Preadaptation, hybridisation, and breeding system shape the invasion of three *Rumex* species in New Zealand

---

A thesis
submitted in partial fulfilment
of the requirements for the Degree of
Doctor of Philosophy

at
Lincoln University
by
Sandra E. Savinen

Lincoln University
2022

Abstract of a thesis submitted in partial fulfilment of the requirements for the Degree of Doctor of Philosophy.

Preadaptation, hybridisation, and breeding system shape the invasion of three *Rumex* species in New Zealand

by

Sandra E. Savinen

Research has previously shown that *Rumex* (Polygonaceae) species introduced to New Zealand have multiple potential drivers behind their success, such as phenotypic plasticity, enemy release, and niche shift. However, it is not known whether these changes were caused by post-introduction evolution, or what other drivers could explain the success of these agricultural weeds. I combined demo-genetic traits and processes to assess how hybridisation and introgression, genetic differentiation, and breeding system contribute to the invasiveness of three introduced *Rumex* species. I compared plants from the species' native (Europe, mainly the UK) and introduced (New Zealand) range and assessed whether the success is more likely due to prior adaptation or post-introduction evolution.

Ploidy is associated with increased invasiveness, and if a species has multiple geo-cytotypes, higher ploidies are often found within the introduced range. Similarly, self-compatibility can help introduced populations to counter mate limitation and mixed mating can introduce new alleles to populations. I found no differences in genome sizes or chromosome numbers between plants from the two ranges using flow cytometry and manual chromosome counts. In addition, a comparison between bagged and unbagged *Rumex conglomeratus* plants showed no consequences from selfing, indicating mixed mating strategies. Surprisingly, the overall seed viability was lower for provenances from the introduced range compared to the native range. Hybridisation and introgression can increase genetic variation and help with adaptation to new environments. In a field survey, hybrid plants were found in New Zealand. However, the majority were likely first-generation hybrids, making introgression an unlikely driver behind the invasiveness. In addition, the parent species co-occurrence was lower in New Zealand compared to the UK. Lastly, genetic differentiation can indicate the origin of the introduction, as well as how likely a post-introduction evolution is a driver behind the invasiveness. A minimal differentiation was revealed by genotyping-by-sequencing both within but also between the native and introduced ranges. The population genetic analyses suggest that the UK is

a likely origin for these species but admixture from elsewhere was also found. This would have likely helped the introduced populations to maintain comparable level of genetic variation to the native populations.

As limited differences were found between the native and introduced populations, the investigated traits and processes are unlikely to explain the invasiveness in New Zealand. Rather, the success of these species is likely caused by prior adaptation. In addition, as these species are primarily agricultural weeds within both provenances, anthropogenically induced adaptation to invade is likely the main driver behind the success of these species. This method of adaptation, likely coupled with jack-of-all-trades genotypes, have allowed the species to thrive in manmade habitats, all around the globe. Thus, similar weeds within these habitats need to be carefully monitored to prevent further invasions in the future.

Keywords: anthropogenically induced adaptation to invade, genetic differentiation, genetic variation, genotyping-by-sequencing, hybridisation, introduced species, introgression, invasive species, plant invasions, pollen:ovule ratio, population genetics, preadaptation, *Rumex*, selfing, self-pollination, weed.

Acknowledgements

In 2016, I was travelling around New Zealand for a few weeks and completely fell in love with the country. I decided then that I would like to pursue my career as a scientist here and come back for a PhD. With huge amounts of luck and some persistence, I managed to do that, and now the PhD journey is nearing its end. There are so many amazing and wonderful people, who have helped me make this thesis a reality, and I hope that I have managed to name at least most of you.

My biggest thanks will go to my wonderful, amazing supervisor Phil Hulme, who granted me the wish to do a PhD in New Zealand and thus made my dreams come true. I am eternally grateful for Phil's guidance and wisdom during my PhD, he truly is a great supervisor, and it would be difficult to find a better one! I would also like to thank my associate supervisor, Chris Winefield for sharing his knowledge and time with me. He said in one of our early meetings to always shoot for the stars, so that even if that goal was out of reach, I would still end up landing as high as I could. I have cherished his words on this crazy PhD journey and will continue trying to always aim high.

I would also like to thank my collaborators in AgResearch: especially Jeanne Jacobs, Ken Dodds, and Rudiger Brauning, and in Manaaki Whenua – Landcare Research: Gary Houliston, Caroline Mitchell, Murray Dawson, and Ines Schonberger. Thank you all so much for sharing your expertise and your time. You – quite literally – made my thesis what it is, and helped me to achieve my goals, and for that I am extremely grateful. Thank you to the staff and support team at Lincoln Uni, especially Brent Richards and the greenhouse team for keeping my plants alive and letting them take over the greenhouse for so many years, as well as Roger Cresswell for all the help with the quarantining of the imported *Rumex* seeds.

Thank you to Bio-Protection Research Centre, for funding my research, and giving me a sense of family, my home away from home. I am so thankful to Andrew Holyoake, Sandy Wilson, Claire Tee, and Michelle Boyle for all the hard work, support, and advice on admin related (or really any) issues. Thank you to Jo Narciso, for all the love and baked goods, thank you for always asking me how I was doing – in Finnish!! Huge thanks to the BPRC Tea Room Puzzle Team, so many puzzles were built, both in the real tearoom and online during the pandemic lockdowns. Thank you to Hoss Alizadeh, for your kind words and help with freeze-drying my samples. Thank you to Brian Kwan, for providing support for the lab work – your office is a treasure trove! Thank you to everyone else in BPRC, you will forever be my family.

Huge thanks to all the Weed Wing people over the years, you are all amazing. Dr Rowan Sprague, Dr Thomas Carlin, Dr Andrei Costan, thank you for making me feel so welcome not only when I joined the office but also welcoming me into your lives. Thank you to Jennifer Bufford, Sarah Wyse, and Will Godsoe for sharing your expertise, help, support, and kind words. Thank you to Maddie Marshall, Rob Gibson II, Jesse Rubenstein, and

Nikki Wells for all the fun and all the support. If BPRC is my family, the Weed Wing are my siblings – including the bitter rivalry of who could get the most smiley faces from Phil in his emails. Lastly, thank you to Dr Roseanna Gamlen-Greene, the honorary member of Weed Wing, the passionate and kind scientist. It was so great to share an office with Roseanna and get to talk about genetics and to struggle over the methodology together.

I could not have done all this without the love and support from my friends – a lot of whom are already listed above. In addition, I am forever indebted to Léonie Mollet, my uni(corn) sis and an amazing Swiss powerhouse who I will always look up to, as well as to Alana Thurston and Amy Vaughan, the crazy progesterone plant squad. Your support and love have been paramount to my wellbeing, and I love you all very, very much. Thank you to my silly nerds, Rob, Holly, Connor, Vikki, Nathan, and Sarah. It has been such a privilege to play board games and DnD with you over the years, and I am looking forward to all the future gaming! Thank you to Johanna and Julian, for providing me a first proper home in New Zealand.

And thank you to my family and friends, back in Finland. Kiitos äiti, että kasvatit minusta itsenäisen – ja itsepäisen. Kiitos, että aina kannustit tekemään sitä, mitä halusin ja mikä teki onnelliseksi. Kiitos kaikesta isä, mummi, Kirsi, Siina ja Eelis. Kiitos kaikille ihanille, rakkaille ystävilleni koto-Suomessa, olette usein ajatuksissa, kiitos, että olette olleet tällä matkalla mukana tsemppaamassa. Erityisesti Bio-jengi, Kimi ja Anna, sekä Monica, isot halit ja kiitokset, olette hurjan rakkaita. Kiitos myös kandin- ja maisterintutkintojeni ohjaajille: Riitta Nissinen, Leena Lindström, Eeva-Riikka Vehniäinen ja Aigi Margus, kiitos, kun autoitte minut alkuun tieteilijänä.

Finally, Tim, I don't know what I would have done without you by my side. Thank you so much for being my rock, being the greatest support and the biggest cheerleader for the past three years. Thank you for helping me out with my field work and listening to me complain about silly things (like stats and dock weeds) and thank you for providing so many silly moments filled with laughter and love. Also, I'm terribly sorry you'll now see *Rumex* everywhere, including TV shows and movies. Love you very much, and I cannot wait to start the next chapter with you. I'm also very thankful to Tim's family, for all the care and support. Special thanks to Tim's mom Kathryn for all the APA wisdom!

The past four-and-a-bit years have been truly magical, and despite some ups and downs (and more ups and very big downs), most of my PhD journey has been amazing. I have enjoyed practicing being a researcher and I am excited to see where the journey takes me. I want to leave you with this great quote that I found as I was writing up the thesis, and that still makes me chuckle every time I read it:

"Docks are not the sort of plants that attract the attention of most naturalists and field botanists."
~ J.K. Morton, 1989

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1. Plant invasions

Globalisation and with it the movement of people and trade goods has enabled thousands of plant species to spread to new areas, and some of them have managed to establish, multiply, and become invasive (Bossdorf et al., 2005; Barrett et al., 2008; Seebens et al., 2017). More research is needed to understand the traits and processes behind invasiveness as predictions show no plateau in the number of introductions in the coming decades. A multitude of studies since Elton's *The Ecology of Invasions by Animals and Plants* (Elton, 1958; Kolar & Lodge, 2001; Richardson & Pyšek, 2008) have added to the understanding of what makes some species invasive and despite some contrasting findings, there are recurring theories and themes (e.g., Bossdorf et al., 2005; van Kleunen, Dawson, et al., 2015; van Kleunen et al., 2018). The enemy release hypothesis (ERH)(Keane & Crawley, 2002) in tandem with evolution of increased competitive ability (EICA)(Blossey & Notzold, 1995) and the shifting defence hypothesis (SDH)(Müller-Schärer et al., 2004) state that species often lack specialised herbivores and pathogens in the introduced range which allows them to allocate resources towards growth rather than defence (for example Mounger et al., 2021). An increased phenotypic plasticity within the introduced populations compared to the native ones (Catford et al., 2009) allows these "general-purpose genotypes" to effectively respond to their new environmental conditions. Alternatively, rapid increase of genetic variation, especially under multiple introductions and/or hybridisation, can lead to an increased adaptation ability. Lastly, niche shifts can allow the species to occupy new habitats in the introduced range compared to the native range, often attributed to rapid adaptation or phenotypic plasticity. This is not an exhaustive list, and other aspects, such as propagule pressure (Kolar & Lodge, 2001; Colautti et al., 2006) and some plant ecological traits, such as high seed set and fast growth-rate (van Kleunen, Dawson, et al., 2015), have been found to be important during introductions. However, despite being intensely studied, invasion biology still lacks knowledge regarding the interplay between the aspects behind invasiveness, and any individual trait is unlikely to explain invasiveness alone (Catford et al., 2009; Sherpa & Després, 2021). Thus, the lack of studies comparing these aspects using multiple species with varying invasiveness and both within the native and introduced regions mean that our ability to see broader patterns is still limited.

In this thesis, I will focus on the genetic traits and processes potentially driving invasiveness, such as genome size and ploidy, intra- and interspecific hybridisation, the amount of genetic variation, as well as demographic traits and processes, such as plant breeding system and specifically, self-compatibility and potential costs from selfing. These traits and processes, hereafter "demo-genetic traits and processes" (after Sherpa & Després, 2021), go hand-in-hand: genetic variation can influence species' adaptive ability, but also, variation itself can be influenced by ploidy and both intra- and interspecific hybridisation, as well as the breeding system (FIGURE 1.1). My thesis is a part of broader research investigating the traits behind invasiveness using non-ornamental

weed species from their native and introduced ranges. Three species from the genus *Rumex* L. (Polygonaceae) were used as study species. A comprehensive look into the genetic aspects and their interplay in invasiveness will help us to broaden our understanding of invasive species and add to the existing knowledge from these species. My thesis is the first study investigating the genetics of these hermaphrodite *Rumex* species, and a first study observing their hybrids in New Zealand (NZ).

Previous research in NZ on *Rumex* shows that enemy release could explain the success of *R. crispus* in NZ, but not the other two species, despite plants from NZ having a decrease in herbivory compared to plants from the UK (Costan et al., 2022). Furthermore, both *R. conglomeratus* and *R. obtusifolius* show high phenotypic plasticity, but only *R. obtusifolius* shows higher tolerance of drought in the introduced range compared to the native range, suggesting that plasticity could have contributed to its invasion success (Bufford & Hulme, 2021a). In addition to a high niche similarity between NZ and the UK, all three species occupy different niches in NZ compared to the UK, showing a niche shift towards wetter climates with less variation in precipitation (Carlin et al., 2022). Lastly, all three species show no seed size-number trade-offs in NZ, whereas the trade-offs do exist in the native range (Bufford & Hulme, 2021b). Including a genetic perspective to these studies will further help integrate these findings, which is something that invasion ecology still lacks (Catford et al., 2009; Sherpa & Després, 2021). In addition, there is a strong historical reason to assume that the introductions originated from the UK, but as the genetic differentiation between the provenances has not been assessed, this has remained unconfirmed. Identifying the source population is important in order to set a baseline onto which the introduced populations can be compared (Estoup et al., 2016). Without a baseline, it is difficult to assess whether the observed differences in *Rumex* – or in fact any species – are due to pre- or post-adaptation (Ellstrand, 2009).

The terminology in this thesis follows the definitions by Richardson et al. (2000) and is as follows: "Naturalised species" refers to a non-native species, which has been introduced to a new area and has overcome a geographic barrier for migration and has subsequently been able to establish and sustain a population without further introductions. The term "invasive species" refers to naturalised species that have further overcome barriers for spread and have been able to spread over large regions outside the site of introduction. Lastly, the term "weed" refers to plant species that cause harm for humans and animals. The three study species can thus be classified as naturalised, invasive weeds.

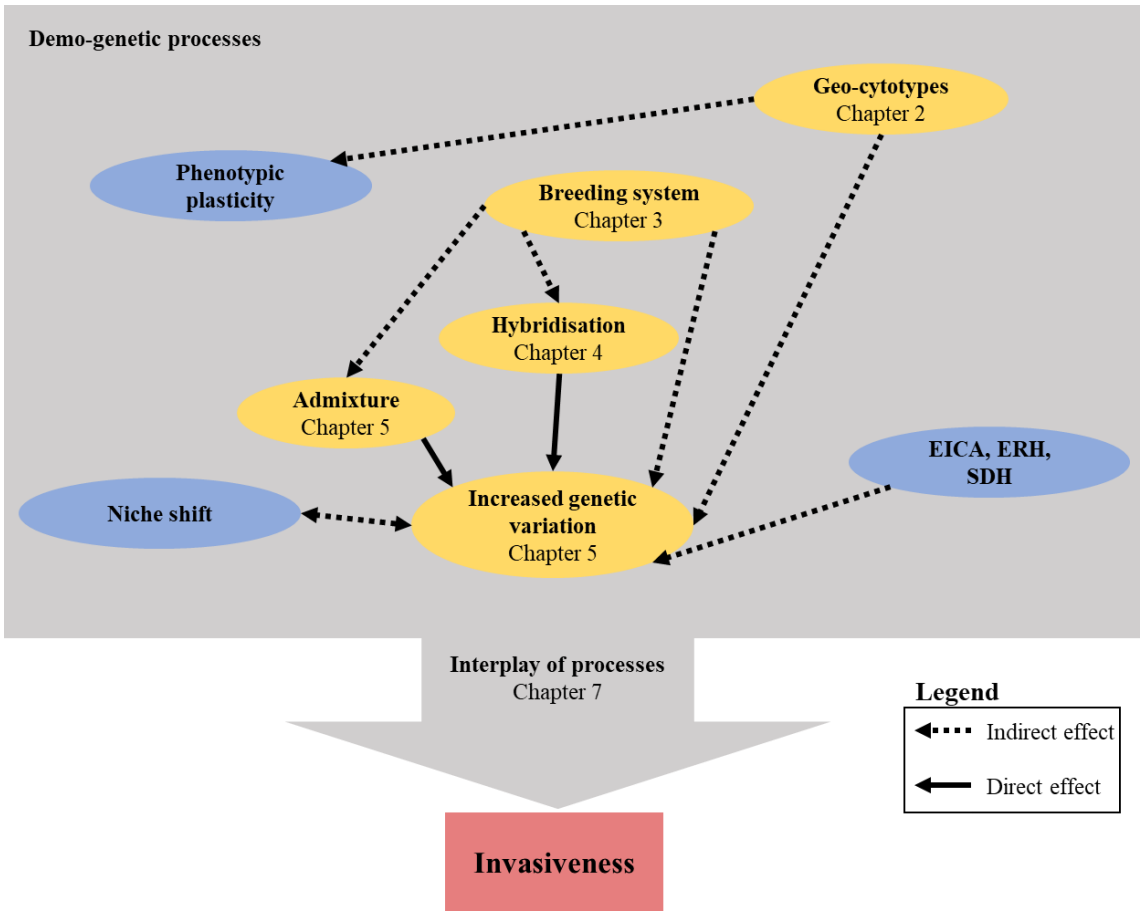

FIGURE 1.1: An overview of the different demo-genetic traits and processes assessed in this thesis (yellow), and their interplay with processes investigated before (blue), and whether they contribute to the increased invasiveness within New Zealand. In the figure, hybridisation refers to interspecific hybridisation, whereas admixture refers to intraspecific hybridisation. EICA: evolution of increased competitive ability; ERH: enemy release hypothesis; SDH: shifting defence hypothesis.

1.1.1. Demo-genetic traits and processes behind invasiveness

1.1.1.1. Breeding system

Outcrossing (xenogamy) and self-fertilisation (cleistogamy) are considered the main breeding methods in plants (Briggs & Walters, 2016), although rather than being fully separate classes, they form a continuum from facultative xenogamy to facultative cleistogamy, with mixed breeding systems in between (e.g., Cruden, 1977). A third breeding class, apomixis or breeding without fertilisation, covers both seed apomixis and vegetative apomixis (Briggs & Walters, 2016), the former of which is outside the scope of this thesis. In vegetative apomixis, or clonal reproduction, plants reproduce via rhizomes or runners, for example, or regenerate from root or shoot fragments. In addition to these, the plant sex distribution can be dioecious, monoecious, or hermaphrodite, or separate female and male plants, separate female and male flowers, or flowers with female and male structures, respectively (Briggs & Walters, 2016). Similarly to the breeding systems, the sex distribution can also have mixed modes, such as gynodioecy (female and

hermaphrodite plants), androdioecy (male and hermaphrodite plants), and even gynomonodioecy (gynodioecious and monoecious plants). Lastly, hermaphrodite flowers can have protandrous or protogynous maturation patterns, where either male or female reproductive structures mature first, respectively. In addition, both temporal and spatial variation of these breeding systems exist (e.g., Friedman & Barrett, 2011). For example, some species of *Viola* produce xenogamic flowers in the spring with insect-pollination allowing outcrossing, and cleistogamic flowers later in the summer (Briggs & Walters, 2016), likely guaranteeing reproductive output for the year (Barrett, 2013).

The breeding system can have a huge impact on species establishing into new areas. For example, if a plant is self-compatible or clonal, even a single individual can establish a new population (Baker, 1955; Cruden, 1977; Barrett et al., 2008). This phenomenon of selfers as more successful colonisers over mate-dependent out-crossers has been termed Baker's Law. However, there can be consequences of decreased fitness from inbreeding, called inbreeding depression (e.g., Charlesworth & Willis, 2009), such as slower or limited population growth or a lower adaptation ability (Allendorf & Lundquist, 2003). This fitness reduction is said to be pronounced under stressful environments (Fox & Reed, 2011). Consequently, plants with mixed breeding system might have the biggest advantage during colonisations, as they can establish in low numbers, but prevent inbreeding depression by outbreeding (Briggs & Walters, 2016). For example, beggarticks *Bidens frondosa* L. (Asteraceae) is capable of both self- and cross-pollination, and this mixed mating system coupled with high seed production and high germination rate is thought to contribute to its invasiveness (Yan et al., 2016).

According to Cruden (1977), plants have optimised costly pollen production to produce enough pollen for their reproductive needs. This means that the number of pollen grains produced correlates with the plant's breeding system: selfing plants need fewer pollen grains than outbreeding plants, especially if the outbreeding plant is also wind pollinated. Thus, the pollen:ovule (P:O) ratio, can be used as a proxy for the breeding system (Cruden, 1977). Multiple studies have found this relationship: Michalski & Durka (2009) found that the P:O ratios correlated with the breeding system in 107 angiosperms. However, the correlation was stronger for wind-pollinated than for animal-pollinated species, and weaker in woody perennials.

1.1.1.2. Hybridisation and admixture

Both hybridisation within species (intraspecific hybridisation or admixture) and interspecific hybridisation can benefit species during colonisation, establishment, and subsequent spread (Ellstrand & Schierenbeck, 2006). Hybridisation and admixture in general can add genetic variation and new alleles into populations that can be beneficial. These positive effects can be caused by multiple mechanisms, such as genetic rescue, where the introduction of new alleles purges deleterious homozygous alleles, and overdominance (also known as heterozygote advantage), where heterozygous alleles have a fitness advantage over homozygous alleles (summarised in Barker et al., 2018). Similarly to overdominance, hybridisation can lead to heterosis or hybrid vigour, where the offspring has higher fitness than either of the parent species (Rieseberg & Carney, 1998; Bar-Zvi et al., 2017), and also purges the genetic load (genetic rescue; Ellstrand & Schierenbeck, 2006). However, both methods can have negative consequences (Barker et al., 2018) as well as waning benefits in later generations (Rieseberg & Carney, 1998).

Hybridisation and admixture between too distant individuals can break up beneficial adaptation similarly to outcrossing, and if interspecific hybrids have lower fitness or are sterile, hybridisation can use important resources and hinder establishment. In yellow starthistle *Centaurea solstitialis* L. (Asteraceae) admixture both increased and decreased the fitness of the progeny, depending on the genetic distance between the parents (Barker et al., 2018), whereas seep monkeyflower *Erythranthe guttata* Fisc. Ex. DC (Phrymaceae; previously *Mimulus guttatus*) benefited from admixture between native and introduced lineages (van Kleunen, Röckle, et al., 2015). This highlights the important role of admixture during invasions but also the complicated and inconsistent consequences from it: without explicitly investigating impacts from admixture, it is difficult to know whether it benefits or hinders introduced populations.

Depending on the question, hybridisation and the effects of admixture can be studied using field surveys, cross-pollination experiments (e.g., van Kleunen, Röckle, et al., 2015; Barker et al., 2018) or genetic tools, as well as assessing the invasion history of the species (e.g., do we know how many times a species was introduced and/or from where). Genetic tools, namely sequencing, such as genotyping-by-sequencing (GBS)(Elshire et al., 2011), and subsequent analyses, such as Principal Components Analysis, can suggest hybridisation or admixture between the species or samples studied.

1.1.1.3. Genome size and ploidy

Studies have shown that a smaller genome size is more often associated with increased weediness and a higher likelihood in naturalisation compared to a larger genome size (Suda et al., 2015; Pyšek et al., 2018). This was shown to be true with weed species in the UK (Bennett et al., 1998) as well as in the Czech Republic (Kubešová et al., 2010). Smaller genomes have been linked with faster reproduction, which might partially explain the relationship. However, there is also a link between polyploidy (or multiple chromosome sets) and invasiveness, with higher ploidies being associated with invasive species. This might not be as contrasting as it seems, as polyploids are known for downsizing genomes (Leitch & Bennett, 2004). This means that the amount of DNA within one set of chromosomes (basic genome or 1Cx value) decreases with increasing ploidy, still allowing the basic genome to replicate faster in polyploid species. The genome size and ploidy are quantified using C-values (Greilhuber et al., 2005; Pandit et al., 2014). According to the terminology by Greilhuber et al. (2005), 1C-value refers to the gametic genome size (haploid, *haplóos*, 'simple'), 2C to the nuclear genome size (holoploid, *hólos*, 'whole'), and 1Cx-value, where x denotes the ploidy, refers to the genome size within one set of chromosomes (monoploid, *mónos*, 'single').

There are two methods for polyploid creation: autopolyploidy, where one chromosome set multiplies, and allopolyploidy, where, after hybridisation, two different chromosome sets come together (Briggs & Walters, 2016). In this thesis, my focus is mainly on allopolyploids, although many of the same processes apply to both types of polyploids. The beneficial effects from ploidy are due to multiple allele copies, often leading to fixed heterozygosity (Soltis & Soltis, 2000; te Beest et al., 2012). In addition, polyploid species often have higher environmental tolerance, faster growth-rate, and bigger seeds, although the evidence is not conclusive (Soltis & Soltis, 2000; te Beest et al., 2012; Briggs & Walters, 2016). Taken together, this can greatly increase the species' colonisation ability. In a study by Moura et al. (2021), the authors found that species with

higher ploidies showed an increased invasiveness via higher environmental tolerance compared to lower ploidies. However, this study did not find a link between genome size and invasiveness, but this might be due to not taking ploidy into account when assessing the genome size (i.e., haploid genome size instead of monoploid genome size). The shorthand for writing the chromosome number of the species is $2n = 2x = 20$, or $2n = 4x = 40$, indicating that the non-reduced (somatic as opposed to gametic) chromosome number of a diploid (2x) or tetraploid (4x) individual is 20 or 40, respectively (see for example Greilhuber et al., 2005).

A study comparing the haploid and monoploid genome sizes of over three thousand plants found that the smaller the 1C-value, the higher the likelihood for the species being identified as a weed (Chen et al., 2010). However, the authors found that monoploid genome size is a better metric of the invasiveness than 1C-value is, as 1C value can increase with the ploidy, but 1Cx-value often decreases, thus linking ploidy and genome size. Chen et al. (2010) suggested that the increased invasiveness due to small genomes could be due to a short juvenile phase, short generation time, and a fast production of small seeds, all of which are linked to small genome size. Pandit et al. (2014) reached a similar conclusion in their study and suggest that ploidy adds to the relationship through heterosis and increased phenotypic variation, for example.

The traditional method for assessing the chromosome number is karyotyping or manual count and assessment of the chromosomes (Windham et al., 2020). This 'squash' technique can be time-consuming and requires fresh, growing tissue, such as root tips (Dawson, 1993), but it allows counting and morphological observations, as well as storing the samples for later, if required. In short, the preferred fresh tissue is stained, and mounted on a glass slide. Next, using a light microscope (×1000), the meiotic cells in a correct state (diakinesis or metaphase, for example) are assessed (Dawson, 1993; Windham et al., 2020).

Flow cytometry is the most used method as well as the gold standard, for assessing the genome size and ploidy using DNA (DNA-ploidy)(Doležel et al., 2007; Sliwinska, 2018; Pellicer & Leitch, 2020). In short, the sample nuclei are stained with a fluorochrome (propidium iodide for genome size, DAPI for ploidy analyses) and their fluorescence is measured and recorded with a flow cytometer. The mean fluorescence value can be then compared to that of a standard with a known genome size or ploidy and calculated using the ratio of the sample/standard values. Flow cytometry is fast and reliable, and suitable standards exist from very small to very large genomes (Doležel et al., 2007). Ploidy assessed using flow cytometry should be called DNA-ploidy, to differentiate from manual chromosome counts, and as chromosome number does not always correlate with genome size, manual chromosome counts should be performed for a baseline/verification (Suda et al., 2006; Sliwinska, 2018; Windham et al., 2020). In addition, when measuring the DNA-ploidy, it is recommended that the standard is a known individual within the same species compared to the study individual (Doležel et al., 2007).

1.1.1.4. Genetic variation

Genetic variation is one of the most studied and argued aspects behind invasiveness. Genetic variation is ultimately due to mutations in the DNA and reflected as different allele compositions between individuals (e.g., Ward et al., 2008). It is generally assumed that some level of genetic variation is a prerequisite for adaptation, although the interplay

with other demo-genetic traits and processes can counter low genetic variation (Ward et al., 2008). However, it is difficult to concretely say what exactly is the sufficient amount of genetic variation needed, and it can be difficult to disentangle whether the invasion success is due to pre-adaptation or selection after colonisation (Ward et al., 2008; Bock et al., 2015).

There are multiple factors that can decrease or increase the genetic variation within a species, most notably bottlenecks and multiple introductions, respectively. The success of invasive species despite bottlenecks is called the invasion paradox (Allendorf & Lundquist, 2003; Frankham, 2005; Estoup et al., 2016), and it puzzled scientist for several years. The belief was that introduced populations almost always undergo a genetic bottleneck during invasions which severely decreases the genetic variation available (Allendorf & Lundquist, 2003). However, nowadays, we know that populations with very limited genetic variation, such as clonal populations, have managed to establish and become invasive (e.g., *Fallopia japonica*; Hollingswort & Bailey, 2000; but see Bzdega et al., 2016). This can be due to multiple factors, such as genetic variation decreasing in genomic areas that are not required for adaptation, introductions occurring in an environment similar enough to the original populations, or due to buffering effects from ploidy (reviewed in Estoup et al., 2016). Most often, however, species are introduced multiple times and never undergo a genetic bottleneck in the first place (Frankham, 2005; Bock et al., 2015; Estoup et al., 2016).

A plethora of methods exist for measuring genetic variation, such as microsatellites, single-nucleotide-polymorphisms (SNPs), or whole-genome-sequencing. By comparing the genetic diversity between native and introduced populations, we can construct invasion histories and learn whether the introduced populations have undergone bottlenecks or admixture (Ward et al., 2008). For the method chosen for this study, see Section 1.1.2.1 below.

1.1.1.5. The interplay between processes and traits

Demo-genetic traits and processes are known to influence each other. According to early studies regarding life-history traits and genetic diversity by Hamrick & Godt (1996), breeding system affects genetic variation and its partitioning within and between populations. They found out that outcrossing populations tend to have lower genetic differentiation among populations compared to inbreeding populations. However, inbreeding is known to decrease genetic variation on population level (Briggs & Walters, 2016). This tends to mean a slower accumulation of potentially useful new mutations. On the other hand, inbreeding species can fix beneficial mutations faster than outcrossing species (Charlesworth & Willis, 2009). Polyploidy, in turn, is known to reduce inbreeding depression (te Beest et al., 2012 and the references within), and polyploid species are more commonly selfing compared to their diploid counterparts (Barringer, 2007). However, this buffering against selfing might not be long-term and the decrease is more pronounced in autopolyploids compared to allopolyploids (Barringer, 2007). Despite that and especially when coupled with other traits and processes, such as increased genetic variation, the benefits from ploidy can greatly aid invasive species. Furthermore, ploidy is also known to increase genetic variation. In giant goldenrod *Solidago gigantea* Aiton (Asteraceae), the hexaploid populations had higher heterozygosity compared to the tetraploid populations (Nagy et al., 2018). However, the tetraploid populations had a

better performance in a common garden study compared to the hexaploid populations, and the tetraploid populations are found within the introduced regions.

As stated above, both admixture and hybridisation can increase genetic variation by introducing new allele combinations to populations (Ellstrand & Schierenbeck, 2006; Ward et al., 2008). However, even when the hybrid offspring themselves are not strong competitors, they can produce offspring with the parent species and introduce genetic variation from one species to the other, creating new, potentially beneficial allele combinations (introgression; Rieseberg & Carney, 1998; Bock et al., 2015). Similarly, admixture via multiple introductions (propagule pressure) can bring together distant populations and add new allele combinations to populations (Bock et al., 2015). This increased genetic variation can help populations to better adapt to the new environment or to keep up with changing environmental conditions, such as climate change (e.g., Schierenbeck, 2017). However, like decreased genetic variation, increased genetic variation alone is unlikely to explain the success of invasive species (Bock et al., 2015).

Beyond the demo-genetic traits and processes listed above, both adaptive phenotypic plasticity and epigenetics can counter the loss of genetic variation during invasions and help in establishment (Estoup et al., 2016). These are outside the scope of this thesis, but I will assess their role in the three *Rumex* species in Chapter 7. In the cordgrass genus *Spartina* Shreb. (Poaceae), hybridisation between *S. alterniflora* Loisel and *S. maritima* (Curtis) Fernald led to the creation of two sterile hybrids, one of which, *S.* × *townsendii* H.Groves & J.Groves, underwent epigenetic changes and a polyploidisation leading to the formation of *S. anglica* C.E.Hubb. (Ainouche et al., 2009). Today *S. anglica*, is considered to be one of the worst weeds in the world (Invasive Species Specialist, n.d.) and it is more vigorous than either of the parent species, disperses both via seeds and vegetatively, and has higher ecological tolerance than the parents coupled with high phenotypic plasticity, leading to a niche shift (Ainouche et al., 2009; Le Roux, 2022). Thus, *S. anglica* is a good example of the complicated interplay between the demo-genetic traits and processes and highlights the importance of a more holistic approach in understanding invasions. We need to compare multiple demo-genetic traits and processes in tandem in order to determine the main driver and create a framework for understanding invasions (Catford et al., 2009; Briggs & Walters, 2016).

1.1.2. Invasion genetics

> *"-- invasion genetics is the study of the historical, ecological and demographic processes responsible for the patterns of genetic diversity in populations and their influence on invasion success and contemporary evolution during biological invasion"*

Invasion genetics, as defined above by Barrett (2015), is a relatively young field of study drawing from and bridging the gap between invasion biology and invasion ecology, but also aids in management of invasive species (Colautti et al., 2005; Barrett, 2015; Bock et al., 2015). Invasion genetics aims to assess the genetic variation present in native and introduced populations, identify the source of the introductions via genetic comparisons, and understand the amount and type of variation required for establishment (Colautti et al., 2005; Barrett, 2015). It helps us to disentangle whether the invasion success due to

niche expansions, for example, is enabled by standing genetic variation or whether the established populations have gone through local adaptations. Answering these questions often starts by identifying the source population (Estoup et al., 2016). This then works as a baseline to which the introduced populations can be compared to.

Various genetic tools (i.e., sequencing) have been utilised in invasion genetics. The information provided by sequencing can be paired with common garden experiments or field studies to compare heritable differences between native and introduced, or multiple introduced populations (Ward et al., 2008; Barrett, 2015). Measures such as fixation index (F_{ST})(Wright, 1943) can be used to inform of the extent of differentiation between the study populations (e.g., Ward et al., 2008). This can be further paired with analyses such as analysis of molecular variation (AMOVA)(Excoffier et al., 1992) to suggest whether there is geographical structuring between the study populations, or whether gene flow has kept differentiation minimal.

1.1.2.1. Genotyping-by-sequencing

A multitude of methods exist for analysing the genetic variation within species, and more and more methods are developed each year (Andrews et al., 2016; Grover & Sharma, 2016; Scheben et al., 2017). However, the field of population genetics was revolutionised with the development of next-generation-sequencing techniques, which made sequencing easier, cheaper, and more accurate. With next-generation-sequencing came reduced representation sequencing (RRS) methods, which allow – as the name suggest – for a reduction of the complexity of the genomes (Altshuler et al., 2000). This is often done using restriction enzymes that find short, repeating sequences within the genomes and cut them into shorter pieces, allowing for the sequencing of the flanking regions.

Genotyping-by-sequencing (GBS) is a RRS method that uses one or two restriction enzymes to cut the genome into smaller pieces (Elshire et al., 2011; Poland & Rife, 2012). The areas around the restriction sites are then sequenced as opposed to the whole genome, which makes sequencing cheaper and faster. In contrast to microsatellite studies, this method yields thousands or tens of thousands of SNPs, which increases the capacity to detect patterns between individuals and populations (Attard et al., 2018; Lemopoulos et al., 2019). In addition, GBS does not require prior knowledge of the genome, unlike microsatellites, and is thus usable in non-model organisms (Elshire et al., 2011; Poland et al., 2012; Poland & Rife, 2012). Contrast to microsatellites, however, GBS can be computationally demanding and requires some level of bioinformatics skills, as well as careful consideration of the SNP-calling pipelines and downstream analyses chosen. It has been successfully used on many invasive species in assessing their invasion history and origin, admixture, and genetic variation within and between the native and introduced ranges. For example, common ragweed *Ambrosia artemisiifolia* L. (Asteraceae) was found to have more geographic structuring than previous microsatellite studies found (Martin et al., 2016). This newfound structure suggests a phylogeographic shift within North America, possibly due to environmental selection and adaptation to local conditions. In a different study, *A. artemisiifolia* was found to have been introduced multiple times from North America to Europe and as a bridgehead invasion to Australia (van Boheemen et al., 2017). The authors suggest that the populations were admixed within the native range prior to introductions, rather than introductions to Europe from multiple sources leading to admixture.

1.2. *Rumex* spp. as study species

In this thesis, three species from the genus *Rumex* L. (dock weeds, family Polygonaceae) were used as study organisms. The species were *R. conglomeratus* Murr., *R. obtusifolius* L., and *R. crispus* L. They are among the world's worst weeds (Holm, 1977), but despite – especially the latter two – being intensely studied (Zaller, 2004), we know nothing about their genetic variation, or the genetic traits behind their success. In addition, these species of *Rumex* are not considered ornamentals and thus have not been under cultivation and trait selection. This means that all genetic differences observed are due to natural adaptation to the surrounding environment, which makes them ideal candidates for weed research.

These species were introduced to NZ in the early to mid-1800s as grass seed contaminants and moved around via uncleaned tools and machinery (Thomson, 1922; Allan, 1937). Already by the 1830s, the previously weedless Māori gardens were taken over by European weeds, including both *R. obtusifolius* and *R. crispus* (Darwin, 1839; Leach, 2005). Indeed, in 1900, all introduced *Rumex* species were listed in the Noxious Weeds list, and their entry to the country was banned (The Noxious Weed Act 1908). Given the colonisation history at the time, the assumed origin of the species introductions is the UK, but this has not been verified using genetics. Both Darwin (1839, p. 511) and Colenso (1844) say that the Englishman had sold dock seeds to Māori as tobacco plants, highlighting the UK as at least one of the sources for the early introductions.

Given the introduction method, the study species were likely introduced multiple times and suffered no genetic bottlenecks during the introductions. This suggests that the genetic variation has been maintained in a steady level via admixture. Despite this, if the plants came from a small area, the genetic diversity within NZ could still be relatively low, compared to the native range. These species – and especially *R. crispus* and *R. obtusifolius* – are known contaminants of multiple crop and pasture species, such as Chewing's fescue *Festuca rubra* ssp. *commutata* Gaud. (Poaceae) and red clover *Trifolium pratense* L. (Fabaceae) (Reddy et al., 1998; Rubenstein et al., 2021; Stewart et al., 2022). However, *Rumex* species can be managed with certain herbicides (Reddy et al., 1998; Harrington, 2019), and given the small number of contaminants found, as well as the rigorous seed detection protocol in NZ (see for example Rubenstein et al., 2021), these reintroductions after the Noxious Weed Act are unlikely to contribute to the weediness of NZ *Rumex* today.

We know that within their native range, these species hybridise often and freely (Williams, 1971), and despite indications of lower hybrid fitness (Cavers & Harper, 1964), there are reports of hybrid swarms as well as introgression back to the parent species (Cavers & Harper, 1964; Williams, 1971; Ziburski et al., 1986; Akeroyd, 2014; Takahashi & Hanyu, 2015). No official reports of hybrids have been made in NZ, indicating that it is not known whether hybridisation could affect their success.

All three species are fast-growing, have high seed-set, and are primarily wind-pollinated hermaphrodites (Cavers & Harper, 1964; Grime et al., 2007; Akeroyd, 2014). They are of different ploidies and records of geo-cytotypes exist (Rice et al., 2015; Pellicer & Leitch, 2020), but the base chromosome number is 10 (Ichikawa et al., 1971), and first-generation hybrids show chromosome numbers half-way between the parents (Williams, 1971). These species, and *Rumex* species in general, can be difficult to identify based on morphology alone, due to their highly varying characteristics (Löve & Kapoor,

1967). The best identifier is the structures in the mature fruits, namely the shape and margins of the valves, but leaf characteristics as well as the shape of the inflorescence can aid in identification. However, as the fruits mature, the rosette leaves tend to wilt making them unsuitable for identification.

1.2.1. *Rumex conglomeratus*

This species is the least weedy out of the three study species and has considerably less literature available. It is a diploid species usually found in damp areas such as ditches or riversides, rarer than the other two species, but still commonly found all over the world (Harper & Chancellor, 1959; Stace, 1991; Akeroyd, 2014). The leaves of *R. conglomeratus* are lanceolate with a rounded base and smooth margins, the inflorescence has branches on up to 90º angle to the main stem, flowers are in clusters, and after the fruit matures, the valves are long and narrow with clear tubercles (Morton, 1989; Stace, 1991; Akeroyd, 2014). These plants have wind-pollinated hermaphroditic flowers (Navajas-Pérez et al., 2005) that are assumed protandrous (Webb et al., 1988). The reproduction of *R. conglomeratus* is said to be both clonal via ramets and sexual via seeds (Kołodziejek, 2019), and hybrids with other *Rumex* species are abundant in the UK (Akeroyd, 2014). Research shows that *R. conglomeratus* seeds are buoyant (Boland, 2017), and thus can spread long distances via waterways. They have also been found occasionally in ballast waters (Nelson, 1917).

Previous studies suggest that *R. conglomeratus* exhibit high phenotypic plasticity in response to drought and flooding, but plants from the UK and NZ do not differ from each other (Bufford & Hulme, 2021a). This suggests that phenotypic plasticity is unlikely to explain the success of *R. conglomeratus* in NZ. Similarly, the NZ plants showed a decrease in both root and shoot herbivory, but this did not lead to increased biomass, suggesting that enemy release hypothesis is unlikely behind invasiveness (Costan et al., 2022). However, *R. conglomeratus* exhibits seed size-number trade-offs in the species native range, but not in the introduced range (Bufford & Hulme, 2021b). The authors suggest that this is more likely to be due to available resources rather than evolution. This was not confirmed using genetics, but nevertheless, a potential higher nutrient availability within NZ is an important aspect to take into consideration. A study by Carlin (2021) indicates that *R. conglomeratus* might benefit from climate change and shift its range towards colder regions, such as Northern Europe, where it is currently absent. This study also highlights potential environmental limitations rather than ecological limitations.

1.2.2. *Rumex obtusifolius* and *Rumex crispus*

These two species are more similar to each other compared to *R. conglomeratus* and are both widely studied (Cavers & Harper, 1966; Zaller, 2004; Grime et al., 2007). Both *R. obtusifolius* and *R. crispus* are said to be among the world's worst weeds, and *R. crispus* one of the most wide-spread species in the world (Cavers & Harper, 1964; Holm, 1977). Zaller (2004) reviewed published information and found over 500 articles published between 1973 and 2003 researching the two species. No mentions of genetic studies exist, and apart from a study assessing the taxonomy of the genus using the ITS region and

chloroplast sequence (Navajas-Pérez et al., 2005) and 5S and 45S rDNA sites (Kim et al., 2006), this holds true to this day. In addition, while morphological studies are common, the majority of the information comes from the species' native range. Both species have multiple subspecies and varieties (Cavers & Harper, 1964), but for simplicity, poor information of some of them, and no information of the subspecies in NZ, I have not discriminated between them in this thesis.

These species grow in paddocks, wastelands, roadsides, as well as in ditches and riversides, and *R. crispus* in particular has a preference for wetter habitats (Cavers & Harper, 1964; Zaller, 2004). Both species are wind-pollinated, although occasional insect visits have been recorded (Grime et al., 2007), hermaphroditic, self-compatible, and known to hybridise freely (Cavers & Harper, 1964; Grime et al., 2007). However, the extent of selfing or inbreeding depression is not known. Hybrids are often sterile but reports of hybrid swarms and introgression exist (Cavers & Harper, 1964; Ziburski et al., 1986; Takahashi & Hanyu, 2015).

The leaves of *R. obtusifolius* are oblong in shape with a chordate base and smooth margins, whereas *R. crispus* has long, narrow lanceolate leaves with a truncate base and wavy margins (Akeroyd, 2014). The inflorescence of *R. obtusifolius* has multiple branches at roughly in a 45º angle, compared to *R. crispus* which has fuller inflorescence with fewer branches on a smaller angle to the main stem. Valves in *R. obtusifolius* have teeth and are triangular, tubercles often not very prominent, whereas in *R. crispus*, the valves are wide with smooth margins and have well-developed tubercles.

Under both drought and flooding, *R. obtusifolius* shows increased phenotypic plasticity, but only under drought conditions is it higher in the introduced region compared to the native region (Bufford & Hulme, 2021a). Both species show a decrease in herbivory damage within the introduced range compared to the native range, but only *R. crispus* shows increased mass within the introduced range (Costan et al., 2022). This can be a consequence from enemy release, which could contribute to its invasion success in NZ. Similarly to *R. conglomeratus*, neither of these species show seed size-number trade-offs in NZ, while these trade-offs exist in the native range (Bufford & Hulme, 2021b). While the extent of selfing is highly variable (Cavers & Harper, 1966), *R. crispus* does not show a decreased seed yield when distance to conspecifics increases, which could indicate that selfing acts as a buffer when mates are rare (Friedman & Barrett, 2009).

1.2.3. Sample collection

Previously collected seed samples were used in this thesis (for full methods see Bufford & Hulme, 2021a, 2021b; Carlin et al., 2022; and Costan et al., 2022). The seed material originated from four regions in both the UK (Cornwall, Cambridgeshire, Glasgow, and Edinburgh) and NZ (Canterbury, Southland, Otago, and Westland) and from four populations within each region (FIGURE 1.2). Seeds were collected from 10 individual plants from each population, totalling at 960 seed families (seeds from one mother plant), or 320 seed families per species. The regions were chosen to represent the available habitats within both countries and were separated by tens or hundreds of kilometres and sometimes by a mountain range (e.g., Canterbury and Westland)(Bufford & Hulme,

2021a). The populations within the regions were tens of kilometres from each other and had around one hundred fruiting individuals within a radius of 25 metres.

The seeds were collected from plants that were screened for hybrids, and samples were taken only from true species using seed and leaf morphology as an identifier (Bufford & Hulme, 2021a, 2021b; Costan et al., 2022). However, this method does not rule out hybrid seeds, because the pollen donor is not known. *Rumex* hybrids are also known to introgress back to the parent species (Ziburski et al., 1986; Takahashi & Hanyu, 2015), which can make it hard to identify later generation hybrids.

The plants were grown from seeds in Lincoln University's greenhouse in Canterbury, NZ. Plants were re-potted several times during growing and cut back during re-potting after they reached maturity. They were sprayed with insecticide when needed and before sample collections, to ensure fresh, healthy growth.

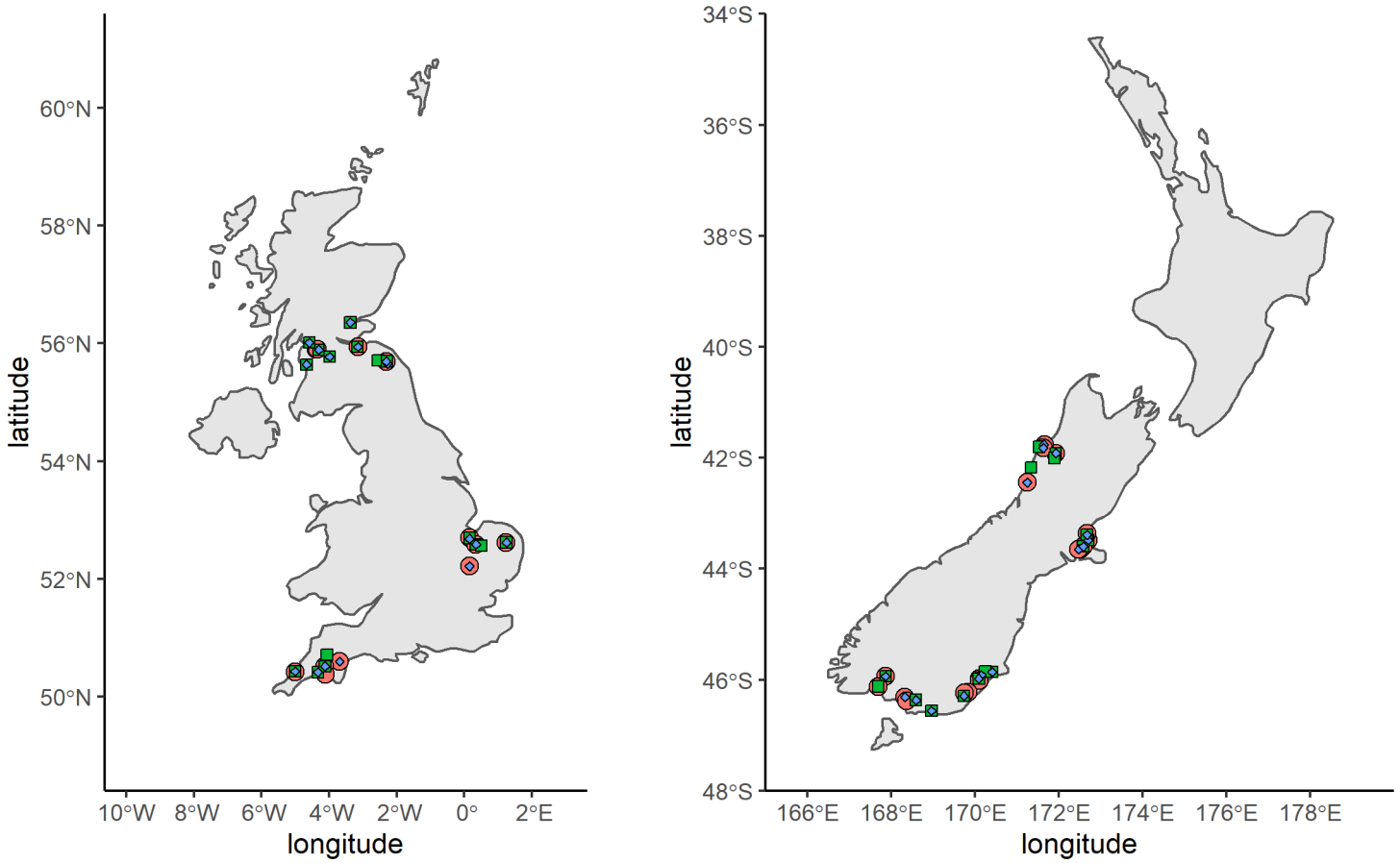

FIGURE 1.2: Maps showing the sampling locations for *Rumex* species within the United Kingdom (left) and New Zealand (right). Red circles indicate *Rumex conglomeratus* populations, blue diamonds indicate *R. obtusifolius* populations, and green squares indicate *R. crispus* populations. Reproduced after Bufford & Hulme (2021b) with a permission.

1.3. Thesis outline

As we have seen, certain traits, such as self-compatibility and ploidy, can give an advantage to a species during colonisation by countering a limited number of mates and increasing the genetic variation, respectively. In addition, both intra- and interspecific hybridisation can increase genetic variation and introduce new allele combinations to the introduced populations. This study is combining these demo-genetic traits and processes and will shed light on whether they can contribute to invasiveness in *Rumex*. A comparison between native and introduced populations was made to help assess the origin of the invasions and to establish a baseline for comparison. Studying more than one closely related species and several populations will allow a better understanding of the underlying processes behind invasiveness.

In the first two data chapters I studied a range of demo-genetic traits and processes, namely ploidy, genome size, and the breeding system of the study species within and between the native and introduced regions. The following two data chapters focused more on invasion genetics, and I assessed the hybridisation in NZ as well as compared the genetic variation of the study species within their introduced and native ranges. The last data chapter focused on the methodology chosen for the sequencing, and I assessed what is commonly done within invasion genetics and compared the two approaches in *Rumex*. Finally, in Chapter 7 I synthesised the findings of this thesis and link the results back to the theory presented in this introduction, as well as assessed the broader interplay between the demo-genetic traits and processes (FIGURE 1.1).

1.3.1. Chapter 2: Are there geo-cytotypes in New Zealand which could explain the higher adaptive plasticity in New Zealand compared to the UK?

I assessed the genome sizes of plants from NZ compared to the plants from the UK. Invasion biology shows that often species that have multiple geo-cytotypes within the native range, have mainly higher chromosome numbers within the introduced range (te Beest et al., 2012). Higher ploidies are also linked to increased phenotypic plasticity (Monty et al., 2010; te Beest et al., 2012; Hahn et al., 2012), and as *R. obtusifolius* is known to show higher plasticity in the introduced range (Bufford & Hulme, 2021a), I wanted to know whether geo-cytotypes could explain this. To assess the variation in ploidy, I conducted a literature search to estimate the geo-cytotypes of my three study species and how the results are split between native and introduced regions. In addition, I calculated the genome size of 50 plants per species to assess whether my material shows any indications of geo-cytotypes. Chromosome counts were done to match genome sizes to a ploidy level, and flow cytometry was used to estimate the genome sizes. These findings provided a basis for Chapters 3 and 4 and provided information on the geo-cytotypes within the three *Rumex* species, both in literature and from the collected data.

The cytology and chromosome count work were done as a collaboration with Manaaki Whenua – Landcare Research, where Dr. Gary Houliston and Caroline Mitchell

shared their equipment and expertise and taught me to conduct the flow cytometric experiments, and Murray Dawson conducted the chromosome counts.

1.3.2. Chapter 3: Is *Rumex conglomeratus* self-compatible and are the costs from selfing different between the native and introduced regions?

In this chapter, I investigated the reproductive ecology of my study species and used pollen:ovule ratios (Cruden, 1977) as a proxy of the breeding system of *R. conglomeratus*. I also studied the effects of selfing and out-crossing on seed weight and germination. The reproductive ecology has important implications from an invasion ecology standpoint, and it can tell us how a species overcomes the barriers in establishment and subsequent spread (Barrett et al., 2008; Ward et al., 2008; Barrett, 2013; Razanajatovo et al., 2016). Information about the reproductive method could also be used to understand the genetics of the species better and help to set up expectations for Chapter 5. The original plan was to investigate all three species; however, flowering was not induced but rather plants were sampled opportunistically while flowering in the greenhouse. All three species have successfully flowered in the greenhouse in the past, but at the time of sampling, mainly *R. conglomeratus* flowered while the other two species had to be dropped out from the experiment due to low numbers. In addition, during the flowering season NZ entered a lockdown due to COVID-19, and I lost access to the greenhouse. Thus, the originally planned comparisons between ploidy levels and generalisations based on broader patterns within the genus could not be made. However, assessing the breeding system as well as costs from selfing in *R. conglomeratus* provides an important first look into the three species.

I used a haemocytometer to count the pollen grains, and the methodology followed Dafni (1992) and Cruden (1977). In the selfing experiment, half of the flowering study plants in the greenhouse were bagged and half were un-bagged, and later seeds were collected, weighed, and germinated on Petri dishes to assess the fitness consequences from selfing.

1.3.3. Chapter 4: How does the hybrid distribution differ between the native and introduced regions and is there evidence of introgression in New Zealand?

