INTRODUCTION

Human activity has destroyed natural ecosystems over vast areas, compromising their function, and leaving many plant species at risk of extinction (IPBES, 2019; IUCN, 2020). Ex situ collections of plants are becoming increasingly important in our efforts to prevent extinction (Mounce et al., 2017), and in providing material for ecological restoration at landscape scales (Nevill et al., 2016). For threatened or...
endangered plants, ex situ collections reduce the probability of sudden extinction due to the loss of remaining in situ populations (Abeli et al., 2020). For species that are used commonly in re-vegetation, ex situ collections cultivated as seed orchards can provide germplasm efficiently and in large volumes, without depleting local material and impacting natural populations (Broadhurst et al., 2008, 2015). In both cases, we should strive to assemble ex situ collections that are genetically diverse and representative of genetic variation in the target species (Hoban, 2019). The goals are to promote the capacity of populations derived from the material to adapt to changing conditions (Burger & Lynch, 1995; Stange et al., 2021), and particularly for endangered species, to reduce the probability of inbreeding depression by avoiding the use of close relatives in collections (Charlesworth & Willis, 2009).

Often it might also be advantageous to select for desirable properties in populations that are assembled ex situ. This might include, for example, promoting disease resistance in an endangered species, or choosing material for revegetation that exhibits adaptations to climatic conditions that are predicted to become more prevalent (Breed et al., 2019; Prober et al., 2015). There are growing prospects for identifying alleles linked to desirable traits, and enriching for these in different management contexts (Borrell et al., 2019; Browne et al., 2019; MacLachlan et al., 2021). However, there is potential for conflict between the goals of imposing selection and promoting diversity (Chivers et al., 2016; Gaitán-Espitia & Hobday, 2021; Kardos & Shafer, 2018). This is because selection for a particular trait not only reduces diversity in that trait, but also associated genetic variation. This is recognized as a substantial challenge in conventional breeding, and strategies have been formulated to deal with it, including for conservation breeding programmes (de Cara et al., 2013; Fernandez & Toro, 1999; Mascher et al., 2019; Meeuissen, 1997). We showed previously how this conflict might manifest in the design of an ex situ collection of an endangered species, where an attempt to enrich for an idealized trait came at the cost of increased genetic similarity among individuals in the collection (Bragg et al., 2020).

We can determine the amount of potential conflict between choosing collections that are diverse (adaptable) and selected for a particular trait (adapted). In conservation, such tradeoffs are often quantified and managed while formulating plans of action. The tradeoffs can relate to the magnitude of conservation activities and their costs (McShane et al., 2011). Examples include the land areas set aside for conservation (Ball et al., 2009). Others involve the size of populations that are protected to preserve genetic diversity, or the effort that is taken to collect seed widely for their establishment or for representative germplasm collections (e.g., Brachi et al., 2011; Griffith & Husby, 2010; Hoban & Schlarbaum, 2014; Marshall & Brown, 1975; Reeves & Richards, 2017; Richards et al., 2007; Schlotthfeldt et al., 2015; Shoen & Brown, 1993). For some applications it is possible and useful to characterize the tradeoff between objectives, commonly known as a Pareto front (described in Box 1), using an approach such as multiobjective optimization. This can potentially identify actions that perform well for both objectives if the tradeoff is relaxed (Box 1), and if the tradeoff is stringent (Box 1) it can facilitate the adoption of actions in a way that is informed by the tradeoff between them, in combination with an understanding of their relative importance. Here the tradeoff represents the cost of imposing selection for a favourable trait, in terms of neutral genetic diversity and associated adaptive capacity.

Broadly, there are many factors that potentially influence whether the tradeoff between selection and genetic diversity is relaxed or stringent for a particular species and trait. These include the ways in which genetic variation and trait variation are distributed across the landscape (Loiselle et al., 1995), and the genetic architecture (complexity) of the trait. If we could characterize these factors, and use them to make predictions about the strength of tradeoffs, it might provide opportunities to design conservation actions more efficiently. For example, if we predicted that selecting for a particular trait potentially had small costs in terms of neutral diversity, and was otherwise likely to substantially affect the fitness of individuals in the future (e.g., an allele conferring resistance to a virulent pathogen), an investment in the costs of assaying the trait in many individuals, or performing a study to find associated loci, might be highly worthwhile. On the other hand, if selecting for a trait were likely to bring great costs in terms of genetic diversity, and the link between the trait and fitness were relatively weak, we might instead focus on maximizing neutral genetic diversity and save the extra costs associated with performing research to inform selection based on traits or associated markers.

Here, we examined approaches for designing ex situ collections using landscape genetic data sets for two exemplar species. The collections represent groups of individuals that could be used in seed production areas for restoration. However, we note that the approaches and observations are also applicable to the design of ex situ germplasm collections for widely distributed species. In each case, we contemplate the design of collections that are optimized for neutral diversity and the design of collections that are adapted to warm temperatures. We characterize tradeoffs between these objectives, and observe differences between the two species that can be related to contrasting patterns in the way that neutral diversity is associated with the environmental driver.

