Genetic assessment for translocation of *Hibbertia* sp. Bankstown

June 2018

FINAL REPORT

The Royal Botanic Gardens & Domain Trust

The Royal
BOTANIC GARDEN
Sydney

Acknowledgements

We acknowledge the Traditional Custodians of the land on which the plant species in this study are found on, and pay respects to Elders past and present. We acknowledge all collectors who conducted the field sampling.

Citation

Bragg JG, Rossetto M (2018) Genetic assessment for translocation of Hibbertia sp. Bankstown: final report. Research Centre for Ecosystem Resilience, Royal Botanic Garden Sydney.

Genetic assessment for translocation of *Hibbertia* sp. Bankstown: final report

Royal Botanic Garden and Domain Trust, Sydney – June 2018

Background

Hibbertia sp. Bankstown (syn. *Hibbertia puberula* subsp. *glabrescens*) consists of one population, distributed across four sites at Bankstown airport in southwest Sydney, and is critically endangered.

Summary of the brief provided by OEH

OEH is planning to perform a translocation, and engaged the Royal Botanic Garden, Sydney, to perform a genetic study to inform decisions related to the translocation. The study had several aims, which are summarised below.

The role of genetics in translocation

In general, genetic analyses performed prior to translocation should aim to:
(a) provide an understanding the genetic health, population structure and genetic diversity patterns of the target species across the geographic range;
(b) identify genetic diversity at the individual level and whether genetic patterns of divergence are consistent across the related populations;
(c) determine whether the subject plant populations have low fitness and low genetic variability that would reduce their potential to persist and adapt to future environmental changes.

Specific aims for a genetic study of *Hibbertia* sp. Bankstown

In our study of *Hibbertia* sp. Bankstown, we set out to determine if there is genetic differentiation among the four sub-populations, and to examine the extent of clonality and inbreeding. These analyses will evaluate levels of genetic diversity across the target species, and also how this diversity is distributed across the sub-populations, and across individuals within the sub-populations. This corresponds to aims **(a)-(c)** outlined above. Additionally, we wanted to complete analyses to inform the design of translocated populations, in terms of the composition of the material that is used. That is, we identify mixtures of individuals from the natural subpopulations that maximize genetic diversity, and reduces the risk of inbreeding in the newly established populations. Finally, we comment on the patterns of relatedness in the population of *Hibbertia* sp. Bankstown in comparison to two congeners, *Hibbertia fumana* and *Hibbertia aspera*.

Scope

Below, we report the methods that were used to generate data used in this study, and the analyses of these data that were undertaken. We then describe the results of the analyses. These are formatted to describe a series of *Observations* concerning the distribution of genetic variation, followed by their *Implications* for the design of translocation populations.

Methods

Sampling

We collected 46 leaf tissue samples from plants in the natural population of *Hibbertia* sp. Bankstown. In each case, we removed approximately 2 g of fresh leaf tissue for genetic analyses. In collecting samples, we aimed to collect a representative sample across the four clusters of individuals. We did not sample every single ramet in order to reduce unnecessary impact on the population. However, in several cases, we collected individuals very close together, which we expected had reproduced vegetatively, and to help validate our ability to detect clonal individuals using genetic data. For each individual, geographic location was estimated using a high resolution GPS instrument, and the

location was marked so that genotyped individuals could be accessed at a later date for subsequent collections of seeds or cuttings.
We collected samples from individuals of *Hibbertia fumana* and *Hibbertia aspera* at several sites. At these sites, we sampled individuals across relatively small spatial areas, so our observations are comparable to *Hibbertia* sp. Bankstown.

Genotyping
Each tissue sample was genotyped using a reduced representation sequencing method called DArTseq. DArTseq was performed by Diversity Arrays Technology Australia (Canberra, Australia). Details of this method are described elsewhere (Sansaloni et al. 2010). Briefly, the approach consists of a restriction digest of the genome, followed by sequencing of the digestion products using an Illumina instrument. Here we describe results from a DArTseq analysis of the 46 samples of *Hibbertia* sp. Bankstown as well as several populations of *Hibbertia aspera* and *Hibbertia fumana*.

