protected areas. The results show that more protected areas are needed to fully conserve species' realized niches, either by expanding protected areas along environmental gradients or by increasing the coverage of the species' realized niche. This chapter also identifies a suite of species that have very low niche protection, and highlights places where conservation action would lead to rapid increases in niche protection.

5.1 Abstract

As we draw closer to 2020, when the world's governments will meet to coordinate global conservation efforts over 2021–2030, policy makers need to begin developing plans for the protection of evolutionary processes. The rate of environmental change is rapidly accelerating and many species' will need to adapt to survive. By conserving populations across a range of environmental conditions, and in turn, safeguarding local adaptations and the habitats that create them, protected areas can foster adaptive evolutionary processes. Yet the extent to which species' environmental niches are covered by protected areas is unknown. Based on a global data set containing 19,937 vertebrate species, we show that 4,862 (93.6 %) amphibian, 9,016 (93.2 %) bird, and 4,810 (94.9 %) mammal species do not have their environmental niche adequately represented in protected areas. Considering that 5,517 (27.5 %) of the species investigated here have an adequate proportion of their overall habitat inside protected areas, these statistics suggest that ignoring intra-specific variation may lead to a serious over-estimation of how well species are conserved. To guide future efforts, we identified priority locations for expanding niche representation by protected areas. These occurred worldwide, although many were concentrated in global biodiversity hotspots, including Colombia, Papua New Guinea, South Africa, and southwest China. Conventionally, data on species' environmental niches has not been used to identify priority areas. We found that priority areas resulting from conventional approaches that omit environmental niche data would fail to adequately represent the environmental niches for 7,989 (40.1 %) species. Moving forward, we urge policy makers to mandate the protection of evolutionary processes to minimize global extinctions.

5.2 Introduction

In response to the biodiversity crisis (Ceballos *et al.*, 2017), 195 governments have signed the Convention on Biological Diversity (`https://cbd.int`). Signatories have pledged to conserve habitats that support "key evolutionary processes" (Article 8) and protect at least 17 % of the Earth's land and 10 % of its oceans through ecologically representative systems of protected areas and other effective area-based conservation measures (Aichi Target 12). Since habitats can foster local adaptations to their local conditions (Kawecki & Ebert, 2004), a well designed protected area system could conserve the areas needed to maximize species' persistence and adaptation by safeguarding populations in suitable habitats that contain different environmental conditions (i.e. a representative sample of the species' realized niche; Hutchinson, 1957; Moritz, 2002; Sgro *et al.*, 2011; Hanson *et al.*, 2017a). However, despite progress in understanding the total amount of habitat covered by protected areas (Rodrigues *et al.*, 2004; Beresford *et al.*, 2011; Buchanan *et al.*, 2011), very little is known about how well protected areas cover the diversity of habitat conditions that affect species' adaptive potential (Breiner *et al.*, 2017). With 2020 approaching, and consideration already being given to the framing of a new Strategic Plan for Biodiversity, it is imperative that conservation scientists and policy makers know how much progress has been made towards protecting biodiversity and how to shape future global conservation targets.

Here we aimed to investigate the performance of the global protected area system in representing species' realized climatic niches (hereafter referred to as niches). To achieve this, we obtained extent of suitable habitat maps for most of the world's known bird (n = 9,670; Beresford *et al.*, 2011; Buchanan *et al.*, 2011), terrestrial mammal (n = 5,070; Rondinini *et al.*, 2011) and amphibian species (n = 5,197; Ficetola *et al.*, 2015) and split each species' habitat map into multiple climatic partitions (Carvalho *et al.*, 2011b) using global climatic data (Hijmans *et al.*, 2005). We also aimed to identify priority areas for expanding protected area coverage to conserve species' niches. Finally, we compared conventional approaches that neglect evolutionary processes with the novel approaches developed in this study to understand how explicitly accounting for species' niches might alter conservation priorities. Failing to account for evolutionary processes during conservation assessments and planning might lead to serious underestimation of threat status, insufficient conservation interventions, and ultimately extinctions (Moritz, 2002; Breiner *et al.*, 2017).

5.3 Methods

5.3.1 Data

We compiled a map delineating the global distribution of protected areas (Appendix Y). To achieve this, we obtained data from the World Database on Protected Areas (WDPA; IUCN & UNEP-WCMC, 2016), and prepared it for analysis following best practices (code available at `https://github.com/jeffreyhanson/global-protected-areas`; Runge *et al.*, 2015). First, we reprojected the database to an equal-area coordinate system (World Behrman; `ESRI:54017`). Second, we excluded UNESCO Biosphere Reserves (Coetzer *et al.*, 2014) and sites with unknown or proposed designation status. Third, we extracted the protected areas only represented with a point locality, reprojected them to an equidistant coordinate system (World Equidistant Cylindrical; `ESRI:54002`), buffered them according to their reported area, and merged them with the main dataset ($n = 14,304$ with a mean reported area of 160 m^2 $\pm$ 1677.7 SD). Fourth, we spatially dissolved overlapping areas and removed slivers. After cleaning the data, we overlaid the protected areas onto a grid (5 $\times$ 5 km resolution) and identified which cells had more than half of their land inside protected areas. This operation had a negligible effect on the analysis, with only 1.8 % (346,974 km^2) of the global amount of land inside protected areas omitted. All statistical and computational processing was performed using the `R` language for statistical computing (version 3.4.1; R Core Team, 2016). Spatial analyses were conducted using `ArcMap` (version 10.3), `python` (version 2.7), and the `sp` (Pebesma & Bivand, 2005) and `raster R` (Hijmans, 2016) packages.

We obtained maps describing the extent of suitable habitat for most of the world's amphibian (Ficetola *et al.*, 2015), avian (Beresford *et al.*, 2011; Buchanan *et al.*, 2011), and mammalian species (excluding marine mammals Rondinini *et al.*, 2011). These maps were produced using the species' global distribution maps (BirdLife International & NatureServe, 2017) and removing areas that are outside the elevational limits of each species or that contained unsuitable land cover (determined from the habitat coded as suitable in their IUCN Red List assessments). All maps were reprojected to an equal-area coordinate system (World Behrman; `ESRI:54017`) and standardized to the same spatial grid by calculating the proportion of suitable habitat in each pixel (5 $\times$ 5 km resolution). Although we were unable to accommodate the migration patterns of migratory mammals into our analysis due to insufficient data, we accommodated the migratory patterns of 1,834 bird species into our analysis by processing and analyzing each of their breeding, non-breeding, and passage distributions separately (following Runge *et al.*, 2015). Specifically, we treated bird species as migratory if they were associated multiple seasonal distributions (following BirdLife International & NatureServe, 2017). We then used the most poorly covered seasonal distribution to assess the overall level of protection for each migratory bird species.

We split the species' habitat maps into climatic partitions using global climatic data (19 layers at 30″ resolution; obtained from `http://worldclim.org`; Hijmans *et al.*, 2005). To achieve this, we initially aggregated and resampled the climatic maps to match the spatial extent and resolution of the species' habitat maps. For each species, we then overlaid the climatic layers with their habitat map, masked the climatic data to places that contained suitable habitat, and subjected the masked climatic data to a principal components analysis. On average, the first two principal components explained 95.98 % $\pm$ 4.25 SD of the climatic variation inside each species' extent of suitable habitat (Table S1). Next, we used the scores from the first two principal components to split the habitat maps into multiple climatic partitions (see below for rationale). Specifically, we used the *k*-means cluster analysis to split each species' distribution into 10 climatic partitions (see below for rationale).

Although more advanced clustering analyses are readily available, we used the *k*-means cluster analysis to create climatic partitions for each species. This is because the *k*-means cluster analysis would split the species' niches into multiple spherical sub-units that each have roughly the same niche volume. By representing habitat associated with each of these sub-units (i.e. the species' climatic partitions), a prioritization is guaranteed—to some extent depending on the number of clusters—to spread out conservation effort across a species' niche. As a consequence, more advanced clustering methods that can identify non-spherical clusters (Ester *et al.*, 1996) or an optimal number of clusters (Fraley & Raftery, 2002) would not be well-suited because a prioritization based on these clusters would not necessarily spread out conservation effort evenly across the species' niches.

5.3.2 Analysis

We created representation targets for each species based on the unpartitioned and the partitioned habitat maps. The targets for the unpartitioned habitat maps were set following standard practices in global gap analyses and prioritizations (Rodrigues *et al.*, 2004; Runge *et al.*, 2015). Briefly, species with less than 1,000 km^2 of suitable habitat were assigned a 100 % target for their unpartitioned map (1,802 amphibian, 966 bird, 645 mammalian species), species with more than 250,000 km^2 were assigned a 10 % target for their unpartitioned map (34 amphibian, 338 bird, 110 mammalian species). Species with an intermediate amount of suitable habitat were assigned a target by log-linearly interpolating values between the previous two thresholds (3,467 amphibian, 11,383 bird, 4,445 mammalian species, note that migratory bird species were assigned targets for each seasonal distribution so the numbers reported here and in the previous sentence do not sum to the total number of species investigated in the study). Additionally, to prevent species with incredibly large range sizes from requiring excessively large amounts of area to be protected, the targets for species with range sizes larger than 10,000,000 km^2 were capped at 1,000,000 km^2. These caps only affected 206 (1.03 %) of the species in our analysis. After calculating distribution-level

targets for the species, we then calculated partition-level targets for the species. For a given species, the targets for their partitioned maps were set by multiplying the target for the species' unpartitioned map by the amount of suitable habitat in each partition. Thus the total amount of habitat required for each species' geographic distribution did not differ between the targets set for the partitioned and unpartitioned maps—though setting targets to capture a representative sample of each species' niche will likely increase the size of the solution (Pressey & Logan, 1994).

We generated prioritizations using the minimum set formulation of the reserve selection problem (Rodrigues *et al.*, 2000). First, we generated a prioritization using the species' partitioned habitat maps and their corresponding targets to identify priority areas for conservation. Second, we generated a prioritization using the species' unpartitioned habitat maps and their corresponding targets to understand how failing to account for species' niches can alter priorities. Both prioritizations were solved to within 10 % of optimality using `Gurobi` (version 7.5.1) and the `prioritizr R` package (Hanson *et al.*, 2017d).

We calculated how well each species is represented in (i) the global protected area system, (ii) the prioritization based on the partitioned habitat maps, and (iii) the prioritization based on the unpartitioned habitat maps. Firstly, we assessed which species are adequately represented under each of the three maps using the species' partitioned habitat maps and their corresponding targets. Secondly, we assessed which species are adequately represented under each of the three maps according to the species' unpartitioned habitat maps and their corresponding targets.

5.4 Results and Discussion

Protected areas are failing to represent the climatic conditions found across species' niches (Figure 5.1). Alarmingly, 18,688 species do not have all of their climatic partitions adequately represented in protected areas, and 8,921 species have less than 10 % of their climatic partitions adequately represented in protected areas. Of these species, 2,266 are listed as globally threatened on the International Union for Conservation of Nature Red List (IUCN, 2017). Although challenges remain in detecting local adaptations (Hoban *et al.*, 2016), some of these imperiled species show signs of local adaptation across their geographic range. For example, foothill populations of the globally vulnerable Italian agile frog (*Rana latastei*) seem to exhibit adaptations to their colder surroundings that are not present in lowland populations (Ficetola & Bernardi, 2005). With none of its climatic partitions adequately protected, further deforestation could limit the ability of this frog to adapt to changing conditions. Further east, populations of the globally vulnerable giant panda (*Ailuropoda melanoleuca*) show marked geographic variation in parts of their genome that underpin disease resistance (Chen *et al.*, 2013). Despite 44.2 % of the panda's suitable habitat laying inside protected areas, we

found that none of its climatic partitions are adequately covered. On average, only 37.01 % of the habitat in each of its climatic partitions currently overlap with protected areas—far less than the 88.47 % target required for each partition—meaning that populations with better disease resistance could be at risk. With such poor niche representation in protected areas, it is crucial that future protected areas are established in places that will have the greatest positive impact.

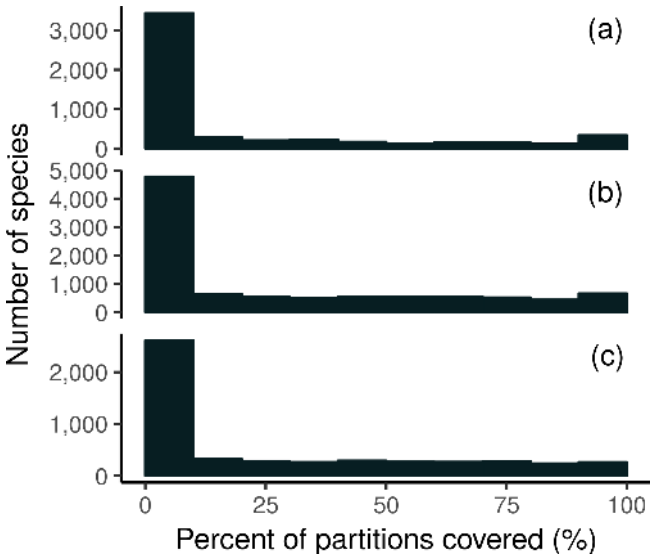


Figure 5.1 The percent of climatic partitions adequately represented in protected areas for each species.

We identified priority areas for protected area expansion to represent species in suitable habitats that span across their niches (Figure 5.2a). These priority areas encompassed 34.1 % of the Earth's land—far exceeding the 17 % target mandated in the Strategic Plan for Biodiversity 2011–2020 (Aichi Target 12). They likely cover a significant proportion of the Earth's remaining habitat given the staggering amount of habitat that has been lost in recent history (Hansen *et al.*, 2013; Venter *et al.*, 2016). Not surprisingly, some of the highest concentrations of priority areas (shown in red; Figure 5.2a) are located in global biodiversity hotspots (Myers *et al.*, 2000). The Andes, for example, stands out as critical for protected area expansion. Here, steep environmental gradients and complex topographies have ignited evolutionary processes (Fjeldså & Irestedt, 2009), resulting in altitudinally zoned habitats that feature unparalleled levels of biodiversity (Myers *et al.*, 2000). The Cape Floristic Region—a floral diversity hotspot located on the southernmost tip of Africa with a distinct temperate Mediterranean-type climate (Verboom *et al.*, 2009)–and Madagascar—brimming with intra-specific genetic variation (Boumans *et al.*, 2007)—are evolutionary powerhouses that also contain many priority areas. Further east, a large number of priority areas are sited along the Himalayan mega-mountain system—an evolutionary epicenter for Chinese bird species (Lei *et al.*, 2015)—and throughout the species rich island nations of Southeastern Asia and the Pacific. Outside of biodiversity hotspots, many more priority areas are located across Africa, Asia, Europe, and Northern America. Although these priority areas mostly oc-

cur in smaller concentrations (shown in green; Figure 5.2a), they are still needed to secure a comprehensive coverage of species' niches.

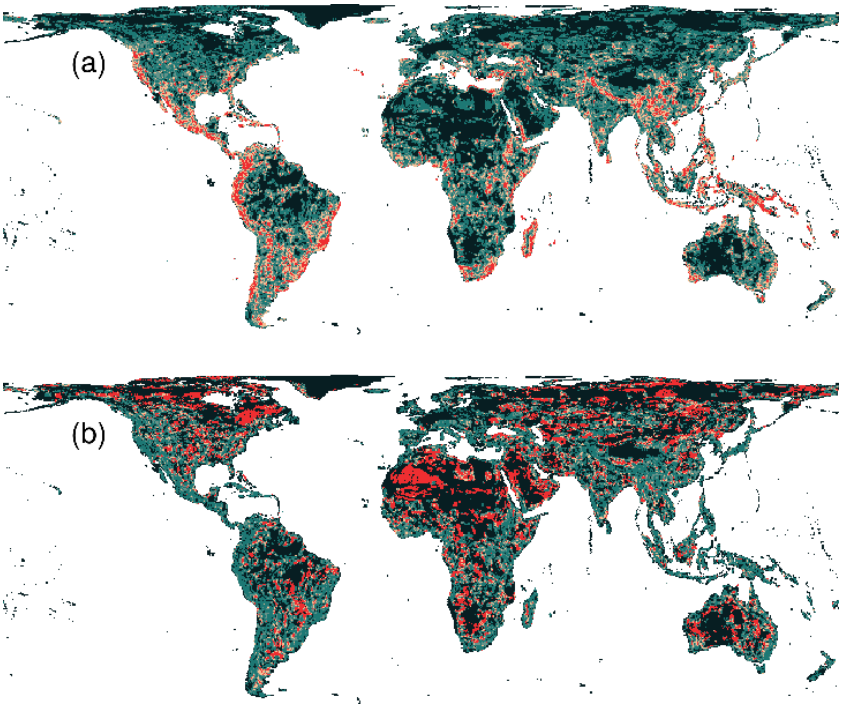


Figure 5.2 Maps show (a) the spatial prioritization for expanding the current global protected area system to represent species' niches and (b) the fraction of priority areas in this prioritization that are missing when intra-specific variation is not used to guide reserve selection. To aid visual interpretation, data show the proportion of areas within 50 km^2 cells selected for protection. Colors correspond to 0 % (■), 1–25 % (■), 25–50 % (■), 50–75 % (■), and 75–100 % (■).

Most prioritizations, to date, have failed to account for species' adaptive potential when identifying priority areas for expanding protected area systems (Carvalho *et al.*, 2011b; Crandall *et al.*, 2000). To understand how accounting for evolutionary processes alters conservation priorities, we created a second prioritization following conventional approaches using the species' unpartitioned maps. This second prioritization encompassed 30.5 % of the Earth's land (Appendix Z) and is comparable to previous global prioritizations (Rodrigues *et al.*, 2004). Although this prioritization contained 82.5 % of the priority areas identified in the first prioritization (shown in green; Figure 5.2b), it also omitted 328,319 of the priority areas identified in the first prioritization (shown in red; Figure 5.2b). As a consequence, the second prioritization failed to adequately represent all of the climatic partitions associated with 7,989 of the 19,937 species.

5.5 Conclusion

Our results show clearly that Article 8 of the Convention of Biological Diversity has been insufficiently implemented to conserve evolutionarily significant habitats. Moving forward, one strategy to address this shortfall would be to set targets under the Strategic Plan for Biodiversity 2021–2030 for the protection of evolutionary processes (Laikre, 2010). Over the

past seven years, the targets set under the Strategic Plan for Biodiversity 2011–2020 have rallied global conservation efforts. For instance, Aichi Target 12, which simply requires that at least 17 % of the Earth's land and 10% of its marine areas are conserved by 2020, has led to the establishment of thousands of new protected areas (H.M. *et al.*, 2015). But we caution that targets must not be too simplistic, noting that Aichi Target 12 has encouraged the establishment of protected areas in places that offer little benefit (Butchart *et al.*, 2012; H.M. *et al.*, 2015). Therefore, it is crucial that targets created for the protection of evolutionary processes are readily understood, firmly grounded in scientific consensus, and not too simple that they enable in perverse outcomes (Maxwell *et al.*, 2015).

As a basis for developing targets to protect evolutionary processes, there are several important considerations. First, some habitats may be conserved through "other effective area-based conservation measures" (Aichi Target 12), such as indigenous reserves and community-managed areas outside the formal protected area system. Although we were unable to analyze these areas because comprehensive global-scale data sets are not yet available (Gannon *et al.*, 2017), future targets should continue to recognize their importance. Second, different habitats are evolutionarily significant for different species. We recognized this by splitting species' distributions into multiple climatic partitions, though more accurate and robust methods are available for narrower geographic and taxonomic scales (Hanson *et al.*, 2017b). Third, different habitats are important for different evolutionary processes, including habitats that facilitate gene flow or provide refuge from rapidly changing climates (Sgro *et al.*, 2011). Fourth, both genetic and habitat factors are important, and although techniques do not yet exist to assess genetic diversity cheaply enough for large-scale deployment, there is some evidence that environmental conditions can be effective surrogates for genetic variation (Hanson *et al.*, 2017a). To conserve evolutionary processes, international policy must motivate the conservation of existing genetic variation (Laikre, 2010) and also the habitats that foster genetic variation (Moritz, 2002).

The Convention on Biological Diversity highlights the importance of conserving habitats that promote evolutionary processes. Despite this, we found very little evidence to suggest that these habitats are adequately protected. We show that plans for expanding protected area systems must account for evolutionary processes, otherwise they risk failing to conserve them. We call on policy makers and conservation practitioners to pay greater attention to evolutionary processes by considering species' environmental niches, so that today's species are given the best chance to survive tomorrow's challenges.

5.6 Acknowledgements

JOH was supported by an Australian Government Research Training Program (RTP) Scholarship. RAF had an Australian Research Council Future Fellowship.

CHAPTER 6

General discussion

This chapter presents a synthesis of this thesis. It integrates the key findings, discusses limitations, and highlights avenues for future research.

6.1 Overview

The overarching goal of conservation is to maximize the long-term persistence of biodiversity and the processes that sustain it (Moritz, 2002). However, as outlined in **Chapter 1**, evolutionary processes—the processes that manufacture new patterns of biodiversity (reviewed in Sobel *et al.*, 2010)—are rarely considered during the evaluation of existing protected areas or during the creation of plans for new protected areas. As a consequence, protected areas may fail to conserve key evolutionary habitats (Cowling & Pressey, 2001; Sgro *et al.*, 2011), halt the erosion of genetic diversity (Frankel, 1974; Ponce-Reyes *et al.*, 2014), or maintain levels of gene flow among populations (Hedrick, 1995; McRae *et al.*, 2012). Here, in this thesis, I have presented novel research and built a greater understanding of how evolutionary processes can be incorporated into conservation planning. I developed (**Chapter 2**) and evaluated (**Chapters 3 and 4**) methods for generating spatial prioritizations that foster evolutionary processes. Building on these ideas, I also assessed how well the global protected area system may be conserving species' adaptive evolutionary processes (**Chapter 5**). This thesis provides new tools, understanding, and directions to safeguard the future patterns of biodiversity.

6.2 New methods

6.2.1 Research summary

Conservation planners need a decision support tool that enables them to explicitly target intra-specific variation. In **Chapter 2**, I developed the `raptr R` package to fulfill this need. The `raptr R` package, unlike previous attempts (e.g. Faith & Walker, 1996; Bonin *et al.*, 2007; Diniz-Filho *et al.*, 2012), can account for acquisition costs, species' habitat requirements, reserve fragmentation, and any continuous intra-specific variation of interest—all in a multi-species conservation planning context. I applied the `raptr R` package to various simulated and case-study species. I showed that prioritizations that are not designed to target intra-specific variation—such as patterns of environmental or genetic variation found across species' geographic ranges—may fail to site reserves in places that would conserve it. By building protected areas that capture intra-specific patterns of biodiversity, conservation efforts stand a greater chance of achieving the long-term persistence of biodiversity.

The `raptr R` package uses novel formulations of the reserve selection problem to generate prioritizations. These formulations aim to minimize the cost of the total reserve system whilst ensuring that solutions adequately protect each feature. When using the `raptr R` package, users set amount-based targets to ensure that an adequate amount of habitat is protected for each feature. Users also set space-based targets to ensure that the solution conserves

a minimum level of the intra-specific variation across the features' geographic distributions. Additionally, users can set fragmentation penalties to increase connectivity in the solution. The `raptr R` package is able to provide this functionality by building on ideas from decision support tools that have come before it (Faith & Walker, 1996; Ball *et al.*, 2009) and can accommodate uncertainty using recent advances in optimization research disciplines (Cui *et al.*, 2010).

6.2.2 Limitations

The `raptr R` package is not without limitations. Foremost, the `raptr R` package was developed to enhance the functionality of the `R` statistical computing environment (R Core Team 2016; see R Core Team 1999 for more information on `R` packages), and as a consequence, users will require some proficiency in computer programming to use the `raptr R` package effectively. Although this limitation means that some potential users will require training before they can use the tool (Pevzner & Shamir, 2009; Barraquand *et al.*, 2014), there are many associated benefits with developing this tool as an `R` package. These benefits include the fact that (i) existing `R` users will be able to use it with minimal training, (ii) users will be able to generate prioritizations in a manner that is reproducible and transparent (Peng, 2011; Sandve *et al.*, 2013), (iii) programming languages offer unparalleled flexibility to customize analyses, and (iv) there is a rich ecosystem of contributed `R` packages that users can combine with the `raptr R` package to tackle complex problems (see the `CRAN` task views for curated topics; `https://cran.r-project.org/web/views`).

The second main limitation is the fact that the `raptr R` package uses the commercial exact algorithm solver `Gurobi` to generate prioritizations. This design choice is, perhaps, the most limiting because although academics have access to a special license to use `Gurobi` for no cost, the licenses available for governmental and non-governmental organization are currently expensive (e.g. single user licenses cost $7,500 USD; see `http://www.gurobi.com/products/licensing-pricing/commercial-pricing`). Despite this, building on the `Gurobi` platform means that the `raptr R` package can solve much larger problems than could otherwise be achieved during the course of a single PhD project. An earlier version of the `raptr R` package implemented a manually coded algorithm to solve problems (simulated annealing; Kirkpatrick *et al.*, 1983), but upon testing this version, it was soon evident that it could not solve even moderately sized problems within a feasible period of time. Indeed, Beyer *et al.* (2016) recently demonstrated that commercial exact algorithm solvers need to be used to solve even relatively simple conservation planning problems. The capabilities of the `Gurobi` platform have improved enormously over the last few years (Mittelmann, 2017), and by building on this platform, users will experience substantial improvements to the `raptr R` package in the coming years.

One of the potential hurdles in using the `raptr R` package pertains to the selection of de-

mand points used to generate prioritizations. Demand points represent regions in an attribute space—for instance, specific environmental conditions in a climatic attribute space—that are desirable for conservation. The concept of using demand points to guide reserve selection originated during the mid-1990s—with the environmental diversity (ED) approach (Faith & Walker, 1996). Since then, there has been considerable debate concerning how they should be generated. On the one hand, Faith (2003) advocated that demand points should reflect the underlying variation that needs to be conserved (termed the "continuous ED" approach). On the other hand, Araújo *et al.* (2001) generated demand points based on the range of options that are available for conservation (termed the "discrete ED" approach). Since prioritizations based on demand points from continuous ED approaches differ from those based on discrete ED approaches (Faith, 2003, 2011; Beier & Albuquerque, 2015; but see Hortal *et al.*, 2009), it is important that conservation planners use the correct approach for generating demand points—but what is the correct approach?

It is my view that the correct approach for generating demand points is entirely context dependent. For instance, imagine a conservation planning scenario involving hundreds of species and, due to budgetary constraints, only a small fraction of the environmental variation found across each species' geographic distribution can be conserved. In this scenario, the discrete ED approach for generating demand points might be more useful because (i) this approach is sufficient when only a small fraction of intra-specific variation needs to be conserved and (ii) this approach will require fewer demand points which, in turn, will reduce the time needed to solve the reserve selection problem. On the other hand, imagine a scenario involving a few threatened species. Although these species once occurred across a range of patches—including high quality and low quality habitats—these species are now only found in low quality habitats. Here, the continuous ED approach may be more useful. By using demand points that reflect the full range of suitable conditions and including those that no longer exist—i.e. the species' fundamental niche—the solution could emphasize habitats that have environmental conditions that are most similar to the lost high quality patches. Thus conservation planners will need to have a detailed understanding of how demand points affect solutions, and should construct demand points in a manner that will fulfill conservation objectives.

6.2.3 Future research directions

Exact algorithms have much to offer conservation planning. In the past, they have generally been regarded as having little utility for conservation planning because they could not solve large problems (Pressey *et al.*, 1996). Today, however, exact algorithms are able to solve much larger and more complex problems than other methods (Billionnet, 2013; Beyer *et al.*, 2016). Indeed, we are beginning to see their reappearance in the literature (e.g. Stigner *et al.*, 2016; Law *et al.*, 2017; Schuster *et al.*, 2017). Future research should aim

to develop more decision support tools that use exact algorithms to solve reserve selection problems. These tools have the potential to be highly customizable—as evidenced by broad range constraints discussed in Billionnet (2013)—so that conservation planners can create prioritizations fulfilling their specific objectives. Concerted efforts by governmental and conservation organizations to partner with commercial organizations that develop exact algorithm solvers (e.g. IBM and Gurobi Optimization Inc) might help to provide conservation planners with access to these solvers when working outside of research institutions.

6.3 Surrogates for genetic data

6.3.1 Research summary

Conservation planners need surrogates for preserving evolutionary processes in spatial prioritizations. Ideally, they would have access to sufficient genetic data to generate spatial prioritizations, but genetic data require extensive resources to obtain and analyze. In broad-scale conservation planning exercises, involving hundreds of species (e.g. Klein *et al.*, 2009a; Fuller *et al.*, 2010; Buchanan *et al.*, 2011; H.M. *et al.*, 2015), it is simply not feasible to collect genetic data for all species of interest. To investigate potential surrogates, I obtained genetic data for 27 alpine plant species from the IntraBioDiv project (Taberlet *et al.*, 2012a,b). I then generated prioritizations using potential surrogate data, and evaluated their performance using the genetic data. In **Chapter 3**, I showed that environmental and geographic data could be used to guide the reserve selection process and generate prioritizations that conserve intra-specific adaptive and neutral genetic variation. I also found that prioritizations that do not spread reserves evenly across a species' geographic distribution may fail to conserve patterns of neutral variation. In **Chapter 4**, I showed that conventional approaches for increasing connectivity—clustering priority areas together and representing species in contiguous areas of suitable habitat—might not be effective strategies for maintaining gene flow. However, by representing species in contiguous areas that offer little impediment to gene flow, I was able to generate prioritizations that were capable of facilitating strong levels of gene flow. These findings demonstrate that fundamental concepts in evolutionary biology can be operationalized to deliver more effective plans for protected areas.

6.3.2 Limitations

The work conducted in **Chapters 3 and 4** has several limitations. These limitations largely stem from the kinds of species used in the analyses, and how the findings reported in these chapters may not generalize to other species or biomes. Below, I discuss the implications of these limitations in more detail.

My analyses did not include any species known to be at serious risk of extinction. As a consequence, the effectiveness of these surrogates might not extend to threatened species. This is because the spatial patterns of genetic variation in threatened species may reflect particularly recent—and not historic—events in their evolutionary history. For instance, the remaining populations of the endangered New Zealand mōhua (*Mōhoua ochrocephala*) show remarkably high levels of spatial structure in their genetic variation due to recent extirpations and bottlenecks (Tracy & Jamieson, 2011). This means that the broad-scale patterns of adaptive and neutral genetic variation in the New Zealand mōhua might not correlate with broad-scale patterns of environmental and geographic variation across its remaining distribution. But just because these surrogates may not be effective for threatened species does not mean that they are useless. The protection of functional habitats and species with stable population trajectories are an important part of modern conservation (Possingham *et al.*, 2002; Bottrill *et al.*, 2008; Runge *et al.*, 2015). For instance, many conservation organizations have proactive strategies that seek to establish protected areas in places before they experience substantial degradation (Brooks *et al.*, 2006). Overall, these results provide insight into conserving evolutionary processes where the goal is to maintain evolutionary processes in species where they have not been substantially disrupted. Future work is needed to verify the effectiveness of these surrogates when attempting to recover evolutionary processes that have been disrupted.

Although this data set is one of the most formidable genetic data sets available—in terms of its spatial coverage and diverse array of species (Gugerli *et al.*, 2008)—it is still restricted to a single biome. It would seem likely that the environmental surrogate variables would be most effective in biomes such as this, which contain steep environmental gradients that foster local adaptation, and might be less effective in more homogenous biomes. Additionally, the geographic distance surrogate variables would be expected to be most effective in systems where dispersal barriers—such as mountain ranges—occur regularly throughout the study area, and not in places where they occur sporadically throughout the landscape (e.g. in systems where highways are significant barriers to dispersal; Dudaniec *et al.*, 2016). Nevertheless, my findings pave the way for incorporating evolutionary processes that operate at the intra-specific level into conservation planning.

6.3.3 Future research directions

The paucity of genetic data available for conservation planning and the need to make urgent and cost effective decisions in conservation (Barnosky *et al.*, 2011) have driven the search for effective surrogates for genetic data. Ideally, surrogates should not be needed and conservation planners should utilize the best quality data to guide their decisions. Although the costs associated with processing genetic material are decreasing (e.g. Abdelkrim *et al.*, 2009), the costs associated with obtaining genetic data remain a substantial barrier to its

widespread usage. In stark contrast to genetic data, geographic distribution data are commonly used to guide conservation planning (e.g. Kiester *et al.*, 1996; Rodrigues *et al.*, 2004; Fuller *et al.*, 2010). This results, in part, from a wealth of freely available distribution data online—for example, from The Atlas of Living Australia (ALA; `https://www.ala.org.au`) and The Global Biodiversity Information Facility (GBIF; `http://www.gbif.org`)—that are fed by citizen science projects—such as eBird (`http://ebird.org`), and iNaturalist (`https://www.inaturalist.org`)—that facilitate curation of data collected by the public at large. Therefore, one potential solution is to develop similar data repositories and citizen science projects for genetic data collection and storage.

Recently, there have been several initiatives designed to host genetic data and engage citizen scientists in its collection. For instance, The Genomic Observatories Metadatabase (GeOMe; `http://www.geome-db.org`) allows researchers to deposit genetic data (Deck *et al.*, 2017). This database also stores geographic and ecological information along side genetic data sets to help researchers find genetic data relevant to their study area or taxon. Since developing plans for protected area establishment is inherently spatial, this feature will be very useful for conservation planners. Another example, The Pieris Project uses citizen scientists to collect cabbage white butterfly samples (*Pieris rapae*). These samples are then sent to geneticists for processing, and the resulting data is used monitor the evolution of this butterfly across space and time (`http://www.pierisproject.org`). To date, this project has received over 2,000 samples from over 32 different countries (Pieris Project, 2017). Citizen scientists need not actually capture wild animals. For instance, volunteers could sample genetic material passively from the surrounding environment given recent advances in environmental DNA (eDNA; reviewed in Bohmann *et al.*, 2014). Until the costs of obtaining genetic data are substantially reduced, genetic data will rarely be used to inform conservation planning.

The tangible conservation benefits gained from genetic data are often perceived to be low, but this is incorrect (Hendry *et al.*, 2011). One source for this misconception is that conservation genetics studies have traditionally failed to provided results in formats that are compatible with decision support tools. For instance, many conservation genetics studies endeavor to establish evolutionarily significant units (ESUs; Moritz, 1994), and yet they fail to publish spatially referenced data that actually delineate them (e.g. shapefiles and accompanying metadata). Similarly, many conservation genetics studies report statistics pertaining to genetic diversity (e.g. F_{st} statistics; reviewed in Holsinger & Weir, 2009), but they do not delineate the spatial extent to which the statistics pertain. Although significant methodological advances are being made to rectify these issues (Shryock *et al.*, 2015; Tarroso *et al.*, 2015), they have not been implemented in decision support tools and are not yet easily integrated into existing conservation decision frameworks.