2 | MATERIALS AND METHODS

2.1 | Overview

We are undertaking landscape genomic studies of multiple plant species across the state of New South Wales, Australia, to provide information about genetic provenance for ecological restoration based on analyses of genetic and environmental data (Rossetto et al., 2019). Data for two species, Westringia fruticosa and Wilkiea huegeliana, are examined here. These species occupy different habitats (coastal fringe and rainforest, respectively),
BOX 1  Tradeoffs in plant collection design

Imagine we have a set of plants and want to create a collection from a subset of them. We want the collection to be genetically diverse (containing individuals that are dissimilar at neutral markers), and we also want it to be enriched with individuals having some desirable property (e.g., a good trait, or good alleles). If one set of plants cannot best satisfy both goals, then a set of different collections will form a tradeoff curve between the two goals, called the Pareto front. If the tradeoff curve is convex, there are collections that are very good for both goals, and we refer to it as a “relaxed” tradeoff (a). If the tradeoff curve is straight (or even concave) it is harder to find collections that are very good for both objectives, and we refer to it as a “stringent” tradeoff (a). We use idealized examples (b and c) to illustrate how the distribution of desirable properties in space might affect the stringency of tradeoffs between the diversity of candidate collections, and the prevalence of the desirable property. There are six plants, and we want to choose a collection consisting of two plants. We assume that the genetic dissimilarity among the plants is related to their spatial distance from each other, and contemplate scenarios where a desirable phenotype (grey shading) is distributed narrowly (b) or widely (c) in space, leading to stringent and relaxed tradeoffs (respectively). In panel (b), the collection that has both plants with the desirable property also has plants that are close together, and not genetically dissimilar. The most diverse collection possible, C3, does not feature the desirable property. Another collection, C2, is intermediate between C1 and C3 in terms of diversity and the representation of the desirable property. Here choosing a collection requires making a substantial compromise on one objective, or both. (c) The collection featuring both individuals with the desirable property, C1, has slightly higher diversity (relative to collection C1 in b). Also, the relaxed tradeoff in (c) means there is a collection (C2) where intermediate trait values are obtained at a small cost to neutral diversity, which might represent a useful compromise.
We sampled at varying elevations in rainforest habitats of eastern Australia. \textit{Wilkiea huegeliana} (Tul.) A.DC. (Monimiaceae) is a shrub or small tree that occurs across their ranges in New South Wales (Figure 1). It naturally occurs only along the coast in eastern Australia. \textit{Wilkiea huegeliana} has small leaves and produces white flowers all year. It naturally occurs only along the coast in eastern Australia.

\textit{Westringia fruticosa} (Willd.) Druce (Lamiaceae) is a woody shrub that has small leaves and produces white flowers all year. It naturally occurs only along the coast in eastern Australia. \textit{Westringia fruticosa} (Willd.) Druce (Lamiaceae) is a woody shrub that has small leaves and produces white flowers all year.

and therefore potentially exhibit contrasting patterns of association between genetic variation and climate (Figure 1). Analyses of landscape genetic data for these species suggested that genetic variation across space was reasonably continuous, rather than strongly structured (for \textit{We. fruticosa}, see Rossetto et al., 2019; Figure 2b,e), simplifying the search for loci that were putatively associated with environmental variables. Below, we provide a brief description of the study species, and the methods that were used to generate these exemplar landscape genomic data sets. We then provide a detailed description of the approaches we used to design collections of these species that were both genetically diverse and adapted to warm temperatures, and the ways in which we investigated the mechanistic basis for these tradeoffs.

2.2 Study species and sampling

\textit{Westringia fruticosa} (Willd.) Druce (Lamiaceae) is a woody shrub that has small leaves and produces white flowers all year. It naturally occurs only along the coast in eastern Australia. \textit{Wilkiea huegeliana} (Tul.) A.DC. (Monimiaceae) is a shrub or small tree that occurs at varying elevations in rainforest habitats of eastern Australia. We sampled \textit{We. fruticosa} and \textit{Wi. huegeliana} at 27 and 28 sites (respectively) across their ranges in New South Wales (Figure 1). At each sampling site, leaf tissue samples were collected from six individual plants.

2.3 Molecular methods and preliminary analyses

Leaf tissue samples were genotyped using DArTseq, a reduced representation sequencing method that is provided on a commercial basis (Diversity Arrays Technology Pty Ltd). DArTseq involves a restriction digest of sample DNA, followed by high throughput sequencing (Sansaloni et al., 2010), and is therefore similar to RADseq or genotyping by sequencing approaches (Andrews et al., 2016). DArTseq produced genotype data for thousands of single nucleotide polymorphism (SNP) loci for each sample. We filtered these data for each species, using several criteria. This included a reproducibility score for each locus provided by DArT, which is an index of the concordance of genotype scores between technical replicate samples. We used a threshold of 0.96 for this reproducibility score. We also removed loci that were missing genotypes in >1% of samples, or that had a minor allele frequency of <3%. These filters were quite stringent, and different to those used in our previous analyses (Rossetto et al., 2019), because we wanted to reduce the chances of inferring spurious associations between allelic variation and climate due to missing data.

For each species, we performed a set of preliminary analyses that explored associations between genetic, spatial, and environmental distance. To do this, we estimated pairwise $F_{st}$ between sampling sites (method of Weir & Hill, 2002, implemented in R package SNPRelate, Zheng et al., 2012), and tested how this measure of genetic differentiation varied as a function of spatial and environmental (mean annual temperature) distance using Mantel tests (99,999 permutations, R package vegan, Oksanen et al., 2016). We also used a clustering method to characterise population structure, and to estimate the membership of individuals in different ancestral populations. For this we used sparse non-negative matrix factorization (sNMF, implemented in R package LEA; Frichot & François, 2015), because this approach directly informed a method of testing genotype–environment associations (latent factor mixed models [LFMM], Frichot & François, 2015, see below).

