Article (refereed) - postprint

Gray, Alan; Perry, Annika; Cavers, Stephen; Eastwood, Antonia; Biermann, Michelle; Darlow, Andrew; Thomas, Vanessa; Lambdon, Phil. 2017. **Hybrid plants preserve unique genetic variation in the St Helena endemic trees Commidendrum rotundifolium DC Roxb. and C. spurium (G.Forst.) DC.** *Conservation Genetics*, 18 (1). 241-246. 10.1007/s10592-016-0884-8

Title: Hybrid plants preserve unique genetic variation in the St Helena endemic trees *Commidendrum rotundifolium* (Roxb.) DC and *C. spurium* (G.Forst.) DC.

Alan Gray[1], Annika Perry[1], Stephen Cavers[1], Antonia Eastwood[2], Michelle Biermann[1], Andrew Darlow[3], Vanessa Thomas[3], and Phil Lambdon[4]

[1] NERC Centre for Ecology and Hydrology, Bush Estate, Penicuik, Edinburgh EH26 0QB, UK

[2] The James Hutton Institute, Craigiebuckler, Aberdeen, AB15 8QH, Scotland, UK

[3] Environmental Management Division, Environment and Natural Resources Directorate, St. Helena Government Scotland, St. Helena STHL IZZ

[4] Royal Botanic Gardens Kew, Richmond, Surrey, TW9 3AB

Corresponding Author: Alan Gray

alangray@ceh.ac.uk

Tel: +44 (0) 131 445 8471

Fax: +44 (0) 131 445 3943

Abstract

The island of St Helena in the South Atlantic Ocean has a rich endemic flora, with 10 endemic genera and 45 recognised endemic species. However, populations of most endemic species have undergone dramatic reductions or extinction due to over-exploitation, habitat destruction and competition from invasive species. Consequently, endemic species are likely to have lost genetic variation, in some cases to extreme degrees. Here, the entire extant wild populations and all planted trees in seed orchards, of two critically endangered species in the endemic genus *Commidendrum* (Asteraceae), *C. rotundifolium* and *C. spurium*, were sampled to assess levels of genetic variation and inbreeding. Six new microsatellite loci were developed from next-generation sequence data, and a total of 190 samples were genotyped. Some seed orchard trees contained alleles from both wild *C. rotundifolium* and *C. spurium* indicating they could be hybrids and that some backcrossing may have occurred. Some of these trees were more similar to *C. rotundifolium* than *C. spurium* both genetically and morphologically. Importantly, allelic variation was detected in the putative hybrids that was not present in wild material. *C. rotundifolium* is represented by just two individuals one wild and one planted and *C. spurium* by seven, therefore the seed orchard trees comprise an important part of the total remaining genetic diversity in the genus *Commidendrum.*

Keywords: allelic variation, genetic conservation, rarity, breeding programme, endemic plants

Introduction

Islands make a disproportionate contribution to global biodiversity as they house distinct evolutionary lineages of endemic species, and many are biodiversity hotspots (Myers et al. 2000; Emerson & Kolm 2005). Island floras are widely threatened by invasive species, exploitation, habitat degradation and climate change, the results of which can cause severe reductions in population sizes (Cronk 1986; Glen et al. 2013; Courchamp et al. 2014). Small population sizes and fragmentation can reduce genetic diversity, and disrupt gene flow and inbreeding, with consequential declines in fitness (Ellstrand & Elam 1993). Another risk, for small plant populations in particular, is hybridisation either through exposure to larger populations of closely related species (e.g. *Hyacinthoides* spp. in the UK, see Kohn et al. 2009) or where previously geographically-separated close relatives are brought together (e.g. *Trochetiopsis* on St Helena, see Cronk 1995). In combination, these threats highlight that island biodiversity is in urgent need of assessment and conservation before genetic variation is lost forever.