Currently, conservation scientists and practitioners have little justification to prioritize the collection of genetic data. This is because very few studies have empirically shown that the

information gained from genetic data can substantially improve conservation decision making processes (but see Carvalho *et al.*, 2011b; Ponce-Reyes *et al.*, 2014). To address this, future studies could conduct value of information analyses (reviewed in Canessa *et al.*, 2015) to understand how genetic data sets of varying quality and completeness could potentially improve conservation planning. For instance, using these methods, Grantham *et al.* (2008) demonstrated that short-term surveys had great potential to improve conservation decisions despite diminishing returns. Additionally, if genetic data have been used to inform conservation decisions in the past, impact assessments could be used to verify their utility (reviewed in Ahmadia *et al.*, 2015). These assessments could use counter factual data—pertaining to places or periods of time where genetic data were absent (Ferraro, 2009)—to estimate the benefits for conservation actions (e.g. Claudet *et al.*, 2006; Barnes *et al.*, 2015). Overall, further studies are needed to determine whether genetic data should be used in particular conservation decision making processes and evaluate their cost-effectiveness.

6.4 Evaluating the conservation of adaptive processes

6.4.1 Research Summary

The Convention on Biological Diversity acknowledges the importance of preserving areas that foster evolutionary processes. However, very little is known about how well protected areas are achieving this goal. One strategy to promote a species' capacity to adapt to changing conditions is to protect populations in a climatically diverse set of areas—a representative sample of the species' realized niche (Hutchinson, 1957)—to maximize the range of potential local adaptations that may arise (Moritz, 2002). Based on this idea, I measured how well protected areas are conserving species' realized niches in **Chapter 5**. I found that protected areas are failing to protect species' realized climatic niches. This work is the first global assessment of how well species' realized niches are being conserved, and identifies priority areas for conserving evolutionary processes.

6.4.2 Limitations

In the work detailed in **Chapter 6**, I used broad-scale climatic variation to map the range of selection pressures that species are currently experiencing (using data obtained from Hijmans *et al.*, 2005). This approach does not account for that fact that there are many other dimensions to species' realized niches—such as biotic interactions (Araújo & Luoto, 2007; Susanne *et al.*, 2013)—that can shape their evolution (e.g. Ratcliffe & Nydam, 2008; Mochida, 2011). For instance, highly toxic invasive cane toads (*Bufo marinus*) have caused susceptible Australian snakes (e.g. *Pseudechis porphyriacus* and *Dendrelaphis punctulatus*) to evolve larger body sizes and relatively smaller head sizes (Phillips & Shine, 2004).

These traits are thought to be advantageous because larger sized individuals have a greater ability to process toxins and individuals with smaller gapes are unable to consume larger sized toads (Phillips & Shine, 2004). Additionally, my approach did not account for variation in micro-habitat conditions across species' geographic distributions. Organisms can use micro-habitat conditions to escape broad-scale selection pressure (Milanovich *et al.*, 2017), and so populations living in areas with broadly different climate regimes may experience similar selection pressures (Rull, 2009, 2010; Dobrowski, 2011). In most cases though, it seems likely that my results overestimate how well protected areas are sampling the range of selection pressures that are shaping species' evolutionary futures.

6.4.3 Future research directions

The effectiveness of international conservation policy could be enhanced by including well-defined targets for conserving evolutionary processes. The Conventional on Biological Diversity currently focuses on the conservation of genetic diversity in managed or cultivated populations (Aichi Target 13). The only mention of conserving evolutionary processes across the broader landscape pertains to the protection of habitats containing "key evolutionary processes" (Article 8). However, given the importance of conserving genetic diversity (Reed & Frankham, 2003), this is simply not enough (Laikre, 2010; Laikre *et al.*, 2010). It is becoming increasingly apparent that new protected areas are being established—as mandated in Aichi Target 11—in places that offer little benefit to biodiversity (Di Minin & Toivonen, 2015). Research is desperately needed to provide targets that pertain to evolutionary processes in a way that can be readily incorporated into the Strategic Plan for Biodiversity 2021–2030.

The threat status of species recognized under the International Union for Conservation of Nature (IUCN) has important ramifications for their conservation (Rodrigues *et al.*, 2006). Currently, the IUCN guidelines for listing species at serious risk of extinction contain very little reference to the species' evolutionary processes (Mace *et al.*, 2008; Laikre, 2010). Although the utility of incorporating genetic data into threat listing processes has been demonstrated (e.g. Reed & Frankham, 2003), this has yet to be achieved. Perhaps one reason for this is that Red List criteria require clear and well-defined thresholds—similar to the geographic range size thresholds (IUCN, 2001)—and determining appropriate thresholds can be difficult. However, the thresholds required for international policy may indeed exist in nature with evidence of threshold relationships existing between extinction risk and inbreeding (e.g. Frankham, 1995). Conservation scientists will rarely have enough ecological or genetic data to estimate species' extinction risks with a high degree of precision (Coulson *et al.*, 2001), and so estimates of extinction risk often have to be made with the best available data.

6.5 Research perspectives

This thesis reveals that conservation planning methodologies must account for evolutionary processes to protect biodiversity. My research chapters (**Chapters 2–5**) demonstrated that prioritizations need to be designed to capture evolutionary processes. This result held true in prioritizations generated for simulated species (**Chapter 2**), alpine plant species (**Chapters 3 and 4**), and nearly every single terrestrial vertebrate species on the planet (**Chapter 5**). Additionally, this thesis reveals that, in some cases, prioritizations may be able to capture evolutionary processes for relatively little additional cost. This finding was evident in my chapters investigating genetic diversity (**Chapters 2–3**) and the global protected area system (**Chapter 5**). All of my research chapters highlighted the importance of incorporating evolutionary processes into conservation planning, either through genetic or surrogate data sets.

Genetic data offer many insights into species' evolutionary histories that can guide conservation decisions. In this thesis, I investigated potential surrogates for representing genetic diversity and maintaining gene flow in prioritizations—just two potential applications of genetic data. As discussed earlier, I found evidence that freely available surrogate data could be used to capture intra-specific genetic diversity (**Chapter 3**), and that conventional strategies for increasing connectivity in prioritizations may not actually maintain gene flow (**Chapter 4**). These results may seem conflicting. However, each of these chapters deals with a different application of genetic data, a different set of surrogate data sets, and a different metric for assessing solution quality. It may simply be that conservation planners may be able to utilize surrogate data for some applications of genetic data and not other applications. Finally, though I investigated genetic diversity and gene flow separately, real-world conservation decisions should ideally be based on multiple applications of genetic data—e.g. using maps showing patterns of genetic diversity, resistance to gene flow, and refugial habitats in combination—to account for all of the evolutionary processes that shape future patterns of biodiversity.

This thesis provides insights into how conservation planning exercises can incorporate evolutionary processes. Firstly, conservation planners can formally recognize and target intra-specific variation. This means that when managers are evaluating candidate sites for expanding a local-scale protected area system, they can use the tools developed here to evaluate how well the populations in the candidate sites complement those already represented in protected areas—both at the local and regional scales. Ideally, managers would have access to the genetic data needed to determine if any of the candidate sites contain populations which exhibit local adaptations or populations have markedly different evolutionary histories compared to those already inside protected areas. Although most real-world conservation planning exercises do not have access to genetic data, the research presented in this thesis suggests that candidate sites may be valuable for enhancing evolutionary processes if they

contain populations in distinct environmental settings compared to the populations already in protected areas, or contain populations that are geographically distant from populations already in protected areas (**Chapter 3**). Secondly, conservation planners can incorporate the fact that species have different dispersal capabilities (**Chapter 4**). In the context of real-world conservation planning (e.g. by national or regional governments), this means that when designing corridors to link up large protected areas it is important to ensure that the habitat present in the corridors can facilitate dispersal for all species of conservation interest. Otherwise, if the corridors only contain a single type of habitat then the corridors will only restore gene flow for the subset of species which can disperse through this habitat type. Thirdly, conservation planning exercises need to formally recognize that profoundly more areas are needed to preserve biodiversity. This thesis found that approximately 34.1 % of the Earth needs to covered by protected areas to adequately represent nearly twenty thousand vertebrate species (**Chapter 5**). Although strategically managing land outside of protected areas can benefit biodiversity, ultimately, more land needs to be set aside for nature.

6.6 Concluding remarks

Conservation actions must foster evolutionary processes to halt further declines in global biodiversity. This thesis developed a new decision support tool to help conserve evolutionary processes. It also investigated strategies for conserving evolutionary processes when genetic data are not available. Finally, it provided the first global assessment for how well protected areas are representing the adaptive evolutionary processes that will shape the future of biodiversity. The next research steps are to enhance this tool, verify the effectiveness the surrogates in other biomes, and further understand the link the between conserving species' adaptive processes and their long-term persistence.

References

Abdelkrim, J., Robertson, B.C., Stanton, J.A.L. & Gemmell, N.J. (2009) Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. *BioTechniques*, **46**, 185.

Ahmad, M.S.A., Ashraf, M. & Ali, Q. (2010) Soil salinity as a selection pressure is a key determinant for the evolution of salt tolerance in Blue Panicgrass (*Panicum antidotale Retz.*). *Flora - Morphology, Distribution, Functional Ecology of Plants*, **205**, 37 – 45.

Ahmadia, G.N., Glew, L., Provost, M., Gill, D., Hidayat, N.I., Mangubhai, S., Purwanto & Fox, H.E. (2015) Integrating impact evaluation in the design and implementation of monitoring marine protected areas. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **370**, 20140275.

Aiello-Lammens, M.E., Boria, R.A., Radosavljevic, A., Vilela, B. & Anderson, R.P. (2015) spThin: An R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography*, **38**, 541–545.

Al-Hiyaly, S.A.K., McNeilly, T., Bradshaw, A.D. & Mortimer, A.M. (1993) The effect of zinc contamination from electricity pylons. Genetic constraints on selection for zinc tolerance. *Heredity*, **70**, 22–32.

Alford, R.A. & Richards, S.J. (1999) Global amphibian declines: A problem in applied ecology. *Annual Review of Ecology & Systematics*, **30**, 133–165.

Allan, J.R., Venter, O., Maxwell, S., Bertzky, B., Jones, K., Shi, Y. & Watson, J.E. (2017) Recent increases in human pressure and forest loss threaten many natural world heritage sites. *Biological Conservation*, **206**, 47–55.

Alroy, J. (2017) Effects of habitat disturbance on tropical forest biodiversity. *Proceedings of the National Academy of Sciences of the United States of America*, **114**, 6056–6061.

Alvarez, N., Thiel-Egenter, C., Tribsch, A., Holderegger, R., Manel, S., Schönswetter, P., Taberlet, P., Brodbeck, S., Gaudeul, M., Gielly, L., Küpfer, P., Mansion, G., Negrini, R., Paun, O., Pellecchia, M., Rioux, D., Schüpfer, F., Van Loo, M., Winkler, M., Gugerli, F. & IntraBioDiv Consortium (2009) History or ecology? Substrate type as a major driver of spatial genetic structure in Alpine plants. *Ecology Letters*, **12**, 632–640.

Antoine, G., Reid, T., B., B.J., Ilona, N.L., R., S.P., T., T.A.I., J., R.T., Lluis, B., Eve, M.M., Chrystal, M.P., G., M.T., R., R.J., Ramona, M., A., S.S., Jane, E., W., S.M., A., W.B., Olivier, B., Mike, A., Simon, F., R., K.M., P., P.H. & M., B.Y. (2013) Predicting species distributions for conservation decisions. *Ecology Letters*, **16**, 1424–1435.

Araújo, M., Humphries, C., Densham, P., Lampinen, R., Hagemeijer, W., Mitchell-Jones, A. & Gasc, J. (2001) Would environmental diversity be a good surrogate for species diversity? *Ecography*, **24**, 103–110.

Araújo, M.B. & Luoto, M. (2007) The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology & Biogeography*, **16**, 743–753.

Arroyo, M.T.K., Primack, R. & Armesto, J. (1982) Community studies in pollination ecology in the high temperate Andes of central Chile. I. Pollination mechanisms and altitudinal variation. *American Journal of Botany*, **69**, 82–97.

Auerbach, N.A., Wilson, K.A., Tulloch, A.I., Rhodes, J.R., Hanson, J.O. & Possingham, H.P. (2015) Effects of threat management interactions on conservation priorities. *Conservation Biology*, **29**, 1626–1635.

Avise, J.C. (2009) Phylogeography: Retrospect and prospect. *Journal of Biogeography*, **36**, 3–15.

Babu, S., Love, A. & Babu, C.R. (2009) Ecological restoration of lantana-invaded landscapes in Corbett Tiger Reserve, India. *Ecological Restoration*, **27**, 467–477.

Ball, I. (2000) *Mathematical applications for conservation ecology: The dynamics of tree hollows and the design of nature reserves*. PhD Dissertation, Departments of Applied Mathematics, Environmental Science and Management, University of Adelaide, Australia.

Ball, I., Possingham, H. & Watts, M.E. (2009) Marxan and relatives: Software for spatial conservation prioritisation. A. Moilanen, K.A. Wilson & H. Possingham, eds., *Spatial conservation prioritisation: Quantitative methods & computational tools*, pp. 185–189. Oxford University Press, Oxford, UK.

Barnes, M. (2015) Aichi targets: Protect biodiversity, not just area. *Nature*, **526**, 195.

Barnes, M., Szabo, J.K., Morris, W.K. & Possingham, H. (2015) Evaluating protected area effectiveness using bird lists in the Australian wet tropics. *Diversity & Distributions*, **21**, 368–378.

Barnosky, A.D., Matzke, N., Tomiya, S., Wogan, G.O., Swartz, B., Quental, T.B., Marshall, C., McGuire, J.L., Lindsey, E.L., Maguire, K.C., Mersey, B. & Ferrer, E.A. (2011) Has the Earth's sixth mass extinction already arrived? *Nature*, **471**, 51.

Barraquand, F., Ezard, T.H., Jørgensen, P.S., Zimmerman, N., Chamberlain, S., Salguero-Gomez, R., Curran, T.J. & Poisot, T. (2014) Lack of quantitative training among early-career ecologists: A survey of the problem and potential solutions. *PeerJ*, **2**, e285.

Barton, N.H. (2000) Genetic hitchhiking. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **355**, 1553–1562.

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**, 1–48.

Bedward, M., Pressey, R.L. & Keith, D.A. (1992) A new approach for selecting fully representative reserve networks: Addressing efficiency, reserve design and land suitability with an iterative analysis. *Biological Conservation*, **62**, 115–125.

Beger, M., Selkoe, K.A., Treml, E., Barber, P.H., von der Heyden, S., Crandall, E.D., Toonen, R.J. & Riginos, C. (2014) Evolving coral reef conservation with genetic information. *Bulletin of Marine Science*, **90**, 159–185.

Beier, P. & Albuquerque, F.S. (2015) Environmental diversity as a surrogate for species representation. *Conservation Biology*, **29**, 1401–1410.

Benjamini, Y. & Hochberg, Y. (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **57**, 289–300.

Beresford, A.E., Buchanan, G.M., Donald, P.F., Butchart, S.H.M., Fishpool, L.D.C. & Rondinini, C. (2011) Poor overlap between the distribution of protected areas and globally threatened birds in Africa. *Animal Conservation*, **14**, 99–107.

Beyer, H.L., Dujardin, Y., Watts, M.E. & Possingham, H.P. (2016) Solving conservation planning problems with integer linear programming. *Ecological Modelling*, **328**, 14–22.

Billionnet, A. (2013) Mathematical optimization ideas for biodiversity conservation. *European Journal of Operational Research*, **231**, 514–534.

BirdLife International & NatureServe (2017) *Bird species distribution maps of the world*. BirdLife International, Cambridge, UK. Available at `http://datazone.birdlife.org`.

Blonder, B., Lamanna, C., Violle, C. & Enquist, B.J. (2014) The *n*-dimensional hypervolume. *Global Ecology & Biogeography*, **23**, 595–609.

Blount, Z.D., Borland, C.Z. & Lenski, R.E. (2008) Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 7899–7906.

Boag, P.T. & Grant, P.R. (1981) Intense natural selection in a population of Darwin's finches (Geospizinae) in the Galápagos. *Science*, **214**, 82–85.

Bohmann, K., Evans, A., Gilbert, M.T.P., Carvalho, G.R., Creer, S., Knapp, M., Douglas, W.Y. & De Bruyn, M. (2014) Environmental dna for wildlife biology and biodiversity monitoring. *Trends in Ecology & Evolution*, **29**, 358–367.

Bonin, A., Nicole, F., Pompanon, F., Miaud, C. & Taberlet, P. (2007) Population adaptive index: A new method to help measure intraspecific genetic diversity and prioritize populations for conservation. *Conservation Biology*, **21**, 697–708.

Bornand, C., Gygax, A., Juillerat, P., Jutzi, M., Möhl, A., Rometsch, S., Sager, L., Santiago, H. & Eggenberg, S. (2016) *Rote Liste Gefässpflanzen. Gefährdete Arten der Schweiz*. Bundesamt für Umwelt, Bern und Info Flora, Genf. Umwelt-Vollzug. Nr. 1621: 178 S.

Bothwell, H., Bisbing, S., Therkildsen, N., Crawford, L., Alvarez, N., Holderegger, R. & Manel, S. (2013) Identifying genetic signatures of selection in a non-model species, alpine gentian (*Gentiana nivalis L.*), using a landscape genetic approach. *Conservation Genetics*, **14**, 467–481.

Bottrill, M.C., Joseph, L.N., Carwardine, J., Bode, M., Cook, C., Game, E.T., Grantham, H., Kark, S., Linke, S., McDonald-Madden, E., Pressey, R.L., Walker, S., Wilson, K.A. & Possingham, H.P. (2008) Is conservation triage just smart decision making? *Trends in Ecology & Evolution*, **23**, 649–654.

Boumans, L., Vieites, D.R., Glaw, F. & Vences, M. (2007) Geographical patterns of deep mitochondrial differentiation in widespread Malagasy reptiles. *Molecular Phylogenetics & Evolution*, **45**, 822–839.

Breiner, F.T., Guisan, A., Nobis, M.P. & Bergamini, A. (2017) Including environmental niche information to improve IUCN Red List assessments. *Diversity & Distributions*, **23**, 484–495.

Brooks, T.M., Mittermeier, R.A., da Fonseca, G.A., Gerlach, J., Hoffmann, M., Lamoreux, J.F., Mittermeier, C.G., Pilgrim, J.D. & Rodrigues, A.S. (2006) Global biodiversity conservation priorities. *Science*, **313**, 58–61.

Bryce, R., Irvine, K.N., Church, A., Fish, R., Ranger, S. & Kenter, J.O. (2016) Subjective well-being indicators for large-scale assessment of cultural ecosystem services. *Ecosystem Services*, **21**, 258–269.

Buchanan, G.M., Donald, P.F. & Butchart, S.H.M. (2011) Identifying priority areas for conservation: A global assessment for forest-dependent birds. *PLoS ONE*, **6**, 1–10.

Bunce, R.G.H. & Shaw, M.W. (1973) Standardized procedure for ecological survey. *Journal of Environmental Management*, **1**, 239–258.

Buri, P. (1956) Gene frequency in small populations of mutant *Drosophila*. *Evolution*, **10**, 367–402.

Butchart, S.H.M., Scharlemann, J.P.W., Evans, M.I., Quader, S., Aricò, S., Arinaitwe, J., Balman, M., Bennun, L.A., Bertzky, B., Besançon, C., Boucher, T.M., Brooks, T.M., Burfield, I.J., Burgess, N.D., Chan, S., Clay, R.P., Crosby, M.J., Davidson, N.C., De Silva,

N., Devenish, C., Dutson, G.C.L., Fernández, D.F.D.z., Fishpool, L.D.C., Fitzgerald, C., Foster, M., Heath, M.F., Hockings, M., Hoffmann, M., Knox, D., Larsen, F.W., Lamoreux, J.F., Loucks, C., May, I., Millett, J., Molloy, D., Morling, P., Parr, M., Ricketts, T.H., Seddon, N., Skolnik, B., Stuart, S.N., Upgren, A. & Woodley, S. (2012) Protecting important sites for biodiversity contributes to meeting global conservation targets. *PloS ONE*, **7**, e32529.

Calenge, C. (2006) The package adehabitat for the R software: Tool for the analysis of space and habitat use by animals. *Ecological Modelling*, **197**, 1035.

Camm, J.D., Polasky, S., Solow, A. & Csuti, B. (1996) A note on optimal algorithms for reserve site selection. *Biological Conservation*, **78**, 353–355.

Canessa, S., Guillera-Arroita, G., Lahoz-Monfort, J.J., Southwell, D.M., Armstrong, D.P., Chadès, I., Lacy, R.C. & Converse, S.J. (2015) When do we need more data? A primer on calculating the value of information for applied ecologists. *Methods in Ecology & Evolution*, **6**, 1219–1228.

Carvalho, S.B., Brito, J.C., Crespo, E.G., Watts, M.E. & Possingham, H.P. (2011a) Conservation planning under climate change: Toward accounting for uncertainty in predicted species distributions to increase confidence in conservation investments in space and time. *Biological Conservation*, **144**, 2020–2030.

Carvalho, S.B., Brito, J.C., Crespo, E.J. & Possingham, H.P. (2011b) Incorporating evolutionary processes into conservation planning using species distribution data: A case study with the western Mediterranean herpetofauna. *Diversity & Distributions*, **17**.

Carvalho, S.B., Velo-Antón, G., Tarroso, P., Portela, A.P., Barata, M., Carranza, S., Moritz, C. & Possingham, H.P. (2017) Spatial conservation prioritization of biodiversity spanning the evolutionary continuum. *Nature Ecology & Evolution*, **1**, 0151.

Carwardine, J., Klein, C.J., Wilson, K.A., Pressey, R.L. & Possingham, H.P. (2009) Hitting the target and missing the point: Target-based conservation planning in context. *Conservation Letters*, **2**, 4–11.

Ceballos, G., Ehrlich, P.R. & Dirzo, R. (2017) Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proceedings of the National Academy of Sciences of the United States of America*, **114**, E6089–E6096.

Ceballos, G., Ehrlich, P.R., Soberón, J., Salazar, I. & Fay, J.P. (2005) Global mammal conservation: What must we manage? *Science*, **309**, 603–607.

Cerdeira, J.O., Pinto, L.S., Cabeza, M. & Gaston, K.J. (2010) Species specific connectivity in reserve-network design using graphs. *Biological Conservation*, **143**, 408 – 415.

Chamberlain, S. (2017) *rgbif: Interface to the Global Biodiversity Information Facility API*. R package version 0.9.7. Available at `https://CRAN.R-project.org/package=rgbif`.

Chen, Y.Y., Zhu, Y., Wan, Q.H., Lou, J.K., Li, W.J., Ge, Y.F. & Fang, S.G. (2013) Patterns of adaptive and neutral diversity identify the Xiaoxiangling Mountains as a refuge for the giant panda. *PloS ONE*, **8**, e70229.

Chernomor, O., Minh, B.Q., Forest, F., Klaere, S., Ingram, T., Henzinger, M. & von Haeseler, A. (2015) Split diversity in constrained conservation prioritization using integer linear programming. *Methods in Ecology & Evolution*, **6**, 83–91.

CIESEN, Columbia University, International Food Policy Research Institute, The World Bank & Centro Internacional de Agricultura Tropical (2011) Global rural-urban mapping project, Version 1 (GRUMP v1): Urban extents grid. NASA Socioeconomic Data and Applications Center, Palisades, NY. Available at: `http://dx.doi.org/10.7927/H4GH9FVG`.

Ciofi, C., Beaumontf, M.A., Swingland, I.R. & Bruford, M.W. (1999) Genetic divergence and units for conservation in the Komodo dragon *Varanus komodoensis*. *Proceedings of the Royal Society of London B: Biological Sciences*, **266**, 2269–2274.

Clarke, R.T., Rothery, P. & Raybould, A.F. (2002) Confidence limits for regression relationships between distance matrices: Estimating gene flow with distance. *Journal of Agricultural, Biological, & Environmental Statistics*, **7**, 361.

Claudet, J., Pelletier, D., Jouvenel, J.Y., Bachet, F. & Galzin, R. (2006) Assessing the effects of marine protected area (MPA) on a reef fish assemblage in a northwestern Mediterranean marine reserve: Identifying community-based indicators. *Biological Conservation*, **130**, 349–369.

Clavero, M., Brotons, L., Pons, P. & Sol, D. (2009) Prominent role of invasive species in avian biodiversity loss. *Biological Conservation*, **142**, 2043–2049.

Cocks, K. & Baird, I. (1989) Using mathematical programming to address the multiple reserve selection problem: An example from the Eyre Peninsula, South Australia. *Biological Conservation*, **49**, 113 – 130.

Coetzer, K.L., Witkowski, E.T.F. & Erasmus, B.F.N. (2014) Reviewing Biosphere Reserves globally: Effective conservation action or bureaucratic label? *Biological Reviews*, **89**, 82–104.

Comes, H.P., Tribsch, A. & Bittkau, C. (2008) Plant speciation in continental island floras as exemplified by Nigella in the Aegean Archipelago. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **363**, 3083–3096.

Coulson, T., Mace, G.M., Hudson, E. & Possingham, H. (2001) The use and abuse of population viability analysis. *Trends in Ecology & Evolution*, **16**, 219–221.

Cowling, R.M. & Pressey, R.L. (2001) Rapid plant diversification: Planning for an evolutionary future. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 5452–5457.

Cowling, R., Pressey, R., Lombard, A., Desmet, P. & Ellis, A. (1999) From representation to persistence: Requirements for a sustainable system of conservation areas in the species-rich mediterranean-climate desert of southern Africa. *Diversity & Distributions*, **5**, 51–71.

Crandall, K.A., Bininda-Emonds, O.R.P., Mace, G.M. & Wayne, R.K. (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology & Evolution*, **15**, 290–295.

Creech, T.G., Epps, C.W., Monello, R.J. & Wehausen, J.D. (2014) Using network theory to prioritize management in a desert bighorn sheep metapopulation. *Landscape Ecology*, **29**, 605–619.

Crowley, T.J. (2000) Causes of climate change over the past 1000 years. *Science*, **289**, 270–277.

Csardi, G. & Nepusz, T. (2006) The igraph software package for complex network research. *InterJournal, Complex Systems*, p. 1695.

Csuti, B., Polasky, S., Williams, P.H., Pressey, R.L., Camm, J.D., Kershaw, M., Kiester, A.R., Downs, B., Hamilton, R., Huso, M. & Sahr, K. (1997) A comparison of reserve selection algorithms using data on terrestrial vertebrates in Oregon. *Biological Conservation*, **80**, 83–97.

Cui, T.T., Ouyang, Y.F. & Shen, Z.J.M. (2010) Reliable facility location design under the risk of disruptions. *Operations Research*, **58**, 998–1011.

Cushman, S.A. & Landguth, E.L. (2010) Spurious correlations and inference in landscape genetics. *Molecular Ecology*, **19**, 3592–3602.

Daugherty, C., Cree, A., Hay, J. & Thompson, M. (1990) Neglected taxonomy and continuing extinctions of tuatara (*Sphenodon*). *Nature*, **347**, 177.

de Villemereuil, P., Frichot, E., Bazin, E., François, O. & Gaggiotti, O.E. (2014) Genome scan methods against more complex models: When and how much should we trust them? *Molecular Ecology*, **23**, 2006–2019.

Dechamps, C., Noret, N., Mozek, R., Escarré, J., Lefèbvre, C., Gruber, W. & Meerts, P. (2008) Cost of adaptation to a metalliferous environment for *Thlaspi caerulescens*: A field reciprocal transplantation approach. *New Phytologist*, **177**, 167–177.

Deck, J., Gaither, M.R., Ewing, R., Bird, C.E., Davies, N., Meyer, C., Riginos, C., Toonen, R.J. & Crandall, E.D. (2017) The Genomic Observatories Metadatabase (GeOMe): A new

repository for field and sampling event metadata associated with genetic samples. *PLoS Biology*, **15**, e2002925.

Deutsch, C.A., Tewksbury, J.J., Huey, R.B., Sheldon, K.S., Ghalambor, C.K., Haak, D.C. & Martin, P.R. (2008) Impacts of climate warming on terrestrial ectotherms across latitude. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 6668–6672.

Di Marco, M., Visconti, P., Santini, L., Mortelliti, A., Boitani, L. & Rondinini, C. (2016) Using habitat suitability models to scale up population persistence targets. *Hystrix, Italian Journal of Mammalogy*, **27**, DOI: 10.4404/hystrix–27.1–11660.

Di Minin, E. & Toivonen, T. (2015) Global protected area expansion: Creating more than paper parks. *BioScience*, **65**, 637–638.

Dilkina, B., Houtman, R., Gomes, C.P., Montgomery, C.A., McKelvey, K.S., Kendall, K., Graves, T.A., Bernstein, R. & Schwartz, M.K. (2017) Trade-offs and efficiencies in optimal budget-constrained multispecies corridor networks. *Conservation Biology*, **31**, 192–202.

Diniz, J.A.F. & Telles, M. (2006) Optimization procedures for establishing reserve networks for biodiversity conservation taking into account population genetic structure. *Genetics & Molecular Biology*, **29**, 207–214.

Diniz-Filho, J.A.F., Melo, D.B., de Oliveira, G., Collevatti, R.G., Soares, T.N., Nabout, J.C., Lima, J.d.S., Dobrovolski, R., Chaves, L.J., Naves, R.V., Loyola, R.D. & Telles, M.P.d.C. (2012) Planning for optimal conservation of geographical genetic variability within species. *Conservation Genetics*, **13**, 1085–1093.

Dobrowski, S.Z. (2011) A climatic basis for microrefugia: The influence of terrain on climate. *Global Change Biology*, **17**, 1022–1035.

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 of the United States of America*, **113**, 11261–11265.

Dudaniec, R.Y., Worthington Wilmer, J., Hanson, J.O., Warren, M., Bell, S. & Rhodes, J.R. (2016) Dealing with uncertainty in landscape genetic resistance models: A case of three co-occurring marsupials. *Molecular Ecology*, **25**, 470–486.

Duforet-Frebourg, N., Bazin, E. & Blum, M.G. (2014) Genome scans for detecting footprints of local adaptation using a Bayesian factor model. *Molecular Biology & Evolution*, **31**, 2483–2495.

Durand, P., Wada, K.T. & Blanc, F. (1993) Genetic variation in wild and hatchery stocks of the black pearl oyster, *Pinctada margarififera*, from Japan. *Aquaculture*, **110**, 27–40.

Edwards, L.J., Muller, K.E., Wolfinger, R.D., Qaqish, B.F. & Schabenberger, O. (2008) An R^2 statistic for fixed effects in the linear mixed model. *Statistics in Medicine*, **27**, 6137–6157.

Elith, J., Graham, C., Anderson, R., Dudík, M., Ferrier, S., Guisan, A., J Hijmans, R., Huettmann, F., R Leathwick, J., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M.M., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberón, J., Williams, S., Wisz, M.S. & Zimmermann, N.E. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.

Ellstrand, N.C. (2014) Is gene flow the most important evolutionary force in plants? *American Journal of Botany*, **101**, 737–753.

Ellstrand, N.C. & Elam, D.R. (1993) Population genetic consequences of small population size: Implications for plant conservation. *Annual Review of Ecology & Systematics*, **24**, 217–242.

Endo, T., Ikeo, K. & Gojobori, T. (1996) Large-scale search for genes on which positive selection may operate. *Molecular Biology & Evolution*, **13**, 685–690.

Endy, D., You, L., Yin, J. & Molineux, I.J. (2000) Computation, prediction, and experimental tests of fitness for bacteriophage T7 mutants with permuted genomes. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 5375–5380.

Engelbrecht, I., Robertson, M., Stoltz, M. & Joubert, J.W. (2016) Reconsidering environmental diversity (ED) as a biodiversity surrogacy strategy. *Biological Conservation*, **197**, 171 – 179.

Epps, C.W., Palsbøll, P.J., Wehausen, J.D., Roderick, G.K., Ramey, R.R. & McCullough, D.R. (2005) Highways block gene flow and cause a rapid decline in genetic diversity of desert bighorn sheep. *Ecology Letters*, **8**, 1029–1038.

Ester, M., Kriegel, H.P., Sander, J. & Xu, X. (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proc. 2nd International Conference on Knowledge Discovery & Data Mining* (eds Simoudis, E., Han, J. & Fayyad, H.). AAAI Press.

Faith, D.P. (2003) Environmental diversity (ED) as surrogate information for species-level biodiversity. *Ecography*, **26**, 374–379.

Faith, D.P., Ferrier, S. & Walker, P.A. (2004) The ED strategy: How species-level surrogates indicate general biodiversity patterns through an "environmental diversity" perspective. *Journal of Biogeography*, **31**, 1207–1217.

Faith, D.P. & Walker, P.A. (1993) *DIVERSITY: A software package for sampling phylogenetic and environmental diversity.* Reference and user's guide (version 1.0). CSIRO Division

of Wildlife and Ecology, Canberra. Available at http://australianmuseum.com/uploads/documents/faith%20diversity.pdf.

Faith, D.P. & Walker, P.A. (1996) Environmental diversity: On the best-possible use of surrogate data for assessing the relative biodiversity of sets of areas. *Biodiversity & Conservation*, **5**, 399–415.

Faith, D.P. (1992) Conservation evaluation and phylogenetic diversity. *Biological Conservation*, **61**, 1–10.

Faith, D.P. (2011) Attempted tests of the surrogacy value of the ED environmental diversity measures highlight the need for corroboration assessment of surrogacy hypotheses. *Ecological Indicators*, **11**, 745–748.

Falush, D., Stephens, M. & Pritchard, J.K. (2007) Inference of population structure using multilocus genotype data: Dominant markers and null alleles. *Molecular Ecology Notes*, **7**, 574–578.

Ferraro, P.J. (2009) Counterfactual thinking and impact evaluation in environmental policy. *New Directions for Evaluation*, **2009**, 75–84.

Ferrier, S., Manion, G., Elith, J. & Richardson, K. (2007) Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity & Distributions*, **13**, 252–264.

Ficetola, G.F. & Bernardi, F. (2005) Supplementation or *in situ* conservation? Evidence of local adaptation in the Italian agile frog *Rana latastei* and consequences for the management of populations. *Animal Conservation*, **8**, 33–40.