2.4 Genotype environment association analyses

We used the SNP data for each species to examine approaches for incorporating adaptive alleles into seed orchard design. To do this, we identified loci that were putatively associated with mean annual temperature (eMAST Bio01, resolution 0.01 degrees, Xu & Hutchinson, 2013). We did this using two statistical approaches that test for these associations while controlling for neutral genetic variation. The first approach tested associations with LFMM (implemented in R package LEA; Frichot & François, 2015), which controls for neutral genetic variation using inferred ancestral population proportions. We began by estimating the best supported number of ancestral populations for each data set ($K = 4$ for \textit{We. fruticosa} and $K = 3$ for \textit{Wi. huegeliana}, Figure S1) using sNMF (implemented in R package LEA; Frichot & François, 2015). We then ran LFMM to test associations between genotypes and mean annual temperature (10
replicate runs). We identified loci that were associated with mean annual temperature based on a false discovery rate criterion (see Results). The second approach was based on environmental and genetic covariance (Coop et al., 2010, implemented in BayPass, Gautier, 2015). It tests environmental associations for each locus, while controlling for neutral genetic covariance among sites. We ran BayPass (standard mode, nval = 1000, numpilot = 20, pilotlength = 1000) and identified loci associated with mean annual temperature according to a threshold value of the Bayes factor (see Results).

Robust identification of loci that mediate selection on complex traits is difficult (Bragg et al., 2015), and we note caveats associated with these analyses. Our data sets were modest in size, and we used reasonably permissive criteria to identify loci whose association with mean annual temperature was sufficiently great to suggest it resulted from selection. We stress that our primary goal was to find a panel of loci that are associated with mean annual temperature, so we can contemplate how these might be included in seed orchard design, and not to validate their adaptive function. At each of these loci, we identified the allele that was associated with warm temperatures, and will refer to these as temperature correlated alleles (TempCA) to emphasize that they exhibit patterns of variation consistent with temperature adaptation, but we do not necessarily assert that they are all functionally adaptive loci. To complement these analyses, and to remedy their limitations, we also simulated loci having prescribed associations with warmer mean annual temperatures. We simulated these loci under the BayPass (Gautier, 2015) model that was fitted to each data set. Here, it is possible to simulate allele frequencies and genotypes that are influenced by the combined effects of neutral genetic variation (represented by the covariance matrix among sites, \( \Omega \)), and direct local adaptation to a nominated environmental variable (with prescribed effect sizes on different loci) (Gautier, 2015; see below for details).

2.5 | Optimization

We next designed collections (for example, seed orchards) for *Westringia fruticosa* and *Wilkiea huegeliana* that were optimized...
according to different criteria (genetic diversity and warm-temperature adaptation). Each collection was set to a size of 30 plants, assuming constraints on available space. First, we aimed to design a population of each species that was as genetically representative as possible. We therefore used the simulated annealing algorithm (Kirkpatrick et al., 1983) to choose a subset of the available genotyped plants (size $N_p = 30$) that maximized a measure of genetic diversity (see Simulated Annealing Analyses below, and Appendix S1 for more details). For this we used expected heterozygosity at SNP markers that were segregating among the sampled individuals (Nei, 1973). Previously we observed that optimizing this measure for plant collections also resulted in the capture of a relatively high proportion of the available alleles (Bragg et al., 2021).

Second, we wanted to explore the process of designing “climate-ready” collections by using individuals adapted to warm temperatures. We did this in two ways. First, we used the mean annual temperature of the site of origin as a surrogate for the temperature tolerances of the individual plants, and designed a phenotypically selected collection. This was the subset of plants from our sample set that maximized the average mean annual temperature of the site of origin. Second, we designed a genomically selected collection that was enriched for a subset of alleles that were found to be common in environments with high mean annual temperature. In this case, we used simulated annealing to choose a subset of individuals that maximized the average frequency of the high mean annual temperature correlated alleles. Finally, we investigated tradeoffs between the objectives of maximizing genetic diversity and deliberately selecting either plants from warm sites, or plants with alleles that were correlated with high mean annual temperatures, and we used multi-objective simulated annealing to characterize these tradeoffs.

### 2.6 Investigating the mechanistic basis of tradeoffs

We wanted to understand whether there were factors that influenced the magnitude of tradeoffs between the goal of increasing genetic diversity and the goal of imposing selection for warm temperatures. We hypothesized that tradeoffs would be greatest when neutral genetic variation across space was strongly associated with environmental distances across space. The reason is, if organisms are distributed along an environmental gradient, and there is also isolation by distance along the gradient, it will often be difficult to choose a set of individuals that are highly genetically diverse and highly adapted to conditions at a particular point on the gradient in environmental conditions (Box 1b,c). This is because choosing individuals to promote genetic diversity would usually require the selection of individuals that are dispersed widely in space, while individuals that are adapted to a particular condition will usually be clumped at a location where that condition prevails. If this were correct, we predicted that the tradeoff would be smaller if we randomly shuffled the temperature data – breaking up the association between neutral genetic variation and environment – and repeated the optimization analyses.

For our examination of alleles that were associated with mean annual temperature, we performed simulations to confirm the observations we made for the empirical data set, and then to test the notion that a strong tradeoff might be attributable to covariation between neutral genetic variation and environment. We performed simulations using a model of genetic covariance among populations (Coop et al., 2010) that is implemented in BayPass (Gautier, 2015). To do this, we simulated loci under a model parameterized with the covariance matrix that was inferred using BayPass (Q), and effect sizes governing the association between each locus and mean annual temperature. These data sets were the same size as our empirical data sets, in terms of the number of samples and loci. The strengths of association were drawn from two uniform distributions, the first bounded by values $-0.25$ to $-0.15$, and the second bounded by values $0.15$–$0.25$ (“beta.coef” in the “simulate.baypass” function). For each data set, we simulated many more associated loci than were required, and chose SNPs with allele frequencies matched to the empirical data. We repeated the multiobjective optimization procedure, replacing the observed TempCA with the simulated ones, to confirm that patterns observed in our empirical data sets were also present in data where the associations with the environmental variable were prescribed. Next, we hypothesized that covariance between the environmental variable and neutral population genetic variation might contribute substantially to the observed tradeoff. To test this hypothesis, we simulated more adapted SNPs in a way that was analogous, except the simulated loci were associated with a permuted set of mean annual temperature value. We then performed the multiobjective optimization using loci associated with permuted mean annual temperature.