Data filtering and analysis
Analyses of genetic relationships among *Hibbertia* sp. Bankstown individuals were then undertaken using a dataset consisting of 4262 single nucleotide polymorphisms (SNPs), including a principle component analysis (PCA), a clustering analysis (estimation of a UPGMA tree), and the estimation of pairwise kinship coefficients between individuals. Finally, we performed analyses to identify possible translocation populations that were optimized for different properties. That is, given the opportunity to establish a population of 296 individuals, using cuttings from the 37 genotyped individuals (including one representative of each clone), we identified mixtures with optimal levels of mean pairwise kinship, or genetic diversity, or both. To perform this optimization, we developed an approach based on the established Simulated Annealing algorithm.

We also estimated genetic relationships between individuals in the natural population of *Hibbertia* sp. Bankstown and individuals grown from seed collected in previous visits to the site. Finally, we compared genetic diversity within *Hibbertia* sp. Bankstown to *H. aspera* and *H. fumana*.

All analyses were performed in the statistical computing package R, using publically available packages including SNPRelate (kinship), adegenet (PCA) and phangorn (UPGMA), as well as two packages developed at the RBGDT, Sydney: RRtools (data processing and filtering) and OptGenMix (optimization of mixtures).

The population genetics of *Hibbertia puberula glabrescens*

Genetic variation: an overview

Observations
Our analyses have identified genetic variability among individuals of *Hibbertia* sp. Bankstown. This variation is distributed across individual plants and across space in ways that inform the design of translocation populations. The individuals of *Hibbertia* sp. Bankstown occur in four discrete clusters in space (Figure 1), and genetic variation among the individuals largely reflects this spatial arrangement (Figures 2, 3). However, there is a difference in that the two eastern populations (populations 1 and 2), which are separated in space, together form a single genetic cluster (Figures 2, 3). We see evidence of inbreeding and clonality, and these features also have a spatial arrangement. We calculated a coefficient of inbreeding, F, based on the proportion of loci that are observed to be homozygous. Positive values ($0 < F < 1$) of this coefficient indicate a surplus of homozygous sites, relative to the number that is expected in the absence of inbreeding. For *Hibbertia* sp. Bankstown, we observe $F \sim 0.3$, suggesting a moderate level of inbreeding. Additionally, we observed variation in inbreeding coefficients among individuals, with samples collected from population 3 exhibiting relatively high individual coefficients of inbreeding.

Implications
The individuals of *Hibbertia* sp. Bankstown form three genetic clusters (roughly, three family groups). This suggests that, other things being equal, the genetic diversity of a translocation

population is likely to be maximized by drawing material (cuttings or seeds) from these clusters in roughly equal proportions, or in proportion to the amount of genetic diversity within the respective clusters.

Kinship analyses

Observations
A number of samples had pairwise kinship values of approximately 0.5, which is indicative of being genetically identical, or clonal. Individuals that appeared to be clones (kinship >0.45) were almost exclusively close together in space. This included one pair of ramets from a single clone in population 1, a set of six ramets from a clone in population 2, and ramets from two sets of clones in population 3 (Appendix 1). Other than clonal individuals, we observed a spectrum of different levels of relatedness ($0 < \text{kinship} < 0.45$; Figure 3), with most non-zero kinship values being shared by individuals that were placed within the same genetic cluster. This is consistent with the observed positive inbreeding coefficient.

Implications
In general terms, current literature suggests that a good way to design diverse populations for conservation is to use the principle of minimizing mean kinship between individuals (Frankham et al 2017). As a result, we recommend that the representation of plants that share high levels of kinship with other individuals is minimised in translocated populations. The reason is that these highly related plants can only contribute to genetic diversity in a way that is similar to a single individual. Below, we describe strategies for designing populations of *Hibbertia* sp. Bankstown (from cuttings) based on this principle.

Optimal design of translocation populations

Observation and Implications
We sought to identify mixtures of genotyped individuals (with each clone represented by one sample only) that resulted in optimal translocation populations. We identified optimal mixtures of 296 individuals, corresponding to the sizes of the planned translocation population. We identified mixtures of genotypes using three different optimization scenarios (S1-S3) (see Table 1, Figure 4):

- **S1**: Minimization of mean kinship among population members, with the additional constraint that no single genotyped individual can represent more than 5.4% of the translocation population
- **S2**: Minimization of mean allelic covariance among population members, with the additional constraint that no single genotyped individual can represent more than 5.2% of the translocation population
- **S3**: Minimization of both mean kinship and mean allelic covariance, with the constraint that no individual could represent more than 5.4% of the translocation population.

