A framework for maximising the capture of genetic diversity in sampling for ex situ conservation

Linda E Neaves¹ & UK Conservation Genetics Exchange Project

¹Royal Botanic Garden Edinburgh, 20A Inverleith Row, EH35LR, Edinburgh, UK

Abstract

The rapid and extensive loss of biodiversity globally has resulted in an increased urgency to capture and conserve the diversity which is present, including genetic diversity within species. However, for many species there is currently no detailed genetic data available to inform the collection and use of material held in *ex situ* collections and this can hamper the consideration of genetic issues and reduce the likelihood collection represent the diversity present. Even in the absence of direct genetic data, however, it is possible to consider genetic issues using the existing theoretical and empirical evidence-based and biological, ecological and demographic data for a given species. Here a framework to facilitate the consideration of genetic diversity and genetic issues, even where genetic data is lacking, is presented.

Introduction

Three quarters of the terrestrial environment has been significantly altered by humans, many of them negatively (https://knowledge.unccd.int/publications/ipbes-2019-global-assessment-report-biodiversity-and-ecosystem-services). As a result of these pressure many species are under threat and/or declining. Global targets for plant restoration and management of at least 15% (Global Strategy for Plant Conservation 2011-2020; https://www.cbd.int/gspc) for each ecological region or vegetation type and at least 75% of threatened plant species are contained within ex situ collections by 2020, have aimed to protect and conserve the remaining diversity.

The establishment of *ex situ* collections to conserve plant diversity and provide material for future ecological restoration are increasingly part of conservation programmes. Intra-specific genetic diversity has an intrinsic value as a measurable component of biological diversity and the amounts and types of genetic diversity can influence the success of restoration, as it underpins the fitness of individuals and their ability to adapt to new or changing environmental conditions.

There are two key situations where 'genetic thinking' can enhance ecological restoration programmes;

- 1. Genetic diversity may be lost during seed collection from natural populations, reducing the suitability of these collections for future reintroductions
- 2. Selection of donor material to ensure populations have adequate diversity to persist, are suitably adapted to the environmental conditions they are moved to; and mixing of incompatible genetic lineages is avoided.

However, detailed genetic data for many species is lacking, and this can hamper the consideration of key genetic issues in the collection and use of material. Despite the decreasing costs of generating genetic data, significant expertise and investment, particularly in the analytical aspects, is still required and for many species it is unlikely that genetic data will be available in the immediate future. Yet, there is an abundant evidence-base and expertise that can be used to guide collection and use of material in the absence of genetic data (and incorporate where possible). Below a

generalised framework is present to facilitate the evaluation of the risks associated with sampling and the likelihood of genetic problems arising during future restoration. The aim is to simplify and standardise the assessment process for use by non-specialists and provide guidance on situations where detailed data on genetic diversity are lacking. The format is to outline the general working principles and the logic flow for evaluating risks, a description of the information used in the evaluation and a detailed explanation of risk categories and recommendations

A framework for assessing and managing genetic issues during sampling

Assessing the risks

Overview:

There are two steps in capturing genetic diversity and reintroducing it into the landscape. (1) Establishing *ex situ* collections that are representative of the overall diversity within species and suitable for use in future restoration; and (2) the appropriate selection of donor material for reintroduction into a given location (that minimises the negative impacts on indigenous populations).

Both steps require an understanding of the organisation of genetic diversity within and among populations. Figure 1 shows a schematic of the process. Ideally information from detailed genetic studies would be available but in the absence of such data there are some general working principles that can be applied.