This chapter investigated the hybridisation of *Rumex* in NZ, as no official records of hybrids exist, but the hybrids are very common in the native range (Williams, 1971; Akeroyd, 2014). Hybrids were identified using morphology and flow cytometry, as the hybrids have an intermediate genome size compared to their parents (Williams, 1971). In addition, a literature search was conducted to compare the density of *Rumex* populations within NZ and the UK. This was done using vegetation survey records from both countries.

The flow cytometry work was done as a collaboration with Gary Houliston and Caroline Mitchell from Manaaki Whenua – Landcare Research. The vegetation surveys were a courtesy of NZ National Vegetation Survey Databank (Manaaki Whenua -

Landcare Research, n.d.) and the Botanical Society of Britain and Ireland (Botanical Society of Britain & Ireland, n.d.).

1.3.4. Chapter 5: What is the origin of the introductions to New Zealand and were the species introduced multiple times?

I assessed the genetic variation in my study species and compared the differentiation within and between the native and introduced regions. Genetic information is important for understanding the invasion history and invasion ecology – such as bottlenecks – of the species in question (Ward et al., 2008; Neinavaie et al., 2021; Sherpa & Després, 2021). In addition, genetic variation is generally thought to be a prerequisite for invasion success, so assessing the variation is important for understanding the species. Comparing the amount of variation within the introduced range to that of the native range and to that of the congenerics can help us see patterns that might be used in extrapolating characteristics to other invasive species (Bossdorf et al., 2005; Sherpa & Després, 2021).

This work was done as a collaboration with AgResearch: Jeanne Jacobs served as a point of contact, Craig Anderson, Won Hong, and Anna Larking guided and contracted the DNA extractions and library making, whereas Ken Dodds and Rudiger Brauning conducted bioinformatics and guided in population genetics analyses.

1.3.5. Chapter 6: How often are reference genomes used in the field of biological invasions, and how successful is the *de novo* approach in comparison in *Rumex*?

This chapter studied commonly used methodology in next-generation-sequencing and specifically, restriction site associated sequencing (RSS) methods. I highlighted the similarities and differences between *de novo* and reference-based analyses and looked for common questions associated with each method within the invasion literature. Lastly, a comparison of the approach suitability in *Rumex* was conducted, and success assessed using metrics commonly found within the literature. This work was based on the work done in Chapter 5 as well as a literature search.

1.4. References

Ainouche, M. L., Fortune, P. M., Salmon, A., Parisod, C., Grandbastien, M.-A., Fukunaga, K., Ricou, M., & Misset, M.-T. (2009). Hybridization, polyploidy and invasion: Lessons from Spartina (Poaceae). Biological Invasions, 11(5), 1159–1173. https://doi.org/10.1007/s10530-008-9383-2

Akeroyd, J. R. (2014). Docks and knotweeds of Britain and Ireland (2nd ed.). Botanical Society of Britain & Ireland.

Allan, H. H. (1937). The origin and distribution of the naturalized plants of New Zealand. Linnean Society of London, 150(1), 25–46.

Allendorf, F. W., & Lundquist, L. L. (2003). Introduction: Population biology, evolution, and control of invasive species. Conservation Biology, 17(1), 24–30. https://doi.org/10.1046/j.1523-1739.2003.02365.x

Altshuler, D., Pollara, V. J., Cowles, C. R., Van Etten, W. J., Baldwin, J., Linton, L., & Lander, E. S. (2000). An SNP map of the human genome generated by reduced representation shotgun sequencing. Nature, 407(6803), 513–516. https://doi.org/10.1038/35035083

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. Nature Reviews Genetics, 17(2), 81–92. https://doi.org/10.1038/nrg.2015.28

Attard, C. R. M., Beheregaray, L. B., & Möller, L. M. (2018). Genotyping-by-sequencing for estimating relatedness in nonmodel organisms: Avoiding the trap of precise bias. Molecular Ecology Resources, 18(3), 381–390. https://doi.org/10.1111/1755-0998.12739

Baker, H. G. (1955). Self-compatibility and establishment after "long-distance" dispersal. Evolution, 9(3), 347–349. https://doi.org/10.2307/2405656

Barker, B. S., Cocio, J. E., Anderson, S. R., Braasch, J. E., Cang, F. A., Gillette, H. D., & Dlugosch, K. M. (2018). Potential limits to the benefits of admixture during biological invasion. Molecular Ecology, 28(1), 100–113. https://doi.org/10.1111/mec.14958

Barrett, S. C. H. (2013). IV.8. evolution of mating systems: Outcrossing versus selfing. In J. B. Losos, D. A. Baum, D. J. Futuyma, H. E. Hoekstra, R. E. Lenski, A. J. Moore, C. L. Peichel, D. Schluter, & M. C. Whitlock (Eds.), The Princeton guide to evolution (pp. 356–362). Princeton University Press. https://doi.org/10.1515/9781400848065-050

Barrett, S. C. H. (2015). Foundations of invasion genetics: The Baker and Stebbins legacy. Molecular Ecology, 24(9), 1927–1941. https://doi.org/10.1111/mec.13014

Barrett, S. C. H., Colautti, R. I., & Eckert, C. G. (2008). Plant reproductive systems and evolution during biological invasion. Molecular Ecology, 17(1), 373–383. https://doi.org/10.1111/j.1365-294X.2007.03503.x

Barringer, B. C. (2007). Polyploidy and self-fertilization in flowering plants. American Journal of Botany, 94(9), 1527–1533. https://doi.org/10.3732/ajb.94.9.1527

Bar-Zvi, D., Lupo, O., Levy, A. A., & Barkai, N. (2017). Hybrid vigor: The best of both parents, or a genomic clash? Current Opinion in Systems Biology, 6, 22–27. https://doi.org/10.1016/j.coisb.2017.08.004

Bennett, M. D., Leitch, I. J., & Hanson, L. (1998). DNA amounts in two samples of angiosperm weeds. Annals of Botany, 82(suppl_1), 121–134. https://doi.org/10.1006/anbo.1998.0785

Blossey, B., & Notzold, R. (1995). Evolution of increased competitive ability in invasive nonindigenous plants: A hypothesis. Journal of Ecology, 83(5), 887–889. https://doi.org/10.2307/2261425

Bock, D. G., Caseys, C., Cousens, R. D., Hahn, M. A., Heredia, S. M., Hübner, S., Turner, K. G., Whitney, K. D., & Rieseberg, L. H. (2015). What we still don't know about invasion genetics. Molecular Ecology, 24(9), 2277–2297. https://doi.org/10.1111/mec.13032

Boland, J. M. (2017). Linking seedling spatial patterns to seed dispersal processes in an intermittent stream. Madroño, 64(2), 61–70.

Bossdorf, O., Auge, H., Lafuma, L., Rogers, W. E., Siemann, E., & Prati, D. (2005). Phenotypic and genetic differentiation between native and introduced plant populations. Oecologia, 144(1), 1–11. https://doi.org/10.1007/s00442-005-0070-z

Botanical Society of Britain & Ireland. (n.d.). The BSBI Distribution Maps Scheme. https://bsbi.org/maps-and-data

Briggs, D., & Walters, S. M. (2016). Plant variation and evolution (4th ed.). Cambridge University Press.

Bufford, J. L., & Hulme, P. E. (2021a). Increased adaptive phenotypic plasticity in the introduced range in alien weeds under drought and flooding. Biological Invasions, 23(8), 2675–2688. https://doi.org/10.1007/s10530-021-02532-5

Bufford, J. L., & Hulme, P. E. (2021b). Seed size-number trade-offs are absent in the introduced range for three congeneric plant invaders. Journal of Ecology, 109, 3849–3860. https://doi.org/10.1111/1365-2745.13761

Bzdega, K., Janiak, A., Książczyk, T., Lewandowska, A., Gancarek, M., Sliwinska, E., & Tokarska-Guzik, B. (2016). A survey of genetic variation and genome evolution within the invasive fallopia complex. PLoS ONE, 11(8), Article e0161854. https://doi.org/10.1371/journal.pone.0161854

Carlin, T. F. (2021). Assessing range limits and niche shifts in invasive weeds [Doctoral Thesis, Lincoln University]. Research@Lincoln. https://hdl.handle.net/10182/14274

Carlin, T. F., Bufford, J. L., Hulme, P. E., & Godsoe, W. K. (2022). Global assessment of three Rumex species reveals inconsistent climatic niche shifts across introduced ranges. Research Square. https://doi.org/10.21203/rs.3.rs-1358439/v1

Catford, J. A., Jansson, R., & Nilsson, C. (2009). Reducing redundancy in invasion ecology by integrating hypotheses into a single theoretical framework. Diversity and Distributions, 15(1), 22–40. https://doi.org/10.1111/j.1472-4642.2008.00521.x

Cavers, P. B., & Harper, J. L. (1964). *Rumex obtusifolius* L. and *R. crispus* L. Journal of Ecology, 52(325), 737–766.

Cavers, P. B., & Harper, J. L. (1966). Germination polymorphism in *Rumex crispus* and *Rumex obtusifolius*. Journal of Ecology, 54(2), 367–382. https://doi.org/10.2307/2257955

Charlesworth, D., & Willis, J. H. (2009). The genetics of inbreeding depression. Nature Reviews. Genetics, 10(11), 783–796. https://doi.org/10.1038/nrg2664

Chen, G.-Q., Guo, S.-L., & Yin, L.-P. (2010). Applying DNA C-values to evaluate invasiveness of angiosperms: Validity and limitation. Biological Invasions, 12(5), 1335–1348. https://doi.org/10.1007/s10530-009-9550-0

Colautti, R. I., Grigorovich, I. A., & MacIsaac, H. J. (2006). Propagule pressure: A null model for biological invasions. Biological Invasions, 8(5), 1023–1037. https://doi.org/10.1007/s10530-005-3735-y

Colautti, R. I., Manca, M., Viljanen, M., Ketelaars, H. a. M., Bürgi, H., Macisaac, H. J., & Heath, D. D. (2005). Invasion genetics of the Eurasian spiny waterflea: Evidence for bottlenecks and gene flow using microsatellites. Molecular Ecology, 14(7), 1869–1879. https://doi.org/10.1111/j.1365-294X.2005.02565.x

Colenso, W. (1844). Excursion in the Northern Island of New Zealand. http://api.digitalnz.org/records/37333369/source

Costan, C.-A., Godsoe, W. K., Bufford, J. L., Marris, J. W. M., & Hulme, P. E. (2022). Can the enemy release hypothesis explain the success of *Rumex* (Polygonaceae) species in an introduced range? Biological Invasions. https://doi.org/10.1007/s10530-022-02810-w

Cruden, R. W. (1977). Pollen-ovule ratios: A conservative indicator of breeding systems in flowering plants. Evolution, 31(1), 32–46. https://doi.org/10.2307/2407542

Dafni, A. (1992). Pollination ecology: A practical approach. IRL Press; Oxford University Press.

Darwin, C. (1839). Narrative of the surveying voyages of His Majesty's Ships Adventure and Beagle between the years 1826 and 1836, describing their examination of the southern shores of South America, and the Beagle's circumnavigation of the globe (Vol. 3). Henry Colburn.

Dawson, M. I. (1993). Contributions to a chromosome atlas of the New Zealand flora—31 *Clematis* (Ranunculaceae). New Zealand Journal of Botany, 31(1), 91–96. https://doi.org/10.1080/0028825X.1993.10419536

Doležel, J., Greilhuber, J., & Suda, J. (2007). Estimation of nuclear DNA content in plants using flow cytometry. Nature Protocols, 2(9), 2233–2244. http://dx.doi.org/10.1038/nprot.2007.310

Ellstrand, N. C. (2009). Evolution of invasiveness in plants following hybridization. Biological Invasions, 11(5), 1089–1091. https://doi.org/10.1007/s10530-008-9389-9

Ellstrand, N. C., & Schierenbeck, K. A. (2006). Hybridization as a stimulus for the evolution of invasiveness in plants? Euphytica, 148, 35–46. https://doi.org/10.1007/s10681-006-5939-3

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high

diversity species. PLoS ONE, 6(5), Article e19379. https://doi.org/10.1371/journal.pone.0019379
Elton, C. S. (1958). The ecology of invasions by animals and plants (1st ed.). Methuen & Co.
Estoup, A., Ravigné, V., Hufbauer, R., Vitalis, R., Gautier, M., & Facon, B. (2016). Is there a genetic paradox of biological invasion? Annual Review of Ecology, Evolution, and Systematics, 47, 51–72. https://doi.org/10.1146/annurev-ecolsys-121415
Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics, 131(2), 479–491. https://doi.org/10.1093/genetics/131.2.479
Fox, C. W., & Reed, D. H. (2011). Inbreeding depression increases with environmental stress: An experimental study and meta-analysis. Evolution, 65(1), 246–258. https://doi.org/10.1111/j.1558-5646.2010.01108.x
Frankham, R. (2005). Resolving the genetic paradox in invasive species. Heredity, 94(4), Article 385. https://doi.org/10.1038/sj.hdy.6800634
Friedman, J., & Barrett, S. C. H. (2009). Wind of change: New insights on the ecology and evolution of pollination and mating in wind-pollinated plants. Annals of Botany, 103(9), 1515–1527. https://doi.org/10.1093/aob/mcp035
Friedman, J., & Barrett, S. C. H. (2011). Genetic and environmental control of temporal and size-dependent sex allocation in a wind-pollinated plant. Evolution, 65(7), 2061–2074. https://doi.org/10.1111/j.1558-5646.2011.01284.x
Greilhuber, J., Doležel, J., Lysák, M. A., & Bennett, M. D. (2005). The origin, evolution and proposed stabilization of the terms “genome size” and “C-value” to describe nuclear DNA contents. Annals of Botany, 95(1), 255–260. https://doi.org/10.1093/aob/mci019
Grime, J. P., Hodgson, J. G., & Hunt, R. (2007). Comparative plant ecology: A functional approach to common British species (2nd ed.). Castlepoint Press.
Grover, A., & Sharma, P. C. (2016). Development and use of molecular markers: Past and present. Critical Reviews in Biotechnology, 36(2), 290–302. https://doi.org/10.3109/07388551.2014.959891
Hahn, M. A., van Kleunen, M., & Müller-Schärer, H. (2012). Increased phenotypic plasticity to climate may have boosted the invasion success of polyploid *Centaurea stoebe*. PLoS ONE, 7(11), Article e50284. https://doi.org/10.1371/journal.pone.0050284
Hamrick, J. L., & Godt, M. J. W. (1996). Effects of life history traits on genetic diversity in plant species. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 351(1345), 1291–1298. https://doi.org/10.1098/rstb.1996.0112
Harper, J. L., & Chancellor, A. P. (1959). The comparative biology of closely related species living in the same area: IV. *Rumex*: Interference between individuals in populations of one and two species. Journal of Ecology, 47(3), 679–695. https://doi.org/10.2307/2257298
Harrington, K. C. (2019). Weed management in New Zealand pastures. Agronomy, 9(8), Article 448. https://doi.org/10.3390/agronomy9080448
Hollingswort, M. L., & Bailey, J. P. (2000). Evidence for massive clonal growth in the invasive weed *Fallopia japonica* (Japanese Knotweed). Botanical Journal of the Linnean Society, 133(4), 463–472. https://doi.org/10.1111/j.1095-8339.2000.tb01589.x
Holm, L. G. (1977). The world’s worst weeds: Distribution and biology. University Press of Hawaii.
Ichikawa, S., Sparrow, A. H., Frankton, C., Nauman, A. F., Smith, E. B., & Pond, V. (1971). Chromosome number, volume and nuclear volume relationships in a polyploid series (2x-20x) of the genus *Rumex*. Canadian Journal of Genetics and Cytology, 13, 842–863. https://doi.org/10.1139/g71-119
Invasive Species Specialist. (n.d.). Global Invasive Species Database. Retrieved May 6, 2022, from http://www.iucngisd.org/gisd/
Keane, R. M., & Crawley, M. J. (2002). Exotic plant invasions and the enemy release hypothesis. Trends in Ecology & Evolution, 17(4), 164–170. https://doi.org/10.1016/S0169-5347(02)02499-0

Kim, S. Y., Choi, H. W., Koo, D. H., Lee, W. K., Lee, J., & Bang, J. W. (2006). Characterization of eight *Rumex* species by FISH (fluorescence in situ hybridization) and 5S rDNA spacer sequences. Korean Journal of Genetics, 28(3), 243–251.

Kolar, C. S., & Lodge, D. M. (2001). Progress in invasion biology: Predicting invaders. Trends in Ecology & Evolution, 16(4), 199–204. https://doi.org/10.1016/S0169-5347(01)02101-2

Kołodziejek, J. (2019). Growth and competitive interaction between seedlings of an invasive *Rumex confertus* and of co-occurring two native *Rumex* species in relation to nutrient availability. Scientific Reports, 9(1), Article 3298. https://doi.org/10.1038/s41598-019-39947-z

Kubešová, M., Moravcová, L., Suda, J., Jarošík, V., & Pyšek, P. (2010). Naturalized plants have smaller genomes than their non-invading relatives: A flow cytometric analysis of the Czech alien flora. Preslia, 82(1), 81–96.

Le Roux, J. J. (2022). The Evolutionary Ecology of Invasive Species. Academic Press.

Leach, H. (2005). Gardens without weeds? Pre-European Māori gardens and inadvertent introductions. New Zealand Journal of Botany, 43(1), 271–284. https://doi.org/10.1080/0028825X.2005.9512954

Leitch, I. J., & Bennett, M. D. (2004). Genome downsizing in polyploid plants. Biological Journal of the Linnean Society, 82(4), 651–663. https://doi.org/10.1111/j.1095-8312.2004.00349.x

Lemopoulos, A., Prokkola, J. M., Uusi-Heikkilä, S., Vasemägi, A., Huusko, A., Hyvärinen, P., Koljonen, M.-L., Koskiniemi, J., & Vainikka, A. (2019). Comparing RADseq and microsatellites for estimating genetic diversity and relatedness—Implications for brown trout conservation. Ecology and Evolution, 9(4), 2106–2120. https://doi.org/10.1002/ece3.4905

Löve, Á., & Kapoor, B. M. (1967). A chromosome atlas of the collective genus *Rumex*. Cytologia, 32, 328–342.

Manaaki Whenua - Landcare Research. (n.d.). New Zealand National Vegetation Survey Databank. Retrieved May 28, 2021, from https://nvs.landcareresearch.co.nz/

Martin, M. D., Olsen, M. T., Samaniego, J. A., Zimmer, E. A., & Gilbert, M. T. P. (2016). The population genomic basis of geographic differentiation in North American common ragweed (*Ambrosia artemisiifolia* L.). Ecology and Evolution, 6(11), 3760–3771. https://doi.org/10.1002/ece3.2143

Michalski, S. G., & Durka, W. (2009). Pollination mode and life form strongly affect the relation between mating system and pollen to ovule ratios. New Phytologist, 183(2), 470–479. https://doi.org/10.1111/j.1469-8137.2009.02861.x

Monty, A., Maurice, S., & Mahy, G. (2010). Phenotypic traits variation among native diploid, native tetraploid and invasive tetraploid *Senecio inaequidens* DC. (Asteraceae). Biotechnologie, Agronomie, Société et Environnement, 14(4), 627–632.

Morton, J. (1989). The clustered dock, *Rumex conglomeratus* (Polygonaceae), in Canada. Canadian Field-Naturalist, 103, 86–88.

Mounger, J., Ainouche, M. L., Bossdorf, O., Cavé-Radet, A., Li, B., Parepa, M., Salmon, A., Yang, J., & Richards, C. L. (2021). Epigenetics and the success of invasive plants. Philosophical Transactions of the Royal Society B: Biological Sciences, 376(1826), Article 20200117. https://doi.org/10.1098/rstb.2020.0117

Moura, R. F., Queiroga, D., Vilela, E., & Moraes, A. P. (2021). Polyploidy and high environmental tolerance increase the invasive success of plants. Journal of Plant Research, 134(1), 105–114. https://doi.org/10.1007/s10265-020-01236-6

Müller-Schärer, H., Schaffner, U., & Steinger, T. (2004). Evolution in invasive plants: Implications for biological control. Trends in Ecology & Evolution, 19(8), 417–422. https://doi.org/10.1016/j.tree.2004.05.010

Nagy, D. U., Stranczinger, S., Godi, A., Weisz, A., Rosche, C., Suda, J., Mariano, M., & Pal, R. W. (2018). Does higher ploidy level increase the risk of invasion? A case study with two geo-cytotypes of *Solidago gigantea* Aiton (Asteraceae). Journal of Plant Ecology, 11(2), 317–327. https://doi.org/10.1093/jpe/rtx005

Navajas-Pérez, R., de la Herrán, R., López González, G., Jamilena, M., Lozano, R., Ruiz Rejón, C., Ruiz Rejón, M., & Garrido-Ramos, M. A. (2005). The evolution of reproductive systems and sex-determining mechanisms within *Rumex* (Polygonaceae) inferred from nuclear and chloroplastidial sequence data. Molecular Biology and Evolution, 22(9), 1929–1939. https://doi.org/10.1093/molbev/msi186

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

Nelson, J. C. (1917). The introduction of foreign weeds in ballast as illustrated by ballast-plants at Linnton, Oregon. Torreya, 17(9), 151–160.

Pandit, M. K., White, S. M., & Pocock, M. J. O. (2014). The contrasting effects of genome size, chromosome number and ploidy level on plant invasiveness: A global analysis. New Phytologist, 203(2), 697–703. https://doi.org/10.1111/nph.12799

Pellicer, J., & Leitch, I. J. (2020). The Plant DNA C-values database (release 7.1): An updated online repository of plant genome size data for comparative studies. New Phytologist, 226(2), 301–305. https://doi.org/10.1111/nph.16261

Poland, J. A., Brown, P. J., Sorrells, M. E., & Jannink, J. L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS ONE, 7(2), Article e32253. https://doi.org/10.1371/journal.pone.0032253

Poland, J. A., & Rife, T. W. (2012). Genotyping-by-sequencing for plant breeding and genetics. The Plant Genome, 5(3), 92–102. https://doi.org/10.3835/plantgenome2012.05.0005

Pyšek, P., Skálová, H., Čuda, J., Guo, W.-Y., Suda, J., Doležal, J., Kauzál, O., Lambertini, C., Lučanová, M., Mandáková, T., Moravcová, L., Pyšková, K., Brix, H., & Meyerson, L. A. (2018). Small genome separates native and invasive populations in an ecologically important cosmopolitan grass. Ecology, 99(1), 79–90. https://doi.org/10.1002/ecy.2068

Razanajatovo, M., Maurel, N., Dawson, W., Essl, F., Kreft, H., Pergl, J., Pyšek, P., Weigelt, P., Winter, M., & van Kleunen, M. (2016). Plants capable of selfing are more likely to become naturalized. Nature Communications, 7, Article 13313. https://doi.org/10.1038/ncomms13313

Reddy, K., Rolston, M. P., & Archie, W. J. (1998). *Phalaris* seed production: Control of *Rumex* and herbicide tolerance. Arable Crops, 51, 80–83. https://doi.org/10.30843/nzpp.1998.51.11654

Rice, A., Glick, L., Abadi, S., Einhorn, M., Kopelman, N. M., Salman-Minkov, A., Mayzel, J., Chay, O., & Mayrose, I. (2015). The Chromosome Counts Database (CCDB) – a community resource of plant chromosome numbers. New Phytologist, 206(1), 19–26. https://doi.org/10.1111/nph.13191

Richardson, D. M., & Pyšek, P. (2008). Editorial: Fifty years of invasion ecology: The legacy of Charles Elton. Diversity and Distributions, 14(2), 161–168.

Richardson, D. M., Pyšek, P., Rejmánek, M., Barbour, M. G., Panetta, F. D., & West, C. J. (2000). Naturalization and invasion of alien plants: Concepts and definitions. Diversity and Distributions, 6(2), 93–107. https://doi.org/10.1046/j.1472-4642.2000.00083.x

Rieseberg, L. H., & Carney, S. E. (1998). Plant hybridization. New Phytologist, 140(4), 599–624. https://doi.org/10.1046/j.1469-8137.1998.00315.x

Rubenstein, J. M., Hulme, P. E., Buddenhagen, C. E., Rolston, M. P., & Hampton, J. G. (2021). Weed seed contamination in imported seed lots entering New Zealand. PLOS ONE, 16(8), Article e0256623. https://doi.org/10.1371/journal.pone.0256623

Scheben, A., Batley, J., & Edwards, D. (2017). Genotyping-by-sequencing approaches to characterize crop genomes: Choosing the right tool for the right application. Plant Biotechnology Journal, 15(2), 149–161. https://doi.org/10.1111/pbi.12645

Schierenbeck, K. A. (2017). Population-level genetic variation and climate change in a biodiversity hotspot. Annals of Botany, 119(2), 215–228. https://doi.org/10.1093/aob/mcw214

Seebens, H., Blackburn, T. M., Dyer, E. E., Genovesi, P., Hulme, P. E., Jeschke, J. M., Pagad, S., Pyšek, P., Winter, M., Arianoutsou, M., Bacher, S., Blasius, B., Brundu, G., Capinha, C., Celesti-Grapow, L., Dawson, W., Dullinger, S., Fuentes, N., Jäger, H., … & Essl, F. (2017). No saturation in the accumulation of alien species worldwide. Nature Communications, 8(1), Article 14435. https://doi.org/10.1038/ncomms14435

Sherpa, S., & Després, L. (2021). The evolutionary dynamics of biological invasions: A multi-approach perspective. Evolutionary Applications, 14(6), 1463–1484. https://doi.org/10.1111/eva.13215

Sliwinska, E. (2018). Flow cytometry – a modern method for exploring genome size and nuclear DNA synthesis in horticultural and medicinal plant species. Folia Horticulturae, 30(1), 103–128. https://doi.org/10.2478/fhort-2018-0011

Soltis, P. S., & Soltis, D. E. (2000). The role of genetic and genomic attributes in the success of polyploids. Proceedings of the National Academy of Sciences, 97(13), 7051–7057. https://doi.org/10.1073/pnas.97.13.7051

Stace, C. A. (1991). New flora of the British Isles. Cambridge University Press.

Stewart, A. V., Goldson, S. L., Marris, J. W., Rubenstein, J. M., & Rolston, M. P. (2022). Analytical purity of old New Zealand forage seed samples and detection of fungal and insect contaminants including ryegrass endophyte and Argentine stem weevil. New Zealand Journal of Agricultural Research. https://doi.org/10.1080/00288233.2022.2066140

Suda, J., Krahulcová, A., Trávníek, P., & Krahulec, F. (2006). Ploidy level versus DNA ploidy level: An appeal for consistent terminology. TAXON, 55(2), 447–450. https://doi.org/10.2307/25065591

Suda, J., Meyerson, L. A., Leitch, I. J., & Pyšek, P. (2015). The hidden side of plant invasions: The role of genome size. New Phytologist, 205(3), 994–1007. https://doi.org/10.1111/nph.13107

Takahashi, K., & Hanyu, M. (2015). Hybridization between alien species *Rumex obtusifolius* and closely related native vulnerable species *R. longifolius* in a mountain tourist destination. Scientific Reports, 5, Article 13898. https://doi.org/10.1038/srep13898

te Beest, M., Le Roux, J. J., Richardson, D. M., Brysting, A. K., Suda, J., Kubešová, M., & Pyšek, P. (2012). The more the better? The role of polyploidy in facilitating plant invasions. Annals of Botany, 109(1), 19–45. https://doi.org/10.1093/aob/mcr277

The Noxious Weed Act 1908. http://www.enzs.auckland.ac.nz/docs/1908/1908C133.pdf

Thomson, G. M. (1922). The naturalisation of animals & plants in New Zealand. Cambridge University Press.

van Boheemen, L. A., Lombaert, E., Nurkowski, K. A., Gauffre, B., Rieseberg, L. H., & Hodgins, K. A. (2017). Multiple introductions, admixture and bridgehead invasion characterize the introduction history of *Ambrosia artemisiifolia* in Europe and Australia. Molecular Ecology, 26(20), 5421–5434. https://doi.org/10.1111/mec.14293

van Kleunen, M., Bossdorf, O., & Dawson, W. (2018). The ecology and evolution of alien plants. Annual Review of Ecology, Evolution, and Systematics, 49(1), 25–47. https://doi.org/10.1146/annurev-ecolsys-110617-062654

van Kleunen, M., Dawson, W., & Maurel, N. (2015). Characteristics of successful alien plants. Molecular Ecology, 24(9), 1954–1968. https://doi.org/10.1111/mec.13013

van Kleunen, M., Röckle, M., & Stift, M. (2015). Admixture between native and invasive populations may increase invasiveness of *Mimulus guttatus*. Proceedings of the Royal Society B: Biological Sciences, 282(1815), Article 20151487. https://doi.org/10.1098/rspb.2015.1487

Ward, S. M., Gaskin, J. F., & Wilson, L. M. (2008). Ecological genetics of plant invasion: What do we know? Invasive Plant Science and Management, 1(1), 98–109. https://doi.org/10.1614/IPSM-07-022.1

Webb, C. J., Sykes, W. R., & Garnock-Jones, P. J. (1988). Flora of New Zealand, Volume IV: Naturalised pteridophytes, gymnosperms, dicotyledons. Botany Division, Department of Scientific and Industrial Research.

Williams, J. T. (1971). Seed polymorphism and germination. II. The role of hybridization in the germination polymorphism of *Rumex crispus* and *R. obtusifolius*. Weed Research, 11(1), 12–21. https://doi.org/10.1111/j.1365-3180.1971.tb00970.x

Windham, M. D., Pryer, K. M., Poindexter, D. B., Li, F.-W., Rothfels, C. J., & Beck, J. B. (2020). A step-by-step protocol for meiotic chromosome counts in flowering plants: A powerful and economical technique revisited. Applications in Plant Sciences, 8(4), Article e11342. https://doi.org/10.1002/aps3.11342

Wright, S. (1943). Isolation by distance. Genetics, 28(2), 114–138.

Yan, X. H., Zhou, B., Yin, Z. F., Wang, N., & Zhang, Z. G. (2016). Reproductive biological characteristics potentially contributed to invasiveness in an alien invasive plant Bidens frondosa. Plant Species Biology, 31(2), 107–116. https://doi.org/10.1111/1442-1984.12092

Zaller, J. G. (2004). Ecology and non-chemical control of *Rumex crispus* and *R. obtusifolius* (Polygonaceae): A review. Weed Research, 44, 414–432. https://doi.org/10.1111/j.1365-3180.2004.00416.x

Ziburski, A., Kadereit, J. W., & Leins, P. (1986). Quantitative aspects of hybridization in mixed populations of *Rumex obtusifolius* and *Rumex crispus* Polygonaceae. Flora, 178, 233–242.

Chapter 2

Absence of geo-cytotype variation in three *Rumex* species within the UK and New Zealand

2.1. Introduction

According to theories on plant invasions, often if a species has multiple geo-cytotypes (multiple chromosomal numbers or ploidies) in its native range after introduction the individuals with higher ploidies are better at establishing in the introduced ranges (te Beest et al., 2012). There are multiple examples in the literature of this happening (Treier et al., 2009; Schlaepfer et al., 2010; Rosche et al., 2018). For example, purple loosestrife *Lythrum salicaria* L. (Lythraceae) has four different cytotypes (2x, 3x, 4x, 6x) present within the native range, whereas only one (4x) is found within the introduced range in North America (Kubátová et al., 2008). In addition, research shows that polyploids are highly efficient colonisers (Pandit et al., 2006; te Beest et al., 2012; Pyšek et al., 2018), which might explain their advantage over diploid conspecifics. Nevertheless, Rutland et al. (2021) point out that there is simply not enough evidence to conclusively say that polyploids are weedier than their diploid congeners. In addition, a link between polyploidisation and increased phenotypic plasticity exists (Raycheva, 2005; Bomblies & Madlung, 2014; Nagy et al., 2018) but few studies investigate geo-cytotypes in tandem with other aspects behind invasiveness. These comparisons could, however, help us to better understand biological invasions, which is paramount with the number of invasive species only increasing (Seebens et al., 2017; Seebens, Bacher, et al., 2021; Seebens, Blackburn, et al., 2021).

For years, scientists have debated over the advantages of polyploidy in plants. It was originally – and sometimes still is – called an evolutionary dead end (Wagner, 1970; Stebbins, 1971; Mayrose et al., 2011). The arguments claiming that polyploids are an evolutionary dead end, mainly refer to the sterile hybrids as a waste of reproductive resources. Mayrose et al. (2011, 2015) argue that when diversification rates are compared between related diploid and polyploid lineages, polyploid lineages are slower to form new species and have higher extinction rates than their diploid counterparts. Nevertheless, polyploidy is a widespread phenomenon, especially in plants (e.g., Wood et al., 2009).

Phenotypic plasticity and increased genetic diversity are often caused by polyploidisation (Raycheva, 2005; Bomblies & Madlung, 2014; Nagy et al., 2018). Phenotypic plasticity means the ability of an individual to change its response based on the changes in the environment, or one genotype leading to multiple phenotypes depending on the abiotic cues (Travis, 1994). Phenotypic plasticity can be especially useful in new environments (reviewed in Ghalambor et al., 2007). Indeed, similarly to higher ploidies, phenotypic plasticity is expected to be of importance in invasive species (Colautti et al., 2017). In invasive populations of the spotted knapweed *Centaurea stoebe* L. (Asteraceae), the tetraploid individuals show broader plasticity responses to climate variables compared to the diploids (Hahn et al., 2012). In general, higher genetic variation is considered a great aid for establishment of introduced species (Allendorf & Lundquist, 2003; Bossdorf et al., 2005; Schrieber & Lachmuth, 2016; van Kleunen et al., 2018), and

the process of introduction can lead to genetic bottlenecks reducing the amount of variation present in the new range. Polyploidy could offer an important countermeasure against the decrease in genetic variation by providing extra copies of the genome (Soltis & Soltis, 2000; Schrieber & Lachmuth, 2016; Van Drunen & Johnson, 2022).

The dock weed species *Rumex conglomeratus* Murr., *R. obtusifolius* L., and *R. crispus* L. (Polygonaceae) are among the world's worst weeds (Holm, 1977). They are found in all continents apart from Antarctica and have become a nuisance in agricultural settings. The two latter mentioned species are intensely studied weeds, and thus their ecology is relatively well known (Zaller, 2004). These species are highly variable in their morphology, have high seed set and are capable of growing in various habitats (Cavers & Harper, 1964; Holm, 1977; Stace, 1991). They are known to hybridise freely, and some of the hybrids are known to cross back to the parent species (Williams, 1971; Takahashi & Hanyu, 2015). We also know that both *R. conglomeratus* and *R. obtusifolius* have adaptive plastic reactions to flood and drought conditions (Bufford & Hulme, 2021a). However, only in the tetraploid *R. obtusifolius* this plastic reaction is greater in plants from the introduced range, compared to plants from the native range.

Despite extensive research into the species, the information of the genetics, genome size or evolution regarding invasiveness is still sparse (Zaller, 2004). Chromosome counts, and genome size measurements of these species have shown that there is variation within the species (i.e., cytotypes) but it seems to be geographically separated to some extent (Rice et al., 2015; Leitch et al., n.d.). Most of these estimates are also from the species native range, mainly Europe, and not much is known about the introduced range, and to my knowledge, there are no published counts of either genome size or chromosome numbers from New Zealand (NZ). These three species were introduced to NZ around mid–1800s (Thomson, 1922; Allan, 1937). They most likely came from the UK (see Chapter 5) where only one cytotype has been reported for each of the species (Rice et al., 2015; Leitch et al., n.d.). However, we know from other species introduced to NZ that there have been introductions from elsewhere in Europe (Darwin, 1839; Trewick et al., 2004), so other cytotypes establishing to NZ cannot be ruled out without further investigation.

Given the evidence of increased adaptive plasticity in NZ (Bufford & Hulme, 2021a) and multiple geo-cytotypes within the native range, this study set out to investigate whether there is a link between increased plasticity and geo-cytotypes. However, within the genus interspecific hybrids are extremely common, and the hybrids are known to backcross back to the parent species (Jensen, 1936; Williams, 1971; Takahashi & Hanyu, 2015). This introgression could make it harder to fully confirm whether individuals with varying genome size and chromosome numbers are in fact geo-cytotypes or hybrids, especially since morphological identification can be difficult as well. We therefore used flow cytometry to estimate the genome size and chromosome numbers (Doležel et al., 2007) in samples collected from both the native and introduced ranges, from the UK and NZ, respectively. In addition, chromosome counts were used to establish a base level to combine genome sizes to chromosome numbers. This study will help to shed light into the puzzle that is the variation in the genome size and ploidy of these species, and how this might differ between the native and introduced regions. Specifically, I asked:

1) Are there geo-cytotypes in NZ which would explain the higher adaptive plasticity?

2) Do the observed cytotypes reflect patterns within the likely historic source of these samples, i.e., the UK?

Using a combination of these methods yields reliable information of the status of the individuals. Based on previous research, we know that the genome size differs between the ploidy levels (Löve & Kapoor, 1967; Raycheva, 2005), providing a great method for assigning my samples to different cytotypes. It was also hypothesised that the hybrids should be distinguishable from the parent species based on this method.

2.2. Materials and methods

2.2.1. Chromosome counts and genome size estimates in the literature

Literature and database searches were made to assess previous counts and information on cytotypes, especially regarding the country of origin of the sample. Data were extracted from the Chromosome Counts Database (CCDB; Rice et al., 2015) and The Plant C-values Database (Leitch et al., n.d.) and pooled for analyses, while all the duplicate records were removed. To establish the sample origin, original articles collated in the databases were tracked and the sampling location recorded. If the original article was not accessed or if the sampling location was not reported, the datapoint was discarded from the analyses. The Global Invasive Species Database (GISD, 2015), the PLANTS database (USDA, NRCS, n.d.), and the Global Register of Introduced and Invasive Species - Japan (GRIIS, Ikeda et al., 2021) were used to determine the native and introduced ranges for each species.

The genome size estimates were obtained from Plant C-values Database and included for comparison. This database classifies the estimates in 'prime' genome size estimates and 'other' estimates. Pellicer & Leitch (2020) define prime estimate within the database as "the most consistent value obtained under best-practice methods". However, as the list of genome size estimates is not comprehensive, a literature search was done to find more counts. Google was used for the search with a query "*Rumex conglomeratus*/*obtusifolius*/*crispus* genome size", with one species at a time. This was done after Scopus returned only a handful of results, and these were not about the genome size of the species in question. Some genome size information is also published in 'grey literature', conference proceedings, older journals, or not in English, and Google has a higher chance of returning these references.

Because of the difficulty to accurately identify *Rumex* species (e.g., Akeroyd, 2014), it is possible that some of the chromosome counts and genome size estimates come from wrongly identified species or a back-crossed hybrid. Since I only accepted datapoints where I was able to access the original study/report, potential errors in the second-hand collections (i.e., the databases) were less likely to manifest here (see Hollingsworth et al., 1998).

2.2.2. Plant material

Seed samples were collected from four regions in NZ and the UK during respective autumns 2016 and 2017 (Bufford & Hulme, 2021a, 2021b; Costan et al., 2022). Seeds from around 25 mother plants per species per four populations in each of the four regions within each country (= seed family) were harvested, and one plant per seed family was kept and grown up in a greenhouse. Harvesting from four separate regions allowed sampling from a range of climates and habitats (Bufford & Hulme, 2021a; Chapter 1.2.3), which provided a more comprehensive understanding of the three study species in NZ and the UK.

Seeds were only collected from plants that were identified morphologically as true species to avoid sampling hybrids. However, the identity of pollen donors was not known, and thus it was expected that some hybrid plants would be included in the seed material. In addition, there is considerable morphological variation within each *Rumex* species (Cavers & Harper, 1964; Stace, 1991; Zaller, 2004) and these species, especially young plants and hybrids, can closely resemble to each other. The most reliable way to distinguish the species and hybrids from each other, are the valves on the fruits (Löve & Kapoor, 1967). There are some differences in the leaf structure as well as sizes, fullness, and angles of the inflorescence (Akeroyd, 2014). However, some of these characteristics are not always present, i.e., the rosette leaves often die after flowering, which makes having an alternative identification method very appealing.

2.2.3. Flow cytometry

There are multiple methods for obtaining the genome size of an organisms (Bennett & Leitch, 1995). Nowadays, the gold standard method is flow cytometry (Doležel et al., 2007; Pellicer & Leitch, 2020). Flow cytometry is a fast and accurate method and allows for analysing a higher number of nuclei a lot faster and with less effort than some older methods, such as Feulgen microdensitometry, which was the most common method until early 1990s (Bennett & Leitch, 1995). For example, Grime & Mowforth (1982) used Feulgen microdensitometry to study the genome sizes, and they analysed 10 nuclei per sample, whereas Doležel et al. (2007) recommend measuring 5000 nuclei using flow cytometry. Sample preparation time using flow cytometry takes only minutes, whereas staining alone can take over an hour in Feulgen microdensitometry. However, Feulgen microdensitometry offers visual selection of the nuclei, as well as long-lasting samples that can be stored (Doležel et al., 2007). In general, these two methods are thought to produce highly comparable results (Michaelson et al., 1991; Bennett & Leitch, 1995).

Flow cytometry was performed to study the genome size and ploidy of these three *Rumex* species. Fresh, mature, and healthy-looking leaf samples from roughly 50 plants per species (25 from each country, 4–7 per region) were harvested on the same day or maximum 24 hours before flow cytometer analyses. After collection, samples were kept at +2 ºC prior to handling. After screening for suitable standards with known genome size, corn (*Zea mays*, genome size 5.33 pg) was found compatible with *R. obtusifolius* and *R conglomeratus*, whereas pea (*Pisum sativum*, genome size 8.80 pg) was used for *R. crispus*. In addition to corn and pea, tomato and radish were trialled, but proved to be

unsuitable as the standard peaks were overlapping with some of the sample peaks, making it difficult to distinguish between the sample and the standard.

For analyses, the method from Doležel et al. (2007) was used with slight modification. A small sample from both the appropriate standard and one of the *Rumex* species was placed in a Petri dish in ice cold isolation buffer, and the samples were quickly chopped, carefully avoiding crushing the cells. Samples were incubated at room temperature for 2 minutes before filtering the homogenate through a 20-µm fine mesh into a small test tube, after which propidium iodide stain (10 µg/ml) was added. Samples were further incubated at room temperature for another 2 minutes. Samples were analysed with CyFlow Space (Partec) until sufficient number of nuclei were analysed and clearly separated fluorescence intensity histograms were formed. Along a linear fluorescent scale at roughly 500/50 nm, fluorescence vs side scatter was plotted, and to help isolate fluorescence data of interest, convex polygons were drawn around the formed scatterplot (see for example Bainard et al., 2012). After that, peaks were gated, mean peak values and cv-% were recorded, and Gauss peak analysis was performed to confirm the gating. Then, genome size was estimated by calculating the ratio between the standard (corn or pea) and *Rumex* sample fluorescence intensity peaks and dividing the known genome size of the standards by that ratio according to Doležel et al. (2007).

Most samples were analysed only once, but in cases where morphological identification of the sample had proven unsuccessful and the sample was paired with an unsuitable standard, the analysis was repeated with a more suitable standard. This was done to keep the methodology consistent within a species.

Differentiation between the countries was analysed using independent samples t-test. Analyses were done using R (version 4.2.1., R Development Core Team, 2008), figures were created using ggplot2 package (Wickham, 2016).

2.2.4. Chromosome counts

Plants chosen for chromosome counts had been used previously for flow cytometric analyses, which helped to identify samples from each species and suspected hybrids. Samples from all three species and suspected hybrids were collected for the chromosome counts, where in total 10 plants were chosen based on their genome size estimates and morphology. Chromosome counts were done by M. Dawson at Manaaki Whenua – Landcare Research Lincoln, following a method by Dawson (1993) and summarised as follows: Freshly growing root tips were harvested in the morning from plants cultivated in the greenhouse. Several root tips were collected per plant and pre-treated immediately in a mixture of cycloheximide and 8-hydroxyquinoline for 5–6 h hours, before fixing them overnight in methanol-chloroform-propionic acid. After rinsing in distilled water and hydrolysing in HCl, root tips were macerated in pectinase and stained with Feulgen reaction before transferring them to a microscope slide. The meristematic portions were excised out of the surrounding tissue, transferred to a slide, mixed with a drop of lactic-acetic-orcein, and squashed under a coverslip. Several slides were made per plant to find enough suitable cells for the chromosome counts. Chromosomes were counted from all pro-metaphase to full metaphase cells under a light microscope (×1000).

2.3. Results

2.3.1. Multiple geo-cytotypes reported within both ranges

After duplicate entries and entries with missing information were discarded, a total of 30, 27, and 44 datapoints were found for *R. conglomeratus*, *R. obtusifolius*, and *R. crispus*, respectively (Appendix 0). Only around 25 % of the datapoints came from the species' introduced range. The observations within the introduced range were from USA and Canada, and from Japan for *R. conglomeratus* and *R. obtusifolius*, whereas for *R. crispus* Japan is considered the species' native range. No records were found from NZ or Australia from either of the databases.

The somatic (2*n*) chromosome numbers reported in the databases found the same chromosome numbers from both ranges for both *R. conglomeratus* and *R. obtusifolius*, whereas *R. crispus* only had one reported value within the introduced range (FIGURE 2.1). The Plant C-values Database does not report any values for *R. conglomeratus*. The most reported 2*n* values were 20, 40, and 60 for *R. conglomeratus*, *R. obtusifolius*, and *R. crispus*, respectively.

The 2C genome size values reported in the database and literature vary from each other (TABLE 2.1). The Kew Plant C-values database (Leitch et al., n.d.) reports 8.80 pg as a prime estimate for *R. crispus*, and 4.40 pg as a non-prime estimate. Similarly, the prime genome size of *R. obtusifolius* is reported as 3.30 pg. Alternative genome size within the database is provided by Grime & Mowforth (1982) as 2.7 pg. Again, the Plant C-values database does not report any values for *R. conglomeratus*.

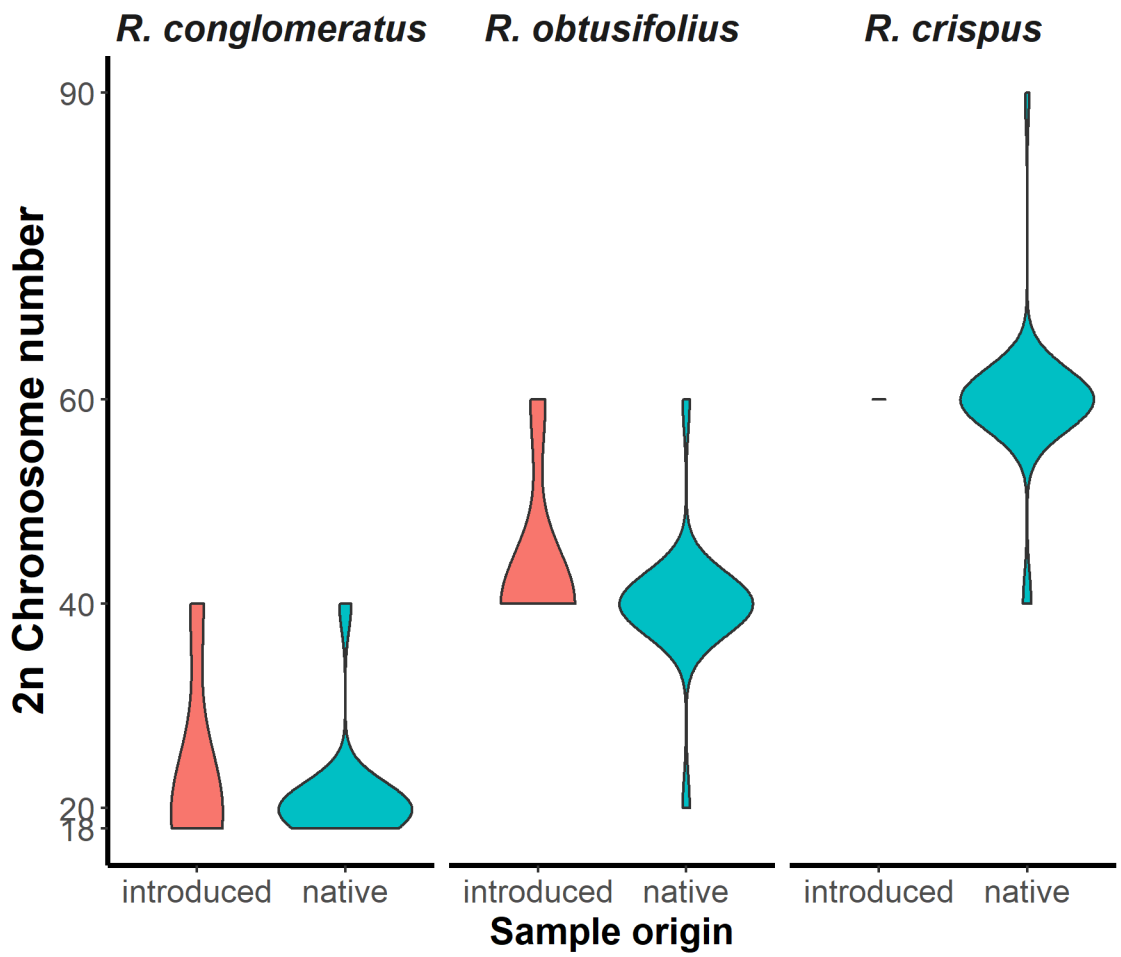

FIGURE 2.1: Comparison of the 2*n* chromosome numbers from CCBD and the Plant C-values database by species and by range. Note, that *Rumex crispus* only has 2*n* = 60 from the introduced range.

TABLE 2.1: The 2C genome size (pg) estimates from the Plant C-value database (denoted with a D superscript) and from the literature for each of the three *Rumex* species. Flow: flow cytometry; Feulgen: Feulgen microdensitometry.

Species	Genome size	Country	Reference	Method
R. conglomeratus				
	1.54	The Netherlands	(Zonneveld, 2019)	Flow
	1.40¥	Czech Republic	(Šmarda et al., 2019)	Flow
R. obtusifolius				
	2.32¥	Czech Republic	(Šmarda et al., 2019)	Flow
	2.64	The Netherlands	(Zonneveld, 2019)	Flow
	2.7	The UK	(Grime & Mowforth, 1982)	Feulgen
	3.0	The UK	(Grime et al., 2007)	Feulgen
	3.1D	Unknown	(Mowforth, 1986 in Leitch et al., n.d.)	Feulgen
R. crispus				
	4.04¥	Czech Republic	(Šmarda et al., 2019)	Flow
	4.44	The Netherlands	(Zonneveld, 2019)	Flow
	4.55D	The USA	(Bai et al., 2012)	Flow
	4.59	Canada	(Bainard et al., 2012)	Flow

	8.8[D]	Unknown	(Conger, 1977 in Leitch et al., n.d.) Feulgen

[¥]Converted from Mbp to pg using 1 pg = 978 Mbp according to Doležel et al. (2003)

2.3.2. No differences in genome sizes between the UK and New Zealand

There was no significant variation in the genome sizes within any of the three species between the UK and NZ (Welch's two sample t-test, equal variances not assumed, $p > 0.25$ in all cases; TABLE 2.2). However, all three species were distinguishable from each other, and the method was able to recognise hybrids (*R.* × *pratensis*) between *R. obtusifolius* and *R. crispus* (see outliers in FIGURE 2.2). Compared to the previous observations in the literature (TABLE 2.1), my results were close to the results obtained using flow cytometry, but not Feulgen microdensitometry.

TABLE 2.2: Parameters from Welch's two sample t-test comparing average genome sizes (pg) between *Rumex* plants from the UK (UK) and New Zealand (NZ). In addition, means of the genome sizes (pg) are given.

Species	t	df	p	Genome size UK	Genome size NZ
R. conglomeratus	-1.1851	44.681	0.2423	1.599	1.612
R. obtusifolius	0.076631	41.561	0.9393	2.675	2.674
R. crispus	-0.73899	36.819	0.3636	4.579	4.602

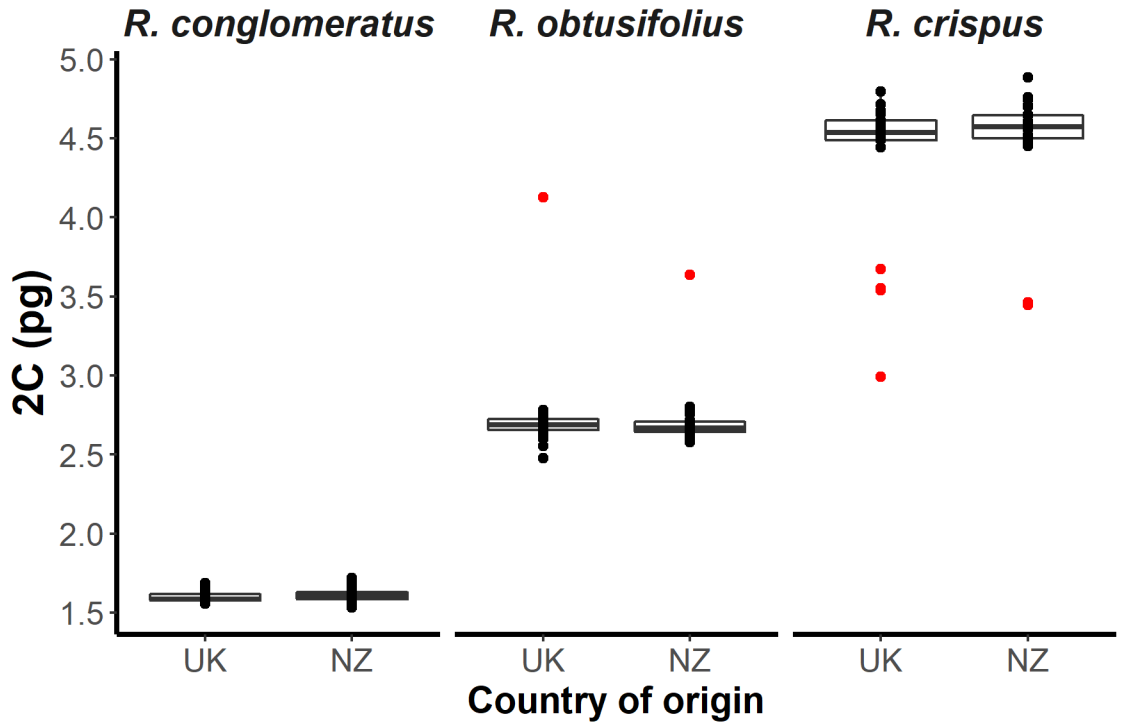

FIGURE 2.2: Box plots of the 2C genome sizes of the three *Rumex* species comparing estimates from the UK and New Zealand (NZ). The outliers (red) between R. obtusifolius and R. crispus are assumed as a hybrid, R. × pratensis.