### 2.7 Simulated annealing algorithms

We optimized collections using implementations of simulated annealing (Kirkpatrick et al., 1983) that have been described in detail previously (Bragg et al., 2020; Bragg et al., 2022; see Appendix S1 for details of parameterization). Briefly, the algorithm works by choosing a subset of the available individuals at random. This becomes the “current” set of individuals. It then completes a large (e.g., 10,000) number of “steps” in a “chain”. At each step, a change is made to the current set by dropping one individual, and replacing it with another, both chosen at random. This becomes the “proposed” set of individuals. If the proposed set of individuals is better than the current set for a nominated objective (e.g., higher genetic diversity), the proposed set is accepted, and becomes the current set. If the proposed set is poorer than the current set, it can still be accepted, with some probability. This probability gets smaller as the number of steps that has already been taken increases, meaning that the algorithm can explore sets quite freely early in the chain, but later, only accepts proposed sets that are better in terms of the objective, so that it moves towards an approximate maximum.
3 | RESULTS

3.1 | Landscape genomic variation

For Westringia fruticosa, the filtered genetic data set consisted of 155 samples from 27 sites, and 2508 loci. For Wilkiae huegeliana, the filtered genetic data set consisted of 152 samples from 28 sites and 6153 SNP loci.

Both species exhibited genetic variation across the landscape consistent with isolation by distance (Figure 2). In each case, genetic differentiation ($F_{ST}$) between pairs of populations increased as a function of their spatial distance apart (Mantel tests, $r = .87, p < 1 \times 10^{-5}$; $r = .95, p < 1 \times 10^{-5}$). However, the two species were strongly contrasted in terms of the association between neutral genetic variation and the climatic variable mean annual temperature. For Westringia fruticosa, there was a strong association between the spatial distance between pairs of sites and the difference (Euclidean distance) in their mean annual temperatures (Mantel test, $r = .79, p < 1 \times 10^{-5}$). Correspondingly, there was a strong tendency for sites that had large differences in mean annual temperature to have large values of $F_{ST}$ (Mantel test, $r = .72, p < 1 \times 10^{-5}$). In contrast, for Wilkiae huegeliana there were substantially weaker associations between the spatial distance between sites and the temperature difference between them (Mantel test, $r = .22, p = .007$), and between the pairwise $F_{ST}$ values between sites and the temperature differences between them (Mantel test, $r = .22, p = .011$).

3.2 | The identification of high temperature correlated alleles

In Westringia fruticosa, we ran LFMM with a K value of 4, and identified 40 SNP loci that were associated with mean annual temperature (false discovery rate = 0.1). With BayPass, we identified 37 SNP loci that were associated with mean annual temperature (Bayes factor >5). Five SNPs were associated with mean annual temperature using both inferential approaches, so that a total of 72 SNPs were found to be associated with mean annual temperature. For optimization analyses, we used 40 loci, after discarding SNPs that had missing data for any sample, and three SNPs with conflicting directions of association between LFMM and BayPass, leaving 40 loci.

In Wilkiae huegeliana, we ran LFMM with a K value of 3, and identified 60 SNP loci that were associated with mean annual temperature (false discovery rate = 0.15). With BayPass, we identified 21 SNP loci that were associated with mean annual temperature (Bayes factor >4). Six SNPs were found to be associated with mean annual temperature using both inferential approaches, such that, collectively, 75 SNPs were found to be associated with mean annual temperature. For optimization analyses, we used 27 loci, after discarding SNPs that had missing data for any sample, and one SNP that had conflicting directions of association between LFMM and BayPass.

3.3 | Optimal population designs

We designed collections for each species that consisted of 30 founder individuals and that approximately maximized genetic diversity. For each species, this set of individuals with large values of genetic diversity was distributed broadly across space (Figure 3a,b). Conversely, when we chose individuals that maximized the mean annual temperature of the site of origin (Figure 3a,b), or that maximized the representation of alleles that were positively associated with mean annual temperature (Figure 4a,b), the chosen individuals tended to be aggregated in warmer parts of the (respective) geographical ranges. In sum, for each species, we had designed a population that was highly diverse, but was not highly adapted to warm temperatures. Conversely, when we designed populations that were adapted to warm temperatures, they were not maximally genetically diverse.

We next characterized the shape of the implied tradeoffs between these objectives, for each species, using multiobjective optimization. For Westringia fruticosa, there was quite a stringent tradeoff between the objectives of maximizing genetic diversity and maximizing the mean annual temperature of the sites of origin (Figure 3c, black circles). That is, the curve formed by the solutions (Pareto front) was only mildly convex. When we performed the same optimization using the randomly shuffled mean annual temperatures, the tradeoff was substantially relaxed (Figure 3c, grey circles). For Wilkiae huegeliana, the tradeoff between the objectives of maximizing genetic diversity and maximizing mean annual temperature was quite relaxed, with a substantially convex Pareto front (Figure 3d, black circles). When the mean annual temperature data were randomly shuffled, and the tradeoff and the optimization was repeated with the shuffled data, the shape of the tradeoff curve was relatively little affected (Figure 3d, grey circles).