Step 1: Capturing genetic diversity

- Populations connect by gene flow contain fewer differences between them [high agreement/robust evidence]
- Local adaptations are associated with divergent ecological/environmental conditions. [high agreement/moderate evidence]
- Populations that have been separated for a prolonged period of time (i.e. glacial refugia)
 may contain unique genetic lineages [moderate agreement/robust evidence]
- Populations in the core of the range contain higher levels of genetic diversity (and fewer differences among them) [high agreement/moderate evidence]
- The levels of genetic diversity contained within populations decrease with decreasing population size and decreasing connectivity [high agreement/robust evidence]

Step 2: Selection of donors

Predicting beneficial outcomes (i.e. avoiding inbreeding depression):

- Population containing 50-100 breeding individuals or fewer benefit from mixing [high agreement/moderate evidence]
- Small populations with fitness problems benefit from the introduction of diversity from elsewhere [high agreement/moderate evidence].
- Mixing populations can increase amounts of genetic diversity enhancing adaptive potential [high agreement, limited evidence]

Predicting negative outcomes (i.e. avoiding outbreeding depression):

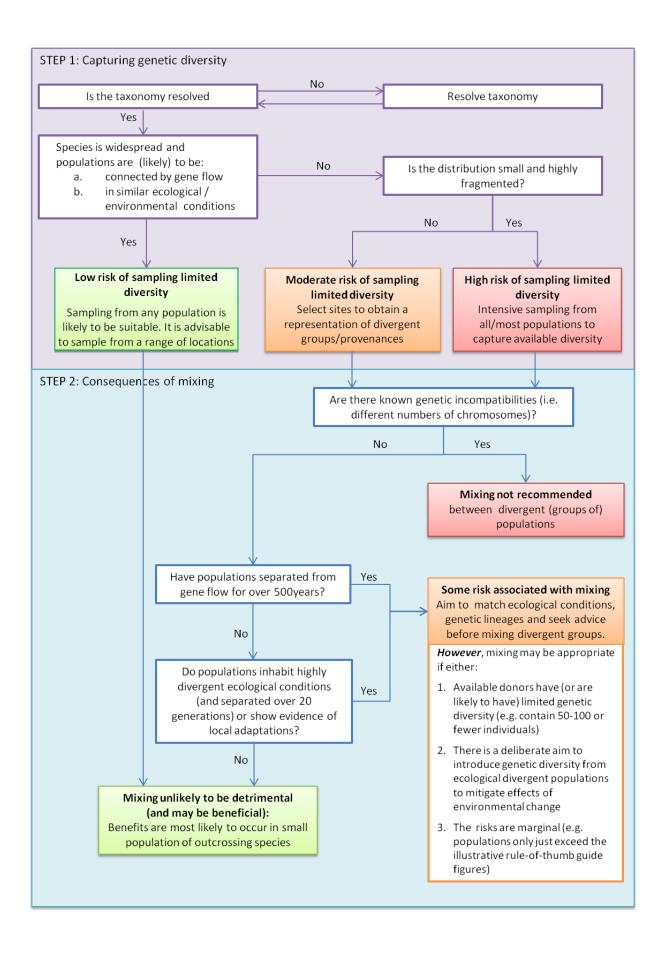
 Populations from ecologically divergent conditions may introduce maladapted genetic variants reducing viability in hybrid offspring (outbreeding depression) [moderate agreement/moderate evidence]

- Populations that have been separated for >500 years (i.e. 100s of generations) may have developed genetic incompatibilities that will lead to outbreeding depression [moderate agreement /limited evidence]
- There is limited risk of outbreeding depression from mixing ecologically similar populations that have been separated for less than 500 years [moderate agreement /limited evidence]
- There is limited risk of outbreeding depression from mixing ecologically different populations that have been separated for less than 20 generations [moderate agreement /limited evidence]

Species traits that affect the organisation of genetic diversity and susceptibility to genetic problems

	Higher	Lower				
Attribute/Trait	diversity/lower differentiation	diversity/higher differentiation	Agreement/Evidence			
ORGANISATION OF GENETIC DIVERSITY						
Dispersal ability	Effective	Poor	high agreement /robust evidence			
Mode of reproduction	Sexual	Asexual / selfing	high agreement /moderate evidence			
Longevity	Long lived	Short lived	moderate agreement /limited evidence			
Attribute/Trait	LESS vulnerable	MORE vulnerable	Agreement/Evidence			
OUTBREEDING DEPRESSION						
Mode of reproduction	Asexual / selfing	Sexual	high agreement /limited evidence			
Mode of reproduction Longevity	Asexual / selfing Long lived	Sexual Short lived				

Figure 1: Decision making process covering (1) sampling strategies for capturing genetic diversity and (2) the selection of donors – based on the consequences of mixing with the indigenous populations (or in restored population).