2.3.3. Chromosome counts disagreed with some morphological identifications

All slide preparations were countable, and fully searched for countable cells. Counts were made from pro-metaphase to full metaphase. The 2*n* = 20 counts were definitive, and the higher ploidies fell into a tight range. Chromosomes were moderately small in size (M. Dawson, personal communication) but chromosome morphology was good especially for the full metaphases. Chromosome counts did not always match with the identification based on plant morphology (TABLE 2.3), especially for the hybrids, which were all assigned incorrectly just based on the plant morphology.

TABLE 2.3: Chromosome counts of 10 *Rumex* spp. samples and their genome sizes (pg) where applicable. Species is then determined based on counts, genome size and plant morphology. Match column tells whether a sample was identified correctly just based on its morphology and its mother plant's morphology.

Sample#	Count (2*n*)	Genome size	Species	Match
1	c. 40 (37–40)	2.78	*R. obtusifolius*	Y
2	60 (56–60)	4.13	*R.* × *pratensis* (× *R. cris*pus)	N
3	c. 50 (47–50)	3.64	*R.* × *pratensis*	N
4	c. 39 (38–39)	N/A	*R. obtusifolius*[¥]	Y[+]
5	60 (57–61)	4.65	*R. crispus*	Y
6	60 (58–60)	4.56	*R. crispus*	Y
7	c. 40 (36–40)	3.50	*R.* × *pratensis* (× *R. obtusifolius*)	N
8	40 (38–40)	2.90	*R.* × *pratensis* (× *R. obtusifolius*)	N
9	20	1.60	*R. conglomeratus*	Y
10	20	1.69	*R. conglomeratus*	Y

[¥]No genome size estimate, species identified based on plant morphology and chromosome count.
[+]Based on mother plant, identification wrong, plant morphology matches the chromosome count.

2.4. Discussion

Polyploidy is known to lead to increased adaptive plasticity (Bomblies & Madlung, 2014; Nagy et al., 2018), and increased plasticity is known to help introduced populations of *R. obtusifolius* in flood and drought conditions compared to the native populations (Bufford & Hulme, 2021a). However, the ploidy and potential geo-cytotypes of *R. conglomeratus*, *R. obtusifolius*, and *R. crispus* in NZ have never been studied before. This provided an opportunity to investigate the potential link of geo-cytotypes and the increase in plasticity. Thus, I compared plants originating from the species' native and introduced ranges, and from four different regions within each range. No geo-cytotypes were found within any of the species, but rather all chromosome counts, and genome sizes pointed to just one cytotype per species, and this was true for both the UK and NZ. In addition, the data did not show support for any differences between the two countries (TABLE 2.2). Flow cytometry was a suitable method for not only identifying the pure species, but also for identifying *R.* × *pratensis*, the hybrid between *R. obtusifolius* and *R. crispus*. In addition, information from chromosome count and genome size databases was collated and analysed and compared to the results from this study.

Hybrid identification in genus *Rumex* is difficult, and sometimes almost impossible, as the plant morphology of the hybrids, especially in later generations, can closely resemble that of their parents (Williams, 1971). Currently, the best way for identifying the hybrids is looking at the mature seed pods (Löve & Kapoor, 1967). However, this is not always feasible depending on the study questions, and often after the seeds mature, the plants do not have any green material left. Unsurprisingly, some of my identifications based on the plant morphology alone were incorrect (TABLE 2.3), and after genome size estimates, the identifications were corrected. In addition, with the assumed back-crossed hybrids even after the genome size estimates, some of the identifications can be challenged. Mostly, however, the hybrid plants were found to have a genome size between the parents (FIGURE 2.2), which makes flow cytometry a good and a fast tool for hybrid screening – especially for the first-generation hybrids – along with the plant morphological estimates.

2.4.1. Most chromosome counts in the databases are from the native range

In contrast to this study, roughly 75 % of the previously published chromosome counts are from the species' native range, and in all three species the counts from the introduced range come from fewer than five countries (Appendix 0). Not one study had compared results between the native and the introduced ranges, and most often reported just one sampling location within the range. Majority of the counts were consistent with the commonly agreed chromosome numbers for each species, and apart from the counts of 20 for *R. obtusifolius* and 40 and 90 for *R. crispus* from just the native range, all other cytotypes were present within both native and introduced ranges (FIGURE 2.1).

The fact that all the cytotypes in *R. conglomeratus* and most in *R. obtusifolius* present in the native range are found within the introduced range when only around 25 % of the counts are from the introduced range, could indicate that the cytotypes have different advantages and disadvantages. It is not uncommon for the cytotypes to have different niches from one another (Sonnleitner et al., 2015; Maguilla et al., 2021). However, since also the lower ploidies were present within the introduced range, we cannot say anything about the superiority of higher ploidies, or that the higher ploidies are contributing to the colonisation abilities of these species. Since most cytotypes were found within both ranges, this likely suggests that selection has not favoured one cytotype over the others and that there is little evidence of bottlenecks. It is also possible that the different cytotypes are an indication of where the introductions originated from and reflect more about the original propagules introduced than advantages for invasions (e.g., Schlaepfer et al., 2008; Dematteis et al., 2020). After all, the initial propagule pressure predicts the success of the invasion better than any other characteristics (reviewed in Levin, 2020).

2.4.2. No geo-cytotypes found in the UK or New Zealand

After each of the genome size estimate paired with a chromosome count, most samples within a species can be associated with just one count. However, eight samples were

found to deviate from this (outliers in FIGURE 2.2; TABLE 2.3). The samples with an intermediate genome size compared to that of *R. obtusifolius* and *R. crispus,* are likely first-generation hybrids, known as *R.* × *pratensis*. The first-generation hybrids are known to have an intermediate genome size and ploidy level compared to the parents (Williams, 1971). As *R.* × *pratensis* is known to introgress back to the parent species thus resembling the respective parent more closely in terms of the genome size and ploidy (Williams, 1971), and as most of the sampled populations would have had more than one species present (J.L. Bufford, personal communication), we can assume that the rest of the outliers are later generation hybrids, instead of geo-cytotypes. In addition, if these samples were geo-cytotypes instead, the number of this second cytotype is small enough to likely not offer any evolutionary advantage or explain the greater plasticity within NZ, as it is so much rarer compared to the main cytotype.

The results in this study agree with the official chromosome counts for these species; 20 for *R. conglomeratus*, 40 for *R. obtusifolius*, and 60 for *R. crispus*. However, the sample sizes were relatively small and by only sampling one country from each region, these results cannot rule out multiple cytotypes elsewhere within the introduced and native regions. For example, in Japan, populations of both *R. conglomeratus* and *R. obtusifolius* exist, where the chromosome counts differ from the typical numbers (Table A.1 and A.2, respectively). In contrast to this, is a more recent study reporting a chromosome number of 40 from *R. obtusifolius* in Korea (Kim et al., 2006), which suggests a potential of multiple geo-cytotypes within Asia. In addition, early introductions from outside the UK to NZ are not unheard of (Darwin, 1839). For example, according to the variation in chloroplast DNA in *Hieracium pilosella,* the samples from NZ had haplotypes that were not present in the UK but rather in Eastern and Northern Europe, suggesting propagules from those areas (Trewick et al., 2004). Indeed, chromosome counts from Russia, Belarus and Sweden show deviations from the most common chromosome counts for the species, as well as do some counts from Germany and Spain (Appendix 0). However, evidence of introductions from regions with differing chromosome counts was not found in this study, and thus the presence of other geo-cytotypes is unlikely. The countries with deviating counts vary between species, and most often there is only one record associated with each count per country.

This study did not assess the accuracy of the methods – or for example how many accessions were used for the count – in the chromosome count studies but given the overlapping morphology of these *Rumex* species and the commonness of the hybrids, there is always a possibility that some of the counts come from different species altogether. Hollingsworth et al. (1998) found several instances of wrongly cited congenerics or wrongly marked chromosome numbers in the chromosome count collections of *Potamogeton* (Potamogetonaceae). These instances seemed to be due to mistakes in the second-hand collections, rather than in the original material itself. Since I verified the species and count listed in the database by referring to the original source, this is not considered an issue in this study. However, if the original material was identified incorrectly, this type of bias would be impossible to identify, especially due to the old age of some of the counts – several counts were made almost a century ago.

2.4.3. My genome size estimates differ from the databases

The reported 'prime' estimates in the Plant C-values database for both *R. obtusifolius* and *R. crispus* differ drastically from my estimates, and from estimates in literature outside the database (TABLE 2.1). As the chromosome counts are the same between this study and the older studies, and the previous genome size estimates were done using older methods (i.e., Feulgen microdensitometry), one can question the accuracy of the older counts in the Plant C-values database. In addition, Pellicer & Leitch (2020) say that flow cytometry is the gold standard in estimating genome size compared to methods such as Feulgen microdensitometry, yet the Plant C-values database still classifies estimates using older methods as 'prime' estimates and estimates using flow cytometry as 'other' estimates. However, as these methods generally produce comparable results (Michaelson et al., 1991; Bennett & Leitch, 1995), these large deviations are likely from geo-cytotypes with higher ploidy and genome size or from hybrids.

The 2C prime estimate for *R. crispus* is 8.80 pg whereas a secondary estimate in the database by Bai et al. (2012) is 4.60 pg, which is more in line with this study (TABLE 2.1). In addition, the second estimate was obtained using the industry gold standard, flow cytometry. The prime estimate of 3.10 pg for *R. obtusifolius* matches with our estimate for the hybrid between *R. obtusifolius* and *R. crispus*, *R.* × *pratensis*, and it is said that the hybrids can be almost indistinguishable from the parents, especially before the seed husks mature (Williams, 1971). No official genome size estimates for *R. conglomeratus* are reported in the databases. However, Zonneveld (2019) reports genome sizes of 1.54 pg, 2.64 pg and 4.44 pg for *R. conglomeratus*, *R. obtusifolius*, and *R. crispus*, respectively, and these match with our findings. In addition, they report an estimate of 3.72 pg for *R.* × *pratensis*, which is also in line with my results (outliers in FIGURE 2.2, TABLE 2.3). Likewise, all of the genome size estimates made using flow cytometry are relatively close to my values (TABLE 2.1).

2.4.4. Conclusions

No geo-cytotypes were found from any of our three study species making it unlikely that differences in ploidy could have contributed to the increased adaptive plasticity of *R. obtusifolius* in NZ but not in the UK (Bufford & Hulme, 2021a). The chromosome counts for each species matched the most often reported counts in the literature: 20 chromosomes in *R. conglomeratus*, 40 chromosomes in *R. obtusifolius*, and 60 chromosomes in *R. crispus*. However, the genome size estimates for *R. obtusifolius* and *R. crispus* differ from what were listed as prime estimates within Plant C-values database but matched with the secondary estimates and other estimates from literature outside the database using flow cytometry.

The historical *Rumex* introductions are assumed to originate from the UK, where only one cytotype has been found. This could explain why no variation in cytotypes was found in NZ. However, as the samples originate only from the South Island, it cannot be confirmed that all the plants in NZ share the same cytotype. Future studies should sample NZ more broadly to rule the presence of multiple cytotypes out.

In addition, this study analysed the genome size of a relatively few samples. However, if other cytotypes were common, our sampling would have likely identified

them despite the relatively low sample numbers, as all the geo-cytotypes were accounted for in relatively few studies from elsewhere within the introduced range. Future research should sample more countries especially within the introduced region, such as Australia and other countries in the Southern Hemisphere, as these countries are not represented in the current databases. Flow cytometry and manual chromosome counts should be utilised to ensure reliable and comparable results.

2.5. References

Akeroyd, J. R. (2014). Docks and knotweeds of Britain and Ireland (2nd ed.). Botanical Society of Britain & Ireland.

Allan, H. H. (1937). The origin and distribution of the naturalized plants of New Zealand. Linnean Society of London, 150(1), 25–46.

Allendorf, F. W., & Lundquist, L. L. (2003). Introduction: Population biology, evolution, and control of invasive species. Conservation Biology, 17(1), 24–30. https://doi.org/10.1046/j.1523-1739.2003.02365.x

Bai, C., Alverson, W. S., Follansbee, A., & Waller, D. M. (2012). New reports of nuclear DNA content for 407 vascular plant taxa from the United States. Annals of Botany, 110(8), 1623–1629. https://doi.org/10.1093/aob/mcs222

Bainard, J. D., Bainard, L. D., Henry, T. A., Fazekas, A. J., & Newmaster, S. G. (2012). A multivariate analysis of variation in genome size and endoreduplication in angiosperms reveals strong phylogenetic signal and association with phenotypic traits. New Phytologist, 196(4), 1240–1250. https://doi.org/10.1111/j.1469-8137.2012.04370.x

Bennett, M. D., & Leitch, I. J. (1995). Nuclear DNA amounts in angiosperms. Annals of Botany, 76(2), 113–176. https://doi.org/10.1006/anbo.1995.1085

Bomblies, K., & Madlung, A. (2014). Polyploidy in the *Arabidopsis* genus. Chromosome Research, 22(2), 117–134. https://doi.org/10.1007/s10577-014-9416-x

Bossdorf, O., Auge, H., Lafuma, L., Rogers, W. E., Siemann, E., & Prati, D. (2005). Phenotypic and genetic differentiation between native and introduced plant populations. Oecologia, 144(1), 1–11. https://doi.org/10.1007/s00442-005-0070-z

Bufford, J. L., & Hulme, P. E. (2021a). Increased adaptive phenotypic plasticity in the introduced range in alien weeds under drought and flooding. Biological Invasions, 23(8), 2675–2688. https://doi.org/10.1007/s10530-021-02532-5

Bufford, J. L., & Hulme, P. E. (2021b). Seed size-number trade-offs are absent in the introduced range for three congeneric plant invaders. Journal of Ecology, 109, 3849–3860. https://doi.org/10.1111/1365-2745.13761

Cavers, P. B., & Harper, J. L. (1964). *Rumex obtusifolius* L. and *R. crispus* L. Journal of Ecology, 52(325), 737–766.

Colautti, R. I., Alexander, J. M., Dlugosch, K. M., Keller, S. R., & Sultan, S. E. (2017). Invasions and extinctions through the looking glass of evolutionary ecology. Philosophical Transactions of the Royal Society B: Biological Sciences, 372(1712), Article 20160031. https://doi.org/10.1098/rstb.2016.0031

Costan, C.-A., Godsoe, W. K., Bufford, J. L., Marris, J. W. M., & Hulme, P. E. (2022). Can the enemy release hypothesis explain the success of *Rumex* (Polygonaceae) species in an introduced range? Biological Invasions. https://doi.org/10.1007/s10530-022-02810-w

Darwin, C. (1839). Narrative of the surveying voyages of His Majesty's Ships Adventure and Beagle between the years 1826 and 1836, describing their examination of the southern shores of South America, and the Beagle's circumnavigation of the globe (Vol. 3). Henry Colburn.

Dawson, M. I. (1993). Contributions to a chromosome atlas of the New Zealand flora—31 *Clematis* (Ranunculaceae). New Zealand Journal of Botany, 31(1), 91–96. https://doi.org/10.1080/0028825X.1993.10419536

Dematteis, B., Ferrucci, M. S., Ortega-Baes, P., & Coulleri, J. P. (2020). Genome size variation between the native and invasive ranges of Senecio madagascariensis (Asteraceae). Systematic Botany, 45(1), 212–218. https://doi.org/10.1600/036364420X15801369352487

Doležel, J., Bartoš, J., Voglmayr, H., & Greilhuber, J. (2003). Letter to the editor. Cytometry Part A, 51A(2), 127–128. https://doi.org/10.1002/cyto.a.10013

Doležel, J., Greilhuber, J., & Suda, J. (2007). Estimation of nuclear DNA content in plants using flow cytometry. Nature Protocols, 2(9), 2233–2244. http://dx.doi.org/10.1038/nprot.2007.310

Ghalambor, C. K., McKay, J. K., Carroll, S. P., & Reznick, D. N. (2007). Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. Functional Ecology, 21(3), 394–407. https://doi.org/10.1111/j.1365-2435.2007.01283.x

Grime, J. P., Hodgson, J. G., & Hunt, R. (2007). Comparative plant ecology: A functional approach to common British species (2nd ed.). Castlepoint Press.

Grime, & Mowforth, M. A. (1982). Variation in genome size-an ecological interpretation. Nature, 299(9), 151–153.

Hahn, M. A., van Kleunen, M., & Müller-Schärer, H. (2012). Increased phenotypic plasticity to climate may have boosted the invasion success of polyploid *Centaurea stoebe*. PLoS ONE, 7(11), Article e50284. https://doi.org/10.1371/journal.pone.0050284

Hollingsworth, P. M., Preston, C. D., & Gornall, R. J. (1998). Euploid and aneuploid evolution in *Potamogeton* (Potamogetonaceae): A factual basis for interpretation. Aquatic Botany, 60(4), 337–358. https://doi.org/10.1016/S0304-3770(97)00101-0

Holm, L. G. (1977). The world's worst weeds: Distribution and biology. University Press of Hawaii.

Ikeda, T., Iwasaki, K., Suzuki, T., Wong, L. J., & Pagad, S. (2021). Global register of introduced and invasive species - Japan. Invasive Species Specialist Group. https://doi.org/10.15468/nt2yla

Invasive Species Specialist Group. (2015). In The Global Invasive Species Database. Version 2015.1. http://www.iucngisd.org

Jensen, H. W. (1936). Meiosis in *Rumex*. I. Polyploidy and the origin of new species. Cytologia, 7(1–2), 1–22.

Kim, S. Y., Choi, H. W., Koo, D. H., Lee, W. K., Lee, J., & Bang, J. W. (2006). Characterization of eight *Rumex* species by FISH (fluorescence in situ hybridization) and 5S rDNA spacer sequences. Korean Journal of Genetics, 28(3), 243–251.

Kubátová, B., Trávníček, P., Bastlová, D., Čurn, V., Jarolímová, V., & Suda, J. (2008). DNA ploidy-level variation in native and invasive populations of *Lythrum salicaria* at a large geographical scale. Journal of Biogeography, 35(1), 167–176. https://doi.org/10.1111/j.1365-2699.2007.01781.x

Leitch, I. J., Johnston, E., Pellicer, J., Hidalgo, O., & Bennett, M. D. (n.d.). Plant DNA C-values Database | Royal Botanic Gardens, Kew. Retrieved 14 October 2021, from https://cvalues.science.kew.org/

Levin, D. A. (2020). Propagule pressure and the establishment of emergent polyploid populations. Annals of Botany, 127(1), 1–5. https://doi.org/10.1093/aob/mcaa187

Löve, Á., & Kapoor, B. M. (1967). A chromosome atlas of the collective genus *Rumex*. Cytologia, 32, 328–342.

Maguilla, E., Escudero, M., Jiménez-Lobato, V., Díaz-Lifante, Z., Andrés-Camacho, C., & Arroyo, J. (2021). Polyploidy expands the range of *Centaurium* (Gentianaceae). Frontiers in Plant Science, 12, Article 373. https://doi.org/10.3389/fpls.2021.650551

Mayrose, I., Zhan, S. H., Rothfels, C. J., Arrigo, N., Barker, M. S., Rieseberg, L. H., & Otto, S. P. (2015). Methods for studying polyploid diversification and the dead end hypothesis: A reply to Soltis et al. (2014). New Phytologist, 206(1), 27–35. https://doi.org/10.1111/nph.13192

Mayrose, I., Zhan, S. H., Rothfels, C. J., Magnuson-Ford, K., Barker, M. S., Rieseberg, L. H., & Otto, S. P. (2011). Recently formed polyploid plants diversify at lower rates. Science, 333(6047), 1257–1257. https://doi.org/10.1126/science.1207205

Michaelson, M. J., Price, H. J., Ellison, J. R., & Johnston, J. S. (1991). Comparison of plant DNA contents determined by Feulgen microspectrophotometry and laser flow cytometry. American Journal of Botany, 78(2), 183–188. https://doi.org/10.2307/2445241

Nagy, D. U., Stranczinger, S., Godi, A., Weisz, A., Rosche, C., Suda, J., Mariano, M., & Pal, R. W. (2018). Does higher ploidy level increase the risk of invasion? A case study with two geo-cytotypes of *Solidago gigantea* Aiton (Asteraceae). Journal of Plant Ecology, 11(2), 317–327. https://doi.org/10.1093/jpe/rtx005

Pandit, M. K., Tan, H. T. W., & Bisht, M. S. (2006). Polyploidy in invasive plant species of Singapore. Botanical Journal of the Linnean Society, 151(3), 395–403. https://doi.org/10.1111/j.1095-8339.2006.00515.x

Pellicer, J., & Leitch, I. J. (2020). The Plant DNA C-values database (release 7.1): An updated online repository of plant genome size data for comparative studies. New Phytologist, 226(2), 301–305. https://doi.org/10.1111/nph.16261

Pyšek, P., Skálová, H., Čuda, J., Guo, W.-Y., Suda, J., Doležal, J., Kauzál, O., Lambertini, C., Lučanová, M., Mandáková, T., Moravcová, L., Pyšková, K., Brix, H., & Meyerson, L. A. (2018). Small genome separates native and invasive populations in an ecologically important cosmopolitan grass. Ecology, 99(1), 79–90. https://doi.org/10.1002/ecy.2068

R Development Core Team. (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org

Raycheva, T. (2005). Chromosome counts in some Bulgarian taxa of subgenus *Rumex* (Polygonaceae). In B. Gruev, M. Nikolova, & A. Donev (Eds.), Proceedings of the Balkan Scientific Conference of Biology (pp. 397–408). http://web.uni-plovdiv.bg/mollov/bio/bscb2005/part1/397-408.pdf

Rice, A., Glick, L., Abadi, S., Einhorn, M., Kopelman, N. M., Salman-Minkov, A., Mayzel, J., Chay, O., & Mayrose, I. (2015). The Chromosome Counts Database (CCDB) – a community resource of plant chromosome numbers. New Phytologist, 206(1), 19–26. https://doi.org/10.1111/nph.13191

Rosche, C., Hensen, I., & Lachmuth, S. (2018). Local pre-adaptation to disturbance and inbreeding–environment interactions affect colonisation abilities of diploid and tetraploid Centaurea stoebe. Plant Biology, 20(1), 75–84. https://doi.org/10.1111/plb.12628

Rutland, C. A., Hall, N. D., & McElroy, J. S. (2021). The impact of polyploidization on the evolution of weed species: Historical understanding and current limitations. Frontiers in Agronomy, 3, Article 626454. https://doi.org/10.3389/fagro.2021.626454

Schlaepfer, D. R., Edwards, P. J., & Billeter, R. (2010). Why only tetraploid *Solidago gigantea* (Asteraceae) became invasive: A common garden comparison of ploidy levels. Oecologia, 163(3), 661–673. https://doi.org/10.1007/s00442-010-1595-3

Schlaepfer, D. R., Edwards, P. J., Semple, J. C., & Billeter, R. (2008). Cytogeography of *Solidago gigantea* (Asteraceae) and its invasive ploidy level. Journal of Biogeography, 35(11), 2119–2127. https://doi.org/10.1111/j.1365-2699.2008.01937.x

Schrieber, K., & Lachmuth, S. (2016). The Genetic paradox of invasions revisited: The potential role of inbreeding × environment interactions in invasion success. Biological Reviews, 92(2), 939–952. https://doi.org/10.1111/brv.12263

Seebens, H., Bacher, S., Blackburn, T. M., Capinha, C., Dawson, W., Dullinger, S., Genovesi, P., Hulme, P. E., van Kleunen, M., Kühn, I., Jeschke, J. M., Lenzner, B., Liebhold, A. M., Pattison, Z., Pergl, J., Pyšek, P., Winter, M., & Essl, F. (2021). Projecting the continental accumulation of alien species through to 2050. Global Change Biology, 27(5), 970–982. https://doi.org/10.1111/gcb.15333

Seebens, H., Blackburn, T. M., Dyer, E. E., Genovesi, P., Hulme, P. E., Jeschke, J. M., Pagad, S., Pyšek, P., Winter, M., Arianoutsou, M., Bacher, S., Blasius, B., Brundu, G., Capinha, C., Celesti-Grapow, L., Dawson, W., Dullinger, S., Fuentes, N., Jäger, H., … & Essl, F. (2017). No saturation in the accumulation of alien species worldwide. Nature Communications, 8(1), Article 14435. https://doi.org/10.1038/ncomms14435

Seebens, H., Blackburn, T. M., Hulme, P. E., van Kleunen, M., Liebhold, A. M., Orlova-Bienkowskaja, M., Pyšek, P., Schindler, S., & Essl, F. (2021). Around the world in 500 years: Inter-regional spread of alien species over recent centuries. Global Ecology and Biogeography, 30(8), 1621–1632. https://doi.org/10.1111/geb.13325

Šmarda, P., Knápek, O., Březinová, A., Horová, L., Grulich, V., Danihelka, J., Veselý, P., Šmerda, J., Rotreklová, O., & Bureš, P. (2019). Genome sizes and genomic guanine+cytosine (GC) contents of the Czech vascular flora with new estimates for 1700 species. Preslia, 91(2), 117–142. https://doi.org/10.23855/preslia.2019.117

Soltis, P. S., & Soltis, D. E. (2000). The role of genetic and genomic attributes in the success of polyploids. Proceedings of the National Academy of Sciences, 97(13), 7051–7057. https://doi.org/10.1073/pnas.97.13.7051

Sonnleitner, M., Hülber, K., Flatscher, R., Escobar García, P., Winkler, M., Suda, J., Schönswetter, P., & Schneeweiss, G. M. (2015). Ecological differentiation of diploid and polyploid cytotypes of *Senecio carniolicus* sensu lato (Asteraceae) is stronger in areas of sympatry. Annals of Botany, 117(2), 269–276. https://doi.org/10.1093/aob/mcv176

Stace, C. A. (1991). New flora of the British Isles. Cambridge University Press.

Stebbins, G. L. (1971). *Chromosomal evolution in higher plants*. Edward Arnold.

Takahashi, K., & Hanyu, M. (2015). Hybridization between alien species *Rumex obtusifolius* and closely related native vulnerable species *R. longifolius* in a mountain tourist destination. Scientific Reports, 5, Article 13898. https://doi.org/10.1038/srep13898

te Beest, M., Le Roux, J. J., Richardson, D. M., Brysting, A. K., Suda, J., Kubešová, M., & Pyšek, P. (2012). The more the better? The role of polyploidy in facilitating plant invasions. Annals of Botany, 109(1), 19–45. https://doi.org/10.1093/aob/mcr277

Thomson, G. M. (1922). The naturalisation of animals & plants in New Zealand. Cambridge University Press.

Travis, J. (1994). Evaluating the adaptive role of morphological plasticity. In P. C. Wainwright & S. M. Reilly (Eds.), Ecological morphology: Integrative organismal biology (pp. 99–122). University of Chicago Press.

Treier, U. A., Broennimann, O., Normand, S., Guisan, A., Schaffner, U., Steinger, T., & Müller-Schärer, H. (2009). Shift in cytotype frequency and niche space in the invasive plant *Centaurea maculosa*. Ecology, 90(5), 1366–1377. https://doi.org/10.1890/08-0420.1

Trewick, S., Morgan-Richards, M., & Chapman, H. (2004). Chloroplast DNA diversity of *Hieracium Pilosella* (Asteraceae) introduced to New Zealand: Reticulation, hybridization, and invasion. American Journal of Botany, 91, 73–85. https://doi.org/10.3732/ajb.91.1.73

USDA, NRCS. (n.d.). USDA Plants Database. The PLANTS Database. Retrieved 26 January 2022, from https://plants.sc.egov.usda.gov/home

Van Drunen, W. E., & Johnson, M. T. J. (2022). Polyploidy in urban environments. Trends in Ecology & Evolution, 37(6), 507–516. https://doi.org/10.1016/j.tree.2022.02.005

van Kleunen, M., Bossdorf, O., & Dawson, W. (2018). The ecology and evolution of alien plants. Annual Review of Ecology, Evolution, and Systematics, 49(1), 25–47. https://doi.org/10.1146/annurev-ecolsys-110617-062654

Wagner, W. H. (1970). Biosystematics and evolutionary noise. Taxon, 19(2), 146–151. https://doi.org/10.2307/1217945

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag. https://ggplot2.tidyverse.org

Williams, J. T. (1971). Seed polymorphism and germination. II. The role of hybridization in the germination polymorphism of *Rumex crispus* and *R. obtusifolius*. Weed Research, 11(1), 12–21. https://doi.org/10.1111/j.1365-3180.1971.tb00970.x

Wood, T. E., Takebayashi, N., Barker, M. S., Mayrose, I., Greenspoon, P. B., & Rieseberg, L. H. (2009). The frequency of polyploid speciation in vascular plants. Proceedings of the National Academy of Sciences, 106(33), 13875–13879. https://doi.org/10.1073/pnas.0811575106

Zaller, J. G. (2004). Ecology and non-chemical control of *Rumex crispus* and *R. obtusifolius* (Polygonaceae): A review. Weed Research, 44, 414–432. https://doi.org/10.1111/j.1365-3180.2004.00416.x

Zonneveld, B. J. M. (2019). The DNA weights per nucleus (genome size) of more than 2350 species of the Flora of The Netherlands, of which 1370 are new to science, including the pattern of their DNA peaks. Forum Geobotanicum, 8, 24–78.

Chapter 3

No inbreeding depression but lower germination within the introduced range in *Rumex conglomeratus*

3.1. Introduction

The breeding system of a plant species can have a big impact on its ability to survive after introduction to a new region and to then establish and thrive in the new area (Baker, 1955; Barrett et al., 2008; Barrett, 2013). In general, certain reproductive characteristics or traits are known to be beneficial during these phases of colonisation. Firstly, self-compatibility is helpful to counter the likely limited number of conspecifics introduced, and thus we might expect a bias towards selfing in successful invaders within the introduced range (Baker, 1955; Barrett et al., 2008). Secondly, the negative consequences of selfing may be countered by intra- and interspecific hybridization as well as through polyploidisation (Soltis & Soltis, 2000). Thirdly, wind-pollination helps to counter the lack of specialised pollinators allowing for species to establish without the need for pollinators (Baker, 1955). Lastly, high seed set allows for a quick increase in numbers solidifying the establishment and enabling the spread of the species (Richardson & Pyšek, 2006; Pyšek & Richardson, 2007; van Kleunen, Dawson, et al., 2015).

Baker's Law states, that selfing species are better colonisers, as they can avoid mate and pollinator limitations, making it possible for a single individual to establish a population (Baker, 1955). In addition to countering low mate abundance, selfing can benefit by purging mutation load that would otherwise threaten small populations (but see also Szövényi et al., 2014; summarised in Jullien et al., 2021), but it also allows for a rapid fixing of beneficial mutations (Glémin, 2012). However, as outcrossing populations accumulate novel combinations faster than selfers, some level of outcrossing can be highly beneficial (Hartfield et al., 2017). Indeed, it is not uncommon to have populations with primarily outcrossing individuals at the core and selfing individuals at the range edges, where conspecifics might be scarce (summarised in Pannell, 2015). However, if selfing is only as a way to counter mate limitation, the population is expected to shift to outcrossing relatively quickly after the population number goes up, whereas if the selfing individuals have a fitness advantage, the opposite can happen.

Polyploid species are known to have increased levels of genetic variation and decreased inbreeding depression indicating that they can tolerate higher levels of selfing without negative consequences (summarised in Soltis & Soltis, 2000). Two invasive rose species (Rosaceae), *Rosa rubiginosa* L. and *R. canina* L. exhibit mixed mating patterns likely contributing to their invasion success (Mazzolari et al., 2017). The number of fruits both species produce is similar between selfing and outcrossing individuals, potentially suggestive of no reproductive cost associated with selfing. Similarly, cocklebur *Xanthium italicum* Moretti (Compositae) shows mixed mating, and both selfing and out-crossing lead to high seed-set (Jiang et al., 2022). In addition, when compared to its native, less weedy congener, *X. sibiricum* Patr. et Widd, *X. italicum* has a longer flowering period and produces more seeds.

The pollen:ovule (P:O) ratio can be used as a simple proxy of the reproductive system, because plants have optimised their pollen production to ensure sufficient fecundity without wasting costly resources for excess amounts of pollen (Cruden, 1977). According to Cruden (1977), plants reproducing by selfing will require a smaller amount of pollen for pollinating their own flowers, whereas out-crossers need more pollen for guaranteeing that at least some of it will pollinate the conspecifics successfully. In short, when the number of pollen grains produced and thus the P:O ratio increases, so does the likelihood for outcrossing. However, Cruden (1977) notes that there are exceptions, and P:O ratios will also reflect the successional stage, with selfing benefiting in early successional stages with small population numbers, and vice versa. In general, wind-pollinated species show higher P:O ratios than animal-pollinated species, but according to Friedman & Barrett (2009), this might be explained by pollen competition. In addition, a negative correlation between P:O ratio and pollen size has been found, although this relationship decreases when the pollen grain number is taken into account (Götzenberger et al., 2007). This size-number trade-off is said to reflect the resources allocated to 'male function', and theories suggest that selfing species allocate less resources to male function than outcrossing species. In a study investigating 40 species of *Pedicularis* L. (Orobanchaceae), Yang & Guo (2004) found a negative relationship between pollen size and number.

The traits behind invasiveness in general have been in the centre of research attention for decades, especially with the search for the 'ideal weed' (e.g., Rejmánek & Richardson, 1996; Pyšek & Richardson, 2007; van Kleunen, Dawson, et al., 2015). Multispecies studies are important for broad generalisations (Pyšek & Richardson, 2007; van Kleunen, Dawson, et al., 2015), but to understand small-scale patterns, more research is needed to compare these traits between individuals found in the native and introduced ranges to see if the reproductive methods has shifted following introduction. I studied a circumpolar weed species belonging to the genus *Rumex* L. (Polygonaceae): *R. conglomeratus* Murr. This species is native to Europe and some parts of Asia but has been introduced to all continents apart from Antarctica (Löve & Kapoor, 1967; Holm, 1977; Akeroyd, 2014; Carlin et al., 2022). This species is highly variable in its morphology, and it has become a persistent weed within agricultural landscapes, such as damp paddocks and ditches (Stace, 1991; Akeroyd, 2014; Harrington, 2019). *Rumex conglomeratus* also possesses multiple traits associated with the ideal weed characteristics (Rejmánek & Richardson, 1996), but there are still many unknowns, and studies comparing traits of individuals from their native and introduced ranges are rare (Costan et al. 2022).

In terms of the breeding system, *R. conglomeratus* is a wind-pollinated hermaphrodite, and in New Zealand (NZ) is assumed to be protandrous (Webb et al., 1988), but the extent of selfing is not well known. According to Cavers & Harper (1964), the congeneric *R. crispus* L. populations in the UK were 25–100 % selfing. However, protandry in *R. crispus* is said to prevent self-pollination (Grime et al., 2007), but Friedman & Barrett (2009) report that despite this *R. crispus* has overlapping male and female functions actually favouring selfing. Studies show that these species can produce tens of thousands of flowers every year, although they often start flowering in their second year (Cavers & Harper, 1964). Pollen numbers are not reported from *R. conglomeratus* before and investigating the P:O ratios could thus help to understand the degree of selfing. As *R. conglomeratus* only produces one ovule per flower, the number of pollen grains produced is the same as P:O ratio.

The purpose of this study was to investigate the propensity for selfing measured as the P:O ratio (Cruden, 1977; Dafni, 1992) and how it differs between the native and introduced ranges, and to understand the potential costs from selfing vs outcrossing as measured by differences in seed production and performance in self-pollinated and outcrossed offspring. Specifically, I asked:

1) Are there differences in the P:O ratios reflecting different breeding systems between congeners from the native and introduced regions?
2) Are the seed weights and germination probabilities similar between the bagged and unbagged treatments in plants from the native and introduced regions?

I expected the plants from the introduced range to produce less pollen and thus show a preference towards selfing. I conducted a selfing experiment by bagging half of the plants and leaving half of them unbagged. I then calculated the P:O ratios to estimate the breeding system and specifically the likelihood for selfing (Cruden, 1977) and compared the seed set and germination probability between the treatments. I expected the bagged progeny from the introduced range to be heavier and have a higher germination percentage compared to the unbagged progeny from the introduced range as an adaptation to selfing due to the shorter residency time. Minimal differentiation was expected between the treatments in the native progeny. A cross-pollination experiment would have provided a deeper understanding of the selfing-outcrossing-patterns, but the small flower size and highly varying floral characteristics rendered it logistically difficult.

3.2. Methods

3.2.1. Pollen:ovule ratio in *Rumex conglomeratus*

Rumex conglomeratus seeds were collected during the respective summers of 2016–2017 from four regions in the UK and NZ (full sampling protocol in Bufford & Hulme, 2021a, 2021b). Then, roughly ten seeds originating from each seed family (= seeds from one individual) were germinated in a greenhouse and one plant was kept and grown up. From plants flowering in the greenhouse in 2020, ten buds about to open were collected and stored in EtOH in 2.5 ml Eppendorf tubes in a dark cabinet until measurement. Three of those buds were chosen at random, and the whole bud with anthers was squashed on a glass slide using a flat side of a blade. After that, it was transferred to a 200 µl tube containing 60 µl of 80 % EtOH, 30 µl of detergent diluted in dH2O and 10 µl of fuchsin (2 % in ddH2O), and vortexed for 30 seconds rotating the tube every few seconds. Then 10 µl of the suspension was pipetted on either slot of the haemocytometer creating two replicates of each bud, and the number of pollen grains were estimated using a Neubauer-Improved haemocytometer (Marienfeld, Germany). Pollen grains were counted from the corner and middle squares using a microscope (Olympus BX40CY) with a 40x / 0.65 (Olympus Ach, Japan) lens and a 10x ocular piece. Average number of pollen grains per square over all six counts were used to calculate the number of pollen grains in a flower for estimating the P:O ratio:

$$Total\ \#pollen\ per\ flower = \frac{number\ of\ pollen\ in\ a\ square \times volume\ in\ eppendorf}{volume\ on\ haemocytometer\ square} = \frac{x \times 100\ \mu l}{0.1\ \mu l}$$

In addition, the pollen size of 40 pollen grains was measured per plant from the pollen count samples, this time using two out of the three buds. Grains were measured using the ocular graticule and a ×60 / 0.80 objective (Olympus Ach, Japan). The ocular was calibrated using a 2 mm calibration graticule to get a more appropriate size estimation in µm rather than in arbitrary units. Pollen size was measured to half a graticule step accuracy. I searched for references to the pollen sizes within existing literature and databases to compare the estimates obtained in this study. No pollen size measurements were found for *R. conglomeratus* from commonly used search engines (i.e., Scopus; Google Scholar; search term: ["*Rumex conglomeratus*" OR "*R. conglomeratus*"] AND "pollen size") nor pollen databases, PalDat – Palynological Database (https://www.paldat.org/) and The Global Pollen Project (Martin & Harvey, 2017).

Lastly, I tested the relationship between pollen grain number and pollen grain size and compared slopes of the regression lines between the native and introduced regions. Findings of negative relationships between pollen grain number and pollen grain size exist (e.g., Cruden, 2000; but see Götzenberger et al., 2007). This relationship might be less important in wind-pollinated plants compared to animal-pollinated plants and can reflect more the changes in either flower or inflorescence size (summarised in Friedman & Barrett, 2009). All statistical comparisons in this chapter were made using R (version 4.2.1., R Development Core Team, 2008) and figures were drawn using the ggplot2 package (Wickham, 2016).

3.2.2. Selfing vs open-pollination experiment

The same plants used for the P:O measurements were also utilised for this experiment. Flowering plants were either bagged with a plastic freezer bag and tied below the inflorescence (selfing/bagged) or just tagged (control/unbagged). Flowering was not induced, but rather all plants flowering at the time were utilised.

After the seeds had matured, a sample of seeds was collected in a zip-lock bag for later measurements. To estimate the potential cost from selfing and inbreeding depression, seed weight and germination percentages were recorded and compared between the two bagged and unbagged treatments. Where possible, up to 30 seeds were weighed after removing the seed husk, and mass was averaged over one seed for comparison. All seeds were then put on damp filter paper in open Petri dishes and kept on room temperature in indirect light to monitor germination for 30 days, watering by misting a few times a day when the filter paper started drying out. Seeds that started germinating were removed and the germination date was recorded. Total germination percentage was calculated for each plant.

3.3. Results

3.3.1. Pollen:ovule ratios show no differences between the ranges

A total of 46 plants flowered in the greenhouse during the time of collections. Out of these, 19 *R. conglomeratus* from the UK and 14 from NZ were used in the analyses, as the rest of the plants did not yield large enough datasets for comparisons between the native and introduced ranges. An average of 10 914 and 10 958 pollen grains per flower were counted in plants from UK and NZ, respectively (FIGURE 3.1). The countries were compared using Welch's two sample t-test without assuming equal variances, and no evidence was found to support differences between the countries. I measured 40 pollen grains from 19 UK and 14 NZ plants using two flowers per plant. The average sizes were 9.7 and 9.4 µm for UK and NZ, respectively (FIGURE 3.1). The data showed little evidence for size differences between the two countries.

To test whether the pollen grain number decreases with increasing pollen grain size, I tested the relationship using linear regression (FIGURE 3.2). Firstly, the data shows no support for a decreasing pollen grain number with increasing pollen size, if anything, the trend is opposite, but even this is not statistically significant ($F(1,31) = 3.101$, $p = 0.09$). Secondly, I compared the differences in pollen grain size and number between samples from the UK and NZ, and found limited differences between the countries (ANCOVA, $F(1, 30) = 0.160$, $p = 0.69$).

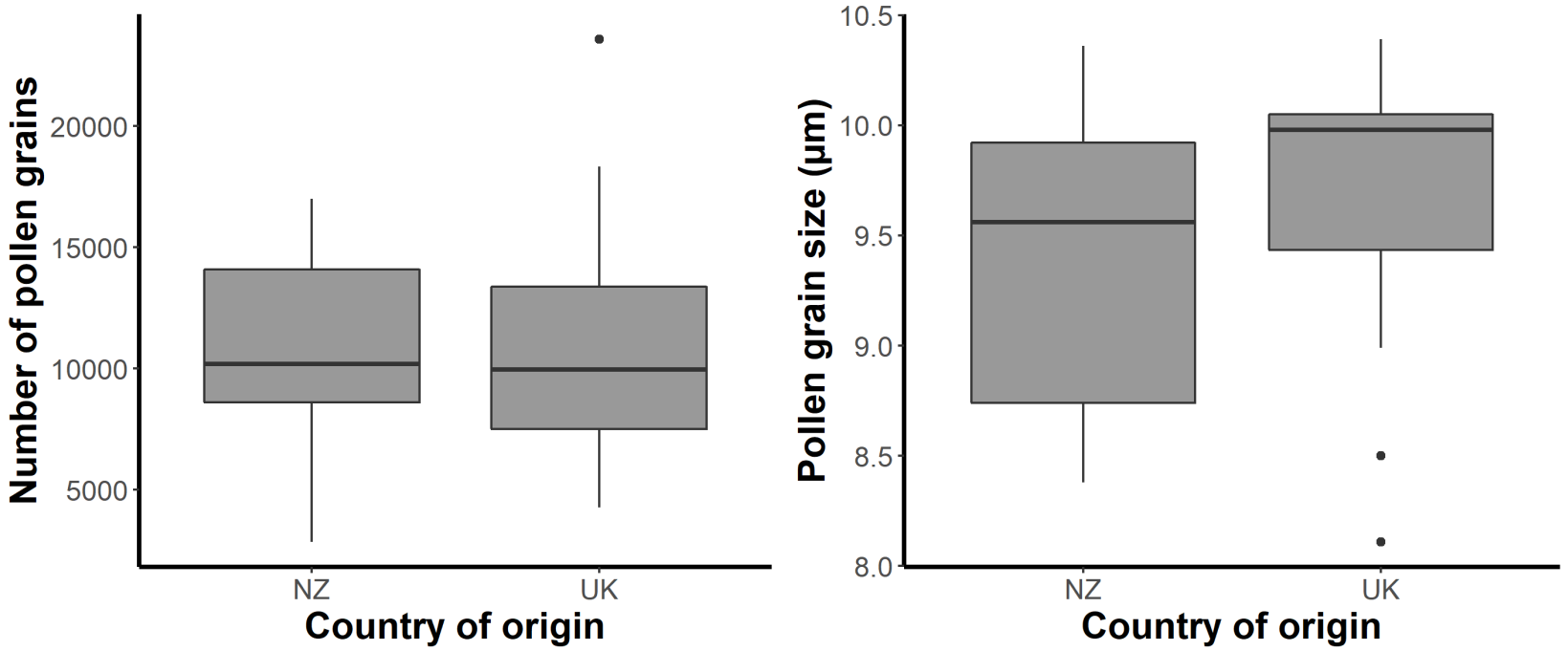

FIGURE 3.1: Boxplots of pollen grain numbers per flower (left) and average pollen grain diameters in µm (right) in *R.* conglomeratus between New Zealand (NZ) and the UK. The pollen grain numbers did not differ significantly between countries (Welch's two sample t-test, t = -0.028195, df = 29.932, p = 0.9777), nor did the pollen grain sizes (Welch's t-test, t = 1.1759, df = 26.651, p = 0.25). The box represents the interquartile range with the median represented with a thick line and the whiskers extending from the 5th to the 95th percentile, with values outside this range represented as dots.

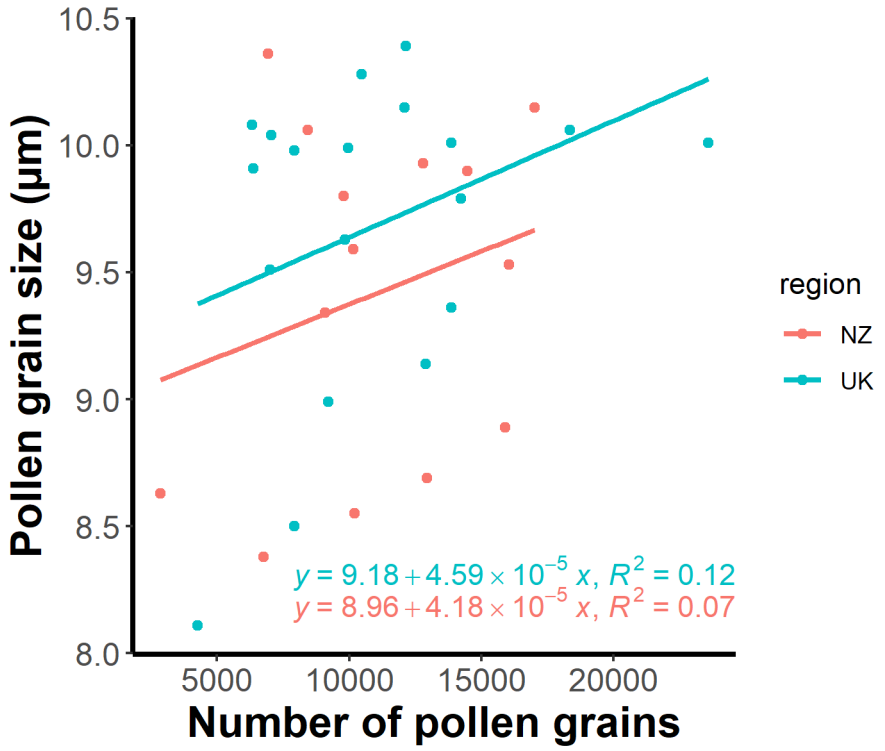

FIGURE 3.2: The relationship between pollen grain size and the number of pollen grains in *R. conglomeratus* samples from New Zealand (NZ) and the UK. Equations for the respective regression lines and R^2 values are given in the figure.

3.3.2. Bagged and unbagged treatments yield viable seeds with no weight differences

In total 11 flowering plants were bagged, 3 from NZ and 8 from UK, and for comparison 14 unbagged plants were tagged, 3 from NZ and 11 from UK. All bagged plants developed seeds within the bags, as did all the unbagged plants. After seeds had matured, samples were collected. Due to bagging the plants using freezer bags (to keep external pollen out) and a high humidity in the greenhouse, some bagged seeds were lost to mould, and thus the whole inflorescence could not be weighed. Instead, a sample of up to 30 seeds were collected from all but one UK plant that did not yield more than 16 seeds. Given the unbalanced group sizes, ANOVA (type III) was used to measure the effect of treatment and country on seed weight. Both treatment and country affected the seed weight, as did their interaction (TABLE 3.1, FIGURE 3.3). On average, the mass of the unbagged seeds from NZ were around 27 % heavier and the bagged seeds around 22 % lighter than the UK seeds.

At least one seed germinated from all of the seed families, and in over 40 % of the seed families all of the seeds germinated. The germination probability data was not normally distributed (Shapiro-Wilk normality test: $W = 0.68531$, $p < 0.001$), so non-parametric Kruskal-Wallis rank sum tests were used to measure the effect of the treatment and country on the germination probability. The treatment did not have a significant effect on the germination probability (Kruskal-Wallis chi-squared = 0.003, df = 1, p-value = 0.957), but the country did (Kruskal-Wallis chi-squared = 4.361, df = 1, p-value = 0.037).

The UK seeds had, on average, over 20 % higher germination probabilities in both treatments compared to the NZ seeds (FIGURE 3.4).

TABLE 3.1: ANOVA (Type III) looking at the effect of country and treatment (unbagged vs. bagged) on seed weight (mg) in *Rumex conglomeratus* from New Zealand and the UK.

	Sum Sq	**Df**	**F value**	**p value**
(Intercept)	3.4409	1	274.6678	< 0.001
Treatment	0.1418	1	11.3202	0.003
Country	0.1245	1	9.9381	0.005
Treatment:country	0.2210	1	17.6409	0.004

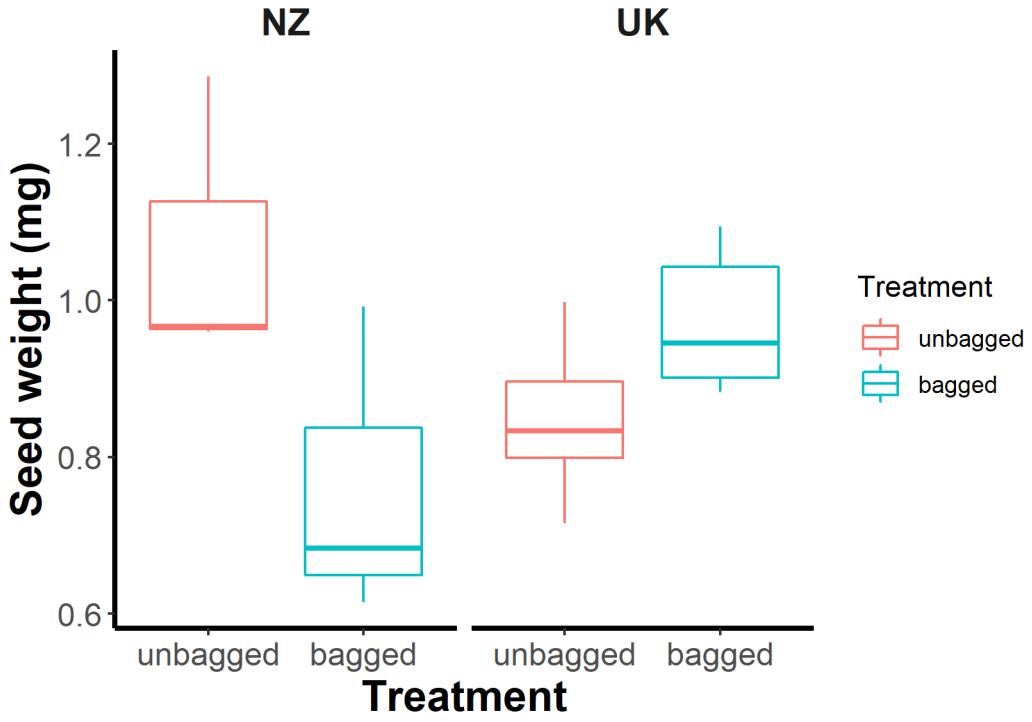

FIGURE 3.3: Boxplots of average seed weights between selfing (bagged) and outcrossing (control) treatments of *R. conglomeratus* samples from New Zealand (NZ) and the UK. The box represents the interquartile range with the median represented with a thick line and the whiskers extending from the 5th to the 95th percentile.

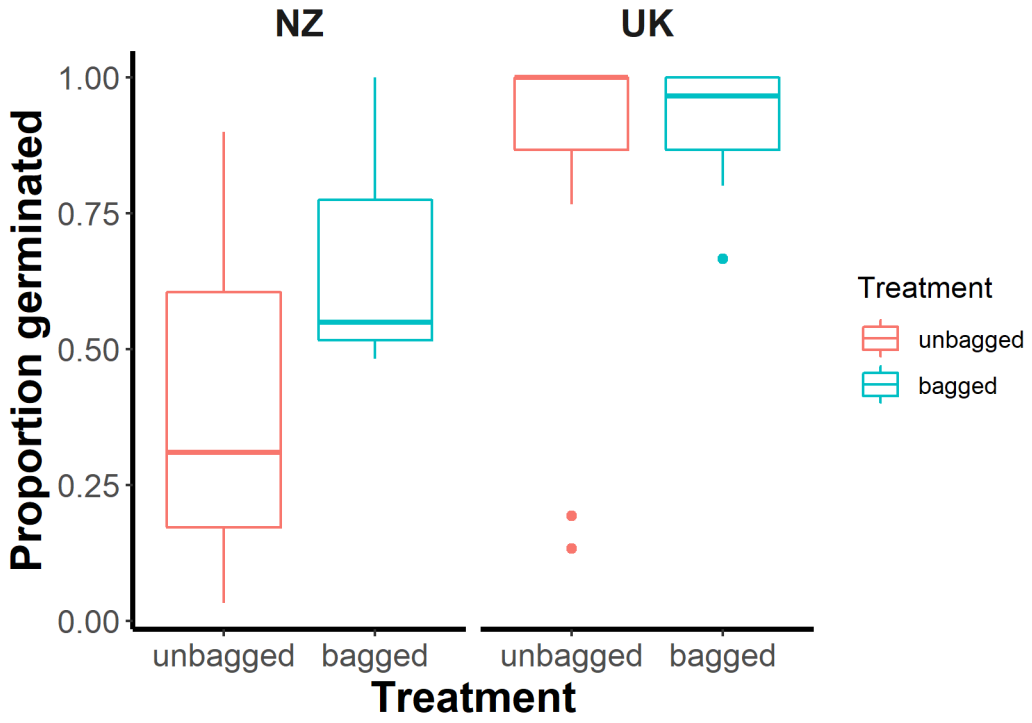

FIGURE 3.4: Boxplots showing the average total proportion of germinated seeds in unbagged (control) and bagged (selfing) treatments of *R. conglomeratus* samples from New Zealand (NZ) and UK. The box represents the interquartile range with the median represented with a thick line and the whiskers extending from the 5th to the 95th percentile, with values outside this range (outliers) represented as dots.