Ficetola, G.F., Rondinini, C., Bonardi, A., Baisero, D. & Padoa-Schioppa, E. (2015) Habitat availability for amphibians and extinction threat: A global analysis. *Diversity & Distributions*, **21**, 302–311.

Fjeldså, J. & Irestedt, M. (2009) Diversification of the South American avifauna: Patterns and implications for conservation in the Andes. *Annals of the Missouri Botanical Garden*, **96**, 398–409.

Foll, M. & Gaggiotti, O. (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, **180**, 977–993.

Foody, G. (2008) GIS: Biodiversity applications. *Progress in Physical Geography*, **32**, 223–235.

Fortuna, M.A., Albaladejo, R.G., Fernández, L., Aparicio, A. & Bascompte, J. (2009) Networks of spatial genetic variation across species. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 19044–19049.

Fournier-Level, A., Korte, A., Cooper, M.D., Nordborg, M., Schmitt, J. & Wilczek, A.M. (2011) A map of local adaptation in *Arabidopsis thaliana*. *Science*, **334**, 86–89.

Fraley, C. & Raftery, A.E. (2002) Model-based clustering, discriminant analysis and density estimation. *Journal of the American Statistical Association*, **97**, 611–631.

Frankel, O.H. (1974) Genetic conservation: Our evolutionary responsibility. *Genetics*, **78**, 53–65.

Frankham, R. (1995) Inbreeding and Extinction: A threshold effect. *Conservation Biology*, **9**, 792–799.

Frankham, R. (2005) Genetics and extinction. *Biological Conservation*, **126**, 131–140.

Frankham, R. (2010) Challenges and opportunities of genetic approaches to biological conservation. *Biological Conservation*, **143**, 1919–1927.

Franks, S.J., Sim, S. & Weis, A.E. (2007) Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 1278–1282.

Fuller, R.A., McDonald-Madden, E., Wilson, K.A., Carwardine, J., Grantham, H.S., Watson, J.E., Klein, C.J., Green, D.C. & Possingham, H.P. (2010) Replacing underperforming protected areas achieves better conservation outcomes. *Nature*, **466**, 365.

Fuller, T. & Sarkar, S. (2006) LQGraph: A software package for optimizing connectivity in conservation planning. *Environmental Modelling & Software*, **21**, 750–755.

Gannon, P., Seyoum-Edjigu, E., Cooper, D., Sandwith, T., de Souza Dias, B.F., Palmer Paşca, C., Lang, B., Ervin, J. & Gidda, S. (2017) Status and prospoects for achieving Aichi Biodiversity Target 11: Implications of national commitments and priority actions. *PARKS*, **23.2**, 9–22.

Gaston, K.J., Jackson, S.E., Cantu-Salazar, L. & Cruz-Pinon, G. (2008) The ecological performance of protected areas. *Annual Review of Ecology, Evolution & Systematics*, **39**, 93–113.

Gehlbach, F.R. (1975) Investigation, evaluation, and priority ranking of natural areas. *Biological Conservation*, **8**, 79–88.

Goldsmith, F.B. (1975) The evaluation of ecological resources in the countryside for conservation purposes. *Biological Conservation*, **8**, 89–96.

Gomez-Mestre, I., Tejedo, M. & Merilä, J. (2003) Local adaptation of an anuran amphibian to osmotically stressful environments. *Evolution*, **57**, 1889–1899.

Gower, J.C. (1971) A general coefficient of similarity and some of its properties. *Biometrics*, **27**, 857–871.

Grady, K.C., Ferrier, S.M., Kolb, T.E., Hart, S.C., Allan, G.J. & Whitham, T.G. (2011) Genetic variation in productivity of foundation riparian species at the edge of their distribution: Implications for restoration and assisted migration in a warming climate. *Global Change Biology*, **17**, 3724–3735.

Grant, P.R. & Grant, R. (1992) Hybridization of bird species. *Science*, **256**, 193.

Grantham, H.S., Moilanen, A., Wilson, K.A., Pressey, R.L., Rebelo, T.G. & Possingham, H.P. (2008) Diminishing return on investment for biodiversity data in conservation planning. *Conservation Letters*, **1**, 190–198.

Greenacre, M. & Primicerio, R. (2014) *Multivariate analysis of ecological data*. BBVA Foundation, Madrid, Spain.

Gugerli, F., Englisch, T., Niklfeld, H., Tribsch, A., Mirek, Z., Ronikier, M., Zimmermann, N., Holderegger, R. & Taberlet, P. (2008) Relationships among levels of biodiversity and the relevance of intraspecific diversity in conservation—a project synopsis. *Perspectives in Plant Ecology, Evolution & Systematics*, **10**, 259–281.

Gurobi Optimization Inc. (2016) Gurobi Optimizer Reference Manual. Version 7.0.2. Available at `https://gurobi.com`.

Hairston, N.G., Ellner, S.P., Geber, M.A., Yoshida, T. & Fox, J.A. (2005) Rapid evolution and the convergence of ecological and evolutionary time. *Ecology Letters*, **8**, 1114–1127.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O. & Townshend, J.R.G. (2013) High-resolution global maps of 21st-century forest cover change. *Science*, **342**, 850–853.

Hanson, J.O., Rhodes, J.R., Riginos, C. & Fuller, R.A. (2017a) Environmental and geographic variables are effective surrogates for genetic variation in conservation planning. *Proceedings of the National Academy of Sciences*, **114**, 12755–12760.

Hanson, J., Rhodes, J., Possingham, H. & Fuller, R. (2017b) raptr: Representative and adequate prioritisation toolkit in R. *Methods in Ecology & Evolution*, **9**, 320–330.

Hanson, J., Schuster, R., Morell, N., Strimas-Mackey, M., Watts, M., Arcese, P., Bennett, J. & Possingham, H. (2017c) *prioritizr: Systematic conservation prioritization in R*. R package version 4.0.0.9. Available at `https://github.com/prioritizr/prioritizr`.

Hanson, J., Schuster, R., Morell, N., Strimas-Mackey, M., Watts, M., Arcese, P., Bennett, J. & Possingham, H. (2017d) *prioritizr: Systematic conservation prioritization in R*. R package version 3.0.3.3. Available at `https://github.com/prioritizr/prioritizr`.

Hedgecock, D. (1994) Does variance in reproductive success limit effective population sizes of marine organisms? *Genetics & Evolution of Aquatic Organisms*, **122**.

Hedrick, P.W. (1982) Genetic hitchhiking: A new factor in evolution? *BioScience*, **32**, 845–853.

Hedrick, P.W. (1995) Gene flow and genetic restoration: The Florida panther as a case study. *Conservation Biology*, **9**, 996–1007.

Helliwell, D.R. (1969) Valuation of wildlife resources. *Regional Studies*, **3**, 41–47.

Hendry, A.P., Kinnison, M.T., Heino, M., Day, T., Smith, T.B., Fitt, G., Bergstrom, C.T., Oakeshott, J., Jorgensen, P.S., Zalucki, M.P., Gilchrist, G., Southerton, S., Sih, A., Strauss, S., Denison, R.F. & Carroll, S.P. (2011) Evolutionary principles and their practical application. *Evolutionary Applications*, **4**, 159–183.

Hendry, A.P., Lohmann, L.G., Conti, E., Cracraft, J., Crandall, K.A., Faith, D.P., Hauser, C., Joly, C.A., Kogure, K., Larigauderie, A., Magallon, S., Moritz, C., Tillier, S., Zardoya, R., Prieur-Richard, A.H., Walther, B.A., Yahara, T. & Donoghue, M.J. (2010) Evolutionary biology in biodiversity science, conservation, and policy: A call to action. *Evolution*, **64**, 1517–1528.

Hijmans, R.J. (2016) *raster: Geographic Data Analysis and Modeling*. R package version 2.5-8. Available at `https://CRAN.R-project.org/package=raster`.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hirao, A. & Kudo, G. (2004) Landscape genetics of alpine-snowbed plants: Comparisons along geographic and snowmelt gradients. *Heredity*, **93**, 290.

H.M., B.S., Martin, C., J., S.R., E., S.R., P.W., S.J., Mike, H., M., B.G., Ariadne, A., Andrew, B., Bastian, B., M., B.T., E., C.K., T., C.R.M., John, C., Francesco, F.G., D.C., F.L., A., F.R., Jonas, G., Heather, H., Craig, H.T., Michael, H., Ackbar, J., Lucas, J., Naomi, K., Ian, M., Amy, M., Beth, P., Gina, R., Nadia, R., Carlo, R., B., S.D., Benjamin, S., D., S.M., N., S.S., Andy, S., Joseph, T., Piero, V., E.M., W.J., Louisa, W. & D., B.N. (2015) Shortfalls and solutions for meeting national and global conservation area targets. *Conservation Letters*, **8**, 329–337.

Hoban, S., Kelley, J.L., Lotterhos, K.E., Antolin, M.F., Bradburd, G., Lowry, D.B., Poss, M.L., Reed, L.K., Storfer, A. & Whitlock, M.C. (2016) Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. *The American Naturalist*, **188**, 379–397.

Hoffmann, A., Griffin, P., Dillon, S., Catullo, R., Rane, R., Byrne, M., Jordan, R., Oakeshott, J., Weeks, A., Joseph, L., Lockhart, P., Borevitz, J. & Sgrò, C. (2015) A framework for incorporating evolutionary genomics into biodiversity conservation and management. *Climate Change Responses*, **2**, 1.

Hoffmann, A.A. & Sgrò, C.M. (2011) Climate change and evolutionary adaptation. *Nature*, **470**, 479–485.

Holderegger, R., Kamm, U. & Gugerli, F. (2006) Adaptive vs. neutral genetic diversity: Implications for landscape genetics. *Landscape Ecology*, **21**, 797–807.

Holsinger, K.E. & Weir, B.S. (2009) Genetics in geographically structured populations: Defining, estimating and interpreting FST. *Nature reviews Genetics*, **10**, 639.

Hortal, J., Araújo, M.B. & Lobo, J.M. (2009) Testing the effectiveness of discrete and continuous environmental diversity as a surrogate for species diversity. *Ecological Indicators*, **9**, 138–149.

Hothorn, T., Bretz, F. & Westfall, P. (2008) Simultaneous inference in general parametric models. *Biometrical Journal*, **50**, 346–363.

Hutchinson, G.E. (1957) Concluding remarks. *Cold Spring Harbor Symposium on Quantitative Biology*, **22**, 415–427.

Innes, J., Hay, R., Flux, I., Bradfield, P., Speed, H. & Jansen, P. (1999) Successful recovery of North Island kōkako *Callaeas cinerea wilsoni* populations, by adaptive management. *Biological Conservation*, **87**, 201–214.

Isaac, N.J., Turvey, S.T., Collen, B., Waterman, C. & Baillie, J.E. (2007) Mammals on the EDGE: Conservation priorities based on threat and phylogeny. *PLoS ONE*, **2**, e296.

IUCN (2001) *IUCN Red List Categories and Criteria. Version 3.1*. IUCN Species Survival Commission, Gland, Switzerland, and Cambridge, United Kingdom.

IUCN (2017) *The IUCN Red List of Threatened Species*. Version 2017-1. Available at `http://iucnredlist.org`.

IUCN & UNEP-WCMC (2016) *The World Database on Protected Areas (WDPA)*. Available at: http://dx.doi.org/10.4225/13/511C71F8612C3 (accessed August 2016).

Jackman, S. (2015) *pscl: Classes and Methods for R Developed in the Political Science Computational Laboratory, Stanford University*. R package version 1.4.9. Department of Political Science, Stanford University, Stanford, California. Available at `http://pscl.stanford.edu`.

Jaeger, B. (2016) *r2glmm: Computes R squared for mixed (multilevel) models*. R package version 0.1.1. Available at `https://CRAN.R-project.org/package=r2glmm`.

Jakobsson, M. & Rosenberg, N.A. (2007) CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.

Jarvis, A., Reuter, H., Nelson, A. & E., G. (2008) Hole-filled srtm for the globe. Version 4.

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.

Kawecki, T.J. (2008) Adaptation to marginal habitats. *Annual Review of Ecology, Evolution, & Systematics*, **39**, 321–342.

Kawecki, T.J. & Ebert, D. (2004) Conceptual issues in local adaptation. *Ecology letters*, **7**, 1225–1241.

Keeley, A.T.H., Beier, P. & Gagnon, J.W. (2016) Estimating landscape resistance from habitat suitability: Effects of data source and nonlinearities. *Landscape Ecology*, **31**, 2151–2162.

Keller, D., Holderegger, R., van Strien, M.J. & Bolliger, J. (2015) How to make landscape genetics beneficial for conservation management? *Conservation Genetics*, **16**, 503–512.

Keyghobadi, N. (2007) The genetic implications of habitat fragmentation for animals. *Canadian Journal of Zoology*, **85**, 1049–1064.

Kiester, A.R., Scott, J.M., Csuti, B., Noss, R.F., Butterfield, B., Sahr, K. & White, D. (1996) Conservation prioritization using gap data. *Conservation Biology*, **10**, 1332–1342.

Kimura, M. (1954) Process leading to quasi-fixation of genes in natural populations due to random fluctuation of selection intensities. *Genetics*, **39**, 280.

King, T., Chamberlan, C. & Courage, A. (2012) Assessing initial reintroduction success in long-lived primates by quantifying survival, reproduction, and dispersal parameters: Western Lowland gorillas (*Gorilla gorilla gorilla*) in Congo and Gabon. *International Journal of Primatology*, **33**, 134–149.

Kirkpatrick, J.B. (1983) An iterative method for establishing priorities for the selection of nature reserves: An example from Tasmania. *Biological Conservation*, **25**, 127–134.

Kirkpatrick, S., Gelatt, C.D. & Vecchi, M.P. (1983) Optimization by simulated annealing. *Science*, **220**, 671–680.

Klappert, K. & Reinhold, K. (2003) Acoustic preference functions and sexual selection on the male calling song in the grasshopper *Chorthippus biguttulus*. *Animal Behaviour*, **65**, 225–233.

Klein, C., Wilson, K., Watts, M., Stein, J., Berry, S., Carwardine, J., Smith, M.S., Mackey, B. & Possingham, H. (2009a) Incorporating ecological and evolutionary processes into continental-scale conservation planning. *Ecological Applications*, **19**, 206–217.

Klein, C.J., Wilson, K.A., Watts, M., Stein, J., Carwardine, J., Mackey, B. & Possingham, H.P. (2009b) Spatial conservation prioritization inclusive of wilderness quality: A case study of Australia's biodiversity. *Biological Conservation*, **142**, 1282–1290.

Krosby, M., Breckheimer, I., John Pierce, D., Singleton, P.H., Hall, S.A., Halupka, K.C., Gaines, W.L., Long, R.A., McRae, B.H., Cosentino, B.L. & Schuett-Hames, J.P. (2015) Focal species and landscape "naturalness" corridor models offer complementary approaches for connectivity conservation planning. *Landscape Ecology*, **30**, 2121–2132.

Kruuk, L.E., Slate, J., Pemberton, J.M., Brotherstone, S., Guinness, F., Clutton-Brock, T. & Houle, D. (2002) Antler size in red deer: Heritability and selection but no evolution. *Evolution*, **56**, 1683–1695.

Kuempel, C.D., Chauvenet, A.L. & Possingham, H.P. (2016) Equitable representation of ecoregions is slowly improving despite strategic planning shortfalls. *Conservation Letters*, **9**, 422–428.

Kuittinen, H., Sillanpää, M. & Savolainen, O. (1997) Genetic basis of adaptation: Flowering time in *Arabidopsis thaliana*. *Theoretical & Applied Genetics*, **95**, 573–583.

Laikre, L. (2010) Genetic diversity is overlooked in international conservation policy implementation. *Conservation Genetics*, **11**, 349–354.

Laikre, L., Allendorf, F.W., Aroner, L.C., Baker, C.S., Gregovich, D.P., Hansen, M.M., Jackson, J.A., Kendall, K.C., McKelvey, K., Neel, M.C., Olivieri, I., Ryman, N., Schwartz, M.K., Bull, R.S., Stetz, J.B., Yallmon, D.A., Taylor, B.L., Vojta, C.D., Waller, D.M. & Waples, R.S. (2010) Neglect of genetic diversity in implementation of the convention of biological diversity. *Conservation Biology*, **24**, 86–88.

Land, A.H. & Doig, A.G. (1960) An automatic method for solving discrete programming problems. *Econometrica*, **28**, 497–520.

Land Resources Management Unit, Institute for Environment Sustainability, European Commission & Joint Research Centre (2010) *Map of Soil pH in Europe*. Available at `https://esdac.jrc.ec.europa.eu/content/soil-ph-europe`.

Lande, R. (1976) Natural selection and random genetic drift in phenotypic evolution. *Evolution*, **30**, 314–334.

Law, E.A., Bryan, B.A., Meijaard, E., Mallawaarachchi, T., Struebig, M.J., Watts, M.E. & Wilson, K.A. (2017) Mixed policies give more options in multifunctional tropical forest landscapes. *Journal of Applied Ecology*, **54**, 51–60.

Lei, F., Qu, Y., Song, G., Alström, P. & Fjeldså, J. (2015) The potential drivers in forming avian biodiversity hotspots in the East Himalaya Mountains of Southwest China. *Integrative Zoology*, **10**, 171–181.

Leiss, K., Vrieling, K. & Klinkhamer, P. (2004) Heritability of nectar production in *Echium vulgare*. *Heredity*, **92**, 446–451.

Lesica, P. & Allendorf, F.W. (1995) When are peripheral populations valuable for conservation? *Conservation Biology*, **9**, 753–760.

Liu, C., Berry, P.M., Dawson, T.P. & Pearson, R.G. (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, **28**, 385–393.

Lombard, A., Nicholls, A. & August, P. (1995) Where should nature reserves be located in South Africa? A snake's perspective. *Conservation Biology*, **9**, 363–372.

Losos, J.B., Warheit, K.I. & Schoener, T.W. (1997) Adaptive differentiation following experimental island colonization in *Anolis* lizards. *Nature*, **387**, 70.

Luu, K., Blum, M.G. & Duforet-Frebourg, N. (2016) *pcadapt: fast principal component analysis for outlier detection*. R package version 3.0. Available at `https://CRAN.R-project.org/package=pcadapt`.

Mace, G.M., Collar, N.J., Gaston, K.J., Hilton-Taylor, C., Akçakaya, H.R., Leader-Williams, N., Milner-Gulland, E.J. & Stuart, S.N. (2008) Quantification of extinction risk: IUCN's system for classifying threatened species. *Conservation Biology*, **22**, 1424–1442.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M. & Hornik, K. (2015) *cluster: Cluster Analysis Basics and Extensions*. R package version 2.0.3. Available at `http://CRAN.R-project.org/package=cluster`.

Mahieu, S., Soussou, S., Cleyet-Marel, J.C., Brunel, B., Mauré, L., Lefèbvre, C. & Escarré, J. (2013) Local adaptation of metallicolous and non-metallicolous *Anthyllis vulneraria* populations: Their utilization in soil restoration. *Restoration ecology*, **21**, 551–559.

Manel, S. & Holderegger, R. (2013) Ten years of landscape genetics. *Trends in ecology & evolution*, **28**, 614–621.

Margules, C. & Usher, M.B. (1981) Criteria used in assessing wildlife conservation potential: A review. *Biological Conservation*, **21**, 79–109.

Margules, C.R. & Pressey, R.L. (2000) Systematic conservation planning. *Nature*, **405**, 243–253.

Masel, J. (2011) Genetic drift. *Current Biology*, **21**, R837–R838.

Mateo-Sánchez, M.C., Balkenhol, N., Cushman, S., Pérez, T., Domínguez, A. & Saura, S. (2015) Estimating effective landscape distances and movement corridors: Comparison of habitat and genetic data. *Ecosphere*, **6**, 1–16.

Mateo-Sánchez, M.C., Balkenhol, N., Cushman, S., Pérez, T., Domínguez, A. & Saura, S. (2015) A comparative framework to infer landscape effects on population genetic structure: Are habitat suitability models effective in explaining gene flow? *Landscape Ecology*, **30**, 1405–1420.

Maxwell, S.L., Milner-Gulland, E.J., Jones, J.P., Knight, A.T., Bunnefeld, N., Nuno, A., Bal, P., Earle, S., Watson, J.E. & Rhodes, J.R. (2015) Being smart about SMART environmental targets. *Science*, **347**, 1075–1076.

McCabe, J. & Partridge, L. (1997) An interaction between environmental temperature and genetic variation for body size for the fitness of adult female *Drosophila melanogaster*. *Evolution*, **51**, 1164–1174.

McCarthy, D.P., Donald, P.F., Scharlemann, J.P.W., Buchanan, G.M., Balmford, A., Green, J.M.H., Bennun, L.A., Burgess, N.D., Fishpool, L.D.C., Garnett, S.T., Leonard, D.L., Maloney, R.F., Morling, P., Schaefer, H.M., Symes, A., Wiedenfeld, D.A. & Butchart, S.H.M. (2012) Financial costs of meeting global biodiversity conservation targets: Current spending and unmet needs. *Science*, **338**, 946–949.

McKay, J.K., Bishop, J.G., Lin, J.Z., Richards, J.H., Sala, A. & Mitchell-Olds, T. (2001) Local adaptation across a climatic gradient despite small effective population size in the rare sapphire rockcress. *Proceedings of the Royal Society of London B: Biological Sciences*, **268**, 1715–1721.

McNeely, J.A. (1994) Protected areas for the 21st-century: Working to provide benefits to society. *Biodiversity & Conservation*, **3**, 390–405.

McRae, B.H., Dickson, B.G., Keitt, T.H. & Shah, V.B. (2008) Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology*, **89**, 2712–2724.

McRae, B.H., Hall, S.A., Beier, P. & Theobald, D.M. (2012) Where to restore ecological connectivity? Detecting barriers and quantifying restoration benefits. *PLoS ONE*, **7**, e52604.

Meineri, E., Skarpaas, O. & Vandvik, V. (2012) Modeling alpine plant distributions at the landscape scale: Do biotic interactions matter? *Ecological Modelling*, **231**, 1 – 10.

Meirmans, P., Goudet, J., IntraBioDiv Consortium & Gaggiotti, O. (2011) Ecology and life history affect different aspects of the population structure of 27 high-alpine plants. *Molecular Ecology*, **20**, 3144–3155.

Milanovich, J.R., Struecker, B.P., Warcholek, S.A. & Harden, L.A. (2017) Thermal environment and microhabitat of ornate box turtle hibernacula. *Wildlife Biology*, pp. wlb–00295.

Mitchell, R.J. & Shaw, R.G. (1993) Heritability of floral traits for the perennial wild flower *Penstemon centranthifolius* (Scrophulariaceae): Clones and crosses. *Heredity*, **71**, 185–192.

Mittelmann, H. (2017) *Gurobi 7.5 Performance Benchmarks*. Gurobi Optimization. Available at: `https://gurobi.com/pdfs/benchmarks.pdf`.

Mochida, K. (2011) Combination of local selection pressures drives diversity in aposematic signals. *Evolutionary Ecology*, **25**, 1017.

Moilanen, A., Franco, A.M.A., Early, R.I., Fox, R., Wintle, B. & Thomas, C.D. (2005) Prioritizing multiple-use landscapes for conservation: Methods for large multi-species planning problems. *Proceedings of the Royal Society of London B: Biological Sciences*, **272**, 1885–1891.

Moilanen, A. (2007) Landscape Zonation, benefit functions and target-based planning: Unifying reserve selection strategies. *Biological Conservation*, **134**, 571–579.

Moilanen, A., Anderson, B.J., Arponen, A., Pouzols, F.M. & Thomas, C.D. (2013) Edge artefacts and lost performance in national versus continental conservation priority areas. *Diversity & Distributions*, **19**, 171–183.

Moilanen, A. & Wintle, B.A. (2007) The boundary-quality penalty: A quantitative method for approximating species responses to fragmentation in reserve selection. *Conservation Biology*, **21**, 355–364.

Møller, A.P. (1990) Nest predation selects for small nest size in the blackbird. *Oikos*, **57**, 237–240.

Moritz, C. (1999) Conservation units and translocations: Strategies for conserving evolutionary processes. *Hereditas*, **130**, 217–228.

Moritz, C. (2002) Strategies to protect biological diversity and the evolutionary processes that sustain it. *Systematic Biology*, **51**, 238–254.

Moritz, C. (1994) Defining evolutionarily significant units for conservation. *Trends in Ecology & Evolution*, **9**, 373–375.

Murray, N.J., Clemens, R.S., Phinn, S.R., Possingham, H.P. & Fuller, R.A. (2014) Tracking the rapid loss of tidal wetlands in the Yellow Sea. *Frontiers in Ecology & the Environment*, **12**, 267–272.

Muscarella, R., Galante, P., Soley-Guardia, M., Boria, R., Kass, J., Uriarte, M. & Anderson, R. (2014) ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for ecological niche models. *Methods in Ecology & Evolution*, **5**, 1198–1205.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., Da Fonseca, G.A. & Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.

NASA LP DAAC (2015) *MCD12Q1 Land Cover Type Yearly L3 Global 500 m SIN Grid*. Version 051. Sioux Falls, South Dakota. Available at `https://lpdaac.usgs.gov`.

Nei, M. (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590.

Nicholls, A.O. & Margules, C.R. (1993) An upgraded reserve selection algorithm. *Biological Conservation*, **64**, 165–169.

Nicotra, A.B., Beever, E.A., Robertson, A.L., Hofmann, G.E. & O'leary, J. (2015) Assessing the components of adaptive capacity to improve conservation and management efforts under global change. *Conservation Biology*, **29**, 1268–1278.

Nielsen, E.S., Beger, M., Henriques, R., Selkoe, K.A. & von der Heyden, S. (2017) Multispecies genetic objectives in spatial conservation planning. *Conservation Biology*, **31**, 872–882.

O'Grady, J.J., Brook, B.W., Reed, D.H., Ballou, J.D., Tonkyn, D.W. & Frankham, R. (2006) Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. *Biological Conservation*, **133**, 42–51.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H. & Wagner, H. (2015) *vegan: community ecology package*. R package version 2.3-2. Available at `http://CRAN.R-project.org/package=vegan`.

Oliver, M.A. & Webster, R. (1990) Kriging: A method of interpolation for geographical information systems. *International Journal of Geographical Information Systems*, **4**, 313–332.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P. & Kassem, K.R. (2001) Terrestrial ecoregions of the world: A new map of life on Earth. *BioScience*, **51**, 933–938.

Önal, H. & Briers, R.A. (2006) Optimal selection of a connected reserve network. *Operations Research*, **54**, 379–388.

Orians, G.H. (1993) Endangered at what level? *Ecological Applications*, **3**, 206–208.

Owen, S.H. & Daskin, M.S. (1998) Strategic facility location: A review. *European Journal of Operational Research*, **111**, 423–447.

Paradis, E., Claude, J. & Strimmer, K. (2004) APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics*, **20**, 289–290.

Pebesma, E.J. & Bivand, R.S. (2005) Classes and methods for spatial data in R. *R News*, **5**, 9–13.

Peng, R.D. (2011) Reproducible research in computational science. *Science*, **334**, 1226–1227.

Peterman, W.E. (2018) ResistanceGA: An R package for the optimization of resistance surfaces using genetic algorithms. *Methods in Ecology & Evolution*, **In press**. Available at `http://dx.doi.org/10.1111/2041-210X.12984`.

Pevzner, P. & Shamir, R. (2009) Computing has changed biology—biology education must catch up. *Science*, **325**, 541–542.

Phillips, B.L. & Shine, R. (2004) Adapting to an invasive species: Toxic cane toads induce morphological change in Australian snakes. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 17150–17155.

Phillips, S.J., Dudík, M. & Schapire, R.E. (2017) *Maxent software for modeling species niches and distributions*. Version 3.4.0. Available at `http://biodiversityinformatics.amnh.org/open_source/maxent`.

Pieris Project (2017) *The Pieris Project*. Available at: `http://www.pierisproject.org/` (accessed November 2017).

Polasky, S., Nelson, E., Lonsdorf, E., Fackler, P. & Starfield, A. (2005) Conserving species in a working landscape: land use with biological and economic objectives. *Ecological applications*, **15**, 1387–1401.

Ponce-Reyes, R., Clegg, S.M., Carvalho, S.B., McDonald-Madden, E. & Possingham, H.P. (2014) Geographical surrogates of genetic variation for selecting island populations for conservation. *Diversity & Distributions*, **20**, 640–651.

Possingham, H.P., Andelman, S.J., Burgman, M.A., Medellın, R.A., Master, L.L. & Keith, D.A. (2002) Limits to the use of threatened species lists. *Trends in Ecology & Evolution*, **17**, 503–507.

Potts, A.J., Hedderson, T.A. & Cowling, R.M. (2013) Testing large-scale conservation corridors designed for patterns and processes: Comparative phylogeography of three tree species. *Diversity & Distributions*, **19**, 1418–1428.

Pough, F.H. (1989) Organismal performance and Darwinian fitness: Approaches and interpretations. *Physiological Zoology*, **62**, 199–236.

Pressey, R.L. & Logan, V.S. (1994) Level of geographical subdivision and its effects on assessments of reserve coverage: A review of regional studies. *Conservation Biology*, **8**, 1037–1046.

Pressey, R.L. & Nicholls, A.O. (1989) Application of a numerical algorithm to the selection of reserves in semi-arid New South Wales. *Biological Conservation*, **50**, 263–278.

Pressey, R.L., Possingham, H.P. & Day, J.R. (1997) Effectiveness of alternative heuristic algorithms for identifying indicative minimum requirements for conservation reserves. *Biological Conservation*, **80**, 207–219.

Pressey, R.L., Possingham, H.P. & Margules, C.R. (1996) Optimality in reserve selection algorithms: When does it matter and how much? *Biological Conservation*, **76**, 259–267.

Pressey, R., Johnson, I. & Wilson, P. (1994) Shades of irreplaceability: Towards a measure of the contribution of sites to a reservation goal. *Biodiversity & Conservation*, **3**, 242–262.

Pulido, F., Berthold, P., Mohr, G. & Querner, U. (2001) Heritability of the timing of autumn migration in a natural bird population. *Proceedings of the Royal Society of London B: Biological Sciences*, **268**, 953–959.

R Core Team (1999) *Writing R extensions*. R Foundation for Statistical Computing. Available at `https://cran.r-project.org/doc/manuals/r-release/R-exts.html`.

R Core Team (2016) R: A language and environment for statistical computing. R Foundation for Statistical Computing. Available at: `http://www.R-project.org`.

Ratcliffe, D. (1971) Criteria for the selection of nature reserves. *Advancement of Science*, **27**, 294–296.

Ratcliffe, J.M. & Nydam, M.L. (2008) Multimodal warning signals for a multiple predator world. *Nature*, **455**, 96.

Rayfield, B., Moilanen, A. & Fortin, M.J. (2009) Incorporating consumer-resource spatial interactions in reserve design. *Ecological Modelling*, **220**, 725–733.

Raymond, B., VanDerWal, J. & Belbin, L. (2015) *ALA4R: Atlas of Living Australia (ALA) Data and Resources in R*. R package version 1.25. Available at `https://github.com/AtlasOfLivingAustralia/ALA4R`.

Rebelo, A.G. & Siegfried, W.R. (1990) Protection of Fynbos vegetation: Ideal and real-world options. *Biological Conservation*, **54**, 15–31.

Rebelo, A.G. & Siegfried, W.R. (1992) Where should nature reserves be located in the Cape Floristic Region, South Africa? Models for the spatial configuration of a reserve network aimed at maximizing the protection of floral diversity. *Conservation Biology*, **6**, 243–252.

Reed, D.H. & Frankham, R. (2003) Correlation between fitness and genetic diversity. *Conservation Biology*, **17**, 230–237.

Reed, J.M., Doerr, P.D. & Walters, J.R. (1988) Minimum viable population size of the red-cockaded woodpecker. *The Journal of Wildlife Management*, **52**, 385–391.

Reyers, B., O'Farrell, P.J., Nel, J.L. & Wilson, K. (2012) Expanding the conservation toolbox: conservation planning of multifunctional landscapes. *Landscape Ecology*, **27**, 1121–1134.

Reznick, D.N., Shaw, F.H., Rodd, F.H. & Shaw, R.G. (1997) Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*). *Science*, **275**, 1934–1937.

Rodrigues, A.S., Orestes Cerdeira, J. & Gaston, K.J. (2000) Flexibility, efficiency, and accountability: Adapting reserve selection algorithms to more complex conservation problems. *Ecography*, **23**, 565–574.

Rodrigues, A.S.L., Akçakaya, H.R., Andelman, S.J., Bakarr, M.I., Boitani, L., Brooks, T.M., Chanson, J.S., Fishpool, L.D.C., Da Fonseca, G.A.B., Gaston, K.J., Hoffmann, M., Marquet, P.A., Pilgrim, J.D., Pressey, R.L., Schipper, J., Sechrest, W., Stuart, S.N., Underhill, L.G., Waller, R.W., Watts, M.E.J. & Yan, X. (2004) Global gap analysis: Priority regions for expanding the global protected-area network. *BioScience*, **54**, 1092–1100.

Rodrigues, A.S., Pilgrim, J.D., Lamoreux, J.F., Hoffmann, M. & Brooks, T.M. (2006) The value of the IUCN Red List for conservation. *Trends in Ecology & Evolution*, **21**, 71–76.

Rondinini, C., Di Marco, M., Chiozza, F., Santulli, G., Baisero, D., Visconti, P., Hoffmann, M., Schipper, J., Stuart, S.N., Tognelli, M.F., Amori, G., Falcucci, A., Maiorano, L. & Boitani, L. (2011) Global habitat suitability models of terrestrial mammals. *Proceedings of the Royal Society of London B: Biological Sciences*, **366**, 2633–2641.

Rouget, M., Cowling, R.M., Pressey, R.L. & Richardson, D.M. (2003) Identifying spatial components of ecological and evolutionary processes for regional conservation planning in the Cape Floristic Region, South Africa. *Diversity & Distributions*, **9**, 191–210.

Ruiz-Lopez, M., Barelli, C., Rovero, F., Hodges, K., Roos, C., Peterman, W. & Ting, N. (2016) A novel landscape genetic approach demonstrates the effects of human disturbance on the Udzungwa red colobus monkey (*Procolobus gordonorum*). *Heredity*, **116**, 167–176.

Rull, V. (2009) Microrefugia. *Journal of Biogeography*, **36**, 481–484.

Rull, V. (2010) On microrefugia and cryptic refugia. *Journal of Biogeography*, **37**, 1623–1625.

Runge, C.A., Watson, J.E., Butchart, S.H., Hanson, J.O., Possingham, H.P. & Fuller, R.A. (2015) Protected areas and global conservation of migratory birds. *Science*, **350**, 1255–1258.