We also performed optimizations where we relaxed the constraint that no individual could represent more than 5.4% of the translocation population. We do not present these, because a large representation of one individual can make populations more vulnerable to infection, or other pests.

Of the three scenarios, we prefer S3, as it uses an established approach, with strong precedent in the literature (minimization of mean kinship), but also produces translocation populations with higher levels of genetic diversity that does optimization on the basis of kinship alone.

Seeds or cuttings?

These results reported above apply to propagation of individuals for translocation populations from cuttings, as cuttings are easy to obtain and provide a way of controlling genetic composition,

especially within a species with such high levels of relatedness. If plants grown from seeds are used in translocation, it is harder to predict the genetic composition of those populations, due to insufficient information about both parents. In general, the data suggest inbreeding is occurring, which would seem to suggest that seeds, on the average, will provide a smaller genetic spectrum than the parent plants. If material grown from seeds is used, the numbers supplied here might still serve as a reasonable guide to the expected diversity of the seeds. That is, we would recommend that seeds drawn from populations 3, 4, and a combination of 1 and 2 ought to be used in roughly equal proportions. However, this would be modified slightly, in that population 3 exhibits substantial inbreeding, such that the representation of individuals from population 3 might be deliberately reduced.

Genetic relationships of previously collected material

We sampled DNA from 32 plants growing at Mt Annan. In Appendix 2, we present kinship coefficients between these plants and members of the natural population of *Hibbertia* sp. Bankstown. An example to illustrate how this Appendix can be used: a number of the individuals under cultivation (e.g. individuals in pots marked DNA002, DNA003, DNA006, etc) are quite closely related (kinship >= 0.3) to sample NSW1024421 from the natural population. We note that close relatives of several individuals from the natural population are prevalent among the individuals growing at Mt Annan.

Comparisons among congeners

We performed analyses of genetic relationships among members of populations of *H. fumana* and *H. aspera* (see Appendices 3 and 4). In particular, *H. aspera* is relatively widely distributed and common. We sampled this species quite intensively over a small spatial scale (~150 m) at a site at Heathcote National Park , with the aim of comparing diversity at this site to *Hibbertia* sp. Bankstown.

Individuals of the species *H.* sp. Bankstown, *H. aspera* and *H. fumana* tended to cluster together with their conspecifics based on genetic data (Appendices 3 and 4). At Heathcote National Park, many of the sampled individuals of *H. aspera* appeared to be clonal. In comparison, clonal growth appeared to be less extensive in *Hibbertia* sp. Bankstown than in a population of *H. aspera* distributed over a similar spatial scale.

Conclusions

Here we provide advice on the design of translocation populations for *Hibbertia* sp. Bankstown. We find that there is genetic variation among individuals at the site. There is also evidence of clonality, and inbreeding. We find that there are three genetic clusters of individuals, corresponding to (1) population 3, (2) population 4, and (3) populations 1 and 2. More specifically, we provide recommendations for how many cuttings of each individual might be included in a translocation population 296 individuals. For smaller translocation populations, we recommend using a fraction of the suggested numbers. We make recommendations based on several different sets of criteria. In particular we recommend S3, which appears to strike a good balance between minimizing kinship among individuals, and maximizing genetic diversity.

Final remarks

We developed a method for optimizing the genetic composition of translocation populations that is configured to accept numerical limits on the number of clones that are available of a specific individual, and that can be run with target translocation populations of different sizes. We can

informally generate new recommendations if plans for the translocation population sizes change, or if the recommended numbers of particular clones cannot be cultivated. We also recommend that when translocation populations are planted, that care be taken to avoid placing genetically identical individuals close together at the translocation site. We are happy to consult on spatial arrangement of planted individuals as plans progress.

Literature Cited

Frankham R, Ballou JD, Ralls K, Eldridge M, Dudash MR, Fenster CB, Lacy RC, Sunnucks P 2017 *Genetic management of fragmented animal and plant populations*. Oxford.