3.4. Discussion

Research indicates that plants capable of self-pollination have a higher likelihood for establishing in introduced regions (Baker, 1955; Barrett et al., 2008; Barrett, 2013). In addition, if a species has a mixed mating system, it can establish when the density is low, but as the conspecifics get more common, out-crossing can introduce genetic variation to populations (Briggs & Walters, 2016). The purpose of this study was to compare the breeding system of *R. conglomeratus* originating from the UK and NZ. More specifically, I assessed the breeding system using P:O ratios as a proxy, and then conducted a selfing experiment, to investigate costs from selfing, and how these differ between the native and introduced region. I found no support that the native and introduced ranges differ from each other in either P:O ratios or pollen grain sizes. In the selfing experiment all bagged plants produced seeds, but both the seed weight and the germination probability were affected by the country, and the seed weight also by the treatment. The plants from the introduced range had significantly lower germination probability compared to the plants from the native range, but the probability was increased in the selfing treatment compared to the control treatment. This suggests that selfing is a viable option for *R. conglomeratus* and that there are limited consequences from selfing.

The P:O ratios reported here are among the first reports for the hermaphrodite *Rumex* species and certainly first for *R. conglomeratus*. It is suggested that introduced

plants may prefer self-pollination as it is not dependent on the number of plants introduced, but rather a single plant can produce offspring and survive in the new environment (Cruden, 1977). In addition, populations in their early stages of invasion are assumed to have higher rates of selfing due to mate limitation. Surprisingly, the P:O ratios indicated obligate xenogamy as the reproductive method in both countries. In addition, as *R. conglomeratus* is assumed to have protandrous flowering patterns in NZ (Webb et al., 1988), which is associated with self-incompatibility (Routley et al., 2004). However, this is refuted by the results from the selfing experiments. This could have more than one explanation: for example, wind-pollinated plants are known to produce more pollen grains compared to animal-pollinated plants (Cruden, 1977; Friedman & Barrett, 2009), and P:O ratios can be more accurate within taxa (Cruden, 1977). In addition, pollen grains are not necessarily that costly to produce, and as the plants were grown in the greenhouse, conditions might have been relatively good for growth in general. Finally, the samples originated from well-established populations, and the species has been in NZ for almost two centuries (Thomson, 1922), and thus the plants in the study were not in the early stages of establishment. Studies show that some plants have plasticity in the breeding system based on the invasion/colonisation stage (Pyšek & Richardson, 2007). For example, devil's backbone *Euphorbia tithymaloides* L. (Euphorbiaceae) shows a reduction in the P:O ratio in populations towards the population expansion front, which could indicate a switch towards selfing from normally xenogamous method of reproduction (Cacho & José-Zacatula, 2020). It would be interesting to compare the P:O ratios between *Rumex* plants from the middle of a population (stable population) and the leading edge of a population (spreading population). This would help to assess whether *R. conglomeratus* shows adaptation of breeding system based on the invasion/colonisation stage.

The pollen size measurements for *R. conglomeratus* (FIGURE 3.1) were likewise among the first reported for the species and thus comparisons to other measurements cannot be made. However, the shape of the pollen grains matched that of *R. obtusifolius* pollen grains reported in the databases, only the size was much smaller than that of *R. obtusifolius* (Martin & Harvey, 2017; Sam & Auer, 2021). However, the methods used for analyses is known to affect pollen size (Reitsma, 1969; Meudt, 2016), which means that direct size comparisons between methods are often unreliable. In any case, this study provides important information of the pollen sizes using methods detailed in Dafni (1992) and can be used as future reference. In addition, a positive, albeit statistically non-significant, correlation between the pollen size (measured as a diameter of the pollen grain) and the number of pollen grains produced by a flower was found. This pattern was similar in both the native and introduced ranges in (FIGURE 3.2). The opposite trend is commonly true in angiosperms (Cruden, 2000), but the relationship is stronger in animal-pollinated species compared to wind-pollinated species. However, similar positive relationships have been found in other species. The genus *Tarasa* Phil. (Malvaceae) showed a significant positive correlation between pollen grain size (measured as the surface volume) and P:O ratios of 24 measured species ranging from diploids to tetraploids (Tate & Simpson, 2004).

Self-compatibility has been associated with colonization ability, as reproduction is not density-dependent, but rather a single plant can set seed and thrive (Cruden, 1977). Despite the P:O ratio classifying the breeding mode as facultative xenogamy, I found strong evidence for selfing. In general, the selfed progeny had on average 11 % higher

germination probability compared to the control treatment (FIGURE 3.4), and all bagged plants produced viable seeds. On average, the NZ seeds had around 20 % lower germination probability compared to the UK seeds, but the treatments did not differ from each other statistically. This likely indicates that *R. conglomeratus* has a mixed breeding system, with a slight bias towards selfing. This could help the species during colonisation or range expansion, when the number of conspecifics is low. These results have been confirmed in field conditions within the hermaphrodite congenerics: Friedman & Barrett (2011) found limited pollen loads with increasing distance from conspecifics in wind-pollinated dioecious *Rumex* plants, *R. nivalis* and *R. acetosella*, but this was not true for the hermaphrodite *R. crispus*. The authors hypothesize that this could be due to buffering effects from self-pollination in *R. crispus*. Conversely, as the plant density increases, there can be a shift towards cross-pollination. As seed size and germination rates found in this study were equal between bagged and unbagged treatments (TABLE 3.1), mixed mating is in *R. conglomeratus* is likely. Previous studies regarding interspecific hybrids and how common they are do suggest that out-crossing is a widespread phenomenon (see Chapter 3). Furthermore, I found low within and between population genetic differentiation in *R. conglomeratus* from both countries (see Chapter 5), which indicates that at least some level of out-crossing is likely.

However, the germination probability of the introduced provenance was significantly lower than that of the native provenance. Similar results have been found before in *R. conglomeratus*, but also in the congenerics, *R. obtusifolius* and *R. crispus* in seeds from NZ when compared to the seeds from the UK (J. Bufford & T. Carlin, personal communication, May 27, 2022). This could indicate inbreeding or outbreeding depression, or adaptation to slightly different condition between the countries. However, *R. conglomeratus* shows high and similar plasticity to both drought and flooding (Bufford & Hulme, 2021a), so the greenhouse conditions are unlikely to have affected here. In addition, minimal genetic differentiation was found between populations from NZ and the UK (Chapter 5) and thus it is likewise unlikely to explain the differences. To properly assess the differences in germination, a larger scale selfing vs outcrossing experiment should likely be conducted.

Lastly, this study confirmed the difficulties related to a cross-pollination experiment. The flowers of this species of *Rumex* are very small, only around 1.5 mm wide and 2.5 mm long (Akeroyd, 2014), so emasculating the anthers without destroying the whole flower was not feasible. Some individuals had female flowers in addition to hermaphrodite flowers, but this was not consistent and thus just removing hermaphroditic flowers did not provide a good alternative for emasculating the flowers. If these difficulties are overcome, a cross-pollination study with differing genetic/geographic distances to the pollen donor could be conducted to fully estimate the costs from both selfing and outcrossing (i.e., admixture). For example, van Kleunen et al. (2015) studied fitness consequences from admixture in *Erythrante guttata* (before *Mimulus guttatus*; Phrymaceae) by creating seeds from selfing and three different levels of out-crossing (within population, within range and between the native and introduced range). They found out that both native and introduced populations suffered from inbreeding depression and conversely benefited from between-range-crossing. Similar effects could be expected from *R. conglomeratus* as well.

3.4.1. Conclusions

This study showed that *R. conglomeratus* exhibits mixed mating, despite being classified as facultatively xenogamous in both NZ and the UK based on the P:O ratio. In addition, the plants produced viable seeds under bagged and unbagged treatments, but both the treatment and the country affected the seed weight. The germination probability was significantly lower in the introduced provenance in both treatments, compared to the native provenance. However, the sample sizes here are small, a more comprehensive study using field collected samples could provide more conclusive answers on how germination is affected by the treatment, and whether the germination probability is consistently lower in the introduced region. Nevertheless, the selfed progeny in both countries did show increased germination probability, suggesting that selfing could aid the species during colonisation. These traits, and others such as a high seed-set and seed longevity, hybridisation, allelopathic tendencies, tolerance to grazing and insect damage (Zaller, 2004; Akeroyd, 2014) make this species very successful and invasive. However, even with low inbreeding depression at the germination stage, inbreeding depression could still manifest in later life-stages. Future research should focus on broadening the study to include more countries within the introduced region and to perform a more comprehensive selfing experiment to see if fitness differences remain absent past the germination stage.

3.5. References

Akeroyd, J. R. (2014). Docks and knotweeds of Britain and Ireland (2nd ed.). Botanical Society of Britain & Ireland.

Baker, H. G. (1955). Self-compatibility and establishment after 'long-distance' dispersal. Evolution, 9(3), 347–349. https://doi.org/10.2307/2405656

Barrett, S. C. H. (2013). IV.8. evolution of mating systems: Outcrossing versus selfing. In J. B. Losos, D. A. Baum, D. J. Futuyma, H. E. Hoekstra, R. E. Lenski, A. J. Moore, C. L. Peichel, D. Schluter, & M. C. Whitlock (Eds.), The Princeton guide to evolution (pp. 356–362). Princeton University Press. https://doi.org/10.1515/9781400848065-050

Barrett, S. C. H., Colautti, R. I., & Eckert, C. G. (2008). Plant reproductive systems and evolution during biological invasion. Molecular Ecology, 17(1), 373–383. https://doi.org/10.1111/j.1365-294X.2007.03503.x

Briggs, D., & Walters, S. M. (2016). Plant variation and evolution (4th ed.). Cambridge University Press.

Bufford, J. L., & Hulme, P. E. (2021a). Increased adaptive phenotypic plasticity in the introduced range in alien weeds under drought and flooding. Biological Invasions, 23(8), 2675–2688. https://doi.org/10.1007/s10530-021-02532-5

Bufford, J. L., & Hulme, P. E. (2021b). Seed size-number trade-offs are absent in the introduced range for three congeneric plant invaders. Journal of Ecology, 109, 3849–3860. https://doi.org/10.1111/1365-2745.13761

Cacho, N. I., & José-Zacatula, D. (2020). Geographic patterns in pollen production in the plant ring species *Euphorbia tithymaloides* in the Caribbean. Systematic Botany, 45(4), 845–853. https://doi.org/10.1600/036364420X16033962925321

Carlin, T. F., Bufford, J. L., Hulme, P. E., & Godsoe, W. K. (2022). Global assessment of three *Rumex* species reveals inconsistent climatic niche shifts across introduced ranges. Research Square. https://doi.org/10.21203/rs.3.rs-1358439/v1

Cavers, P. B., & Harper, J. L. (1964). *Rumex obtusifolius* L. and *R. crispus* L. Journal of Ecology, 52(325), 737–766.
Costan, C.-A., Godsoe, W. K., Bufford, J. L., Marris, J. W. M., & Hulme, P. E. (2022). Can the enemy release hypothesis explain the success of *Rumex* (Polygonaceae) species in an introduced range? Biological Invasions. https://doi.org/10.1007/s10530-022-02810-w
Cruden, R. W. (1977). Pollen-ovule ratios: A conservative indicator of breeding systems in flowering plants. Evolution, 31(1), 32–46. https://doi.org/10.2307/2407542
Cruden, R. W. (2000). Pollen grains: Why so many? Plant Systematics and Evolution, 222(1–4), 143–165. http://dx.doi.org/10.1007/BF00984100
Dafni, A. (1992). Pollination ecology: A practical approach. IRL Press; Oxford University Press.
Friedman, J., & Barrett, S. C. H. (2009). Wind of change: New insights on the ecology and evolution of pollination and mating in wind-pollinated plants. Annals of Botany, 103(9), 1515–1527. https://doi.org/10.1093/aob/mcp035
Friedman, J., & Barrett, S. C. H. (2011). The evolution of ovule number and flower size in wind-pollinated plants. The American Naturalist, 177(2), 246–257. https://doi.org/10.1086/657954
Glémin, S. (2012). Extinction and fixation times with dominance and inbreeding. Theoretical Population Biology, 81(4), 310–316. https://doi.org/10.1016/j.tpb.2012.02.006
Götzenberger, L., Durka, W., Kühn, I., & Klotz, S. (2007). The relationship between the pollen–ovule ratio and pollen size: Another comparative test of a sex allocation hypothesis. Evolutionary Ecology Research, 9, 1145–1161.
Grime, J. P., Hodgson, J. G., & Hunt, R. (2007). Comparative plant ecology: A functional approach to common British species (2nd ed.). Castlepoint Press.
Harrington, K. C. (2019). Weed management in New Zealand pastures. Agronomy, 9(8), Article 448. https://doi.org/10.3390/agronomy9080448
Hartfield, M., Bataillon, T., & Glémin, S. (2017). The evolutionary interplay between adaptation and self-fertilization. Trends in Genetics, 33(6), 420–431. https://doi.org/10.1016/j.tig.2017.04.002
Holm, L. G. (1977). The world's worst weeds: Distribution and biology. University Press of Hawaii.
Jiang, M., Ma, M., Lin, H., & Ma, X. (2022). Reproductive strategies of *Xanthium italicum* differ from those of native *Xanthium sibiricum*, and they are key to its invasiveness. Plant Ecology, 223(4), 453–463. https://doi.org/10.1007/s11258-021-01219-0
Jullien, M., Ronfort, J., & Gay, L. (2021). How and when does outcrossing occur in the predominantly selfing species *Medicago truncatula*? Frontiers in Plant Science, 12, Article 619154. https://www.frontiersin.org/article/10.3389/fpls.2021.619154
Löve, Á., & Kapoor, B. M. (1967). A chromosome atlas of the collective genus *Rumex*. Cytologia, 32, 328–342.
Martin, A. C., & Harvey, W. J. (2017). The global pollen project: A new tool for pollen identification and the dissemination of physical reference collections. Methods in Ecology and Evolution, 8(7), 892–897. https://doi.org/10.1111/2041-210X.12752
Mazzolari, A. C., Marrero, H. J., & Vázquez, D. P. (2017). Potential contribution to the invasion process of different reproductive strategies of two invasive roses. Biological Invasions, 19(2), 615–623. https://doi.org/10.1007/s10530-016-1315-y
Meudt, H. (2016). Pollen morphology and its taxonomic utility in the Southern Hemisphere bracteate-prostrate forget-me-nots (*Myosotis*, Boraginaceae). New Zealand Journal of Botany, 54(4), 475–497. https://doi.org/10.1080/0028825X.2016.1229343
Pannell, J.R. (2015), Evolution of the mating system in colonizing plants. Molecular Ecology, 24, 2018-2037.
Pyšek, P., & Richardson, D. M. (2007). Traits associated with invasiveness in alien plants: Where do we stand? In W. Nentwig (Ed.), *Biological Invasions* (pp. 97–125). Springer. https://doi.org/10.1007/978-3-540-36920-2_7
R Development Core Team. (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org

Reitsma, Tj. (1969). Size modification of recent pollen grains under different treatments. Review of Palaeobotany and Palynology, 9(3), 175–202. https://doi.org/10.1016/0034-6667(69)90003-7
Rejmánek, M., & Richardson, D. M. (1996). What attributes make some plant species more invasive? Ecology, 77(6), 1655–1661. https://doi.org/10.2307/2265768
Richardson, D. M., & Pyšek, P. (2006). Plant invasions: Merging the concepts of species invasiveness and community invasibility. Progress in Physical Geography: Earth and Environment, 30(3), 409–431. https://doi.org/10.1191/0309133306pp490pr
Routley, M. B., Bertin, R. I., & Husband, B. C. (2004). Correlated evolution of dichogamy and self-incompatibility: A phylogenetic perspective. International Journal of Plant Sciences, 165(6), 983–993. https://doi.org/10.1086/423881
Sam, S., & Auer, W. (2021). *Rumex obtusifolius*. PalDat - A Palynological Database. https://www.paldat.org/pub/Rumex_obtusifolius/305151
Soltis, P. S., & Soltis, D. E. (2000). The role of genetic and genomic attributes in the success of polyploids. Proceedings of the National Academy of Sciences, 97(13), 7051–7057. https://doi.org/10.1073/pnas.97.13.7051
Stace, C. A. (1991). New flora of the British Isles. Cambridge University Press.
Szövényi, P., Devos, N., Weston, D. J., Yang, X., Hock, Z., Shaw, J. A., Shimizu, K. K., McDaniel, S. F., & Wagner, A. (2014). Efficient purging of deleterious mutations in plants with haploid selfing. Genome Biology and Evolution, 6(5), 1238–1252. https://doi.org/10.1093/gbe/evu099
Tate, J. A., & Simpson, B. B. (2004). Breeding system evolution in *Tarasa* (Malvaceae) and selection for reduced pollen grain size in the polyploid species. American Journal of Botany, 91(2), 207–213. https://doi.org/10.3732/ajb.91.2.207
Thomson, G. M. (1922). The naturalisation of animals & plants in New Zealand. Cambridge University Press.
van Kleunen, M., Dawson, W., & Maurel, N. (2015). Characteristics of successful alien plants. Molecular Ecology, 24(9), 1954–1968. https://doi.org/10.1111/mec.13013
van Kleunen, M., Röckle, M., & Stift, M. (2015). Admixture between native and invasive populations may increase invasiveness of *Mimulus guttatus*. Proceedings of the Royal Society B: Biological Sciences, 282(1815), Article 20151487. https://doi.org/10.1098/rspb.2015.1487
Webb, C. J., Sykes, W. R., & Garnock-Jones, P. J. (1988). Flora of New Zealand, Volume IV: Naturalised pteridophytes, gymnosperms, dicotyledons. Botany Division, Department of Scientific and Industrial Research.
Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag. https://ggplot2.tidyverse.org
Yang, C.-F., & Guo, Y.-H. (2004). Pollen size-number trade-off and pollen-pistil relationships in *Pedicularis* (Orobanchaceae). Plant Systematics and Evolution, 247(3/4), 177–185.
Zaller, J. G. (2004). Ecology and non-chemical control of *Rumex crispus* and *R. obtusifolius* (Polygonaceae): A review. Weed Research, 44, 414–432. https://doi.org/10.1111/j.1365-3180.2004.00416.x

Chapter 4

Rumex species co-occurrence and hybrid occurrence lower in the introduced range compared to the native range

4.1. Introduction

Interspecific hybridisation is an important mechanism where new allele combinations are formed when two distinct species interbreed (Lee, 2002; Ellstrand & Schierenbeck, 2006). Hybridisation can aid in invasions by introducing new genetic material to help in adaptation and by purging deleterious alleles that can accumulate, especially in small populations (Ellstrand & Schierenbeck, 2006; Mesgaran et al., 2016). Indeed, there are many cases where interspecific hybrids, either between two introduced species or between an introduced and a native species, are more invasive than either parent species (Abbott et al., 2008; Ainouche et al., 2009; Walls, 2010; Mesgaran et al., 2016).

An important consequence of hybridisation is introgression or the transferring of alleles between parent species via a hybrid intermediate (Rieseberg & Carney, 1998). Introgression resulting from backcrossing between hybrids and the parent species can introduce new alleles and increase the genetic variation present in the parent species (Anderson & Stebbins, 1954; Barker et al., 2018). Introgression can provide a long-term benefit (adaptive introgression) or it can detrimental due to outbreeding depression where the hybrid progeny have lower fertility than the parents (Rhymer & Simberloff, 1996). However, where new hybrids possess beneficial alleles that help them to adapt to the environment, introgression back to the parent species will occur (Barker et al., 2018). This, in turn, can increase the invasiveness of the parent species by broadening their environmental tolerance or increasing their growth rate (Whitney et al., 2010). The increased adaptation potential can help established species respond to new environments but can be especially important when introduced species lack sufficient genetic variation (Ward et al., 2008).

Three species belonging to the genus *Rumex* L. (Polygonaceae), *R. conglomeratus* Murr., *R. obtusifolius* L., and *R. crispus* L., are among some of the world's worst weeds (Holm, 1977). They are also known to hybridise freely and often with each other in their native range (Cavers & Harper, 1964; Clapham et al., 1987) and some of the hybrids can become weeds that are difficult to manage (Ziburski et al., 1986; Bond et al., 2007). Out of these, *R. obtusifolius* and *R. crispus* are known for their high fitness hybrid, *R.* × *pratensis* Mert. & W.D.J. Koch. The hybrids are also known to introgress back to the parent species even when hybrid numbers are low, both in their native (Williams, 1971) and their introduced ranges (Takahashi & Hanyu, 2015; Uemura et al., 2022). Williams (1971) found introgression from *R. obtusifolius* to *R. crispus* to be more common than vice versa but found evidence of introgression both ways.

All three species were introduced to New Zealand (NZ) from the UK in the mid-1800s as grass seed contaminants (Thomson, 1922; Allan, 1937) and have since naturalised and become persistent weeds, especially in agricultural landscapes (Esler & Astridge, 1987). They are wind-pollinated, usually perennial species, primarily found in pastures, wastelands, roadside ditches and riversides (Cavers & Harper, 1964; Clapham

et al., 1987; Akeroyd, 2014). These species are good invaders due to their high seed production, up to 60 000 seeds a year, and difficult to eradicate because of their tap root and ability to grow from root fragments (summarised in Zaller, 2004).

In contrast to the UK where hybrids among the three *Rumex* species are common (Stace et al., 2015), there is only one brief mention of any hybrids in NZ (Allan, 1929). There are no vouchered specimens, and these hybrids are not recorded as naturalised in the country (Howell & Sawyer, 2006). The absence of any hybrids either reflect a significant oversight in the botanical recording in NZ or that opportunities for hybridisation are more limited. Limited opportunities for hybridisation might reflect a stronger propensity for selfing in NZ populations, lower viability of hybrid seed, or shifts in ecological niche that results in infrequent co-occurrence of two or more parental species. To determine the reasons for an apparent absence of hybrids in NZ, this study asked:

1) How frequently do parent species co-occur in close proximity in NZ compared to that in the UK?
2) Are there hybrids in NZ and what determines the frequencies of hybrids in a population?
3) Is there any evidence of introgression?

I used a combination of field surveys and vegetation survey results to answer to these questions. The species identifications were made using plant morphology and flow cytometry.

4.2. Materials & Methods

4.2.1. *Rumex* spp.

This study looked at three *Rumex* species: *R. obtusifolius*, *R. crispus* and *R. conglomeratus*, and their hybrids (TABLE 4.1). The parent species are common throughout NZ. The species differ from each other in their leaf morphology, such as leaf base and margin, as well as fruit characteristics (Akeroyd, 2014). All of the hybrids are intermediate to the parents in these characteristics (Stace et al., 2015); however, later generation hybrids and backcrosses can be difficult to identify based simply on their leaf and fruit morphology (Williams, 1971; Takahashi & Hanyu, 2015). The species also differ in their chromosome numbers and thus genome sizes. According to Löve (1942), the somatic (2*n*) chromosome numbers for *R. conglomeratus*, R. *obtusifolius* and *R. crispus* are 20, 40 and 60, respectively, and chromosome sizes 1.5 pg, 2.6 pg and 4.4 pg, respectively (Bai et al., 2012; Zonneveld, 2019). The hybrids, again, are intermediate between the parents, for example, *R.* × *pratensis* has a genome size of 3.72 pg (Zonneveld, 2019). These chromosome numbers and genome sizes were confirmed with NZ material in Chapter 2. This makes flow cytometry a valuable tool on top of morphology to confirm parent and hybrid identification and potentially separate

introgressed individuals that have started to resemble their parents more closely in their genome size.

The three possible hybrids, *R.* × *pratensis*, *R.* × *abortivus* Ruhmer, and *R.* × *schulzei* Hausskn. are varying in their distribution within the native range (FIGURE 4.1), with *R.* × *pratensis* being the most common of the three and of all *Rumex* hybrids in general (Stace et al., 2015). This likely reflects the co-occurrence of the parent species, with *R. obtusifolius* and *R. crispus* having a wider distribution than *R. conglomeratus*, thus providing more opportunities for hybridisation. All of the three hybrids are reported to have low fertility, and if seeds are produced, they have decreased viability (Stace et al., 2015). Despite this, introgression is known to happen (Williams, 1971; Takahashi & Hanyu, 2015; Uemura et al., 2022). Both *R.* × *pratensis* and *R.* × *schulzei* have been reported from Australia where parents co-occur (Rechinger, 1984), but these hybrids are not considered naturalised due to their low viability and thus inability to persist past first generation (Keighery & Longman, 2004). Allan (1929) mentioned all three hybrids, but only *R.* × *pratensis* appears in more than one location. However, Allan specifies that while these hybrids are observed, they have yet to be studied in detail. No other records were found from literature or herbarium records.

TABLE 4.1: Three possible hybrids between *Rumex conglomeratus*, *R. obtusifolius*, and *R. crispus*, and their genome sizes and chromosome numbers. Genome sizes and chromosome numbers are obtained by averaging the genome sizes and chromosome numbers of the parent species (see Chapter 2), apart from *R.* × *pratensis*, where references exist.

Parent species	*R. crispus*	*R. conglomeratus*
R. obtusifolius	***R.* × *pratensis*** Genome size 3.72 pg (Zonneveld, 2019) Chromosome number around 50 (Ziburski et al., 1986; Stace et al., 2015) Can be more vigorous than the parents (Williams, 1971) Found in Australia (Rechinger, 1984)	***R.* × *abortivus*** Genome size around 2 pg, Chromosome number around 30
R. conglomeratus	***R.* × *schulzei*** Genome size around 3 pg Chromosome number around 40 Found in Australia (Rechinger, 1984)	

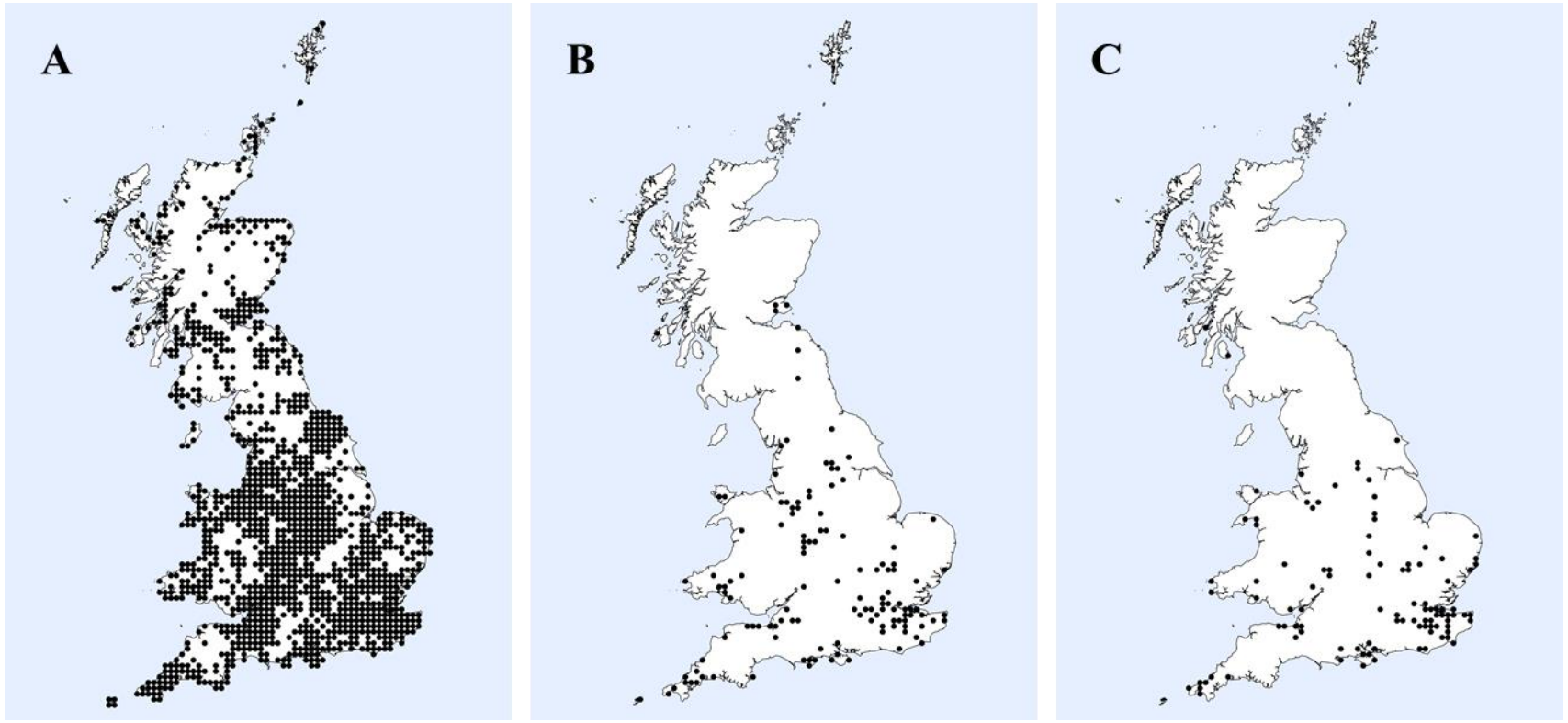

FIGURE 4.1: Hybrid occurrence in Britain. A): *Rumex* × *pratensis*; B): *R.* × *abortivus*; C) *R.* × *schulzei*. Distribution maps courtesy of Botanical Society of Britain & Ireland (retrieved from https://bsbi.org/, June 2022).

4.2.2. Parent species co-occurrence in survey records in the native and introduced ranges

National Vegetation Surveys (NVS) and UKCEH Countryside Survey (CS) records were used to access information on parent species co-occurrence within NZ and the UK, respectively. The NVS is a databank maintained by Manaaki Whenua - Landcare Research and it contains information from vegetation plots for over 70 years (Manaaki Whenua - Landcare Research, n.d.). It lists both exotic and native plants and covers a broad range of habitats. Similarly, the CS, maintained by the Centre for Ecology & Hydrology, collects information from vegetation plots spanning across a wide range of habitats throughout the UK (Centre for Ecology & Hydrology, n.d.). In addition, both surveys are used to inform about changes in biodiversity, as well as collect additional information, such as soil quality data. The NVS adds data from various projects each year, whereas the CS is done roughly every five years and it uses more permanent plots.

A search was made through the NVS website for records of two or more *Rumex* parent species co-occurring in close proximity, making hybridisation possible. The parent species *R. conglomeratus*, *R. obtusifolius*, and *R. crispus* were included as search words, and a manual search was carried out on all datasets reporting a match. Datasets were then split into categories based on the list of reported species (one or multiple *Rumex* identified) and access (open access or permission required). A more thorough search was then carried out to quantify the species co-occurrence within 20 × 20 m plots in datasets with multiple *Rumex,* and requests were made to access private datasets. Parent species co-occurrence was also estimated in UK using a CS from 2007. This CS comprised of over 17 000 surveyed plots which are mapped within 1000 × 1000 m squares all around UK and are of varying sizes depending on the surveyed vegetation type (for example river and stream edges were assessed in 1 × 100 m plots).

Since NVS and CS differ slightly in their plot sizes, parent species co-occurrence was calculated separately for all the plot size classes in CS to see the average variation across all size classes. The likelihood for species co-occurrence was calculated by dividing the number of plots with two or more *Rumex* species with the total number of *Rumex* observations over all plots. These were then compared between NZ and UK to see whether co-occurrence, and thus likelihood for hybridisation, is similar in the introduced range compared to that of the native range.

4.2.3. Field survey methods

To assess the frequency of hybridisation, mixed populations comprising of more than one parent species were selected. Leaf samples were collected from each population around Canterbury and the West Coast, NZ, during the summers of 2019–2020 and 2020–2021. A population was deemed suitable for the study if it had more than one species and a plant density high enough for obtaining a sufficient number of samples. In practice, this meant at least a plant every two meters, to get a minimum of 25 plants along a 50 meter transect line. This ruled out many of the visited sites, as the density of the population was much lower and/or only one species was present. Some of the single species populations were included in the sampling as controls.

On all of the sites, one 50-m transect line or multiple shorter lines adding up to 50 m were placed along the major axis of the population, and 25 contiguous 2 × 2 m plots were mapped along the transect. All *Rumex* plants in each quadrat were counted on the transect line and *Rumex* species were identified based on morphological characteristics to get an estimate of both species density and hybrid frequency. If the rosettes or stems were roughly within 10 cm from each other and were identified as the same species, they were classified as clones and only counted once as multiple stems can grow from the roots (Bond et al., 2007). Then leaf samples were collected for flow cytometric analyses from up to 20 individuals per parent species and per hybrid class.

Identification based on morphology followed characteristics listed in multiple references, for example, Ziburski et al. (1986) and Jensen (Figure 12, 1936). However, since flow cytometry requires fresh, green material, identification based on fruit valve or perianth morphology was not possible. Instead, identification was based on rosette leaf morphology, and if the plants were flowering, the fullness of inflorescences and the angle of branches were taken into account (FIGURE 4.2).

To see if parent species proportions within a population affected the hybrid frequency, parent species proportions were plotted against hybrid percentage over the whole population. In addition, seed weights were compared between the two parent species and their hybrid, as seed weight has been found to correlate with seedling survival (summarized in Leishman et al., 2000), size (Jakobsson & Eriksson, 2000), and establishment (Leishman et al., 2000; Murray et al., 2005), especially under stress. Weight differences were analysed with Kruskal-Wallis rank sum test, as the variances were not equal. Pairwise comparisons were then done using Wilcoxon rank sum test to see which groups of species differed from each other.

FIGURE 4.2: A picture showing the differences in the fullness of the inflorescence in *Rumex obtusifolius*, *R. crispus*, and *R. conglomeratus* as pictured from left to right.

4.2.4. Flow cytometry

Flow cytometry was performed to estimate the genome size of the *Rumex* species. Fresh, mature and healthy-looking leaf samples were harvested on the same day, or a maximum of 48 hrs, before flow cytometric analyses. After collection, samples were kept in a cooler box and then in fridge at +4 ºC until analysed. For analyses, the method from Dolezel et al. (2007) was used with slight modification. After screening for suitable standards with known genome size, corn (*Zea mays*, genome size 5.33 pg) and pea (*Pisum sativum*,

genome size 8.80 pg) were used as standards. Selecting the standard was dependent on the availability and condition of the standard plants in any given day and the assumed taxonomic identity of the samples, as with *R. conglomeratus* the genome of corn is closer in size and thus more accurate, whereas *R. crispus* and corn have genome sizes too similar and thus pea is a more appropriate standard.

A small piece of leaf, around 1 cm^2, from both the standard and one of the *Rumex* species was placed in a Petri dish in ice-cold Otto I isolation buffer, and the samples were then quickly and carefully chopped with a razor blade to avoid crushing the cells. Samples were incubated at room temperature for 2 min before filtering the homogenate through 20 µm mesh into a test tube, after which propidium iodide (10 µg/ml) stain was added. Samples were further incubated at room temperature for another 2 min, before analysing them with Partec CyFlow Space instrument (Partec GmbH., Münster, Germany) until a sufficient number of nuclei were examined and clearly separated fluorescence intensity histograms formed. Using FloMax software, the fluorescent intensity was plotted as a histogram. Along a linear fluorescent scale, fluorescence vs side scatter was plotted, and to help isolate fluorescence data of interest, polygons were drawn around the formed scatterplot to exclude non-intact cells (Bainard et al., 2012). After that, peaks were gated, Gauss peak analysis was performed, and mean peak values and CV% were recorded. Then, the sample genome size was estimated by calculating the ratio between the standard (corn or pea) and *Rumex* sample fluorescence intensity peaks and dividing the known genome size of the standards by that ratio (Doležel et al., 2007).

4.3. Results

4.3.1. Parent species co-occurrence lower in New Zealand compared to the UK

A search within NVS resulted in 132 datasets that included *R. conglomeratus*, *R. obtusifolius*, or *R. crispus* in the species list. These datasets consisted of over 25 000 plots of 20 × 20 m distributed throughout NZ and ranged from low to high altitudes. Habitats surveyed ranged from urban to wetlands to lake- and riversides. Out of these 132 surveys, 21 were freely accessible and the rest required permission. Because even the restricted datasets provided access to the species lists while omitting plot level information, a search was conducted to see how many species were listed in each survey. Out of these, only six datasets listed two or more species and since access was not granted to view plot level information, they were left out of this study.

After removing surveys which studied same plots in multiple years and the six datasets with no access, 76 surveys were kept. These consisted of 21 506 surveyed plots, 403 plots with *Rumex* and 9 plots with two or more species co-occurring. Thus 2.2 % of the total number of plots had a *Rumex* plant observation. This comes with a caveat, however, as often *Rumex* species were recorded just at genus level instead of species level, with the exception of *R. acetosella*, which occurs in over 10 000 plots within 870 datasets. However, *R. acetosella* is outside of the scope of this study due to its altogether different chromosome system (Kihara & Ono, 1926). The search "*Rumex* species" yields

more datasets and more plots than all three study species combined, in total 8 745 plots over 883 datasets. In some cases, a study had both unidentified *Rumex* species and one of the study species. The identification "*Rumex* species" was observed in 17 % of the total number of plots.

The CS had 17 000 plots in total, out of these 2 929 plots had a *Rumex* plant listed and out of these 278 plots had two or more species co-occurring. Co-occurrence in the UK was around 8.6 % over all different plot sizes. This varied from 5 % co-occurrence within 2 × 2 m plots to 8 % in 10 × 1 m plots, 0 % in 30 × 1 m plots, 16 % in 100 × 1 m plots, and 14 % within 200 × 200 m plots.

A chi-squared contingency test was performed to compare differences between plots with no *Rumex*, one *Rumex*, and *Rumex* co-occurring in NZ compared to UK. As plot sizes between countries are not equal, data from 2 × 2, 10 × 1, 30 × 1 and 100 × 1 m plots were combined from CS and compared to data from 20 × 20 m plots from NVS. NZ has significantly lower co-occurrence compared to the UK, and I found significantly fewer individuals of *Rumex* species and lower frequency of co-occurrence than expected given the bigger plot size and more surveyed plots ($\chi^2 = 2737$, df = 2, $p < 0.05$).

4.3.2. Field surveys identify two *Rumex* hybrids in New Zealand

Samples were collected from 20 different sites, which included multiple habitats, namely stream and river edges, roadside ditches, paddocks, and wastelands. Between 25 to over 300 plants were identified per site (mean = 124 plants), with almost 2 500 plants identified in total based on their morphology. The most frequent and abundant parent species was *R. obtusifolius*, and two of the three possible hybrids were found – these were *R.* × *pratensis* and *R.* × *abortivus*.

Hybrids were present in all of the sites where two or more parent species co-occurred (n = 16), and on one site where only one parent was found on the transect line along with the hybrid. Hybrid abundance within the sampled populations ranged from 0 % to 33 %, averaging in around 5.7 % (SD 7.81). No clear relationship between the parent species proportions and the hybrid abundance was found (linear regression: F = 1.269, df = 18, $p > 0.27$; FIGURE 4.3). In addition, the hybrid seeds were significantly smaller from that of either of the parents (Kruskal-Wallis rank sum test: $\chi^2 = 19.672$, df = 2, $p < 0.05$; FIGURE 4.4).

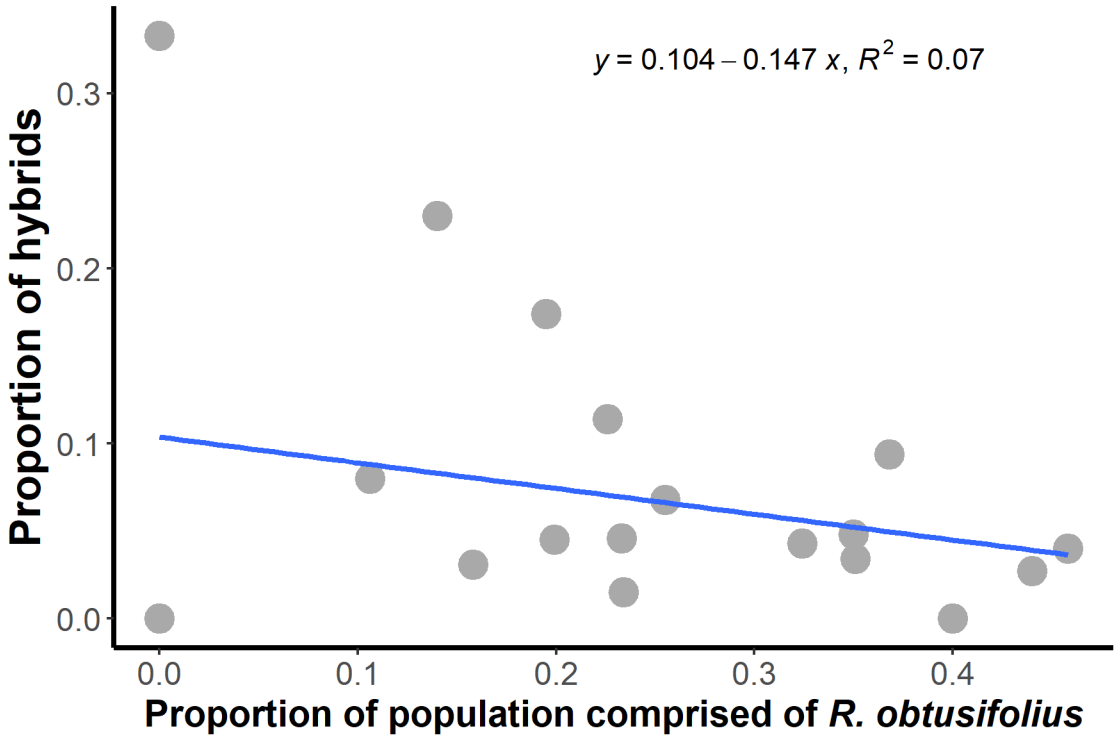

FIGURE 4.3: Proportion of the population comprised of *Rumex obtusifolius* individuals to the proportion of hybrid plants within a population. Regression line equation is given in the figure.

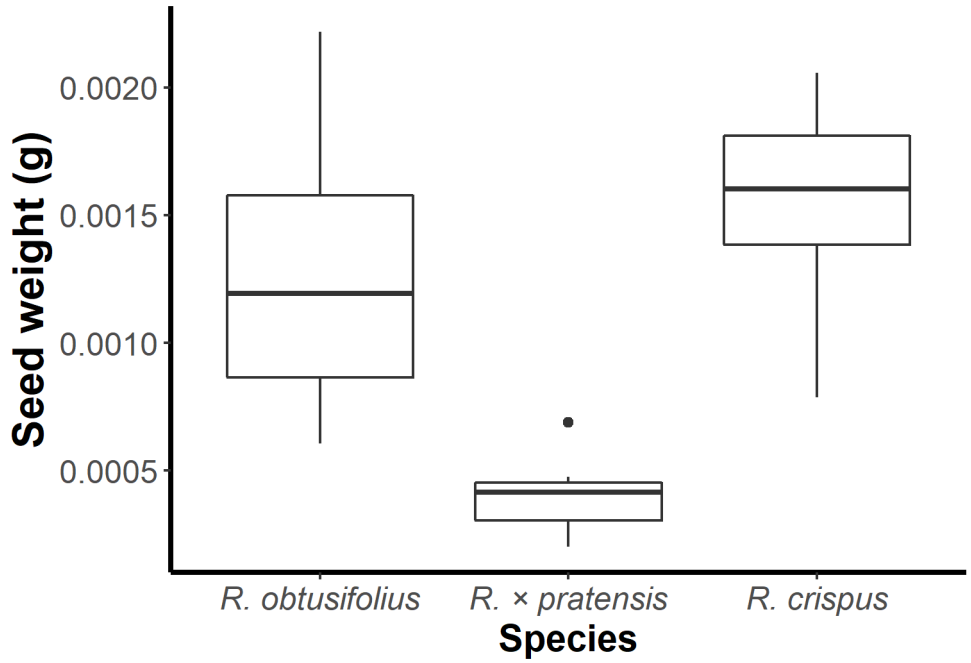

FIGURE 4.4: Boxplot of seed weights in *Rumex × pratensis* and its parents, *R. obtusifolius* and *R. crispus*. The box represents the interquartile range with the median represented with a thick line and the whiskers extending from the 5th to the 95th percentile, with values outside this range represented as dots.

4.3.3. Most hybrids have intermediate genome size compared to the parent species

Around 850 plants were identified using flow cytometry, and similarly to the morphological identifications, the majority of these samples were from *R. obtusifolius*. Similar to Chapter 2, the observed genome sizes matched previous reports by Zonneveld (2019) but differed from old records in the databases (see detailed description in Chapter 2). There was minor variation in the genome sizes within the species, possibly highlighting the introgressed individuals (FIGURE 4.5). The CV% values in this study were slightly higher than optimal (> 5), but as peaks were rather even, the geometric mean of a peak should stay roughly the same, even if peak itself is wide (G. Houliston, personal communication, 2020).

From the samples analysed with flow cytometry, 85.6 % were identified correctly based on their morphology. In most incorrect identifications, a plant was identified as a hybrid when in fact it was a true species based on the genome size. Conversely, around 5 % of the hybrids based on their genome size were misidentified based on the morphology. There was a strong association between the identification of hybrids using morphological and flow cytometry methods ($\chi^2 = 177.11$, df = 1, $p < 0.05$).

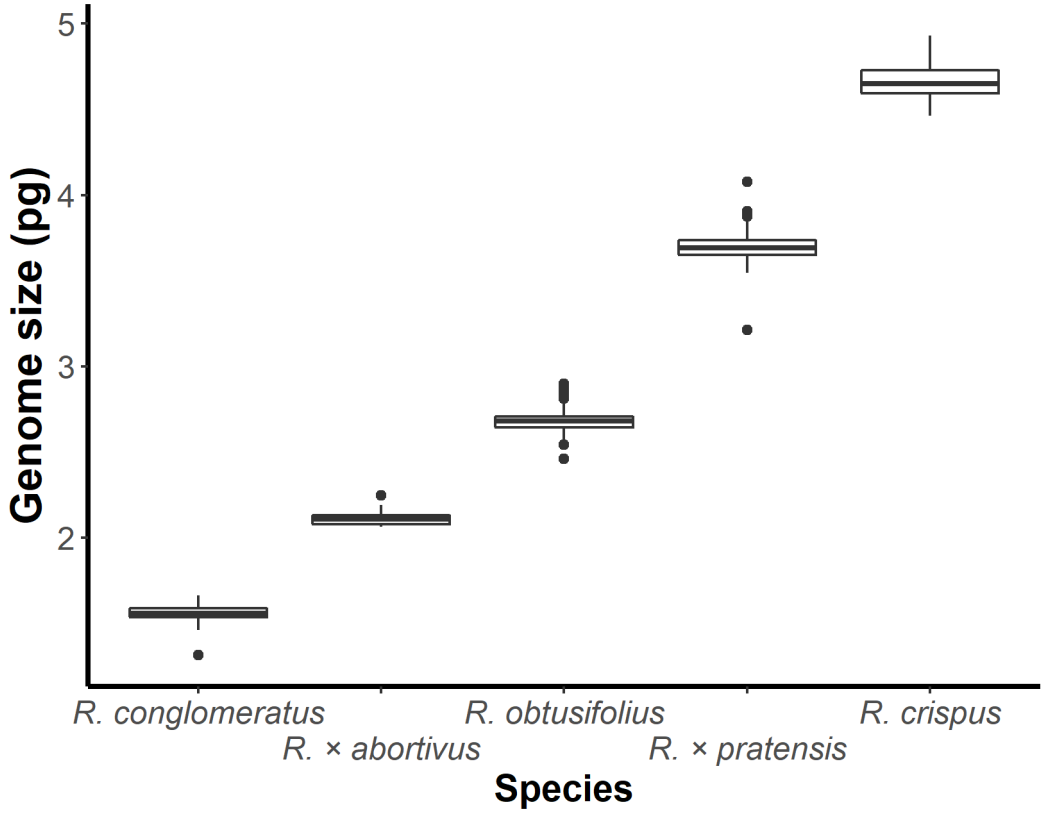

FIGURE 4.5: Boxplots displaying the variation in genome sizes in *Rumex* species and their hybrids. Hybrids are placed between the parent species to highlight the intermediate genome size compared to the parents. Samples with CV% > 10 are filtered out. The box represents the interquartile range with the median represented with a thick line and the whiskers extending from the 5th to the 95th percentile, with values outside this range represented as dots.

4.4. Discussion

The three *Rumex* species are known to hybridise freely (Akeroyd, 2014; Stace et al., 2015), and since their niche requirements overlap, hybridisation is known to be a relatively common phenomenon in their native range. To this day, there are no confirmed records of hybrids in NZ. However, as hybridisation can be an important driver behind invasiveness (Ellstrand & Schierenbeck, 2006), I set out to investigate the parent species co-occurrence within vegetation survey records from the UK and NZ, and the hybrid plant occurrence in the wild in the South Island, NZ. The parent species co-occurrence was lower in NZ compared to the UK according to the vegetation survey records. The field surveys showed that hybrids do exist whenever parent species co-occur. This study found two out of the three possible hybrids. The hybrids had an intermediate genome size compared to the parents, and thus were likely first-generation hybrids. This indicates that introgression is rare and likely does not contribute to the increased invasiveness of these species in NZ.

4.4.1. Parent species co-occur less in the introduced range

For hybridisation to occur, gene flow must happen between two parent species. At the scale of 2 × 2 m plots, *Rumex* species capable of hybridising occurred at twice the frequency in the UK than in NZ. Co-occurrence may have been underestimated in the NVS records, as often *Rumex* species were not identified to the species level, but my own field surveys highlight the rarity of co-occurrence. The majority of the visited field sites were unsuitable for the field surveys, as the populations consisted of only one species or only few individuals.

4.4.2. Hybrids are present where parents are co-occurring

Two hybrids, *R.* × *abortivus* and *R.* × *pratensis*, were found in several field sites in NZ. The overall hybrid frequency, however, was low, on average around 6 %, even with relatively even parent species frequencies within the surveyed field sites. This frequency of hybrids when both *R. obtusifolius* and *R. crispus* are present is similar (3–10.3 %) to that found in the native range (Ziburski et al., 1986). Surprisingly, a similar frequency of the two parents within a population did not lead to an increase in the hybrid frequency (FIGURE 4.3), and hybrid frequency was independent of parental frequency.

Low hybrid frequency could be partially explained by high selfing rates reported in the literature. Both *R. obtusifolius* and *R. crispus* are selfing: In the UK, it is estimated that *R. crispus* populations can be nearly fully selfing (Cavers & Harper, 1964; Akeroyd & Briggs, 1983; Bond et al., 2007). With higher selfing rates it is understandable that opportunities for hybridisation are rare, even when the species co-occur. Selfing rates have not been studied in NZ, and it is possible that the rate differs between the two ranges. However, the genetic analyses found low between population differentiation in both provenances and all three species (Chapter 5). This suggests of relatively high percentage of outbreeding, whereas high selfing would increase the differentiation between populations. Nevertheless, another study investigating the costs from selfing in *R.*

conglomeratus showed no decrease in seed size or viability when compared to out-crossed progeny (Chapter 2). In addition, seed weights showed a clear difference between parent seed and hybrid seed weights. With lower seed weight and a decrease in seed formation, the germination of hybrid plants can be lower than that of the parent species. However, seed size is found to have varying effects in *Rumex*: small seed size can decrease the germination in *R. crispus* but actually increase germination in *R. obtusifolius* (Cideciyan & Malloch, 1982).

Rumex hybrids do not experience post-zygotic barriers (Ziburski et al., 1986), but instead, the hybrid offspring themselves often appear sterile (Cavers & Harper, 1964). Low hybrid pollen viability could hinder introgression opportunities, especially with lower hybrid frequency. In addition, most of the hybrids found had an intermediate genome size compared to that of the parents, which indicates that they are likely first-generation hybrids (Williams, 1971; Ziburski et al., 1986). There were some outliers, and these individuals – based on the genome size – could be later generation introgressed hybrids, as they were intermediate in their genome size to a parent species and a F2 hybrid. This is, however, difficult to quantify without manual chromosome counts.

4.4.3. Identifying plants to species level is not always easy

Around 85 % of the identifications based on leaf morphology were correct. Based on morphology of the leaves only, misidentifying the species is likely, especially if the plants are still young and fruit characteristics cannot be used for identification. This result highlights the importance of more than one identification method. It is especially true with introgressed individuals, as they can match one of the parents both in morphology and sometimes in genome size as well (Williams, 1971). Ziburski et al. (1986) found that 4 out of 11 studied second generation *R.* × *pratensis* hybrids had chromosome number of 40, matching the parent *R. obtusifolius*, and a few other individuals had chromosome numbers close to this. The authors mention that some of the F2 hybrids were also morphologically indistinguishable from *R. obtusifolius* (Ziburski et al., 1986). This suggests that even with flow cytometry, there can be some underestimation of the true number of hybrids and especially of the introgressed individuals. Consequently, it is important to note that around 5 % of the hybrids were missed based solely on plant morphology, making it likely that the hybrids have been missed in the past as well. Especially since those 5 % were missed even when actively looking for the hybrids, and with a specialised knowledge of the *Rumex* species in NZ.

Since hybrids were encountered in 85 % of the surveyed populations and in all populations where parents were co-occurring, it seems likely that they have been overlooked in the past. There can be multiple reasons behind this, such as not requiring a species level identification or not being able to identify plants correctly based on morphology due to lack of knowledge of the genus or wrong time of the year (i.e., very young seedlings only). There are also multiple subspecies within all species in the genus (Cavers & Harper, 1964; Williams, 1971; Bond et al., 2007) and sometimes the species can be hard to distinguish from each other especially given similar habitat requirements. However, limited knowledge of the subspecies present in NZ exists, and thus I have not tried to distinguish between them.

For the genus, the perianth morphology is the most accurate way to distinguish the species from each other, but inflorescences are not always present during field surveys. According to Williams (1971), the morphological characteristics form a continuum, and especially considering the backcrossed plants, dividing individuals into species can be difficult. This is pronounced if the surveyor has limited experience with these species. The challenges in identification generally coupled with the high frequency of plants identified only to the genus level in the NVS further highlights the likelihood of overlooking the hybrids in the past.

4.4.4. Conclusions

This study has shown that the three most common *Rumex* species do hybridise in NZ in the wild. However, the rate of hybridisation is low, and when hybrids were encountered, they made up only around 6 % of the populations, even when both parents co-occurred in similar numbers. The parent species have lower co-occurrence in the introduced range compared to the native range, and thus, opportunities for hybridisation were rarer compared to what is seen in the native range. Potential high rates of selfing and low quality of hybrid seeds can also help to explain the low degree of hybrids, since without outcrossing or with limited outcrossing, there is a low chance for hybridisation. In addition, with a lower viability in hybrid seeds or hybrid pollen, chances for introgression are even lower. However, the rates of selfing within NZ are not known, but based on the genetic analyses outbreeding rates seem to be relatively high (Chapter 5). Nevertheless, selfing experiment with *R. conglomeratus* showed that selfing does not seem to lower viability, rather vice versa (Chapter 3). Furthermore, even with relatively low frequency, hybridisation and introgression can benefit introduced species and increase their invasiveness, especially in conjunction with plasticity. To confirm how and if hybridisation can aid invasiveness or increase adaptation potential, viability, and fertility of the hybrids, both pollen and seeds should be examined, along with conducting field surveys more widely within the introduced provenance.