Safi, K., Armour-Marshall, K., Baillie, J.E. & Isaac, N.J. (2013) Global patterns of evolutionary distinct and globally endangered amphibians and mammals. *PLoS ONE*, **8**, e63582.

Sanderson, E.W., Segan, D.B. & Watson, J.E.M. (2015) Global status of and prospects for protection of terrestrial geophysical diversity. *Conservation Biology*, **29**, 649–656.

Sandve, G.K., Nekrutenko, A., Taylor, J. & Hovig, E. (2013) Ten simple rules for reproducible computational research. *PLoS Computational Biology*, **9**, e1003285.

Sarkar, S., Aggarwal, A., Garson, J., Margules, C.R. & Zeidler, J. (2002) Place prioritization for biodiversity content. *Journal of Biosciences*, **27**, 339–346.

Saura, S. & Pascual-Hortal, L. (2007) A new habitat availability index to integrate connectivity in landscape conservation planning: Comparison with existing indices and application to a case study. *Landscape & Urban Planning*, **83**, 91–103.

Savolainen, V., Anstett, M.C., Lexer, C., Hutton, I., Clarkson, J.J., Norup, M.V., Powell, M.P., Springate, D., Salamin, N. & Baker, W.J. (2006) Sympatric speciation in palms on an oceanic island. *Nature*, **441**, 210–213.

Schoville, S.D., Bonin, A., Francois, O., Lobreaux, S., Melodelima, C. & Manel, S. (2012) Adaptive genetic variation on the landscape: Methods and cases. *Annual Review of Ecology Evolution & Systematics*, **43**, 23–43.

Schrage, L. (1984) *Linear, integer, and quadratic programming with LINDO: User's manual*. The Scientific Press Palo Alto, CA.

Schuster, R., Law, E.A., Rodewald, A.D., Martin, T.G., Wilson, K.A., Watts, M., Possingham, H.P. & Arcese, P. (2017) Tax shifting and incentives for biodiversity conservation on private lands. *Conservation Letters*, **11**, e12377.

Seto, K.C., Güneralp, B. & Hutyra, L.R. (2012) Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 16083–16088.

Sgro, C.M., Lowe, A.J. & Hoffmann, A.A. (2011) Building evolutionary resilience for conserving biodiversity under climate change. *Evolutionary Applications*, **4**, 326–337.

Shryock, D.F., Havrilla, C.A., DeFalco, L.A., Esque, T.C., Custer, N.A. & Wood, T.E. (2015) Landscape genomics of *Sphaeralcea ambigua* in the Mojave Desert: A multivariate, spatially-explicit approach to guide ecological restoration. *Conservation genetics*, **16**, 1303–1317.

Skerratt, L.F., Berger, L., Speare, R., Cashins, S., McDonald, K.R., Phillott, A.D., Hines, H.B. & Kenyon, N. (2007) Spread of chytridiomycosis has caused the rapid global decline and extinction of frogs. *EcoHealth*, **4**, 125.

Slarkin, M. (1985) Gene flow in natural populations. *Annual Review of Ecology & Systematics*, **16**, 393–430.

Smith, T.B., Kark, S., Schneider, C.J., Wayne, R.K. & Moritz, C. (2001) Biodiversity hotspots and beyond: The need for preserving environmental transitions. *Trends in Ecology & Evolution*, **16**, 431.

Sobel, J.M., Chen, G.F., Watt, L.R. & Schemske, D.W. (2010) The biology of speciation. *Evolution*, **64**, 295–315.

Sørensen, J.G., Schou, M.F. & Loeschcke, V. (2017) Evolutionary adaptation to environmental stressors: A common response at the proteomic level. *Evolution*, **71**, 1627–1642.

Steiner, C.C., Weber, J.N. & Hoekstra, H.E. (2007) Adaptive variation in beach mice produced by two interacting pigmentation genes. *PLoS Biology*, **5**, 1–10.

Stigner, M.G., Beyer, H.L., Klein, C.J. & Fuller, R.A. (2016) Reconciling recreational use and conservation values in a coastal protected area. *Journal of Applied Ecology*, **53**, 1206–1214.

Storey, J.D., Bass, A.J., Dabney, A. & Robinson, D. (2015) *qvalue: Q-value estimation for false discovery rate control*. R package version 2.2.2. Available at `http://github.com/jdstorey/qvalue`.

Sugama, K., Taniguchi, N. & Umeda, S. (1988) An experimental study on genetic drift in hatchery population of red sea bream. *Nippon Suisan Gakkaishi*, **54**, 739–744.

Susanne, W.M., Julien, P., Daniel, K.W., Loïc, P., Jonathan, L., F., D.C., F., D.C., C., F.M., John-Arvid, G., Antoine, G., K., H.R., T., H.T., Ingolf, K., Miska, L., Luigi, M., Marie-Charlotte, N., Signe, N., Öckinger Erik, M., S.N., Mette, T., Allan, T., A., W.D., Peter, A. & Jens-Christian, S. (2013) The role of biotic interactions in shaping distributions and realised assemblages of species: Implications for species distribution modelling. *Biological Reviews*, **88**, 15–30.

Taberlet, P., Zimmermann, N.E., Englisch, T., Tribsch, A., Holderegger, R., Alvarez, N., Niklfeld, H., Coldea, G., Mirek, Z., Moilanen, A., Ahlmer, W., Marsan, P.A., Bona, E., Bovio, M., Choler, P., Cieślak, E., Colli, L., Cristea, V., Dalmas, J.P., Frajman, B., G., L., Gaudeul, M., Gielly, L., Gutermann, W., Jogan, N., Kagalo, A.A., Korbecka, G., Küpfer, P., Lequette, B., Letz, D.R., Manel, S., Mansion, G., Marhold, K., Martini, F., Negrini, R., Niño, F., Paun, O., Pellecchia, M., Perico, G., Piękoś-Mirkowa, H., Prosser, F., Puşcaş, M., Ronikier, M., Scheuerer, M., Schneeweiss, G.M., Schönswetter, P., Schratt-Ehrendorfer, L., Schüpfer, F., Selvaggi, A., Steinmann, K., Thiel-Egenter, C., van Loo, M., Winkler, M., Wohlgemuth, T., Wraber, T., Gugerli, F. & IntraBioDiv Consortium (2012a) Genetic diversity in widespread species is not congruent with species richness in alpine plant communities. *Ecology Letters*, **15**, 1439–1448.

Taberlet, P., Zimmermann, N., Englisch, T., Tribsch, A., Holderegger, R., Alvarez, N., Niklfeld, H., Mirek, Z., Moilanen, A., Ahlmer, W., Ajmone, Marsan, P., Bona, E., Bovio, M., Choler, P., Cieślak, E., Coldea, G., Colli, L., Cristea, V., Dalmas, J., Frajman, B., Garraud, L., Gaudeul, M., Gielly, L., Gutermann, W., Jogan, N., Kagalo, A., Korbecka, G., Küpfer, P., Lequette, B., Letz, D., Manel, S., Mansion, G., Marhold, K., Martini, F., Negrini, R.,

Niño, F., Paun, O., Pellecchia, M., Perico, G., Piękoś-Mirkowa, H., Prosser, F., Puşcaş, M., Ronikier, M., Scheuerer, M., Schneeweiss, G., Schönswetter, P., Schratt-Ehrendorfer, L., Schüpfer, F., Selvaggi, A., Steinmann, K., Thiel-Egenter, C., van, Loo, M., Winkler, M., Wohlgemuth, T., Wraber, T., Gugerli, F. & IntraBioDiv Consortium (2012b) *Data from: Genetic diversity in widespread species is not congruent with species richness in alpine plant communities*. Dryad Digital Repository. Available at `https://dx.doi.org/10.5061/dryad.s4q6s`.

Tarroso, P., Velo-Antón, G. & Carvalho, S.B. (2015) phylin: An R package for phylogeographic interpolation. *Molecular Ecology Resources*, **15**, 349–357.

Thiel-Egenter, C., Gugerli, F., Alvarez, N., Brodbeck, S., Cieślak, E., Colli, L., Englisch, T., Gaudeul, M., Gielly, L., Korbecka, G., Negrini, R., Paun, O., Pellecchia, M., Rioux, D., Ronikier, M., Peter, S., Fanny, S., Pierre, T., Andreas, T., Marcela, V.L., Manuela, W., Rolf, H. & IntraBioDiv Consortium (2009) Effects of species traits on the genetic diversity of high-mountain plants: A multi-species study across the Alps and the Carpathians. *Global Ecology & Biogeography*, **18**, 78–87.

Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C., Erasmus, B.F., Ferreira de Siqueira, M., Grainger, A., Hannah, L., Hughes, L., Huntley, B., van Jaarsveld, A.S., Midgley, G.F., Miles, L., Ortega-Huerta, M.A., Peterson, A.T., Phillips, O.L. & Williams, S.E. (2004a) Extinction risk from climate change. *Nature*, **427**, 145–148.

Thomas, J.A., Telfer, M.G., Roy, D.B., Preston, C.D., Greenwood, J.J.D., Asher, J., Fox, R., Clarke, R.T. & Lawton, J.H. (2004b) Comparative losses of British butterflies, birds, and plants and the global extinction crisis. *Science*, **303**, 1879–1881.

Thomassen, H.A., Buermann, W., Milá, B., Graham, C.H., Cameron, S.E., Schneider, C.J., Pollinger, J.P., Saatchi, S., Wayne, R.K. & Smith, T.B. (2010) Modeling environmentally associated morphological and genetic variation in a rainforest bird, and its application to conservation prioritization. *Evolutionary Applications*, **3**, 1–16.

Title, P.O. & Bemmels, J.B. (2017) ENVIREM: An expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. *Ecography*, **41**, 291–307.

Torrubia, S., McRae, B.H., Lawler, J.J., Hall, S.A., Halabisky, M., Langdon, J. & Case, M. (2014) Getting the most connectivity per conservation dollar. *Frontiers in Ecology & the Environment*, **12**, 491–497.

Tracy, L.N. & Jamieson, I.G. (2011) Historic DNA reveals contemporary population structure results from anthropogenic effects, not pre-fragmentation patterns. *Conservation Genetics*, **12**, 517–526.

Tubbs, C.R. & Blackwood, J.W. (1971) Ecological evaluation of land for planning purposes. *Biological Conservation*, **3**, 169–172.

Tufto, J. (2001) Effects of releasing maladapted individuals: A demographic-evolutionary model. *The American Naturalist*, **158**, 331–340.

Turner, T.L., Bourne, E.C., Von Wettberg, E.J., Hu, T.T. & Nuzhdin, S.V. (2010) Population resequencing reveals local adaptation of *Arabidopsis lyrata* to serpentine soils. *Nature Genetics*, **42**, 260–263.

Underhill, L.G. (1994) Optimal and suboptimal reserve selection algorithms. *Biological Conservation*, **70**, 85–87.

van Etten, J. (2017) R package gdistance: Distances and routes on geographical grids. *Journal of Statistical Software*, **76**, 1–21.

van Oers, K., Drent, P.J., De Goede, P. & Van Noordwijk, A.J. (2004) Realized heritability and repeatability of risk-taking behaviour in relation to avian personalities. *Proceedings of the Royal Society of London B: Biological Sciences*, **271**, 65–73.

Vane-Wright, R.I., Humphries, C.J. & Williams, P.H. (1991) What to protect? Systematics and the agony of choice. *Biological Conservation*, **55**, 235–254.

Vasconcelos, R., Brito, J.C., Carvalho, S.B., Carranza, S. & Harris, D.J. (2012) Identifying priority areas for island endemics using genetic versus specific diversity—the case of terrestrial reptiles of the Cape Verde Islands. *Biological Conservation*, **153**, 276–286.

Venter, O., Sanderson, E.W., Magrach, A., Allan, J.R., Beher, J., Jones, K.R., Possingham, H.P., Laurance, W.F., Wood, P., Fekete, B.M., Levy, M.A. & Watson, J.E.M. (2016) Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation. *Nature Communications*, **7**, 12558.

Verboom, G.A., Archibald, J.K., Bakker, F.T., Bellstedt, D.U., Conrad, F., Dreyer, L.L., Forest, F., Galley, C., Goldblatt, P., Henning, J.F., Mummenhoff, K., Linder, H.P., Muasya, A.M., Oberlander, K.C., Savolainen, V., Snijman, D.A., van der Niet, T. & Nowell, T.L. (2009) Origin and diversification of the Greater Cape flora: Ancient species repository, hot-bed of recent radiation, or both? *Molecular Phylogenetics & Evolution*, **51**, 44–53.

Violle, C., Navas, M.L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I. & Garnier, E. (2007) Let the concept of trait be functional! *Oikos*, **116**, 882–892.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee, T.v.d., Hornes, M., Friters, A., Pot, J., Paleman, J., Kuiper, M. & Zabeau, M. (1995) AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Research*, **23**, 4407–4414.

Wake, D.B. & Vredenburg, V.T. (2008) Are we in the midst of the sixth mass extinction? A view from the world of amphibians. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 11466–11473.

Wang, J. (2005) Estimation of effective population sizes from data on genetic markers. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **360**, 1395–1409.

Wang, Y. & Önal, H. (2016) Optimal design of compact and connected nature reserves for multiple species. *Conservation Biology*, **30**, 413–424.

Wang, Y.H., Yang, K.C., Bridgman, C.L. & Lin, L.K. (2008) Habitat suitability modelling to correlate gene flow with landscape connectivity. *Landscape Ecology*, **23**, 989–1000.

Watson, J.E., Dudley, N., Segan, D.B. & Hockings, M. (2014) The performance and potential of protected areas. *Nature*, **515**, 67–73.

Watts, M.E., Ball, I.R., Stewart, R.S., Klein, C.J., Wilson, K., Steinback, C., Lourival, R., Kircher, L. & Possingham, H.P. (2009) Marxan with Zones: Software for optimal conservation based land- and sea-use zoning. *Environmental Modelling & Software*, **24**, 1513 – 1521.

Whitlock, M.C. & Mccauley, D.E. (1999) Indirect measures of gene flow and migration: F_{ST} $\neq$ 1 /(4Nm + 1). *Heredity*, **82**, 117.

Wichman, H., Badgett, M., Scott, L., Boulianne, C. & Bull, J. (1999) Different trajectories of parallel evolution during viral adaptation. *Science*, **285**, 422–424.

Wilke, C.O., Wang, J.L., Ofria, C., Lenski, R.E. & Adami, C. (2001) Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature*, **412**, 331–333.

Williams, J.C. & ReVelle, C.S. (1998) Reserve assemblage of critical areas: A zero-one programming approach. *European Journal of Operational Research*, **104**, 497–509.

Wilson, K.A., Underwood, E.C., Morrison, S.A., Klausmeyer, K.R., Murdoch, W.W., Reyers, B., Wardell-Johnson, G., Marquet, P.A., Rundel, P.W., McBride, M.F., Pressey, R.L., Bode, M., Hoekstra, J.M., Andelman, S., Looker, M., Rondinini, C., Kareiva, P., Shaw, M.R. & Possingham, H.P. (2007) Conserving biodiversity efficiently: What to do, where, and when. *PLoS Biology*, **5**, e223.

Wilson, K.A., Westphal, M.I., Possingham, H.P. & Elith, J. (2005) Sensitivity of conservation planning to different approaches to using predicted species distribution data. *Biological Conservation*, **122**, 99 – 112.

Wright, D. (1977) A site evaluation scheme for use in the assessment of potential nature reserves. *Biological Conservation*, **11**, 293 – 305.

Wright, J.W., Stanton, M.L. & Scherson, R. (2006) Local adaptation to serpentine and non-serpentine soils in *Collinsia sparsiflora*. *Evolutionary Ecology Research*, **8**, 1–21.

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

Wright, S. (1929) The evolution of dominance. *The American Naturalist*, **63**, 556–561.

Wright, S. (1931) Evolution in Mendelian populations. *Genetics*, **16**, 97–159.

Wright, S. (1932) The roles of mutation, inbreeding, crossbreeding, and selection in evolution. *Proceedings of the Sixth International Congress of Genetics*, **1**, 356–366.

Zeller, K.A., McGarigal, K. & Whiteley, A.R. (2012) Estimating landscape resistance to movement: a review. *Landscape Ecology*, **27**, 777–797.

Appendix A

The reliable formulation explicitly considers the probability that planning units are inhabited by features. As a consequence, it may deliver prioritizations that sufficiently conserve an attribute space even if the features do not inhabit several of the planning units when the prioritization is implemented. This behavior is achieved by siting back-up planning units near selected planning units with low occupancy probabilities in the attribute space(s). To ensure that the solution is robust against multiple planning units being uninhabited, demand points are treated as being conserved by different planning units at different backup levels.

Backup levels are defined as r-levels (similar to failure levels in Snyder & Daskin 2005). The first backup r-level is used to calculate how well the demand points are conserved assuming that all the selected planning units are occupied by all $f \in F$. For this scenario, the closest selected planning unit to each demand point i for attribute space s is assigned at r-level = 0. This scenario essentially serves as Y_{fsij} in the unreliable formulation. The second backup r-level is used to assess how well each demand point is conserved when the closest planning unit to each demand point i is unoccupied. For this scenario, the second closest planning units are assigned at r-level = 1. The third backup r-level is used to assess how well demand points are captured when the first two closest planning units are unoccupied. The third closest planning units are assigned at r-level = 2 . Continuing on, in this manner, the selected planning units in a prioritization are assigned to each demand point $i \in I$, attribute space $s \in S$, and each feature $f \in F$ at an r-level.

A final backup r-level ($r = R$) is used to describe how well demand points are captured by a solution when the features $f \in F$ do not occupy any selected planning units. Each demand point $i \in I$ for each $s \in S$ and $f \in F$ is assigned to an "imaginary" planning unit $j = J$ at $r = R$. The distance variables associated with this imaginary planning unit d_{fsiJ} denote the loss of biological value associated with failing to secure a representative sample of feature f in attribute space s. However, the d variables are in distance units which effectively meaningless in this context. Thus these variables are calculated using a failure multiplier (M) and the maximum distance between the planning units and the demand points for $f \in F$, $s \in S$ (eqn 1).

$$d_{fsiJ} = M \max_{0 \le i \le I-1, 0 \le j \le J-1} d_{fsij} \qquad \forall 0 \le f \le F-1, \; 0 \le s \le S-1 \qquad \text{eqn 1}$$

Conservation planning problems often involve several hundred or more planning units. It is currently not feasible to solve problems this large when considering all possible failure scenarios (see Appendix B). To reduce computational burden, users can set a smaller number to denote the maximum number of backup levels (R). For instance, when $R = 3$ only 2 backup levels are considered in addition to the final backup level. Cui *et al.* (2010) found that $R = 5$ yields comparable solutions to when R is much larger. However, in most cases the decision maker will be limited to $R = 1$ to obtain prioritizations in a feasible amount of time.

The control variables for the reliable formulation are the B (eqn 2.1a), T_s (eqn 2.1b), τ_{fs} (eqn 2.1c), R, and M variables.

$R =$ number of backup levels — eqn 2a

$M =$ failure multiplier — eqn 2b

The decision variables are the X_j (eqn 2a), Y_{fsijr}, P_{fsijr} variables.

$$Y_{fsijr} = \begin{cases} 1, & \text{if demand point } i \text{ for feature } f \text{ in space } s \text{ is conserved by planning unit } j \text{ at back-up level } r. \\ 0, & \text{otherwise} \end{cases} \qquad \text{eqn 3a}$$

$P_{fsijr} =$ probability that demand point i for feature f in space s will need to be conserved by planning unit j at back-up level r — eqn 3b

The reliable formulation (RRAP) is a multi-objective optimization problem.

(RRAP) Min (eqn 2.3a)

s.t. (eqn 2.3b)

$$1 - \frac{\sum_{i=0}^{I-1}\sum_{j=0}^{J-1}\lambda_{fsi}{d_{fsij}}^2 P_{fsijr}Y_{fsij}}{\sum_{i=0}^{I-1}\lambda_{fsi}{\delta_{fsi}}^2} \geq T_{fs} \qquad \forall 0 \leq f \leq F-1, \quad 0 \leq s \leq S-1 \qquad \text{eqn 4a}$$

$$\sum_{j=0}^{J-1} Y_{fsijr} = 1 \qquad \forall 0 \leq f \leq F-1, \quad 0 \leq s \leq S-1, \quad 0 \leq i \leq I-1, \quad 0 \leq r \leq R \qquad \text{eqn 4b}$$

$$\sum_{r=0}^{R} Y_{fsijr} = 1 \qquad \forall 0 \leq f \leq F-1, \quad 0 \leq s \leq S-1, \quad 0 \leq i \leq I-1, \quad 0 \leq j \leq J \qquad \text{eqn 4c}$$

$$\sum_{r=0}^{R-1} Y_{fsijr} \leq X_j \qquad \forall 0 \leq f \leq F-1, \quad 0 \leq s \leq S-1, \quad 0 \leq i \leq I-1, \quad 0 \leq j \leq J-1 \qquad \text{eqn 4d}$$

$$Y_{fsiJR} = 1 \qquad \forall 0 \leq f \leq F-1, \quad 0 \leq s \leq S-1, \quad 0 \leq i \leq I-1 \qquad \text{eqn 4e}$$

$$P_{fsij0} = q_{fj} \qquad \forall 0 \leq f \leq F-1, \quad 0 \leq s \leq S-1, \quad 0 \leq i \leq I-1, \quad 0 \leq j \leq J \qquad \text{eqn 4f}$$

$$P_{fsijr} = (1-)\sum_{k=0}^{J-1} \frac{1-q_k}{q_k} P_{f,s,i,k,r-1}Y_{f,s,i,k,r-1} \qquad \forall 0 \leq f \leq F-1, \quad 0 \leq s \leq S-1, \quad 0 \leq i \leq I-1, \quad 0 \leq j \leq J, \quad 1 \leq r \leq R \qquad \text{eqn 4g}$$

$$X_j, Y_{fsijr} \in 0, 1 \qquad \forall 0 \leq f \leq F-1, \quad 0 \leq s \leq S-1, \quad 0 \leq i \leq I-1, \quad 0 \leq j \leq J, \qquad \text{eqn 4h}$$

$$0 \leq r \leq R$$

The objective function for the reliable formulation is the same as for the unreliable formation (eqn 2.3a). Similar to the unreliable formulation, constraints (eqn 2.3b) and (eqn 4a) ensure that the amount-based and space-based targets are met. Constraints (eqns 4b–4c) ensure that each planning unit is only assigned to one backup r-level for $i \in I$. Constraints (eqn 4d) ensure that only selected planning units are assigned to demand points $i \in I$. Constraints (eqn 4e) ensure that the imaginary planning unit is always assigned to the highest backup r-level. Constraints (eqns 4f–4g) determine the probability that planning unit j will be used to sample demand point $i \in I$ for $s \in S$ and $f \in F$ (see Cui *et al.* 2010 for more information). Constraints (eqn 4h) ensure that the X and Y variables are binary.

The reliable formulation is non-linear. However, the non-linear components can be linearized. First—as discussed in the main text—the expression $X_j X_k$ in (eqn 2.3a) can be linearized using methods described by Beyer et al. (2016). Second, the expression $P_{fsijr} Y_{fsijr}$ in (eqn 4a) can be linearized using techniques described by Sherali and Alameddine (1992) as implemented in Cui et al. (2010).

Appendix B

Symbols and descriptions of terms used in the formulation of the unreliable representative and adequate prioritization (URAP) problem and reliable version (RRAP).

Symbol	Description
A_j	area of planning unit j—note this can also be some other measure of conservation utility
B	boundary length modifier (BLM)
d_{fsij}	distance between demand point i and planning unit j for feature f in attribute space s
e_{jk}	shared edge between planning unit j and planning unit k. Note that when $j == k$ this used to parametrize exposed edges with no neighbours.
F	set of biodiversity features (indexed by f)
I_{fsi}	set of demand points (indexed by i) for a feature f in attribute space s
J	set of planning units (indexed by j)
M	failure multiplier
P_{fsijr}	probability that demand point i for feature f in space s will need to be conserved by planning unit j at back-up level r
q_{fj}	probability of feature f occupying planning unit j
R	number of backup levels (indexed by r)
S	set of attribute spaces (indexed by S)
T_f	amount target for feature f
X_j	binary decision variable controlling if a planning unit is selected for preservation (1) or discarded (0)
Y_{fsij}	binary decision variable indicating if planning unit j is assigned to demand point i for species s in attribute space s when determining the amount of the attribute space sampled by the selected planning units (unreliable version)
Y_{fsijr}	binary decision variable indicating if planning unit j is assigned to demand point i for species s in attribute space s at r-level r when determining the amount of the attribute space sampled by the selected planning units (reliable version)
δ_{fsi}	the distance between each demand point i and the centroid of the demand points I for feature f in attribute space s
λ_{fsi}	set of weights for demand point i for feature f in attribute space s
τ_{fs}	space-based target for feature f in attribute space s

Appendix C

We conducted a benchmark analysis to understand how much time would be required to solve different sized problems using the `raptr R` package.

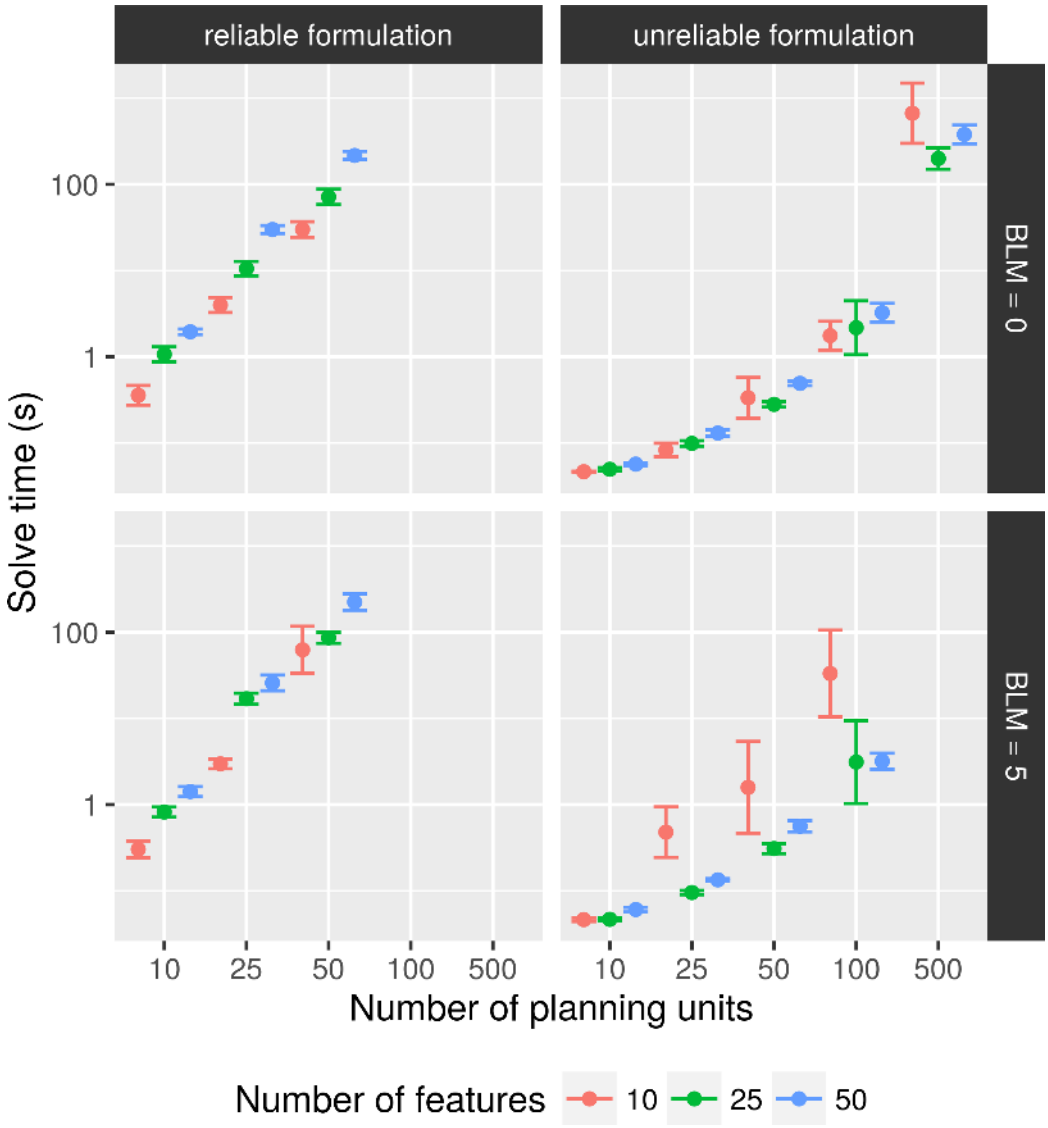
Methods

We simulated conservation planning data sets containing different numbers of features and planning units (see figure below for numbers). For a given number of planning units and features, a data set was simulated using the following steps. First, the planning units were simulated as squares that were arranged in a grid. Second, the planning unit costs were set to one. Third, the spatial distribution of each feature in the planning units was simulated using binary, spatially autocorrelated random fields (using the `RandomFields R` package; Schlather *et al.* 2015, 2016). Fourth, if a feature was found to occupy less than two planning units, the feature was set to occupy two randomly selected planning units. All occupied planning units were set to have a probability of occupancy of 0.9 for all features. Fifth, similar to the simulated data sets described in the main text, a geographic attribute space was generated for each feature. The features' demand points were set as the centroids of the planning units that they occupied, and were weighted by the units' probability of occupancy. After simulating the data sets, solutions were generated using the `raptr R` package and `Gurobi` (version 7.0.2).

We generated solutions for the conservation planning data sets using different formulations and parameters. Specifically, the solutions were generated using the unreliable and reliable formulation, and two different boundary length multipliers (0 and 5). The solutions were generated using amount-based targets of 20 % and space-based targets of 50 %. To assess variation in solve times, we generated and solved 4 random data sets per combination of number of features and planning units. All problems were solved to within 10 % of optimality and were solved using a single thread on a system with 64 GB RAM and an Intel(R) Xeon (R) CPU (E5-2630 v3; 2.40 GHz).

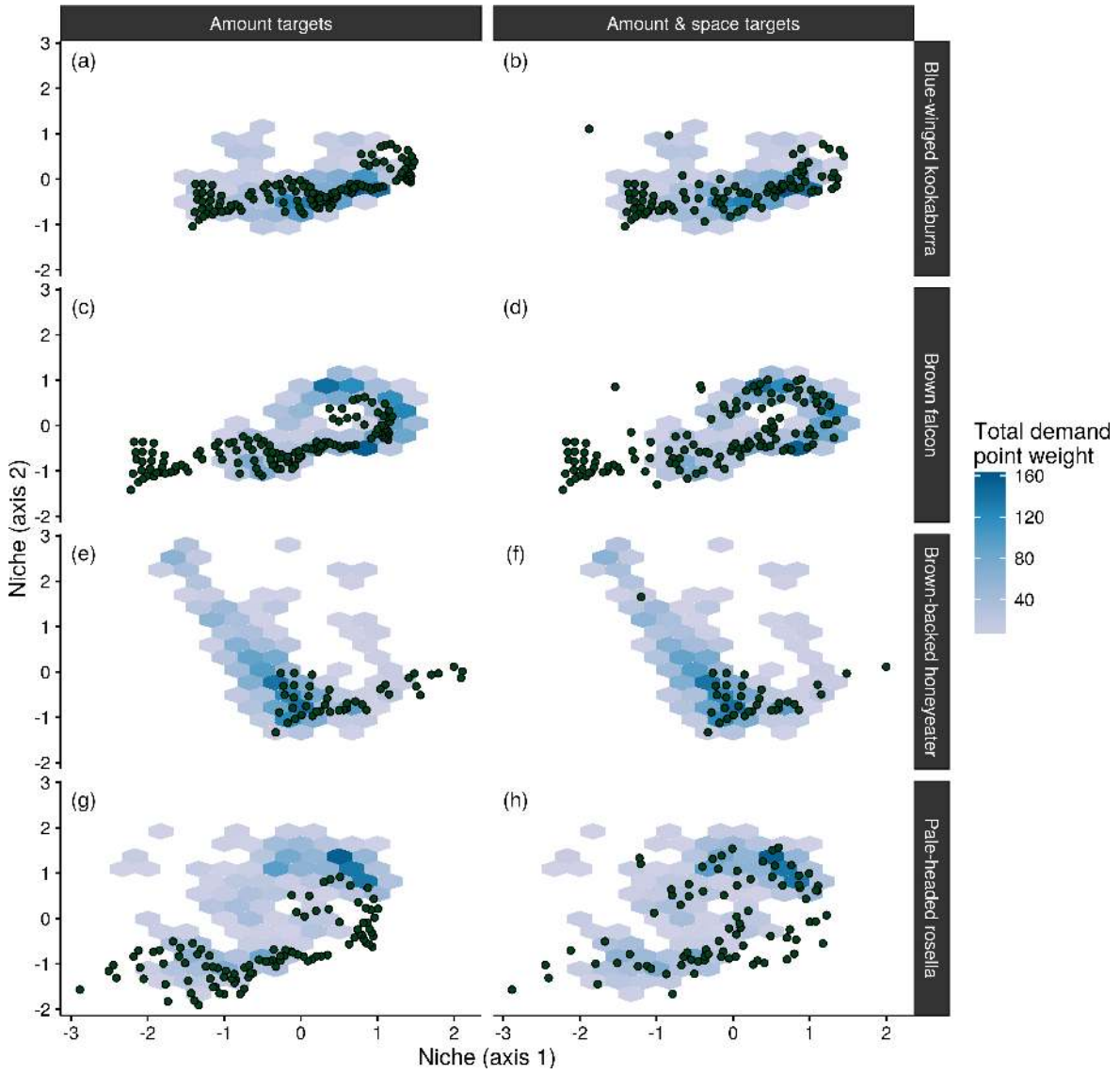
Results and discussion

The benchmark analysis revealed that moderately-sized conservation problems can be solved in a relatively short period of time. Problems were solved much quicker when using the unreliable formulation than when using the reliable formulation. The solve times could be further reduced by using more threads. We were not able to determine how long bigger problems would take to solve using the reliable formulation because the system did not have enough RAM. Future research could involve implementing more advanced methods to solve the problems (eg. as described in Cui *et al.* 2010; Aboolian *et al.* 2013).