Sansaloni CP, Petroli CD, Carling J, Hudson CJ, Steane DA, Myburg AA, Grattapaglia D, Vaillancourt RE, Kilian A 2010 A high-density Diversity Arrays Technology (DArT) microarray for genome-wide genotyping in *Eucalyptus. Plant Methods* 2010 6:16.

Table 1. For each sample (clones excluded – see Appendix 2), we have tabulated the numbers of individuals that would be used in translocation of 296 individuals that are optimized using different criteria. We found optimum numbers for three scenarios – S1: minimize mean kinship; S2: minimize mean covariance; S3: minimize mean kinship and minimize mean covariance. In each scenario, we do not let any individual represent more than 5.4% of total plantings. We have highlighted S3 (green), as it is based on a principle that has strong precedent in the conservation genetics literature, and which does a good job at conserving high levels of genetic diversity.

Sample	S1	S2	S3
NSW1024401	14	0	4
NSW1024243	5	4	3
NSW1024273	0	15	2
NSW1024206	9	3	8
NSW1024211	16	16	16
NSW1024421	15	16	16
NSW1024375	11	16	16
NSW1024397	8	3	6
NSW1024471	12	1	5
NSW1024485	9	2	4
NSW1024480	0	9	0
NSW1024409	0	15	9
NSW1024264	16	2	12
NSW1024424	0	2	1
NSW1024176	16	8	13
NSW1024249	0	16	6
NSW1024425	13	5	7
NSW1024221	8	7	8
NSW1024278	0	16	15
NSW1024171	16	4	11
NSW1024472	9	4	11
NSW1024244	0	0	1
NSW1024404	1	2	0
NSW1024403	2	0	1
NSW1024259	16	16	16
NSW1024487	16	0	12
NSW1024402	1	0	2
NSW1024410	13	0	7
NSW1024417	16	12	16
NSW1024396	4	6	3
NSW1024488	4	16	10
NSW1024269	0	16	0
NSW1024398	0	9	3
NSW1024399	16	16	16
NSW1024172	16	11	16
NSW1024268	0	12	4
NSW1024415	14	16	16

Figure 1. A map of the range of *Hibbertia* sp. Bankstown in a field at the aerodrome. Dots represent sampling locations.