When we performed analyses where we increased the representation of alleles that were associated with mean annual temperature (instead of the temperature at the site of origin), we made broadly similar observations. That is, in Westringia fruticosa, there was quite a stringent tradeoff between choosing collections that maximized genetic diversity and that maximized the representation of alleles that were associated with warmer temperatures (Figure 4c, black circles). This was the case both when we used warm adapted loci identified in the empirical data set (by LFMM or BayPass), and when we used loci that were simulated to have prescribed associations with temperature (Figure 4c, black squares). However, when we performed the same optimization for loci that were simulated to have prescribed associations with the randomly shuffled temperature data, there was a relaxation in the tradeoff (the Pareto front was more convex) (Figure 4c, grey circles). That is, the tradeoff was relatively stringent for alleles that were associated with an environmental variable (mean annual temperature) that was linked with neutral genetic variation (Figure 2c), and more relaxed for alleles that were associated with an environmental variable (shuffled mean annual temperature) that did not have the same link to neutral genetic variation. In Wilkiae huegeliana, when we characterized tradeoffs between
maximizing genetic diversity and maximizing the representation of warm adapted alleles, we observed little difference in their shapes (convexity) when we used empirical data (Figure 4d, black circles), or data that were simulated with prescribed associations with temperature (Figure 4d, black squares), or data that were simulated with prescribed associations with randomly shuffled temperature (Figure 4d, grey circles).

**FIGURE 3** Optimized collection designs for *Westringia fruticosa* (a and c) and *Wilkiea huegeliana* (b and d). Collections of each species were designed that were optimized for neutral genetic diversity and maximum mean annual temperature. (a and b) The latitude and mean annual temperature of sampled individuals (black circles) are shown. For each species, plants that were selected in the populations optimized for mean annual temperature and genetic diversity are indicated with semi-transparent red and blue circles, respectively (where points are overplotted, the colour of points becomes more intense). (c and d) The values for single objective optimizations of mean annual temperature (large red point) and neutral genetic diversity (large blue point) are illustrated, along with the tradeoff line (Pareto front, black circles) that was inferred with multiobjective optimization. For comparison, we randomly shuffled the mean annual temperature, and show the outcome of a multiobjective optimization for neutral genetic diversity and the shuffled temperature variable (grey circles).

4 | DISCUSSION

4.1 | Plant collections: can they be adapted and adaptable?

When we design a collection of plants for conservation or restoration, we usually want it to be genetically diverse and to represent genetic variation across a nominated geographic area. Sometimes, we would also like the collection to be enriched with a desirable trait. However, we know imposing selection for a trait might reduce neutral diversity, resulting in a tradeoff between these goals. Here, our aim was to explore factors that influence variation in the intensity, or stringency, of such tradeoffs by examining possible collections of two exemplar species, *Westringia fruticosa* and *Wilkiea huegeliana* (Box 1).

For both species, when collections were designed with the single goal of maximizing genetic diversity, individuals were chosen from sites across the range, consistent with proposed strategies to maximize adaptive potential (Broadhurst et al., 2008; Bucharova et al., 2019). In *Westringia fruticosa*, there was a stringent tradeoff between the goal of obtaining a genetically diverse collection, and the goal of obtaining a collection that was potentially adapted to warm temperatures. *Westringia fruticosa* is a coastal species, occurring at low elevations in coastal habitats. As such, sites that were far apart tended to be warmer in the north and cooler in the south, and had larger differences in mean annual temperature. In combination with a pattern of isolation by distance, this led to an association between neutral genetic variation and variation in mean annual temperature. This resulted in a stringent tradeoff between the genetic diversity of possible collections and the mean value of their temperatures of origin. This is supported by the observation that the tradeoff relaxed (the Pareto frontier was less convex) when we randomly permuted the mean annual temperature data, disrupting the association...
between mean annual temperature and neutral genetic variation. Conversely, for Wilkiea huegeliana, there was a substantially weaker association between the distance between sites and the difference in their temperatures. This is probably because it occurs in rainforest habitats inland, where a range of local factors influence mean annual temperature, including topography and the distance to the coast. This meant that, despite a pattern of isolation by distance, there was little association between neutral genetic variation and mean annual temperature. Correspondingly, the tradeoff between designing a genetically diverse collection and a collection adapted to warm temperatures was observed to be relaxed, and did not become substantially more relaxed when the temperature of sampling sites was randomly permuted.

Taken together, these observations show how tradeoffs can make it difficult to design a collection that is both highly adaptable (genetically diverse) and adapted (to a specific driver), and that the strength of these tradeoffs can vary substantially among species or traits. Aspects of these relationships have long been understood and anticipated (Reeves & Richards, 2018; Shoen & Brown, 1993), and here we extend this understanding by: (i) examining species that exhibit contrasting levels of covariation between neutral genetic variation and an environmental driver, and (ii) manipulating this covariation using simulations and permutation. Collectively, the results support the notion that a landscape level association between neutral genetic variation and the distribution of a desirable trait is likely to promote the stringency of the tradeoff between these variables in collection design.