4.5. References

Abbott, R. J., Brennan, A. C., James, J. K., Forbes, D. G., Hegarty, M. J., & Hiscock, S. J. (2008). Recent hybrid origin and invasion of the British Isles by a self-incompatible species, Oxford ragwort (*Senecio squalidus* L., Asteraceae). Biological Invasions, 11(5), Article e1145. https://doi.org/10.1007/s10530-008-9382-3

Ainouche, M. L., Fortune, P. M., Salmon, A., Parisod, C., Grandbastien, M.-A., Fukunaga, K., Ricou, M., & Misset, M.-T. (2009). Hybridization, polyploidy and invasion: Lessons from *Spartina* (Poaceae). Biological Invasions, 11(5), 1159–1173. https://doi.org/10.1007/s10530-008-9383-2

Akeroyd, J. R. (2014). Docks and knotweeds of Britain and Ireland (2nd ed.). Botanical Society of Britain & Ireland.

Akeroyd, J. R., & Briggs, D. (1983). Genecological studies of *Rumex crispus* L. I. Garden experiments using transplanted. New Phytologist, 94(2), 309–323.

Allan, H. H. (1929). A list of supposed wild hybrids among the naturalized plants of New Zealand. New Zealand Journal of Science and Technology, 11(260), 255–261.

Allan, H. H. (1937). The origin and distribution of the naturalized plants of New Zealand. Linnean Society of London, 150(1), 25–46.
Anderson, E., & Stebbins, G. L. (1954). Hybridization as an evolutionary stimulus. Evolution, 8(4), 378–388. https://doi.org/10.2307/2405784
Bai, C., Alverson, W. S., Follansbee, A., & Waller, D. M. (2012). New reports of nuclear DNA content for 407 vascular plant taxa from the United States. Annals of Botany, 110(8), 1623–1629. https://doi.org/10.1093/aob/mcs222
Bainard, J. D., Bainard, L. D., Henry, T. A., Fazekas, A. J., & Newmaster, S. G. (2012). A multivariate analysis of variation in genome size and endoreduplication in angiosperms reveals strong phylogenetic signal and association with phenotypic traits. New Phytologist, 196(4), 1240–1250. https://doi.org/10.1111/j.1469-8137.2012.04370.x
Barker, B. S., Cocio, J. E., Anderson, S. R., Braasch, J. E., Cang, F. A., Gillette, H. D., & Dlugosch, K. M. (2018). Potential limits to the benefits of admixture during biological invasion. Molecular Ecology, 28(1), 100–113. https://doi.org/10.1111/mec.14958
Bond, W., Davies, G., & Turner, R. (2007). The biology and non-chemical control of broad-leaved dock (*Rumex obtusifolius* L.) and curled dock (*R. crispus* L.) [Internal Report]. HDRA. https://www.gardenorganic.org.uk/sites/www.gardenorganic.org.uk/files/organic-weeds/dock-review.pdf
Cavers, P. B., & Harper, J. L. (1964). *Rumex obtusifolius* L. and *R. crispus* L. Journal of Ecology, 52(325), 737–766.
Centre for Ecology & Hydrology. (n.d.). Countryside Survey. Countryside Survey. Retrieved May 28, 2021, from https://countrysidesurvey.org.uk/
Cideciyan, M. A., & Malloch, A. J. C. (1982). Effects of seed size on the germination, growth and competitive ability of *Rumex crispus* and *Rumex obtusifolius*. Journal of Ecology, 70(1), 227–232. https://doi.org/10.2307/2259875
Clapham, A. R., Tutin, T. G., & Moore, D. M. (1987). Flora of the British Isles (3rd ed.). Cambridge University Press.
Doležel, J., Greilhuber, J., & Suda, J. (2007). Estimation of nuclear DNA content in plants using flow cytometry. Nature Protocols, 2(9), 2233–2244. http://dx.doi.org/10.1038/nprot.2007.310
Ellstrand, N. C., & Schierenbeck, K. A. (2006). Hybridization as a stimulus for the evolution of invasiveness in plants? Euphytica, 148, 35–46. https://doi.org/10.1007/s10681-006-5939-3
Esler, A. E., & Astridge, S. J. (1987). The naturalisation of plants in urban Auckland, New Zealand 2. Records of introduction and naturalisation. New Zealand Journal of Botany, 25(4), 523–537. https://doi.org/10.1080/0028825X.1987.10410084
Holm, L. G. (1977). The world's worst weeds: Distribution and biology. University Press of Hawaii.
Howell, C. J., & Sawyer, J. W. D. (2006). New Zealand naturalised vascular plant checklist. New Zealand Plant Conservation Network. https://www.nzpcn.org.nz/site/assets/files/0/13/474/naturalised-list-06-new.pdf
Jakobsson, A., & Eriksson, O. (2000). A comparative study of seed number, seed size, seedling size and recruitment in grassland plants. Oikos, 88(3), 494–502. https://doi.org/10.1034/j.1600-0706.2000.880304.x
Jensen, H. W. (1936). Meiosis in *Rumex*. I. Polyploidy and the origin of new species. Cytologia, 7(1–2), 1–22.
Keighery, G., & Longman, V. (2004). The naturalized vascular plants of Western Australia 1: Checklist, environmental weeds and distribution in IBRA regions. Plant Protection Quarterly, 19(1), 12–32.
Kihara, H., & Ono, T. (1926). Chromosomenzahlen und systematische Gruppierung der *Rumex*-Arten. Zeitschrift für Zellforschung und Mikroskopische Anatomie, 4(3), 475–481. https://doi.org/10.1007/BF00391215
Lee, C. E. (2002). Evolutionary genetics of invasive species. Trends in Ecology & Evolution, 17(8), 386–391. https://doi.org/10.1016/S0169-5347(02)02554-5

Leishman, M., Wright, I. J., Moles, A., & Westoby, M. (2000). The evolutionary ecology of seed size. In M. Fenner (Ed.), Seeds: The ecology of regeneration in plant communities (pp. 31–59). CABI.
Löve, Á. (1942). Cytogenetic studies in *Rumex*. Hereditas, 28(3–4), 289–296. https://doi.org/10.1111/j.1601-5223.1942.tb03281.x
Manaaki Whenua - Landcare Research. (n.d.). New Zealand National Vegetation Survey Databank. Retrieved May 28, 2021, from https://nvs.landcareresearch.co.nz/
Mesgaran, M. B., Lewis, M. A., Ades, P. K., Donohue, K., Ohadi, S., Li, C., & Cousens, R. D. (2016). Hybridization can facilitate species invasions, even without enhancing local adaptation. Proceedings of the National Academy of Sciences, 113(36), 10210–10214. https://doi.org/10.1073/pnas.1605626113
Murray, B. R., Kelaher, B. P., Hose, G. C., Figueira, W. F., Leishman, M. R., & Eriksson, O. (2005). A meta-analysis of the interspecific relationship between seed size and plant abundance within local communities. Oikos, 110(1), 191–194.
Rechinger, K. H. (1984). *Rumex* (Polygonaceae) in Australia: A reconsideration. Nyutsia, 5(1), 75–122.
Rhymer, J. M., & Simberloff, D. (1996). Extinction by hybridization and introgression. Annual Review of Ecology and Systematics, 27, 83–109.
Rieseberg, L. H., & Carney, S. E. (1998). Plant hybridization. New Phytologist, 140(4), 599–624. https://doi.org/10.1046/j.1469-8137.1998.00315.x
Stace, C. A., Preston, C. D., & Pearman, D. (2015). Hybrid Flora of the British Isles. Botanical Society of Britain & Ireland.
Takahashi, K., & Hanyu, M. (2015). Hybridization between alien species *Rumex obtusifolius* and closely related native vulnerable species *R. longifolius* in a mountain tourist destination. Scientific Reports, 5, Article 13898. https://doi.org/10.1038/srep13898
Thomson, G. M. (1922). The naturalisation of animals & plants in New Zealand. Cambridge University Press.
Uemura, R., Asakawa, A., Fujii, S., Matsuo, A., Suyama, Y., & Maki, M. (2022). Can *Rumex madaio* (Polygonaceae) be threatened by natural hybridization with an invasive species in Japan. Nordic Journal of Botany, 2022(5), Article e03543. https://doi.org/10.1111/njb.03543
Walls, R. L. (2010). Hybridization and plasticity contribute to divergence among coastal and wetland populations of invasive hybrid Japanese knotweed s.l. (*Fallopia* spp.). Estuaries and Coasts, 33(4), 902–918. https://doi.org/10.1007/s12237-009-9190-8
Ward, S. M., Gaskin, J. F., & Wilson, L. M. (2008). Ecological genetics of plant invasion: What do we know? Invasive Plant Science and Management, 1(1), 98–109. https://doi.org/10.1614/IPSM-07-022.1
Whitney, K. D., Randell, R. A., & Rieseberg, L. H. (2010). Adaptive introgression of abiotic tolerance traits in the sunflower *Helianthus annuus*. New Phytologist, 187(1), 230–239. https://doi.org/10.1111/j.1469-8137.2010.03234.x
Williams, J. T. (1971). Seed polymorphism and germination. II. The role of hybridization in the germination polymorphism of *Rumex crispus* and *R. obtusifolius*. Weed Research, 11(1), 12–21. https://doi.org/10.1111/j.1365-3180.1971.tb00970.x
Zaller, J. G. (2004). Ecology and non-chemical control of *Rumex crispus* and *R. obtusifolius* (Polygonaceae): A review. Weed Research, 44, 414–432. https://doi.org/10.1111/j.1365-3180.2004.00416.x
Ziburski, A., Kadereit, J. W., & Leins, P. (1986). Quantitative aspects of hybridization in mixed populations of *Rumex obtusifolius* and *Rumex crispus* Polygonaceae. Flora, 178, 233–242.
Zonneveld, B. J. M. (2019). The DNA weights per nucleus (genome size) of more than 2350 species of the Flora of The Netherlands, of which 1370 are new to science, including the pattern of their DNA peaks. Forum Geobotanicum, 8, 24–78.

Chapter 5

Low genetic differentiation between the native and introduced ranges, but evidence of multiple introductions to New Zealand in three *Rumex* spp.

5.1. Introduction

The movement of people and commodities over the world has spread thousands of species to new areas and some of them have become invasive (Seebens et al., 2017). While many hypotheses exist to explain why some introduced species become invasive (Catford et al., 2009), an important subset addresses details relating to the degree of intraspecific genetic variation. However, from a genetic perspective, invasion represents a paradox (see for example Allendorf & Lundquist, 2003; Pérez et al., 2006; Schrieber & Lachmuth, 2016; Estoup et al., 2016). First, colonising populations often go through a drastic size reduction following introduction since usually only a handful of individuals colonize new areas. This should lead to decreased genetic variation (genetic bottleneck) and hence lower fitness. Greater genetic variation is often linked to better ability to adapt to new environment whereas small population size is linked to harmful chance effects (e.g., drift load) and inbreeding, which could further decrease genetic variation (Dlugosch & Parker, 2008). Thus, the success of introduced species challenges the idea that small populations and reduced genetic variation limits species expansion. Second, native populations are assumed much better adapted to their environments than introduced species, yet introduced species sometimes still outcompete them (Oduor et al., 2016).

Several reasons might help to explain this paradox. Bottlenecks may be insufficient to remove the important variation in quantitative traits (Dlugosch & Parker, 2008; Alexander et al., 2009; Barrett, 2015), whereas phenotypic plasticity (see for example Ward et al., 2008; Estoup et al., 2016) might explain the success of colonisations of some species, and admixture by multiple introductions could explain novel allele combinations and rapid adaptation (Barrett, 2015). Indeed, research has shown that often successful colonisers are introduced multiple times, thus keeping genetic variation at similar levels to the native range (Bossdorf et al., 2005; Roman & Darling, 2007; Estoup et al., 2016; Schrieber & Lachmuth, 2016). In addition, if species are introduced multiple times and especially if introductions are from multiple locations, this genetic admixture can lead to increased adaptation potential and create unintended consequences where species become invasive (Dlugosch & Parker, 2008; Keller & Taylor, 2010).

Intraspecific genetic admixture is a phenomenon where gene flow and interbreeding occur between two or more genetically distinct populations of the same species (Ellstrand & Schierenbeck, 2006; Rius & Darling, 2014). Intraspecific admixture can either benefit the population by increasing genetic variation or masking fixed deleterious alleles (Keller & Taylor, 2010; Verhoeven et al., 2011; van Kleunen et al., 2015), it can be neutral indicating that the populations do not suffer from deleterious alleles (Tayeh et al., 2013), or lead to outbreeding depression (Montesinos et al., 2012). These effects can be short-term or long-term, and admixture can in some cases boost the invasiveness of the species

(Verhoeven et al., 2011; Rius & Darling, 2014). De Carvalho et al. (2010) found admixture of differentiated postglacial lineages of common aspen *Populus tremula* L. (Salicaceae) across Europe and believe that it has contributed to adaptation by adding new alleles to the populations. Similarly, Vallejo-Marín et al. (2021) found that UK populations of common yellow monkeyflower *Erythrante guttata* (Fisch. ex DC.) G.L. Nesom (Phrymaceae) were admixed from various regional populations from the species' native range. These admixed plants were then spread further within the introduced range along with more introductions from the native range, further creating new intraspecific hybrids. Despite the ongoing research effort put into admixture studies, we do not know the exact role genetic variation plays in plant invasions (Hufbauer, 2017) and this makes it hard to know the consequences that multiple introductions could have for introduced species.

Several species of dock weeds, *Rumex* L. (Polygonaceae), were introduced to New Zealand (NZ) in the early to mid-1800s (Thomson, 1922). As they were often introduced as grass seed contaminants as well as in the soil of agricultural machinery, it is very likely that they were introduced multiple times from multiple locations. According to Thomson (1922), these early *Rumex* introductions are thought to have come from the UK with the early European settlers. Already in 1830s reports state that Māori horticultural fields were struggling to keep *Rumex* species at bay. And indeed, today *R. conglomeratus* Murr., *R. obtusifolius* L. and *R. crispus* L. are all naturalised and thriving, especially in agricultural habitats, but also in wastelands, roadsides and next to water bodies. These species are among the world's worst weeds (Holm, 1977), as they are fast growing, have high seed yields and are known to hybridise freely (e.g., Akeroyd, 2014). In NZ they were listed as noxious weeds and seeds in "The Noxious Weeds Act, 1900" (summarised in The Noxious Weed Act 1908). However, the extent to which the species experienced a bottleneck or have benefited from genetic admixture following introduction to NZ is unknown.

The three *Rumex* species have wind-pollinated hermaphrodite flowers, and are self-compatible (Les, 2017) but can also spread vegetatively from root and leaf pieces (Zaller, 2004), making them hard to eradicate. These species are reported as being highly variable and having several different ecotypes (Grime et al., 2007) as well as having polymorphic seeds with different germination requirements (Cavers & Harper, 1966; Assche et al., 2002). *Rumex* plants produce high numbers of seeds, up to 40 000 per year for *R. crispus* and for *R. obtusifolius* up to 60 000 per year (Cavers & Harper, 1964; Bufford & Hulme, 2021b), and the seeds can stay viable in the soil for tens of years (Zaller, 2004). Some studies report that the structure of *R. crispus* flowers prevents autogamy (self-pollination) (Grime et al., 2007), whereas others say that it is highly self-fertile, some populations even fully selfing (Cavers & Harper, 1964). In addition, while *R. conglomeratus* is a diploid, *R. obtusifolius* is an allotetraploid and *R. crispus* an allohexaploid, hybrids and subsequent backcrosses between all three species are nevertheless common (Williams, 1971; Takahashi & Hanyu, 2015).

To date, the population genetics of *Rumex* have received limited attention. Genetic methods using molecular markers, such as single nucleotide polymorphisms (SNPs), are often used to study population demographics: bottlenecks and founder effect, the level of admixture, as well as in- and outbreeding within and between populations and species. Genotyping-by-sequencing (GBS)(Elshire et al., 2011) utilises restriction enzymes to reduce genome complexity and can be used to obtain high numbers of SNPs. Reducing

genome complexity leads to higher depth of coverage, that is, the number of times a certain genomic area is sequenced is increased, which increases genotyping accuracy. Due to its speed and low cost, GBS is now a widely used method proven to be a useful tool both when there already is an existing reference genome sequence and for *de novo* discovery (Poland & Rife, 2012). GBS has been successfully used in population structure and admixture studies as well as in interspecies hybridisation studies (see for example Stetter et al., 2017; Zlonis & Gross, 2018; Pina-Martins et al., 2018). Recently, GBS was used to estimate the introduction pathway of common yellow monkeyflower *Erythrante guttata* using samples from the species' native and introduced ranges as well as information on the species' invasion history (Vallejo-Marín et al., 2021). However, GBS is still fairly rarely used in invasion biology despite the higher number of markers it provides and less prior knowledge it requires compared to microsatellite analysis (Grover & Sharma, 2016).

I compared plants belonging to the three *Rumex* species collected from their native range (UK, France, the Czech Republic, and two of them also from Finland) to plants within the introduced range (NZ) and asked:

1) Is there a difference in how genetic variation is partitioned within and between the native and introduced ranges?
2) Is there evidence of admixture within the introduced range such that genetic variation is greater than in the native range?
3) Could the UK be the source for the introductions in NZ?

The following potential scenarios were considered for the first question: variation is different in native and introduced regions, with the introduced region exhibiting lower variation likely as a result of bottlenecks; variation is different in native and introduced regions, with relatively high variation within the introduced region likely due to admixture and multiple introductions; variation is similar in both ranges due to admixture and multiple introductions. In addition, more differentiation within the native range was expected, given more time to adapt to local conditions and higher initial genetic variation. Furthermore, the following scenarios were considered for the third question: populations sampled in NZ are genetically similar to the populations in UK, which suggests UK as the source of the introductions; some but not all populations are genetically similar to the populations in the UK, which suggests that UK is one of the sources, but propagules originated elsewhere as well; the populations in NZ do not match that of the populations in UK, indicating that the UK is not a source of the introductions.

5.2. Methods

5.2.1. Sampling protocol

Seeds from all three species were collected from ten plants (seed families) in four populations within four climatically matching regions in NZ and the UK during 2016-2017 (Bufford & Hulme, 2021a, and a map with sampling locations in 2021b). In addition, seeds from up to four populations from one region were collected from each of

the Czech Republic (Středočeský, Prague area), France (Ile-De-France, Paris area), and Finland (Uusimaa, Helsinki area) in 2019 to get a broader picture from the native range and to assess potential introductions from mainland Europe to NZ (for sample numbers, see TABLE 5.1). These three countries will be referred as the outgroup from here on. Since *R. conglomeratus* is not present in Finland, only *R. obtusifolius* and *R. crispus* were collected from that country. Ten seeds from each seed family were germinated in a glasshouse at Lincoln University and leaf material from one seedling per seed family was sampled once the first true leaf had been produced. Each sample, thus, comes from one individual plant, and here I use the words sample, plant and individual interchangeably. Seeds were imported into NZ under Ministry for Primary Industries permits 201661142, C2019/466334, and C2019/387324.

TABLE 5.1: The number of samples collected from each country belonging to each of the three species, as well as the number of SNPs, a mean sample depth and a proportion of missing genotypes after filtering through KGD. Country abbreviations: FI: Finland; FR: France; NZ: New Zealand; UK: the United Kingdom; CZ: the Czech Republic.

	UK	**NZ**	**FR**	**FI**	**CZ**	**Total**	**#SNP**	**Depth**	**Missing Genotypes**
R. conglomeratus	136	168	15	-	9	328	791	49.6	0.22
R. obtusifolius	152	126	10	20	21	329	977	62.0	0.24
R. crispus	159	153	5	9	29	355	1276	50.8	0.25

5.2.2. DNA extraction & sequencing

From each plant of each study-species young, growing leaves were collected as samples, and 50 mg of fresh tissue was then freeze-dried for 48 hours and kept in a fridge or on ice until DNA extractions. DNA was extracted from freeze-dried samples using a modified CTAB method by Anderson et al. (2018) with the addition of polyvinylpyrrolidone-10 to help remove excess carbohydrates. Library making and sequencing followed the method by Elshire et al. (2011) with modifications by Dodds et al. (2015). In short, DNA samples were digested with PstI restriction enzyme (NEB R140L and R0106L, New England Biolabs, Ipswich, United States), and uniquely barcoded adapters were ligated to the fragments. Subsequently, all samples were pooled into multiplexed libraries, each consisting of 94 samples. Libraries were cleaned and run through PCR before size selection (193-318 bp) with Pippin Prep (SAGE Science, Beverly, Massachusetts, United States) and finally sequencing with Illumina HiSeq 2500 using single-end reads for 101 cycles in high-output mode (v4 chemistry) at AgResearch Invermay Agricultural Centre, NZ.

5.2.3. Quality controls

After sequencing, all raw reads were subjected to quality controls based on Dodds et al. (2015), using FastQC tool created for high throughput sequencing data (v.0.10.1, https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). This included checking samples for complete bar codes and making sure that all the samples were represented

relatively well. Samples with very low quality were sequenced again, and the duplicate with the highest sample depth was kept. Reads with partially missing barcodes were discarded from the downstream analyses.

I trialled both *de novo* and reference-genome-based approaches. In the reference-based method *Rumex hastatulus* Baldw. was used as a reference genome, as it is the closest related species sequenced to date. Both analyses yielded similar results, but as *R. hastatulus* differs in both its chromosomal number and structure compared to the study species (Navajas-Pérez et al., 2005; Grabowska-Joachimiak et al., 2015) and as just 17 % of the reads mapped to the reference genome, only *de novo* results are presented here (see Appendix B).

De-multiplexing, clean-up and SNP calling were done using reference-free pipeline, UNEAK (Tassel version 3.0.173; Lu et al., 2013) allowing one SNP per location. The settings for UNEAK followed Dodds et al. (2015): -UFastqToTagCountPlugin -c 1 -e PstI; -UMergeTaxaTagCountPlugin -m 600000000 -x 100000000 -c 3; -UTagCountToTagPairPlugin -e 0.03; -UMapInfoToHapMapPlugin -mnMAF 0.03 -mxMAF 0.5 -mnC 0.1.

5.2.4. Filtering in KGD

After SNP calling, HapMap.hmc-files generated by UNEAK were read using KGD (Dodds et al., 2015; https://github.com/AgResearch/KGD) in R software (Version 4.1.0., R Development Core Team, 2008). The KGD software was created for calculating genomic relatedness matrices, and it can handle some QC and data splitting, as well as creating VCF files with sample information that can be used in downstream analyses. In the first step, 'sampdepth.thresh' was set to 0, and SNPs with mean depth under 0.01 and with a minor allele frequency (MAF) = 0 were discarded. After that, SNPs with Hardy-Weinberg disequilibrium (HWD) < – 0.05 were discarded as in Dodds et al. (2015), and only the samples with a call rate of > 0.4 were kept. This was done to ensure that the amount of missing data was kept relatively low, and that only good quality, representative SNPs were used. In addition, the HWD filter removes samples with very high heterozygosity, which could indicate pooling of repetitive sequences or homeologs (K.G. Dodds, personal communication, 2022). After filtering, HapMap.hmc-files were converted as VCF files for population genetics analyses.

5.2.5. Population genetics analyses

After filtering through KGD, data were analysed with population genetics methods, namely AMOVA, DAPC and F_{ST}. AMOVA and DAPC were performed using poppr, adegenet and ade4 packages (Chessel et al., 2004; Jombart, 2008; Jombart et al., 2010; Kamvar et al., 2014) and F_{ST} using KGD in R. All analyses were done within species, and, after QC, all SNPs were treated as if they were diploid.

Two analyses were undertaken for all three species; the first examined country level information from NZ and UK but also from all of the outgroups (hereafter referred to as "the full dataset"). The second analysis focused only on NZ and UK and consisted of

region information on top of the country level information (hereafter referred to as "the region dataset"). Graphs were created using ggplot2 package in R (Wickham, 2016).

Analysis of Molecular Variance (AMOVA)(Excoffier et al., 1992) was chosen to analyse how genetic variation was partitioned, i.e., to look for genetic and geographic structure in the data. In AMOVA, hierarchy for the samples came from the sampling locations, i.e., individuals within regions within countries. AMOVAs were run for each species with the flag *within = FALSE* and significance was tested with 'randtest' function within the *ade4* packages using 999 permutations.

Discriminant Analysis of Principal Components (DAPC)(Jombart et al., 2010) uses known population information and helps visualise patterns in the data. It minimizes the differences within each population while maximising differences between populations, and it uses information from Principal Component Analysis (PCA) as input for a Discriminant Analysis (DA). However, retaining too many PCA axes will lead to overfitting of the data (Jombart & Collins, 2015), thus it is important to optimise the number of axes retained. For DAPC, a-optimisation (function 'optim.a.score') within *adegenet* was performed to retain the optimal number of PCA axes in the analyses. It evaluates the proportion of successful group assignments for every number of PCA axes retained (Jombart & Collins, 2015). Optimisation was undertaken separately for the full and region datasets. DAPC was also performed by retaining PCA axes explaining 80 % of the variation, and results were compared. As the patterns were similar, only the a-optimised results are shown here. The overlap of the 95 % probability ellipses was quantified using SIBER package in R (Jackson et al., 2011). The methods followed the default settings listed in the manual for running the 'bayesianOverlap' function. Overlap was explained as a proportion of overlapping area, with 0 meaning no overlap and 1 meaning a complete overlap.

Within the *adegenet* package, STRUCTURE-like posterior probability graphs were created for allowing a closer look into admixture within samples. The argument 'dapcfile$posterior' saves information of probabilistic value that each sample has in belonging to each of the population, or in our case, country. An individual was considered admixed if it had less than 0.90 assignment probability to any one country. In addition, average posterior probabilities were calculated for each country by taking an average probability of each individuals' assignment to all countries.

Lastly, pairwise F_{ST} values were calculated between pairs of countries and regions to estimate the amount of genetic variance that can be explained by population structure. Wright's F_{ST} (Wright, 1943) was used, which allows values between 0 and 1, indicating no differentiation and a complete differentiation, respectively. For each species, the region dataset was used to compare differentiation within and between NZ and UK, as samples were collected from four regions from both countries. Since only one region was sampled within France (FR), the Czech Republic (CZ) and Finland (FI), comparisons using the full dataset for each species were done at the country level instead.

5.3. Results

5.3.1. Sequencing details and the number of single-nucleotide-polymorphisms

Sequencing yielded 3,124 M raw reads, and after quality control, 2,746 M reads remained; out of these *R. conglomeratus* yielded 939 M reads, *R. obtusifolius* 840 M reads, and *R. crispus* 966 M reads. This equated to around 2 M reads per sample after the quality control.

After filtering in KGD, roughly 750–1300 SNPs were found to be of sufficient quality for downstream analyses (TABLE 5.1). In addition, based on a minimum call rate > 0.4, 21 samples were removed leaving 1012 samples for downstream analyses. These samples had a mean depth of around 50, and around 25 % missing genotypes (TABLE 5.1).

5.3.2. Partitioning of the variation showed high variation within individuals and low variation between countries

To compare partitioning of genetic variation within and between different hierarchy levels, AMOVAs were run for each species. The full datasets were used to assess genetic structure within and between countries.

Most of the variation, around 85 %, was found within the individuals (i.e., high heterozygosity), around 10 % of variation within countries, and no variation between countries (TABLE 5.2). The permutation test showed that within individual variation and within country variation were statistically significant in all species, whereas between country variation was not significant in any species. This indicates little genetic structure between the countries, but some structure within them.

TABLE 5.2: AMOVA results showing partitioning of variations within different levels of hierarchy (Sigma) as well as percentage of the total variation, using the full dataset for each *Rumex* species. *: < 0.05, **: < 0.01, ***: < 0.001.

	R. conglomeratus		***R. obtusifolius***		***R. crispus***	
Comparison	*Sigma*	*%*	*Sigma*	*%*	*Sigma*	*%*
Between countries	-1.191	-0.463	-25.816	-6.785	-2.668	-0.610
Between individuals within countries	35.313	13.747***	82.903	21.789***	52.270	11.960***
Within individuals	222.757	86.717***	323.4	84.996***	387.450	88.651***
Total variation	256.879	100	380.487	100	437.051	100

5.3.3. New Zealand provenances differed little from all European provenances

Discriminant analyses of principal components were undertaken with the full dataset to portray overlap between the countries. A-optimisation was done to determine the best number of PCA axes to retain, and 14, 15 and 13 axes were kept for *R. conglomeratus*, *R. obtusifolius* and *R. crispus*, respectively. Results show that NZ and UK have varying degrees of overlap with each other and with the other European outgroups (FIGURE 5.1, TABLE 5.3). In order to make comparisons easier, data were plotted along the first two axes in all three species. The highest overlap was found in *R. obtusifolius* between UK and NZ with 49 % overlap between the 95 % probability ellipses, for *R. conglomeratus* and *R. crispus* the overlap was 1 % and 24 %, respectively. The overlap in *R. conglomeratus* was highest between NZ and FR, around 18 %.

The posterior probability analyses (FIGURE 5.2, TABLE 5.4, TABLE 5.5) showed varying levels of similarities – or admixture – with non-origin countries. The average assignment probabilities varied drastically between countries, with plants from NZ and UK averaging over 0.75 probability matching with the country of origin in all three species (TABLE 5.4). But on the other hand, both *R. obtusifolius* and *R. crispus* plants from NZ showed a high association with UK, 0.19 and 0.22, respectively. In addition, plants from FR showed a high association with NZ in all three species.

A majority of the samples produced posterior assignment probabilities of 0.90 matching with the country where the samples originated from, indicating a genetic match with the country of origin (TABLE 5.5). However, almost half of *R. obtusifolius* and *R. crispus* from NZ showed admixture with other countries, mainly with UK, but also with FR. In contrast, only 7 % of *R. conglomeratus* samples from NZ suggested admixture. Similarly, the majority of the UK samples designated as admixed showed high proportions of similarities with NZ, but the proportion of admixed individuals was smaller, 3 %, 38 % and 33 % for *R. conglomeratus*, *R. obtusifolius*, and *R. crispus*, respectively.

The genetic differentiation between all countries was analysed using pairwise F_{ST}. Apart from Finland in *R. obtusifolius*, all of the pairwise comparisons have a value of less than 0.05 (TABLE 5.6), indicating very low differentiation between the countries. In addition, in all three species, NZ had less differentiation when compared to FR than to UK.

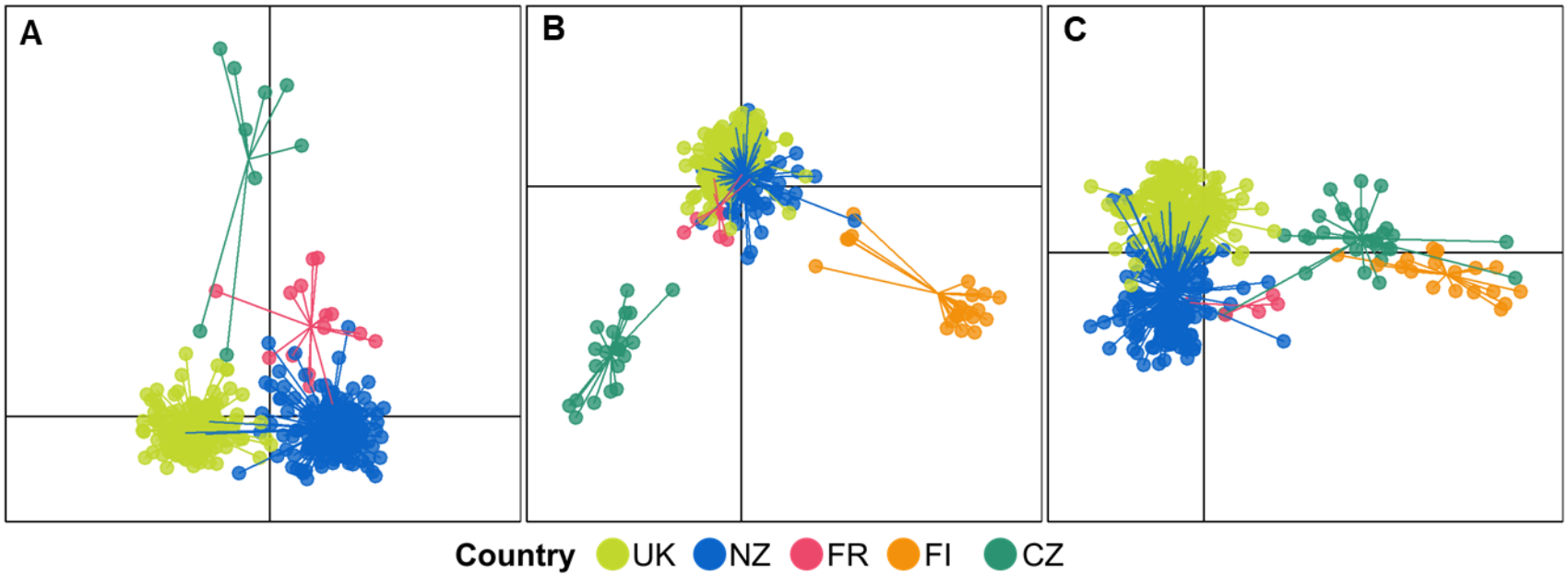

FIGURE 5.1: Discriminant Analysis of Principal Components (DAPC) highlighting hierarchy based on genetic information and *a priori* information of the populations, the country of origin in this case, for A) *Rumex conglomeratus*; B) *R. obtusifolius*; C) *R. crispus*. After a-optimisation, 14, 15 and 13 PCA axes were kept for *R. conglomeratus*, *R. obtusifolius* and *R. crispus*, respectively. Individuals are plotted along the first and second DA axes. Country abbreviations: FI: Finland; FR: France; NZ: New Zealand; UK: the United Kingdom; CZ: the Czech Republic.

TABLE 5.3: The overlap between 95 % probability ellipses in the corresponding two-dimensional DAPC (FIGURE 5.1) calculated between New Zealand and all the countries within the native region for all three species of *Rumex*. Overlap of 1 equals full overlap, and 0 equals no overlap. Country abbreviations: FI: Finland; FR: France; NZ: New Zealand; UK: the United Kingdom; CZ: the Czech Republic.

Pair	***R. conglomeratus***	***R. obtusifolius***	***R. crispus***
UK-NZ	0.012	0.493	0.244
FR-NZ	0.179	0.374	0.121
FI-NZ	N/A	< 0.001	< 0.001
CZ-NZ	0.002	< 0.001	0.010

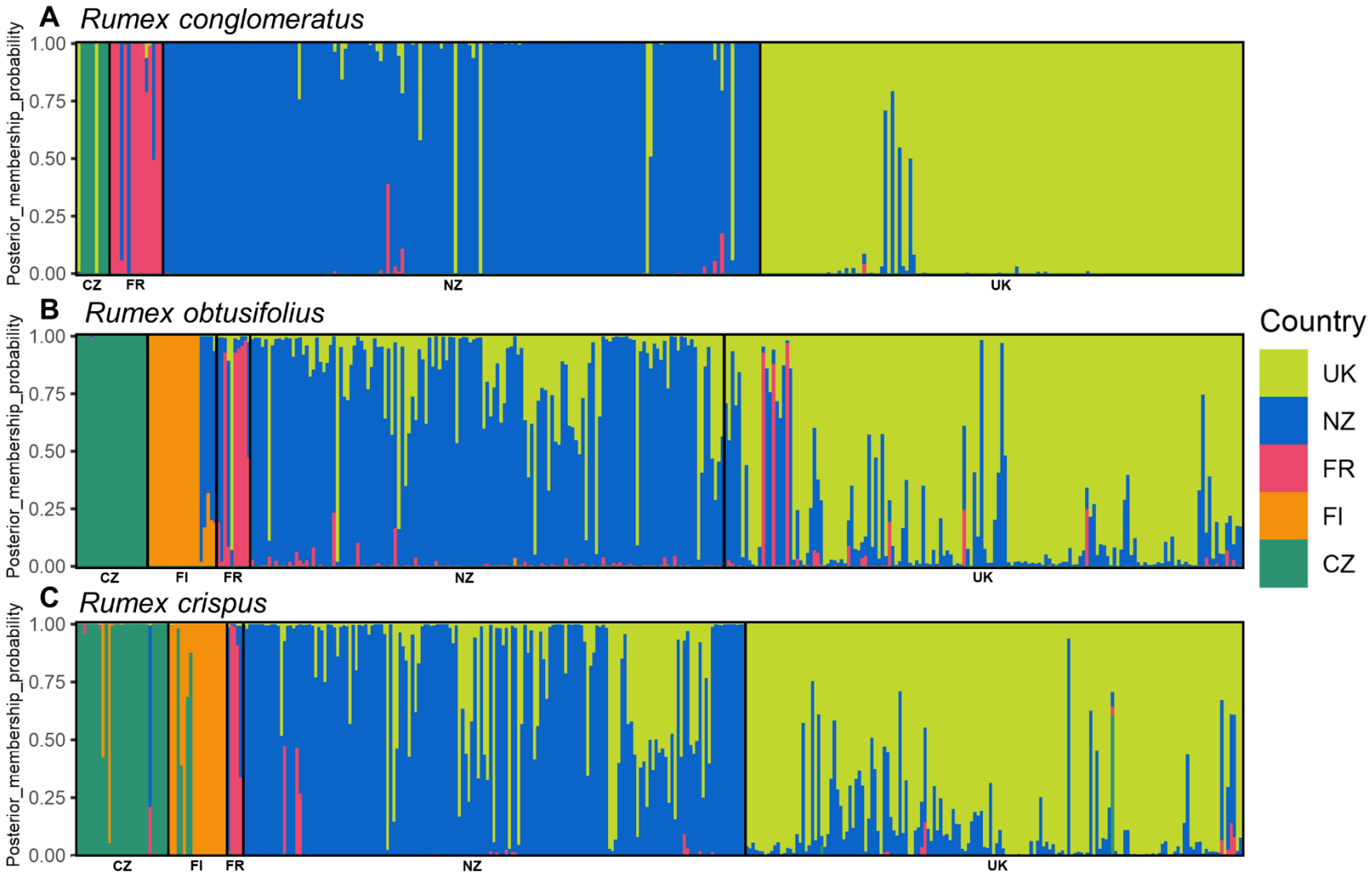

FIGURE 5.2: Posterior probability plots from Discriminant Analysis of Principal Components (DAPC) showing each individual as a vertical line and its probability to be assigned to each country, i.e., samples with multiple colours suggest similar genetic makeup to more than one country (admixture). Samples are clustered based on the sample's country of origin. Country abbreviations: FI: Finland; FR: France; NZ: New Zealand; UK: the United Kingdom; CZ: the Czech Republic.

TABLE 5.4: The average posterior assignment probabilities of plants within each species to the country of origin and a proportion of association each plant has with countries other than the country of origin, i.e., the proportion admixture. *Rumex conglomeratus*: Rco; *R. obtusifolius*: Rob; *R. crispus*: Rco. Country abbreviations: FI: Finland; FR: France; NZ: New Zealand; UK: the United Kingdom; CZ: the Czech Republic.

Origin	**NZ**			**UK**			**FR**			**FI**		**CZ**		
Species	*Rco*	*Rob*	*Rcr*	*Rco*	*Rob*	*Rcr*	*Rco*	*Rob*	*Rcr*	*Rob*	*Rcr*	*Rco*	*Rob*	*Rcr*
Association														
NZ	0.96	0.80	0.77	0.02	0.16	0.12	0.17	0.34	0.34	0.20	<0.01	<0.01	<0.01	0.03
UK	0.04	0.19	0.22	0.98	0.81	0.87	0.01	0.11	0.01	<0.01	<0.01	0.22	0.00	<0.01
FR	0.01	0.01	0.01	<0.01	0.03	<0.01	0.82	0.55	0.65	<0.01	0.00	<0.01	<0.01	0.01
FI	N/A	<0.01	0.00	N/A	0.00	0.00	N/A	0.00	0.00	0.79	0.85	N/A	0.00	0.05
CZ	0.00	0.00	<0.01	0.00	0.00	<0.01	0.00	0.00	<0.01	0.00	0.15	0.78	1.00	0.91

TABLE 5.5: Proportion of *Rumex* samples within each species with under 0.9 posterior assignment probability, i.e., admixed samples with mixed assignment to at least one country other than the country of origin. Country abbreviations: FI: Finland; FR: France; NZ: New Zealand; UK: the United Kingdom; CZ: the Czech Republic.

	R. conglomeratus	***R. obtusifolius***	***R. crispus***
Sample origin			
NZ	7 %	49 %	41 %
UK	3 % (4 plants)	38 %	33 %
FR	27 % (4 plants)	50 % (5 plants)	40 % (2 plants)
FI	-	25 % (5 plants)	21 % (4 plants)
CZ	22 % (2 plants)	0 %	10 % (3 plants)

TABLE 5.6: Pairwise F_{ST} values showing differentiation between the countries using the full dataset. a) *Rumex conglomeratus*; b) *R. obtusifolius*; c) *R. crispus*. Country abbreviations: FI: Finland; FR: France; NZ: New Zealand; UK: the United Kingdom; CZ: the Czech Republic.

		FI	**FR**	**NZ**	**UK**
a. *R. conglomeratus*					
	CZ	-	0.045	0.042	0.030
	FR	-		0.018	0.016
	NZ	-			0.025
b. R. obtusifolius					
	CZ	0.138	0.038	0.035	0.027
	FI		0.128	0.143	0.109
	FR			0.010	0.008
	NZ				0.014
c. R. crispus					
	CZ	0.046	0.022	0.029	0.028
	FI		0.047	0.029	0.024
	FR			0.006	0.008
	NZ				0.013

5.3.4. Closer comparison between the UK and New Zealand showed less differentiation within the countries than between

Lastly, I wanted to get a better understanding of the differences between the native and introduced ranges and took a closer look using region datasets. This time individual regions were compared to see finer patterns within and between the ranges in the three *Rumex* species.

When UK and NZ were analysed without the outgroups, the variation between countries increased in *R. conglomeratus* to explain around 5 % of the variation and a bit less in the other two species (TABLE 5.7). Most of the variation was still within individuals in all three species. All but the between country level variation in *R. crispus* were significant according to the randomisation test.

Similar to the other analyses, the pairwise F_{ST} values using just the region datasets showed little genetic differentiation between the countries or regions (FIGURE 5.3). All F_{ST} values were below 0.05, which is considered low differentiation (Hartl & Clark, 1997; Frankham et al., 2002). However, *R. conglomeratus* and *R. obtusifolius* showed similar patterns: NZ regions differed less from each other than UK regions did, whereas for *R. crispus* the differentiation within countries was more similar.

TABLE 5.7: AMOVA results showing partitioning of variations within different levels of hierarchy in *Rumex* using the region datasets. *: < 0.05, **: < 0.01, ***: < 0.001.

	R. conglomeratus		***R. obtusifolius***		***R. crispus***	
Comparison	*Sigma*	*%*	*Sigma*	*%*	*Sigma*	*%*
Between countries	12.606	4.973***	7.588	2.356***	7.210	1.649
Between individuals within countries	24.871	9.812***	18.424	5.722***	36.496	8.346***
Within individuals	216.002	85.215***	295.992	91.922***	393.580	90.005***
Total variations	253.479	100	396.165	100	437.286	100

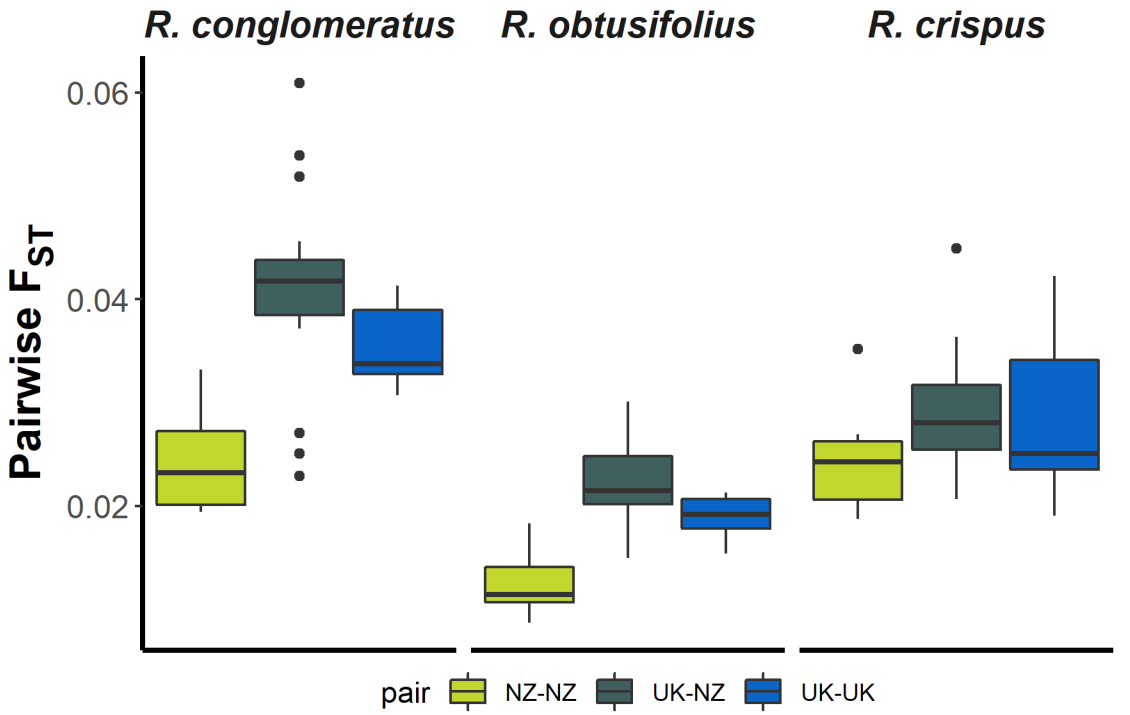

FIGURE 5.3: Pairwise F_{ST} analysis shows little differentiation within and between the regions in UK and New Zealand (NZ) in all three *Rumex* species. For each country, four regions were compared to each other to get the pairwise estimates. Similarly, for the country comparison the four regions within one country were compared to the four regions within the other country. The box represents the interquartile range with the median represented with a thick line and the whiskers extending from the 5^{th} to the 95^{th} percentile, with values outside this range (outliers) represented as dots.

5.4. Discussion

The purpose of this study was to investigate genetic variation and its partitioning in three species of *Rumex* in both the species' native range as well as the introduced range. I asked if variation is partitioned differently within and between the ranges, if there is evidence of admixture within the introduced range and whether the NZ introduction could have originated from the UK. Samples were collected from the UK and NZ from climatically

matching regions. In addition, a smaller set of samples were collected from three other European countries (FI, CZ, and FR) to see how these other regions within the native range compared to the introduced range. Three potential scenarios were considered: 1) Variation is different in native and introduced regions, with introduced regions exhibiting lower variation likely telling of bottlenecks; 2) Variation is different in native and introduced regions, with relatively high variation within introduced range likely due to admixture and multiple introductions; and 3) Variation is similar in both ranges due to admixture and multiple introductions. The results support scenario 3) minimal differentiation was found both within and between the native and introduced ranges. In addition, this likely indicates that the UK is a source of the introductions to NZ. Due to low differentiation between the ranges, prior adaptation likely explains the success of these species in NZ, rather than post-introduction evolution.

5.4.1. The partitioning of genetic variation similar between the native and introduced ranges

Overall, most of the genetic variation was within individuals in all three species, indicating a lack of genetic structure in the data when all countries were considered. This supports scenario 3, where native and introduced regions exhibit similar partitioning of variation and low overall differences between the countries. According to the AMOVAs using the full datasets, there was minimal geographic structuring in all of the species, and none of the variation was attributed to between the countries (TABLE 5.2), so genetically the native and introduced ranges are hard to separate from each other. Within country level variation explained around 10–25 % of the total variation in all three species, and was statistically significantly higher than expected by chance, based on the randomisation test. This indicates some genetic structure within the countries; however, around 85 % of the variation was within individuals.

The high variation within individual plants can reflect the breeding system of the species. Favouring less related individuals in mating (i.e., outcrossing) can lead to higher within individual variation compared to between individuals variation by increasing heterozygosity (Fox, 2005). Polyploidy also often increases heterozygosity (te Beest et al., 2012), which could help to explain the high within individual variation. On the other hand, especially *R. obtusifolius* and *R. crispus* are known for high rates of selfing (Cavers & Harper, 1964; Akeroyd & Briggs, 1983; Grime et al., 2007), which would decrease the variation within individuals (e.g., homozygosity) and increase the variation between individuals. However, since only 12–22 % of the variation can be accounted for between the individuals, this does not necessarily suggest high selfing rates. Furthermore, both my work with *R. conglomeratus* (Chapter 3) and a study by Friedman & Barrett (2009) with *R. crispus* indicate that these species likely have a mixed breeding system, with some combination of selfing and outcrossing. This is likely reflected in the AMOVA results as well. However, as all samples were analysed as diploids, if the SNP calling did not differentiate between the subgenomes, this could show as an increase in the heterozygosity.

Lastly, between country level comparisons produced negative values, but as these values are only slightly negative, they should be considered as zero, rather than indications of issues in the data or sampling effort (Meirmans, 2006).

5.4.2. The species were likely introduced multiple times creating admixture

Admixture is known to protect species from negative consequences due to low genetic variation, and to help introduced species in establishing (Ellstrand & Schierenbeck, 2006). As the AMOVAs revealed little geographic structure between the countries, introductions from multiple countries within the native range to the introduced range (i.e., admixture) was expected to be the reason. DAPCs using the full datasets show that there is a considerable amount of overlap between all countries in all three species (TABLE 5.3). Most of the overlap is seen between UK and NZ in all three species, but samples from France show overlap with NZ as well (FIGURE 5.1, TABLE 5.3). Indeed, the overlap in *R. conglomeratus* is higher between NZ and FR than between NZ and UK, and while it is higher between NZ and UK for the other two species, especially in *R. obtusifolius* we can still see a considerable amount of overlap between NZ and FR. DAPC is a powerful tool for mapping overlap between *a priori* defined populations, as it minimises within cluster variation while maximising between cluster variation (Jombart et al., 2010).

In addition, the data highlights many similarities not only between NZ and UK but also between NZ and FR, in all three species, as well as low differentiation between NZ and CZ according to F_{ST} (TABLE 5.6). The lack of geographic structure even when the outgroup countries are considered is likely an indication of introductions and subsequent admixture from outside the UK. Indeed, pairwise F_{ST} shows less differentiation between NZ and FR compared to differentiation between NZ and UK (TABLE 5.6), which can partially reflect the low sampling effort in these countries but suggests introduction from FR to NZ. However, it is important to keep in mind the small numbers of samples collected from the outgroup countries, especially given only one sampling region in each outgroup country. The pattern is similar in all three species, though, which seems unlikely if it was just a chance effect caused by a small number of samples.

The posterior probability analyses (FIGURE 5.2, TABLE 5.4, TABLE 5.5) can provide us information on the sample level, as opposed to country level. However, as a part of the DAPC analyses and thus similarly to them, the output is dependent on the number of PCA axes retained, and thus should only provide a proxy of the admixture within the samples. It does provide a good indication of whether the patterns of DAPC arise from a few fully admixed individuals or whether more samples share moderate amounts of similarities to other countries. The latter seems to be the case for both *R. obtusifolius* and *R. crispus* (TABLE 5.5), with well over third of the samples showing mixed assignments. In addition, the analyses showed similarities between NZ and FR in all three species, supporting observations from the other analyses. And indeed, these results further solidify the scenario 3, by highlighting the similarities and admixture between the countries and between the native and introduced range.

F_{ST} values for polyploids can appear lower than what they are in reality, if estimated using methods for diploids (Meirmans et al., 2018). However, as F_{ST} values were extremely low and similar across all three species, and similar patterns were produced by DAPC and AMOVA, this is not considered an issue. The species also hybridise freely despite the different ploidy levels, and the hybrids are known to backcross back to the

parent species. If anything, the more common *R. obtusifolius* and *R. crispus* might be expected to have less differentiation than *R. conglomeratus* given their broader range.

5.4.3. The UK is a likely source for the introductions to New Zealand

When I took a closer look using regional data from UK and NZ, the AMOVAs still found most of the variation within individuals, but now the country level variation was significant and higher than before (TABLE 5.7). In addition, according to the F_{ST} (FIGURE 5.3), the native range has a higher differentiation between regions than the introduced range, especially in *R. conglomeratus* and *R. obtusifolius*. This was expected, as native range has had longer time to adapt to local conditions and is often found to have higher genetic variation than the introduced range (Ward et al., 2008; Barrett, 2015). In addition, both *R. conglomeratus* and *R. obtusifolius* showed higher differentiation when UK regions are compared to NZ regions, whereas for *R. crispus* this is less pronounced. However, with multiple introductions or gene flow from UK to NZ, it is expected that the differentiation between the countries is low.

In conclusion, the very low differentiation between NZ and UK supports the assumption of multiple introductions, and the historical assumption of early introductions from the UK to NZ. In addition, the NZ regions barely differ from each other (FIGURE 5.3), which could indicate a gene flow between the regions.

Harrop et al. (2020) found similar results indicating gene flow and low differentiation within NZ in Argentine stem weevil *Listronotus bonariensis* Kuschel (Coleoptera: Curculionidae). The authors found F_{ST} values between stem weevil populations within NZ to be lower than 0.05, similar to this study. However, despite low F_{ST} values within the whole country, the authors did not find gene flow between the two major clusters they identified. Rather, according to the introduction models by Harrop et al. (2020), it is likely that *L. bonariensis* was introduced multiple times from the same source populations to these two major clusters, explaining the low differentiation despite the lack of gene flow.

As I did not explicitly model gene flow, it is difficult to confirm whether the results are an indication of recent gene flow or of historical introductions from the same sources. It needs to be noted, that *Rumex* seeds can travel relatively long distances assisted by rivers, humans and animals, soil and machinery (Holm, 1977; Grime et al., 2007). However, the samples were collected from sites that are tens if not hundreds of kilometres apart (Bufford & Hulme, 2021b) and sometimes separated by a mountain range, which could limit gene flow between the regions.

In addition, F_{ST} values were indicating that some NZ regions differentiated less from some UK regions than from other NZ regions (data not shown). For example, plants from Southland NZ were less differentiated from Cambridge UK than from Canterbury NZ in all three species. This could indicate introductions from similar sources rather than gene flow, as this would allow for regional differences. However, this could also be an indication of adaptation to local conditions.