Hybridization among plants is an important evolutionary mechanism with the origin of 40–80% of angiosperms estimated to involve either hybridisation or changes in ploidy (Stebbins 1950; Stace 1975; Rieseberg et al. 1993; Rhymer & Simberloff 1996). However, hybridization is a conservation risk for rare and/or endangered species, potentially threatening their genetic integrity (Levin et al. 1996). Gene flow between related species can compromise fitness by the wastage of reproductive effort (Levin, Francisco-Ortega et al. 1996). It can also be a threat where it occurs between differently-adapted populations of a single species by disrupting co-adapted gene complexes (Rhymer & Simberloff 1996). Of particular concern is when hybrids display greater fitness than either or both of the parental species (hybrid vigour

or heterosis) causing competition with the parental species (Rhymer & Simberloff 1996; Emms & Arnold 1997). Hybridization is more likely where there are limited options for out-breeding (Rhymer & Simberloff 1996; Kothera et al. 2007) or where isolation barriers between two previously isolated species are broken (Ellstrand & Schierenbeck 2000). On the other hand, where species have become critically endangered to the extent that only a few individuals remain, hybridization may be the only means to preserve alleles that would be lost to extinction (Fant et al. 2010), especially where outbreeding is obligate due to mechanisms for self-incompatibility.

On St Helena, a small island (122 km^2) in the South Atlantic (15° 58'S and 5°43'W, Suppl. Figure S1a) several endemic species are at risk. The endemic genus *Commidendrum* DC. (Asteraceae), the 'gumwoods', contains four very closely related species (Eastwood et al. 2004), all severely threatened in the main by introduced species. Hybridisation has also been found among *Commidendrum* species (Eastwood 2003), but was not thought to be widespread. *Commidendrum rotundifolium* (Roxb.) DC. was classified by IUCN as Extinct in the Wild until recently as it was rediscovered at the top of a cliff edge (Suppl. Figure S1b) but remains Critically Endangered and is likely to again be classified as Extinct in the Wild when this individual dies unless other individuals are discovered. *Commidendrum spurium* (G. Forst.) DC. is Critically Endangered with the largest population currently comprising just seven individuals. The other two species, *C. rugosum* (Dryand) DC. and *C. robustum* DC., are slightly more widespread and have larger populations (approximately 35,000 and 680 individuals respectively). The latter species is almost exclusively confined to a single site but has been the focus of a successful community woodland restoration project (Figure S1b). The extremely small sizes of these populations, allied with the self-incompatibility of the species (Eastwood 2003), places substantial barriers to establishment of self-sustaining populations.

Until very recently all extant *C. rotundifolium* were the progeny of one individual tree which has subsequently died. At least nine trees were established from this individual and grown in a seed orchard (at Pounceys, Suppl. Figure S1b). In 1998, seedlings were raised from the Pounceys seed orchard and planted in a second seed orchard at Scotland (Suppl. Figure S1b). In 2002, more seedlings were raised and planted at a third seed orchard at Barren Ground (Suppl. Figure S1b). All but one of the original nine progeny at Pounceys have since died, leaving this individual and the seed orchard stock as the entire surviving *C. rotundifolium* population at the time. However, as the seed orchard trees at Scotland and Barren Ground matured, morphological ambiguity suggested that these may be of hybrid origin. Several *C. spurium* trees grew adjacent to the original planting site (Pounceys), and are likely to be the co-parental species. To inform decisions for the recovery and re-introduction of *C. rotundifolium* and *C. spurium*, it was necessary to establish the hybrid status and levels of extant genetic diversity in the seed orchards for both species. In this study, we specifically aimed to:

1. establish the possible hybrid status of seed orchard trees, and
2. identify any pure *C. rotundifolium* or *C spurium* plants for subsequent conservation breeding.

Methods

Collection of samples

With the exception of samples of *C. spurium* taken from the living collection at the Royal Botanic Gardens Edinburgh, all samples were collected on St Helena from wild and seed orchard populations (Figure 1). A total of 191 individuals were collected including all four *Commidendrum* species (*C. spurium*, *C. rotundifolium*, *C. rugosum* and *C. robustum*) and the putative hybrid samples from the seed orchards at Scotland and Barren Ground and a few trees planted at the George Benjamin Arboretum (GBA). Leaf samples were collected into polythene bags containing silica gel, between 01/06/2010 and 15/07/2010. *C. rotundifolium* came from the single planted individual at Pounceys and a wild plant near Botley's and four seedlings planted at Drummond Point. Wild *C. spurium* was collected from the seven individuals at Mount Vessey and included fresh material of *C. spurium* was donated by the Royal Botanic Gardens Edinburgh (RBGE Accession number 20000247E; collected 05/10/2012) from which initial sequences were generated. Additional material of *C. robustum* (Peak Dale, Thompsons Wood, Deep Valley and Millenium Forest) and *C. rugosum* (Man and Horse, Horse Point and Blue Point) was also included for comparison.

