Conservation genomics of an Australian cycad, cycas calcicola and the absence of key genotypes in botanic gardens

Conservation genomics of an Australian cycad, *Cycas calcicola* and the Absence of Key Genotypes in Botanic Gardens

James A. R. Clugston[1,2,3*], Markus Ruhsam[2*], Gregory J. Kenicer[2], Murray Henwood[4], Richard Milne[3], and Nathalie S. Nagalingum[5]

[1.] National Herbarium of New South Wales, Royal Botanic Gardens and Domain Trust, Mrs Macquaries Rd, Sydney NSW 2000, Australia

[2.] Royal Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh, Scotland, EH35LR, United Kingdom

[3.] The University of Edinburgh, College of Science and Engineering, School of Biological Sciences, Institute of Molecular Plant Sciences, The Kings Building, West Mains Road, Edinburgh, Scotland, EH9 3JN, United Kingdom

[4.] The University of Sydney, School of Life and Environmental Sciences, Heydon-Laurence Building A08, NSW, 2006, Australia

[5.] California Academy of Sciences, 55 Music Concourse Dr, San Francisco, CA 94118, USA

* These authors contributed equally

Corresponding authors: James A. R. Clugston james.clugston@rbgsyd.nsw.gov.au and Nathalie S. Nagalingum nnagalingum@calacademy.org

ORCID ID:

Markus Ruhsam 0000-0002-8457-345X

James A. R. Clugston 0000-0002-3653-6953

Nathalie S. Nagalingum 0000-0002-0221-9650

Richard Milne 0000-0001-7949-7539

Murray Henwood 0000-0002-7066-8191

Abstract

Understanding the genetic diversity of wild populations is fundamental to conserving species *in-situ* and *ex-situ*. To aid conservation plans and to inform *ex-situ* conservation, we examined the genetic diversity of the cycad *Cycas calcicola* (Cycadaceae). Samples were collected from wild populations in the Litchfield National Park and Katherine regions in the Northern Territory, Australia. Additional samples were obtained from botanic garden plants that were originally collected in the Katherine region, Daly River and Spirit Hills in the Northern Territory, Australia. Using RADseq we recovered 2271 informative genome-wide SNPs, revealing low to moderate levels of gene diversity (uH_e=0.037 to 0.135), very low levels of gene flow, and significant levels of inbreeding (mean F_{IS}=0.491). Population structure and multivariate analysis showed that populations fall into two genetic groups (Katherine vs Litchfield + Daly River + Spirit Hills). Genetic differentiation was twice as high between populations of the Katherine and Litchfield regions (F_{ST}~0.1) compared to within these two regions (F_{ST}~0.05). Increasing population fragmentation together with high levels of inbreeding and very little gene flow are concerning for the future adaptability of this species. The results indicated that the *ex-situ* collections (1) had significantly lower genetic diversity than the wild populations, and (2) only partly capture the genetic diversity present, because the Litchfield National Park populations are not represented. We recommend that *ex-situ* collections be expanded to incorporate the genetic diversity found in Litchfield National Park, and that *in-situ* populations from the Katherine and Greater Litchfield regions be conserved as separate management units.

Keywords: RADseq, next generation sequencing, population genetics, genomics, Cycadaceae, Cycadales, Cycas, *ex-situ* conservation, *in-situ* conservation.

Declarations and Data Accessibility Statement

Competing Interests

There are no competing interests within this manuscript by any listed authors

Funding information

1. Australian Flora Foundation - Research grant – Australia
2. The Australasian Systematic Botany Society and The Nature Conservatory–The Thomas Foundation for an Australian Conservation Taxonomy Award – Australia
3. Biotechnology and Biological Sciences Research Council (BBSRC) UK and the EASTBIO Doctoral Training Partnership – UK, Ref: 1429569

Authors' contributions

- James A. R. Clugston – planning, conducted fieldwork. research, lab work, bioinformatics, main writer.
- Markus Ruhsam – major contributor second writer, bioinformatics, statistics, manuscript structure.
- Gregory J. Kenicer – PhD supervisor, advice and manuscript feedback.
- Murray J. Henwood - PhD supervisor, manuscript editor, advice and input into research aims and outputs.
- Richard Milne - PhD supervisor, advice and manuscript feedback.
- Nathalie S. Nagalingum - PhD supervisor, obtained funding, conceived, planned and oversaw project, planned and conducted fieldwork, collected cultivated samples, manuscript editor, and provided feedback and overall guidance on manuscript.

Conflicts of interest/Competing interests

No conflicts of interest to disclose

Data Availability and material

RAW fastq sequinning files, assemblies and will be uploaded to NCBI GenBank after the

initial review process has been completed, to be accessible for publication.

Code availability

Not applicable

Animal research and ethics approval

Not applicable

Consent to participate (ethics)

Not applicable

Consent for publication (ethics)

Not applicable

Plant Reproducibility

Not applicable

Clinical Trails Registration

Not applicable

Gels and Blots/image manipulation

No gels or images were manipulated for this manuscript

Introduction

The risk of extinction for plant species is increasing worldwide due to habitat fragmentation, climate change, land clearance, competition with invasive species and, in some cases, over-collection (Vilà et al. 2011; Newbold et al. 2016). The conservation of native plant populations is, therefore, becoming ever more important to help preserve biodiversity (Hefley et al. 2016). Conservation genetics provides a framework to guide both conservation and restoration to help minimise the extinction risk of species (Frankham et al. 2004; Kramer & Havens 2009) with the aim of determining if populations contain enough genetic variation for adaptation, expansion, and re-establishment (Hedrick & Miller 1992; Paz-Vinas et al. 2018; Yoder et al. 2018). Conservation genetics has also informed many *in-situ* conservation plans by inferring the overall dynamics of populations, such as decreases in population size, past bottlenecks, and sex-specific gene flow (Ahrens et al. 2017; Zhang et al. 2018), and has been used to identify populations with high levels of genetic diversity as conservation priorities (Drury et al. 2017; Hou et al. 2018; Rodríguez-Rodríguez et al. 2018; Wu et al. 2020).