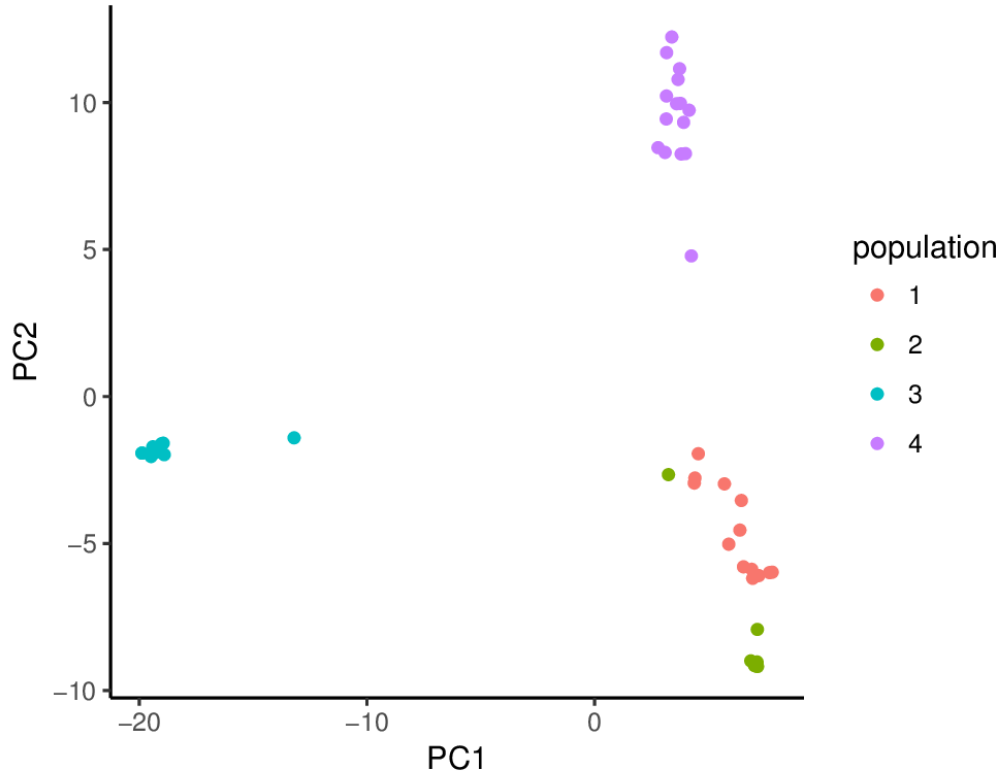


Figure 2. A principle component analysis was performed to summarize genetic variation across the 46 individuals in a small number of principle components (PCs). The scores for the two principle components for each sample are plotted. These are PC1, which explains 34% of genetic distance among the samples, and PC2, which explains 18% of genetic variation among the samples. Data points are coloured according to the spatial population (see Figure 1). Three clusters of points can be observed, corresponding to P3, P4, and the final cluster corresponds to P1 and P2.