Laboratory Methods

DNA was extracted from leaf tissue using DNeasy 96 and Mini plant kits (Qiagen), following manufacturer's instructions.

To identify microsatellite loci, over 48 million bases of genomic DNA sequence were obtained from *C. spurium* by 454 sequencing using a GS FLX (GATC Biotech). The sequence was searched for 3, 4, 5 and 6 base pair repeat sequences using msatcommander (Rozen & Skaletsky

2000; Faircloth 2008) and primers were designed for 48 potential marker loci. In all cases an M13 sequence tag was added to the 5' end of the forward primer. Potential markers were used in polymerase chain reaction (PCR) amplification in 8 individuals from the sample set and those showing consistent amplification and potential for diagnostic purposes were amplified in a further subset of 28 individuals. Five trinucleotide and one tetranucleotide microsatellite loci were chosen and the whole set of 191 samples were genotyped.

All microsatellites were amplified using 10 µl PCR reactions, each comprising 1µl of genomic DNA (diluted from original elution to 1:10), 1.5mM MgCl2, 1 X PCR Buffer, 200 µM each dNTP, 0.2 µM each primer, 0.2 µM IRD fluorescent labelled M13 primer (700 or 800), 20% v/v BSA and 1 U Taq DNA polymerase. Reactions were run on a Hybaid MBS thermocycler using the following protocol for all loci: 5 min at 95 °C, then 10 cycles of 30 sec at 94 °C, 1 min at 57 °C, 30 sec at 72 °C, followed by 22 cycles of 30 sec at 94 °C, 30 sec at 55 °C, 30 sec at 72 °C, followed by 10 min at 72 °C. PCR products were then separated on an 8% denaturing polyacrylamide gel (25 cm), and visualised using a LI-COR 4200 IR2 automated genotyper. PCR products were run out alongside a standard and fragment sizes were scored by eye.

Data Analyses

All summary statistics (number of different alleles; number of effective alleles; observed heterozygosity; expected heterozygosity; unbiased expected heterozygosity; fixation index) were calculated using GenAlex (version 6.501) (Peakall & Smouse 2006; Peakall & Smouse 2012). A Principal Coordinate Analysis (PCoA) was also computed based on the pairwise genetic distance matrix to examine relatedness among samples. Pairwise genetic distance was estimated among all pairs of samples using the squared distance method for codominant genotypes as implemented in GenAlEx.

Results and Discussion

Across all species, very little variation was evident in marker screening, and only 6 loci had any variation, probably due to the extremely small population sizes and inbred nature of the extant trees. Although several loci were monomorphic within species, loci were retained on the basis that they showed polymorphism either within or among species. It was not possible to determine whether monomorphism at a locus was due to null alleles or homozygosity, but as we had completely sampled all plants of the extant populations, these markers were nevertheless useful for species resolution, and to indicate gene pool variation (comparing species with putative hybrids).

Overall, levels of genetic variation within and among species were very low, with mean numbers of alleles per locus, N_a = 1 - 2.5 (Table 1), especially for *C. rotundifolium* (N_a = 1). Seed orchard plants had slightly higher levels of variation (N_a = 2.5) than wild populations, except for *C. rugosum* where sample size was small. One allele was present in some of the hybrids that was not found in any of the extant parent plants (Locus 6, allele 162, Supplementary Table S1). Most samples shared the majority of alleles with *C. spurium*, but a few samples had more allelles in common with *C. rotundifolium* (one sample shared 75%, Supplementary Table S2, and displayed leaf morphology closer to *C. rotundifolium*). The first 3 axes of the PCoA explained over 80% of the variation in the data (Figure 2). All of the seed orchard samples from Barren Ground, Scotland and GBA had alleles found in both *C. rotundifolium* and *C. spurium* supporting their putative hybrid origin (Figure 2). The distribution of hybrid samples on axes 1 & 2, range from being close to a putative parents or somewhat intermediate between the two, suggesting both first generation hybrids and hybrid-parent backcrossed progeny may be present (Figure 2).