Species information

This section outlines the information that is evaluated to obtain an assessment. Complete information will not be available in all cases and approximations based on the available evidence may be necessary.

Life history traits: Species differ in how their genetic diversity is distributed and their susceptibility to different genetic problems. Factors influencing this include dispersal ability, mode of reproduction, longevity, time to first reproduction, reproductive output and ecological amplitude. Many of these factors are non-independent and it is difficult to accurately predict the scale of their impacts in any given situation. Nevertheless, despite the inevitable imprecision of making generic statements at this level, some broad generalisations can be made and this provides context to the assessment (see above).

Information on life history traits can be obtained from:

- Ecological Flora of the British Isles: http://www.ecoflora.co.uk
- Online Atlas of the British and Irish Flora: http://www.brc.ac.uk/plantatlas/.
- The LEDA traitbase: http://www.leda-traitbase.org/LEDAfactsheet/index.jsp
- Genetic Flora of the British Isles: http://elmer.rbge.org.uk/geneticflora/gflora.php
- BSBI Cytology database: http://rbg-web2.rbge.org.uk/BSBI/cytsearch.php

Dispersal distance: The movement of organisms and their genes between populations reduces the differences between them. Hence effectively dispersed species are predicted to be less likely to possess unique variation or incompatibilities that may result in genetic problems if mixed (e.g. outbreeding depression or loss of local adaptations) compared with poorly dispersed species. There is, however, a greater chance of unintentional mixing with indigenous populations in effectively dispersed species.

Mode of reproduction: Species that reproduce predominantly asexually or via self-fertilisation rarely mate with other individuals. This means populations are more likely to contain differences but limited outcrossing also greatly reducing the likelihood of outbreeding depression occurring should previously isolated lineages be brought together.

Longevity (and generational overlap): Long lived species (e.g. >10yrs) are more likely to survive inappropriate mixing of lineages in translocation programmes as their longevity provides a mechanism for 'riding out' periods of unfavourable conditions.

Reproductive output: Species with low reproductive output may be less likely to persist if outbreeding depression occurs.

Number of sets of chromosomes (ploidy): Species which possess populations with different levels of ploidy are more susceptible to outbreeding depression if these populations are mixed as chromosomal difference typically cause genetic incompatibilities.

Ecological/genetic data: The organisation of genetic diversity is influenced by levels of gene flow, adaptations to local conditions, past and contemporary demographic and distributional changes and hybridisation with closely related species. Ideally genetic data would be evaluated but inferences about the organisation of genetic diversity can also be made from other sources (see below).

Species' distributions can be obtained from the National Biodiversity Network (https://data.nbn.org.uk) and information on ecological and genetical differences among

populations can be obtained from the literature (e.g. Genetic Flora of the British Isles: http://elmer.rbge.org.uk/geneticflora/gflora.php; Google Scholar; http://scholar.google.com/)

Extent and fragmentation of distribution: Populations that are isolated in the landscape are (a) less likely to exchange genes and hence are more likely to contain unique genetic variation, (b) more likely to have genetic differences that can lead to outbreeding depression, and (c) more likely to have low levels of genetic diversity that may compromise their use in restoration.

Where populations are separated (and unlikely to be exchanging genes) the length of separation should also be considered. This is often difficult to determine but the key is to identify situations where populations are likely to have been separated for more than 500 years, or fewer than 20 generations. This is because populations separated for over 500 years are more likely to contain genetic incompatibilities/unique lineages, while those separated for only a short period of time are unlikely to contain differences (even if they inhabit divergent conditions). The presence of divergent genetic lineages (see below) suggests prolonged isolation, although lack of divergence at neutral markers may mean mixing has subsequently occurred. This pattern has been widely documented in UK plant populations.