Cycads are at the highest risk of extinction of any plant group (Donaldson 2003). Their leaves, sap, and seeds are poisonous to livestock (Norstog & Nicholls 1997), which has led to the clearing of cycads from arable land in order to prevent accidental poisoning (Hall & McGavin 1968; Hall & Walter 2014). Cycads are also highly prized in horticulture, with some species being sold for thousands of US dollars (Donaldson 2003). The ornamental appeal of cycads has generated a great demand on the world market, which has led to over-collection and illegal removal from the wild (Pérez-Farrera et al. 2006; Torgersen 2017). Many cycad populations have declined in size (González-Astorga et al. 2008; Shuguang et al. 2006; Octavio-Aguilar et al. 2009; Da Silva et al. 2012; Cabrera-Toledo et al. 2012), with many

125 species surviving in small and fragmented populations with low genetic diversity (Long-Qian
126 et al. 2004; Meerow et al. 2012), especially in Africa (Ekué et al. 2008; Da Silva et al. 2012)
127 and North America (Cabrera-Toledo et al. 2010). However, genetic diversity in cycads is not
128 always correlated with population size; for example, small and isolated populations of *Cycas*
129 *multipinnata* C.J.Chen & S.Y.Yang were found to have high levels of genetic diversity (Gong
130 et al. 2014). This is likely to be due to the long generation times in cycads delaying the
131 genetic effects of inbreeding and bottleneck effects (González-Astorga et al. 2008; Cibrián-
132 Jaramillo et al. 2010; James et al. 2018).
133
134 *Ex-situ* living plant collections are a safeguard for species threatened with extinction, and
135 help to preserve their genetic diversity (Dosmann 2006; Cibrian-Jaramillo et al. 2013). The
136 *Ex-situ* conservation of plants is carried out either through the use of seed banks or by
137 growing plants in botanic gardens, either of which have the potential to replenish depleted
138 natural populations though reintroductions (Fant et al. 2016; Volis 2017). Of these, seed
139 banks have the advantage of being able to store a large number of individuals in a relatively
140 small space, making them more cost- and space-efficient in the long term (Hamilton 1994),
141 and giving them a higher probability of containing greater genetic diversity compared to living
142 plant collections (Schoen and Brown 2011). However, seed banks are not an option for
143 cycads, because their seed are recalcitrant and only viable for about one year, making long
144 term seed storage very challenging (Calonje et al. 2011; Mondoni et al. 2011; Nadarajan et
145 al. 2018). This means that living plant collections are at present the only option to conserve
146 cycads *ex-situ*. As the number of living individuals that can be held in *ex-situ* collections is
147 significantly lower compared to seeds, these individuals must be carefully selected based on
148 the genetic diversity of the species and the distribution of genetic diversity among
149 populations (Hurka 1994; Hoban et al. 2020). Additionally, although different species within a
150 genus may have similar traits, they may have different patterns of genetic diversity due
151 differences in population size or range. Therefore, they may require different collection sizes

to safeguard their genetic diversity in *ex-situ* collections (Hoban et al. 2020).

Cycas L. is the largest extant genus of Cycadales consisting of 117 currently recognised species (Calonje et al. 2020), and occurs throughout Madagascar, Asia to Indonesia, Australia and New Caledonia (Chaw et al. 2005). Australia has many large and undisturbed populations of *Cycas* (Liddle 2009), but little is known about their genetic diversity. Of the 38 *Cycas* species endemic to Australia, the only species represented by a conservation genetic study is *Cycas megacarpa* K.D.Hill (James et al. 2018). Populations of this species in Queensland (Australia) showed low to moderate levels of gene diversity, which did not correlate with population size. There was little genetic differentiation among populations over broad geographic regions, perhaps because historical geneflow was detected (James et al. 2018).

Cycas calcicola Maconochie is an endemic Australian cycad species (Figure 1) with populations that are believed to be largely undisturbed. The populations occur in four main areas, all within the Northern Territory in Australia: Daly River (>7000 plants), Spirit Hills (>5000), Litchfield National Park (>5000) and Katherine region (> 1500) [population sizes based on estimates from the Parks and Wildlife Commission, 1994, 1995 and 1996; cited in Liddle (2009)]. The most recent IUCN Red List conservation assessment for *C. calcicola* was carried out in 2010 and classified this species as Least Concern due to the size of the populations (IUCN 2019). However, there is evidence of recent population contraction which has caused disjunctions between some populations due to increased burning, and habitat clearing for farmland or roads. In particular the populations in the Katherine region show evidence of a decline in the number of individuals due to uncontrolled over-collection and fire damage (Liddle 2009). For these reasons the IUCN assessment of this species is in need of updating and it is likely that the conservation prospects for *C. calcicola* have deteriorated in the last ten years.

Here we aim to determine the levels of genetic diversity of *Cycas calcicola* populations throughout its distribution to answer the following questions: (1) How is genetic diversity distributed among populations? (2) how much gene flow is there between populations and regions? (3) is genetic diversity in wild *C. calcicola* populations captured by existing *ex-situ* collections?

Materials and methods

Study species. All known *Cycas calcicola* populations occur in large but disjunct populations in four main areas in the Northern Territory (mostly): (1) Litchfield region (includes all populations in the Litchfield National Park), (2) the Daly River, (3) Spirit Hills conservation site (on the border of Northern Territory and Western Australia) and (4) Katherine region (includes populations from Katherine and the surrounding area) (Figure 2) (Hill 1996; Jones 2002). This species usually occurs on or near limestone in open bush or rocky outcrops. *Cycas calcicola* has an arborescent trunk typically ≤ 5 m in height, and is easily distinguished from other Australian *Cycas* species by its dark green leaflets with recurved margins and leaflets covered in silvery-grey hairs (Hill 1996). Like other cycads, it is dioecious and likely to be insect pollinated (Kono & Tobe 2007) although the pollinators of this species have not yet been documented (Liddle 2009). Some *C. calcicola* populations occur in close proximity (within 10 km) to *Cycas armstrongii* Miq., which is known to be pollinated by two species of beetle in the Tenebrionidae (Ornduff 1992). Although *C. calcicola* is not known to hybridise with *C. armstrongii* it is likely that the species share pollinators due to similar phenological patterns (Liddle 2009). Seed dispersal distances of *Cycas armstrongii* growing in the Northern Territory are rarely greater than 3 m from a mother plant (Watkinson & Powell 1997), and it is likely that *C. calcicola* seeds disperse similar distances.

205

206 **Sampling strategy**. Silica-dried leaflets of *C. calcicola* were collected from wild populations
207 growing in Litchfield National Park and the Katherine region in the Northern Territory,
208 Australia (Table 1). Populations were selected based on the collection sites of herbarium
209 specimens recorded online using The Australasian Virtual Herbarium
210 (https://avh.chah.org.au), accessed 12th January 2015). A total of 60 individuals were
211 sampled from six populations: three populations from Litchfield National Park and three from
212 the Katherine region (Figure 2, Table 1). For each population, ten individuals were sampled
213 from plants of varying ages, but where possible bearing microsporangiate strobili or
214 megasporophylls. In addition, a further 13 samples were obtained from all known individuals
215 cultivated in *ex-situ* collections: George Brown Darwin Botanic Garden (Darwin, Northern
216 Territory, Australia) and Montgomery Botanical Center (Miami, Florida, USA). The *ex-situ*
217 botanic garden material represented plants of known wild origin from Katherine (n=7 from
218 type population Katherine TT), Daly River (n=4), and Spirit Hills (n=2) (Note: there was no
219 known individual in botanic garden collected from Litchfield National Park). In addition to the
220 tissue sampling, we also recorded basic population demographics for each sampled
221 population (Table 1).
222