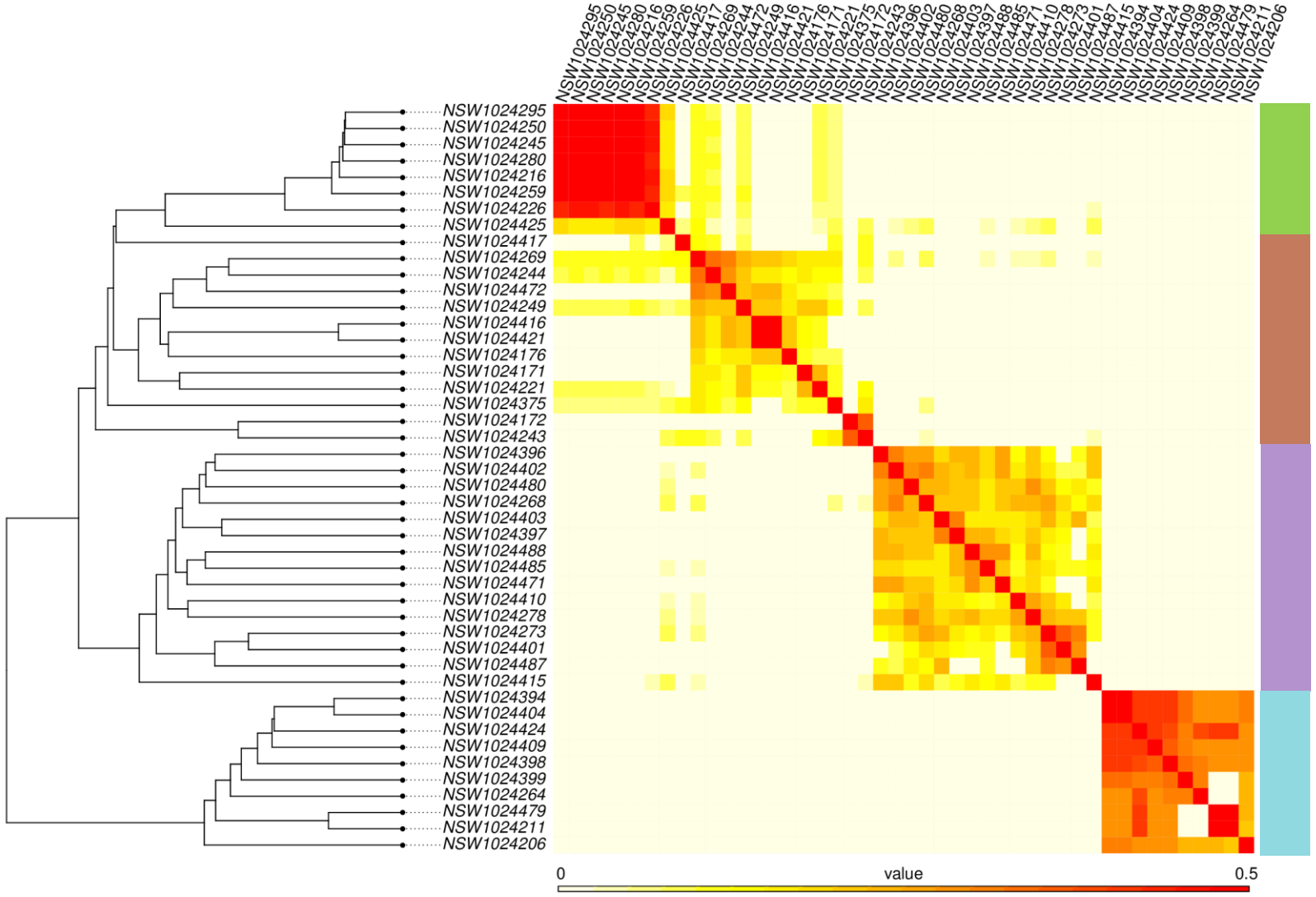

NSW1024295
NSW1024250
NSW1024245
NSW1024280
NSW1024216
NSW1024259
NSW1024226
NSW1024425
NSW1024417
NSW1024269
NSW1024244
NSW1024472
NSW1024249
NSW1024416
NSW1024421
NSW1024176
NSW1024171
NSW1024221
NSW1024375
NSW1024172
NSW1024243
NSW1024396
NSW1024402
NSW1024480
NSW1024268
NSW1024403
NSW1024397
NSW1024488
NSW1024485
NSW1024471
NSW1024410
NSW1024278
NSW1024273
NSW1024401
NSW1024487
NSW1024415
NSW1024394
NSW1024404
NSW1024424
NSW1024409
NSW1024398
NSW1024399
NSW1024264
NSW1024479
NSW1024211
NSW1024206
0
value
0.5

Figure 3. Heatmap displays kinship estimated for each pair of samples from genetic data. Red corresponds to high kinship. The maximum value is 0.5, and corresponds to kinship between an individual and itself (or between ramets of a clone). On the diagonal are kinship values between individuals and themselves, and these cells are therefore red. Off the diagonal, red cells indicate high levels of kinship, including clonality, between samples. Note that the bottom left triangle of the heatmap is reflected in the upper right triangle. The tree at left is the result of applying a UPGMA clustering algorithm to the genetic data. Samples that are joined by a node that is shallow in the tree (nodes further to the right) tend to be closely related and within the same cluster (e.g. NSW1024416 and NSW1024421 are clonal), and samples that are joined at a nodes deeper in the tree (nodes further to the left) belong to different clusters and tend to be genetically distant (e.g. NSW1024269 and NSW1024424). Colours on right indicate population (see colours Figure 1).