The data indicated a very low level of genetic diversity in these threatened species as expected from the extremely small extant population sizes. They also support the suspected hybrid origin of the seed orchard plants, as had been suggested by previous studies (Eastwood 2003). Level of heterozygosity in the putative hybrid samples were higher than expected (H_o = 0.77, H_e = 0.51, mean F = -0.52, Table 1), possibly indicative of combination of gene pools as would occur in hybridisation. None of the seed orchard plants were pure *C. rotundifolium* or *C. spurium* but alleles were discovered in some of the hybrids that were not present in any of the wild or wild derived individuals. These alleles may be derived from now-extinct *C. rotundifolium* or *C. spurium* parent populations. The hybrid plants may therefore represent a repository of genetic variation, which merits careful conservation on St Helena given the extremely limited genetic variation present in both *C. rotundifolium* and *C. spurium* in the wild. This potential repository is significant due to the self-incompatibility system in both species, which results in limited seed production from mating between closely-related individuals.

Conservation Implications

The genus *Commidendrum* is endemic to St Helena and is a unique part of global plant diversity. All species in the genus currently face extinction, being threatened by invasive species, exploitation, habitat degradation and climate change. A further difficulty is effective propagation due to self-incompatibility mechanisms. For species such as these, hybrid plants may represent a valuable source of variation that would otherwise be lost via extinction (Fant et al. 2010). In *C. rotundifolium* and *C. spurium*, reproductive success is dependent on mate availability for cross-pollination success and the limited genetic diversity in extant populations of *C. rotundifolium* and *C. spurium* will undoubtedly impede the recovery programme. Our results show that seed orchard trees contain variation not found in the wider population; these

trees should therefore be considered a resource for a controlled breeding programme or genetic rescue (Whiteley et al. 2015). To support this work, as well as continuing to develop the record of their unique genetic variation through wider genomic sequencing, additional studies of inter-fertility and propagation are urgently required for the *Commidendrum* species. However, any moves to implement genetic rescue or hybrid breeding should take careful account of the ethical questions that arise when dealing with highly threatened species. We recommend that the natural populations are maintained 'as is' but that other mixed putative hybrid populations should be established. In such threatened populations we suggest that the conservation of *Commidendrum* should focus on all genetic diversity and this is as much of a priority as conserving taxonomic species.

Acknowledgments

We would like to thank Katrine Herian, Lourens Mallen, Matti Niissalo, Neil Henry, Jeremy Henry and David Francis who all collected samples in the field. We also thank Sadie Barber, Fiona Inches, Greg Kenicer and Suzanne Cubey from RBGE for facilitating and granting access to the living collections of *C. spurium* at RBGE. We would like to acknowledge the financial support from the Darwin Initiative, St Helena Agriculture and Natural Resources Department and the Natural Environment Research Council. Finally we would like to the editor and the anonymous reviewers for comments and suggestions that greatly improved the manuscript.

References

Courchamp F, Hoffmann BD, Russell JC, Leclerc C, Bellard C (2014) Climate change, sea-level rise, and conservation: keeping island biodiversity afloat. *Trends in Ecology & Evolution*, **29**, 127-130.

Cronk Q (1995) A new species and hybrid in the St Helena endemic genus Trochetiopsis. *Edinburgh Journal of Botany*, **52**, 205-213.

Cronk QCB (1986) The decline of the St Helena gumwood Commidendrum robustum. *Biological Conservation*, **35**, 173-186.

Eastwood A (2003) Use of molecular tools in the conservation biology of the St Helenan flora. University of Edinburgh, Edinburgh.

Eastwood A, Gibby M, Cronk QCB (2004) Evolution of St Helena arborescent Astereae (Asteraceae): relationships of the genera Commidendrum and Melanodendron. *Botanical Journal of the Linnean Society*, **144**, 69-83.

Ellstrand NC, Elam DR (1993) Population Genetic Consequences of Small Population Size: Implications for Plant Conservation. *Annual Review of Ecology and Systematics*, **24**, 217-242.

Ellstrand NC, Schierenbeck KA (2000) Hybridization as a stimulus for the evolution of invasiveness in plants? *Proceedings of the National Academy of Sciences*, **97**, 7043-7050.

Emerson BC, Kolm N (2005) Species diversity can drive speciation. *Nature*, **434**, 1015-1017.

Emms SK, Arnold ML (1997) The Effect of Habitat on Parental and Hybrid Fitness: Transplant Experiments With Louisiana Irises. *Evolution*, **51**, 1112-1119.

Faircloth B (2008) MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources*, **8**, 92-94.