223 **DNA extraction and quantification**. Approximately 0.05 g of silica-dried leaflets were
224 ground to a fine powder using a TissueLyser (Qiagen, Hilden, Germany). When present in
225 large amounts (common with *C. calcicola*), trichomes were removed using a wire brush to
226 improve DNA yield. High molecular weight genomic DNA was extracted using a Qiagen
227 DNeasy Plant DNA Extraction Mini Kit (Qiagen, Hilden, Germany). DNA extractions were
228 quantified using an Invitrogen Qubit fluorometer (3.0 BR DNA assay; Invitrogen, Life
229 Technologies, Carlsbad, CA, USA) with a target concentration of 17 µg/mL (enough to obtain
230 500 ng within 42 µL of solution); any sample that yielded less than 17 µg/mL was either re-
231 extracted or concentrated using a 1:1 ratio of Agencourt AMPure XP sample purification

beads (Beckman Coulter, Inc.) by combining multiple extractions from the same sample.

DNA normalisation and restriction digest reaction. For a full protocol, see Clugston et al. (2019). First, genomic DNA was normalised to a concentration of 500 ng in 42 µL total volume (0.01 µg/mL). Second, 5 µL of NEB 10x CutSmart buffer (New England Biolabs, Ipswich, MA) and 1 µL of Bovine Serum Albumin (BSA) was added to each well. Samples were then held at 4°C for a minimum of five hours before adding restriction enzymes—this five hour of incubation aided in the cutting action of the restriction enzymes. Next, double digest reactions were carried out using 1 µL each of the restriction enzymes EcoR1-HF and Mse1. Reactions were then placed into a thermocycler for three hours at 37°C with a final 20-minute enzyme deactivation step at 65°C. The reactions were checked on a 2% agarose gel for quality of digestion. Last, reactions were cleaned using 1.8:1 ratio of AMPure XP beads to sample (90 µL of AMPure XP beads to 50 µL of digested DNA) and quantified using a Qubit (3.0 HS DNA assay, Invitrogen, Life Technologies, Carlsbad, CA, USA).

Library preparation. Libraries were prepared using an Illumina TruSeq nano high-throughput dual index library preparation kit (Illumina Inc., CA, USA). We followed the ezRAD v3 modified protocol (Toonen et al. 2013) using half of the recommended volumes of the kit to save costs (Clugston et al. 2019). Following Clugston et al. (2019), the final steps of library preparation were modified from the ezRAD protocol by modifying the final bead clean with a 0.8:1 ratio of AMPure XP beads to remove adapter dimer. The libraries were validated using a LabChip, cleaned using a ratio of 0.9:1 AMPure XP beads, and quantified using a Qubit high sensitivity kit. Libraries were then normalised to a concentration of 10 nM, after which 5 µL of each library was pooled for sequencing.

Sequencing. Following Clugston et al. (2019), we aimed to capture ~1 GB of sequence data per sample (in a run of 95 libraries including 73 samples of *C. calcicola*). Our goal was to

obtain adequate coverage of the large genome of *C. calcicola* to ensure a good read depth to improve SNP calling accuracy during de-novo assembly. Sequencing was completed using an Illumina NextSeq 500, with 150 bp paired-end high throughput reagents kit (HT) on a single flow cell, spiked with 10% PhiX sequencing control V3.

Bioinformatics

Quality control and filtering of sequence reads. The NextSeq 500 generated eight raw fastq files for each sample: four forward files and four reverse files. For downstream analysis, the four forward files were combined into one single fastq file, and the four reverse files into another. Illumina reads were assessed for quality using FastQC 0.11.4 (Andrews et al. 2014). Trimmomatic 0.36 (Bolger et al. 2014) was used to filter reads according to their quality, remove Illumina adapter sequences and cut sites (the first six base pairs of reads), and then crop reads to 120 bp in length (because reads dropped in quality after 120 bp). A sliding window approach was used to remove low quality reads with a 'PhredQ score' less than 20, and all reads less than 50 bp were discarded.

Assembly of RADseq data. De novo assembly of the paired-end reads was performed with ipyrad 0.7.18 (Eaton & Overcast 2020) using a high-performance online instance with Amazon Web Services through the California Academy of Sciences. For ipyrad, various settings were tested following guidance from the ipyrad development team. In ipyrad most parameters were set to default, except that bases with a 'PhredQ score' less than 30 were converted to 'N', reads with 15 'uncalled bases' were discarded, and 'data type' was set to 'pairgbs'. Reads were further filtered for adapter sequences, adapters were trimmed, and reads were discarded if they were less than 40 bp after trimming.

The maximum number of uncalled bases in consensus sequences was set to 10 in both forward and reverse reads. The minimum depth for statistical base calling and majority rule

base calling were both set to ‘6’ and the setting for ‘maximum shared heterozygotes per locus’ was left at 0.5 (default) to reduce the effects of paralogs. The ‘maximum heterozygotes in consensus sequences’ were set at eight for both forward and reverse sequences, and the ‘minimum number of samples per locus’ was set to 36, so each SNP would be present across a minimum of 36 samples, which corresponded to at least 50% of the samples (one sample failed to meet quality threshold for assembly). This ensures effective population genotyping (Shafer et al. 2016). The maximum SNPs per locus was set to ‘20’ and the maximum number of indels per locus to 8 forward and 8 reverse reads.

Population genetic statistics. We used GenALEx 6.5 (Peakall & Smouse 2012) to estimate the number of alleles (N_a), the effective number of alleles per locus (N_e), observed heterozygosity (H_o), expected heterozygosity (H_e), and unbiased expected heterozygosity (uH_e=2n/(2n-1)*H_e); the latter has been shown to be a better estimator of gene diversity if sample numbers are small (Nei 1978). Level of genetic differentiation among and between populations was inferred using an analysis of molecular variance (AMOVA), and pairwise F_{ST} (the fixation index) values were calculated with 999 permutations and the ‘Codom-Allelic’ option selected, with data being portioned for nine populations and four regions (Table 1), also using GenALEx. F_{IS} (the inbreeding co-efficient) was calculated using the R package diveRsity 1.9.90 (Keenan et al. 2013).