4.2 | Applications: conservation and restoration on the Pareto frontier

The speed of climate change threatens to overwhelm the capacity for evolutionary adaptation in many species. This has increased our motivation to select for climate readiness in conservation and restoration collections. Advances in genomic methods are providing new opportunities to do this with "marker assisted" and genomic selection approaches, which could improve the fitness of conserved or restored populations. However, there has been useful debate concerning the value of targeting particular gene variants in conservation (Kardos & Shafer, 2018), and discussion of possible unintended evolutionary consequences (Gaitán-Espitia & Hobday, 2021). Our study contributes here by illustrating a way to quantify some of the possible disadvantages of

---

**FIGURE 4** Optimized collection designs for *Westringia fruticosa* (a and c) and *Wilkiea huegeliana* (b and d), with collections optimized for neutral genetic diversity and enrichment with high temperature correlated alleles. (a and b) The latitude and mean annual temperature of all sampled individuals (black circles) are shown. For each species, plants that were selected in the populations optimized for high temperature correlated alleles and genetic diversity are indicated with semi-transparent green and blue circles, respectively (where points are overplotted, the colour of points becomes more intense). (c and d) The values for single objective optimizations of high temperature correlated alleles (large green point) and neutral genetic diversity (large blue point) are illustrated along with the tradeoff line (Pareto front, black circles) that was inferred with multiobjective optimization. For comparison, we also performed multiobjective optimizations to enrich for sets of loci that were simulated to have prescribed associations with mean annual temperature (black squares), and prescribed associations with randomly shuffled values of mean annual temperature (grey circles).
imposing selection for desirable traits or markers. It identifies a set of candidate collections that form a tradeoff curve (Pareto frontier, Box 1) between the two conflicting goals, providing a range of options for managers to consider. Here the curve characterizes the loss in neutral genetic diversity, relative to the maximum that could be obtained, as a consequence of increasing the representation of a desirable trait or alleles. This is useful especially when considered in relation to some of the arguments that have been posed against the imposition of selection for specific genes (Kardos & Shafer, 2018). For example, there is a risk that targeted alleles will turn out not to be adaptive, leading to no tangible benefit. This risk might be tolerable if there is also a small cost in terms of genetic diversity, and an intermediate solution might be chosen (e.g., Box 1c, collection C2). Otherwise, it might be better to maintain as much genome wide diversity and representativeness as possible, to promote the likelihood that populations can adapt to future challenges involving different environmental changes and traits.

Our study also sheds light on factors that influence tradeoffs between adapted and adaptable collections. This might prove useful in promoting efficiency in our conservation programs, by prioritizing intense experimentation for circumstances where it promises the greatest applied impacts. For example, there can be strong links between neutral diversity and desirable traits or markers where there is a genetic pattern of isolation by distance, and the trait varies along a gradient in space. Moreover, realized tradeoffs might be harder to avoid when the trait is influenced by many genes of small effect. If a strong tradeoff seems highly likely, one might adopt a strategy of maximizing neutral diversity, and avoiding the costs of an expensive (whole genome scale) study aimed at identifying associated alleles. Alternatively, this framework provides insights into circumstances where it might be highly worthwhile to identify desirable alleles, and impose selection for them. A conspicuous example might include alleles conferring resistance to emerging pathogens, which potentially have large fitness consequences, often have substantial effect sizes for genotype-phenotype associations, and potentially are distributed patchily across the landscape in ways that might be little associated with neutral genetic diversity. In this case it might be possible to select for the desired trait or alleles while also maintaining high levels of neutral diversity and representativeness.

4.3 | Paths forward

This study has presented approaches to characterizing tradeoffs between designing adapted and adaptable collections, illustrated variation among species in the stringency of these tradeoffs, and provided insights into the mechanistic basis of this variation via simulations. However, there is more to be done to understand the generality of patterns described here, and how best to translate these observations into effective and efficient conservation action. For example, there are a number of additional factors that could influence the stringency of tradeoffs, potentially in complex ways. These include the size of the optimized collection and the size of the sample of individuals from which it is drawn. Also, we have previously shown (Bragg et al., 2020) that the frequency of desirable alleles in the initial population affects the cost of enriching for them in a collection. The present study carefully controlled for these factors (in the simulations), but we suggest that in future it might be useful to vary them systematically, and to examine their effect on the quality of collections that can be designed, and the stringency of tradeoffs. This might help inform details of the application of these approaches (e.g., significance or effect size thresholds for loci associated with a desirable trait). We also expect that factors that were not considered here could be important, and might be best investigated with larger data sets. These include the genetic architecture of adaptation, and the tendency for loci that are associated with traits under selection to be physically linked to neutral loci (i.e., the extent of linkage disequilibrium). Ultimately, we hope some general principles will emerge, perhaps in the form of factors that can lead, directly or indirectly, to covariance between markers under selection and neutral genetic variation, in aggregate.

In terms of application, there might also be better approaches for promoting useful compromises between goals. For example, we aimed to maximize the frequency of desirable alleles. In some circumstances, depending on how the collection is used, and the nature of the trait and its consequences for fitness, this could be important. However, there might also be cases where it is useful to make sure desirable alleles are present in the collection, perhaps at a frequency exceeding some threshold, to reduce the probability of rapid loss by drift. This would provide an opportunity for natural selection to increase the frequency of the putatively adaptive alleles, should they prove advantageous. It would be possible to formulate an alternative objective function to reflect this goal (e.g., the proportion of desirable alleles reaching a threshold frequency), rather than maximizing the prevalence of desirable alleles. This is somewhat similar to methods used in assisted migration (Borrell et al., 2019).

These approaches could also be extended in conjunction with the development of collection strategies to achieve specific project goals. For example, in cases where a seed orchard was used to generate plants for translocation populations across different environments, it might be useful to maintain multiple seed orchards, optimized for different objectives, and to draw material from those in different proportions according to management goals (including the implementation of composite or climate-adjusted strategies, Prober et al., 2015). Finally, careful interrogation of optimized sets of individuals from an initial collection could inform additional sampling efforts. For example, if individuals from a particular region contribute inordinately to multiple objectives, and are overrepresented in collections that are adaptable and adapted, this region could be sampled with greater intensity in supplemental collections.