Fant JB, Banai A, Havens K, Vitt P (2010) Hybridization between the threatened plant, Lespedeza leptostachya Englem. and its co-occurring congener Lespedeza capitata Michx.: morphological and molecular evidence. *Conserv Genet*, **11**, 2195-2205.

Glen A, Atkinson R, Campbell K, Hagen E, Holmes N, Keitt B, Parkes J, Saunders A, Sawyer J, Torres H (2013) Eradicating multiple invasive species on inhabited islands: the next big step in island restoration? *Biol Invasions*, **15**, 2589-2603.

Kohn DD, Hulme PE, Hollingsworth PM, Butler A (2009) Are native bluebells (Hyacinthoides non-scripta) at risk from alien congenerics? Evidence from distributions and co-occurrence in Scotland. *Biological Conservation*, **142**, 61-74.

Kothera L, Ward SM, Carney SE (2007) Assessing the threat from hybridization to the rare endemic Physaria bellii Mulligan (Brassicaceae). *Biological Conservation*, **140**, 110-118.

Levin DA, Francisco-Ortega J, Jansen RK (1996) Hybridization and the Extinction of Rare Plant Species. *Conservation Biology*, **10**, 10-16.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853-858.

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research *Molecular Ecology Notes*, **6**, 288-295.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*, **28**, 2537-2539.

Rhymer JM, Simberloff D (1996) Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics*, **27**, 83-109.

Rieseberg LH, Ellstrand NC, Arnold M (1993) What Can Molecular and Morphological Markers Tell Us About Plant Hybridization? *Critical Reviews in Plant Sciences*, **12**, 213-241.

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (eds. Krawetz S, Misener S). Humana Press, Totowa, NJ.

Stace C (1975) *Hybridization and the flora of the British Isles*. Academic Press Inc., London.

Stebbins G (1950) *Variation and evolution in plants*. Columbia University Press, New York.

Whiteley AR, Fitzpatrick SW, Funk WC, Tallmon DA (2015) Genetic rescue to the rescue. *Trends in Ecology & Evolution*, **30**, 42-49.

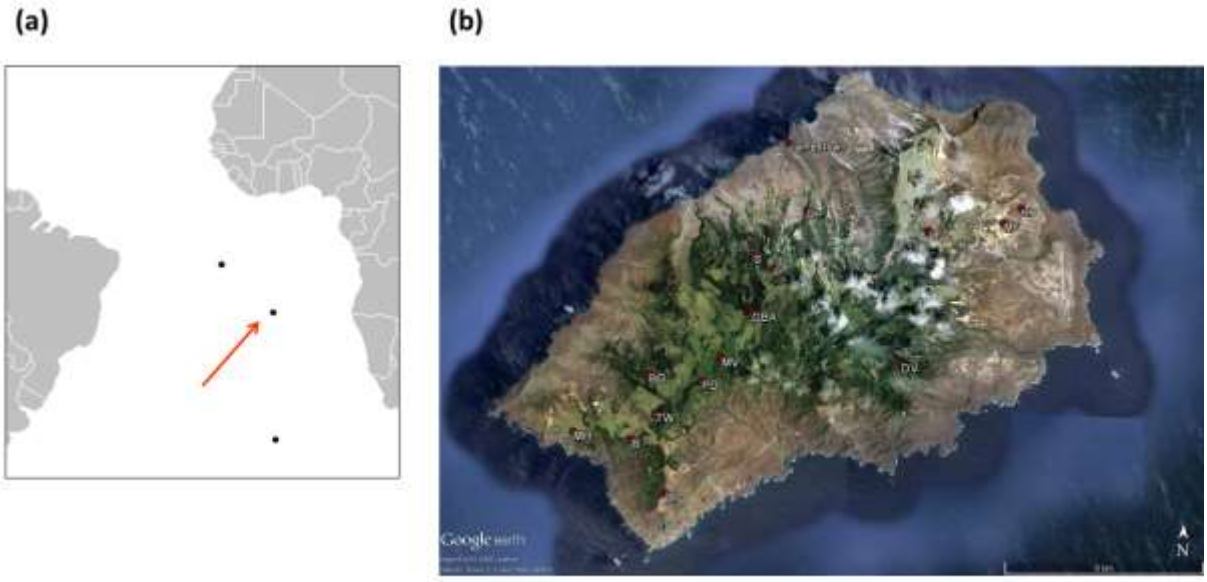

Figure 1: (a) Map showing general locality of St Helena in the south Atlantic Ocean with Ascension Island to the north west and Tristan da Cunha to the south. (b) Location of sites on St Helena mentioned in the main manuscript: BG - Barren Ground, BP - Blue Point, B - Botley's, DV - Deep Valley, DP - Drummond Point, GBA - George Benjamin Arboretum (Casons), HP - Horse Point, LF - Longwood Farm (Picolo), MH - Man and Horse, MF - Millennium Forest, MV - Mount Vessey, PD - Peak Dale, P – Pounceys, S – Scotland, and TW - Thompsons Wood.