Population structure analysis. STRUCTURE v.2.3.4 (Pritchard et al., 2000) was used to explore the genetic structure and identify for the most likely number of distinct genetic groups. STRUCTURE uses a Bayesian algorithm to cluster samples into *K* distinct genetic groups by minimizing deviations from Hardy–Weinberg and linkage equilibrium within each cluster. The analyses were carried out using only unliked markers (i.e., one SNP per RAD tag was randomly chosen for the analysis) for *K*=1–5 using 500,000 Markov chain Monte Carlo (MCMC) iterations after a burnin of 20,000 steps. Each analysis was repeated 10 times

for each value of *K*. If genetic clusters have widely different sample sizes (unbalanced sampling), then STRUCTURE has been shown to yield poor estimates of both individual ancestry and *K,* if the default settings are used (Wang, 2017). Therefore, we followed Wang's (2017) recommendation and selected the alternative option ('Separate α for each Population') allowing a separate α, which is a measure of the relative admixture level between populations

To identify the most likely number of distinct genetic groups (*K*)*,* two approaches were implemented using the software Kfinder2 (Wang 2019). First, we used the Δ*K* statistic (Evanno et al., 2005), which is based on the rate of change in the log probability of data between successive *K* values. Secondly, we employed the parsimony index, termed PI (Wang 2019), which aims to identify the number of populations (*K*) that consistently yield the minimal admixture estimates of sampled individuals. Additionally, to test for a correlation between the genetic ($F_{ST}/(1 - F_{ST})$) (Rousset 1997) and the log transformed geographic distance, a Mantel test was carried out using GenALEx 6.5.

To visualise the genetic relationships among populations, a Discriminant Analysis of Principal Components (DAPC) was carried out using adegenet 2.1.0 (Jombart & Ahmed 2011) in R (R Core Team, 2019). DAPC shows the number genetic clusters (groups) of samples by using a combination of linear variables (in this case alleles), which have the largest between-group and smallest within-group variance, and provides group membership probabilities for each individual in a population based on the number of retained discriminant functions in the DAPC (Jombart & Ahmed 2011). The optimal number of clusters in the data and the number of principal components (PCAs) to be retained for discriminate analysis were determined using the 'find.clusters' command in combination with the optimal *a*-score. A DAPC scatter plot was used to depict the genetic relationship between individuals.

Gene flow analysis. Gene flow was estimated with the software BA3-SNPS (Mussmann *et al.*, 2019), which is a modification of BayesAss 3.0.4 (Wilson and Rannala, 2003) that permits handling of large SNP datasets generated via methods such as RADseq. BayesAss uses individual multi-locus genotypes and a Bayesian Markov chain Monte Carlo (MCMC) approach to estimate the rates of recent immigration (over the last several generations) among populations. The BA3-SNPS method rests on fewer assumptions compared to other estimators of long-term gene flow like migrate-n (Beerli, 2006) and can be applied to non-stationary populations that are far from migration-drift and Hardy-Weinberg equilibrium (Wilson and Rannala, 2003). Migration rates (m), which are interpreted as the proportion of migrants per generation in one population that are derived from another population, are assumed to be low in BayesAss. MCMCs were run for 50 million generations, with a burnin of five million, sampling every 1000 generations. The mixing parameters for migration rates (m), allele frequencies (a) and inbreeding coefficients (f) were optimised using BA3-SNPs-autotune (m=0.2125, a=0.775, f=0.0625) to achieve the recommended acceptance rates between 0.35 and 0.45 (Mussmann *et al.*, 2019). Five independent replicates with different random starting seeds were carried out, assessing convergence of the combined and individual runs using Tracer v1.5 (Rambaut and Drummond, 2009). Gene flow was estimated among all six populations from the Litchfield and Katherine regions using the full dataset of 2271 SNPs.

Results

Sequencing and de-novo assembly. After filtering the raw data, the number of reads that remained per sample ranged from 1,296,034 to 4,650,176. From the original 73 samples sequenced only 72 samples were processed further as one sample from Katherine CDU2 did not pass quality control (uploaded to NCBI GenBank BioProject ID: PRJNA746394 see

Supplementary Table 1). De-novo assembly of the reads using ipyrad generated 1,296,034 to 3,037,283 (average cluster depth = 1552682.44) sequence clusters per sample, with 22,806 to 78,631 (average cluster depth = 42124.16) high depth clusters (defined as containing six or more reads). Considering only loci that were present in at least 50% of all individuals, the final output from ipyrad generated 2,271 SNPs (see Supplementary Table 1).

Population genetic statistics. Unbiased gene diversity (uH_e) ranged from uH_e=0.037 in Spirit Hills to uH_e=0.135 in Litchfield NP1, with a mean of uH_e=0.095 (Table 2). Observed heterozygosity (H_o) ranged from H_o=0.028 in both Spirit Hills and in Daly River (both *ex-situ* conservation populations) to H_o=0.059 in Litchfield NP1, with a mean of H_o=0.039. These results indicate low to moderate levels of diversity in *C. calcicola.* The inbreeding coefficient (F_{IS}) ranged from F_{IS}=-0.244 in Spirit Hills to F_{IS}=0.605 in Katherine CDU1 with an average across all populations of F_{IS}=0.409 (0.015 – 0.425, 95% CI) (Table 2). Levels of inbreeding were highest in the *ex-situ* conservation populations. The Mantel test revealed a significant correlation (R^2=0.42, P=0.000) between the genetic distance (FST /(1 – FST)) and the log transformed geographic distance (see Supplementary Figure 1).

Population differentiation. Analysis of molecular variance (AMOVA) showed low but significant (P=0.001) levels of differentiation among regions (Litchfield, Katherine, Daly River and Spirit Hills) (Phi_{ST}=6%), with an equal amount of genetic differentiation among populations (Phi_{ST}=6%, Table 3). Additionally, pairwise F_{ST} values at the regional level (Table 4) showed that genetic differentiation was greatest when comparing populations from Katherine with those from the Greater Litchfield region (F_{ST}~0.1), and this was about twice as high as within regions (F_{ST}~0.05). At the population level, genetic differentiation was highest among Spirit Hills and all other populations (F_{ST}~0.2) and lowest among populations within regions (F_{ST}~0.05, Table 4).

Population structure analysis. Both the Δ*K* statistic and parsimony index suggested that the most likely number of genetic groups was *K* = 2 (Δ*K* = 1833). Populations from the Litchfield and Katherine regions formed separate genetic clusters with little admixture (Figure 3). Spirit Hills and Daly River plants were mostly genetically closer to those from Litchfield National Park but showed some admixture with the Katherine cluster. Discriminant analysis of principal components (DAPC) equally resolved two (*K*=2) genetic groups, i. e. Litchfield + Spirit Hills + Daly River, and a second group containing Katherine populations (greater Katherine region) (Figure 4).

Gene flow analysis. The five independent BA3-SNPs analyses each yielded effective sample sizes (ESS) that were well above 40,000 for all pairwise migration estimates, indicating adequate sampling of the posterior distribution. All runs resulted in nearly identical very low migration rates which were not significantly different from zero (Table 5).