5 | CONCLUSIONS

We studied landscape genetic patterns of variation in two species, with a view to designing ex situ collections that were genetically diverse and adaptable to changing conditions, as well as adapted to warm temperatures. In one species, Westringia fruticosa, the tradeoff
between designing diverse collections and adapted collections was highly stringent, while in another species, Wilkia huegeliana, the tradeoff was substantially more relaxed. We linked the relative stringency of these tradeoffs to a stronger association between neutral genetic variation and mean annual temperature in We. fruticosa, which made it impossible to identify one collection that performed well for both the nominated objectives. We performed simulations that bolstered these interpretations. In sum, there is variation among species in the strength of tradeoffs encountered when attempting to design collections that are both genetically diverse (adaptable) and selected for a specific trait (adapted). We show that the strength of these tradeoffs varies according to the way genetic and trait (environmental) variation is distributed across the landscape. This might ultimately lead to an improved capacity to understand and predict these tradeoffs, and to plan management activities in light of these predictions.

ACKNOWLEDGEMENTS
We thank Joel Cohen and a team of volunteers for field sampling. Open access publishing facilitated by New South Wales Department of Planning Industry and Environment, as part of the Wiley - New South Wales Department of Planning Industry and Environment agreement via the Council of Australian University Librarians.

CONFLICT OF INTEREST
None declared.

DATA AVAILABILITY STATEMENT
Genetic data have been deposited in a Dryad Repository associated with this publication, https://doi.org/10.5061/dryad.th76hf11. Code used to process genetic data, and perform simulations and optimizations for each species are archived in the same Dryad Repository, https://doi.org/10.5061/dryad.th76hf11 (Bragg et al. (2022)).