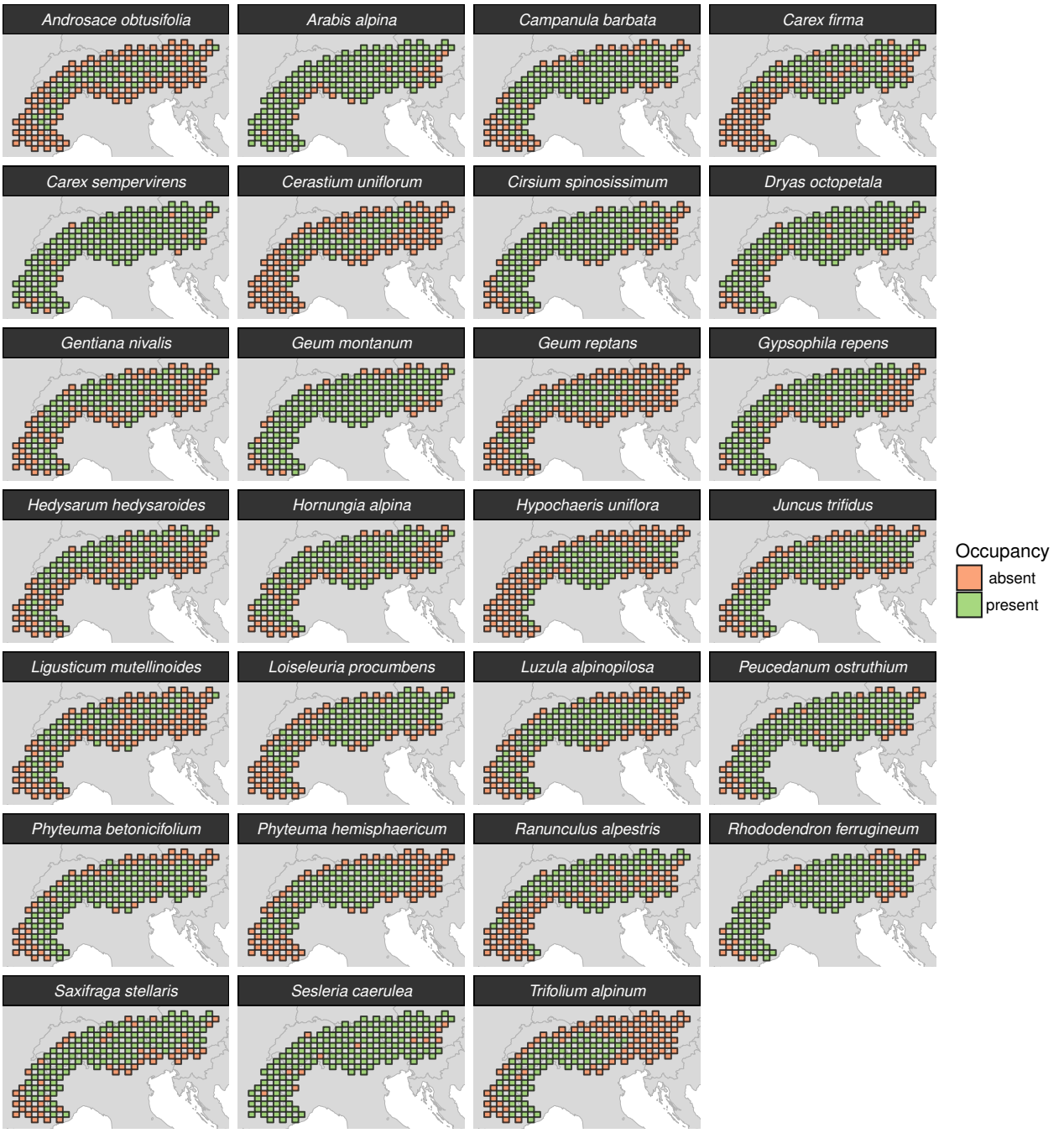
Time required to solve different sized problems using different parameters and formulations. Points denote average times and bars show standard errors. Time is shown on a $\log_{10}$ scale.

Appendix D



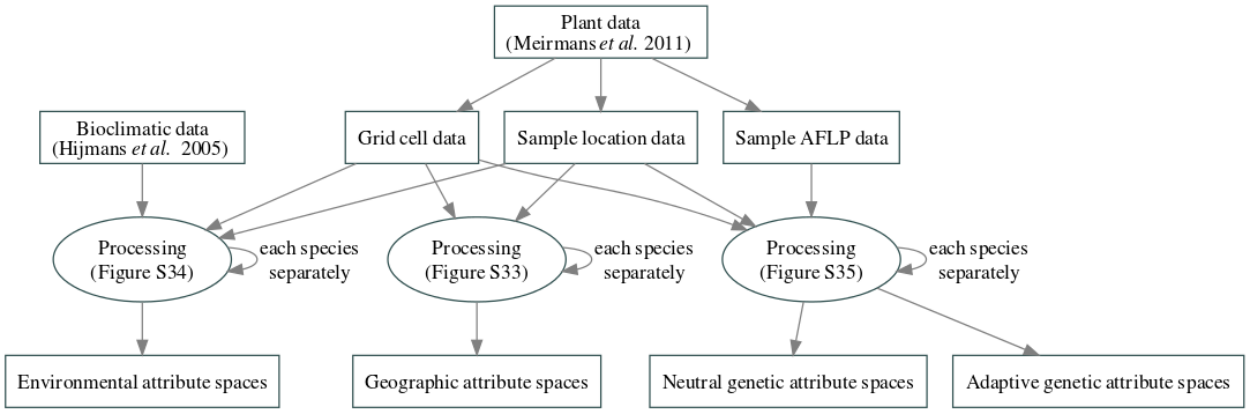
Attribute spaces used in the first case-study. Each panel shows the distribution of a solution in environmental space and how it samples the realized niche for a different species. The left column of panels shows the solution generated using amount targets. The right column of panels shows the solution generated using amount and space targets. Each column of panels corresponds to a different species. Hexagons show the distribution of demand points. The color of each hexagon denotes the weighted frequency of demand points inside it. Points denote the environmental conditions associated with planning units inside the species geographic range that were selected for preservation in a given solution.

Appendix E

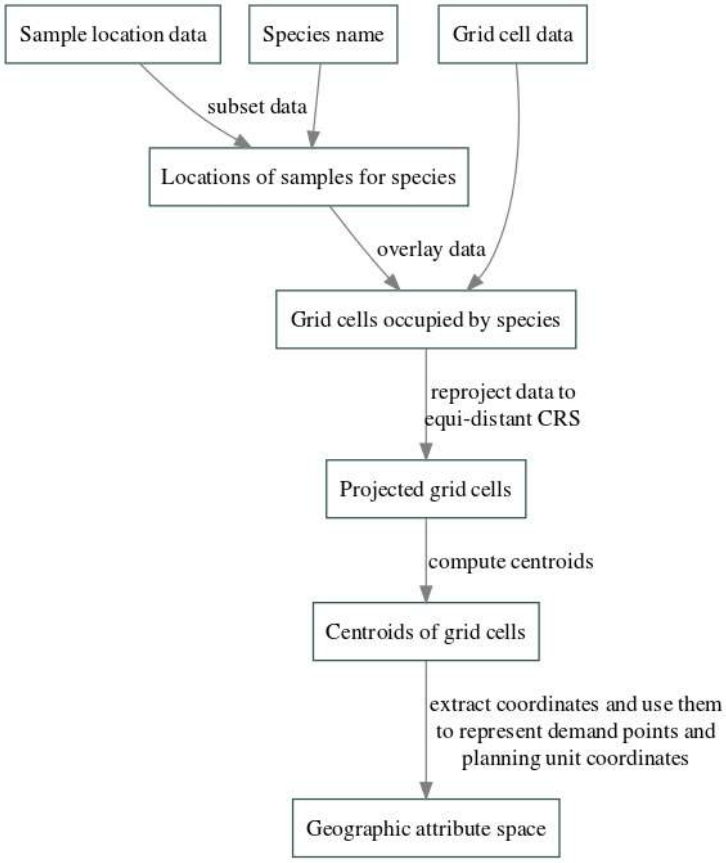


Species distributions. Squares represent planning units. For a given species, planning units that it was found in are shown in green.

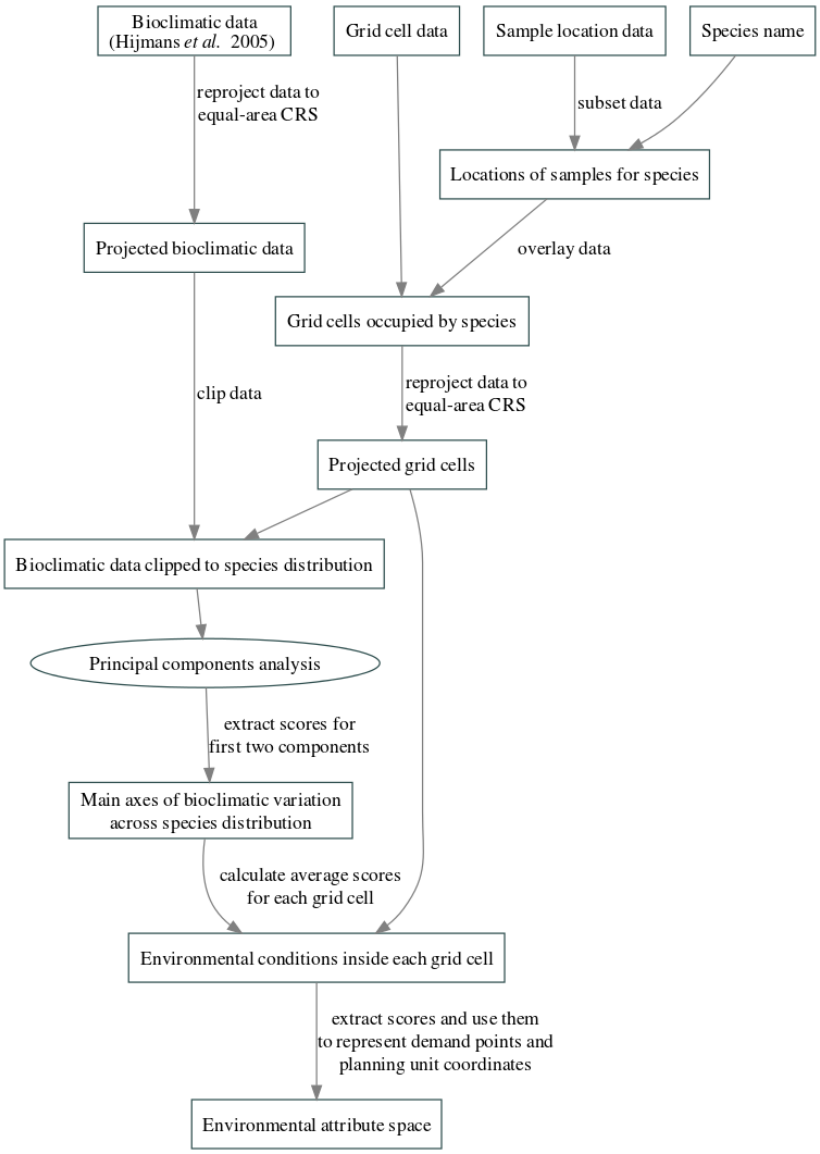
Appendix F



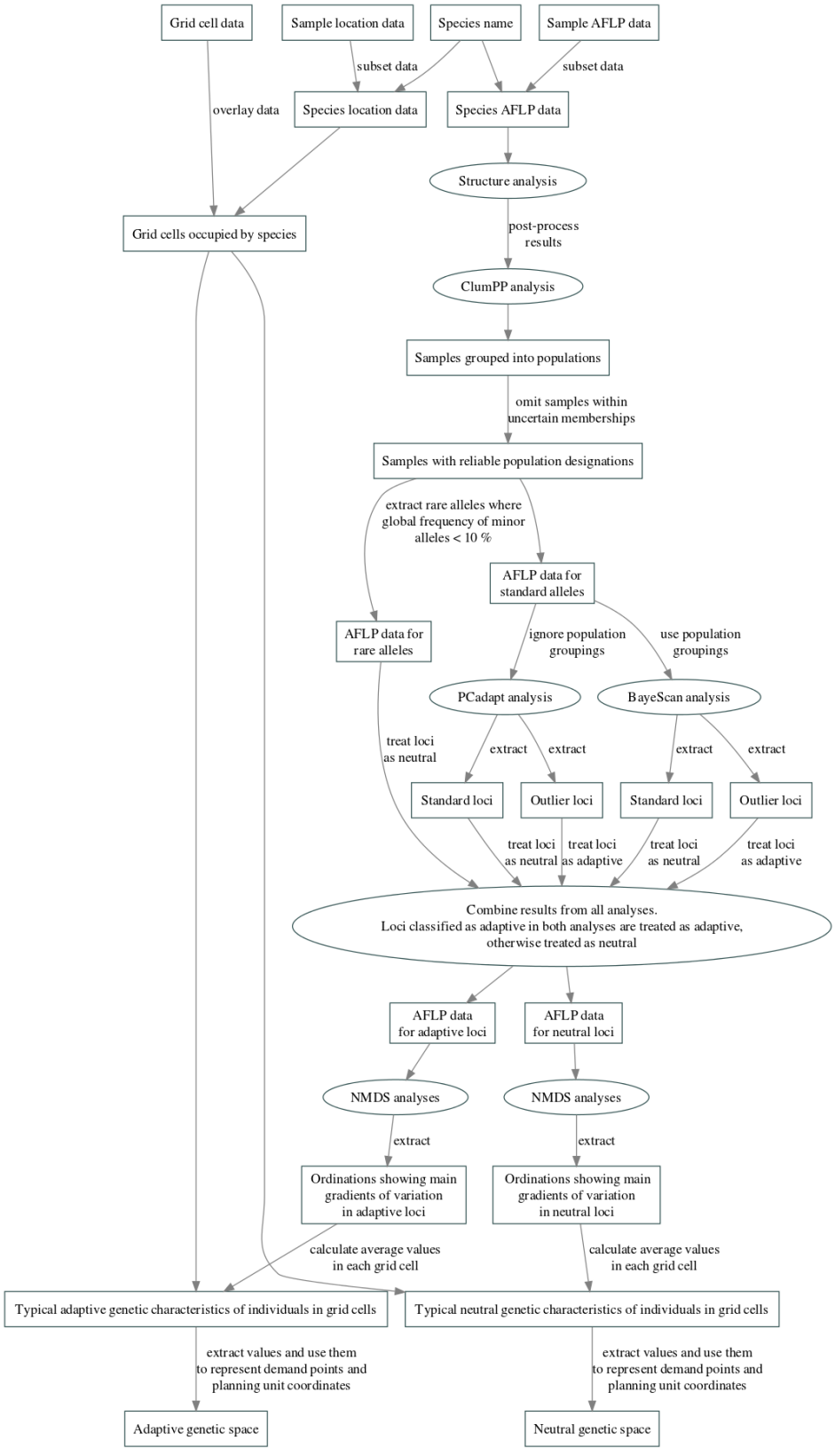
Overall data processing workflow. The flowcharts show the methods used to generate the environmental and geographic surrogate spaces, and the adaptive and neutral genetic spaces. Squares indicate data sets. Ellipses indicate analyses. Lines show relationships between data and analysis, and their captions describe how data was processed.



Geographic surrogate processing workflow. Conventions are detailed the in caption for first figure in Appendix F.



Environmental surrogate processing workflow. Conventions are detailed the in caption for first figure in Appendix F.



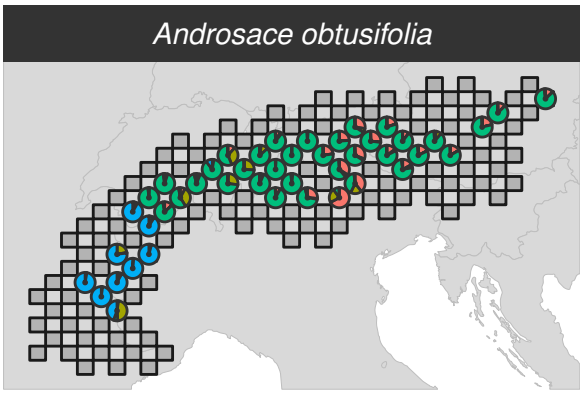
Genetic data processing workflow. Conventions are detailed the in caption for first figure in Appendix F.

Appendix G

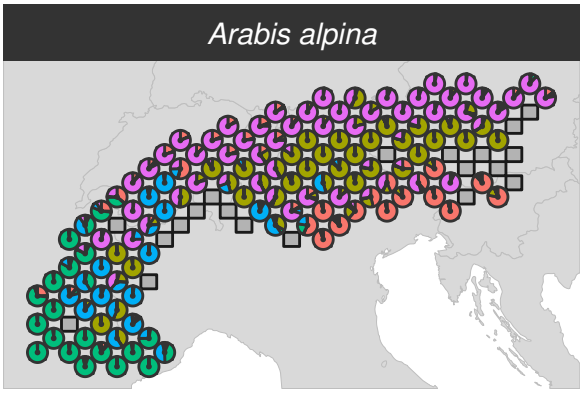
Summary of principal component analyses fit to the climatic variation in the species' ranges. Percentages correspond to the percent of variation described by the first three principle components.

Species	Variation explained (%)
Androsace obtusifolia	90.374
Arabis alpina	91.319
Campanula barbata	89.931
Carex firma	91.017
Carex sempervirens	90.212
Cerastium uniflorum	91.454
Cirsium spinosissimum	89.628
Dryas octopetala	90.150
Gentiana nivalis	90.191
Geum montanum	89.742
Geum reptans	90.255
Gypsophila repens	91.679
Hedysarum hedysaroides	89.719
Hornungia alpina	90.383
Hypochaeris uniflora	93.062
Juncus trifidus	91.062
Ligusticum mutellinoides	88.709
Loiseleuria procumbens	91.679
Luzula alpinopilosa	88.776
Peucedanum ostruthium	89.595
Phyteuma betonicifolium	89.678
Phyteuma hemisphaericum	88.909
Ranunculus alpestris	90.019
Rhododendron ferrugineum	89.779
Saxifraga stellaris	89.484
Sesleria caerulea	90.589
Trifolium alpinum	89.536

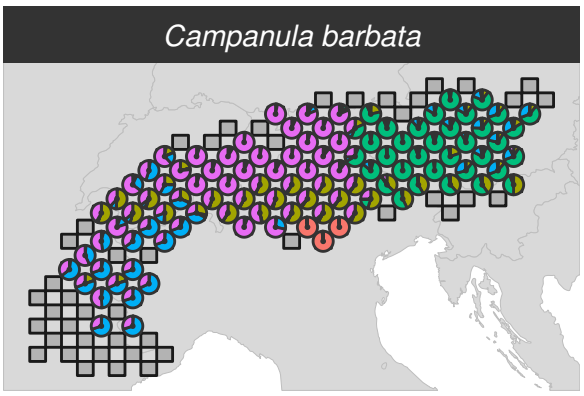
Appendix H



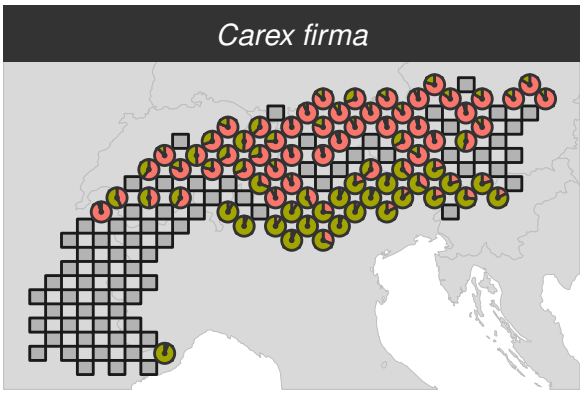
Distribution of *Androsace obtusifolia* populations. Squares denote planning units. Pie charts denote quadrats where individuals were sampled, and colors denote the average probability that individuals in the quadrat belong to different populations on average.



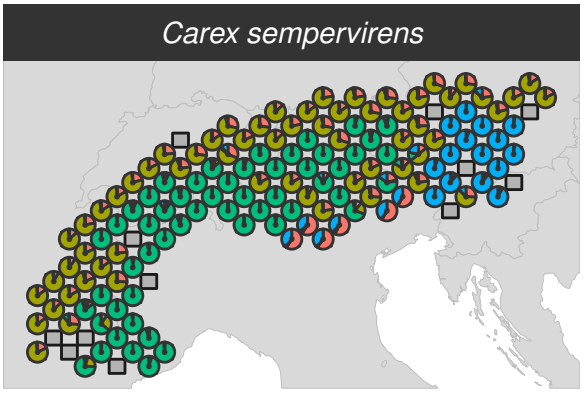
Distribution of *Arabis alpina* populations. See caption for first figure in Appendix H for conventions.



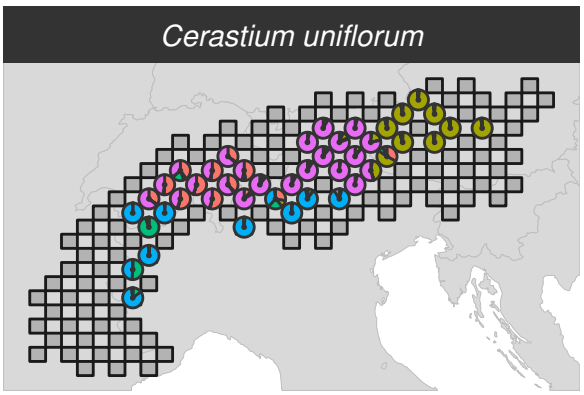
Distribution of *Campanula barbata* populations. See caption for first figure in Appendix H for conventions.



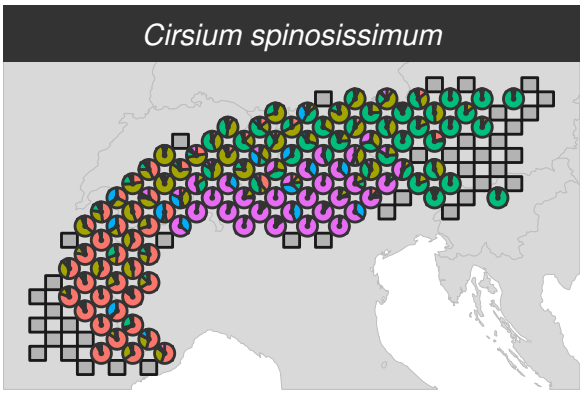
Distribution of *Carex firma* populations. See caption for first figure in Appendix H for conventions.



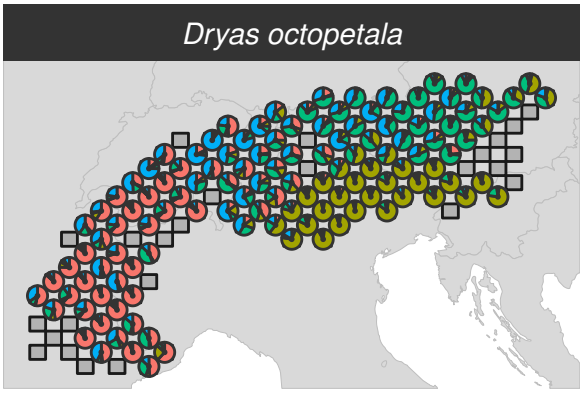
Distribution of *Carex sempervirens* populations. See caption for first figure in Appendix H for conventions.



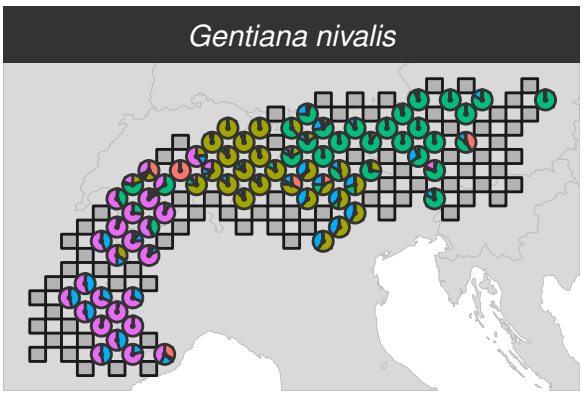
Distribution of *Cerastium uniflorum* populations. See caption for first figure in Appendix H for conventions.



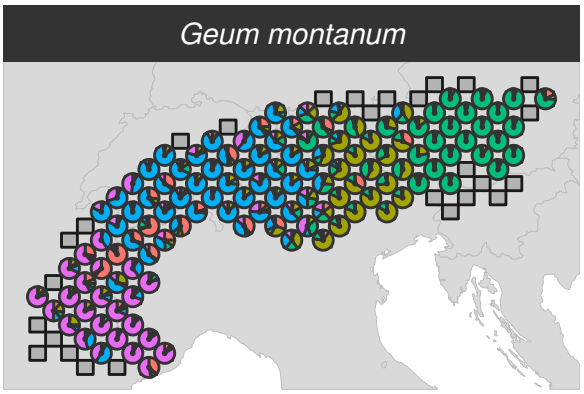
Distribution of *Cirsium spinosissimum* populations. See caption for first figure in Appendix H for conventions.



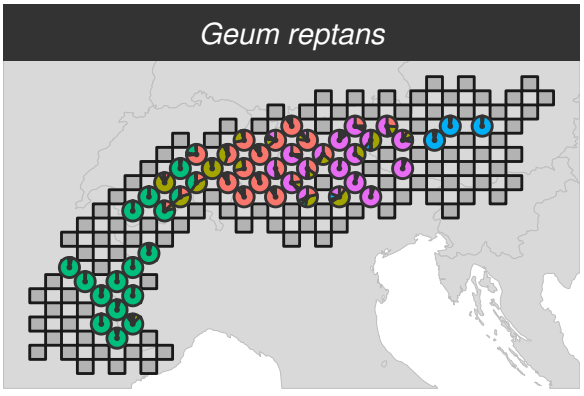
Distribution of *Dryas octopetala* populations. See caption for first figure in Appendix H for conventions.



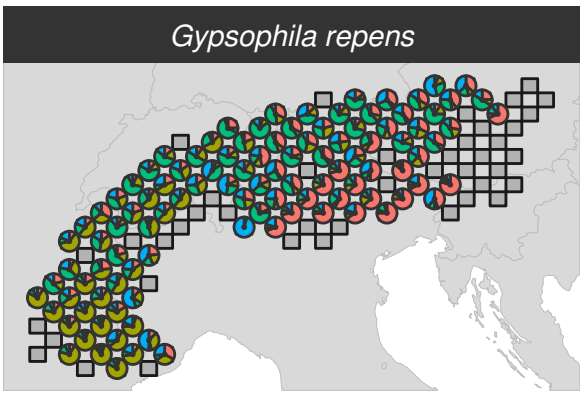
Distribution of *Gentiana nivalis* populations. See caption for first figure in Appendix H for conventions.



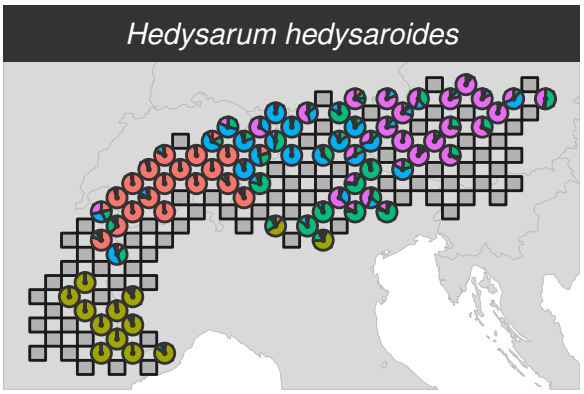
Distribution of *Geum montanum* populations. See caption for first figure in Appendix H for conventions.



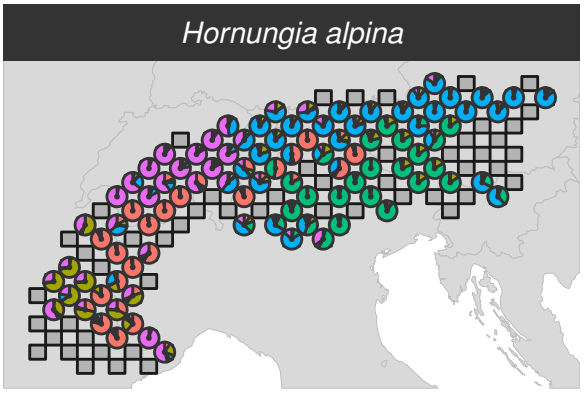
Distribution of *Geum reptans* populations. See caption for first figure in Appendix H for conventions.



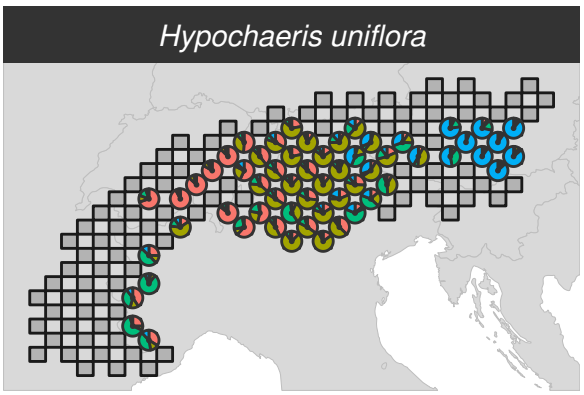
Distribution of *Gypsophila repens* populations. See caption for first figure in Appendix H for conventions.



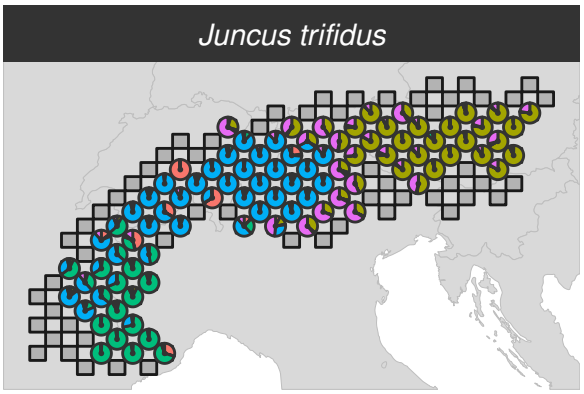
Distribution of *Hedysarum hedysaroides* populations. See caption for first figure in Appendix H for conventions.



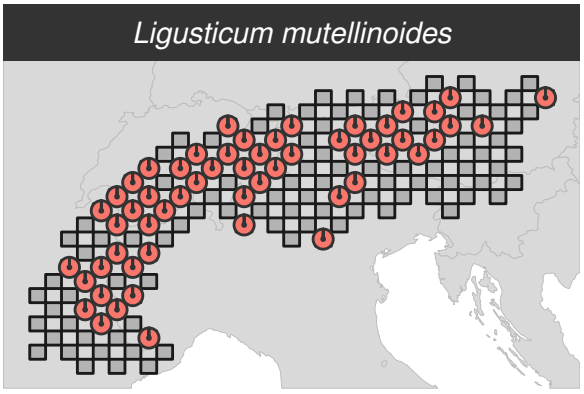
Distribution of *Hornungia alpina* populations. See caption for first figure in Appendix H for conventions.



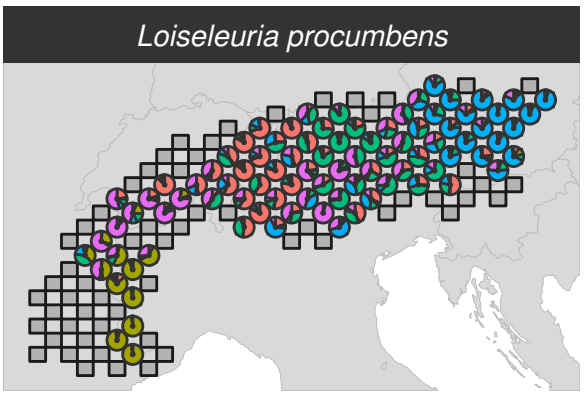
Distribution of *Hypochaeris uniflora* populations. See caption for first figure in Appendix H for conventions.



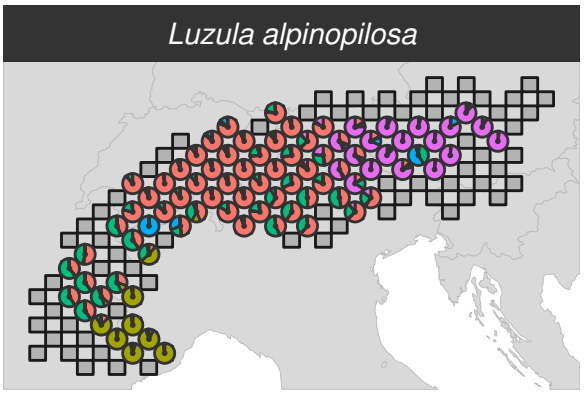
Distribution of *Juncus trifidus* populations. See caption for first figure in Appendix H for conventions.



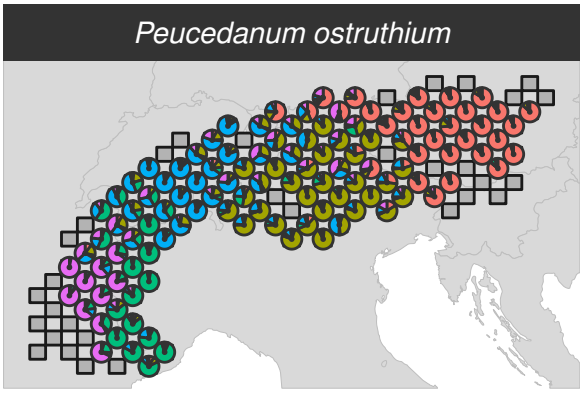
Distribution of *Ligusticum mutellinoides* populations. See caption for first figure in Appendix H for conventions.



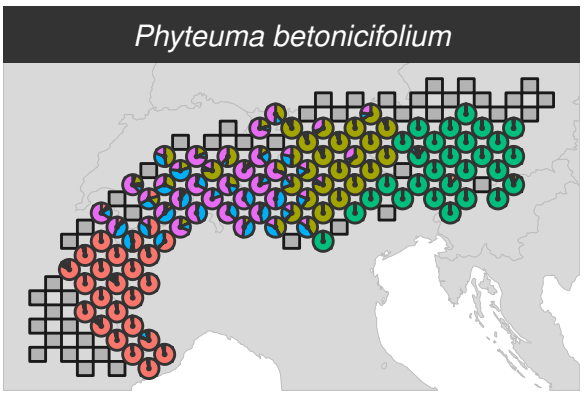
Distribution of *Loiseleuria procumbens* populations. See caption for first figure in Appendix H for conventions.



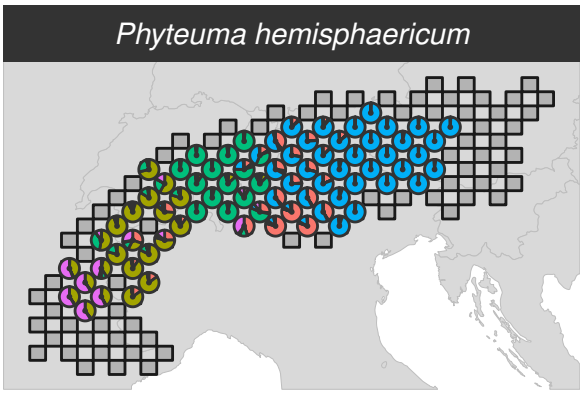
Distribution of *Luzula alpinopilosa* populations. See caption for first figure in Appendix H for conventions.