Discussion

In this study, we investigated levels and patterns of genetic diversity of *Cycas calcicola* using genomic data from RADseq, and assessed if *ex-situ* collections represent the genetic diversity of wild populations. Generally, we found low to moderate levels of genetic diversity in populations of *C. calcicola* and evidence of inbreeding, with genetic differentiation between populations being low, but greater between regions than between populations. We also found that although *C. calcicola* is represented in *ex-situ* botanic garden collections, essential genotypes were missing, and *ex-situ* collections do not represent the genetic diversity of the wild populations.

Genetic diversity. Our results indicated that *C. calcicola* had low levels of gene diversity in both wild (uH_e = 0.100 to 0.135, Table 2) and ex-situ populations (uH_e = 0.037 to 0.085,

Table 2). Comparatively, *Cycas megacarpa* K.D. Hill from Queensland, the only other Australian *Cycas* species whose genetic diversity has been investigated, has genetic diversity nearly three times higher than that of *C. calcicola* (mean H_e = 0.269 based on 12 nuclear microsatellite markers; James et al. 2018). Similarly, *Cycas simplicipinna* (Smitinand) K.D.Hill populations from Laos and China had a mean H_e = 0.447 based on 16 SSR markers (Feng et al. 2014). However, meaningful comparisons between studies using different genetic markers can be difficult to make (Peakall et al. 2003; Hodel et al. 2017; Sunde et al, 2020), and might be questionable even if the same type of genetic markers were used but did not screen homologous loci. In particular, gene diversity (H_e) estimates using microsatellites can be two to three times higher than estimates using SNPs (Fischer et al. 2017; Hodel et al, 2017; Lemopoulos et al. 2019, Sunde et al. 2020). For example, Zimmerman et al. (2020) found that estimates of H_o and H_e in the Gunnison sage-grouse (*Centrocercus minimus*) were two to three times higher using microsatellites compared to SNPs. This is probably due to the high number of alleles per microsatellite locus which increases H_e, i. e. the likelihood of drawing two random alleles from a population that are not identical by descent. Assuming that microsatellite H_e values are generally two to three times higher than SNP H_e values suggests that the levels of genetic diversity are probably comparable between *C. calcicola* and C. *megacarpa*/*C. simplicipinna*. A recent study on *Dioon merolae* De Luca, Sabato & Vázq. Torres in Mexico also using RADseq data (Gutiérrez-Ortega et al., 2020) found even lower levels of gene diversity (H_e = 0.027 to 0.076), indicating that low levels of gene diversity are present in other cycad species as well.

Many Australian cycad populations are considered to be large and healthy (Liddle 2009). Although *Cycas megacarpa* (occurring in Queensland, Australia) has sizeable populations (>250 or even >500 individuals), there are also many which have fewer than 50 individuals (James et al. 2018). Populations of *C. calcicola* show evidence of population contraction (Liddle 2009), perhaps due to the frequent occurrence of anthropogenic fires, which have

446 been a long-time feature of the Australian landscape (Andersen et al. 2005). It is likely that
447 population sizes have decreased through a range of anthropogenic activities since regional
448 population sizes of *C. calcicola* were estimated about 25 years ago by Hill (1996). For
449 example, based on herbarium records (AVH 2019) and our field-observations (Clugston and
450 Nagalingum, pers. obs.), populations of *C. calcicola* are likely to be far more fragmented than
451 has been assumed. During fieldwork we noted that populations known from herbarium
452 records in areas between the two major regions (Litchfield and Katherine) no longer exist.
453 Additionally, we found considerable variation in the number of individuals in each population.
454 Populations at the Charles Darwin University campus in Katherine (Katherine CDU1 and
455 CDU2) and Tolmer Falls (Litchfield NP Tolmer) in Litchfield National Park are sizeable (>200
456 individuals) and showed evidence of recent recruitment (Table 1), but other populations in
457 the Katherine (Katherine TT = >100 individuals) and Litchfield (Litchfield NP1 and Litchfield
458 NP2 = >25 individuals) regions were smaller and did not show evidence of recent
459 recruitment.
460
461 It is concerning that most of the sampled populations showed significant levels of inbreeding
462 (F_{IS} ranged between 0.456 and 0.605; Table 2). The exception to this was the *ex-situ*
463 collections from Sprit Hills with F_{IS}=-0.244. However, as only two individuals were available
464 from this population this is unlikely to be a reliable estimate. Cycad species are dioecious
465 and, therefore, obligate outbreeders, but biparental inbreeding (mating between close
466 relatives) seems to occur at an appreciable frequency according to our F_{IS} results. This is not
467 surprising given that (1) the recorded seed dispersal distances in *Cycas armstrongii* (and
468 other *Cycas* species) in the Northern Territory are rarely greater than 3 m (Watkinson &
469 Powell 1997), and (2) that only a few individuals seem to participate in any given
470 reproductive event for cycads (Vovides et al. 1997; Suinyuy et al. 2009; Terry et al. 2012) .
471 Significant levels of inbreeding have also been reported in other cycad species with values
472 ranging between 0.122 to 0.483 (Keppel et al. 2002; Cibrián-Jaramillo et al., 2010).

473 Inbreeding leads to a reduction in genetic fitness (inbreeding depression), which is common
474 in many angiosperms (Charlesworth & Charlesworth 1987; Mahy & Jacquemart 1999; Vogl
475 et al. 2002; Bellusci et al. 2009; Ruhsam et al. 2010; Sletvold et al. 2013) and gymnosperms
476 (Kärkkäinen and Savolainen 1993; Durel et al. 1996; Williams & Savolainen 1996).
477 Furthermore, dioecy is likely to maintain lethal inbreeding factors in a species resulting in
478 their slow elimination from the gene pool (Willi et al., 2006). The low recruitment in some *C.*
479 *calcicola* populations could, therefore, be due to the effects of early inbreeding depression
480 manifesting in poor germination rates and seedling viability. Unfortunately, information on
481 inbreeding depression is not available for *C. calcicola*.
482
483 **Population differentiation**. The small and fragmented populations of many cycad species
484 (Zheng et al. 2017), has resulted in reduced gene flow and high levels of genetic
485 differentiation between populations (Keppel et al. 2002; Meerow & Nakamura 2007; Keppel
486 et al. 2008; Cibrián-Jaramillo et al. 2010; Meerow et al. 2012; Calonje 2013). This is
487 consistent with limited seed-dispersal, as there are few seed dispersal agents for cycads,
488 and seeds rarely disperse more than 3 m (Watkinson & Powell 1997; Hall and Walter 2013).
489 The same applies to pollen-dispersal, as cycad pollinators rarely travel distances greater
490 than 100 m (Norstog & Fawcett 1989). The results from our gene flow analysis indicated that
491 there has been very little, or possibly no recent gene flow among populations (Table 5),
492 which is likely due to the geographic distance between some populations (Figure 2). We also
493 found a significant correlation between genetic and geographic distance, indicating that
494 populations that are geographically closer are genetically more similar, which is to be
495 expected given the high genetic differentiation among populations and substantial biparental
496 inbreeding within them.
497
498 The genetic structure of the populations indicates the existence of two distinct genetic
499 groups, namely Katherine (Katherine CDU1 + Katherine CDU2 + Katherine TT + Katherine