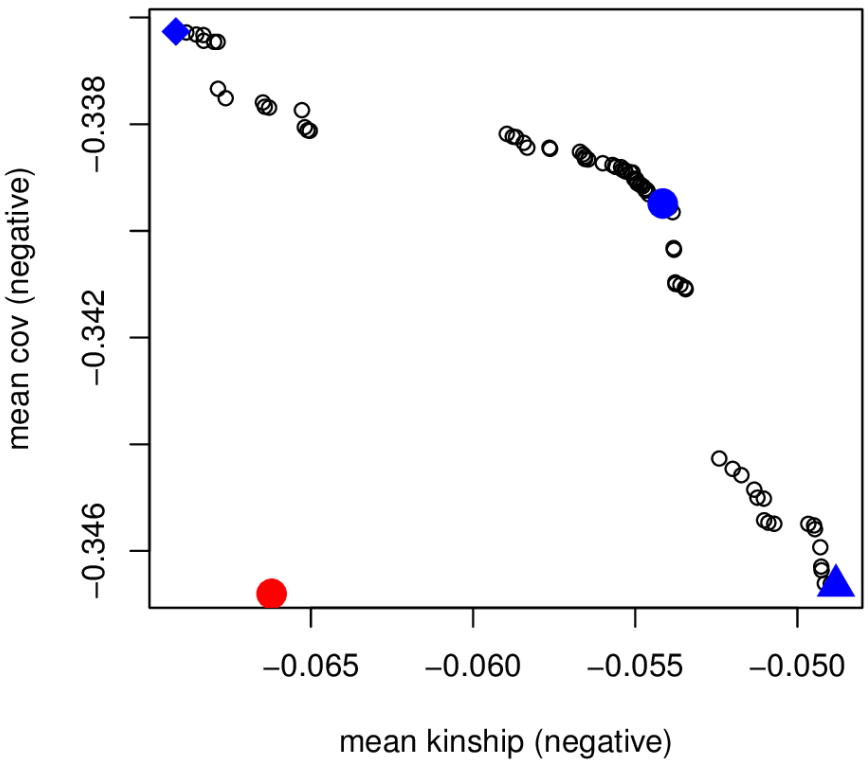


Figure 4. Outcomes of our optimization algorithm. Our goal is to find mixtures of individuals that minimize mean kinship and covariance (maximize the negatives of these values). This plot shows that these two objectives trade-off. Mixtures of individuals with small values of mean kinship, such as scenario S1 (see Table 1, blue triangle) tend to have large values of mean covariance. Mixtures with small values of mean covariance, such as scenario S2 (see Table 1, blue diamond) tend to have large values of mean kinship. We used an algorithm to explore this tradeoff (unfilled circles), and identify scenario S3 (blue circle) as a good compromise. Note, the red circle depicts the values for using an equal number of all individuals in a translocation population.

Appendix 1. Here we indicate which sets of individuals were treated as being clonal. For each population, we list sets of clonal individuals. The individual appearing before the pipe ("|") character was retained in the dataset for optimization analyses. For instance, where we recommend that 10 clones of "**NSW1024259"** be used in an analysis, this individual could be used interchangeably with "NSW1024295," "NSW1024250," etc.

Clonal samples

Population 1:
NSW1024421 | NSW1024416

Population 2:
NSW1024259 | NSW1024295 NSW1024250 NSW1024245 NSW1024280 NSW1024216 NSW1024226

Population 3:
NSW1024404 | NSW1024394
NSW1024211 | NSW1024479

Population 4:
Nil

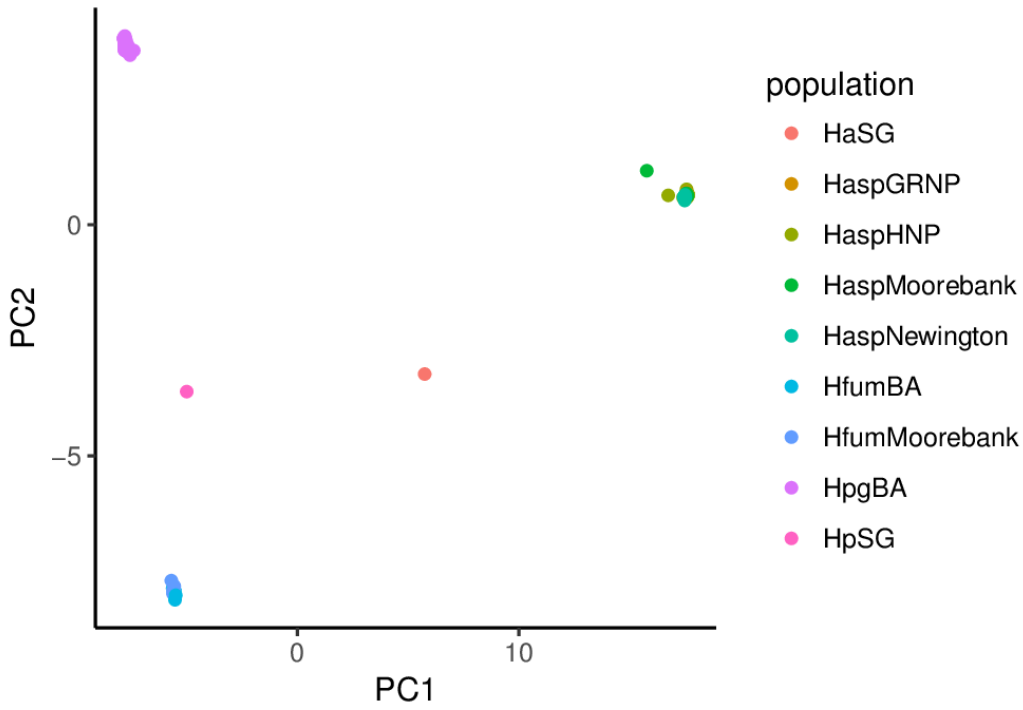
Appendix 2. Coefficients of kinship between the natural population of *Hibbertia* sp. Bankstown (rows) and cultivated plants at Mt Annan (columns). Only non-zero values as shown.