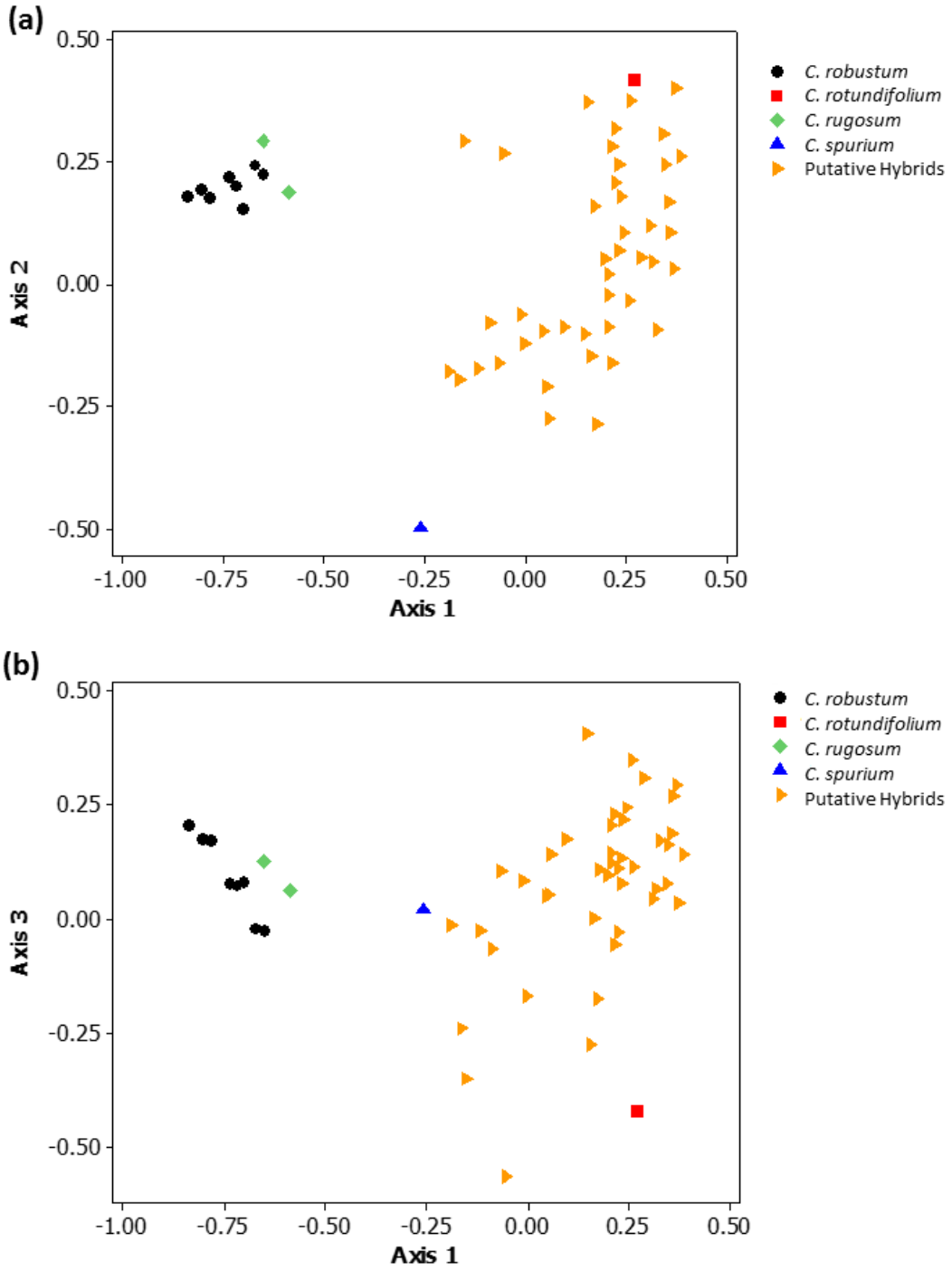

Figure 2: Axes 1 and 2 (a) and axes 1 and 3 (b) from a principal co-ordinates analysis based on genetic distance estimated using 6 microsatellite loci. The percentage of variation explained by these axes was: axis 1 - 47.26 %, axis 2 - 19.26 %, and axis 3 - 15.04 %, cumulative variation – 81.56 %.