CUL) and 'Greater Litchfield' (Litchfield NP1 + Litchfield NP2 + Litchfield NP Tolmer + Daly River + Spirit Hills) (Figures 3 and 4). Although pairwise F_{ST} values (Table 4) were twice as high between regions (Katherine and Greater Litchfield, F_{ST}~0.1) compared to within regions (F_{ST}~0.05), the F_{ST} values seem surprisingly low. Higher differentiation was detected between Spirit Hills and all other populations (F_{ST}~0.2), as well as between Daly River and most other populations (F_{ST}~0.15). However, this is based on only a few assayed individuals from Spirit Hills (n=2) and Daly River (n=4), which are held in *ex-situ* collections. Although it has been shown that RADseq using large number of SNPs (>1500) accurately captures the genetic diversity of populations if only three to eight individuals per population are screened (Qu et al. 2019, Nazareno et al. 2017). The average sample size (number of individuals) per locus for Spirit Hills (n=0.89) and Daly River (n=1.81) is below this number due to missing data, which indicates that that the results need to be interpreted with caution.

Why do we not see higher levels of differentiation between populations? One answer to this question is probably the recency (~100 years) of fragmentation events (Mankga & Yessoufou 2017), the long generation times and the low mutation rates in cycads (Chiang et al. 2009; Mankga & Yessoufou 2017). For example, in some species of South African *Encephalartos* Lehm.*,* the minimum generation time is about 60 years (Da Silva et al. 2012). If this is also the case for *C. calcicola,* then no more than two to three generations would have passed since fragmentation has had an impact on the population dynamics of this species. As a comparison, populations of *Dioon merolae* De Luca, Sabato & Vázq. Torres 1981 exhibited similar levels of geographic disjunction to that found in *C. calcicola*, but with greater levels of differentiation among populations (F_{ST} = 0.184 to 0.647). As cycads can live hundreds of years (Norstog and Nichols 1997), this means that our assayed individuals (mostly adults, but also 4-5 juveniles) are likely to show the genetic signature of a time when populations of *C. calcicola* were much less fragmented. To assess whether fragmentation has an impact on the genetic diversity of a species, comparisons between cohorts of adults and seedlings are

usually carried out. For example, analyses of fragmented *Primula vulgaris* Huds. populations indicated that seedlings had significantly lower genetic diversity (H_e seedlings = 0.436 vs. H_e adults = 0.535) and showed higher genetic differentiation between populations (F_{ST} seedlings = 0.136 vs. F_{ST} adults = 0.060) than mature plants (Van Geert et al. 2008). Similarly, in populations of *Myrtus communis* L.*,* seedlings showed lower genetic diversity when compared to mature plants due to population fragmentation (González-Varo et al. 2010). However, the long generation time and longevity of *C. calcicola* could mean that an effect of recent fragmentation on parameters such as H_e and F_{ST} might be difficult to detect even if seedling-adult comparisons are carried out (Kettle et al 2007).

Do *ex-situ* collections represent wild diversity? *Ex-situ* living plant collections of botanic gardens are critical in the conservation of species, as they can directly help to conserve the genetic diversity of natural populations and safeguard a species from extinction (Fant et al. 2016). However, a recent study highlighted that the percentage of extant genetic diversity conserved *ex-situ* varied between 40% to 95% for the 11 surveyed taxa (Hoban et al 2020), indicating that some ex-situ collections may not be sufficient to preserve the total genetic diversity. We screened all 13 *C. calcicola* individuals that are currently held within *ex-situ* collections to establish whether these collections capture the genetic diversity present in the wild. These individuals were from Katherine, Daly River, and Spirit Hills (Table 1). Based on our STRUCTURE analysis (Figure 3), all assayed populations of *C. calcicola* belong to one of two genetic groups— one group comprising populations of the Katherine region and the other comprising populations from Greater Litchfield National Park region, Daly River and Spirit Hills. However, the DAPC analysis (Figure 4) further indicated that the Litchfield National Park populations form a distinct cluster separate from those of Daly River and Spirit Hills. As *ex-situ* collections do not contain individuals from any of the Litchfield National Park populations, a large and unique part of the genetic diversity of *C. calcicola* is currently not conserved. Litchfield National Park is a major stronghold for *C. calcicola* populations,

554 accounting for at least 5000 individuals of this species (Liddle 2009), and the absence of this
555 region from *ex-situ* collections represents a significant conservation gap.
556
557 *Ex-situ* botanic garden collections often represent only a subset of the genetic diversity found
558 in wild populations (Li et al. 2002; Namoff et al. 2010; Cibrian-Jaramillo et al. 2013; Griffith et
559 al. 2015; Hoban et al. 2020), which is also true for *C. calcicola*. Our results showed that gene
560 diversity was lower in *ex-situ* collections compared to wild populations of *C. calcicola* (Table
561 2). Although no wild populations were assessed from Daly River or Spirit Hills these results
562 need to be interpreted with caution due to the small sample size in the *ex-situ* collection.
563 Many of the cultivated samples originally collected as seeds and young plants (mostly the
564 Montgomery Botanical Center samples) are from the Katherine region (CUL) and showed
565 only slightly (albeit significantly) lower levels of gene diversity (uH_e=0.085), when compared
566 to each of the wild populations from the region (uH_e=0.10 to 0.114, n=10), suggesting that
567 genetic diversity is captured *ex-situ* for the Katherine Region. This is encouraging and is
568 likely that augmenting *ex-situ* collections with perhaps just three or more individuals from
569 Katherine would raise diversity levels of the *ex-situ* collection to levels comparable with that
570 of assayed wild populations. In contrast, genetic diversity among the four *ex-situ* Daly River
571 (uH_e=0.061) collections was half as much as wild Katherine and Litchfield populations, and
572 that for Spirit Hills (uH_e=0.037, n=2) was even lower. However, due to missing data, the
573 average number of samples screened per locus was only 1.8 in Daly River and 0.9 in Spirit
574 Hills (Table 2), which might not be an accurate estimate of the genetic diversity despite the
575 large number of SNPs (2271) used. For logistical reasons we were unable to collect samples
576 from Daly River or Spirit Hills, but the most recent estimates suggest that they contain at
577 least as many individuals (>7000 and >5000, respectively) as Litchfield (>5000) or Katherine
578 (>1500) (Liddle 2009), hence, they could contain at least as much genetic diversity. It is clear
579 that two or four *ex-situ* individuals from the largest known *C. calcicola* populations are very
580 unlikely to capture a large part of the genetic variation found in the wild. Given the genetic

distinctiveness of Daly River (less so with Spirit Hills) populations in the DAPC (Figure 4), we suggest that *ex-situ* collections be expanded to incorporate more representatives from these regions.