ORCID
Jason G. Bragg https://orcid.org/0000-0002-7621-7295

REFERENCES
Abeli, T., Dalrymple, S., Godefroid, S., Mondoni, A., Müller, J. V., Rossi, G., & Orsenigo, S. (2020). Ex situ collections and their potential for the restoration of extinct plants. Conservation Biology, 34, 303–313. https://doi.org/10.1111/cobi.13391
Anders, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. Nature Reviews Genetics, 17, 81–92. https://doi.org/10.1038/nrg.2015.28
Ball, I. R., Possingham, H. P., & Watts, M. (2009). Marxan and relatives: Software for spatial conservation prioritisation. In A. Molanzen, K. A. Wilson, & H. P. Possingham (Eds.), Spatial conservation prioritisation: Quantitative methods and computational tools (pp. 185–195). Oxford University Press.
Borell, J. S., Zohren, J., Nichols, R. A., & Buggs, R. J. A. (2019). Genomic assessment of local adaptation in dwarf birch to inform assisted gene flow. Evolutionary Applications, 13, 161–175. https://doi.org/10.1111/eva.12883
Brachi, B., Morris, G. P., & Borevitz, J. O. (2011). Genome-wide association studies in plants: The missing heritability is in the field. Genome Biology, 12, 232. https://doi.org/10.1186/gb-2011-12-10-232
Bragg, J. G., Cuneo, P., Sherief, A., & Rossetto, M. (2020). Optimizing the genetic composition of a translocation population: Incorporating constraints and conflicting objectives. Molecular Ecology Resources, 20, 54–65. https://doi.org/10.1111/1755-0998.13074
Bragg, J. G., Supple, M. A., Andrew, R. L., & Borevitz, J. O. (2015). Genomic variation across landscapes: Insights and applications. New Phytologist, 207, 953–967. https://doi.org/10.1111/nph.13410
Bragg, J. G., van der Merwe, M., Yap, J. Y. S., Borevitz, J., & Rossetto, M. (2022). Data from – “Plant collections for conservation and restoration: can they be adapted and adaptable?” Dryad, https://doi.org/10.5061/dryad.th76hf11
Bragg, J. G., Yap, J. Y. S., Wilson, T., Lee, E., & Rossetto, M. (2021). Conserving the genetic diversity of condemned populations: Optimizing collections and translocation. Evolutionary Applications, 14, 1225–1238. https://doi.org/10.1111/eva.13192
Breed, M. F., Harrison, P. A., Blyth, C., Byrne, M., Gaget, V., Gellie, N. J. C., Groom, S. V. C., Hodgson, R., Mills, J. G., Prowse, T. A. A., Steane, D. A., & Mohr, J. J. (2019). The potential of genomics for restoring ecosystems and biodiversity. Nature Reviews Genetics, 20, 615–628. https://doi.org/10.1038/s41576-019-0152-0
Broadhurst, L., Hopley, T., Li, L., & Begley, J. (2015). Using seed production areas to meet restoration targets and secure genetic diversity. Australasian Plant Conservation: Journal of the Australian Network for Plant Conservation, 23, 7–8.
Broadhurst, L. M., Lowe, A., Coates, D. J., Cunningham, S. A., McDonald, M., Vesk, P. A., & Yates, C. (2008). Seed supply for broadscale restoration: Maximizing evolutionary potential. Evolutionary Applications, 1, 587–597. https://doi.org/10.1111/j.1755-4571.2008.00045.x
Brown, L., Wright, J. W., Fitz-Gibbon, S., Gugger, P. F., & Sork, V. L. (2019). Adaptational lag to temperature in valley oak (Quercus lobata) can be mitigated by genome-informed assisted gene flow. Proceedings of the National Academy of Sciences of the United States of America, 116, 25179–25185.
Bucharova, A., Bosssdorf, O., Hölzel, H., Kollmann, J., Prasse, R., & Durka, W. (2019). Mix and match: Regional admixture provenance strikes a balance among different seed sourcing strategies for ecological restoration. Conservation Genetics, 20, 7–17. https://doi.org/10.1007/s10592-018-1067-6
Burger, R., & Lynch, M. (1995). Evolution and extinction in a changing environment: A quantitative-genetic analysis. Evolution, 49, 151–163.
Charlesworth, D., & Willis, J. H. (2009). The genetics of inbreeding depression. Nature Reviews Genetics, 10, 783–796. https://doi.org/10.1038/nrg2664
Chivers, I. H., Jones, T. A., Broadhurst, L. M., Mott, I. W., & Larson, S. R. (2016). The merits of artificial selection for the development of restoration-ready plant materials of native perennial grasses. Restoration Ecology, 24, 174–183. https://doi.org/10.1111/rec.12323
Coop, G., Witonsky, D., Di Rienzo, A., & Pritchard, J. K. (2010). Using environmental correlations to identify loci underlying local adaptation. Genetics, 185, 1411–1423. https://doi.org/10.1534/genet.ics.110.114819
de Cara, M. A. R., Villanueva, B., Toro, M. A., & Fernandez, J. (2013). Purging deleterious mutations in conservation programmes: Combining optimal contributions with inbred matings. Heredity, 110, 530–537. https://doi.org/10.1038/hdy.2012.119
Fernandez, J., & Toro, M. A. (1999). The use of mathematical programming to control inbreeding in selection schemes. Journal of Animal Breeding and Genetics, 116, 447–466. https://doi.org/10.1046/j.1439-0388.1999.00196.x
Fritch, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. Methods in Ecology and Evolution, 6, 925–929. https://doi.org/10.1111/2041-210X.12382
Gaitán-Espitia, J. D., & Hobday, A. J. (2021). Evolutionary principles and genetic considerations for guiding conservation interventions
under climate change. Global Change Biology, 27, 475–488. https://doi.org/10.1111/gcb.15359
Gautier, M. (2015). Genome-wide scan for adaptive differentiation and association analysis with population-specific covariates. Genetics, 201, 1555–1579.
Griffith, P., & Husby, C. (2010). The price of conservation: Measuring the mission and its cost. BG Journal, 7, 12–14.
Hoban, S. (2019). New guidance for ex situ gene conservation: Sampling realistic population systems and accounting for collection attrition. Biological Conservation, 235, 199–208. https://doi.org/10.1016/j.biocon.2019.04.013
Hoban, S., & Schlärbaum, S. (2014). Optimal sampling of seeds from plant populations for ex-situ conservation of genetic biodiversity, considering realistic population structure. Biological Conservation, 177, 90–99.
IPBES (2019). Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. In S. Díaz, J. Settele, E. S. Brondizio, H. T. Ngo, M. Guèze, J. Agard, A. Arneth, P. Balvanera, K. A. Brauman, S. H. M. Butchart, K. M. A. Chan, L. A. Garibaldi, K. Ichii, J. Liu, S. M. Subramanian, G. F. Midgley, P. Miloslavich, Z. Molnár, D. Obura, … C. N. Zayas (eds.). IPBES secretariat.
IUCN (2020). The IUCN Red List of Threatened Species. Version 2020-1. Retrieved from https://www.iucnredlist.org
Kardos, M., & Shafer, A. B. A. (2018). The peril of gene-targeted conservation. Trends in Ecology and Evolution, 33, 827–839. https://doi.org/10.1016/j.tree.2018.08.011
Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. Science, 220, 671–680. https://doi.org/10.1126/science.220.4598.671
Loiselle, B. A., Sork, V. L., Nason, J., & Graham, C. (1995). Spatial genetic structure of a tropical understory shrub, Psychotria Officinalis (Rubiaceae). American Journal of Botany, 82, 1420–1425.
MacLachlan, I. R., McDonald, T. K., Lind, B. M., Rieseberg, L. H., Yeaman, S., & Aitken, S. N. (2021). Genome-wide shifts in climate-related variation underpin responses to selective breeding in a widespread conifer. Proceedings of the National Academy of Sciences, 118, e2016900118. https://doi.org/10.1073/pnas.2016900118
Marshall, D. R., & Brown, A. H. D. (1975). Optimum sampling strategies in crop relatives is aided by assessment of genetic markers. Proceedings of the National Academy of Sciences of the United States of America, 90, 10623–10627. https://doi.org/10.1073/pnas.90.22.10623
Stange, M., Barrett, R. D. H., & Hendry, A. P. (2021). The importance of genomic variation for biodiversity, ecosystems and people. Nature Review Genetics, 22, 89–105. https://doi.org/10.1038/s41576-020-00288-7
Weir, B. S., & Hill, W. G. (2002). Estimating F-statistics. Annual Review of Genetics, 36, 721–750. https://doi.org/10.1146/annurev.genet.36.050802.093940
Xu, T., & Hutchinson, M. F. (2013). New developments and applications in the ANUCLIM spatial climatic and bioclimatic modelling package. Environmental Modelling Software, 40(267–2). 79. https://doi.org/10.1016/j.envsoft.2012.10.003
Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics, 28, 3326–3328. https://doi.org/10.1093/bioinformatics/bts606

SUPPORTING INFORMATION
Additional supporting information may be found in the online version of the article at the publisher’s website.

How to cite this article: Bragg, J. G., van der Merwe, M., Yap, J.-Y. S., Borevitz, J., & Rossetto, M. (2022). Plant collections for conservation and restoration: Can they be adapted and adaptable? Molecular Ecology Resources, 22, 2171–2182. https://doi.org/10.1111/1755-0998.13605