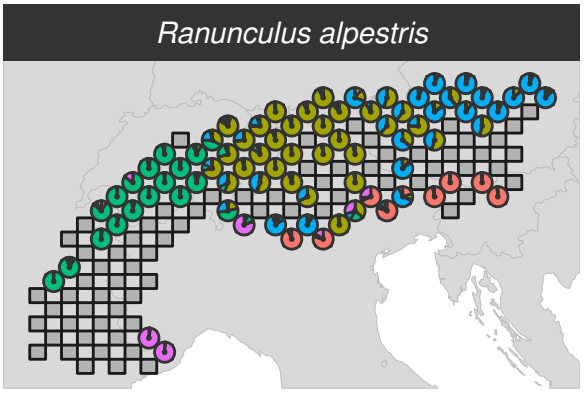
Distribution of *Peucedanum ostruthium* populations. See caption for first figure in Appendix H for conventions.



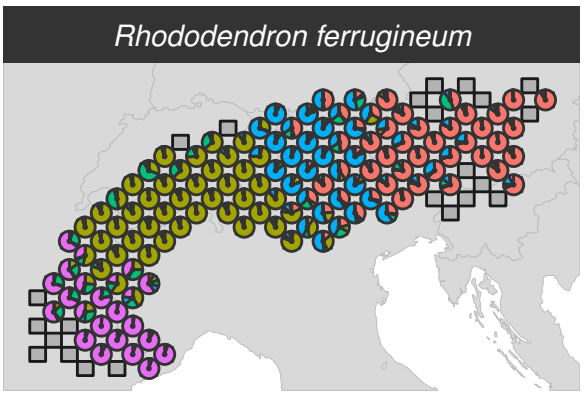
Distribution of *Phyteuma betonicifolium* populations. See caption for first figure in Appendix H for conventions.



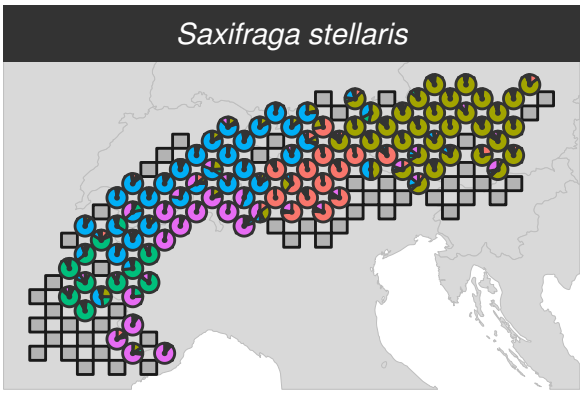
Distribution of *Phyteuma hemisphaericum* populations. See caption for first figure in Appendix H for conventions.



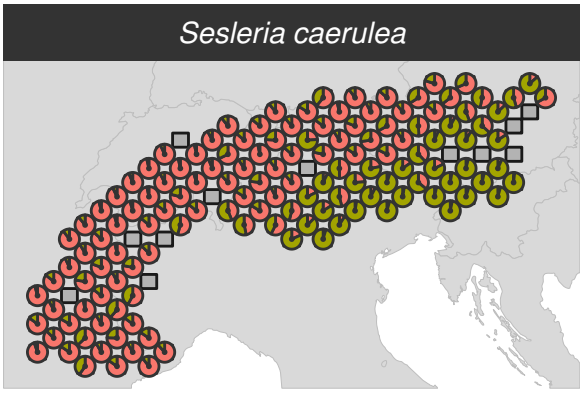
Distribution of *Ranunculus alpestris* populations. See caption for first figure in Appendix H for conventions.



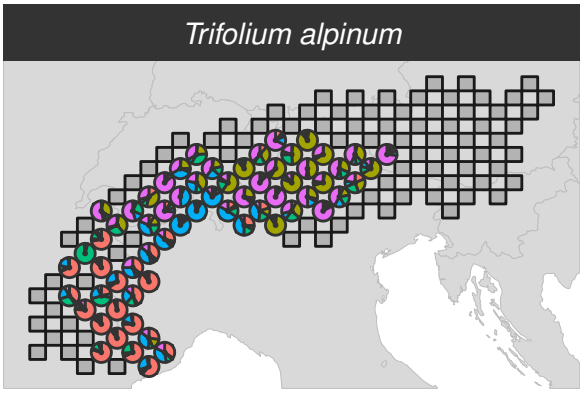
Distribution of *Rhododendron ferrugineum* populations. See caption for first figure in Appendix H for conventions.



Distribution of *Saxifraga stellaris* populations. See caption for first figure in Appendix H for conventions.



Distribution of *Sesleria caerulea* populations. See caption for first figure in Appendix H for conventions.



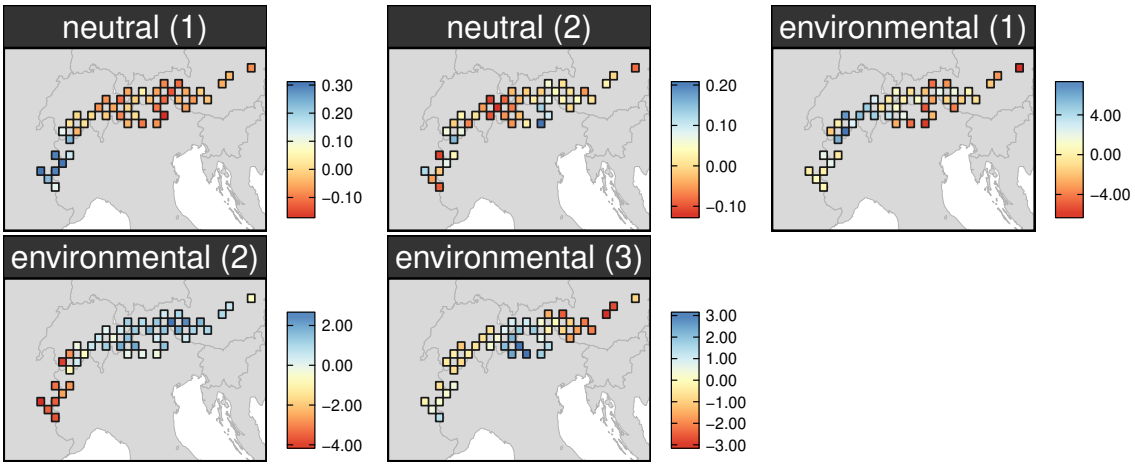
Distribution of *Trifolium alpinum* populations. See caption for first figure in Appendix H for conventions.

Appendix I

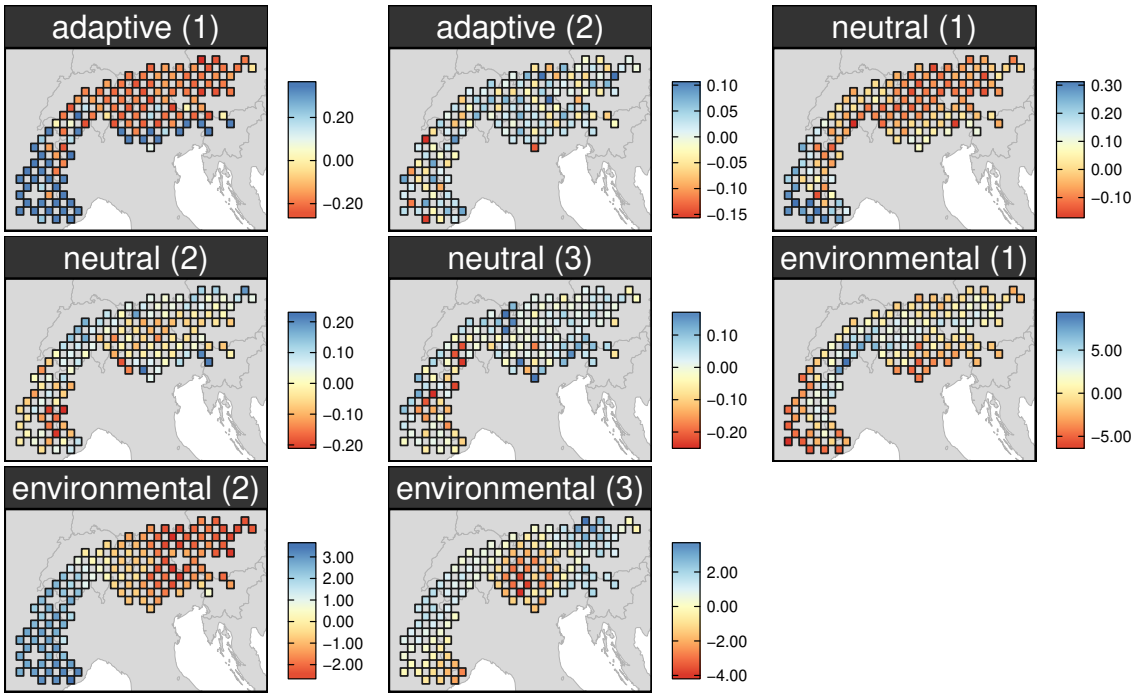
Summary of non-metric multi-dimensional scaling (NMDS) analyses on genetic variation for each species. The analyses were used to identify the main gradients of variation in binary adaptive and neutral loci data. Bold stress values indicate NMDS analyses that have converged.

Species	Adaptive		Neutral	
	K	Stress	K	Stress
Androsace obtusifolia			2	0.218
Arabis alpina	2	<0.001	3	**0.187**
Campanula barbata			2	0.204
Carex firma			3	0.241
Carex sempervirens	2	<0.001	2	0.219
Cerastium uniflorum	2	<0.001	2	0.205
Cirsium spinosissimum	2	<0.001	3	**0.197**
Dryas octopetala			3	0.218
Gentiana nivalis			2	**0.046**
Geum montanum			2	0.163
Geum reptans	2	<0.001	2	0.188
Gypsophila repens			2	**0.06**
Hedysarum hedysaroides	2	<0.001	2	0.197
Hornungia alpina			3	0.2
Hypochaeris uniflora			2	0.241
Juncus trifidus			2	0.151
Ligusticum mutellinoides			4	0.212
Loiseleuria procumbens	2	<0.001	3	0.208
Luzula alpinopilosa			2	**0.051**
Peucedanum ostruthium	2	<0.001	4	0.221
Phyteuma betonicifolium			2	0.172
Phyteuma hemisphaericum	2	0.003	2	0.226
Ranunculus alpestris	2	**0.012**	3	0.221
Rhododendron ferrugineum			2	0.208
Saxifraga stellaris			3	0.201
Sesleria caerulea			4	0.209
Trifolium alpinum			3	0.212

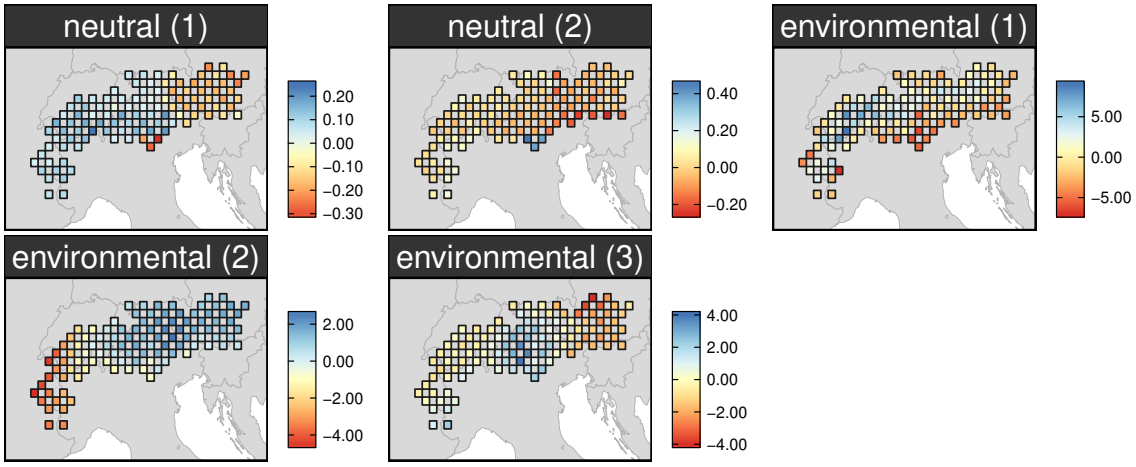
Appendix J



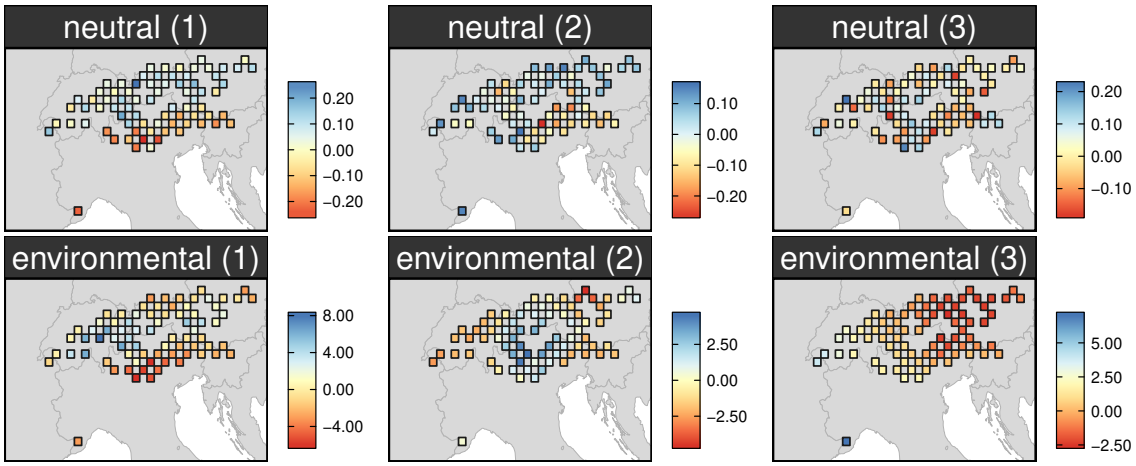
Distribution of environmental variation found across the geographic range of *Androsace obtusifolia*, and the distribution of its adaptive (if detected) and neutral genetic variation in the study area. Each panel shows a different dimension of variation as indicated by its title. Squares represent planning units. Samples collected from the species were subject to NMDS to ordinate the main axes of adaptive (if detected) and neutral genetic variation. The bioclimatic data were clipped to planning units where the species was found, and subject to a principal components analysis to summarize the main gradients of environmental variation. The panels show the spatial distribution of the ordinated environmental and genetic data averaged to the planning-unit level. Planning units where individuals were not observed are shown in gray.



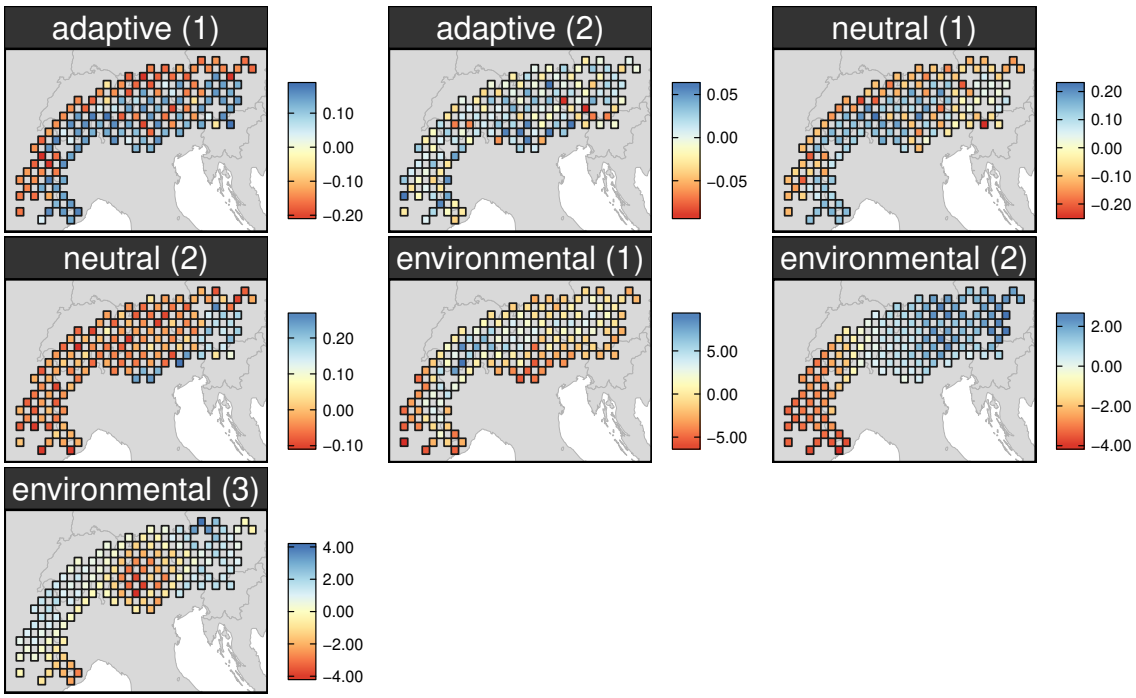
Distribution of intra-specific variation for *Arabis alpina*. See caption for first figure in Appendix J for conventions.



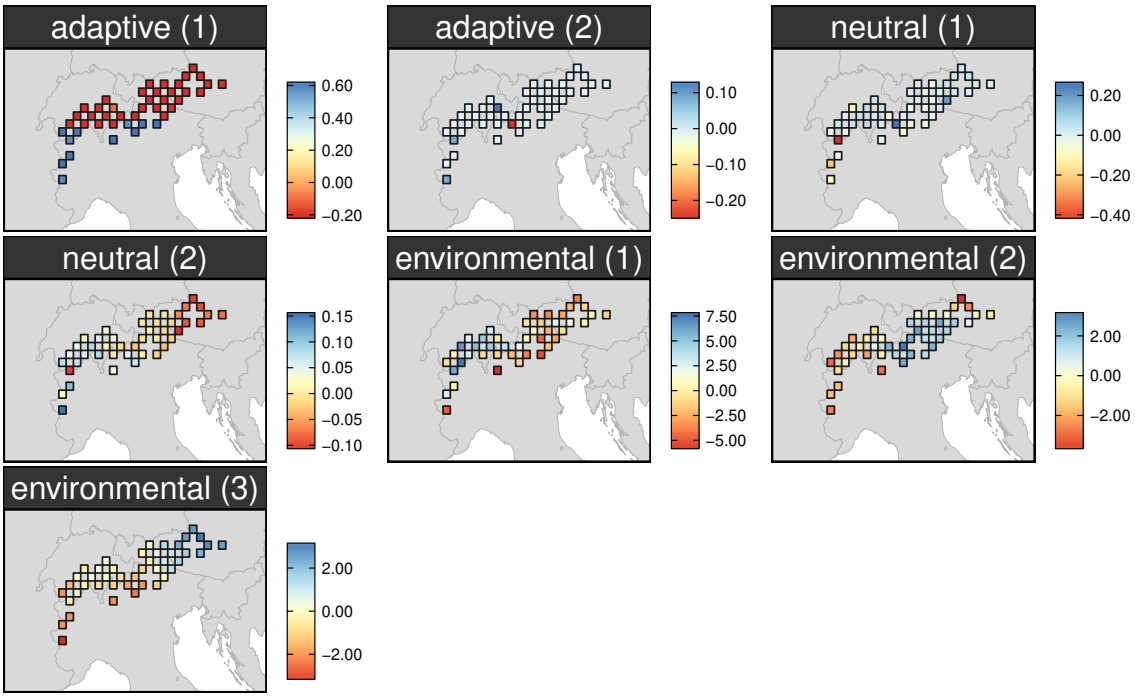
Distribution of intra-specific variation for *Campanula barbata*. See caption for first figure in Appendix J for conventions.



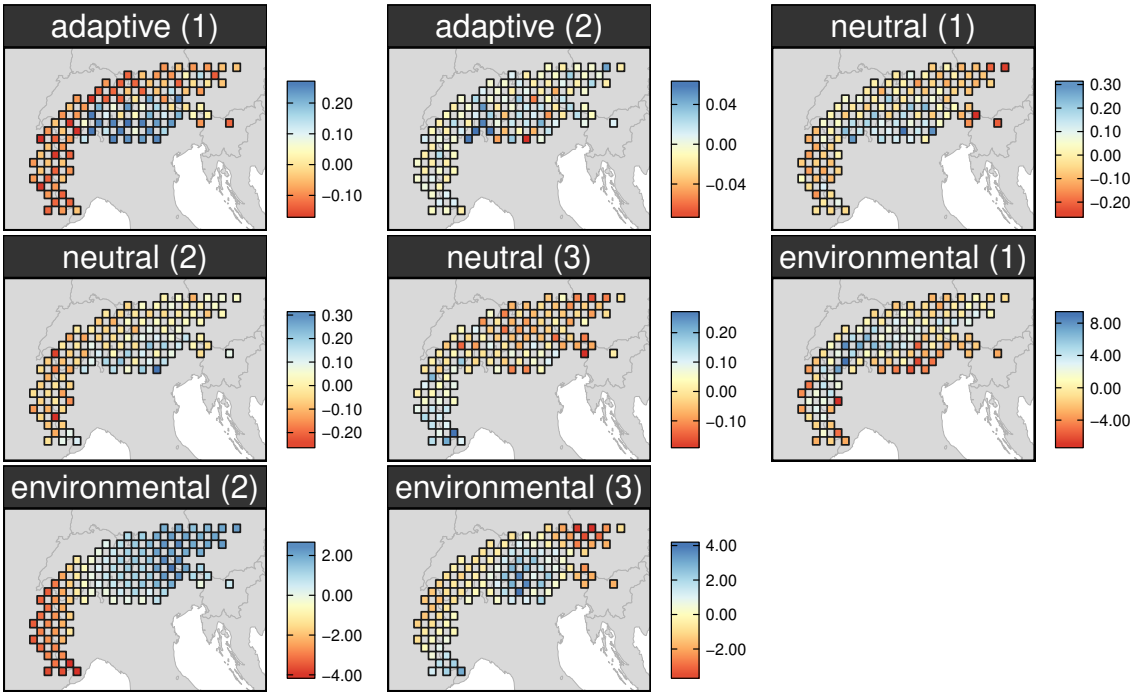
Distribution of intra-specific variation for *Carex firma*. See caption for first figure in Appendix J for conventions.



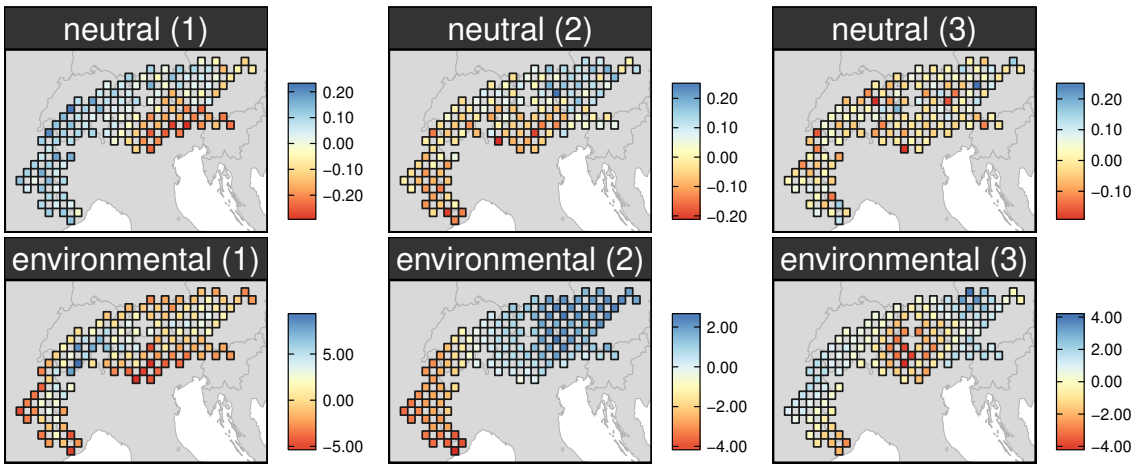
Distribution of intra-specific variation for *Carex sempervirens*. See caption for first figure in Appendix J for conventions.



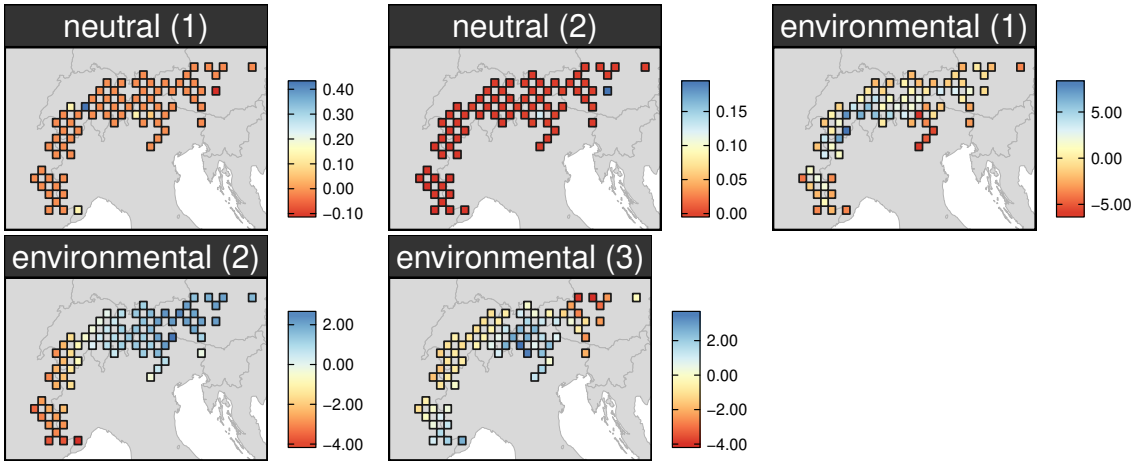
Distribution of intra-specific variation for *Cerastium uniflorum*. See caption for first figure in Appendix J for conventions.



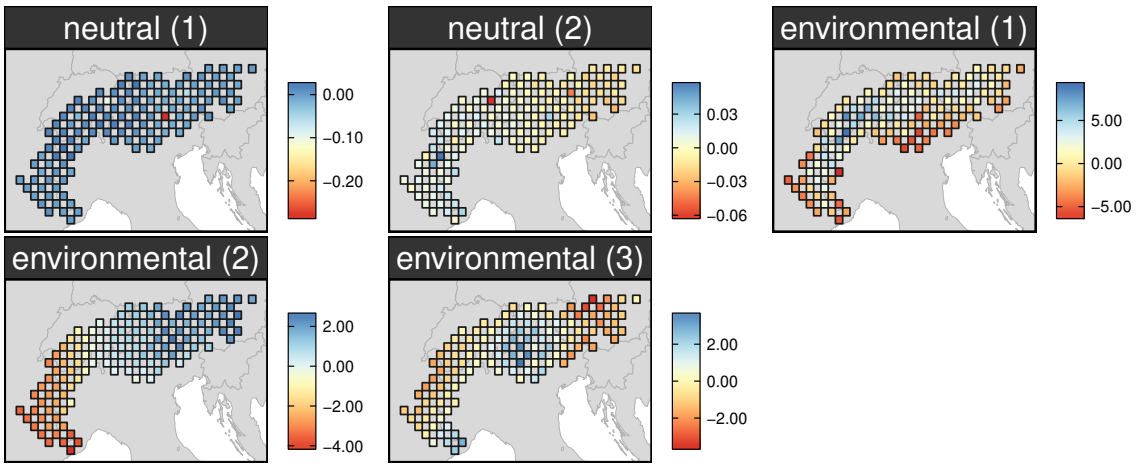
Distribution of intra-specific variation for *Cirsium spinosissimum*. See caption for first figure in Appendix J for conventions.



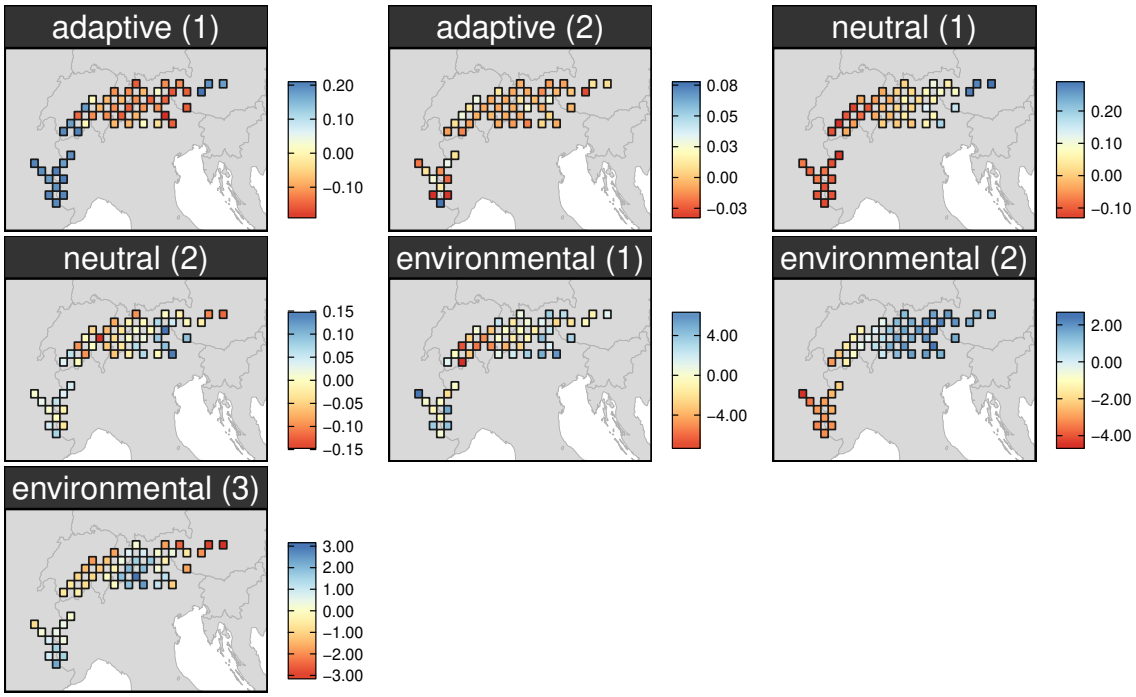
Distribution of intra-specific variation for *Dryas octopetala*. See caption for first figure in Appendix J for conventions.



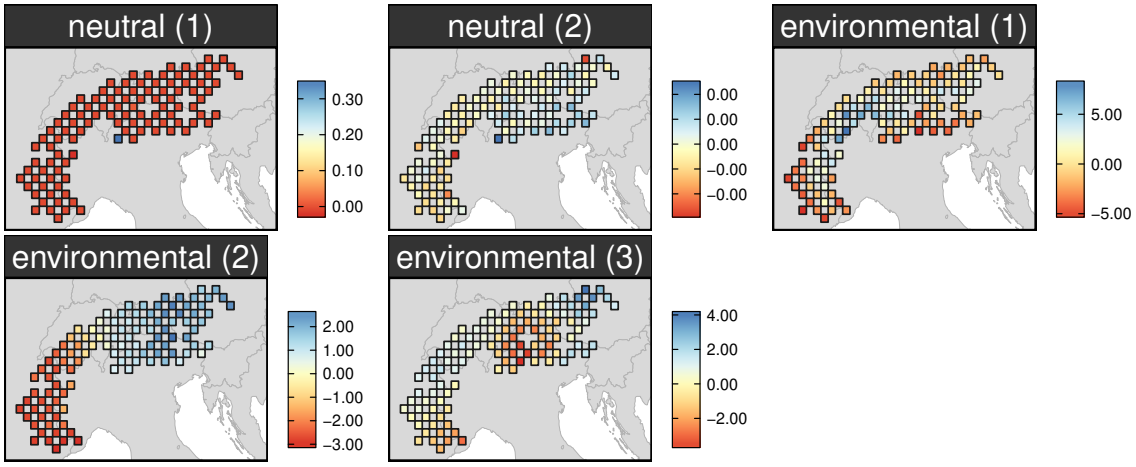
Distribution of intra-specific variation for *Gentiana nivalis*. See caption for first figure in Appendix J for conventions.



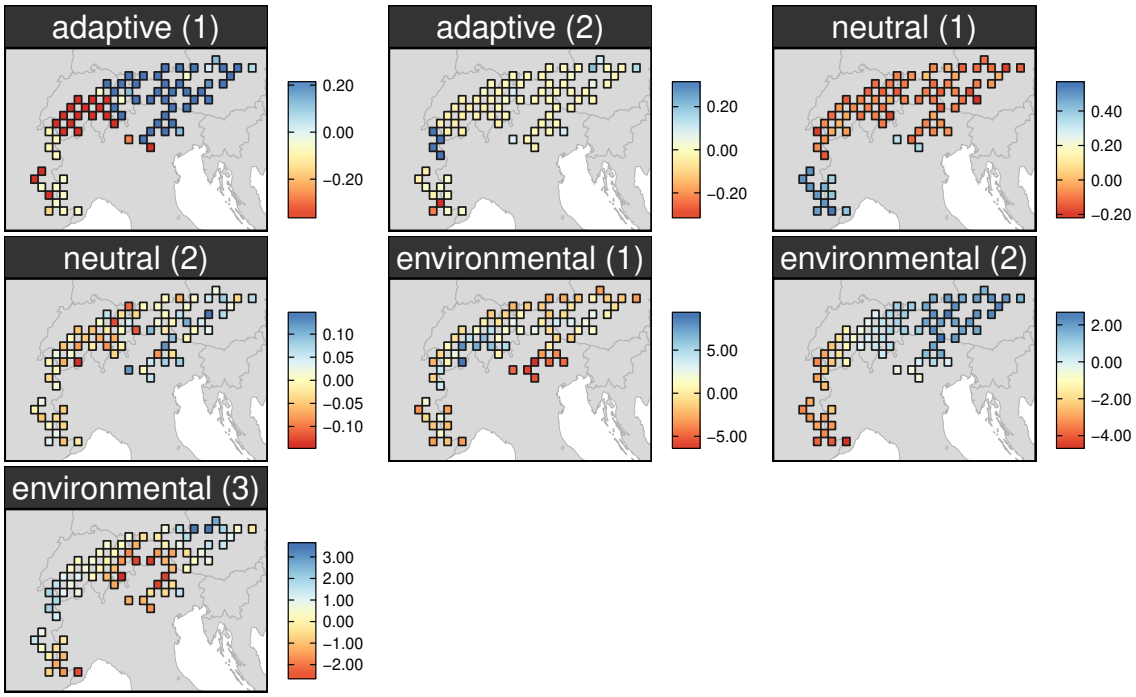
Distribution of intra-specific variation for *Geum montanum*. See caption for first figure in Appendix J for conventions.