5.4.4. Invasion history of *Rumex* spp. to New Zealand

Given the European colonisation history of NZ, most weeds introduced in the 1800s are assumed to have come from the UK (Thomson, 1922). My data supports this theory by showing minimal differentiation between the two countries. However, since NZ plants show high similarities to plants from elsewhere in Europe, namely FR, introductions from other areas are very likely as well. Similar results were shown in *Hieracium pilosella* (Trewick et al., 2004). Trewick et al. (2004) showed more similarities in *H. pilosella* between NZ and both CZ and FI than between NZ and UK, based on chloroplast DNA. Contrasting to their results, this study found support for introductions from UK to NZ, but also similarly to their findings, introductions from other European countries are highly likely. This was especially so for FR, which, in all three species, showed similar or less differentiation when compared to NZ than when compared to UK (TABLE 5.6). In addition, the differentiation between UK and FR was smaller than the differentiation between UK and NZ in two out of three species, which is to be expected when countries within the native region are compared to each other. However, since NZ differed less from FR than to UK, this could indicate introductions from FR to NZ independent of introductions from the UK.

The results show no signs of bottlenecks but rather a strong indication of multiple introductions and subsequent admixture both from the UK but also elsewhere in Europe to NZ. This would likely mean introductions at least from FR, but other European countries outside this study cannot be ruled out. It needs to be considered, that already admixed individuals from UK, outside our sampling efforts, could explain the patterns. However, results were similar in all three species and has been shown before with *H. pilosella* (Trewick et al., 2004), rather indicating introductions from multiple countries.

In addition, when just UK and NZ are compared using AMOVAs, country level variation is significant, but it is not significant when the outgroups are included. This can further suggest that UK is not the only source for introductions in NZ, otherwise the geographic structuring should appear only when the outgroups are included. Alternatively, if the outgroup countries have high amounts of variation, removing them from the analyses can lead to an increase in the between country variation, as the differences get more pronounced. However, as F_{ST} values suggested less differentiation between NZ and FR, it is likely that some introductions originated outside the UK.

5.4.5. Caveats and future analyses

While sample sizes from the outgroup countries were small and not much can be said on country level representativeness, the low differentiation compared to the introduced range gives a good indication of introductions outside of the UK, as well as an indication of a relatively low overall variation within the species' native range.

In addition, genetic analyses in polyploids, and especially allopolyploids, are still not common, especially in non-model species and species without a reference genome. This makes it difficult to estimate the reliability of the results (but see Chapter 6). However, as allopolyploids, and especially tetraploids, have diploid genomes from different ancestors (Bourke et al., 2018), I chose to analyse all three species as diploids, and assume low dosage bias unlike with autopolyploids that have multiple copies of the

same genome, which would bias the allele calls to heterozygous and homozygous. Matias et al. (2019) studied association mapping in *Urochloa* spp. and compared GWAS analysis in diploid and allotetraploid species using diploid and tetraploid configurations in their analyses. While they recommend both for GWAS, the authors also highlight how diploid configuration alone can provide good results in allotetraploid plants. However, this assumes that the sequencing differentiated between the subgenomes and distinguished between homoeologous loci. Not distinguishing between these loci can show as increased sample depth, which was much higher in the reference-based approach compared to the *de novo* approach (Appendix B, Chapter 6).

Analysing allopolyploids as diploids assumes independent sub-genomes with no interaction with each other. While this can be the case, in reality the sub-genomes can interact and homoeologous recombination happens. In addition, the progenitors for *R. obtusifolius* and *R. crispus* are not known, and no prior information about the genomes exists, which can complicate matters especially with the hexaploid *R. crispus*. The only currently sequenced *Rumex* species with an annotated genome (*R. hastatulus*) differs from these species in major ways, having smaller base chromosome number and separate sex chromosomes. In the reference-based method less than 20 % of the SNPs mapped to the genome (Appendix B). Thus, even without the reference genome, our study provides an important starting point in understanding the polyploid *Rumex* species better. In addition, while it is likely that some samples are hybrids and backcrosses, we did not filter these out, as the recognition of hybrids would have required a broad scale flow-cytometry analysis for estimating the genome size of the samples.

The results provided here were well replicated in the reference-based approach (see Appendix B, Chapter 6), giving confidence in the patterns found. Especially since the SNPs found using the *de novo* and reference-based methods were almost completely different from one another: well below 1 % of the available tag pairs matched between the methods (data not shown).

Further research in understanding the colonisation history of *Rumex* should investigate more populations within both the native and the introduced ranges. Vallejo-Marín et al. (2021) found that in *Erythrante guttata* the introductions from the native range to UK had served as a bridgehead for further introductions. This could be the case with the three *Rumex* species, especially from Australia, but as this study did not investigate samples from Australia, it cannot be ruled in or out.

Understanding weed evolution is paramount to the success in eradicating or controlling them. It will help us, for example, in determining the suitable eradication method. Low differentiation between the ranges suggests that similar methods will work in both the native and introduced ranges. In addition, the species are growing in very similar habitats, with mostly other European species, and thus are likely to have prior adaptation to anthropogenic habitats, such as paddocks (Hufbauer et al., 2012). The niche overlap between NZ and UK is relatively low (Carlin et al., 2022), which could indicate some level of local adaptation due to environmental drivers. However, *Rumex* species have shown high phenotypic plasticity (Farris & Schaal, 1983; Zaller, 2004; Kołodziejek, 2019; Bufford & Hulme, 2021a), which favours prior adaptation rather than post-introduction evolution.

5.5. References

Akeroyd, J. R. (2014). *Docks and knotweeds of Britain and Ireland* (2nd ed.). Botanical Society of Britain & Ireland.

Akeroyd, J. R., & Briggs, D. (1983). Genecological studies of *Rumex crispus* L. II. Variation in plants grown from wild-collected seed. *New Phytologist*, *94*(2), 325–343.

Alexander, J. M., Poll, M., Dietz, H., & Edwards, P. J. (2009). Contrasting patterns of genetic variation and structure in plant invasions of mountains. *Diversity and Distributions*, *15*(3), 502–512. https://doi.org/10.1111/j.1472-4642.2008.00555.x

Allendorf, F. W., & Lundquist, L. L. (2003). Introduction: Population biology, evolution, and control of invasive species. *Conservation Biology*, *17*(1), 24–30. https://doi.org/10.1046/j.1523-1739.2003.02365.x

Anderson, C. B., Franzmayr, B. K., Hong, S. W., Larking, A. C., van Stijn, T. C., Tan, R., Moraga, R., Faville, M. J., & Griffiths, A. G. (2018). Protocol: A versatile, inexpensive, high-throughput plant genomic DNA extraction method suitable for genotyping-by-sequencing. *Plant Methods*, *14*(1), Article 75. https://doi.org/10.1186/s13007-018-0336-1

Assche, J. V., Nerum, D. V., & Darius, P. (2002). The comparative germination ecology of nine Rumex species. *Plant Ecology*, *159*, 121–142.

Barrett, S. C. H. (2015). Foundations of invasion genetics: The Baker and Stebbins legacy. *Molecular Ecology*, *24*(9), 1927–1941. https://doi.org/10.1111/mec.13014

Bossdorf, O., Auge, H., Lafuma, L., Rogers, W. E., Siemann, E., & Prati, D. (2005). Phenotypic and genetic differentiation between native and introduced plant populations. *Oecologia*, *144*(1), 1–11. https://doi.org/10.1007/s00442-005-0070-z

Bourke, P. M., Voorrips, R. E., Visser, R. G. F., & Maliepaard, C. (2018). Tools for genetic studies in experimental populations of polyploids. *Frontiers in Plant Science*, *9*, Article 513. https://doi.org/10.3389/fpls.2018.00513

Bufford, J. L., & Hulme, P. E. (2021a). Increased adaptive phenotypic plasticity in the introduced range in alien weeds under drought and flooding. *Biological Invasions*, *23*(8), 2675–2688. https://doi.org/10.1007/s10530-021-02532-5

Bufford, J. L., & Hulme, P. E. (2021b). Seed size-number trade-offs are absent in the introduced range for three congeneric plant invaders. *Journal of Ecology*, *109*, 3849–3860. https://doi.org/10.1111/1365-2745.13761

Carlin, T. F., Bufford, J. L., Hulme, P. E., & Godsoe, W. K. (2022). Global assessment of three *Rumex* species reveals inconsistent climatic niche shifts across multiple introduced ranges. *Biological Invasions*. https://doi.org/10.1007/s10530-022-02893-5

Catford, J. A., Jansson, R., & Nilsson, C. (2009). Reducing redundancy in invasion ecology by integrating hypotheses into a single theoretical framework. *Diversity and Distributions*, *15*(1), 22–40. https://doi.org/10.1111/j.1472-4642.2008.00521.x

Cavers, P. B., & Harper, J. L. (1964). *Rumex obtusifolius* L. and *R. crispus* L. *Journal of Ecology*, *52*(325), 737–766.

Cavers, P. B., & Harper, J. L. (1966). Germination polymorphism in *Rumex crispus* and *Rumex obtusifolius*. *Journal of Ecology*, *54*(2), 367–382. https://doi.org/10.2307/2257955

Chessel, D., Dufour, A.-B., & Thioulouse, J. (2004). The ade4 package – I: One-table methods. *R News*, *4*(1), 5–10.

De Carvalho, D., Ingvarsson, P. K., Joseph, J., Suter, L., Sedivy, C., MacAya-Sanz, D., Cottrell, J., Heinze, B., Schanzer, I., & Lexer, C. (2010). Admixture facilitates adaptation from standing variation in the European aspen (*Populus tremula* L.), a widespread forest tree. *Molecular Ecology*, *19*(8), 1638–1650. https://doi.org/10.1111/j.1365-294X.2010.04595.x

Dlugosch, K. M., & Parker, I. M. (2008). Invading populations of an ornamental shrub show rapid life history evolution despite genetic bottlenecks. *Ecology Letters*, *11*(7), 701–709. https://doi.org/10.1111/j.1461-0248.2008.01181.x

Dodds, K. G., McEwan, J. C., Brauning, R., Anderson, R. M., van Stijn, T. C., Kristjánsson, T., & Clarke, S. M. (2015). Construction of relatedness matrices using genotyping-by-

sequencing data. *BMC Genomics*, *16*(1), Article 1047. https://doi.org/10.1186/s12864-015-2252-3

Ellstrand, N. C., & Schierenbeck, K. A. (2006). Hybridization as a stimulus for the evolution of invasiveness in plants? *Euphytica*, *148*, 35–46. https://doi.org/10.1007/s10681-006-5939-3

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, *6*(5), Article e19379. https://doi.org/10.1371/journal.pone.0019379

Estoup, A., Ravigné, V., Hufbauer, R., Vitalis, R., Gautier, M., & Facon, B. (2016). Is there a genetic paradox of biological invasion? *Annual Review of Ecology, Evolution, and Systematics*, *47*(1), 51–72. https://doi.org/10.1146/annurev-ecolsys-121415-032116

Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, *131*(2), 479–491. https://doi.org/10.1093/genetics/131.2.479

Farris, M. A., & Schaal, B. A. (1983). Morphological and genetic variation in ecologically central and marginal populations of *Rumex acetosella* L. (Polygonaceae). *American Journal of Botany*, *70*(2), 246–255. https://doi.org/10.2307/2443270

Fox, C. W. (2005). Problems in measuring among-family variation in inbreeding depression. *American Journal of Botany*, *92*(11), 1929–1932. https://doi.org/10.3732/ajb.92.11.1929

Frankham, R., Ballou, S. E. J. D., Briscoe, D. A., & Ballou, J. D. (2002). *Introduction to conservation genetics*. Cambridge University Press.

Friedman, J., & Barrett, S. C. H. (2009). Wind of change: New insights on the ecology and evolution of pollination and mating in wind-pollinated plants. *Annals of Botany*, *103*(9), 1515–1527. https://doi.org/10.1093/aob/mcp035

Grabowska-Joachimiak, A., Kula, A., Książczyk, T., Chojnicka, J., Sliwinska, E., & Joachimiak, A. J. (2015). Chromosome landmarks and autosome-sex chromosome translocations in *Rumex hastatulus*, a plant with XX/XY1Y2 sex chromosome system. *Chromosome Research*, *23*(2), 187–197. https://doi.org/10.1007/s10577-014-9446-4

Grime, J. P., Hodgson, J. G., & Hunt, R. (2007). *Comparative plant ecology: A functional approach to common British species* (2nd ed.). Castlepoint Press.

Grover, A., & Sharma, P. C. (2016). Development and use of molecular markers: Past and present. *Critical Reviews in Biotechnology*, *36*(2), 290–302. https://doi.org/10.3109/07388551.2014.959891

Harrop, T. W. R., Le Lec, M. F., Jauregui, R., Taylor, S. E., Inwood, S. N., van Stijn, T., Henry, H., Skelly, J., Ganesh, S., Ashby, R. L., Jacobs, J. M. E., Goldson, S. L., & Dearden, P. K. (2020). Genetic Diversity in Invasive Populations of Argentine Stem Weevil Associated with Adaptation to Biocontrol. *Insects*, *11*(7), Article 441. https://doi.org/10.3390/insects11070441

Hartl, D. L., & Clark, A. G. (1997). *Principles of Population Genetics* (3rd ed.). Sinauer Associates.

Holm, L. G. (1977). *The world's worst weeds: Distribution and biology*. University Press of Hawaii.

Hufbauer, R. A. (2017). Admixture is a driver rather than a passenger in experimental invasions. *Journal of Animal Ecology*, *86*(1), 4–6. https://doi.org/10.1111/1365-2656.12600

Hufbauer, R. A., Facon, B., Ravigné, V., Turgeon, J., Foucaud, J., Lee, C. E., Rey, O., & Estoup, A. (2012). Anthropogenically induced adaptation to invade (AIAI): Contemporary adaptation to human-altered habitats within the native range can promote invasions. *Evolutionary Applications*, *51*(2), 89–101. https://doi.org/10.1111/J.1752-4571.2011.00211.X

Jackson, A. L., Parnell, A. C., Inger, R., & Bearhop, S. (2011). Comparing isotopic niche widths among and within communities: SIBER – Stable Isotope Bayesian Ellipses in R. *Journal of Animal Ecology*, *80*(3), 595–602.

Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129
Jombart, T., & Collins, C. (2015, June 23). *A tutorial for discriminant analysis of principal components (DAPC) using adegenet 2.0.0*. https://adegenet.r-forge.r-project.org/files/tutorial-dapc.pdf
Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, *11*(1), Article 94. https://doi.org/10.1186/1471-2156-11-94
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*, Article e281. https://doi.org/10.7717/peerj.281
Keller, S. R., & Taylor, D. R. (2010). Genomic admixture increases fitness during a biological invasion. *Journal of Evolutionary Biology*, *23*(8), 1720–1731. https://doi.org/10.1111/j.1420-9101.2010.02037.x
Kołodziejek, J. (2019). Growth and competitive interaction between seedlings of an invasive *Rumex confertus* and of co-occurring two native *Rumex* species in relation to nutrient availability. *Scientific Reports*, *9*(1), Article 3298. https://doi.org/10.1038/s41598-019-39947-z
Les, D. (2017). *Aquatic Dicotyledons of North America*. CRC Press. https://doi.org/10.1201/9781315118116
Lu, F., Lipka, A. E., Glaubitz, J. C., Elshire, R., Cherney, J. H., Casler, M. D., Buckler, E. S., & Costich, D. E. (2013). Switchgrass genomic diversity, ploidy, and evolution: Novel insights from a network-based snp discovery protocol. *PLoS Genetics*, *9*(1), Article e1003215. https://doi.org/10.1371/journal.pgen.1003215
Matias, F. I., Vidotti, M. S., Meireles, K. G. X., Barrios, S. C. L., Valle, C. B. do, Carley, C. A. S., & Fritsche-Neto, R. (2019). Association mapping considering allele dosage: An example of forage traits in an interspecific segmental allotetraploid *Urochloa* spp. panel. *Crop Science*, *59*(5), 2062–2076. https://doi.org/10.2135/cropsci2019.03.0185
Meirmans, P. G. (2006). Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution*, *60*(11), 2399–2402.
Meirmans, P. G., Liu, S., & Tienderen, P. (2018). The analysis of polyploid genetic data. *The Journal of Heredity*, *109*(3), 283–296. https://doi.org/10.1093/jhered/esy006
Montesinos, D., Santiago, G., & Callaway, R. M. (2012). Neo-allopatry and rapid reproductive isolation. *The American Naturalist*, *180*(4), 529–533. https://doi.org/10.1086/667585
Navajas-Pérez, R., de la Herrán, R., López González, G., Jamilena, M., Lozano, R., Ruiz Rejón, C., Ruiz Rejón, M., & Garrido-Ramos, M. A. (2005). The evolution of reproductive systems and sex-determining mechanisms within *Rumex* (Polygonaceae) inferred from nuclear and chloroplastidial sequence data. *Molecular Biology and Evolution*, *22*(9), 1929–1939. https://doi.org/10.1093/molbev/msi186
Oduor, A. M. O., Leimu, R., & van Kleunen, M. (2016). Invasive plant species are locally adapted just as frequently and at least as strongly as native plant species. *Journal of Ecology*, *104*(4), 957–968. https://doi.org/10.1111/1365-2745.12578
Pérez, J. E., Nirchio, M., Alfonsi, C., & Muñoz, C. (2006). The biology of invasions: The genetic adaptation paradox. *Biological Invasions*, *8*, 1115–1121. https://doi.org/10.1007/s10530-005-8281-0
Pina-Martins, F., Baptista, J., Pappas, G., & Paulo, O. S. (2018). New insights into adaptation and population structure of cork oak using genotyping by sequencing. *Global Change Biology*, *25*(1), 337–350. https://doi.org/10.1111/gcb.14497
Poland, J. A., & Rife, T. W. (2012). Genotyping-by-sequencing for plant breeding and genetics. *The Plant Genome*, *5*(3), 92–102. https://doi.org/10.3835/plantgenome2012.05.0005
Rius, M., & Darling, J. A. (2014). How important is intraspecific genetic admixture to the success of colonising populations? *Trends in Ecology & Evolution*, *29*(4), 233–242. https://doi.org/10.1016/j.tree.2014.02.003

Roman, J., & Darling, J. A. (2007). Paradox lost: Genetic diversity and the success of aquatic invasions. *Trends in Ecology & Evolution*, *22*(9), 454–464. https://doi.org/10.1016/j.tree.2007.07.002

Schrieber, K., & Lachmuth, S. (2016). The Genetic Paradox of Invasions revisited: The potential role of inbreeding × environment interactions in invasion success. *Biological Reviews*, *92*(2), 939–952. https://doi.org/10.1111/brv.12263

Seebens, H., Blackburn, T. M., Dyer, E. E., Genovesi, P., Hulme, P. E., Jeschke, J. M., Pagad, S., Pyšek, P., Winter, M., Arianoutsou, M., Bacher, S., Blasius, B., Brundu, G., Capinha, C., Celesti-Grapow, L., Dawson, W., Dullinger, S., Fuentes, N., Jäger, H., … & Essl, F. (2017). No saturation in the accumulation of alien species worldwide. *Nature Communications*, *8*(1), Article 14435. https://doi.org/10.1038/ncomms14435

Stetter, M. G., Müller, T., & Schmid, K. J. (2017). Genomic and phenotypic evidence for an incomplete domestication of South American grain amaranth (Amaranthus caudatus). *Molecular Ecology*, *26*(3), 871–886. https://doi.org/10.1111/mec.13974

Takahashi, K., & Hanyu, M. (2015). Hybridization between alien species *Rumex obtusifolius* and closely related native vulnerable species *R. longifolius* in a mountain tourist destination. *Scientific Reports*, *5*, Article 13898. https://doi.org/10.1038/srep13898

Tayeh, A., Estoup, A., Hufbauer, R. A., Ravigne, V., Goryacheva, I., Zakharov, I. A., Lombaert, E., & Facon, B. (2013). Investigating the genetic load of an emblematic invasive species: The case of the invasive harlequin ladybird *Harmonia axyridis*. *Ecology and Evolution*, *3*(4), 864–871. https://doi.org/10.1002/ece3.490

te Beest, M., Le Roux, J. J., Richardson, D. M., Brysting, A. K., Suda, J., Kubešová, M., & Pyšek, P. (2012). The more the better? The role of polyploidy in facilitating plant invasions. *Annals of Botany*, *109*(1), 19–45. https://doi.org/10.1093/aob/mcr277

The Noxious Weed Act 1908. http://www.enzs.auckland.ac.nz/docs/1908/1908C133.pdf

Thomson, G. M. (1922). *The naturalisation of animals & plants in New Zealand*. Cambridge University Press.

Trewick, S., Morgan-Richards, M., & Chapman, H. (2004). Chloroplast DNA diversity of *Hieracium Pilosella* (Asteraceae) introduced to New Zealand: Reticulation, hybridization, and invasion. *American Journal of Botany*, *91*, 73–85. https://doi.org/10.3732/ajb.91.1.73

Vallejo-Marín, M., Friedman, J., Twyford, A. D., Lepais, O., Ickert-Bond, S. M., Streisfeld, M. A., Yant, L., van Kleunen, M., Rotter, M. C., & Puzey, J. R. (2021). Population genomic and historical analysis suggests a global invasion by bridgehead processes in *Mimulus guttatus*. *Communications Biology*, *4*, Article 327. https://doi.org/10.1038/s42003-021-01795-x

van Kleunen, M., Röckle, M., & Stift, M. (2015). Admixture between native and invasive populations may increase invasiveness of *Mimulus guttatus*. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1815), Article 20151487. https://doi.org/10.1098/rspb.2015.1487

Verhoeven, K. J. F., Macel, M., Wolfe, L. M., & Biere, A. (2011). Population admixture, biological invasions and the balance between local adaptation and inbreeding depression. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1702), 2–8. https://doi.org/10.1098/rspb.2010.1272

Ward, S. M., Gaskin, J. F., & Wilson, L. M. (2008). Ecological genetics of plant invasion: What do we know? *Invasive Plant Science and Management*, *1*(1), 98–109. https://doi.org/10.1614/IPSM-07-022.1

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag. https://ggplot2.tidyverse.org

Williams, J. T. (1971). Seed polymorphism and germination. II. The role of hybridization in the germination polymorphism of *Rumex crispus* and *R. obtusifolius*. *Weed Research*, *11*(1), 12–21. https://doi.org/10.1111/j.1365-3180.1971.tb00970.x

Wright, S. (1943). Isolation by distance. *Genetics*, *28*(2), 114–138.

Zaller, J. G. (2004). Ecology and non-chemical control of *Rumex crispus* and *R. obtusifolius* (Polygonaceae): A review. *Weed Research*, *44*, 414–432. https://doi.org/10.1111/j.1365-3180.2004.00416.x

Zlonis, K. J., & Gross, B. L. (2018). Genetic structure, diversity, and hybridization in populations of the rare arctic relict *Euphrasia hudsoniana* (Orobanchaceae) and its invasive congener *Euphrasia stricta*. *Conservation Genetics*, *19*(1), 43–55. https://doi.org/10.1007/s10592-017-0995-x

Chapter 6

Comparing trends in *de novo* versus reference-based sequencing approaches in invasion biology

6.1. Introduction

Reduced representation sequencing (RRS) methods are widely used in population genetics analyses as these methods are becoming cheaper, more accurate, and more accessible. These methods are used in both model and non-model species and with and without prior information of the genome, such as a reference genome (see for example Poland & Rife, 2012). In conservation biology, the use of RRS methods have surpassed the use of microsatellites (DeWoody et al., 2021; although see also Hauser et al., 2021), but in invasion biology microsatellites are still commonly used. In addition, using RRS methods for non-model species, especially of higher ploidies, can still feel like a massive undertaking (Paris et al., 2017). Indeed, despite being widely used, there is a lack of clear consensus when it comes to the gold standard in RRS methods from needing a reference genome to subsequently filtering the single-nucleotide-polymorphisms (SNPs) and all the way to how to perform the correct analyses, and it can be a daunting task to navigate through the various opinions. In this chapter, I will mainly focus on the value of using a reference genome versus using a *de novo* (Latin, "anew" or "from the beginning") approach, e.g., an approach without a reference genome to map the sequence against.

The most common RRS methods are genotyping-by-sequencing (GBS)(Elshire et al., 2011) and restriction-site associated DNA sequencing (RADseq)(Baird et al., 2008), and the two restriction enzymes utilising variations – or double-digest (often "dd", i.e., ddRAD) – of these two methods (Peterson et al., 2012; Poland et al., 2012). Both GBS and RADseq are used interchangeably as the umbrella terms for RRS methods (Andrews et al., 2016; Shafer et al., 2017; Elleouet & Aitken, 2018; Parchman et al., 2018), possibly because the methods do have a lot in common (for comparison of the methods see Andrews et al., 2016; or Ulaszewski et al., 2021). For example, both Elfekih et al. (2022) and Hodkinson et al. (2019) refer to their method as "genotyping by sequencing", but in the methods talk about ddRADseq. The power of these RRS methods lies in the ability to reduce the complexity of the genome using restriction enzymes, which cut genomes to smaller pieces and allow for the sequencing of the areas adjacent to the restriction sites as opposed to the whole genome (Baird et al., 2008). In addition, samples are barcoded with unique short sequences, which allows for the pooling of multiple samples in a single sequencer lane, reducing the cost of sequencing (Elshire et al., 2011; Andrews et al., 2016). The number of the cut pieces of the genome, and thus the number of potential markers, can be altered by selecting an appropriate restriction enzyme(s) (Elshire et al., 2011; Peterson et al., 2012; Poland et al., 2012).

These sequencing methods target SNPs, which are in high density within genomes, roughly once per 100–300 base pairs (Gupta et al., 2001). SNPs are often biallelic, which reduces their power to distinguish pairwise relationships (i.e., between two samples)

compared to the multiallelic microsatellites (Gupta et al., 2001; Glaubitz et al., 2003). At the same time, the utilisation of next-generation-sequencing techniques can provide thousands and sometimes tens of thousands of SNPs, which, for example, exceeds the estimated 100 SNPs needed to resolve parentage reliably (Anderson & Garza, 2006), and thus can help distinguish patterns which are not visible using microsatellites (Jeffries et al., 2016; Hodel et al., 2017).

But what does the current research say about using a reference genome versus a *de novo* approach? The reference genome approach offers – as the name suggests – a reference point, something to compare the results with, and more importantly: a map of physical genomic locations where each of the observed SNPs are located (e.g., in which chromosome or in which gene). Mapping can be especially useful with polyploid species, or species with many repetitive sequences. This is so that individual loci, especially homoeologous loci, can be correctly identified based on the originating sub-genomes, when without a reference genome they can appear to represent a single locus due to their very similar genetic sequences (Glover et al., 2016; Blischak et al., 2018). The same is true, however, if a reference genome has a different chromosome number or a lower ploidy level: when reference genome is missing chromosomes or chromosome copies, this can make homoeologous loci appear as a single locus (Nielsen, 2004). However, using a good quality reference genome will likely also reduce error rates (for example Fountain et al., 2016). Despite this, RRS methods have been used reliably without a suitable reference genome and are known to be capable of producing comparable information to reference based analyses (Rius et al., 2015; Torkamaneh et al., 2016; McCartney et al., 2019). In addition, when a species used as a reference is not closely related to the study species or when populations are genetically very different from each other, the use of a reference genome can also create ascertainment bias (Heslot et al., 2013; Shafer et al., 2015; Lang et al., 2020). This bias arises when some of the variation in the samples (e.g., the rare alleles) is not present in the reference and thus is not captured when the samples are mapped against the reference genome (Clark et al., 2005; Heslot et al., 2013). This can make the study populations appear more similar to each other, and these issues could be countered to some extent with a *de novo* approach (Paris et al., 2017).

Both GBS and RADseq have always advertised their ability to be used *de novo* (Baird et al., 2008; Elshire et al., 2011), contrast to widely used microsatellites, which often require to be developed separately for each study species (Zane et al., 2002; but see Abdelkrim et al., 2009). Indeed, multiple RRS studies comparing both *de novo* and reference-based approaches have reported similar results to each other (e.g., Poland et al., 2012; Stetter & Schmid, 2017; Martin et al., 2020; Kunvar et al., 2021; Omire et al., 2022), highlighting the usability of these methods without a reference genome, such as with non-model species. In addition, several studies report opting for a *de novo* approach, after finding that only a small percentage of SNPs are mapping onto the reference genome (Tripp et al., 2017; Helliwell et al., 2018; Cordeiro et al., 2020). Cordeiro et al. (2020) trialled both reference-based and *de novo* approaches in cotton bollworm *Helicoverpa armigera* Hübner (Lepidoptera) but after only 15 % of their markers were mapped onto the *H. armigera* reference genome, they opted for the *de novo* approach instead. Similarly,

Sherpa et al. (2018) chose the *de novo* approach after trialling three different reference genomes with a diverse range of issues: incomplete assembling (i.e., longer contigs or overlapping sequences instead of a full genome), poor mapping success of less than 25 %, and one of the reference genomes was developed for a laboratory reared strain that would likely introduce ascertainment bias if used.

However, contrasting results from studies that use both a reference-based and a *de novo* approach are also plentiful (e.g., Ilut et al., 2014; Shafer et al., 2015; Yang et al., 2017; Maroso et al., 2018; Ulaszewski et al., 2021). Maroso et al. (2018) suggested that using a reference genome makes the results more reproduceable, but studies comparing different SNP calling pipelines have shown large discrepancies between the methods (Shafer et al., 2015; Torkamaneh et al., 2016), so this might only apply when using the exact same pipeline with the exact same parameters. Some studies argue that *de novo* analyses require higher sample depth to counteract the error rate and amount of missing information naturally in RRS data (see for example Andrews et al., 2016; Fountain et al., 2016).

When it comes to the number of SNPs retained between *de novo* and reference-based approaches, some studies have found more SNPs with *de novo* (Tripp et al., 2017; Maroso et al., 2018) and others have found more with the reference-based approach (Kunvar et al., 2021; Omire et al., 2022), and sometimes it depends on the species, the approach, and the filtering method used (Tripp et al., 2017; Paudel et al., 2018; Ulaszewski et al., 2021). In addition, arguments can be made for "the more SNPs the better" (e.g., Elleouet & Aitken, 2018) and "a small number of reliable SNPs is better" (e.g., Yang et al., 2017), but neither of these arguments is necessarily coupled with just one of the approaches or with conclusive evidence to judge which one is better or produces more reliable results. Several studies indicate that the approach and methods should be tailored for the study question and species (Andrews et al., 2016; Scheben et al., 2017; Díaz-Arce & Rodríguez-Ezpeleta, 2019), but it is harder to then find information on what method suits which question, which can be overwhelming. Lastly, judging the success of the results is hard, if not impossible, especially if no other similar results exist from the study species and thus the results cannot be compared to previous studies.

The lack of consensus is evident, and the list of contrasting findings extends well beyond what is listed above. Keeping this in mind, this study investigated the existing literature and synthesised information from RRS studies in invasion biology with the aim to:

1) Estimate how often reference genomes are used in the field of biological invasions
2) Analyse how closely related the reference genomes are
3) Summarise the studies comparing these methods
4) Assess the value of reference-based and *de novo* approaches for three *Rumex* species

It is difficult to measure the success of the approaches or conclusively say when one is better than the other. However, I defined success in this study using the following common metrics: the number of SNPs, the quality of the SNPs as measured by depth and proportion of missing information, good separation of homoeologous loci from orthologous loci, low ascertainment bias, and good mapping to the reference-genome (> 50 %) if one is used. This chapter provides insight into what kind of reference genome is used, when one is used, and if there are specific questions better suited for *de novo* versus reference-based approaches. My research builds on the work by McCartney et al. (2019), who similarly compared reference-based and *de novo* approaches, but my focus is on the commonly used methods within the field of invasion biology in general, rather than listing the achievements of individual studies.

6.2. Methods

6.2.1. Literature search

A search for literature was made using Scopus, and a query (("genotyping-by-sequencing" OR "genotyping by sequencing" OR “ddRAD” OR “ddRADseq” OR “RAD” OR “RADseq” OR “RAD sequencing” OR “GBS sequencing”) AND ("invasive species" OR “alien species” OR “exotic species”) AND (SNP OR "single nucleotide polymorphism")) was used to search within all fields. The search was undertaken on 31 March 2022 and yielded articles from 2008 to 2022. All articles listed on the search date were extracted as a CSV file. From the chosen articles the main objective, the information of the study species (species and taxa), the reference sequence species if used and its relationship to the study species were recorded. The main study questions were categorised into five categories based on a preliminary overview of the articles to understand the types of questions typically asked: 1) population structure within one range (native or introduced); 2) population structure both within the introduced and native ranges; 3) establishment history; 4) population genetics, such as admixture, hybridisation, effective population size, and loci under selection; and 5) other, such as marker development or a method comparison. Only studies focusing on introduced species were included in the analyses, even when the studies were focusing on the species’ native range, and only when the study had conducted at least part of the sequencing themselves. These excluded studies focusing on breeding and crop development as well as studies using published genomes or *in silico* sequencing.

All species analysed using the *de novo* method were searched through NCBI database (National Library of Medicine, n.d.) to assess if a reference genome was published, and a note was made whether it was published before or after the study was conducted. This was done to estimate how common it is to not use a reference when it exists, but also to get a sense of whether the species were more common research species or not. If the publishing year for the reference genome coincided with that of the article, reference genome was considered unavailable, due to the time it takes to prepare a publication.

6.2.2. Comparison between reference-based and *de novo* approaches in *Rumex*

A comparison between a reference-based sequencing approach using *Rumex hastatulus* Baldw. as a reference and a *de novo* approach was done. The methods are detailed in Chapter 5 and the supplementary material within. The search for a reference genome revealed a lack of information not only from the conspecifics but the congenerics as well. The only species sequenced within the genus is *R. hastatulus*, a dioecious species with- a different base chromosome number compared to the hermaphrodite study species. The genome of a confamilial, buckwheat *Fagopyrum esculentum* Moench (Polygonaceae) has been sequenced as well, but this genome was trialled and only around 5 % of the reads mapped onto it, rendering it unusable for the purpose of this study.

Measuring the success of population genetics analyses is hard, especially without prior studies from the species to compare with. In this study, I considered the following parameters when assessing success: a reasonable number of SNPs is needed to reliably detect smaller differences between populations (i.e., over 100 SNPs is needed for parentage analyses Anderson & Garza, 2006); the sample depth tells us how many times a locus has been sequenced, meaning a higher depth increases reliability that a given locus is correct and not a sequencing error; too much missing information can make populations seem more similar than they are in reality; and low mapping rate to a reference genome can indicate differences between the study individuals and the reference individual, giving rise to ascertainment bias. However, there is no gold standard for any of these parameters and thus more weight should be put on the species' ecology and invasion history when known, and what kind of genetic patterns can be expected based on them. For example, whether it is known that a species was likely introduced multiple times and admixture is expected, which leads to observable intermediate genotypes between the different source population genotypes (Rius & Darling, 2014), or whether the species can reproduce by selfing, which likely decreases the amount of genetic variation (Hamrick & Godt, 1996). In addition, similarities between results from multiple parallel analyses should be taken into account (i.e., the reference-based and *de novo* approaches showing similar results, or F_{ST} values and Principal Components Analysis (PCA) results detailing similar patterns).

Thus, in Chapter 5, and Appendix B, I compared multiple population genetic metrics and considered the ecology of the species to assess the reliability within the *de novo* and reference-based approaches, and in this chapter, I compare a few key findings between the reference-based and *de novo* methods using samples just from the UK and New Zealand (NZ). I report basic summary statistics after filtering, such as the number of SNPs, a mean sample depth, the proportion of missing information, and the percentage mapping to the reference genome. In addition, I report F_{ST} values and show PCA plots, as these will reflect the population structuring and differentiation. F_{ST} values were done according to the methodology in Chapter 5, and PCA using the 'glPca' command from adegenet package (Jombart, 2008) in R (version 4.1.2., R Development Core Team, 2008). Figures were plotted using ggplot2 (Wickham, 2016) in R.

6.3. Results

6.3.1. The *de novo* and reference-based approaches are equally used in invasion biology

The literature search yielded 509 articles, and roughly 30 % or 146 articles were included in this study. This meant that 363 articles were outside the scope of this study. Roughly a third of the discarded articles were researching species that were not considered introduced species, and another 25 % of the articles were review articles. The rest of the articles had methods outside our scope (i.e., microsatellites, transcriptome analyse, SNP arrays, or whole genome sequencing), studies looking into crop or animal breeding, or both.

The 146 articles used in this study sequenced a total of 130 different species ranging from microbes to mammals and seaweed to trees (FIGURE 6.1). Roughly 80 % of the species were studied only once, whereas rainbow trout *Oncorhynchus mykiss* Walbaum (Salmoniformes), brown rat *Rattus norvegicus* Berkenhout (Rodentia), and Asian tiger mosquito *Aedes albopictus* Skuse (Diptera) were studied four, five, and seven times, respectively. A few articles looked at more than one species, and when the research was interested in both species separately (e.g., comparison of the population structure of two invasive species), these were counted as separate analyses, which means that the number of studies or analyses exceeds the number of articles. However, even when the article analysed more than one species, but their purpose was, for example, to estimate introgression within a parent species or when species were used as outgroups for a phylogeny, only the main study species was recorded.

Of the final set of articles, around 55 % used a double-digest and the rest used a single-digest method. Articles were not separated to GBS and RAD, given the similarities between the methods and given that many articles reported their methods vaguely and in uncertain terms. As an example, some articles talked about RADseq but cited articles associated with GBS and vice versa, or sometimes when an external organisation had conducted the sequencing, the articles had barely anything about the method used. The analyses were conducted *de novo* in 50 % of the cases, reference-based in 46 % and the remainder – or 5 studies – used both methods. When a reference was used, the majority used a genome from a conspecific (84 %), followed by congeneric (11 %), and finally a reference from a confamiliar species was used in just 5 % of the studies. A majority, or 68 out of 74 of the *de novo* analysed species had no reference genome available at the time of the studies in question. Out of these, only 17 species have since been sequenced fully, in addition to the 6 species sequenced before the time of the studies.

The main objective of the study was grouped into five categories. Studies assessing the population structure within one range were most common, around 30 % of the studies, followed by studies investigating population genetics, while the rest of the categories were relatively even (FIGURE 6.2). The *de novo* method was used more frequently in population structure analyses and less frequently in population genetics analyses than expected, when compared to the reference-based analyses (Pearson's Chi-squared test, χ^2

= 12.888, df = 4, p = 0.012). The studies comparing both *de novo* and reference-based methods were omitted from this analysis.

Five articles trialled or compared *de novo* and reference-based methods. Out of these, three articles opted for the *de novo* method after a low alignment percentage of under 50 % to the reference genome (Helliwell et al., 2018; Sherpa et al., 2018; Cordeiro et al., 2020). Two articles reported high concordance between the two approaches and presented the reference-based results in the publication (Pichler et al., 2019; Martin et al., 2020). For example, Martin et al. (2020) found very similar F_{ST} values indicating limited differentiation between populations of 0.0107 and 0.0084 with reference-based and *de novo* methods, respectively, whereas Pichler et al. (2019) found no clusters (K = 1) with ADMIXTURE analysis using both the *de novo* and the reference-based methods.

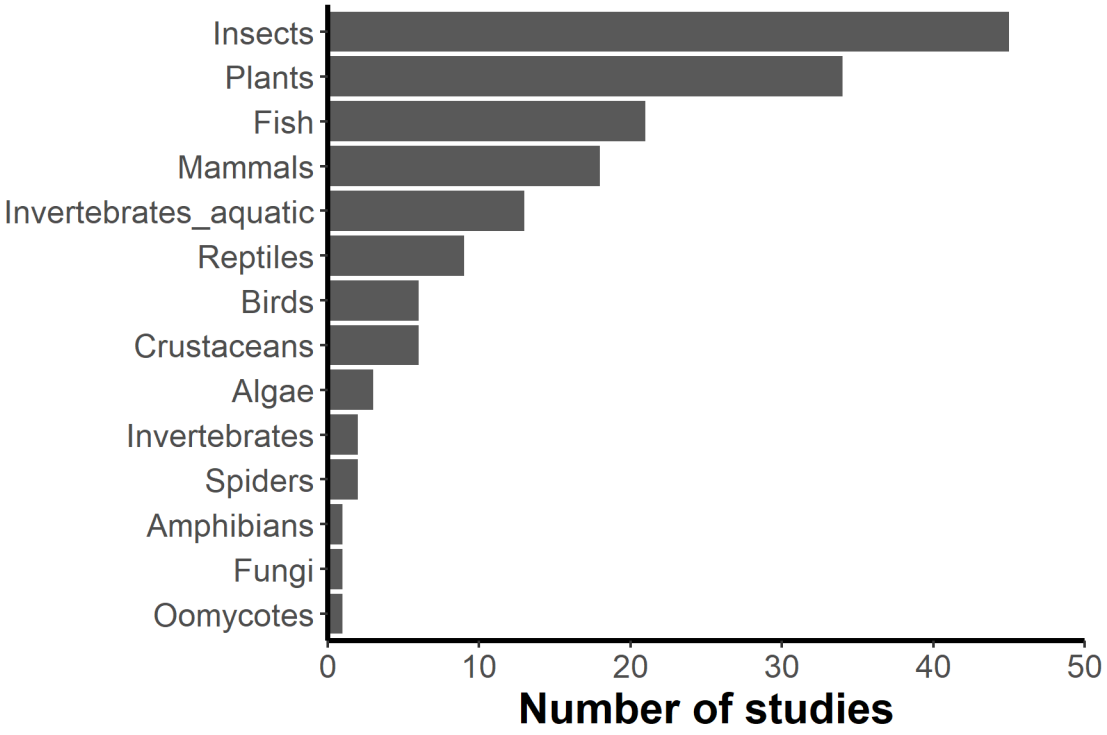

FIGURE 6.1: The wide range of taxa studied by the 146 articles included in this literature review.

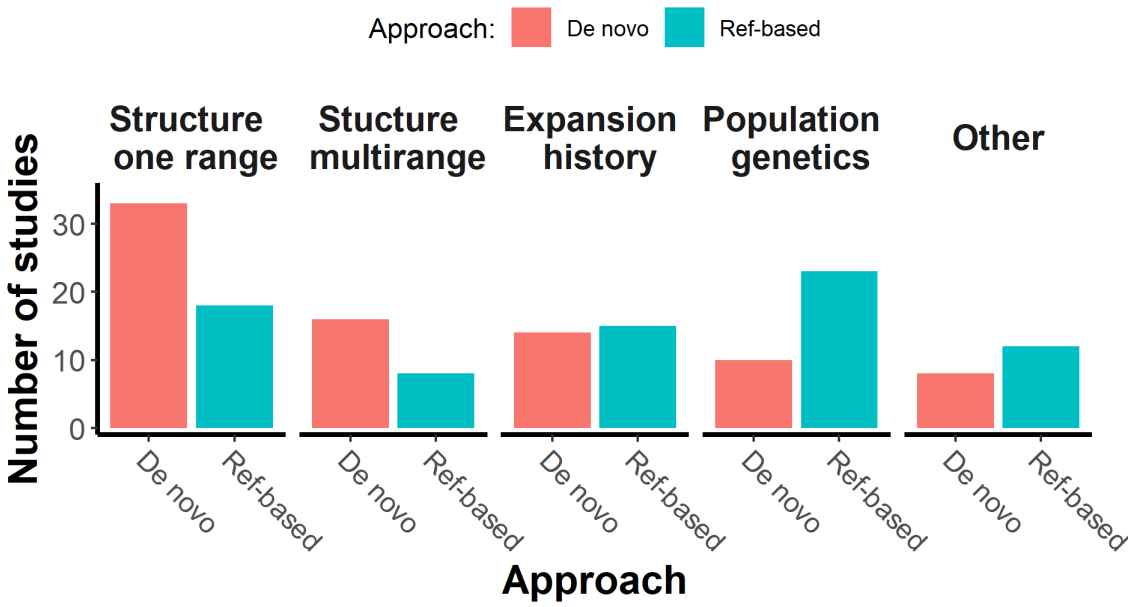

FIGURE 6.2: The main purpose of the compared 146 studies between *de novo* and reference-based methods. The five purpose categories are organised from highest to lowest according to the *de novo* method.

6.3.2. Both reference-based and *de novo* approaches yield similar results in *Rumex*

Based on the sequencing parameters after filtering (TABLE 6.1), the reference-based and the *de novo* approaches differed depending on the parameters considered. The sample sizes were comparable, and apart from *R. crispus*, both methods had the same number of samples. The 10 samples of *R. crispus* were excluded from the *de novo* approach due to the proportion of missing data exceeding the set limit of 0.3 (results not shown). The number of SNPs is roughly half with the *de novo* method compared to the reference-based method, and similarly, the SNP depth in the reference-based approach is more than six times higher. Around 17 % of the reads were mapped successfully onto the reference genome.

According to the F_{ST} values (FIGURE 6.3), the *de novo* approach shows consistently higher values compared to the reference-based analyses. However, as the averages in the two approaches are below 0.05 (i.e., little differentiation)(Hartl & Clark, 1997), this indicates that both the UK and NZ have little differentiation not only within, but also between the countries. This is echoed in the PCA plots comparing the two countries (FIGURE 6.4): There is high overlap between the samples from both countries rather than countries clustering separately.

TABLE 6.1: The number of *Rumex* samples collected from each country belonging to each of the three species, as well as the number of SNPs, a mean sample depth and a proportion of missing genotypes (= Prop. Missing) after filtering using KGD software. Each of the summary statistics is split in two: Ref = reference based and DN = *de novo* approach statistics. For summary before filtering, see Chapter 5 and Appendix B.

	#Samples		**#SNPs**		**Depth**		**Prop. Missing**	
	Ref	DN	Ref	DN	Ref	DN	Ref	DN
R. conglomeratus	328	238	2512	791	382	50	0.18	0.22
R. obtusifolius	329	329	2507	977	439	62	0.13	0.24
R. crispus	365	355	3538	1276	318	51	0.19	0.25

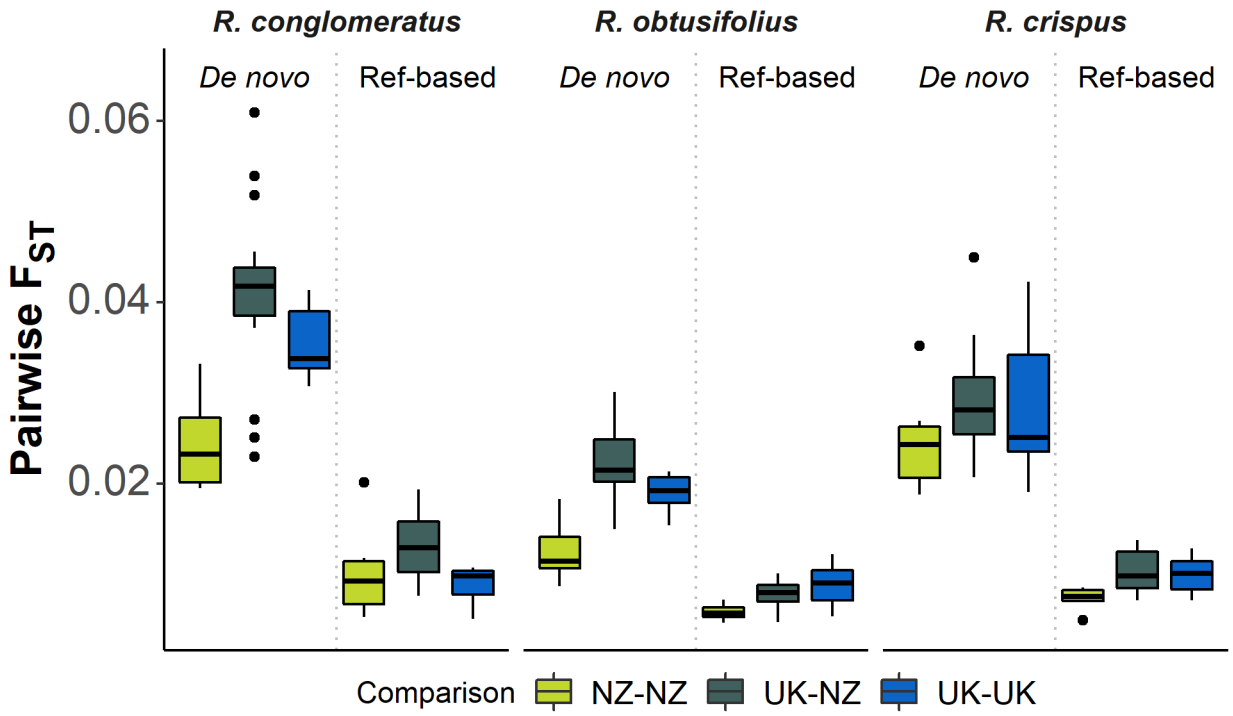

FIGURE 6.3: Boxplots of pairwise F_{ST} values within and between New Zealand and the UK. Comparisons are shown by *Rumex* species and between the *de novo* and reference-based approaches. The box represents the interquartile range with the median represented with a thick line and the whiskers extending from the 5^{th} to the 95^{th} percentile, with values outside this range (outliers) represented as dots.

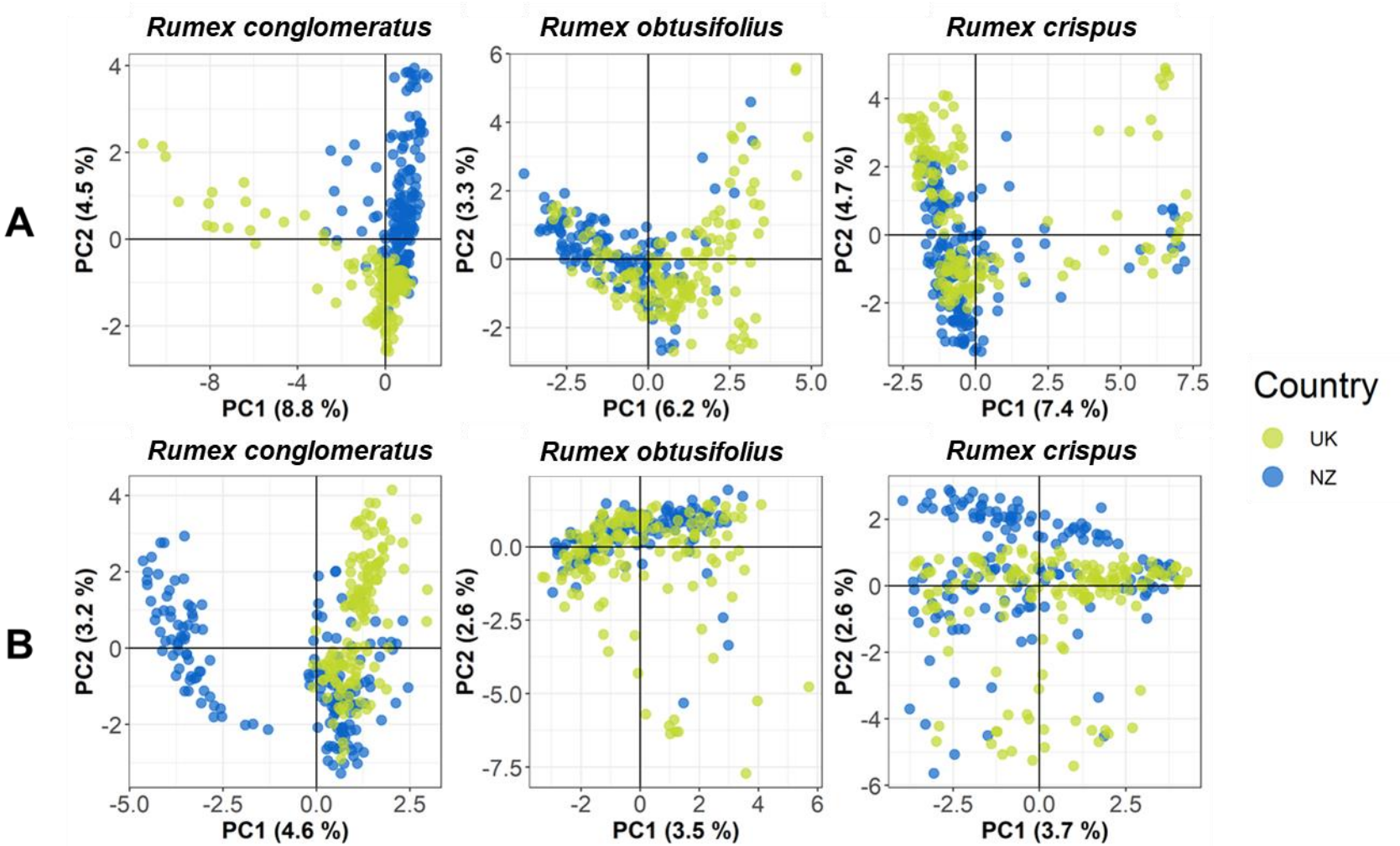

FIGURE 6.4: PCA plots comparing the *de novo* (A) and the reference-based (B) approaches between samples from the UK and New Zealand (NZ). All samples are mapped along the first and second axes, and the variation explained by each axis is provided

6.4. Discussion

The popularity of RRS based methods has made it possible to sequence more and more species more cheaply and with a higher accuracy. But it has not come without issues, one of the biggest being a clear lack of consensus when it comes to the methodology. In this study, I investigated existing literature within invasion biology to assess how often reference genomes are used versus how often studies use a *de novo* approach instead, and when a reference is used, what is the relationship between the study species and the species whose genome is used as a reference.

Roughly half of the investigated studies used a reference-based approach and a half a *de novo* method, and when a reference was used, it was most often from a conspecific. Consequently, usually when the study took a *de novo* approach it was due to no existing sequence for that specific species, and only 6 studies had a form of an existing reference at the time of the study but did not use one. Indeed, this is generally listed as both the biggest advantage and hindrance of RRS methods. It is an advantage, since sequencing of non-model species is possible without prior information (Baird et al., 2008; Elshire et al., 2011; Peterson et al., 2012; Poland et al., 2012), but a hindrance due to challenges in the data validation (see for example Cerca et al., 2021). Out of the six studies which did not use a reference even when it existed, one was creating a phylogeny and thus had other species in the article as well, one only had a commercial strain as a reference genome when in the article a wild type was studied, two had a scaffold assembly rather than a chromosome assembly available and the last reference was published only one calendar year earlier and might have coincided with the publishing process. None of these articles talked about an existing reference, and thus it can only be speculated whether decisions to use the *de novo* methods were made due to poor alignment and potential ascertainment bias (for example Heslot et al., 2013), not knowing that a suitable reference exists or just opting to utilise the potential of these methods without a reference genome.