Species with small, restricted distributions, particularly where they are highly fragmented may only contain low levels of diversity within populations. In these cases special effort may be required to capture any remaining diversity, to avoid creating populations with inadequate diversity in the future.

Ecological amplitude/local adaptations: Species with high ecological amplitude, where they inhabit a range of environmental conditions are more likely contain differentially adapted populations. Differentially adapted populations may contain unique adaptive diversity and hence represent a genetic resource. In addition, mixing of differentially adapted populations may disrupt local adaptation and/or lead to outbreeding depression.

Identifying adaptive genetic differences is not straightforward. Morphological variation between populations may be the result of plastic (non genetic) differences, and the neutral genetic markers typically used to assess genetic diversity are not suitable for assessing adaptive genetic variation. Common garden or reciprocal transplant experiments are most frequently used to establish adaptive differences. In the absence of data, a useful (and precautionary) approach is to assume that populations inhabiting highly divergent environmental conditions (and isolated for at least 20 generations) may contain local adaptations.

Genetic data: The amount of genetic diversity within populations and the differences between them can be measured via genetic markers.

Where individual populations contain low levels of genetic diversity, multiple populations may need to be sampled to maximise the capture diversity, and ensure sufficient diversity is conserved to establish viable populations in the future. Populations that only contain a subset of the diversity found in other populations may represent a lower priority for sampling. Measures of genetic diversity include percent polymorphism, allelic diversity and expected heterozygosity.

Where populations contain genetic differences, these should be incorporated into sampling strategies to maximise the capture of genetic diversity. The presence and extent of differences can be obtained by measuring genetic exchange (e.g. parentage analyses), genetic differentiation (e.g. Fst) or identifying genetic groups (e.g. assignment/clustering methods). In addition, populations may contain unique genetic lineages (typically assessed using cpDNA or mtDNA), reflecting divergent evolutionary histories (i.e. isolation in different Pleistocene

refugia). It can be surprisingly difficult to link the level of genetic differentiation between populations to outbreeding depression, but in general the risks are greatest where populations exhibit chromosomal differences (e.g. ploidy level) or have been isolated from gene flow for 100s of generations (e.g. divergent genetic lineages with no evidence of mixing). In the UK divergent genetic lineages are most likely to be the result of differing colonisation routes from continental refugia.

The larger the number of markers employed the stronger the evidence, and new technologies can provide data from hundreds or thousands of genetic markers. Technical problems with RAPD mean this method is considered limited evidence.

Demography: Declines in population size leads to genetic diversity loss. This may reduce the viability of individual populations, and require consideration in sampling strategies to ensure adequate diversity is conserved for future reintroduction. Where the only available populations are small, there is an increased impetus for allocating sample effort to sampling multiple populations. Where small populations exhibit low fitness (e.g. low seed set) this may be indicative of genetic problems and hence potential limitations in their suitability as a donor population.

1. Sampling representative genetic diversity:

Lower risk

- Well dispersed species in continuous habitat.
- High levels of genetic diversity and high levels of gene flow and inhabiting similar ecological conditions

Higher risk (Intensive sampling)

- High ecological amplitude (i.e. local adaptations) and low gene flow.
- Small number of populations that may contain low levels of genetic diversity

2. Selecting donors

Lower risk

- Species reproducing predominantly via selfing or asexual means
- Species with high gene flow among populations and no evidence of local adaptations
- Species with low ecological amplitude in continuous habitat

Higher risk

- Short lived outcrossing species with isolated populations
- Species with high ecological amplitude or populations showing local adaptations
- Outcrossing species with multiple chromosome races

Categories and recommendations:

This section provides a detailed explanation of the categories used in the assessment and associated recommendations.