	DNA 002	DNA 003	DNA 004	DNA 006	DNA 007	DNA 008	DNA 009	DNA 010	DNA 011	DNA 012	DNA 013	DNA 014	DNA 015	DNA 016	DNA 017	DNA 018	DNA 019	DNA 020	DNA 022	DNA 023	DNA 024	DNA 025	DNA 026	DNA 027	DNA 028	DNA 029	DNA 030	DNA 031	DNA 022 *	DNA 033 *	DNA 034 *	DNA 035 *
	NS W10 24433	NS W10 24434	NS W10 24435	NS W10 24437	NS W10 24438	NS W10 24439	NS W10 24440	NS W10 24441	NS W10 24442	NS W10 24443	NS W10 24444	NS W10 24445	NS W10 24446	NS W10 24447	NS W10 24448	NS W10 24449	NS W10 24450	NS W10 24451	NS W10 24453	NS W10 24454	NS W10 24455	NS W10 24456	NS W10 24457	NS W10 24458	NS W10 24459	NS W10 24460	NS W10 24461	NS W10 24462	NS W10 24463	NS W10 24464	NS W10 24465	NS W10 24468
NS W10 24401																																
NS W10 24243																												0.13				
NS W10 24273																																
NS W10 24206										0.28		0.26															0.26		0.24	0.26	0.21	0.28
NS W10 24211										0.39		0.26															0.5				0.49	0.27
NS W10 24421	0.39	0.38		0.3	0.39	0.35	0.38	0.38	0.37		0.38		0.36	0.12	0.42	0.39		0.38	0.35	0.41	0.39	0.37		0.39		0.4		0.08				
NS W10 24375														0.12											0.12			0.1				
NS W10 24397																	0.17															
NS W10 24471																	0.22															
NS W10 24485																																
NS W10 24480																	0.25															
NS W10 24409										0.32		0.32															0.37		0.32	0.35	0.26	0.35
NS W10										0.37		0.41															0.41		0.44	0.46		0.29

24264																															
NSW10244 24										0.43		0.37															0.42		0.36	0.41	0.4	0.39
NSW10241 76								0.15						0.1		0.15		0.14	0.11	0.13				0.14				0.09				
NSW10242 49	0.06	0.09			0.08		0.09	0.08	0.04					0.19	0.13	0.08		0.08	0.09	0.14	0.08	0.1				0.08		0.2				
NSW10244 25			0.13																									0.02				
NSW10242 21														0.13														0.19				
NSW10242 78																	0.22															
NSW10241 71														0.19														0.14				
NSW10244 72	0.09	0.1			0.11	0.09	0.12	0.13			0.09		0.08	0.12	0.18	0.12		0.1	0.13	0.16	0.11	0.13		0.1		0.1		0.15				
NSW10242 44														0.1	0.08				0.09	0.08			0.12					0.13				
NSW10244 04										0.35		0.34															0.36		0.31	0.35	0.27	0.48
NSW10244 03																	0.16															
NSW10242 59			0.4																									0.06				
NSW10244 87																																
NSW10244 02																	0.2															
NSW10																	0.23															

24410																															
NSW10244 17																							0.4								
NSW10243 96																0.19															
NSW10244 88																0.16															
NSW10242 69	0.09	0.17	0.02	0.1	0.11	0.09	0.13	0.13	0.1		0.11		0.07	0.18	0.17	0.13		0.12	0.14	0.17	0.11	0.14	0.16	0.06	0.03	0.14		0.19				
NSW10243 98										0.4		0.43															0.35		0.35	0.33	0.28	0.39
NSW10243 99										0.41		0.36															0.34		0.34	0.5		0.31
NSW10241 72																															
NSW10242 68																0.16															
NSW10244 15																															

Appendix 3. Principal component analysis of *Hibbertia* sp. Bankstown, *H. aspera*, and *H. fumana*. Different sites for *H. fumana* and *H. aspera* are indicated by the legend. Note two samples collected at Sylvain Grove appear to be genetically intermediate between the focal species. It is possible these samples are different subspecies of *H. puberula* or *H. aspera*, or that they represent the outcome of hybridization between species.



Appendix 4. A cluster diagram of relationships among samples of H. aspera, H. fumana, H. sp. Bankstown. See attached pdf.