When the main purposes of the studies were compared between the reference-based and *de novo* methods, some significant preferences were found (FIGURE 6.2). For example, population structure within the native or the introduced range or within both ranges was studied using a *de novo* method almost two times more often compared to the reference-based method, and in population genetics analyses the opposite was found to be true. This is likely due to spatial structure not being dependent on the exact genomic location of the SNPs, but rather requiring more individuals even with low coverage for population structure studies (e.g., McCartney et al., 2019), whereas the population genetics studies were often more interested in specific areas of the genome, which requires a reference genome.

McCartney et al. (2019) highlight, how the field of invasion genetics would benefit from a concerted effort to obtain well annotated reference genomes to better understand the genomic processes behind invasiveness. However, around 80 % of the species were studied only once, which could indicate that spending resources to create a reference genome is not feasible or even sensible (Neinavaie et al., 2021), as a majority

of these species seem to garner little interest within the field. Indeed, the concept of genetic diversity in invasive species management tends to be difficult, and genetic studies have been expensive compared to the management budget (Cook & Sgrò, 2018; Burgess et al., 2021). Thus, pooling resources could solve some problems, but deciding which species are more important than others is likely a non-trivial task.

The motivation for the approach chosen was rarely shared, but when it was, all *de novo* studies had chosen RRS *de novo* approach over other sequencing methods due to no requirement for prior genetic knowledge of the study species (data not shown). This motivation is not surprising, considering the applicability of these methods specifically in non-model species (Baird et al., 2008; Elshire et al., 2011; Peterson et al., 2012; Poland et al., 2012). However, generally the *de novo* studies did not openly consider the caveats or potential bias introduced by the lack of a reference genome. This perhaps suggests that the researchers are not overly worried about lacking a reference genome. Three out of five studies comparing both approaches opted for the *de novo* method due to low alignment of the sequence data to the reference genome but did not make further comments related to the benefits or issues this might have had. Motivation for the approach in studies using a reference genome was even lower, but the arguments behind the choice varied more. Some studies using the reference-based approach chose to use a reference genome to increase data accuracy, to avoid bias created by homeologs, and lastly, to get a chromosome level information. A few studies chose to create a reference for their RRS data to be mapped on, before conducting the RRS.

The lack of reporting motivation for the chosen approach is unsurprising, as there are already issues regarding reproducibility of the methods (e.g., Miller et al., 2020). Miller et al. (2020) looked into reporting reproducible methods in Discriminant Analysis of Principal Components, a method which assesses population structure (Jombart & Collins, 2015), and found that less than half of the articles reported methods clearly enough for repeatability.

6.4.1. Comparison between the approaches in *Rumex*

Some differences were found between the *de novo* and reference-based approaches in the results in *Rumex*. The SNP numbers differed from each other by the reference-based method yielding more than double compared to the *de novo* method (TABLE 6.1), but both methods yielded enough SNPs to be able to reliably differentiate between individuals (Anderson & Garza, 2006). Missing information in the raw data differed between the methods, and the reference-based approach yielded roughly 50 % less missing information compared to the *de novo* approach (see Chapter 5), but the differences were minimal after filtering (TABLE 6.1). In addition, Hodel et al. (2017) found that missing information will produce inflated estimates of differentiation but say that the commonly used cut-offs (proportions of around 0.3 and less) for missing information are small enough to not cause any issues. The missing values in all three *Rumex* species were below this cut-off using both methods, and the *de novo* approach excluded ten *R. crispus* samples with a higher proportion of missing information.

With less than 20 % of the data mapping to the reference genome, we can assume that the reference might not be the optimal match for the study species. In addition, the SNP depth for the reference-based approach was extremely high, which could indicate an inability to differentiate between repetitive sequences or homoeologous loci, potentially leading to an inflated number of heterozygotes (Nielsen, 2004; Heslot et al., 2013; Li, 2014; You et al., 2018). However, even with less than 15 shared SNPs between the datasets (data not shown), the results themselves show a very similar picture with both the reference-based and the *de novo* methods (FIGURE 6.3; FIGURE 6.4): The study species have limited differentiation and geographic structuring. The F_{ST} values are consistently low, and this measure of differentiation has been said to be the most robust between the varying approaches (Shafer et al., 2015; Hodel et al., 2017). Similarly, the PCA did not reveal strong differences between the countries in any of the three species.

While individual values are different and some species show a slightly different pattern (e.g., *R. obtusifolius* has the highest differentiation between UK and NZ in *de novo* analysis whereas the reference-based analysis shows similar levels of differentiation between the countries and within NZ), the data still supports the UK as a potential source for the introductions in NZ. In addition, the unified results suggest that either of these approaches can successfully produce results in these study species. Based on the metrics of success defined in the methods (higher number of samples, the number of SNPs, SNP depth, and proportion of missing information) the reference-based approach scores higher than the *de novo* approach. However, three out of five studies found in the literature search comparing both approaches chose the *de novo* approach based on low alignment alone (between 15–50 %, compared to 17 % in this study). In addition to the low alignment, the SNP depth in the reference-based approach suggests of poor differentiation of homoeologs or repetitive sequences, and the study species and the reference species differing in their ecology and genetic traits (namely chromosome number and ploidy), I opted for the *de novo* approach in Chapter 5.

6.4.2. Consequences from *de novo* approach

Perhaps the biggest potential risks in the *de novo* approach are severely underestimating the number of loci due to homeologs in polyploid species or the risk of under or overestimating similarities between populations due to missing information. However, there are filtering methods that are specifically developed to counter these biases (Attard et al., 2018), as well as general trouble-shooting guides for dealing with false heterozygotes or false homozygotes (see for example O'Leary et al., 2018). In addition, as the results from the *Rumex* approach comparison show, using a reference genome is not a guarantee of more reliable results.

6.4.3. Conclusions

This study investigated the existing literature related to RRS methods used in introduced and invasive species. As these methods are lacking a gold standard, I wanted to study when researchers have chosen a *de novo* approach and when they have used a reference-

based approach. My data highlight the usability of both approaches in very similar studies, and as the motivation for either of the approaches was rarely given, this question remains unsolved for now. A similar study conducted by McCartney et al. (2019) reached a different conclusion, and instead highlighted the need for good quality, well annotated reference genomes. They do, however, point out that creating these high-quality reference genomes is a big undertaking, and say that researchers should collaborate and focus their efforts on few, important species. This does not provide an immediate solution to invasion genetics, especially considering the vast number of species studied only once over the last 15 years.

In addition, according to the literature and our study comparing both *de novo* and reference-based of the approaches, often the results are very similar, and bigger differences are introduced by the SNP-calling pipelines and filtering methods. In the end, researchers can be fairly confident that either of the methods will work and are widely used, especially regarding population structure. This is especially true if careful decisions are made regarding data filtering, and that parameters (such as the SNP number and the proportion of missing information) are tailored to the study questions. Even then, the research is likely going to face criticism, given how subjective the preferences of certain methods within the scientific world can be.

6.5. 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(3), 185–192. https://doi.org/10.2144/000113084

Anderson, E. C., & Garza, J. C. (2006). The power of single-nucleotide polymorphisms for large-scale parentage inference. Genetics, 172(4), 2567–2582. https://doi.org/10.1534/genetics.105.048074

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. Nature Reviews Genetics, 17(2), 81–92. https://doi.org/10.1038/nrg.2015.28

Attard, C. R. M., Beheregaray, L. B., & Möller, L. M. (2018). Genotyping-by-sequencing for estimating relatedness in nonmodel organisms: Avoiding the trap of precise bias. Molecular Ecology Resources, 18(3), 381–390. https://doi.org/10.1111/1755-0998.12739

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E. U., Cresko, W. A., & Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLOS ONE, 3(10), Article e3376. https://doi.org/10.1371/journal.pone.0003376

Blischak, P. D., Kubatko, L. S., & Wolfe, A. D. (2018). SNP genotyping and parameter estimation in polyploids using low-coverage sequencing data. Bioinformatics, 34(3), 407–415. https://doi.org/10.1093/bioinformatics/btx587

Burgess, B. T., Irvine, R. L., Howald, G. R., & Russello, M. A. (2021). The promise of genetics and genomics for improving invasive mammal management on islands. Frontiers in Ecology and Evolution, 9, Article 704809. https://doi.org/10.3389/fevo.2021.704809

Cerca, J., Maurstad, M. F., Rochette, N. C., Rivera-Colón, A. G., Rayamajhi, N., Catchen, J. M., & Struck, T. H. (2021). Removing the bad apples: A simple bioinformatic method to improve loci-recovery in *de novo* RADseq data for non-model organisms. Methods in Ecology and Evolution, 12(5), 805–817. https://doi.org/10.1111/2041-210X.13562

Clark, A. G., Hubisz, M. J., Bustamante, C. D., Williamson, S. H., & Nielsen, R. (2005). Ascertainment bias in studies of human genome-wide polymorphism. Genome Research, 15(11), 1496–1502. https://doi.org/10.1101/gr.4107905

Cook, C. N., & Sgrò, C. M. (2018). Understanding managers' and scientists' perspectives on opportunities to achieve more evolutionarily enlightened management in conservation. Evolutionary Applications, 11(8), 1371–1388. https://doi.org/10.1111/eva.12631

Cordeiro, E. M. G., Pantoja-Gomez, L. M., de Paiva, J. B., Nascimento, A. R. B., Omoto, C., Michel, A. P., & Correa, A. S. (2020). Hybridization and introgression between *Helicoverpa armigera* and *H. zea*: An adaptational bridge. BMC Evolutionary Biology, 20(1), Article 61. https://doi.org/10.1186/s12862-020-01621-8

DeWoody, J. A., Harder, A. M., Mathur, S., & Willoughby, J. R. (2021). The long-standing significance of genetic diversity in conservation. Molecular Ecology, 30(17), 4147–4154. https://doi.org/10.1111/mec.16051

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

Elfekih, S., Metcalfe, S., Walsh, T. K., Cox, T. E., & Strive, T. (2022). Genomic insights into a population of introduced European rabbits *Oryctolagus cuniculus* in Australia and the development of genetic resistance to rabbit hemorrhagic disease virus. Transboundary and Emerging Diseases, 69, 895–902. https://doi.org/10.1111/tbed.14030

Elleouet, J. S., & Aitken, S. N. (2018). Exploring approximate bayesian computation for inferring recent demographic history with genomic markers in nonmodel species. Molecular Ecology Resources, 18(3), 525–540. https://doi.org/10.1111/1755-0998.12758

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE, 6(5), Article e19379. https://doi.org/10.1371/journal.pone.0019379

Fountain, E. D., Pauli, J. N., Reid, B. N., Palsbøll, P. J., & Peery, M. Z. (2016). Finding the right coverage: The impact of coverage and sequence quality on single nucleotide polymorphism genotyping error rates. Molecular Ecology Resources, 16(4), 966–978. https://doi.org/10.1111/1755-0998.12519

Glaubitz, J. C., Rhodes Jr, O. E., & Dewoody, J. A. (2003). Prospects for inferring pairwise relationships with single nucleotide polymorphisms. Molecular Ecology, 12(4), 1039–1047. https://doi.org/10.1046/j.1365-294X.2003.01790.x

Glover, N. M., Redestig, H., & Dessimoz, C. (2016). Homoeologs: What are they and how do we infer them? Trends in Plant Science, 21(7), 609–621. https://doi.org/10.1016/j.tplants.2016.02.005

Gupta, P., Roy, J., & Prasad, M. (2001). Single nucleotide polymorphisms: A new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plantsPK Gupta, JK Roy, M PrasadCurr Sci 80 (4), 524-35. Current Science, 80, 524–535.

Hamrick, J. L., & Godt, M. J. W. (1996). Effects of life history traits on genetic diversity in plant species. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 351(1345), 1291–1298. https://doi.org/10.1098/rstb.1996.0112

Hartl, D. L., & Clark, A. G. (1997). Principles of Population Genetics (3rd ed.). Sinauer Associates.

Hauser, S. S., Athrey, G., & Leberg, P. L. (2021). Waste not, want not: Microsatellites remain an economical and informative technology for conservation genetics. Ecology and Evolution, 11(22), 15800–15814. https://doi.org/10.1002/ece3.8250

Helliwell, E. E., Faber-Hammond, J., Lopez, Z. C., Garoutte, A., von Wettberg, E., Friesen, M. L., & Porter, S. S. (2018). Rapid establishment of a flowering cline in *Medicago polymorpha* after invasion of North America. Molecular Ecology, 27(23), 4758–4774. https://doi.org/10.1111/mec.14898

Heslot, N., Rutkoski, J., Poland, J., Jannink, J.-L., & Sorrells, M. E. (2013). Impact of marker ascertainment bias on genomic selection accuracy and estimates of genetic diversity. PLoS ONE, 8(9), Article e74612. https://doi.org/10.1371/journal.pone.0074612

Hodel, R. G. J., Chen, S., Payton, A. C., McDaniel, S. F., Soltis, P., & Soltis, D. E. (2017). Adding loci improves phylogeographic resolution in red mangroves despite increased missing data: Comparing microsatellites and RAD-Seq and investigating loci filtering. Scientific Reports, 7(1), Article 17598. https://doi.org/10.1038/s41598-017-16810-7

Hodkinson, T. R., Perdereau, A., Klaas, M., Cormican, P., & Barth, S. (2019). Genotyping by sequencing and plastome analysis finds high genetic variability and geographical structure in *Dactylis glomerata* L. in northwest Europe despite lack of ploidy variation. Agronomy, 9(7), Article 342. https://doi.org/10.3390/agronomy9070342

Ilut, D. C., Nydam, M. L., & Hare, M. P. (2014). Defining loci in restriction-based reduced representation genomic data from nonmodel species: Sources of bias and diagnostics for optimal clustering. BioMed Research International, 2014, Article e675158. https://doi.org/10.1155/2014/675158

Jeffries, D. L., Copp, G. H., Lawson Handley, L., Olsén, K. H., Sayer, C. D., & Hänfling, B. (2016). Comparing RADseq and microsatellites to infer complex phylogeographic patterns, an empirical perspective in the Crucian carp, *Carassius carassius*, L. Molecular Ecology, 25(13), 2997–3018. https://doi.org/10.1111/mec.13613

Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. Bioinformatics, 24(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T., & Collins, C. (2015, June 23). A tutorial for discriminant analysis of principal components (DAPC) using adegenet 2.0.0. https://adegenet.r-forge.r-project.org/files/tutorial-dapc.pdf

Kunvar, S., Czarnomska, S., Pertoldi, C., & Tokarska, M. (2021). In search of species-specific SNPs in a non-model animal (european bison (*Bison bonasus*))-comparison of *de novo* and reference-based integrated pipeline of STACKS using genotyping-by-sequencing (GBS) data. Animals, 11(8), Article 2226. https://doi.org/10.3390/ani11082226

Lang, P. L. M., Weiß, C. L., Kersten, S., Latorre, S. M., Nagel, S., Nickel, B., Meyer, M., & Burbano, H. A. (2020). Hybridization ddRAD-sequencing for population genomics of nonmodel plants using highly degraded historical specimen DNA. Molecular Ecology Resources, 20(5), 1228–1247. https://doi.org/10.1111/1755-0998.13168

Li, H. (2014). Toward better understanding of artifacts in variant calling from high-coverage samples. Bioinformatics, 30(20), 2843–2851. https://doi.org/10.1093/bioinformatics/btu356

Maroso, F., Hillen, J. E. J., Pardo, B. G., Gkagkavouzis, K., Coscia, I., Hermida, M., Franch, R., Hellemans, B., Van Houdt, J., Simionati, B., Taggart, J. B., Nielsen, E. E., Maes, G., Ciavaglia, S. A., Webster, L. M. I., Volckaert, F. A. M., Martinez, P., Bargelloni, L., & Ogden, R. (2018). Performance and precision of double digestion RAD (ddRAD) genotyping in large multiplexed datasets of marine fish species. Marine Genomics, 39, 64–72. https://doi.org/10.1016/j.margen.2018.02.002

Martin, S. L., Benedict, L., Wei, W., Sauder, C. A., Beckie, H. J., & Hall, L. M. (2020). High gene flow maintains genetic diversity following selection for high EPSPS copy number in the weed kochia (Amaranthaceae). Scientific Reports, 10(1), Article 18864. https://doi.org/10.1038/s41598-020-75345-6

McCartney, M. A., Mallez, S., & Gohl, D. M. (2019). Genome projects in invasion biology. Conservation Genetics, 20(6), 1201–1222. https://doi.org/10.1007/s10592-019-01224-x

Miller, J. M., Cullingham, C. I., & Peery, R. M. (2020). The influence of a priori grouping on inference of genetic clusters: Simulation study and literature review of the DAPC method. Heredity, 125(5), 269–280. https://doi.org/10.1038/s41437-020-0348-2

National Library of Medicine. (n.d.). National Center for Biotechnology Information. Retrieved April 13, 2022, from https://www.ncbi.nlm.nih.gov/

Neinavaie, F., Ibrahim-Hashim, A., Kramer, A. M., Brown, J. S., & Richards, C. L. (2021). The genomic processes of biological invasions: From invasive species to cancer metastases

and back again. Frontiers in Ecology and Evolution, 9, Article 681100. https://doi.org/10.3389/fevo.2021.681100
Nielsen, R. (2004). Population genetic analysis of ascertained SNP data. Human Genomics, 1(3), Article 218. https://doi.org/10.1186/1479-7364-1-3-218
O'Leary, S. J., Puritz, J. B., Willis, S. C., Hollenbeck, C. M., & Portnoy, D. S. (2018). These aren't the loci you're looking for: Principles of effective SNP filtering for molecular ecologists. Molecular Ecology, 27(16), 3193–3206. https://doi.org/10.1111/mec.14792
Omire, A., Neondo, J., Budambula, N. L. M., Wangai, L., Ogada, S., & Mweu, C. (2022). Genetic diversity and population structure of doum palm (*Hyphaene compressa*) using genotyping by sequencing. Frontiers in Genetics, 13, Article 762202. https://doi.org/10.3389/fgene.2022.762202
Parchman, T. L., Jahner, J. P., Uckele, K. A., Galland, L. M., & Eckert, A. J. (2018). RADseq approaches and applications for forest tree genetics. Tree Genetics & Genomes, 14(3), Article 39. https://doi.org/10.1007/s11295-018-1251-3
Paris, J. R., Stevens, J. R., & Catchen, J. M. (2017). Lost in parameter space: A road map for stacks. Methods in Ecology and Evolution, 8(10), 1360–1373. https://doi.org/10.1111/2041-210X.12775
Paudel, D., Kannan, B., Yang, X., Harris-Shultz, K., Thudi, M., Varshney, R. K., Altpeter, F., & Wang, J. (2018). Surveying the genome and constructing a high-density genetic map of napiergrass (*Cenchrus purpureus* Schumach). Scientific Reports, 8(1), Article 14419. https://doi.org/10.1038/s41598-018-32674-x
Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. PLoS ONE, 7(5), Article e37135. https://doi.org/10.1371/journal.pone.0037135
Pichler, V., Kotsakiozi, P., Caputo, B., Serini, P., Caccone, A., & della Torre, A. (2019). Complex interplay of evolutionary forces shaping population genomic structure of invasive *Aedes albopictus* in southern Europe. PLOS Neglected Tropical Diseases, 13(8), Article e0007554. https://doi.org/10.1371/journal.pntd.0007554
Poland, J. A., Brown, P. J., Sorrells, M. E., & Jannink, J. L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS ONE, 7(2), Article e32253. https://doi.org/10.1371/journal.pone.0032253
Poland, J. A., & Rife, T. W. (2012). Genotyping-by-sequencing for plant breeding and genetics. The Plant Genome, 5(3), 92–102. https://doi.org/10.3835/plantgenome2012.05.0005
R Development Core Team. (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org
Rius, M., Bourne, S., Hornsby, H. G., & Chapman, M. A. (2015). Applications of next-generation sequencing to the study of biological invasions. Current Zoology, 61(3), 488–504. https://doi.org/10.1093/czoolo/61.3.488
Rius, M., & Darling, J. A. (2014). How important is intraspecific genetic admixture to the success of colonising populations? Trends in Ecology & Evolution, 29(4), 233–242. https://doi.org/10.1016/j.tree.2014.02.003
Scheben, A., Batley, J., & Edwards, D. (2017). Genotyping-by-sequencing approaches to characterize crop genomes: Choosing the right tool for the right application. Plant Biotechnology Journal, 15(2), 149–161. https://doi.org/10.1111/pbi.12645
Shafer, A. B. A., Peart, C. R., Tusso, S., Maayan, I., Brelsford, A., Wheat, C. W., & Wolf, J. B. W. (2017). Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. Methods in Ecology and Evolution, 8(8), 907–917. https://doi.org/10.1111/2041-210X.12700
Shafer, A. B. A., Wolf, J. B. W., Alves, P. C., Bergström, L., Bruford, M. W., Brännström, I., Colling, G., Dalén, L., De Meester, L., Ekblom, R., Fawcett, K. D., Fior, S., Hajibabaei, M., Hill, J. A., Hoezel, A. R., Höglund, J., Jensen, E. L., Krause, J., Kristensen, T. N., … & Zieliński, P. (2015). Genomics and the challenging translation into conservation

practice. Trends in Ecology & Evolution, 30(2), 78–87. https://doi.org/10.1016/j.tree.2014.11.009

Sherpa, S., Rioux, D., Pougnet-Lagarde, C., & Després, L. (2018). Genetic diversity and distribution differ between long-established and recently introduced populations in the invasive mosquito *Aedes albopictus*. Infection, Genetics and Evolution, 58, 145–156. https://doi.org/10.1016/j.meegid.2017.12.018

Stetter, M. G., & Schmid, K. J. (2017). Analysis of phylogenetic relationships and genome size evolution of the *Amaranthus* genus using GBS indicates the ancestors of an ancient crop. Molecular Phylogenetics and Evolution, 109, 80–92. https://doi.org/10.1016/j.ympev.2016.12.029

Torkamaneh, D., Laroche, J., & Belzile, F. (2016). Genome-wide SNP calling from genotyping by sequencing (GBS) data: A comparison of seven pipelines and two sequencing technologies. PLoS ONE, 11(8), Article e0161333. https://doi.org/10.1371/journal.pone.0161333

Tripp, E. A., Tsai, Y.-H. E., Zhuang, Y., & Dexter, K. G. (2017). RADseq dataset with 90% missing data fully resolves recent radiation of *Petalidium* (Acanthaceae) in the ultra-arid deserts of Namibia. Ecology and Evolution, 7(19), 7920–7936. https://doi.org/10.1002/ece3.3274

Ulaszewski, B., Meger, J., & Burczyk, J. (2021). Comparative analysis of SNP discovery and genotyping in *Fagus sylvatica* L. and *Quercus robur* L. using RADseq, GBS, and ddRAD methods. Forests, 12(2), Article 222. https://doi.org/10.3390/f12020222

Wickham, H. (2016). ggplot2: Elegant graphics for data analysis. Springer-Verlag. https://ggplot2.tidyverse.org

Yang, X., Song, J., You, Q., Paudel, D. R., Zhang, J., & Wang, J. (2017). Mining sequence variations in representative polyploid sugarcane germplasm accessions. BMC Genomics, 18, Article 594. https://doi.org/10.1186/s12864-017-3980-3

You, Q., Yang, X., Peng, Z., Xu, L., & Wang, J. (2018). Development and applications of a high throughput genotyping tool for polyploid crops: Single nucleotide polymorphism (SNP) array. Frontiers in Plant Science, 9, Article 104. https://www.frontiersin.org/article/10.3389/fpls.2018.00104

Zane, L., Bargelloni, L., & Patarnello, T. (2002). Strategies for microsatellite isolation: A review. Molecular Ecology, 11(1), 1–16. https://doi.org/10.1046/j.0962-1083.2001.01418.x

Chapter 7

Discussion

7.1. Scope of the thesis

This thesis combined demo-genetic traits and processes to assess how they contribute to invasiveness in three *Rumex* species, *R. conglomeratus*, *R. obtusifolius*, and *R. crispus*. A comparison between native and introduced populations was made to help assess the origin of the invasions and to establish a baseline for all comparisons between the ranges. Generally, it is expected that invasion success can be aided by certain species' traits and processes, such as high genetic variation, polyploidy, multiple introductions and admixture, interspecific hybridisation and introgression, and self-compatibility (Colautti et al., 2006; Barrett et al., 2008; Catford et al., 2009; te Beest et al., 2012; Barker et al., 2018; Sherpa & Després, 2021). This list is by no means exhaustive but summarises the topics I investigated in this thesis. These traits and processes can be viewed within the general theories of plant invasions: propagule pressure, pre-adaptation or prior adaptation, and evolution following introductions (summarised recently in Young et al., 2022).

Based on the genetic similarities between plants from the UK and New Zealand (NZ), I have found strong support for the UK being one of the countries, where the introductions originated. In addition, I found no marked differences between the traits and processes investigated that could account for the higher success of these species in NZ compared to the UK and Europe more broadly. This likely indicates that multiple traits and processes are working in tandem instead of having a single strong driver, and that these species were well adapted before their introduction to NZ. As these species of *Rumex* are important invaders of human-altered environments, a special form of prior adaptation called anthropogenically induced adaptation to invade (Hufbauer et al., 2012) could explain the success of these species. The anthropogenically induced adaptation to invade (AIAI) can aid during invasions when species have adapted to human-altered habitats within their native range. If then similar manmade habitats are present in the introduced range, the species can adapt to successfully invade them. I will expand on the reasoning later in this chapter. Next, I will summarise the main findings of each of my data chapter and go over the interplay between the assessed demo-genetic traits and processes (for an overview of these, see FIGURE 7.1) and link them to previous studies investigating the invasiveness in these *Rumex* species. Finally, I will assess the limitations of the thesis, as well as describe potential future research directions that could help us understand more about the aspects behind invasiveness in *Rumex* in NZ.

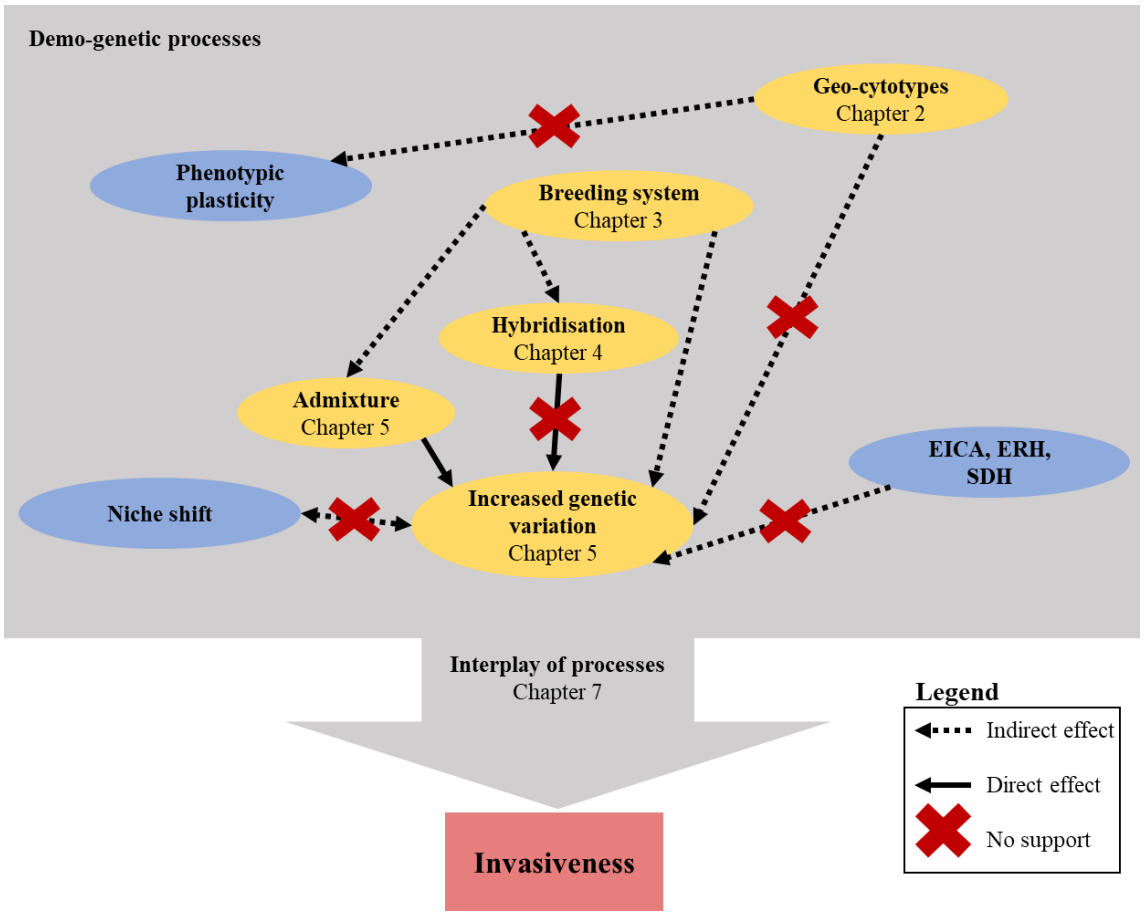

FIGURE 7.1: An overview of the different demo-genetic traits and processes assessed in this thesis (yellow), and their interplay with processes investigated before (blue), and whether they contribute to the increased invasiveness within New Zealand. In the figure hybridisation refers to interspecific hybridisation, whereas admixture refers to intraspecific hybridisation. EICA: evolution of increased competitive ability; ERH: enemy release hypothesis; SDH: shifting defence hypothesis.

7.2. Summaries of major findings

7.2.1. Absence of geo-cytotypes variation in three *Rumex* species within the UK and New Zealand (Chapter 2)

Research suggests that small genome size and polyploidy can benefit invasive species due to faster generation time, higher seed set, fixed heterosis, increased tolerance for inbreeding, and increased phenotypic plasticity (Pandit et al., 2006; Kubešová et al., 2010; te Beest et al., 2012; Bomblies & Madlung, 2014; Nagy et al., 2018). In addition, if a species has multiple geo-cytotypes, often the higher ploidies are found in the introduced range (Kubátová et al., 2008; Schlaepfer et al., 2008; Treier et al., 2009). Research shows that the three study species all have records of geo-cytotypes (Rice et al., 2015; Leitch et al., n.d.), and a study by Bufford & Hulme (2021a) suggests that both *R. conglomeratus* and *R. obtusifolius* show high phenotypic plasticity, but only the latter seems to benefit from it in the introduced range. Thus, I wanted to investigate, whether

the genome size or ploidy differs between plants from the UK and NZ, and whether this could explain the increased plasticity. This was assessed using a combination of flow cytometry and manual chromosome counts for verification. The analyses of roughly 25 plants per species per range yielded no differences in the genome sizes within or between the ranges suggesting that no geo-cytotypes exist in either the UK or in NZ. This suggests that the increased phenotypic plasticity observed in *R. obtusifolius* is not due to ploidy differences between the ranges. The results also hint that the introductions likely originated from countries, where the matching ploidy is the most prevalent one, such as in the UK. Furthermore, I showed that flow cytometry is a good tool differentiating between the species, and can identify hybrids, even when morphological tools fail.

Similar results have been found before, and higher ploidies are not exclusively better competitors than lower ploidies or diploids (Rutland et al., 2021). Tetraploid giant goldenrod *Solidago gigantea* Aiton (Asteraceae) performs better than hexaploid plants in a common garden experiment, despite the hexaploid plants having a significantly higher genetic diversity (Nagy et al., 2018). The tetraploid plants have been found in the introduced range, whereas hexaploid plants only within the native range, but based on the competitive ability, the authors suggest that the higher ploidy does not possess added risks. Similarly in the three species of *Rumex* the higher ploidies present within the native and parts of the introduced range might not show higher competitive ability compared to the cytotypes found in NZ. Alternatively, the differences in the cytotype performance could be more regional. Kumari et al. (2021) found tetraploid giant fescue *Festuca gigantea* L. (Poaceae) to have a wider geographic range within the Himalayas compared to hexaploid plants, but the hexaploid plants are larger than the tetraploids. However, there are indications that the tetraploid *F. gigantea* is more widespread outside the Himalayan region.

7.2.2. No inbreeding depression but lower germination within the introduced range in *Rumex conglomeratus* (Chapter 3)

Even a single plant capable of self-pollination can establish a population after being introduced. This can offer a huge advantage to invasive species, many of which, have been shown to have some level of self-compatibility (Baker, 1955; Barrett et al., 2008; Barrett, 2013). Pollen:ovule ratios can be used as a proxy for the breeding system, because plants have optimised pollen production to match their needs but not waste resources for excess amounts of pollen (Cruden, 1977). In this chapter, I investigated the pollen:ovule ratios of *R. conglomeratus*, as well as performed an experiment to assess the costs associated with selfing. I measured this using seed weight and germination percentage and compared the results between selfed and open-pollinated progeny. No differences were found in the P:O ratios between the ranges, and the ratios suggested that the breeding system of *R. conglomeratus* is facultative xenogamy (outcrossing). Contrastingly, all bagged plants yielded seeds, and there were no differences in the seed weights or germination percentages between the treatments. However, the germination percentage was significantly lower in plants from the introduced range compared to plants from the native range.

Wind-pollinated plants are known to produce more pollen compared to animal-pollinated plants, and variation between taxa exists (Cruden, 1977, 2000; Friedman &

Barrett, 2009, 2011). The lower germination of both treatments but especially of the selfed progeny from the introduced range might be due to the small sample size and chance effects, or fresher seeds, as *Rumex* seeds, including the seeds of *R. conglomeratus*, have been shown to exhibit higher germination percentages after a year of storage compared to fresh mature seeds (Benvenuti et al., 2001; Assche et al., 2002). The out-crossed seeds from NZ had approximately six months shorter storage period, than all the other seeds in the experiment, and their germination percentage ranged between 3–90 % (n = 3). However, this does not explain the lower germination in the selfing treatment, but as there were only three samples in this group as well, it is hard to disentangle the true patterns.

If the pattern observed represents the biology of the species accurately, selfing seems to lead to increased viability within the introduced range. Indeed, this was reflected in the average germination overall: the germination percentage was higher albeit non-significant in the selfed progeny compared to the bagged progeny. This could indicate an adaptation for selfing, which is generally linked to increased invasiveness and larger range (Razanajatovo et al., 2016). A lower overall germination within introduced region could indicate different germination requirements. In addition, in previous studies both *R. obtusifolius* and *R. crispus* have shown great variation in germination percentages, some of which can likely be accounted for differences between populations and seed ripeness (summarised in Cavers & Harper, 1964).

7.2.3. *Rumex* species co-occurrence and hybrid occurrence lower in the introduced range compared to the native range (Chapter 4)

Hybridisation and subsequent introgression can create new allele combinations that can benefit invasive species (Ellstrand & Schierenbeck, 2006; Mesgaran et al., 2016). This can be especially helpful in small, selfing populations to help counter inbreeding depression. The three *Rumex* species are known to hybridise freely and often in the native and introduced ranges, and introgress back to the parent species (Williams, 1971; Takahashi & Hanyu, 2015). However, no information exists regarding hybrids or introgression in NZ. I wanted to investigate whether hybrids do exist in the wild and whether there is any indication of introgression, but also to compare vegetation survey records to assess whether the parent species are co-occurring in NZ as frequently as in the UK. I collected leaf samples from individuals in multi-species populations and identified them using plant morphology and flow cytometry. Up to 20 individuals per species or per hybrid class were analysed from 20 different sites around Canterbury and the West Coast.

The vegetation survey records showed a significantly lower co-occurrence of congeners in NZ compared to the UK. However, in the field surveys hybrids were found in all populations, where the parent species were co-occurring. No clear relationship was found between the parent species co-occurrence and the proportion of hybrids found. In addition, hybridisation occurs in NZ almost as often as in the UK, but since the hybrids had an intermediate genome size compared to the parents and were thus assumed first generation hybrids, no clear indication of introgression was observed. This might be due to the lower viability found in hybrids (Cavers & Harper, 1964), or due to collecting

samples based on morphological identifications, rather than at random or by distance from the hybrid individual towards a nearest individual of the parent species.

However, there are multiple examples of introgression benefitting invasive species, such as increased environmental tolerance or acquired higher resistance against herbivores (Suarez-Gonzalez et al., 2018). In *Fallopia* spp., introgression from *F. sachalinensis* into *F. japonica* via first generation hybrids, is thought to explain some of the vast morphological variation and increased invasiveness of *F. japonica* (Gammon et al., 2007). Similar events in *Rumex* might explain their varying morphology, although my work could not confirm whether benefits from introgression are only pre-introduction, or also within NZ.

7.2.4. Low genetic differentiation between the native and introduced ranges, but evidence of multiple introductions to New Zealand in three *Rumex* spp. (Chapter 5)

Genetic variation is often deemed important for the survival and establishment of introduced species, because increased genetic variation can lead to increased adaptation ability (Dlugosch & Parker, 2008). Introductions were thought to lead to drastic reduction in genetic variation via genetic bottlenecks, but nowadays we know that many species retain high amounts of variation, often due to multiple introductions and admixture (Frankham, 2005; Dlugosch & Parker, 2008). The three *Rumex* species were introduced to NZ around mid-1800s, and as they arrived as grass seed contaminants and with dirty machinery, multiple introductions were likelier than bottlenecks. In addition, based on colonisation history, the introductions were likely originating from the UK (Thomson, 1922), but as this has never been investigated, I wanted to assess whether the UK could be the origin country. I also wanted to compare the genetic differentiation within and between the countries, as well as three outgroup countries from the native range. I sequenced all three species and compared four climatically matching regions within the UK and NZ as well as one region from France, Finland, and Czech Republic using genotyping-by-sequencing and population genetics analyses. The sequencing showed little genetic variation and differentiation within and between the populations. Most of the variation present in NZ was present in the UK, suggesting that the UK is likely the origin. However, some of the introductions could have originated elsewhere in Europe as well, or from bridgehead invasions from elsewhere within the introduced range.

Given the low genetic variation and differentiation within NZ, it is unlikely that there has been a genetic shift post-colonisation. Rather the introduced individuals were likely adapted to similar conditions in the native range. Both in the native and introduced regions, these species of *Rumex* inhabit pastures, ditches, and riversides, or what would mainly constitute of a human-altered habitat. In addition, they would have been introduced to pastures with European species and similar climates (Carlin et al., 2022). Thus, with the variation staying at similar levels compared to the native populations, AIAI likely contributed to their success as invasive species in NZ. Furthermore, the lack of bottlenecks, admixture and prior adaptation together could help to explain the early success of these species in NZ: only around 50 years after their introductions, they were

listed as noxious weeds, and their entry to NZ was prohibited (The Noxious Weed Act 1908).

7.2.5. Comparing trends in *de novo* versus reference-based sequencing approaches in invasion biology (Chapter 6)

In the last data chapter, I investigated whether reference-based or *de novo* approach are more common in invasion biology with restriction-site associated sequencing methods (namely RADseq, GBS). I wanted to see if these approaches were associated with certain types of questions more often compared to the alternate approach. In addition, I compared the approaches in *Rumex*, and assessed the success of the methods using a variety of commonly reported quality metrics. No clear patterns were found in the literature, but the *de novo* approach was more often associated with studies investigating population structure, and the reference-based approach when assessing population genetics metrics. Only five studies compared both methods, and out of these, three opted for the *de novo* approach due to low alignment to the reference genome, and two reported high similarities between the methods. Depending on the metrics used, the success of reference-based and *de novo* approaches in *Rumex* varied. Based on statistical measures such as the number of SNPs, SNP depth and proportion of missing information, the reference-based approached fared better compared to the *de novo* approach. However, when alignment to the reference genome and the taxonomic distance to the reference species were considered, the reference-based approach did not seem as suitable for *R. conglomeratus*, *R. obtusifolius*, and *R. crispus*. However, both approaches yielded very similar results, indicating high similarities between plants from the UK and NZ, and limited differentiation both between the countries but also within the countries.

7.3. Interplay between the demo-genetic traits and processes

The demo-genetic traits and processes investigated in this thesis can work in tandem with or separately from other traits and processes, such as increased phenotypic plasticity, increased environmental tolerance and niche shifts, and evolution of increased competitive ability, enemy release hypothesis, and shifting defence hypothesis (FIGURE 7.1). However, limited support was found for differentiation between the native and introduced populations and most of the investigated traits and processes are unlikely to contribute much to the increased invasiveness in NZ. This likely suggests that the opportunities for broader interplay as a main driver behind the invasiveness are also, diminished.

No geo-cytotypes were found, and thus the increased adaptive plasticity in drought conditions in *R. obtusifolius* (Bufford & Hulme, 2021a) is unlikely to be due to geo-cytotypes. However, as plasticity was only contributing to invasiveness in the tetraploid *R. obtusifolius* from NZ but not in the diploid *R. conglomeratus*, this might be due to ploidy. It has been shown that allopolyploids have higher heterozygosity, which could lead to higher phenotypic plasticity (te Beest et al., 2012). Since no geo-cytotypes were found in *R. obtusifolius*, allopolyploidy on its own is not enough to explain the differences

observed between the ranges, only the differences between the two species. The genetic analyses indicated admixture to European populations outside of the UK, which could help to explain the differences in plasticity between the native and introduced populations, however.

Another study found that all three species experience a significantly lower herbivory damage in NZ compared to the UK, but only *R. crispus* shows increased biomass in NZ (Costan et al., 2022). The authors suggest that enemy release could play a role in explaining the success of *R. crispus* but not the other two species. No drastic changes were observed in the genetic analyses I conducted, however. This can be due to changes happening outside the regions sequenced using genotyping-by-sequencing, or perhaps decreased herbivory leads to epigenetic changes that in turn impact growth. Herrera & Bazaga (2011) examined the effects of herbivory on genetic and non-genetic changes in *Viola cazorlensis* Gand. (Violaceae) and found a strong link between long-term herbivory and epigenetic variation. This study investigated the link between epigenetics and changes in plant defences against herbivory, but similarly when herbivory decreases, changes in epigenetics are likely to happen and impact plant size, for example (Bossdorf et al., 2010).

Previous study observing climate adaptation in native and introduced populations suggest that the native populations are pre-adapted to the analogue and non-analogue climates of the introduced ranges, rather than adaptation to non-analogue climates happening post-introduction in the introduced population (Carlin et al., 2022). My results showing low genetic differentiation support this result and the AIAI theory in general: no geographic differentiation was observed between the ranges, suggesting high similarities between the countries. Post-introduction adaptation to new climates would have likely manifested as a higher differentiation between the ranges, or at least between some of the assessed populations.

7.4. Future research directions

The ways of obtaining a deeper understanding of the invasiveness of the three *Rumex* species are plentiful. My genetic studies could benefit from investigating more countries, especially within the introduced region. One such country could be Australia, which has a similar and shared colonisation history, but more niche overlap with the native range compared with that between NZ and the native range (Carlin et al., 2022). This could further help to understand whether the patterns of admixture and genetic differentiation found within and between NZ and the UK are unique. In addition, increasing the number of countries within the introduced range could help obtaining information about potential bridgehead invasions as well (van Boheemen et al., 2017; Vallejo-Marín et al., 2021).

More research into the breeding system and especially the extent of selfing and mixed mating could better highlight the role of selfing, especially in *R. obtusifolius* and *R. crispus*. I aimed to assess the costs from selfing in all three species and wanted to investigate how ploidy could contribute as a buffer in inbreeding depression. Thus, broadening the bagging experiment to all three species with more samples both from NZ and the UK could help to shed light to this. The bagging experiment could be paired with a systematic assessment of the viability and fitness of hybrid seeds. Understanding the

consequences from both selfing and interspecific hybridisation to fitness during the full life cycle of a plant would better help to explain how these processes could aid during invasions. Furthermore, by comparing outcrossed, selfed and hybrid seeds within multiple ploidy levels would be beneficial in assessing the interplay between these traits and processes.

A cross-pollination experiment could help to understand whether admixture from European populations outside the UK could have contributed to the species' success and especially to the increased adaptive plasticity in *R. obtusifolius* (Bufford & Hulme, 2021a). However, the small size of the flowers and complicated floral patterns of the plants might make a cross-pollination experiment difficult to conduct. Nevertheless, cross-pollination experiments have shown that common yellow monkeyflower *Erythrante guttata* (Fisch. ex DC.) G.L. Nesom (Phrymaceae) benefits from heterosis when the native and introduced plants are crossed (van Kleunen et al., 2015). Admixture can also lead to heterosis when separated populations within the introduced range are crossed, which is observed in garlic mustard *Alliaria petiolata* (M. Bieb.) Cavara and Grande (Brassicaceae) and purple loosestrife *Lythrum salicarium* L. (Lythraceae) (Mullarkey et al., 2013; Shi et al., 2018). This highlights the need to control reintroductions, even when species have fully established. With *Rumex*, accidental introductions were likely limited after the species were added to the Noxious Weed Act in the early 20th century (The Noxious Weed Act 1908), but this legislation likely came after considerable propagule pressure and admixture had already happened.

An alternative method for aiding plants in adaptation, especially in rapidly changing conditions, is epigenetics (Banerjee et al., 2019; Ashapkin et al., 2020; Mounger et al., 2021). Epigenetic variation alters existing genetic variation usually by methylating certain sites in the DNA, which in turn alters the expression of genes (summarised in Ashapkin et al., 2020). This change alters secondary metabolites, which in turn can aid the species in adaptation to its new environment (e.g., Mounger et al., 2021). While some existing genetic variation is required for epigenetics to work quickly, epigenetic, and genetic variation do not have to work in tandem (Banerjee et al., 2019). As summarized in Ashapkin et al. (2020), epigenetic changes can be the first response for altering phenotypic variation, and is sometimes able to distinguish between populations, where genetic variation is not showing any significant difference. In addition, in genetically similar populations, phenotypic variation due to epigenetic changes is common (Banerjee et al., 2019). In addition, both polyploidisation and environmental stressors are known to change DNA methylation and lead to epigenetic changes in gene expression (summarised in Mounger et al., 2021). Thus, investigating epigenetic changes in *Rumex* could provide an important and interesting alternative mechanism behind the invasion success in NZ.

7.5. Conclusions and contributions to invasion biology

There are a multitude of demo-genetic traits and processes that could increase the invasiveness of plants. Scientists have worked hard to try and understand these, and the interplay between traits and environment. It has been suggested that universal traits behind invasiveness might not actually exist, but rather they represent the specific taxa and habitat in question (Radford & Cousens, 2000; Pyšek & Richardson, 2007; Catford

et al., 2009). Moreover, not a single trait or process investigated in my thesis or in the earlier works with these *Rumex* species (Carlin, 2021; Bufford & Hulme, 2021a, 2021b; Costan et al., 2022) could explain the increased invasiveness in all three species. However, there are considerable differences in the habitats these species occupy in NZ compared to the UK (Carlin, 2021), but as genetic differentiation was low, it is likely enabled by high phenotypic plasticity (Farris & Schaal, 1983; Grime et al., 2007; Bufford & Hulme, 2021a) rather than post-introduction evolution. Indeed, the success is likely caused by prior adaptation and the species having evolved into general purpose genotypes or jack-of-all-trades genotypes. Furthermore, anthropogenically induced adaptation to invade (Hufbauer et al., 2012) is likely the driver behind the success of these species.

These species of *Rumex* are not unique in this sense, but rather the phenomena of 'jack-of-all-trades' and anthropogenically induced adaptation to invade are known to invasion biology. Evening primrose *Oenothera biennis* L. (Onagraceae) shows no superiority in any of the characteristics compared to less invasive congeners, and the authors conclude that instead it can be considered a jack-of-all-trades, which might explain its success (Mihulka & Jarošík, 2006). Similarly, a study investigating the invasiveness of waterhemp *Amaranthus tuberculatus* (Moq.) Sauer (Amaranthaceae) suggests that it has prior adaptation to agricultural landscapes it now invades outside of its native range (Waselkov et al., 2020). In addition, the authors suggest that this adaptation to agricultural landscape could also increase the likelihood of fast, long-distance spread as a contaminant of machinery.

These three *Rumex* species are excellent model species for understanding invasiveness. They have been in NZ for almost two centuries (Darwin, 1839; Allan, 1937), which means that any post-introduction evolution would have likely occurred. None of the species are ornamental plants, which indicates that changes in the plants are due to natural selection rather than human selection by breeding. In addition, the species are fast-growing and hardy, they can tolerate drought and high levels of insect damage (Zaller, 2004; Costan et al., 2022), which makes them ideal for growing in the greenhouse. Since these species are classified as naturalised, they are also permitted into NZ, which enables experiments outside quarantine laboratories. However, compared to many model species, there is a considerable gap in knowledge regarding the genetics of these species (Zaller, 2004), and thus it can be difficult to find comparable information within other countries in the introduced range. This can be countered to an extent by studying three congenerics, instead of a single species, but even then, the observed patterns between the species can be inconsistent.

This thesis showed that all three species of *Rumex* are invasive in New Zealand without post-introduction evolution, but admixture from multiple populations have helped by countering potential bottlenecks during invasions. Understanding the drivers behind invasiveness requires investigating multiple traits and processes, the species' genetics being one of the most important. Assessing the genetic variation and partitioning helps not only in identifying the origin, but also in determining differentiation between native and introduced populations. Identifying the origin correctly means that all subsequent experiments are reflecting real differences between populations, rather than pre-existing variation outside the study populations (Sherpa & Després, 2021).

7.6. References

Allan, H. H. (1937). The origin and distribution of the naturalized plants of New Zealand. Linnean Society of London, 150(1), 25–46.

Ashapkin, V. V., Kutueva, L. I., Aleksandrushkina, N. I., & Vanyushin, B. F. (2020). Epigenetic mechanisms of plant adaptation to biotic and abiotic stresses. International Journal of Molecular Sciences, 21(20), Article e7457. https://doi.org/10.3390/ijms21207457

Assche, J. V., Nerum, D. V., & Darius, P. (2002). The comparative germination ecology of nine Rumex species. Plant Ecology, 159, 121–142.

Baker, H. G. (1955). Self-compatibility and establishment after "long-distance" dispersal. Evolution, 9(3), 347–349. https://doi.org/10.2307/2405656

Banerjee, A. K., Guo, W., & Huang, Y. (2019). Genetic and epigenetic regulation of phenotypic variation in invasive plants – linking research trends towards a unified framework. NeoBiota, 49, 77–103. https://doi.org/10.3897/neobiota.49.33723

Barker, B. S., Cocio, J. E., Anderson, S. R., Braasch, J. E., Cang, F. A., Gillette, H. D., & Dlugosch, K. M. (2018). Potential limits to the benefits of admixture during biological invasion. Molecular Ecology, 28(1), 100–113. https://doi.org/10.1111/mec.14958

Barrett, S. C. H. (2013). IV.8. evolution of mating systems: Outcrossing versus selfing. In J. B. Losos, D. A. Baum, D. J. Futuyma, H. E. Hoekstra, R. E. Lenski, A. J. Moore, C. L. Peichel, D. Schluter, & M. C. Whitlock (Eds.), The Princeton guide to evolution (pp. 356–362). Princeton University Press. https://doi.org/10.1515/9781400848065-050

Barrett, S. C. H., Colautti, R. I., & Eckert, C. G. (2008). Plant reproductive systems and evolution during biological invasion. Molecular Ecology, 17(1), 373–383. https://doi.org/10.1111/j.1365-294X.2007.03503.x

Benvenuti, S., Macchia, M., & Miele, S. (2001). Light, temperature and burial depth effects on *Rumex obtusifolius* seed germination and emergence. Weed Research, 41(2), 177–186. https://doi.org/10.1046/j.1365-3180.2001.00230.x

Bomblies, K., & Madlung, A. (2014). Polyploidy in the *Arabidopsis* genus. Chromosome Research, 22(2), 117–134. https://doi.org/10.1007/s10577-014-9416-x

Bossdorf, O., Arcuri, D., Richards, C. L., & Pigliucci, M. (2010). Experimental alteration of DNA methylation affects the phenotypic plasticity of ecologically relevant traits in *Arabidopsis thaliana*. Evolutionary Ecology, 24(3), 541–553. https://doi.org/10.1007/s10682-010-9372-7

Bufford, J. L., & Hulme, P. E. (2021a). Increased adaptive phenotypic plasticity in the introduced range in alien weeds under drought and flooding. Biological Invasions, 23(8), 2675–2688. https://doi.org/10.1007/s10530-021-02532-5

Bufford, J. L., & Hulme, P. E. (2021b). Seed size-number trade-offs are absent in the introduced range for three congeneric plant invaders. Journal of Ecology, 109, 3849–3860. https://doi.org/10.1111/1365-2745.13761

Carlin, T. F. (2021). Assessing range limits and niche shifts in invasive weeds [Doctoral Thesis, Lincoln University]. Research@Lincoln. https://hdl.handle.net/10182/14274

Carlin, T. F., Bufford, J. L., Hulme, P. E., & Godsoe, W. K. (2022). Global assessment of three *Rumex* species reveals inconsistent climatic niche shifts across introduced ranges. Research Square. https://doi.org/10.21203/rs.3.rs-1358439/v1

Catford, J. A., Jansson, R., & Nilsson, C. (2009). Reducing redundancy in invasion ecology by integrating hypotheses into a single theoretical framework. Diversity and Distributions, 15(1), 22–40. https://doi.org/10.1111/j.1472-4642.2008.00521.x

Cavers, P. B., & Harper, J. L. (1964). *Rumex obtusifolius* L. and *R. crispus* L. Journal of Ecology, 52(325), 737–766.

Colautti, R. I., Grigorovich, I. A., & MacIsaac, H. J. (2006). Propagule pressure: A null model for biological invasions. Biological Invasions, 8(5), 1023–1037. https://doi.org/10.1007/s10530-005-3735-y

Costan, C.-A., Godsoe, W. K., Bufford, J. L., Marris, J. W. M., & Hulme, P. E. (2022). Can the enemy release hypothesis explain the success of *Rumex* (Polygonaceae) species in an introduced range? Biological Invasions. https://doi.org/10.1007/s10530-022-02810-w

Cruden, R. W. (1977). Pollen-ovule ratios: A conservative indicator of breeding systems in flowering plants. Evolution, 31(1), 32–46. https://doi.org/10.2307/2407542

Cruden, R. W. (2000). Pollen grains: Why so many? Plant Systematics and Evolution, 222(1–4), 143–165. http://dx.doi.org/10.1007/BF00984100

Darwin, C. (1839). Narrative of the surveying voyages of His Majesty's Ships Adventure and Beagle between the years 1826 and 1836, describing their examination of the southern shores of South America, and the Beagle's circumnavigation of the globe (Vol. 3). Henry Colburn.