Step 1 Sampling to capture genetic diversity

Populations selected for sampling should be large and genetically diverse. Ideally samples should be obtained from a range of sites across the distribution to maximise the capture of genetic diversity. Species are assigned to one of three categories according to the risk that a limited amount of diversity will be captured.

- 1. LOW RISK: The species exhibits (or is predicted to exhibit) little genetic structuring among populations and high diversity within them hence most populations should be suitable. It is still advisable to obtain samples from a range of locations, and environmental space.
- 2. MODERATE RISK: Populations are (likely to be) genetically differentiated (e.g. high Fst values, local adaptations) and sampling should be geographically and/or environmentally stratified to encompass this variation. The level of effort will vary; at one end of the scale only a few populations will exhibit divergence and minimal effort is required, while at the other end of the scale all/most populations may exhibit substantial differences and sampling from many populations may be required.
- 3. HIGH RISK: Genetic diversity within populations is likely to be limited and/or differentiation among populations is high. In these situations there is an increased impetus to sample from multiple populations to ensure the diversity is adequately sampled.

Step 2 Donor selection

The key genetic issues in selecting donors is maximising the likelihood of successful restoration and minimising the potential for negative impacts on indigenous populations. In general mixing populations is considered beneficial, increasing levels of genetic diversity and reducing the risk of inbreeding depression and low evolvability. However, incompatibilities may arise between populations that can lead to outbreeding depression or disruption of local adaptations. To provide guidance on the potential risk of problems arsing due to mixing, species are separated into three categories:

- LOW RISK (and likely to be beneficial): In these situations there is little differentiation between populations that would lead to outbreeding depression/swamping, meaning that genetic material from any population can be used for restoration, or mixed to increase diversity.
- 2. SOME RISKS: In these situations genetic differences that may lead to outbreeding depression have been predicted/documented and mixing is not recommended between divergent groups. However, there are situations where mixing may be considered:
 - a. Available donors have limited genetic diversity (e.g. contain 50-100 or fewer individuals) that may comprise persistence of restored populations
 - b. There is a deliberate aim to introduce genetic diversity from ecological divergent populations to mitigate effects of novel environmental conditions/change
 - c. The risks are marginal (e.g. populations only just exceed the illustrative rule-of-thumb guide figures)
- 3. HIGH RISK: Populations should not be mixed as the degree of differentiation between populations means outbreeding depression is likely.

Dealing with uncertainty

The amount, quality and consistency of the information available will vary among species and hence the confidence in estimates will vary. To provide a quantification of the degree of uncertainty associated with assessments we follow a version of the confidence scales employed by the Intergovernmental Panel on Climate Change (IPCC—AR5). Certainty is separated into the evidence (type, amount, consistency and quality) and expert opinion. Evidence is most robust where it is based on multiple high quality studies with consistent, unambiguous results. Given the complexity of natural processes it is possible that there are some instances where the formal evidence base is weak, but agreement among experts is high.

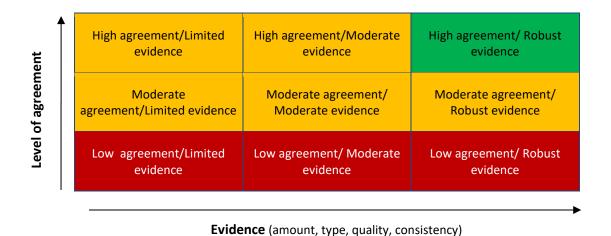


Figure 2. 'Certainty scores' rationale. Shading reflects level of overall certainty

Template and worked example: The following is a template to synthesise the relevant data for assessing the risks of sampling limited genetic diversity and mixing populations inappropriately using the process described above. The template has been completed as a worked example using *Phyteuma spicatum*. It outlines the data acquired and recommendations about sampling strategies obtained using the process described above.