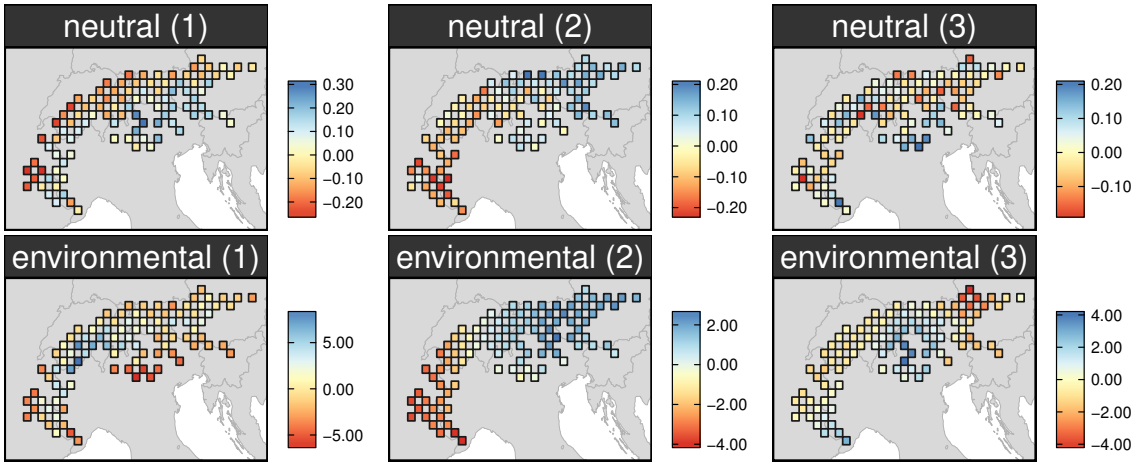
Distribution of intra-specific variation for *Geum reptans*. See caption for first figure in Appendix J for conventions.



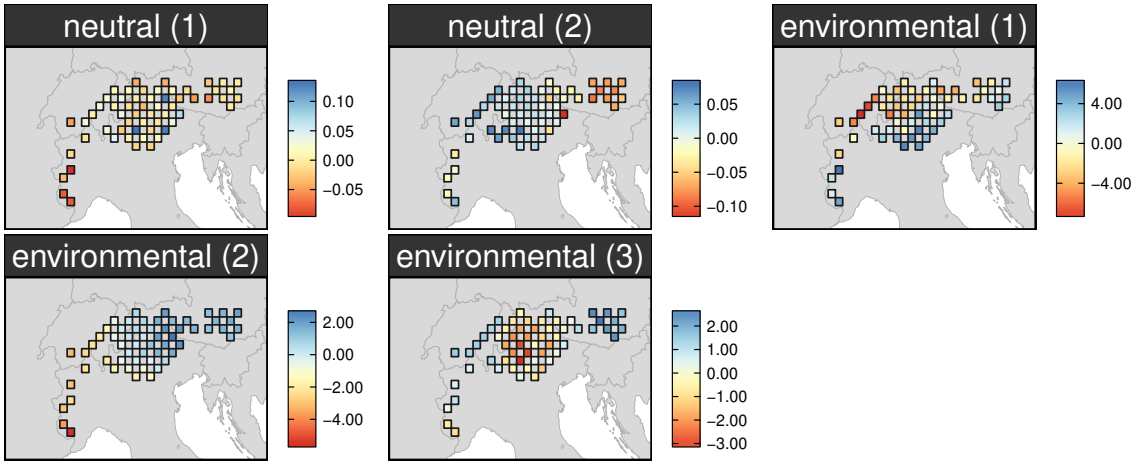
Distribution of intra-specific variation for *Gypsophila repens*. See caption for first figure in Appendix J for conventions.



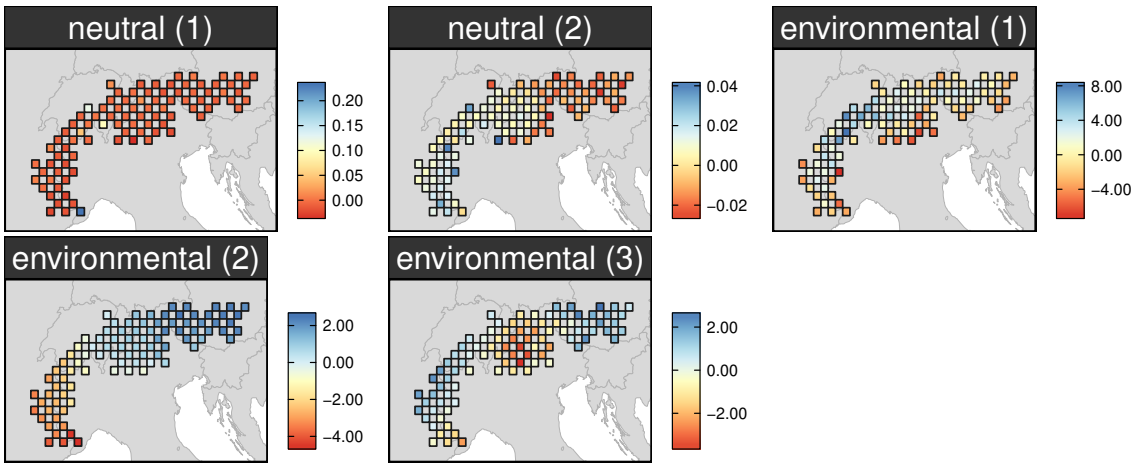
Distribution of intra-specific variation for *Hedysarum hedysaroides*. See caption for first figure in Appendix J for conventions.



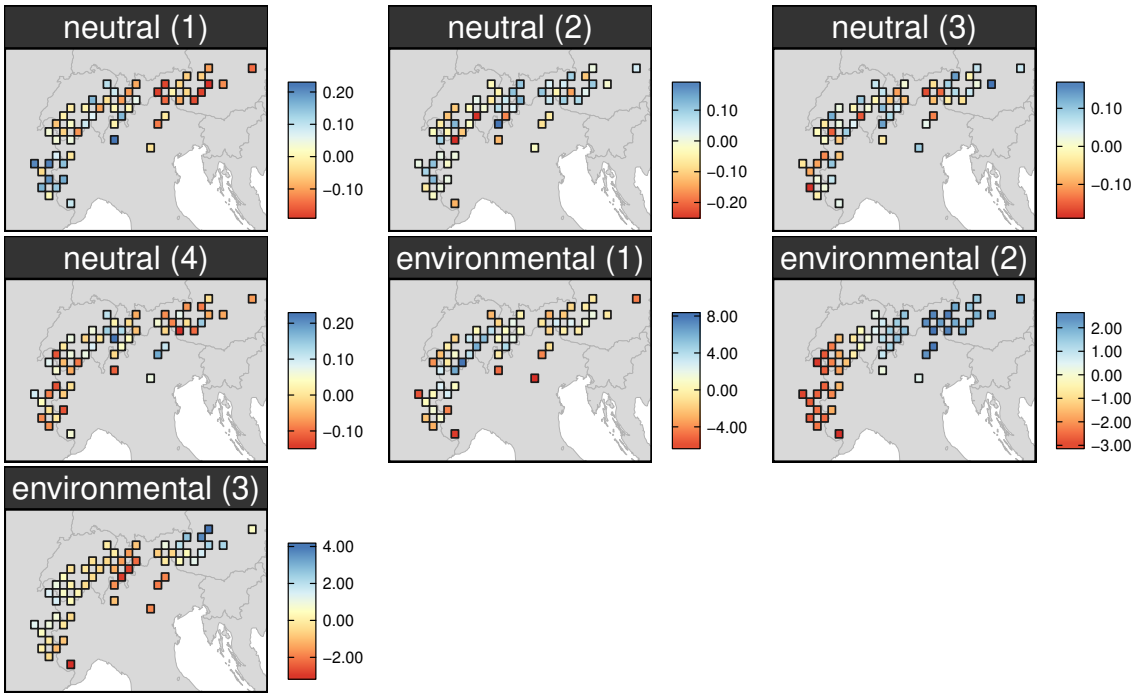
Distribution of intra-specific variation for *Hornungia alpina*. See caption for first figure in Appendix J for conventions.



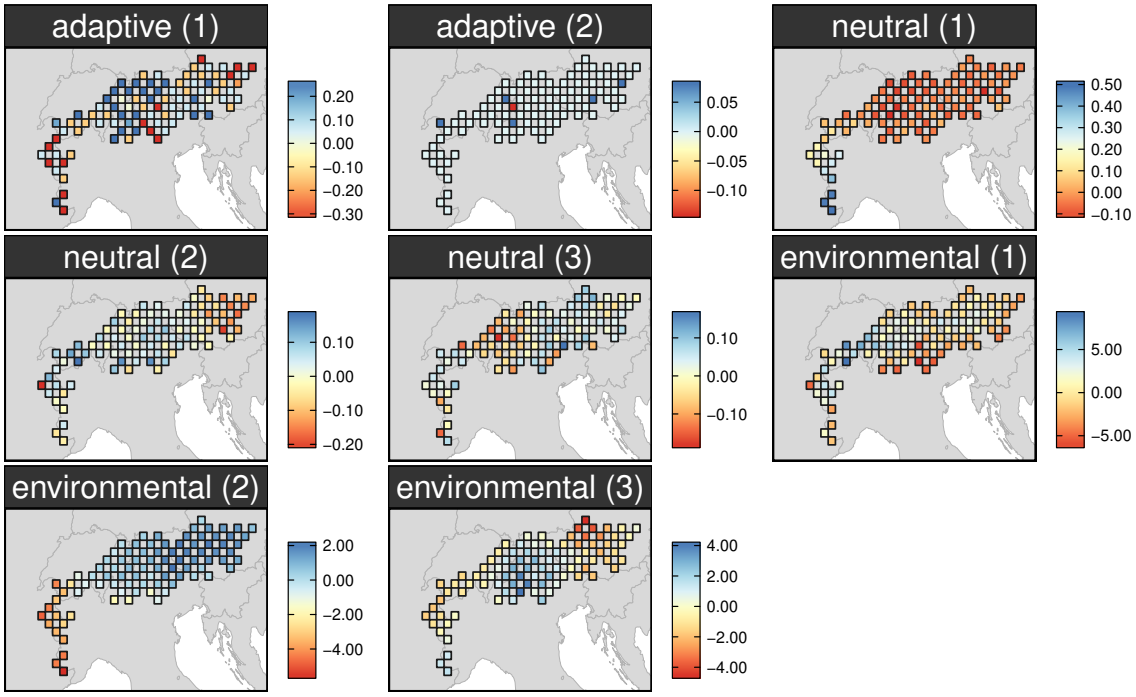
Distribution of intra-specific variation for *Hypochaeris uniflora*. See caption for first figure in Appendix J for conventions.



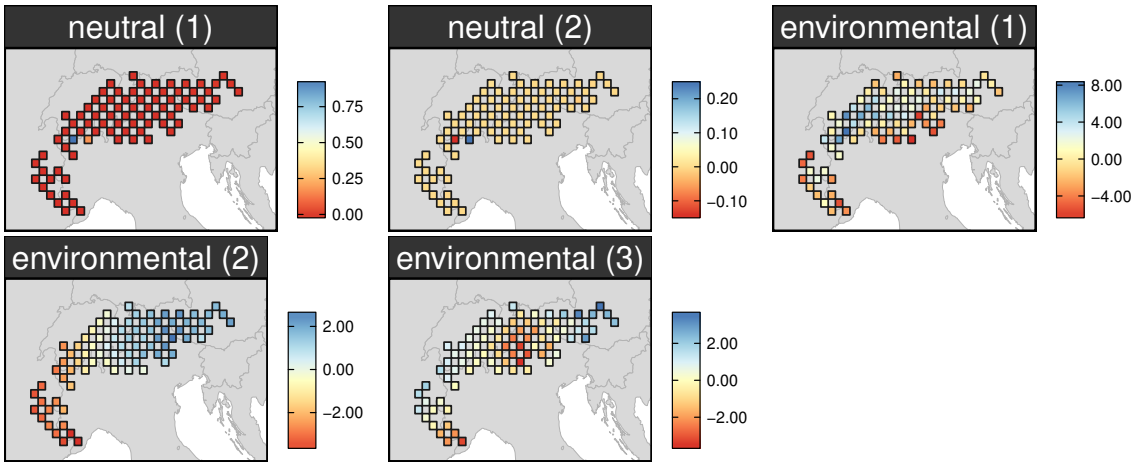
Distribution of intra-specific variation for *Juncus trifidus*. See caption for first figure in Appendix J for conventions.



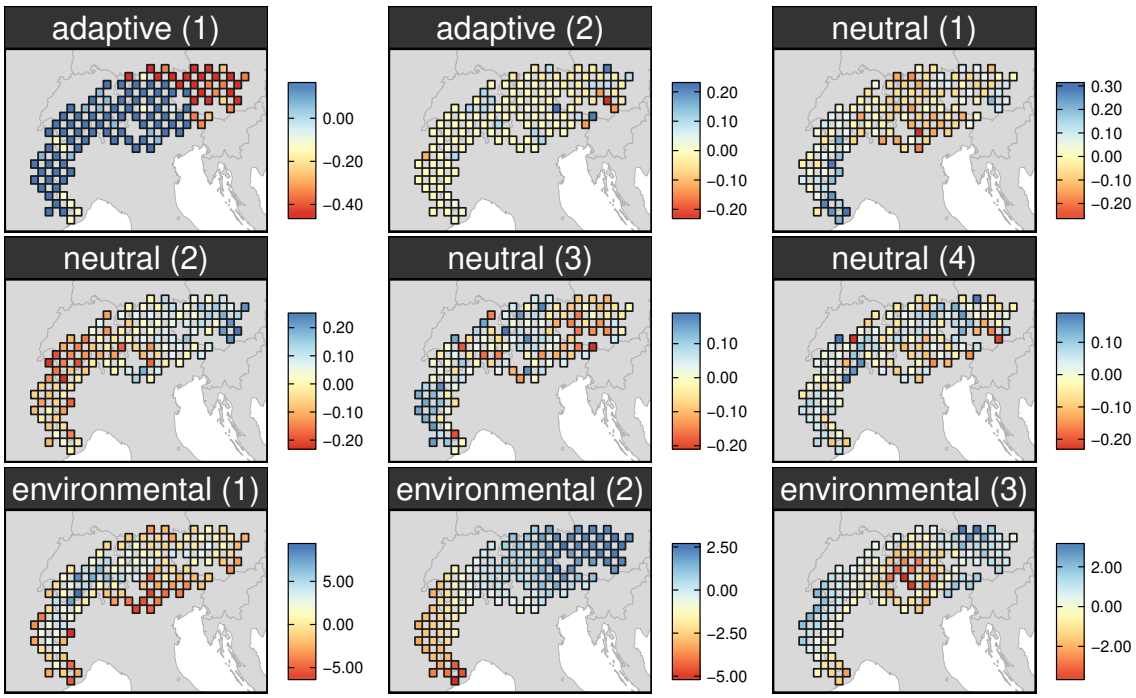
Distribution of intra-specific variation for *Ligusticum mutellinoides*. See Figure S1 caption for conventions.



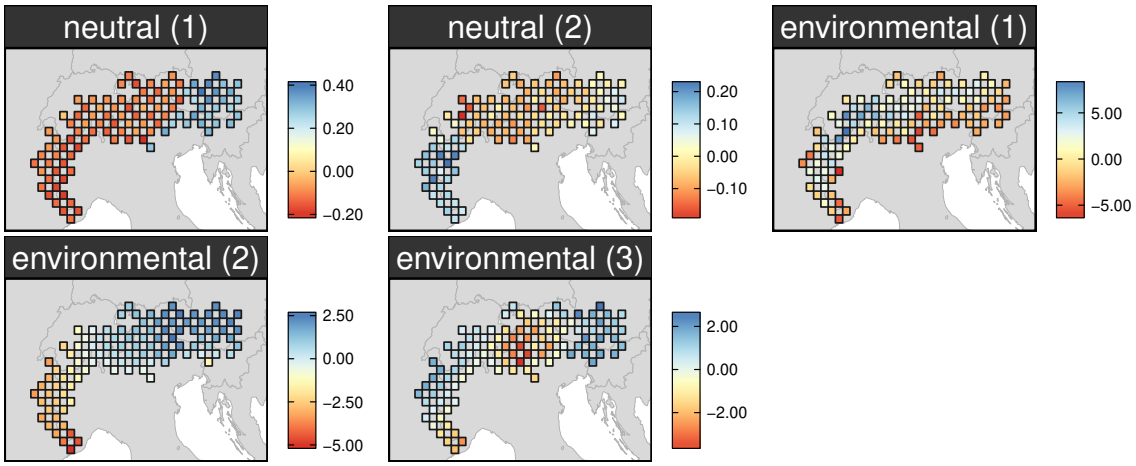
Distribution of intra-specific variation for *Loiseleuria procumbens*. See caption for first figure in Appendix J for conventions.



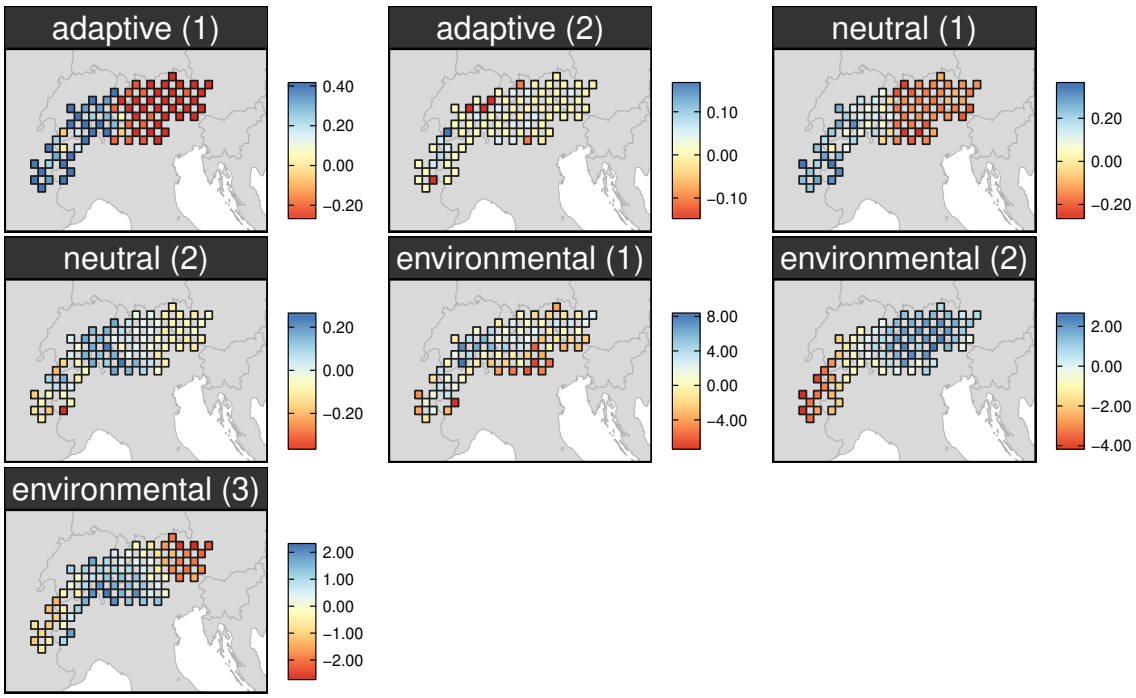
Distribution of intra-specific variation for *Luzula alpinopilosa*. See caption for first figure in Appendix J for conventions.



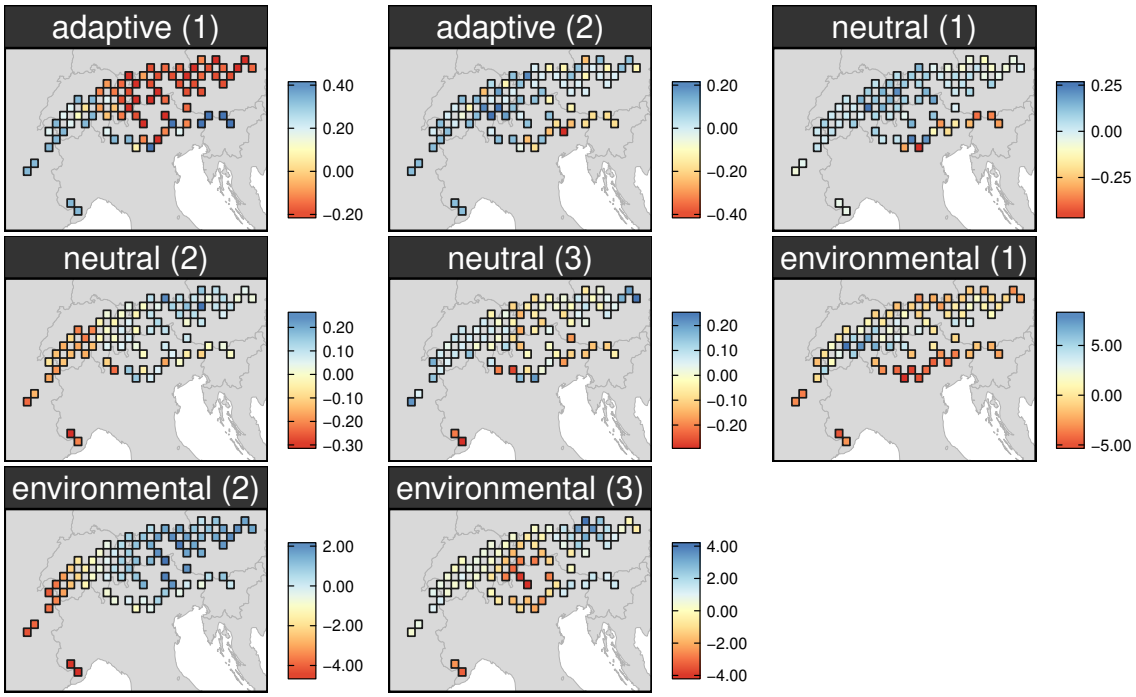
Distribution of intra-specific variation for *Peucedanum ostruthium*. See caption for first figure in Appendix J for conventions.



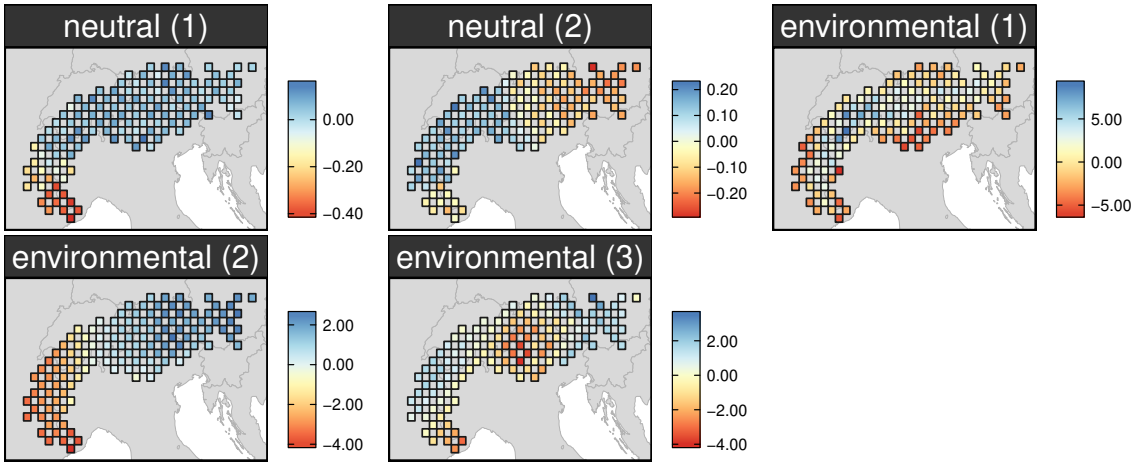
Distribution of intra-specific variation for *Phyteuma betonicifolium*. See Figure S1 caption for conventions.



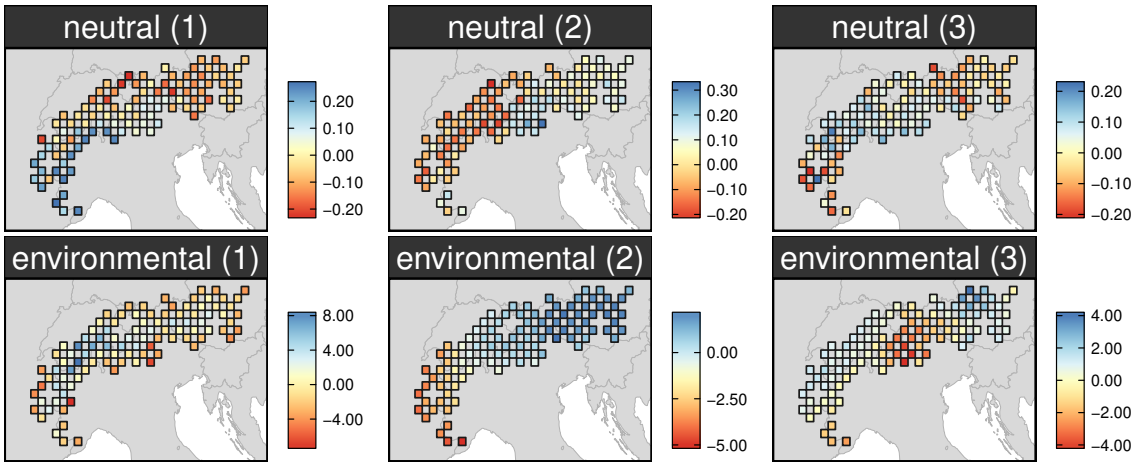
Distribution of intra-specific variation for *Phyteuma hemisphaericum*. See Figure S1 caption for conventions.



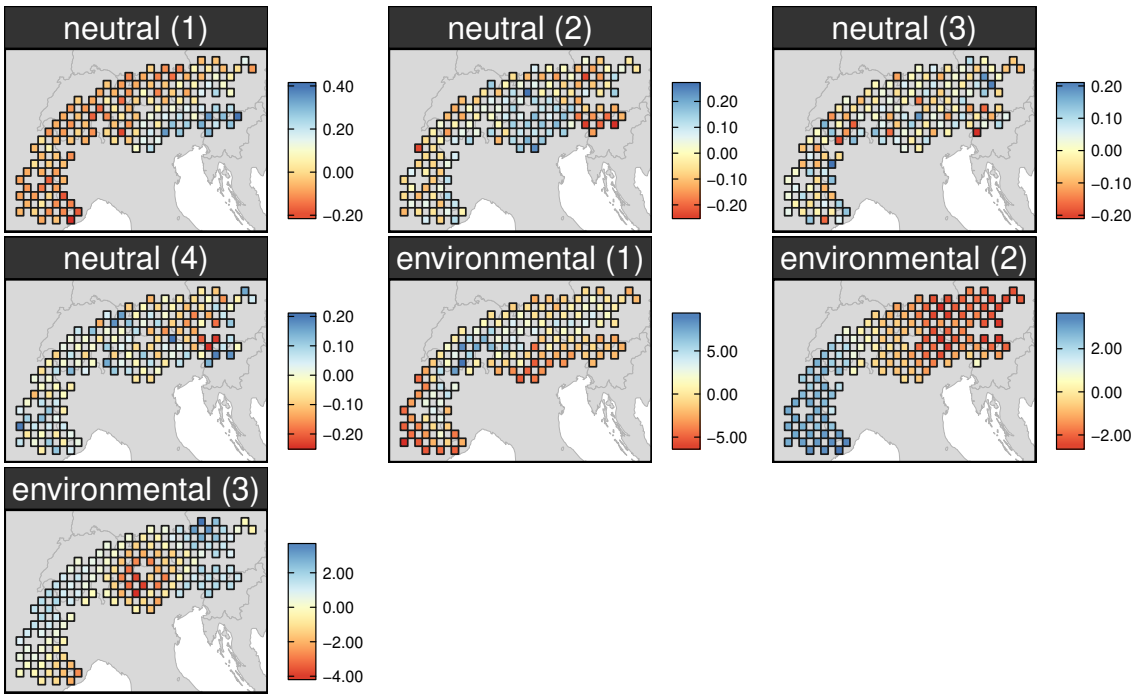
Distribution of intra-specific variation for *Ranunculus alpestris*. See caption for first figure in Appendix J for conventions.



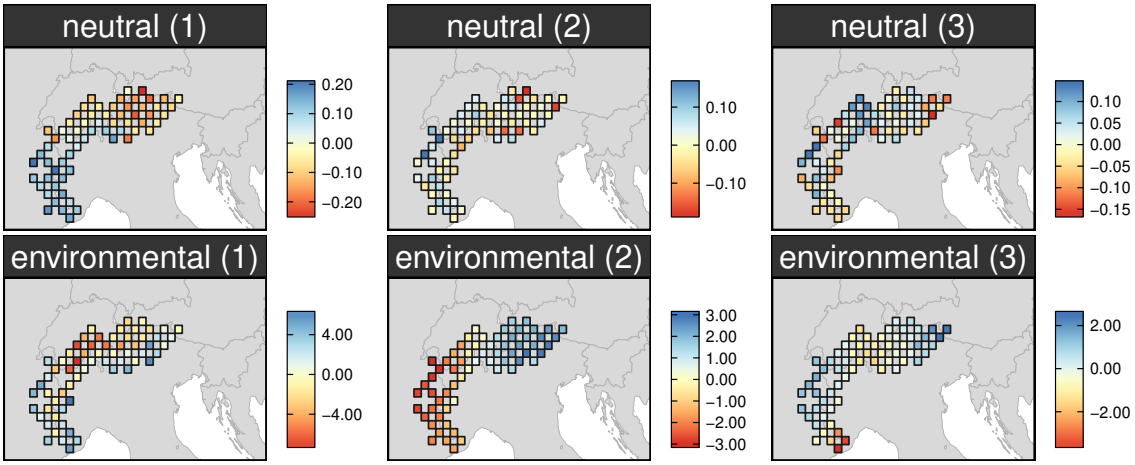
Distribution of intra-specific variation for *Rhododendron ferrugineum*. See Figure S1 caption for conventions.



Distribution of intra-specific variation for *Saxifraga stellaris*. See caption for first figure in Appendix J for conventions.



Distribution of intra-specific variation for *Sesleria caerulea*. See caption for first figure in Appendix J for conventions.



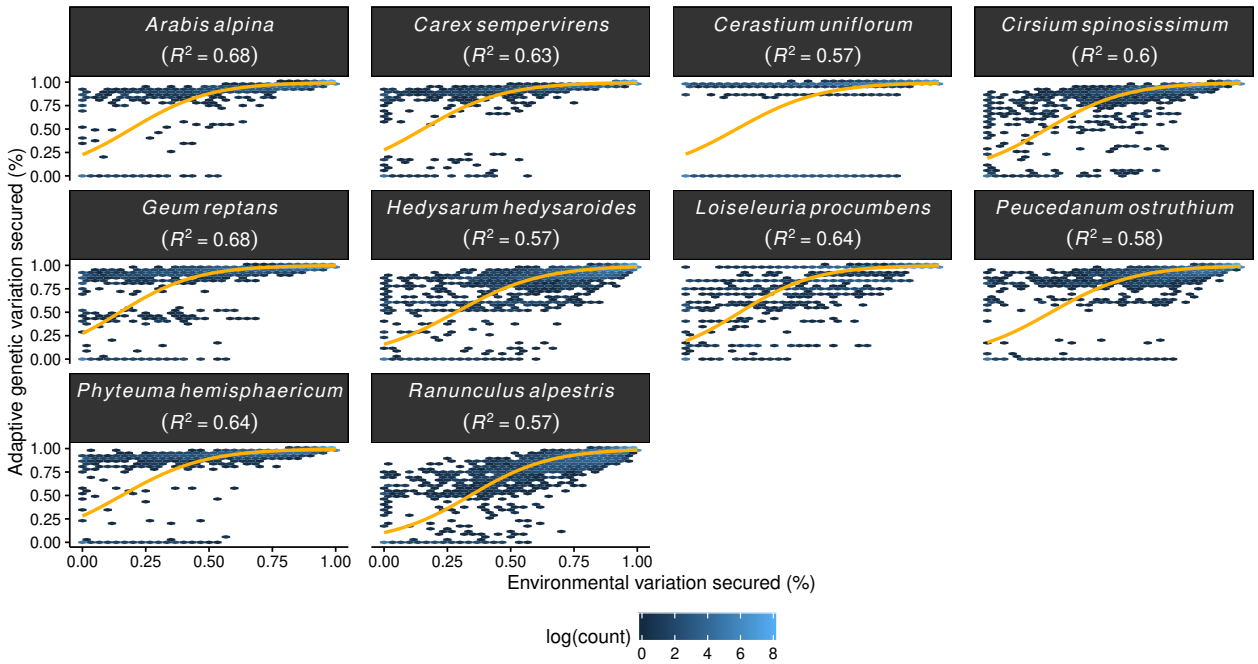
Distribution of intra-specific variation for *Trifolium alpinum*. See caption for first figure in Appendix J for conventions.