Differentiation of *C. calcicola* populations between regions was at least twice as high as within regions, so it is important to include individuals from multiple populations from each region in ex-situ collections. From a conservation perspective we recommend that each geographic region should be regarded as a separate conservation management unit, with *ex-situ* collections consisting of around ten well-spaced individuals from each region. Given the short average dispersal distances of cycad pollen and seeds (Hall & Walter 2013), this would maximise the chance of collecting genetically diverse and unrelated individuals. Offspring produced by cross-pollinating these more or less unrelated individuals would have a lower risk of inbreeding depression and therefore increase the chance of successful reintroductions into the wild (Cohen et al. 1991). Without insight into the genetic diversity of *ex-situ* collections, inbreeding depression due to a narrow genetic base could become a problem among *ex-situ* collections, and any wild populations subsequently established from reintroductions (Enßlin et al. 2011). An added benefit of augmenting *ex-situ* collections is the greater number of specimens that will be accessible for scientific and horticultural research, ultimately aiding the study of physiological and reproductive factors that may have contributed to the decline of species like *C. calcicola* in the wild (Chen et al. 2012).

Although botanic gardens represent safe sites for holding the genetic reserves of wild populations, conserving the species in botanic gardens does not address the processes that have affected the genetic diversity of *C. calcicola* in the first place. Our field observations and the results of this study suggest that the size of populations have been on the decline in recent years, which is not yet reflected on a genomic level. This means that the remaining populations are at risk of an increased loss of genetic diversity in the future and therefore

conservation plans need to factor in both *in-situ* and *ex-situ* reserves to ensure the survival of this species.

Conclusions

Here we have provided new insights into the genetic diversity of the cycad, *Cycas calcicola*. By screening samples from both *in-situ* wild populations and *ex-situ* botanic garden collections, our results suggested low to moderate levels of genetic diversity, with little recent gene flow and high levels of inbreeding in most populations. The results from this study are pertinent for the formulation and implementation of two key conservation strategies: (1) populations from Litchfield National Park, Daly River and Spirit Hills form a genetically distinct group and should be managed as a conservation unit separate from those of the Katherine Region; and (2) plants from Litchfield National Park are genetically differentiation from other regions and are absent from *ex-situ* collections. Consequently, we recommend that priority be given to the acquisition of genetically representative material from this region to aid in the future conservation of the species. Additionally, our results indicate that low genetic diversity could relate to reduced population size and fragmentation, which highlights the importance in understanding generic diversity of threatened and rare species in conservation management assessments. This work demonstrates that ad hoc collections may not successfully capture genetic diversity, and, furthermore, genomic analysis should be considered when developing conservation plans for Australian cycad species.

Acknowledgements

We wish to acknowledge funding received from the Australian Flora Foundation, the Australasian Systematic Botany Society, and The Nature Conservatory - The Thomas Foundation for an Australian Conservation Taxonomy Award. Support from the Biotechnology and Biological Sciences Research Council (BBSRC) UK and the EASTBIO

635 Doctoral Training Partnership studentship at the University of Edinburgh is greatly
636 acknowledged. The Royal Botanic Garden Edinburgh is supported by the Scottish
637 Government's Rural and Environment Science and Analytical Services Division. We wish to
638 thank Lichfield National Park and Charles Darwin University for allowing us to collect
639 samples. Joe Perner of Cycad International for helping us locate population of *C. calcicola*
640 around the Katherine region.
641
642 The staff at the Royal Botanic Gardens and Domain Trust, in particular, Carolyn Connolly is
643 thanked her support in the molecular laboratory, Hannah McPherson for helping with the
644 initial quality testing of data. We also thank the California Academy of Sciences; Joe
645 Russack for his support and assistance in data assembly and Athena Lam and Boni Cruz for
646 their support in the molecular laboratory. The assistance of various botanic gardens is
647 acknowledged for providing *C. calcicola* leaf tissue samples, including Patrick Griffith and
648 Michael Calonje of the Montgomery Botanical Center, Florida, USA and George Brown
649 Darwin Botanic Garden Darwin, Northern Territory, Australia. We wish to thank, Alan Meerow
650 of the United States Department of Agriculture for his support and advice regarding
651 population genetics of cycads.
652

653 **References**

654 Aguilar, R. et al., (2019). Habitat fragmentation reduces plant progeny quality: a global
655 synthesis. J. M. Gomez, ed. *Ecology letters*, 22(7), 1163–1173.

656 Ahrens, C.W. et al., (2017). Genomic diversity guides conservation strategies among rare
657 terrestrial orchid species when taxonomy remains uncertain. *Annals of Botany*, 119(8),
658 1267–1277.

659 Álvarez-Yépiz, J. C., Dovčiak, M., & Búrquez, A. (2011). Persistence of a rare ancient
660 cycad: effects of environment and demography. *Biological Conservation*, 144(1), 122-130.

661

662 Amos, W. & Balmford, A., (2001). When does conservation genetics matter? Heredity,
663 87(3), 257–265.

664

665 Andersen, A. N., Cook, G. D., Corbett, L. K., Douglas, M. M., Eager, R. W., Russell-
666 Smith, J. & Woinarski, J. C. (2005). Fire frequency and biodiversity conservation in
667 Australian tropical savannas: implications from the Kapalga fire experiment. *Austral*
668 *ecology*, 30(2), 155-167.

669

670 Andrews, K.R. et al., (2016). Harnessing the power of RADseq for ecological and
671 evolutionary genomics. *Nature Reviews. Genetics*, 17(2), 81–92.

672

673 Andrews, K.R. et al., (2014). Trade-offs and utility of alternative RADseq methods:
674 Reply to Puritz et al. *Molecular Ecology*, 23(24), 5943–5946.

675

676 Beerli, P. (2006). Comparison of Bayesian and maximum-likelihood inference of
677 population genetic parameters. *Bioinformatics*, 22(3), 341-345.

678

679 Bellusci, F., Pellegrino, G., & Musacchio, A. (2009). Different levels of inbreeding
680 depression between outcrossing and selfing *Serapias* species. *Biologia Plantarum*,
681 53(1), 175-178.

682

683 Bolger, A.M., Lohse, M. & Usadel, B., (2014). Trimmomatic: a flexible trimmer for
684 Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120.