273 Table 1: Summary genetic diversity statistics for the four *Commidendrum* species and

274 putative hybrids. Number of genotypes; N_a - No. of Different Alleles; N_e - No. of Effective

275 Alleles; H_o - Observed Heterozygosity; H_e - Expected Heterozygosity; ${}_uH_e$ - Unbiased

276 Expected Heterozygosity; F - Fixation Index.

Samples	No. genotypes	Locus	N	Na	Ne	Ho	He	uHe	F
Seed Orchard	**18**	**6**	151	3.00	2.64	0.77	0.62	0.62	-0.25
		11	151	3.00	1.99	0.90	0.50	0.50	-0.81
		19	151	2.00	1.98	0.79	0.49	0.50	-0.59
		36	151	2.00	2.00	0.79	0.50	0.50	-0.59
		42	151	3.00	1.98	0.87	0.49	0.50	-0.77
		43	151	2.00	1.85	0.50	0.46	0.46	-0.10
			Mean	2.50	2.07	0.77	0.51	0.51	-0.52
			SE	0.22	0.11	0.06	0.02	0.02	0.12
C. robustum	**4**	**6**	0	0.00	0.00	0.00	0.00	0.00	
		11	17	2.00	1.99	0.94	0.50	0.51	-0.89
		19	18	1.00	1.00	0.00	0.00	0.00	#N/A
		36	19	2.00	1.87	0.32	0.47	0.48	0.32
		42	19	2.00	1.82	0.68	0.45	0.46	-0.52
		43	19	1.00	1.00	0.00	0.00	0.00	#N/A
			Mean	1.33	1.28	0.32	0.24	0.24	-0.36
			SE	0.33	0.31	0.17	0.11	0.11	0.25
C. rotundifolium	**1**	**6**	6	1.00	1.00	0.00	0.00	0.00	#N/A
		11	6	1.00	1.00	0.00	0.00	0.00	#N/A
		19	6	1.00	1.00	0.00	0.00	0.00	#N/A
		36	6	1.00	1.00	0.00	0.00	0.00	#N/A
		42	6	1.00	1.00	0.00	0.00	0.00	#N/A
		43	6	1.00	1.00	0.00	0.00	0.00	#N/A
			Mean	1.00	1.00	0.00	0.00	0.00	
			SE	0.00	0.00	0.00	0.00	0.00	0.00
C. rugosum	**6**	**6**	0	0.00	0.00	0.00	0.00	0.00	
		11	7	4.00	2.51	0.57	0.60	0.65	0.05
		19	4	2.00	1.60	0.50	0.38	0.43	-0.33
		36	4	2.00	1.60	0.00	0.38	0.43	1.00
		42	7	3.00	2.65	0.86	0.62	0.67	-0.38
		43	5	1.00	1.00	0.00	0.00	0.00	#N/A
			Mean	2.00	1.56	0.32	0.33	0.36	0.09
			SE	0.58	0.40	0.15	0.11	0.12	0.26
C. spurium	**1**	**6**	8	1.00	1.00	0.00	0.00	0.00	#N/A
		11	8	2.00	2.00	1.00	0.50	0.53	-1.00
		19	8	1.00	1.00	0.00	0.00	0.00	#N/A
		36	8	1.00	1.00	0.00	0.00	0.00	#N/A
		42	8	2.00	2.00	1.00	0.50	0.53	-1.00
		43	8	1.00	1.00	0.00	0.00	0.00	#N/A
			Mean	1.33	1.33	0.33	0.17	0.18	-1.00
			SE	0.21	0.21	0.21	0.11	0.11	0.00
Across all loci and species			**Mean**	1.63	1.45	0.35	0.25	0.26	-0.39
			SE	0.17	0.12	0.07	0.05	0.05	0.10

277