Dlugosch, K. M., & Parker, I. M. (2008). Founding events in species invasions: Genetic variation, adaptive evolution, and the role of multiple introductions. Molecular Ecology, 17(1), 431–449. https://doi.org/10.1111/j.1365-294X.2007.03538.x

Ellstrand, N. C., & Schierenbeck, K. A. (2006). Hybridization as a stimulus for the evolution of invasiveness in plants? Euphytica, 148, 35–46. https://doi.org/10.1007/s10681-006-5939-3

Farris, M. A., & Schaal, B. A. (1983). Morphological and genetic variation in ecologically central and marginal populations of *Rumex acetosella* L. (Polygonaceae). American Journal of Botany, 70(2), 246–255. https://doi.org/10.2307/2443270

Frankham, R. (2005). Resolving the genetic paradox in invasive species. Heredity, 94(4), Article 385. https://doi.org/10.1038/sj.hdy.6800634

Friedman, J., & Barrett, S. C. H. (2009). Wind of change: New insights on the ecology and evolution of pollination and mating in wind-pollinated plants. Annals of Botany, 103(9), 1515–1527. https://doi.org/10.1093/aob/mcp035

Friedman, J., & Barrett, S. C. H. (2011). The evolution of ovule number and flower size in wind-pollinated plants. The American Naturalist, 177(2), 246–257. https://doi.org/10.1086/657954

Gammon, M. A., Grimsby, J. L., Tsirelson, D., & Kesseli, R. (2007). Molecular and morphological evidence reveals introgression in swarms of the invasive taxa *Fallopia japonica*, *F. sachalinensis*, and *F. ×bohemica* (Polygonaceae) in the United States. American Journal of Botany, 94(6), 948–956. https://doi.org/10.3732/ajb.94.6.948

Grime, J. P., Hodgson, J. G., & Hunt, R. (2007). Comparative plant ecology: A functional approach to common British species (2nd ed.). Castlepoint Press.

Herrera, C. M., & Bazaga, P. (2011). Untangling individual variation in natural populations: Ecological, genetic and epigenetic correlates of long-term inequality in herbivory. Molecular Ecology, 20(8), 1675–1688. https://doi.org/10.1111/j.1365-294X.2011.05026.x

Hufbauer, R. A., Facon, B., Ravigné, V., Turgeon, J., Foucaud, J., Lee, C. E., Rey, O., & Estoup, A. (2012). Anthropogenically induced adaptation to invade (AIAI): Contemporary adaptation to human-altered habitats within the native range can promote invasions. Evolutionary Applications, 51(2), 89–101. https://doi.org/10.1111/J.1752-4571.2011.00211.X

Kubátová, B., Trávníček, P., Bastlová, D., Čurn, V., Jarolímová, V., & Suda, J. (2008). DNA ploidy-level variation in native and invasive populations of *Lythrum salicaria* at a large geographical scale. Journal of Biogeography, 35(1), 167–176. https://doi.org/10.1111/j.1365-2699.2007.01781.x

Kubešová, M., Moravcová, L., Suda, J., Jarošík, V., & Pyšek, P. (2010). Naturalized plants have smaller genomes than their non-invading relatives: A flow cytometric analysis of the Czech alien flora. Preslia, 82(1), 81–96.

Kumari, V., Singh, J., & Singhal, V. K. (2021). A report of intraspecific polyploidy (4x, 6x) in *Festuca gigantea* (L.) Vill (Family: Poaceae) from Western Himalaya, India and comments on its cyto-geography. The Nucleus, 64(2), 219–223. https://doi.org/10.1007/s13237-020-00347-8

Leitch, I. J., Johnston, E., Pellicer, J., Hidalgo, O., & Bennett, M. D. (n.d.). Plant DNA C-values Database | Royal Botanic Gardens, Kew. Retrieved October 14, 2021, from https://cvalues.science.kew.org/

Mesgaran, M. B., Lewis, M. A., Ades, P. K., Donohue, K., Ohadi, S., Li, C., & Cousens, R. D. (2016). Hybridization can facilitate species invasions, even without enhancing local adaptation. Proceedings of the National Academy of Sciences, 113(36), 10210–10214. https://doi.org/10.1073/pnas.1605626113

Mihulka, S., & Jarošík, V. (2006). Invasiveness of *Oenothera* congeners alien to Europe: Jack of all trades, master of invasion? Perspectives in Plant Ecology, Evolution and Systematics, 8(2), 83–96. https://doi.org/10.1016/j.ppees.2006.08.003

Mounger, J., Ainouche, M. L., Bossdorf, O., Cavé-Radet, A., Li, B., Parepa, M., Salmon, A., Yang, J., & Richards, C. L. (2021). Epigenetics and the success of invasive plants. Philosophical Transactions of the Royal Society B: Biological Sciences, 376(1826), Article 20200117. https://doi.org/10.1098/rstb.2020.0117

Mullarkey, A. A., Byers, D. L., & Anderson, R. C. (2013). Inbreeding depression and partitioning of genetic load in the invasive biennial *Alliaria petiolata* (Brassicaceae). American Journal of Botany, 100(3), 509–518. https://doi.org/10.3732/ajb.1200403

Nagy, D. U., Stranczinger, S., Godi, A., Weisz, A., Rosche, C., Suda, J., Mariano, M., & Pal, R. W. (2018). Does higher ploidy level increase the risk of invasion? A case study with two geo-cytotypes of *Solidago gigantea* Aiton (Asteraceae). Journal of Plant Ecology, 11(2), 317–327. https://doi.org/10.1093/jpe/rtx005

Pandit, M. K., Tan, H. T. W., & Bisht, M. S. (2006). Polyploidy in invasive plant species of Singapore. Botanical Journal of the Linnean Society, 151(3), 395–403. https://doi.org/10.1111/j.1095-8339.2006.00515.x

Pyšek, P., & Richardson, D. M. (2007). Traits associated with invasiveness in alien plants: Where do we stand? In W. Nentwig (Ed.), Biological Invasions (pp. 97–125). Springer. https://doi.org/10.1007/978-3-540-36920-2_7

Radford, I. J., & Cousens, R. D. (2000). Invasiveness and comparative life-history traits of exotic and indigenous *Senecio* species in Australia. Oecologia, 125(4), 531–542. https://doi.org/10.1007/s004420000474

Razanajatovo, M., Maurel, N., Dawson, W., Essl, F., Kreft, H., Pergl, J., Pyšek, P., Weigelt, P., Winter, M., & van Kleunen, M. (2016). Plants capable of selfing are more likely to become naturalized. Nature Communications, 7, Article 13313. https://doi.org/10.1038/ncomms13313

Rice, A., Glick, L., Abadi, S., Einhorn, M., Kopelman, N. M., Salman-Minkov, A., Mayzel, J., Chay, O., & Mayrose, I. (2015). The Chromosome Counts Database (CCDB) – a community resource of plant chromosome numbers. New Phytologist, 206(1), 19–26. https://doi.org/10.1111/nph.13191

Rutland, C. A., Hall, N. D., & McElroy, J. S. (2021). The impact of polyploidization on the evolution of weed species: Historical understanding and current limitations. Frontiers in Agronomy, 3, Article 626454. https://doi.org/10.3389/fagro.2021.626454

Schlaepfer, D. R., Edwards, P. J., Semple, J. C., & Billeter, R. (2008). Cytogeography of *Solidago gigantea* (Asteraceae) and its invasive ploidy level. Journal of Biogeography, 35(11), 2119–2127. https://doi.org/10.1111/j.1365-2699.2008.01937.x

Sherpa, S., & Després, L. (2021). The evolutionary dynamics of biological invasions: A multi-approach perspective. Evolutionary Applications, 14(6), 1463–1484. https://doi.org/10.1111/eva.13215

Shi, J., Macel, M., Tielbörger, K., & Verhoeven, K. J. F. (2018). Effects of admixture in native and invasive populations of *Lythrum salicaria*. Biological Invasions, 20, 2381–2393. https://doi.org/10.1007/s10530-018-1707-2

Suarez-Gonzalez, A., Lexer, C., & Cronk, Q. C. B. (2018). Adaptive introgression: A plant perspective. Biology Letters, 14(3), Article 20170688. https://doi.org/10.1098/rsbl.2017.0688

Takahashi, K., & Hanyu, M. (2015). Hybridization between alien species *Rumex obtusifolius* and closely related native vulnerable species *R. longifolius* in a mountain tourist destination. Scientific Reports, 5, Article 13898. https://doi.org/10.1038/srep13898

te Beest, M., Le Roux, J. J., Richardson, D. M., Brysting, A. K., Suda, J., Kubešová, M., & Pyšek, P. (2012). The more the better? The role of polyploidy in facilitating plant invasions. Annals of Botany, 109(1), 19–45. https://doi.org/10.1093/aob/mcr277

The Noxious Weed Act 1908. http://www.enzs.auckland.ac.nz/docs/1908/1908C133.pdf

Thomson, G. M. (1922). The naturalisation of animals & plants in New Zealand. Cambridge University Press.

Treier, U. A., Broennimann, O., Normand, S., Guisan, A., Schaffner, U., Steinger, T., & Müller-Schärer, H. (2009). Shift in cytotype frequency and niche space in the invasive plant *Centaurea maculosa*. Ecology, 90(5), 1366–1377. https://doi.org/10.1890/08-0420.1

Vallejo-Marín, M., Friedman, J., Twyford, A. D., Lepais, O., Ickert-Bond, S. M., Streisfeld, M. A., Yant, L., van Kleunen, M., Rotter, M. C., & Puzey, J. R. (2021). Population genomic and historical analysis suggests a global invasion by bridgehead processes in *Mimulus guttatus*. Communications Biology, 4, Article 327. https://doi.org/10.1038/s42003-021-01795-x

van Boheemen, L. A., Lombaert, E., Nurkowski, K. A., Gauffre, B., Rieseberg, L. H., & Hodgins, K. A. (2017). Multiple introductions, admixture and bridgehead invasion characterize the introduction history of *Ambrosia artemisiifolia* in Europe and Australia. Molecular Ecology, 26(20), 5421–5434. https://doi.org/10.1111/mec.14293

van Kleunen, M., Röckle, M., & Stift, M. (2015). Admixture between native and invasive populations may increase invasiveness of *Mimulus guttatus*. Proceedings of the Royal Society B: Biological Sciences, 282(1815), Article 20151487. https://doi.org/10.1098/rspb.2015.1487

Waselkov, K. E., Regenold, N. D., Lum, R. C., & Olsen, K. M. (2020). Agricultural adaptation in the native North American weed waterhemp, *Amaranthus tuberculatus* (Amaranthaceae). PLOS ONE, 15(9), Article e0238861. https://doi.org/10.1371/journal.pone.0238861

Williams, J. T. (1971). Seed polymorphism and germination. II. The role of hybridization in the germination polymorphism of *Rumex crispus* and *R. obtusifolius*. Weed Research, 11(1), 12–21. https://doi.org/10.1111/j.1365-3180.1971.tb00970.x

Young, S. L., Kimball, S., & Novak, S. J. (2022). Invasion of plant communities. In D. R. Clements, M. K. Upadhyaya, S. Joshi, & A. Shrestha (Eds.), Global plant invasions (pp. 29–51). Springer.

Zaller, J. G. (2004). Ecology and non-chemical control of *Rumex crispus* and *R. obtusifolius* (Polygonaceae): A review. Weed Research, 44, 414–432. https://doi.org/10.1111/j.1365-3180.2004.00416.x

Appendix A

Database search for chromosome numbers (Chapter 2)

In Chapter 2, I assessed the geo-cytotypes of *R. conglomeratus*, *R. obtusifolius*, and *R. crispus*. Below, I present the findings from The Chromosome Counts Database (CCDB; Rice et al., 2015) and The Plant C-values Database (Leitch et al., n.d.) for each species.

A.1. Chromosome numbers for *Rumex conglomeratus*

TABLE A.2: Geo–cytotypes for *Rumex conglomeratus*.

Range	Country/Region	2n	Reference
introduced	Japan	18	Sugiura, T. (1936). A list of chromosome numbers in angiospermous plants. II. Proceedings of the Imperial Academy, 12(5), 144–146.
introduced	Japan	20	Himi, H. (1999). Chromosome numbers of 11 species in Japanese *Rumex* subg. *Rumex* (Polygonaceae). Journal of Phytogeography and Taxonomy, 47, 121–130.
introduced	USA	20	Löve, A., & Löve, D. (1986). IOPB Chromosome number reports 93. Taxon, 35, 897–899.
introduced	Japan	20	Takenaka, Y. (1941). The relation between polyploidy and the size of stoma. I. On the plants of the subgenus Lapathum. Botanical Magazine, Tokyo, 55, 319–323.
introduced	Japan	40	Kihara H. & Ono T. (1923). Cytological studies on *Rumex* L. I. Chromosomes of *Rumex acetosa* L. II. On the relation of chromosome number and sexes in *Rumex acetosa* L. Botanical Magazine, Tokyo, 37, 84–90.
native	Scandinavia	18	Löve, A. (1942). Cytogenetic studies in *Rumex*: III. Some notes on the Scandinavian species of the genus. Hereditas, 28(3-4), 289–296.
native	Germany	18	Tischler G. (1934). Die Bedeutung der Polyploidie für die Verbreitung der Angiospermen, erläutert an den Arten Schleswig–Holsteins, mit Ausblicken auf andere Florengebiets. Botanische Jahrbücher fur Systematik, Pflanzengeschichte und Pflanzengeographie, 67, 1–36.
native	Europe	20	Degraeve, N. (1975). Contribution a L'Etude Cytotaxonomique des *Rumex*—I. Le Genre *Rumex* L. Sensu Stricto. Caryologia, 28(2), 187–201.

native	Austria	20	Dobeš, C., Kiehn, M. & Vitek, E. (1996). Beiträge zur Gefässpflanzen–Flora von Österreich: Chromosomenzählungen III Verhandlungen der Zoologisch–botanischen Gesellschaft, 133, 301–318.
native	Spain	20	Garcia, C., Pastor Díaz, J. E., & Luque Palomo, M. T. (1989). Contribución al estudio cariológico del género *Rumex* (Polygonaceae). Acta Botanica Malacitana, 14, 129–140.
native	Slovakia	20	Hindakova, M. (1976). Index of chromosome numbers of Slovakian flora. (Part 5). Acta Facultatis Rerum Naturalium Universitatis Comenianae. Botanica, 25, 1–18.
native	UK	20	Holyoak, D. T. (2000). Hybridisation between *Rumex rupestris* Le Gall (Polygonaceae) and other docks. Watsonia, 23, 83–92.
native	Switzerland	20	Huber, W. & Baltisberger, M. (1989). IOPB Chromosome Data 1. International Organization of Plant Biosystematists. Newsletter, 13, 19–20.
native	Poland	20	Jankun, A. (1990). In Further studies in chromosome numbers of Polish angiosperms, part 23. Acta Biologica Cracoviensia s. Botanica, 32, 172, 177–179, 181–183.
native	Germany	20	Jaretzky, R. (1928). Histologische and karyologische Studien an Polygonaceen. Jahrbücher für Wissenschaftliche Botanik, 69, 357–490.
native	France	20	Labadie, J. P. (1976). Reports [In Löve, A. (ed.), IOPB chromosome numbers reports LIV]. Taxon 25(5/6), 636–639.
native	Spain	20	Löve, A. & Kjellqvist, E. (1974). Cytotaxonomy of Spanish plants III. Dycotiledons: Salicaceae–Rosaceae. Lagascalia 4(1), 3–32.
native	Scandinavia	20	Löve, A. (1942). Cytogenetic studies in *Rumex*. III. Some notes on the Scandinavian species of the genus. Hereditas, 28, 289–296.
native	Spain	20	Löve, A. (1967). Reports [In Löve, A. (ed.), IOPB chromosome numbers reports XIII]. Taxon, 16(5), 445–461.
native	Sweden	20	Löve, A. (1967). Reports [In Löve, A. (ed.), IOPB chromosome numbers reports XIII]. Taxon, 16(5), 445–461.
native	Scandinavia	20	Löve, A. (1942). in Darlington CD, Wylie AP. *Chromosome atlas of flowering plants*. George Allen and Unwin Ltd.
native	Sweden	20	Lövkvist, B. & Hultgård, U. M. (1999). Chromosome numbers in south Swedish vascular plants. Opera Botanica, 137, 1–42.
native	Slovakia	20	Májovský, J. (1976). Index of chromosome numbers of Slovakian flora (Part 5). Acta Facultatis Rerum Naturalium Universitatis Comenianae. Botanica, 25, 1–18.
native	Slovakia	20	Májovský J., Uhríková A., Javorcíková D., Micieta K., Králik E., Dúbravcová Z., Feráková V., Murín A., Cernušáková D., Hindáková M., Schwarzová T. & Záborský J. (2000). Prvý doplnok karyotaxonomického prehladu flóry Slovenska. Acta Facultatis Rerum Naturalium Universitatis Comenianae. Botanica Supplementum, Nr 1, 1–127.

native	Spain	20	Pavone, P., & Zizza, A. (1981). Números cromosomáticos de plantas occidentales, 113–128. In Anales del Jardín Botánico de Madrid, 38(1), 273–280.
native	Poland	20	Pogan E., Jankun A. & Wcisło, H. (1990). Further studies in chromosome numbers of Polish Angiosperms. Part XXIII. Acta Biologica Cracoviensia, Series Botanica, 32, 171–188.
native	Portugal	20	Queirós, M. (1985). Números cromosómicos para a flora Portuguesa. Boletim da Sociedade Broteriana, Ser. 2, 58, 85–96.
native	Mediterranean	20	Runemark, H. (1996). Mediterranean chromosome number reports 6 (590–678). Flora Mediterranea, 6, 223–243.
native	Belarus	40	Dmitrieva, S. A. (1986). Chisla khromosom nekotorych vidov rastenij Berezinskogo Biosfernogo Zapovednika. Zapovedniki Belorussii Issledovaniia, 10, 24–28.
native	Germany	40	Jaretzky, R. (1927). Einige Chromosomenzahlen aus der Familie der Polygonaceae. Berichte der Deutschen Botanischen Gesellschaft, 45, 48–54.

A.2. Chromosome numbers for *Rumex obtusifolius*

TABLE A.3: Geo-cytotypes for *Rumex obtusifolius*.

Range	Country/Region	2n	Reference
introduced	Japan	40	Himi, H. (1999). Chromosome numbers of 11 species in Japanese *Rumex* subg. *Rumex* (Polygonaceae). Journal of Phytogeography and Taxonomy, 47, 121–130.
introduced	Canada	40	Löve, A. (1986). In Chromosome numbers reports XCII. Taxon 35(3), 611–613.
introduced	Canada	40	Mulligan, G. A. (1959). Chromosome numbers of Canadian weeds II. Canadian journal of botany 37, 81–92.
introduced	Canada	40	Taylor, R. L., & Mulligan, G. A. (1968). *Flora of the Queen Charlotte Islands. Part 2. Cytological aspects of the vascular plants.* Queen's Printer.
introduced	Japan	60	Himi, H. (1999). Chromosome numbers of 11 species in Japanese *Rumex* subg. *Rumex* (Polygonaceae). Journal of Phytogeography and Taxonomy, 47, 121–130.
native	Spain	20	Löve, A. & Kjellqvist, E. (1974). Cytotaxonomy of Spanish plants III. Dycotiledons: Salicaceae–Rosaceae. Lagascalia 4(1), 3–32.
native	UK	40	Al–Bermani, A. K. K. A., Al–Shammary, K. I. A., Gornall, R. J., & Bailey, J. P. (1993). Contribution to a cytological catalogue of the British and Irish flora, 3. Watsonia, 19, 169–171.
native	Europe	40	Degraeve, N. (1975). Contribution a L'Etude Cytotaxonomique des *Rumex*—I. Le Genre *Rumex* L. Sensu Stricto. Caryologia, 28(2), 187–201.
native	Belarus	40	Dmitrieva, S. A. (1986). Chisla khromosom nekotorych vidov rastenij Berezinskogo Biosfernogo Zapovednika. Zapovedniki Belorussii Issledovaniia, 10, 24–28.
native	Netherlands	40	http://wbd.etibioinformatics.nl/bis/flora.php?selected=beschrijving&menuentry=soorten&id=2061
native	Switzerland	40	Huber, W. & Baltisberger, M. (1989). IOPB Chromosome Data 1 (ed. C.A. Stace). International Organization of Plant Biosystematists. Newsletter, 13, 19–20.
native	Poland	40	Jankun, A., 1990. In Further studies in chromosome numbers of Polish angiosperms, part 23. Acta biologica Cracoviensia. Series botanica, 32, 172, 177–179, 181–183.
native	Europe	40	Löve, A. & Löve, D. (1961). Chromosome numbers of central and northwest European plant species. Opera Botanica, 5, 1–581.
native	Scandinavia	40	Löve, A. (1942). Cytogenetic studies in *Rumex*. III. Some notes on the Scandinavian species of the genus. Hereditas, 28, 289–296.

native	Sweden	40	Lövkvist, B. & Hultgård, U. M. (1999). Chromosome numbers in south Swedish vascular plants. Opera Botanica, 137, 1–42.
native	Slovenia	40	Májovský J., Uhríková A., Javorcíková D., Micieta K., Králik E., Dúbravcová Z., Feráková V., Murín A., Cernušáková D., Hindáková M., Schwarzová T. & Záborský J. (2000). Prvý doplnok karyotaxonomického prehladu flóry Slovenska. – Acta Facultatis Rerum Naturalium Universitatis Comenianae. Botanica. Supplementum Nr. 1, 1–127.
native	UK	40	Mowforth, M.A. (1986). *Variation in nuclear DNA amounts in flowering plants: An ecological analysis.* [Doctoral thesis, University of Sheffield].
native	Poland	40	Pogan E., Jankun A. & Wcisło, H. (1990). Further studies in chromosome numbers of Polish Angiosperms. Part XXIII. Acta Biologica Cracoviensia, Series Botanica, 32, 171–188.
native	Spain	40	Bermejo, E. V., & García, J. G. (1976). Notas cariosistemáticas sobre flora española, I. Acta Botanica Malacitana, 2, 39–50.
native	UK	40	Wentworth, J. E., J. P. Bailey & Gornall, R. J. (1991). Contributions to a cytological catalogue of the British and Irish flora, 1. Watsonia, 18, 415–417.
native	UK	40	Williams, J.T. (1971). Seed polymorphism and germination. 2. The role of hybridization in germination polymorphism of *Rumex crispus* and *Rumex obtusifolius*. Weed Research, 11, 12–21.
native	Russia	40	Probatova, N. S. & A. P. Sokolovskaya. 1989. Chromosome numbers in vascular plants from Primorye Territory, the Amur region, Sakhalin, Kamchatka and the Kuril Islands. Botanicheskii Zhurnal (Moscow & Leningrad), 74, 120–123.
native	Greece	40	Strid, A. & Franzen, R. (1981). In Chromosome number reports LXXIII. Taxon, 30, 829–842.
native	Spain	60	Castroviejo, S. (2003). *Flora Iberica. Plantas vascularesde la Península Ibérica e Islas Baleares.* Publicaciones del CSIC Madrid.

A.3. Chromosome numbers for *Rumex crispus*

TABLE A.4: Geo-cytotypes for *Rumex crispus*.

Range	Country/Region	2n	Reference
introduced	Australia	60	Degraeve, N. (1975) Contribution a L'Etude Cytotaxonomique des *Rumex*—I. Le Genre *Rumex* L. Sensu Stricto, Caryologia. International Journal of Cytology, Cytosystematics and Cytogenetics, 28(2), 187–201.
introduced	USA	60	Graham S.A. & Wood, C.E. (1965). The genera of Polygonaceae in the southeastern United States. Journal of the Arnold Arboretum 46(2), 91–121.
introduced	USA	60	Heiser, C.B. & Whitaker, T.W. (1948). Chromosome Number, Polyploidy, and Growth Habit in California Weeds. American Journal of Botany 35(3), 179–186.
introduced	USA	60	Hill, L. M. (1995). IOPB chromosome data 10. International Organization of Plant Biosystematists. Newsletter, 25, 8–9.
introduced	USA	60	Jensen, H. W. (1936). Meiosis in *Rumex*. I. Polyploidy and the origin of new species. Cytologia, 7, 1–22.
introduced	Canada	60	Löve, A. (1986). Reports [In Löve, A. (ed.), IOPB chromosome numbers reports XCII]. Taxon, 35(3), 611–613.
introduced	USA	60	Löve, A. (1967). Reports [In Löve, A. (ed.), IOPB chromosome numbers reports XIII]. Taxon, 16(5), 445–461.
introduced	Canada	60	Mulligan, G. A. (1957). Chromosome numbers of Canadian weeds. I. Canadian Journal of Botany, 35(5), 779–789.
introduced	Canada	60	Taylor, R. L., & Mulligan, G. A. (1968). *Flora of the Queen Charlotte Islands. Part 2. Cytological aspects of the vascular plants.* Queen's Printer.
native	Russia	40	Chepinoga, V. V., Aleksandr A, G., Enushchenko, I. V., & Rosbakh, S. A. (2009). IAPT/IOPB chromosome data 8. Taxon, 58(4), 1281–1314.
native	Russia	40	Probatova, N.S., Seledets, V.P., Rudyka, E.G., Gnutikov, A.A., Kozhevnikova, Z.V. & Barkalov V.Y. (2009). IAPT/IOPB Chromosome Data 8. Taxon, 58(4), 1284–1288.
native	Finland	60	Arohonka, T. (1982). Chromosome counts of vascular plants of the island Seili in Nauvo, SW Finland. Turun Yliopiston Biologian–Laitoksen Julkaisuja, 3, 1–12.
native	Turkey	60	Baltisberger, M. (1991). Cytological investigations of some plants from Turkey. Willdenowia, 21, 225–232.
native	Europe	60	Degraeve, N. (1975). Contribution a L'Etude Cytotaxonomique des *Rumex*—I. Le Genre *Rumex* L. Sensu Stricto. Caryologia, 28(2), 187–201.
native	UK	60	Dempsey, R. E. (1994). Contributions to a cytological catalogue of the British and Irish flora, 4. Watsonia, 20, 63–66.
native	Belarus	60	Dmitrieva, S. A. (1986). Chisla khromosom nekotorych vidov rastenij Berezinskogo Biosfernogo Zapovednika. Zapovedniki Belorussii Issledovaniia, 10, 24–28.

native	Spain	60	Garcia, C., Pastor Díaz, J. E., & Luque Palomo, M. T. (1989). Contribución al estudio cariológico del género *Rumex* (Polygonaceae). Acta Botanica Malacitana, 14, 129–140.
native	Japan	60	Himi, H. (1999). Chromosome numbers of 11 species in Japanese *Rumex* subg. *Rumex* (Polygonaceae). Journal of Phytogeography and Taxonomy, 47, 121–130.
native	Germany	60	Jaretzky, R. (1928). Histologische and karyologische Studien an Polygonaceen. Jahrbücher für Wissenschaftliche Botanik, 69, 357–490.
native	Czechia	60	Měsíček, J., & Javůrková–Jarolímová, V. (1992). *List of chromosome numbers of the Czech vascular plants.* Academia.
native	Germany	60	Kihara, H., & Ono, T. (1926). Chromosomenzahlen und systematische gruppierung der *Rumex*–arten. Zeitschrift für Zellforschung und mikroskopische Anatomie, 4(3), 475–481.
native	Japan	60	Kihara, H., & Ono, T. (1926). Chromosomenzahlen und systematische gruppierung der *Rumex*–arten. Zeitschrift für Zellforschung und mikroskopische Anatomie, 4(3), 475–481.
native	Scandinavia	60	Darlington, C. D. & Wylie, A. P. (1956). Chromosome atlas of flowering plants. Chromosome atlas of flowering plants (2nd ed). George Allen and Unwin Ltd.
native	Spain	60	Löve, A. & Kjellqvist, E. (1974). Cytotaxonomy of Spanish planta. III. Dicotyledons: Salicaceae–Rosaceae. Lagascalia, 4(1), 3–32.
native	Europe	60	Löve, A. & Löve, D. (1956). Cytotaxonomical conspectus of the Icelandic flora. Acta Horti Gothoburgensis, 20(4), 65–291.
native	Scandinavia	60	Löve, A. (1942). Cytogenetic studies in *Rumex*. III. Some notes on the Scandinavian species of the genus. Hereditas 28, 289–296.
native	Sweden	60	Löve, A. (1967). Reports [In Löve, A. (ed.), IOPB chromosome numbers reports XIII]. Taxon, 16(5), 445–461.
native	Sweden	60	Lövkvist, B. & Hultgård, U. M. (1999). Chromosome numbers in south Swedish vascular plants. Opera Botanica, 137, 1–42.
native	Slovenia	60	Majovsky, J. (1970). Index of chromosome numbers of Slovakian flora Part 2. Acta Facultatis Rerum Naturalium Universitatis Comenianae. Botanica, 18, 45–60.
native	Portugal	60	Queirós, M. (1985). Números cromosómicos para a flora Portuguesa. Boletim da Sociedade Broteriana Ser. 2, 58, 85–96.
native	South Africa	60	Nordenstam, B. (1982). In Chromosome numbers of southern African plants: 2. Journal of South African Botany, 48, 273–275.
native	Russia	60	Probatova, N. S. & Sokolovskaya, A. P. (1989). Chromosome numbers in vascular plants from Primorye Territory, the Amur region, Sakhalin, Kamchatka and the Kuril Islands. Botanicheskii Zhurnal (Moscow & Leningrad), 74, 120–123.
native	Portugal	60	Queirós, M. (1983). Números cromossómicos para a Flora Portugesa, 64–85. Boletim da Sociedade Broteriana, 56, 79–98.

native	Germany	60	Rohweder, H. (1937). Versuch zur Erfassung der mengenmässigen Bedeckung des Darss und Zingst mit polyploiden Pflanzen. Ein Beitrag zur Bedeutung der Polyploidie bei der Eroberung neuer Lebensräume. Planta, 27, 501–549.
native	Poland	60	Skalinska, M. (1976). Further studies in chromosome numbers of Polish Angiosperms. Part XI. Acta Biologica Cracoviensia s. Botanica, 19, 107–148.
native	UK	60	Squirrell, J. (1994). Contributions to a cytological catalogue of the British and Irish flora. Watsonia 20.
native	Bulgaria	60	Stoeva, M. P. (1987). Chromosome numbers of Bulgarian Angiosperms. Fitologija, 33, 65–66.
native	Greece	60	Strid, A., & Franzen, R. (1984). Chromosome numbers in flowering plants from Greece. Willdenowia, 13, 329–333.
native	Albania	60	Strid, A. (1971). Chromosome numbers in some Albanian angiosperms. Botaniska Notiser, 124, 490–496.
native	Korea	60	Sun, B. Y., Sul, M. R., Im, J. A., Kim, C. H., & Kim, T. J. (2002). Evolution of endemic vascular plants of Ulleungdo and Dokdo in Korea–floristic and cytotaxonomic characteristics of vascular flora of Dokdo. Korean Journal of Plant Taxonomy, 32(2), 143–158.
native	Japan	60	Takenaka, Y. (1941). The relation between polyploidy and the size of stoma. I. On the plants of the subgenus Lapathum. Botanical Magazine, Tokyo, 55, 319–323.
native	Germany	60	Tischler, G. (1934). Die Bedeutungen der Polyploidie fur die Verbreitung der Angiospermen, erlautert an den Arten Schleswig–Holsteins, mit Ausblicken auf andere Florengebiete. Botanische Jahrbücher fur Systematik, Pflanzengeschichte und Pflanzengeographie, 67, 1–36.
native	Slovakia	60	Uhríková, A. & Murín, A. (1970). Reports (In Májovsky, J. & al. (ed.), Index of chromosome numbers of Slovakian flora. Part 2). Acta Facultatis Rerum Naturalium Universitatis Comenianae. Botanica, 18, 45–60.
native	UK	60	Williams, J. T. (1971). Seed polymorphism and germination II. The role of hybridization in the germination polymorphism of *Rumex crispus* and *R. obtusifolius*. Weed Research, 11(1), 12–21.
native	Mediterranean	60	Runemark, H. (1996). Mediterranean chromosome number reports 6 (590–678). Flora Mediterranea, 6, 223–243.
native	Sweden	90	Lövkvist, B. (1999). Chromosome numbers in south Swedish vascular plants. Opera Botanica, 137, 1–42.

A.4. References

Leitch, I. J., Johnston, E., Pellicer, J., Hidalgo, O., & Bennett, M. D. (n.d.). Plant DNA C-values Database | Royal Botanic Gardens, Kew. Retrieved October 14, 2021, from https://cvalues.science.kew.org/

Rice, A., Glick, L., Abadi, S., Einhorn, M., Kopelman, N. M., Salman-Minkov, A., Mayzel, J., Chay, O., & Mayrose, I. (2015). The Chromosome Counts Database (CCDB) – a community resource of plant chromosome numbers. New Phytologist, 206(1), 19–26. https://doi.org/10.1111/nph.13191

Appendix B

Reference-based sequencing approach: Analyses and results (Chapter 5)

In chapter 5 I assessed genetic variation and differentiation in the native and introduced ranges of three *Rumex* species using *de novo* approach. In this appendix I detail the methods and results from the reference-based approach.

B.1. Methods

After sequencing, all raw reads were subjected to quality controls based on Dodds et al. (2015). This included checking samples for complete bar-codes and making sure that all the samples were represented relatively well. Reads with partially missing barcodes were discarded from the downstream analyses. The reads were then mapped to *R. hastatulus* genome (accession PRJNA638915; Rifkin et al., 2021) with bwa mem (0.7.17-r1188) with default settings. All three *Rumex* species were treated separately throughout the analyses. Around 17 % of the reads mapped to the *R. hastatulus* reference genome. TASSEL-GBS was used to call SNPs, with following parameters: UCreatWorkingDirPlugin; UFastqToTagCountPlugin -c 1 -e PstI; UMergeTaxaTagCountPlugin -t n -c 3; UTagCountToTagPairPlugin-e 0.03; UTagPairToTBTPlugin; UTBTToMapInfoPlugin; UMapInfoToHapMapPlugin -mnMAF 0.001 -mxMAF 0.5 -mnC 0.001 -mxC 1. After SNP calling, roughly 4000–10 000 SNPs were kept per species and information was saved in a Variant Call Format (VCF) file by species.

Filtering using KGD in R and all the analyses were done using the same methods from *de novo* approach (Chapter 5).

B.2. Results and discussion

B.2.1. Sequencing details and the number of single-nucleotide-polymorphisms

Sequencing yielded 2746 M raw reads, 462 M of these were successfully mapped to *R. hastatulus* genome. After quality controls and filtering in KGD, roughly 2500–3500 SNPs were found to be of sufficient quality for downstream analyses (TABLE B.1). This led to 1022 samples remaining, they had less than 20 % missing genotypes on average and an average depth of over 300.

TABLE B.1: The number of samples collected from each country belonging to each of the three species, as well as the number of SNPs, a mean sample depth and a proportion of missing genotypes after filtering in KGD. The UK: UK; New Zealand: NZ; France: FR; Finland: FI; the Czech Republic: CZ.

	UK	**NZ**	**FR**	**FI**	**CZ**	**Total**	**#SNPs**	**Depth**	**Missing genotypes**
R. conglomeratus	136	168	15	N/A	9	328	2512	382	0.18
R. obtusifolius	152	126	10	20	21	329	2507	439	0.13
R. crispus	159	153	5	19	29	365	3538	318	0.19

B.2.2. Partitioning of the variation within and between native and introduced ranges

To compare partitioning of genetic variation within and between different hierarchy levels, AMOVAs were run for each species. The full dataset was used to assess genetic structure within and between the regions.

The AMOVAs showed that almost 100 % of the variation was found within the individuals, a few percentages of variation within countries, and no variation between countries (TABLE B.2). The permutation test from showed that within individual variation and within country variation were statistically significant in all species, whereas between country variation was not significant in any species. This indicates little genetic structure between the countries, but some structure within them.

TABLE B.2: AMOVA results showing partitioning of variations within different levels of hierarchy (Sigma) as well as percentage of the total variation, using the full dataset for each Rumex species. *: < 0.05, **: < 0.01, ***: < 0.001.

	R. conglomeratus		***R. obtusifolius***		***R. crispus***	
Comparison	*Sigma*	*%*	*Sigma*	*%*	*Sigma*	*%*
Between country	1.549	0.382	-2.616	-0.658	-0.725	-0.129
Between samples within country	29.108	7.186***	20.913	5.259***	26.729	4.755***
Within individuals	374.412	92.432***	379.348	95.398***	536.078	95.374***
Total variations	405.068	100	397.646	100	568.081	100

B.2.3. Is there genetic admixture within the introduced range?

As I found lack of genetic structure between the countries, I proceeded to look into admixture that could explain the phenomenon. To estimate genetic admixture within New Zealand (NZ), samples from the UK and all three outgroups (France, Finland and the Czech Republic) were used in analyses to study similarities between the ranges.

DAPCs were done with the full dataset to estimate overlap between the countries. Results show that NZ and UK have varying degrees of overlap with each other and with the other European outgroups (FIGURE B.1). For *R. crispus* the first two discriminant-axes explain most of the variation, whereas for the other two species the first axis holds

most of the information, and results could be shown along that axis only, but in order to make comparisons easier, data was plotted along the first two axes.

The posterior probability analyses (FIGURE B.2, TABLE B.3, TABLE B.4) showed varying levels of similarities – or admixture – with non-origin countries. Around half of *R. obtusifolius* and *R. crispus* samples from NZ show less than 0.9 assignment probability with NZ, and for *R. conglomeratus* this is considerably less, but still every fifth sample can be considered admixed. Based on the average probabilities, *R. obtusifolius* and *R. crispus* share a high assignment with UK, whereas *R. conglomeratus* shows equal assignments to UK and FR.

The genetic differentiation between all countries was analysed using pairwise F_{ST}. All the pairwise comparisons have a value of less than 0.05 (TABLE B.5), indicating very low differentiation between the countries. In *R. conglomeratus* and *R. crispus* NZ shows less differentiation when compared to FR and in *R. conglomeratus* also when compared to CZ than differentiation between NZ and UK, even when the differentiation between UK and FR or CZ is less or equal to NZ and UK.

TABLE B.3: Proportion of *Rumex* samples within each species with under 0.9 posterior assignment probability, i.e., admixed samples with mixed assignment to at least one country other than the country of origin. Number of plants in parenthesis. The UK: UK; New Zealand: NZ; France: FR; Finland: FI; the Czech Republic: CZ.

Sample origin	***R. conglomeratus***	***R. obtusifolius***	***R. crispus***
NZ	22 % (38)	48 % (60)	58 % (88)
UK	13 % (18)	48 % (74)	53 % (84)
FR	73 % (11)	60 % (6)	40 % (2)
FI	-	20 % (4)	47 % (9)
CZ	22 % (2)	29 % (6)	10 % (3)

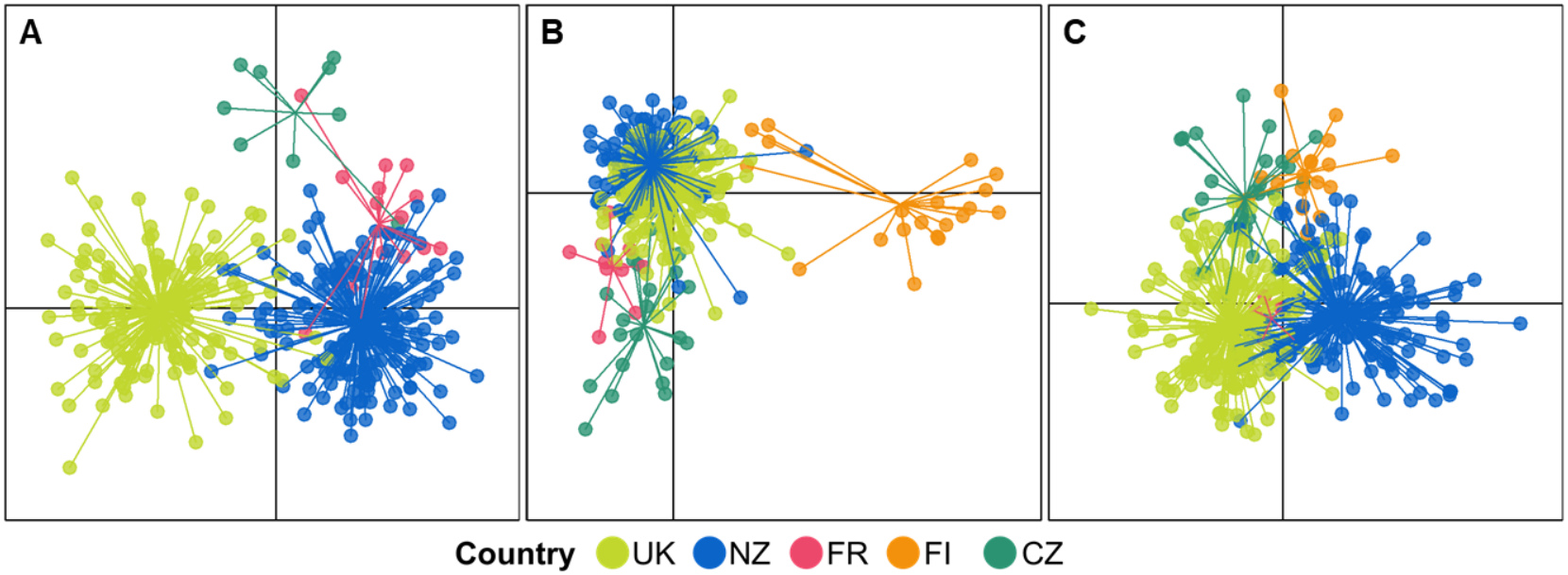

FIGURE B.1: Discriminant Analysis of Principal Components (DAPC) highlighting hierarchy based on genetic information and *a priori* information about the populations for a) *Rumex conglomeratus*; b) *R. obtusifolius*; c) *R. crispus*. PCA accumulation curve shows the amount of variation explained by the retained PCA axes and DA eigenvalues similarly show how much each DA axes contribute to the variation. Country abbreviations: FI: Finland; FR: France; NZ: New Zealand; UK: the United Kingdom; CZ: the Czech Republic.

TABLE B.4: The average posterior assignment probabilities of plants to the country of origin and a proportion of association each plant has with countries other than the country of origin, i.e., the proportion admixture. *Rumex conglomeratus*: Rco; *R. obtusifolius*: Rob; *R. crispus*: Rco. The UK: UK; New Zealand: NZ; France: FR; Finland: FI; the Czech Republic: CZ.

Origin	**NZ**			**UK**			**FR**			**FI**		**CZ**		
Species **Association**	*Rco*	*Rob*	*Rcr*	*Rco*	*Rob*	*Rcr*	*Rco*	*Rob*	*Rcr*	*Rob*	*Rcr*	*Rco*	*Rob*	*Rcr*
NZ	0.91	0.73	0.73	0.07	0.16	0.15	0.38	0.11	0.14	0.04	0.06	0.09	0.07	0.08
UK	0.04	0.24	0.20	0.93	0.81	0.78	<0.01	0.05	0.19	0.16	0.05	<0.01	0.10	0.20
FR	0.05	<0.01	<0.01	<0.01	0.02	<0.01	0.62	0.74	0.67	0.00	<0.00	0.05	0.03	<0.01
FI	N/A	<0.01	0.03	N/A	<0.01	<0.01	N/A	0.00	<0.01	0.80	0.84	N/A	0.00	0.09
CZ	<0.01	0.01	0.02	<0.01	0.02	0.05	<0.01	0.09	<0.01	<0.01	0.05	0.86	0.80	0.62

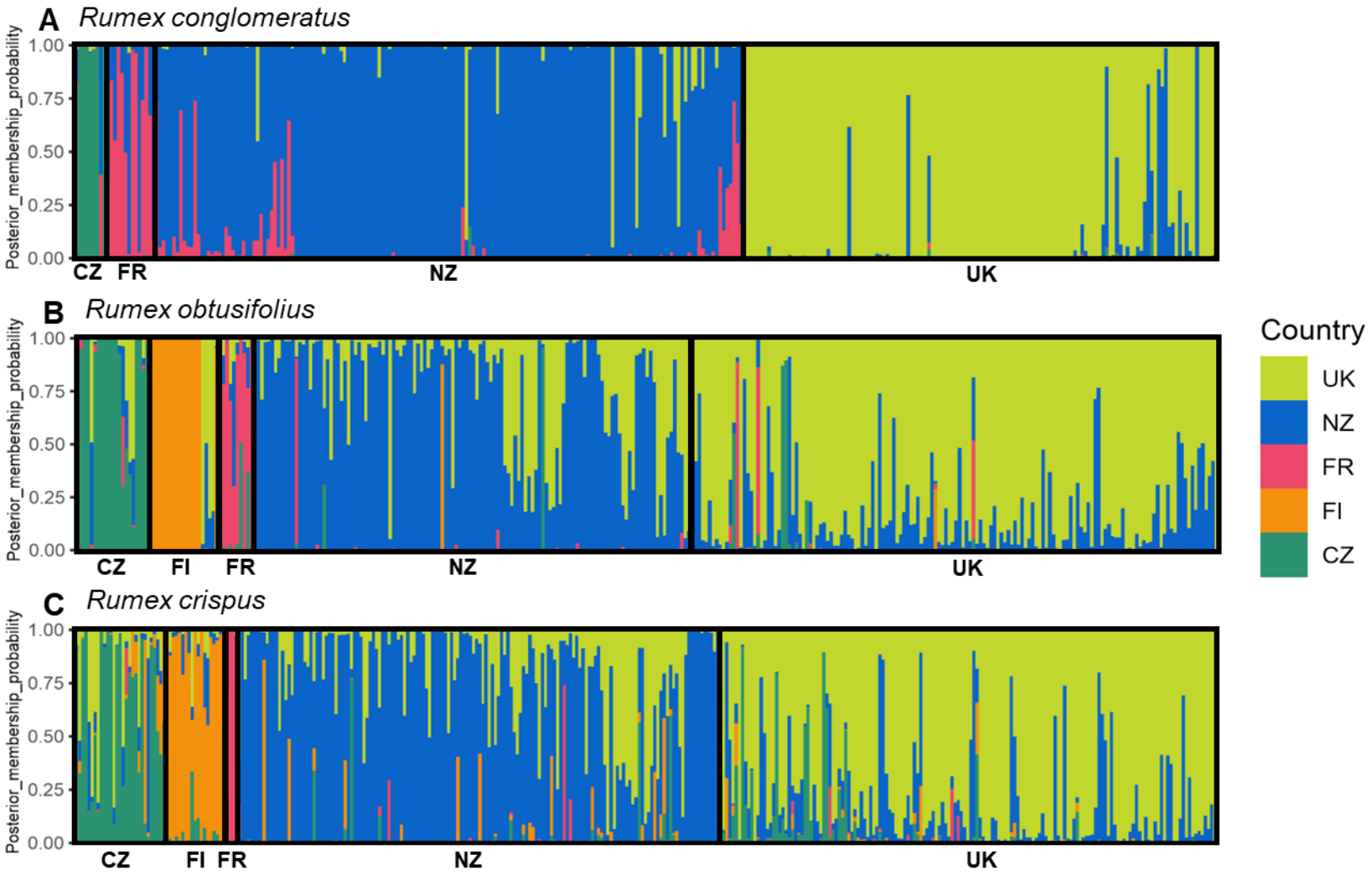

FIGURE B.2: Posterior probability plots from Discriminant Analysis of Principal Components (DAPC) showing each individual as a vertical line and its probability to be assigned to a population, i.e., the country where the sample was collected. Samples are clustered based on the country of origin: 1: the UK; 2: New Zealand; 3: France; 4: Finland; 5: the Czech Republic. Number of PCA axes retained after a-validation: 29, 30, and 24 for *Rumex conglomeratus*, *R. obtusifolius*, and *R. crispus*, respectively.

TABLE B.5: Pairwise F_{ST} values showing differentiation between the countries using the full dataset. a) *Rumex conglomeratus*; b) *R. obtusifolius*; c) *R. crispus*. Country abbreviations: FI: Finland; FR: France; NZ: New Zealand; UK: the United Kingdom; CZ: the Czech Republic.

		FI	**FR**	**NZ**	**UK**
a. R. conglomeratus					
	CZ	-	0.016	0.005	0.003
	FR	-		0.004	0.004
	NZ	-			0.008
b. R. obtusifolius					
	CZ	0.016	0.009	0.005	0.004
	FI		0.021	0.013	0.009
	FR			0.003	0.002
	NZ				0.003
c. R. crispus					
	CZ	0.009	0.007	0.005	0.004
	FI		0.011	0.006	0.006
	FR			0.002	0.002
	NZ				0.004

B.2.4. Closer comparison between the UK and New Zealand

Lastly, I wanted to get a better understanding of the differences between the native and introduced ranges and took a closer look using region datasets. This time individual regions were compared to see finer structures, patterns, and differences within and between the ranges in three *Rumex* species.

To rule out one of the outgroups driving the patterns seen in AMOVAs when all countries were compared, I investigated the differences and similarities between UK and NZ. When UK and NZ were analysed without the outgroups, the variation between countries increased in *R. conglomeratus* to explain around 5 % of the variation and less in the other two species (TABLE B.6). Most of the variation was still within samples in all three species. All but the between country level variation in *R. obtusifolius* were significant according to the randomisation test.

Similar to the other analyses, the pairwise F_{ST} values showed little genetic differentiation between the countries or regions (FIGURE B.3). All F_{ST} values were well below 0.05, which is considered low differentiation (Hartl & Clark, 1997; Frankham et al., 2002). However, *R. obtusifolius* and *R. crispus* showed similar patterns: NZ regions differed less from each other than did UK regions, whereas for *R. conglomeratus* the differentiation within countries was similar.

TABLE B.6: AMOVA results showing partitioning of variations within different levels of hierarchy using the region datasets. *: < 0.05, **: < 0.01, ***: < 0.001.

	R. conglomeratus		***R. obtusifolius***		***R. crispus***	
Comparison	*Sigma*	*%*	*Sigma*	*%*	*Sigma*	*%*
Between countries	23.244	5.463***	0.350	0.088	7.847	1.292*
Between individuals within countries	19.012	4.469***	12.040	3.039***	18.883	3.108***
Within individuals	383.174	90.068***	383.775	96.873***	580.733	95.600***
Total variations	425.429	100	396.165	100	607.463	100

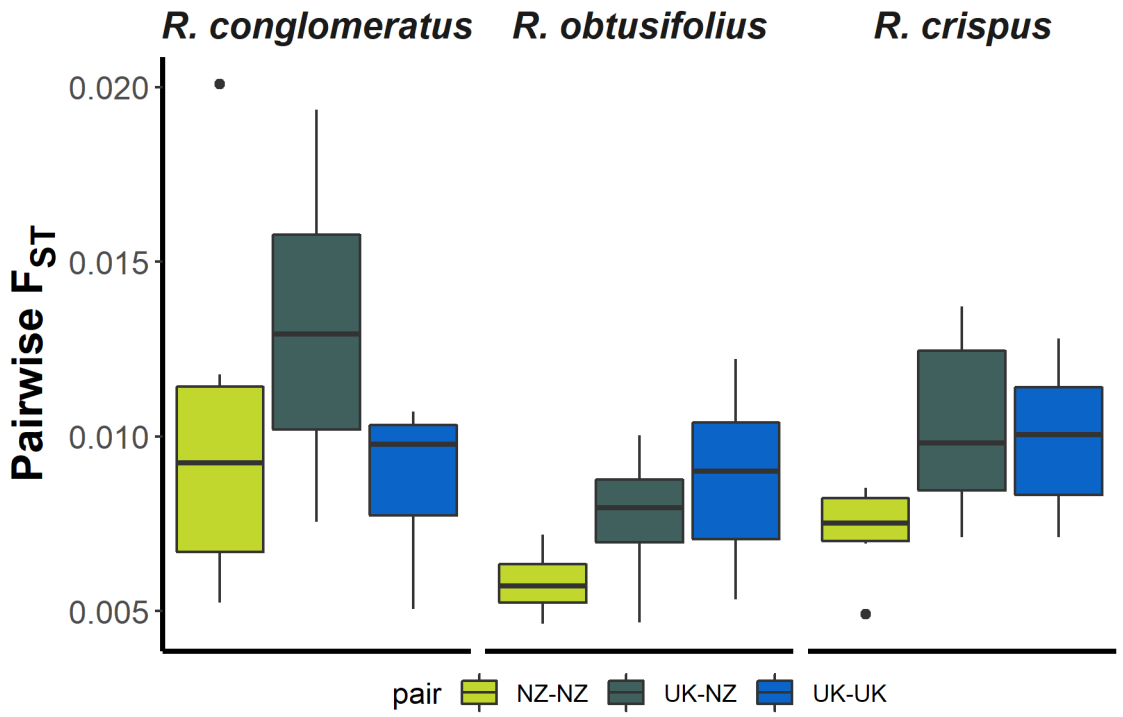

FIGURE B.3: Pairwise F_{ST} analysis shows little differentiation within and between the regions in the UK and New Zealand (NZ) in all three species. For each country, four regions were compared to each other to get the pairwise estimates. Similarly, for the country comparison the four regions within one country were compared to the four regions within the other country. The box represents the interquartile range with the median represented with a thick line and the whiskers extending from the 5^{th} to the 95^{th} percentile, with values outside this range (outliers) represented as dots.

B.2.5. Comparison of reference-based and *de novo* results

The big picture stays the same using either the reference-based or the *de novo* approach, but there are slight variations in the actual values from the analyses. The number of SNPs was almost double in the reference-based approach compared to the *de novo* approach, the depth was considerably higher, but the amount of missing data was relatively similar.

Both approaches show that most of the variation in AMOVAs, can be found within individuals, this is true for both country and region datasets. Similarly, the results from DAPC clusters NZ and UK relatively tightly together with FR in the mix, whereas CZ and FI and slightly further away. The differentiation according to F_{ST} is again similar, but differentiation is now even less than with *de novo* approach. In addition, NZ shows less or equal differentiation to FR than to UK, while FR and UK show little differentiation from each other.

While individual values are different and some species show a slightly different patterns, the data still supports multiple introductions to NZ. It also shows a high degree of similarity between NZ and UK, supporting the UK as one of the origin countries for these species in NZ. Similarly, I found support for introductions from outside the UK as well.

B.3. References

Dodds, K. G., McEwan, J. C., Brauning, R., Anderson, R. M., van Stijn, T. C., Kristjánsson, T., & Clarke, S. M. (2015). Construction of relatedness matrices using genotyping-by-sequencing data. BMC Genomics, 16(1), Article 1047. https://doi.org/10.1186/s12864-015-2252-3

Frankham, R., Ballou, S. E. J. D., Briscoe, D. A., & Ballou, J. D. (2002). Introduction to conservation genetics. Cambridge University Press.

Hartl, D. L., & Clark, A. G. (1997). Principles of Population Genetics (3rd ed.). Sinauer Associates.

Rifkin, J. L., Beaudry, F. E. G., Humphries, Z., Choudhury, B. I., Barrett, S. C. H., & Wright, S. I. (2021). Widespread recombination suppression facilitates plant sex chromosome evolution. Molecular Biology and Evolution, 38(3), 1018–1030. https://doi.org/10.1093/molbev/msaa271