Appendix K

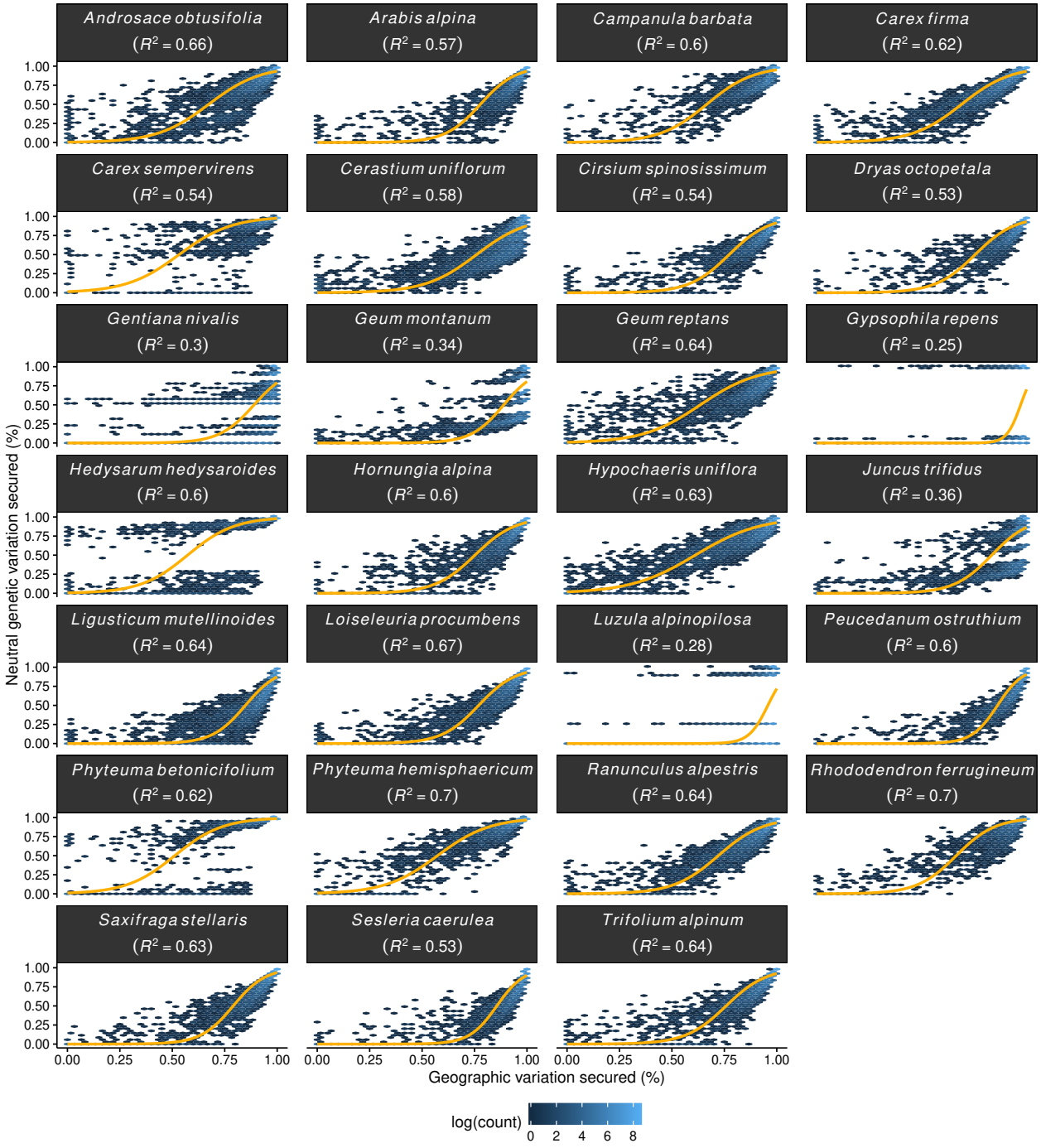
Summary of maximum likelihood population-effects (MLPE) models. Briefly, for each species associated with potentially adaptive loci, a model was fit to correlate dissimilarities among planning units where the species in terms of the adaptive genetic characteristics of individuals in them with dissimilarities in terms of the environmental conditions inside the planning units (denoted in the environmental column group). Similarly, for each species, a model was fit to correlate dissimilarities among the planning units in terms of the neutral genetic characteristics of individuals in them with dissimilarities in terms of the geographic position of the planning units (denoted in the geographic column group). Data shows the results from comparing each model to its corresponding null model using χ^2 tests (1 degree of freedom), and the R^2 (MR^2) and conditional R^2 (CR^2) for each model.

	Environmental				**Geographic**			
Species	MR^2	CR^2	χ^2	*P*	MR^2	CR^2	χ^2	*P*
Androsace obtusifolia					0.28	0.66	408.27	<0.001
Arabis alpina	0.15	0.21	989.5	<0.001	0.16	0.43	1611.53	<0.001
Campanula barbata					0.32	0.69	2880.9	<0.001
Carex firma					0.036	0.35	89.45	<0.001
Carex sempervirens	<0.001	0.011	1.67	>0.99	0.018	0.13	139.74	<0.001
Cerastium uniflorum	0.016	0.28	13.81	0.007	0.14	0.84	436.43	<0.001
Cirsium spinosissimum	0.015	0.24	68.73	<0.001	0.11	0.46	811.87	<0.001
Dryas octopetala					0.22	0.48	2040.82	<0.001
Gentiana nivalis					<0.001	0.94	3.72	>0.99
Geum montanum					0.041	0.93	2496.3	<0.001
Geum reptans	0.093	0.14	91.89	<0.001	0.42	0.75	921.34	<0.001
Gypsophila repens					<0.001	>0.99	234.03	<0.001
Hedysarum hedysaroides	0.13	0.21	291.81	<0.001	0.23	0.55	832.02	<0.001
Hornungia alpina					0.20	0.37	941.47	<0.001
Hypochaeris uniflora					0.22	0.62	422.7	<0.001
Juncus trifidus					0.061	0.88	1311.63	<0.001
Ligusticum mutellinoides					0.088	0.34	142.65	<0.001
Loiseleuria procumbens	0.001	0.12	2.94	>0.99	0.27	0.67	1643.99	<0.001
Luzula alpinopilosa					<0.001	>0.99	34.2	<0.001
Peucedanum ostruthium	0.046	0.31	266.91	<0.001	0.22	0.42	1677.68	<0.001
Phyteuma betonicifolium					0.39	0.52	2488.9	<0.001
Phyteuma hemisphaericum	0.14	0.18	308.32	<0.001	0.58	0.71	2419.26	<0.001
Ranunculus alpestris	0.090	0.22	199.11	<0.001	0.22	0.54	805.66	<0.001
Rhododendron ferrugineum					0.50	0.67	5807.85	<0.001
Saxifraga stellaris					0.28	0.48	1623.69	<0.001
Sesleria caerulea					0.053	0.44	633.32	<0.001
Trifolium alpinum					0.36	0.55	970.49	<0.001

Appendix L



The relationship between the proportion of adaptive genetic variation secured in a prioritization and the proportion of environmental variation it also secures. Points represent values associated with randomly generated prioritizations. Lines show trends fit using generalised linear models with a logit link functions. Each panel corresponds to a different species.



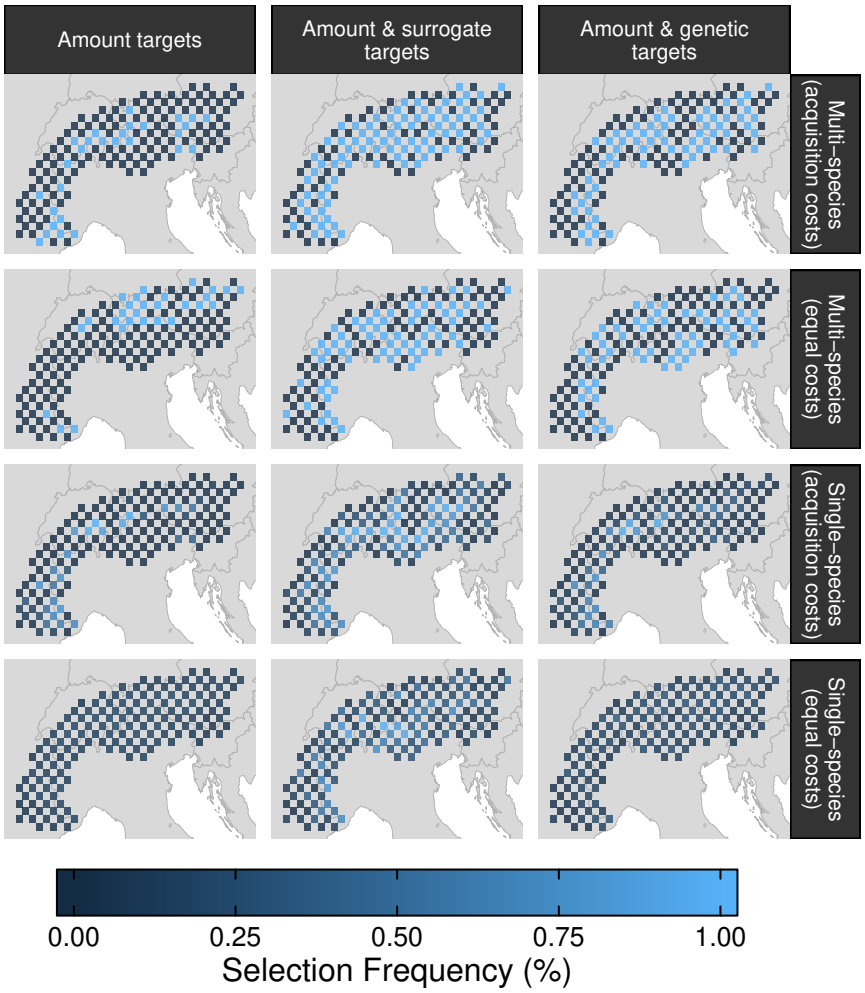
The relationship between the proportion of neutral genetic variation secured in a prioritization and the proportion of geographic variation it also secures. Conventions are detailed in the caption for the first figure in Appendix L.

Appendix M

Summary of *post-hoc* analyses on generalized linear models. Briefly, two generalized linear models were fit to correlate the proportion of adaptive genetic variation and environmental variation secured, and the proportion of neutral genetic variation and geographic variation secured in a prioritization. These models also included a predictor variable indicating the species for which the prioritisations were generated, and an interaction term. Table shows the results from one-sided *post-hoc* analyses that were conducted to determine if the slopes for each species were positive (using the alternative hypothesis that slopes were less than or equal to zero; tests used a single degree of freedom; results reported in the *Z* and *P* columns). The table also shows Cragg and Uhler's pseudo-R^2 values that were calculated to show the proportion of variation explained by the slopes for each species (R^2 column).

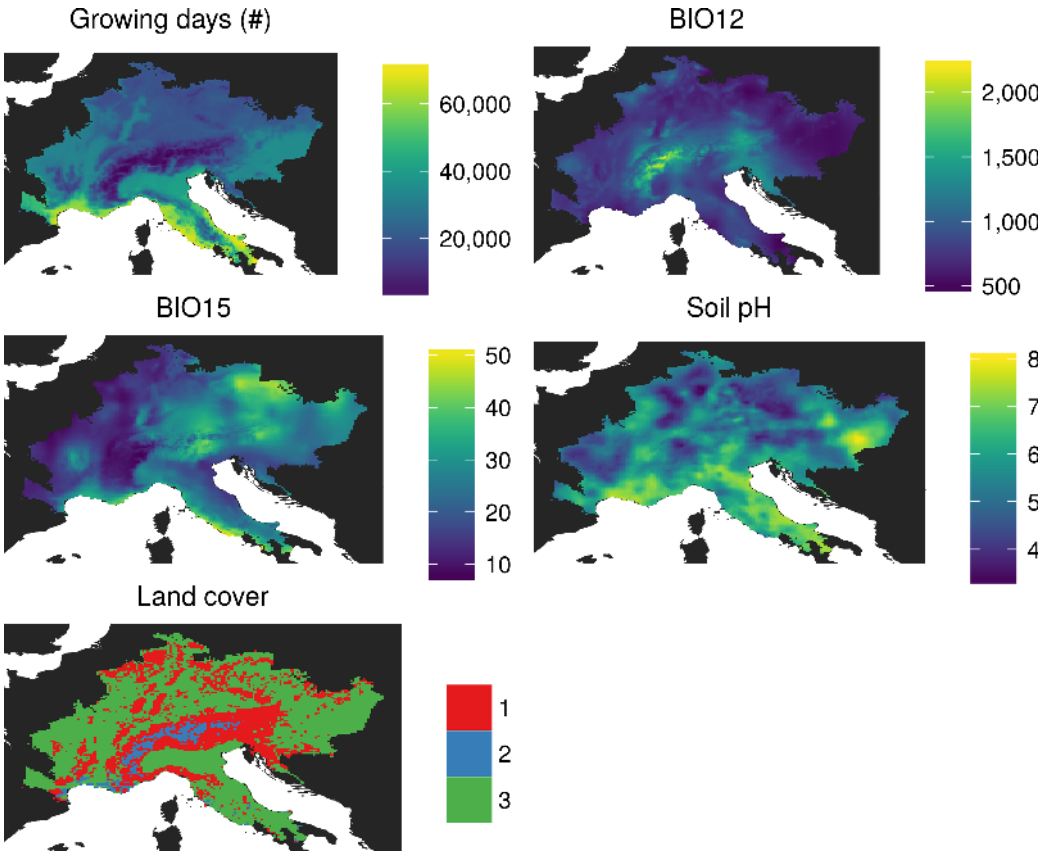
	Environmental			**Geographic**		
Species	*Z*	*P*	R^2	*Z*	*P*	R^2
Androsace obtusifolia				8.18 ± 0.24	<0.001	0.66
Arabis alpina	6.38 ± 0.2	<0.001	0.68	11.98 ± 0.48	<0.001	0.57
Campanula barbata				8.97 ± 0.33	<0.001	0.60
Carex firma				8.53 ± 0.25	<0.001	0.62
Carex sempervirens	6.23 ± 0.21	<0.001	0.63	7.98 ± 0.32	<0.001	0.54
Cerastium uniflorum	5.96 ± 0.15	<0.001	0.57	8.21 ± 0.26	<0.001	0.58
Cirsium spinosissimum	6.27 ± 0.19	<0.001	0.60	10.55 ± 0.39	<0.001	0.54
Dryas octopetala				10.63 ± 0.42	<0.001	0.53
Gentiana nivalis				12.08 ± 0.45	<0.001	0.30
Geum montanum				12.87 ± 0.55	<0.001	0.34
Geum reptans	6.6 ± 0.18	<0.001	0.68	7.3 ± 0.23	<0.001	0.64
Gypsophila repens				30.31 ± 1.08	<0.001	0.25
Hedysarum hedysaroides	5.88 ± 0.16	<0.001	0.57	8.93 ± 0.28	<0.001	0.60
Hornungia alpina				10.44 ± 0.37	<0.001	0.60
Hypochaeris uniflora				6.48 ± 0.18	<0.001	0.63
Juncus trifidus				11.15 ± 0.45	<0.001	0.36
Ligusticum mutellinoides				12.75 ± 0.4	<0.001	0.64
Loiseleuria procumbens	6.65 ± 0.19	<0.001	0.64	11.43 ± 0.37	<0.001	0.67
Luzula alpinopilosa				22.16 ± 0.76	<0.001	0.28
Peucedanum ostruthium	6.08 ± 0.18	<0.001	0.58	16.43 ± 0.59	<0.001	0.60
Phyteuma betonicifolium				8.76 ± 0.32	<0.001	0.62
Phyteuma hemisphaericum	6.06 ± 0.18	<0.001	0.64	8.01 ± 0.25	<0.001	0.70
Ranunculus alpestris	6.33 ± 0.17	<0.001	0.57	9.41 ± 0.28	<0.001	0.64
Rhododendron ferrugineum				10.78 ± 0.4	<0.001	0.70
Saxifraga stellaris				12.96 ± 0.44	<0.001	0.63
Sesleria caerulea				15.32 ± 0.59	<0.001	0.53
Trifolium alpinum				10.39 ± 0.35	<0.001	0.64

Appendix N



Planning unit selection frequencies. Squares represent planning units, and their colors indicate their selection frequency. Each column of panels shows solutions generated using different targets. Each row of panels shows solutions generated under different scenarios. Note that the selection frequencies for the single-species prioritizations are based on their frequency among prioritizations generated for each species. The selection frequencies for the multi-species prioritizations are each based on a single optimal prioritization for all species and so are binary.

Appendix O



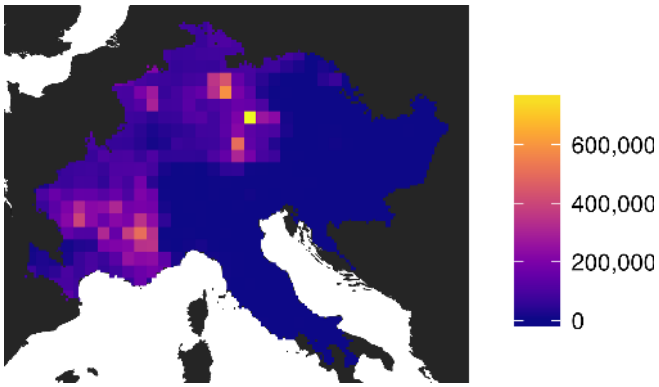
Predictor data used to train environmental niche models and build resistance maps. Panels show (a) number of days in the year with the temperature above 5 °C, (b) annual precipitation (`BIO12`), (c) precipitation seasonality (`BIO15`), (d) soil pH and (e) land cover. Land cover classes represent forest (1), other land-uses (2), and human-modified landscape (3).

Appendix P

Summary of models used to generate habitat suitability maps for each species. Data show the total number of records, subset of cleaned records used to fit the models, model performance (mean and variance in AUC: $\overline{\text{AUC}}$, AUC_{var}), features used to construct the model (Linear, Quadratic, Product, Threshold, Hinge), number of parameters in the models, and regularization penalty (λ). Models were evaluated using ten-fold cross-validation. The mean and variance associated with the area under the receiver operating characteristic curve (AUC) statistics among each fold are presented.

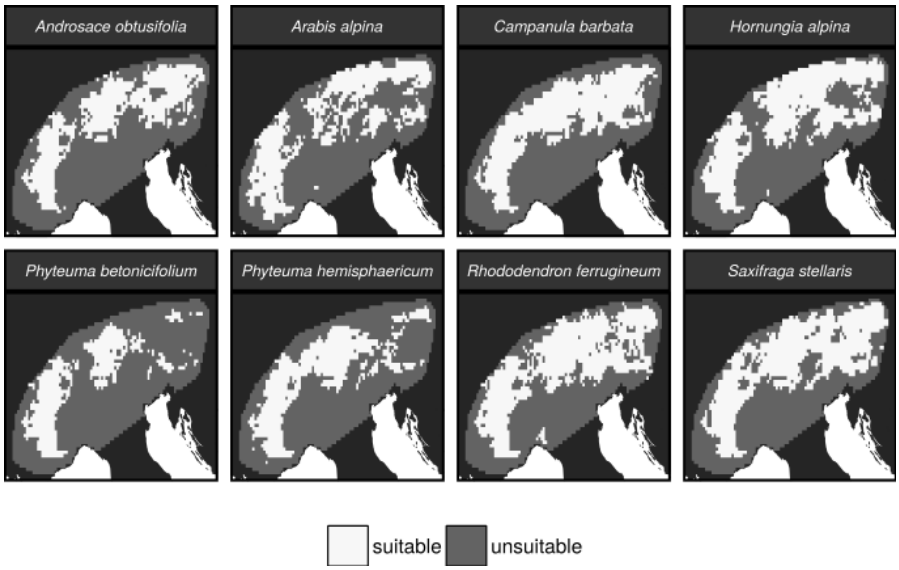
Species	Total records	Subset records	$\overline{\text{AUC}}$	AUC_{var}	Features	Parameters	λ
Androsace obtusifolia	1,013	140	0.974	0.002	LQHP	139	0.1
Arabis alpina	7,508	1,277	0.833	0.004	LQHP	221	0.1
Campanula barbata	2,471	418	0.955	0.002	LQH	119	0.5
Hornungia alpina	4,655	504	0.960	0.000	LQHP	167	0.1
Phyteuma betonicifolium	3,091	276	0.964	0.001	H	106	0.5
Phyteuma hemisphaericum	6,080	257	0.944	0.003	H	94	0.5
Rhododendron ferrugineum	17,292	659	0.946	0.002	LQHP	196	0.1
Saxifraga stellaris	4,738	390	0.923	0.003	LQHP	187	0.1

Appendix Q



Map of sampling bias. Colors denote the density of plant records in the Global Biodiversity Information Facility (GBIF) database collected using human observation and without any reported spatial issues.

Appendix R



Distribution suitable habitat for each species. Squares correspond to planning units.

Appendix S

Summary of maximum likelihood population effects models (MLPE). These models were fit to the genetic differences between sample locations using the accumulation cost associated with least cost paths between the sample locations. The least cost paths were calculated using the near-optimal resistance maps. Data shows the models' fixed effects (intercept and slope), log-likelihood (*LL*), and $R^2\beta$ statistics. Uncertainty is expressed as the 5th and 95th percentiles of $R^2\beta$ statistics obtained from the bootstrap analysis.

Species	*LL*	Intercept	Slope	$R^2\beta$
Androsace obtusifolia	1466.656	0.303	0.124	0.804 (0.762 - 0.828)
Arabis alpina	12351.245	0.253	0.079	0.634 (0.562 - 0.677)
Campanula barbata	5265.647	0.413	0.171	0.564 (0.521 - 0.63)
Hornungia alpina	6482.017	0.298	0.079	0.478 (0.433 - 0.511)
Phyteuma betonicifolium	4387.397	0.394	0.233	0.711 (0.687 - 0.733)
Phyteuma hemisphaericum	2974.141	0.491	0.290	0.728 (0.708 - 0.744)
Rhododendron ferrugineum	10156.353	0.319	0.152	0.712 (0.685 - 0.725)
Saxifraga stellaris	5901.533	0.389	0.158	0.614 (0.588 - 0.634)

Appendix T

We used a novel formulation of the reserve selection to generate prioritizations that represented species in contiguous areas that offered little impediment to gene flow. This same formulation was also used to represent species in contiguous areas of suitable habitat. Let I denote the set of planning units (indexed by i). Each unit i is associated with an acquisition cost C_i. Let F denote the set of conservation features (e.g. species or ecosystem types; indexed by f). Each planning unit is associated with an amount of suitable habitat for each feature H_{fi}. Each feature is associated with a representation target T_f that denotes the total amount of connected suitable habitat that needs to be conserved.

We adopted the connectivity constraints outlined in Önal & Briers (2006) to represent species in contiguous areas. As described in the main text, we generated symmetric graphs to describe rook-style dispersal between adjacent planning units for each feature. These graphs determine which adjacent planning units are capable of potentially facilitating gene flow. In our analysis, we generated these graphs using the species' conductance maps and the species' habitat suitability maps. To avoid circularity issues (see below for more information; discussed at length in Önal & Briers, 2006), we extracted asymmetric sub-graphs $G_f = (I, U_f)$ for each f in F from the symmetric graphs using the planning units I as nodes and U_f to represents edges. These graphs G_f only contained edges from planning units to their eastern and southern neighbors. For convenience, we used A_{fi} to denote the set of planning units connected to planning unit i for feature f in graph G_f. Additionally, we use symbols I and F to also denote the cardinality of these sets.

The control variables in the problem are X_i and T_f.

$$X_i = \begin{cases} 1, & \text{if planning unit } i \text{ is purchased} \\ 0, & \text{otherwise} \end{cases}$$

$T_f =$ target amount of habitat to be conserved for feature f

The decision variables in the problem are Y_{fi} and Z_{fij}.

$$Y_{fi} = \begin{cases} 1, & \text{if the purchased planning unit } i \text{ is allocated to conserve feature } f \\ 0, & \text{otherwise} \end{cases}$$

$$Z_{fij} = \begin{cases} 1, & \text{if connection for feature } f \text{ between the planning units } i \\ & \text{and } j \text{ is being used to form a contiguous reserve} \\ 0, & \text{otherwise} \end{cases}$$

The problem is formulated as an integer programming problem.

$$\text{min.} \quad \sum_{i=0}^{I} X_i \times C_i \qquad \text{eqn 1a}$$

$$\text{st.} \quad Y_{fi} \leq X_i \qquad \forall f \in F, \; i \in I \qquad \text{eqn 1b}$$

$$\sum_{i=0}^{I} Y_{fi} \times H_{fi} \geq T_f \qquad \forall f \in F \qquad \text{eqn 1c}$$

$$Z_{fij} \leq Y_{fi} \qquad \forall f \in F, \; i \in I, \; j \in A_{fi}, \qquad \text{eqn 1d}$$

$$\sum_{i \in A_{fj}} Z_{fij} \leq Y_{fj} \qquad \forall f \in F, \; j \in I \qquad \text{eqn 1e}$$

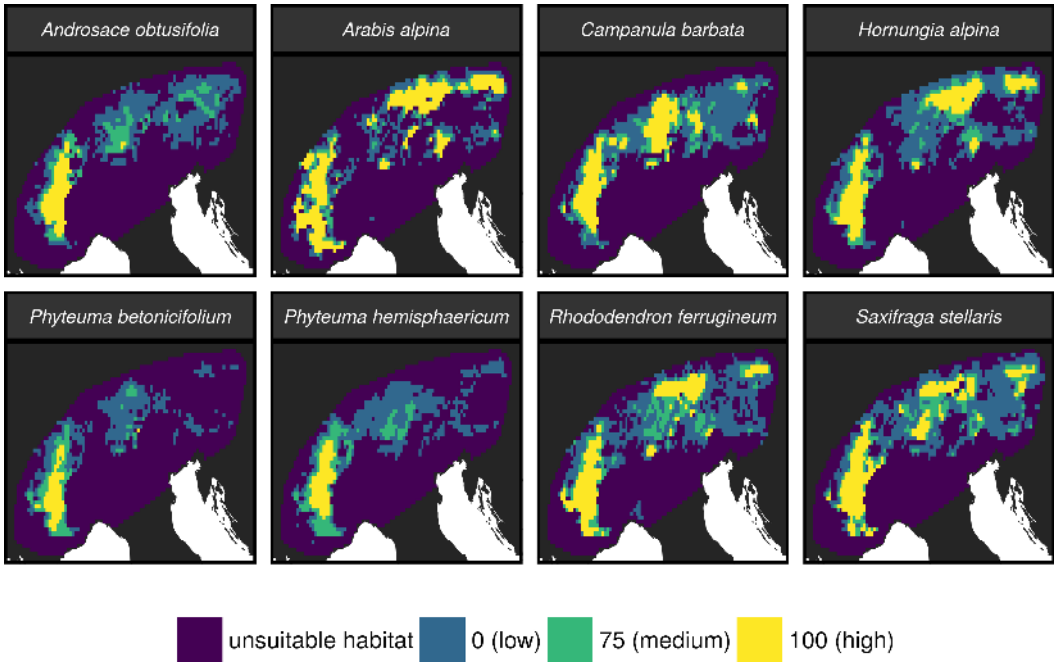
$$\sum_{i=0}^{I} \sum_{j \in A_{fi}} Z_{fij} - \sum_{i=0}^{I} Y_{fi} = -1 \qquad \forall f \in F \qquad \text{eqn 1f}$$

$$X_i, Y_{fi}, Z_{fij} \in (0, 1) \qquad \forall f \in F, \; i \in I, \; j \in A_{fi} \qquad \text{eqn 1g}$$

The objective is to minimize the cost of the reserve network (eqn 1a). Constraints (eqn 1b) ensure that planning unit i cannot be allocated to conserve species f unless the unit has been purchased. Note that the same planning unit can be allocated to multiple species. Constraints (eqn 1c) ensure that the total amount of habitat allocated for feature $f \in F$ is not less than the conservation target T_f. Constraints (eqns 1d–1f) guarantee that purchased planning units form a contiguous reserve. Specifically, constraints (eqn 1d) ensure that the connection from planning unit i to unit j for feature f can only be active if planning unit i is allocated for feature f. Constraints (eqn 1e) ensure that planning unit j has only one incoming connection for each feature $f \in F$. Constraints (eqn 1f) ensure that the number of active connections for each feature $f \in F$ is equal to the number if planning units allocated for the feature. Since asymmetric directed graphs were used to represent possible

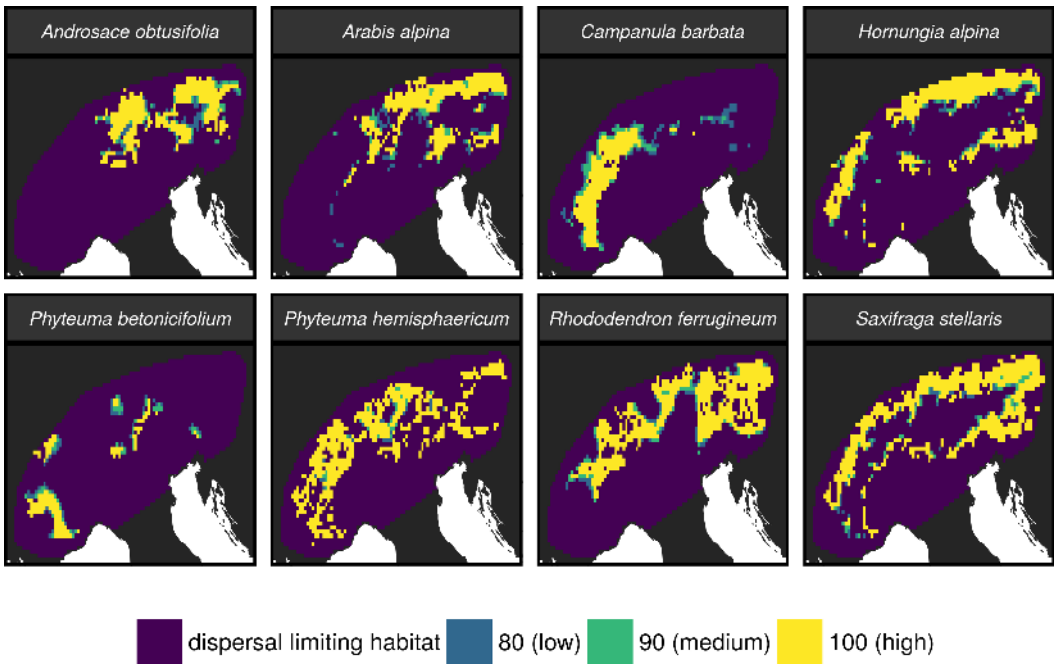
connections between adjacent planning units, no additional constraints are required to avoid circularity (e.g. forbid planning units *A*, *B*, *C*, and *D* forming the circuit $A \rightarrow B \rightarrow C \rightarrow D \rightarrow A$ where *A* is the start and end of the circuit; discussed in Önal & Briers, 2006). Finally, constraints (eqn 1g) specify that the X, Y, and Z variables are binary.

Appendix U



Planning units with habitat suitability values above specified thresholds used for generating prioritizations.

Appendix V



Planning units with conductance values above specified thresholds used for generating prioritizations.

Appendix W

Spearman's correlation tests between genetically derived resistance maps and habitat suitability maps for each species.

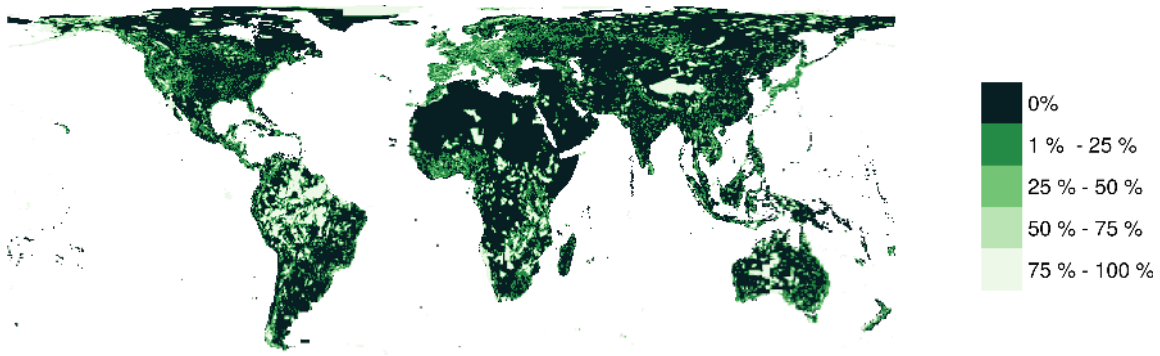
Species	Estimate	Statistic	*P*
Androsace obtusifolia	-0.474	4566789764	< 0.001
Arabis alpina	-0.505	4662169772	< 0.001
Campanula barbata	-0.775	5500254458	< 0.001
Hornungia alpina	-0.054	3266623294	0.003
Phyteuma betonicifolium	0.138	2671830130	> 0.99
Phyteuma hemisphaericum	-0.461	4527484523.231	< 0.001
Rhododendron ferrugineum	-0.571	4868321396	< 0.001
Saxifraga stellaris	-0.081	3349122336	< 0.001

Appendix X

Average percent contributions for predictor layers to the overall resistance map for each species. The layer with the highest contribution for each species is shown in bold.

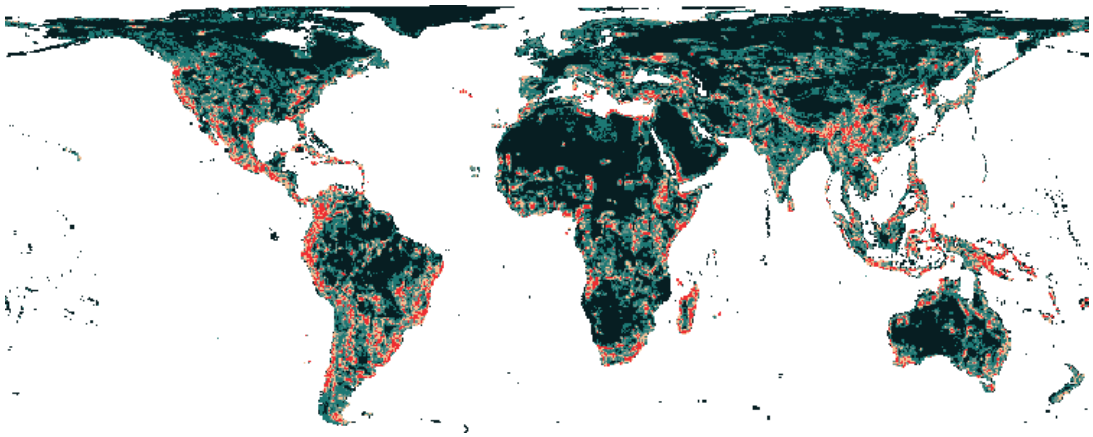
Species	Annual days above 5 °C	Annual precipitation	Precipitation seasonality	Soil pH	Land cover
Androsace obtusifolia	1.817	1.817	**61.673**	1.817	32.876
Arabis alpina	1.161	1.161	12.485	**55.872**	29.321
Campanula barbata	30.359	3.322	**36.097**	3.322	26.9
Hornungia alpina	**70.31**	2.627	0.76	22.779	3.523
Phyteuma betonicifolium	0.805	0.805	0.805	**96.649**	0.937
Phyteuma hemisphaericum	1.198	1.198	1.198	**54.761**	41.644
Rhododendron ferrugineum	0.753	0.753	0.753	**68.346**	29.395
Saxifraga stellaris	**63.907**	1.084	3.198	28.113	3.699

Appendix Y



Global distribution of protected areas. To aid visual interpretation, data show the proportion of areas within 50 km^2 cells covered by protected areas.

Appendix Z



Prioritization for expanding protected areas to cover a target amount of habitat for each species. This prioritization was generated using conventional approaches that do not account for environmental variation across species' distributions. To aid visual interpretation, data show the proportion of areas within 50 km^2 cells selected for protection. Colors correspond to 0 % (■), 1–25 % (■), 25–50 % (■), 50–75 % (■), and 75–100 % (■).