Species: Phyteuma	spicatum			
Recommendations:				Confidence
Sampling:	High risk of sampling limited diversity: The species has a restricted distribution, and many populations are isolated, small in size and declines have been reported. This means genetic diversity may be limited and it is recommended that all/most populations should be sampled to ensure diversity is captured. Non-native populations should be avoided.			High agreement/ Moderate evidence
Donor selection (risk of mixing)	Low risk of mixing: Populations are likely to be (or have recently been) connected by gene flow. In addition, since populations are small and may contain low levels of genetic diversity mixing may be beneficial. However, it is critical to avoid non-native populations.			High agreement/ Moderate evidence
Knowledge <i>gaps</i>	Levels of genetic diversity within and differentiation between populations in the UK			
Information:				
Тахопоту:	Two subspecies are recognised within the UK. while <i>P. s. coeruleum</i> is non-native, the result flower colour ¹ .			
Hybridisation	-			
Life history traits / /	Attributes	Organisation of diversity	Negative outcomes of mixing	Strength of evidence
Dispersal ability	Effective Seed is wind dispersed and both short ¹ and long ² distance movements have reported. Pollination is by bees ¹	Higher diversity/ Lower differentiation	Lower vulnerability	Moderate evidence
Mode of reproduction	Predominantly outcrossing (self incompatible)	Higher diversity/Lower differentiation	Higher vulnerability	Robust evidence

Longevity	Long lived (5-50 years)	Higher diversity/lower differentiation	Lower vulnerability	Robust evidence
First reproduction	1-5 yrs	-	-	Robust evidence
Reproductive output	High (>3000 seeds/plant)	-	Lower vulnerability	Robust evidence
Ploidy	Diploid (2n=22)	-	-	
Range/ Fragmentation:	Restricted distribution with low-moderate fragmentation (of native populations). Native populations are only found in East Sussex county, other populations in England, Scotland and Wales are considered garden escapes.	Lower diversity/ Lower differentiation	Lower vulnerability (but avoid non-natives)	Robust evidence

Ecological	No information on adaptive variation.	Lower	Lower vulnerability	Limited evidence
amplitude	Populations appear to inhabit similar	differentiation	Lower vallierability	Limited evidence
umpiituue	·	unierentiation		
	conditions across range.			
Genetic diversity	No information in UK populations.	Lower	Higher vulnerability	Moderate evidence
	Studies of German populations (using AFLP	diversity/Higher		
	data) indicated low levels of diversity within	differentiation		
	populations and limited gene flow between			
	populations in fragmented landscapes ³ .			
	Population differentiation increased with			
	geographic distance and dispersal up to 20m			
	was estimated ³ . In addition, population size,			
	genetic diversity and fitness (e.g. seedling			
	survival) are positively correlated ^{3,4} .			
	No information on genetic lineages.			
Demography:	Populations small and declining	Lower diversity	Lower vulnerability	Moderate evidence
			(likely beneficial)	
References				
	² Maler A, Emig W, Leins P. 1999. Dispersal Patterns of some <i>Phyteuma</i> Species (Campanulaceae). <i>Plant Biology</i> 1 : 408-417.			
³ Kolb A, Durka W. 2013. Reduced genetic variation mainly affects early rather than late life-cycle stages. <i>Biologic Conservation</i> 159: 367-374.				cle stages. <i>Biological</i>
	⁴ Kolb A. 2005. Reduced reproductive success a Phyteuma spicatum. Journal of Ecology 93 : 1		fragmented populations	s of the forest herb

Bibliography

Angeloni F, Ouborg NJ, Leimu R. 2011. Meta-analysis on the association of population size and life history with inbreeding depression in plants. Biol Conserv. 144: 35–43. doi:10.1016/j.biocon.2010.08.016

Aguilar, R., Ashworth, L., Galetto, L. & Aizen, M. A. 2006. Plant reproductive susceptibility to habitat fragmentation: review and synthesis through a meta-analysis. *Ecology Letters*, 9, 968-980.

Chen, F., Dong, W., Zhang, J., Guo, X., Chen, J., Wang, Z., Lin, Z., Tang, H. & Zhang, L. 2018. The sequenced angiosperm genomes and genome databases. *Frontiers in Plant Sciences*, 9, doi.org/10.3389/fpls.2018.00418.

Frankham, R. 1996. Relationship of population size to genetic variation in wildlife. *Conservation Biology*, 10, 1500-1508.

Frankham, R., Ballou, J., Ralls, K., Eldridge, M. D. B., Dudash, M. R., Fenster, C. B., Lacy, R. & Sunnucks, P. 2017. *Genetic management of fragmented animal and plant populations,* Oxford, Oxford University Press.

Honnay O, Bossuyt B, Jacquemyn H, Shimono A, Uchiyama K. 2008. Can a seed bank maintain the genetic variation in the above ground plant population? Oikos;117: 1–5. doi:10.1111/j.2007.0030-1299.16188.x

Leimu R, Fischer M. 2008. A Meta-Analysis of Local Adaptation in Plants. PLoS One. 3: e4010. Available: http://dx.doi.org/10.1371%2Fjournal.pone.0004010

Leimu R, Fischer M. 2010. Between-Population Outbreeding Affects Plant Defence. PLoS One. 5: e12614. Available: http://dx.doi.org/10.1371%2Fjournal.pone.0012614

Leimu R, Mutikainen PIA, Koricheva J, Fischer M. 2006. How general are positive relationships between plant population size, fitness and genetic variation? J Ecol.94: 942–952. doi:10.1111/j.1365-2745.2006.01150.x

Mimura, M., Yahara, T., Faith, D. P., Vázquez-Domínguez, E., Colautti, R. I., Araki, H., Javadi, F., Núñez-Farfán, J., Mori, A. S., Zhou, S., Hollingsworth, P. M., Neaves, L. E., Fukano, Y., Smith, G. F., Sato, Y.-I., Tachida, H. & Hendry, A. P. 2017. Understanding and monitoring the consequences of human impacts on intraspecific variation. *Evolutionary Applications*, 10, 121-139.

Neaves, L. E., Whitlock, R., Piertney, S. B., Burke, T., Butlin, R. K. & Hollingsworth, P. M. 2015. Implications of climate change for genetic diversity and evolvability in the UK. *Living With Environmental Change Biodiversity Climate Change impacts report card technical paper 15.*

Reed DH, Frankham R. 2001. How Closely Correlated Are Molecular and Quantitative Measures of Genetic Variation ? Evolution. 55: 1095–1103

Todesco, M., Pascual, M. A., Owens, G. L., Ostevik, K. L., Moyers, B. T., Hübner, S., Heredia, S. M., Hahn, M. A., Caseys, C., Bock, D. G. & Rieseberg, L. H. 2016. Hybridization and extinction. *Evolutionary Applications*, 9, 892-908.

Trivedi, C., Cavers, S., Atkinson, N., Clark, J. & Cottrell, J. 2019. A Strategy for UK Forest Genetic Resources, Royal Botanic Gardens, Kew. https://www.kew.org/sites/default/files/2019-06/UK%20Forest%20Genetic%20Resources%20Strategy.pdf

Acknowledgements

This concepts which led to the development of this framework were formed based on the many meetings and workshops undertaken by the UK Conservation Knowledge Exchange Project, which represents a consortium of multiple UK-based institutions. Special thanks to all of its main members without whose contribution this would not have been developed: Peter Hollingsworth, Stuart Piertney, Raj Whitlock, Terry Burke, Andrew Pullin, Jacqualyn Eales as well as Mike Bruford, Richard Ennos, Mike Fay, Barbara Jones, Ian Taylor, Robin Payne, Barbara Smith, Richard Smithers. Many thanks also to Kate Harwick, Clare Trivedi, Ted Chapman and the Millennium Seed Bank Partnership.

The UK Conservation Knowledge Exchange Project, and LN was partially supported by the NERC Knowledge Exchange Grant (NE/H001824/1) to Peter Hollingsworth, Terry Burke, Andrew Pullin.