685

686 Cabrera-Toledo, D. et al., (2010). Allozyme diversity levels in two congeneric *Dioon* spp.

687 (Zamiaceae, Cycadales) with contrasting rarities. *Plant Systematics & Evolution*, 290(1-
688 4), 115–125.
689
690 Cabrera-Toledo, D., González-Astorga, J. & Flores-Vázquez, J.C., (2012). Fine-scale
691 spatial genetic structure in two Mexican cycad species *Dioon caputoi* and *Dioon merolae*
692 (Zamiaceae, Cycadales): Implications for conservation. *Biochemical Systematics &*
693 *Ecology*, 40, 43–48.
694
695 Calonje M, Stevenson DW, Stanberg L. The World List of Cycads, online edition
696 [Internet]. 2013-2020. [cited 2020 August 28]. Available from: http://www.cycadlist.org.
697
698 Calonje, M., Meerow, A.W., Knowles, L., Knowles D., (2013). Cycad biodiversity in
699 the Bahamas Archipelago and conservation genetics of the threatened *Zamia*
700 *lucayana* (Zamiaceae). *Oryx*, 47(2), 190–198.
701
702 Calonje, M., Kay, J. & Griffith, P.M., (2011). Propagation of Cycad Collections from Seed:
703 Applied Reproductive Biology for Conservation. *Sibbaldia: the Journal of Botanic Garden*
704 *Horticulture*, 9, 79–96.
705
706 Cantrill, D. J. (2018). The Australasian Virtual Herbarium: Tracking data usage and
707 benefits for biological collections. *Applications in plant sciences*, 6(2), e1026. Available at:
708 https://avh.chah.org.au (accessed 12th January 2015).
709
710 Charlesworth, D. (2003). Effects of inbreeding on the genetic diversity of populations.
711 Philosophical Transactions of the Royal Society of London. Series B: *Biological Sciences*,
712 358(1434), 1051-1070.

Charlesworth, D., & Charlesworth, B. (1987). Inbreeding depression and its evolutionary consequences. *Annual Review of Ecology and Systematics*, 18(1), 237-268.

Charlesworth, D. & Willis, J.H., (2009). The genetics of inbreeding depression. *Nature Reviews. Genetics*, 10(11), 783–796.

Chaw, S.-M. et al., 2005. A phylogeny of cycads (Cycadales) inferred from chloroplast matK gene, trnK intron, and nuclear rDNA ITS region. *Molecular Phylogenetics and Evolution*, 37(1), pp.214–234.

Chen, C., Mitchell, S. E., Elshire, R. J., Buckler, E. S., & El-Kassaby, Y. A. (2013). Mining conifers' mega-genome using rapid and efficient multiplexed high-throughput genotyping-by-sequencing (GBS) SNP discovery platform. *Tree Genetics & Genomes*, 9(6), 1537-1544.

Chen, S. Y., Zhang, Y. J., Wang, X. L., Sun, J. Y., Xue, Y., Zhang, P., & Qu, L. H. (2012). Extremely low genetic diversity indicating the endangered status of *Ranodon sibiricus* (Amphibia: Caudata) and implications for phylogeography. *PLoS One*, 7(3), e33378.

Chiang, Y. C., Hung, K. H., Moore, S. J., Ge, X. J., Huang, S., Hsu, T. W., ... & Chiang, T. Y. (2009). Paraphyly of organelle DNAs in *Cycas* Sect. Asiorientales due to ancient ancestral polymorphisms. *BMC Evolutionary Biology*, 9(1), 161.

Chybicki, I. J., & Oleksa, A. (2018). Seed and pollen gene dispersal in *Taxus baccata*, a dioecious conifer in the face of strong population fragmentation. *Annals of Botany*, 122(3), 409-421.

740

741 Cibrian-Jaramillo, A., Hird, A., Oleas, N., Ma, H., Meerow, A. W., Francisco-Ortega, J.,
742 & Griffith, M. P. (2013). What is the conservation value of a plant in a botanic garden?
743 Using indicators to improve management of ex situ collections. *The Botanical Review*,
744 79(4), 559-577.

745

746 Cibrián-Jaramillo, A., Daly, A. C., Brenner, E., Desalle, R., & Marler, T. E. (2010). When
747 North and South don't mix: genetic connectivity of a recently endangered oceanic
748 cycad, *Cycas micronesica*, in Guam using EST-microsatellites. *Molecular ecology*,
749 19(12), 2364-2379.

750

751 Clark, D.A. & Clark, D.B., (1987). Temporal and environmental patterns of reproduction in
752 *Zamia skinneri*, a tropical rain forest in cycad. *Journal of Ecology*, 75(1), 135.

753

754 Clugston, J. A. R., R. R. Milne, G. J. Kenicer, I. Overcast, N. S. Nagalingum. (2019).
755 RADseq as a valuable tool for plants with large genomes—a case study in cycads.
756 *Molecular Ecology Resources*. 19, 1610-1622.

757

758 Coates, D.J. & Atkins, K.A., (2001). Priority setting and the conservation of Western
759 Australia's diverse and highly endemic flora. *Biological conservation*, 97(2), 251–263.

760

761 Cohen, J.I. et al., (1991_. Ex Situ Conservation of Plant Genetic Resources: Global
762 Development and Environmental Concerns. *Science*, 253(5022), 866–872.

763

764 Crisp, M.D. & Cook, L.G., (2011). Cenozoic extinctions account for the low diversity of
765 extant gymnosperms compared with angiosperms. *The New Phytologist*, 192(4), 997–
766 1009.

Cousins, S.R., Williams, V.L. & Witkowski, E.T., (2011). Quantifying the Trade in Cycads (*Encephalartos* Species) in the Traditional Medicine Markets of Johannesburg and Durban, South Africa1. *Economic Botany*, 65(4), 356– 370.

De La Torre, A. R., Wilhite, B., & Neale, D. B. (2019). Environmental genome-wide association reveals climate adaptation is shaped by subtle to moderate allele frequency shifts in loblolly pine. *Genome biology and evolution*, 11(10), 2976-2989.

da Silva Carvalho, C., Ribeiro, M. C., Côrtes, M. C., Galetti, M., & Collevatti, R. G., (2015). Contemporary and historic factors influence differently genetic differentiation and diversity in a tropical palm. *Heredity*, 115(3), 216-224.

Da Silva, J. M., Donaldson, J. S., Reeves, G., & Hedderson, T. A., (2012). Population genetics and conservation of critically small cycad populations: a case study of the Albany Cycad, *Encephalartos latifrons* (Lehmann). *Biological Journal of the Linnean Society*, 105(2), 293-308.

Dehgan, B., & Yuen, C. K. K. H., (1983). Seed morphology in relation to dispersal, evolution, and propagation of *Cycas* L. *Botanical Gazette*, 144(3), 412-418.

De Vere, N., Jongejans, E., Plowman, A., & Williams, E., (2009). Population size and habitat quality affect genetic diversity and fitness in the clonal herb *Cirsium dissectum*. *Oecologia*, 159(1), 59-68.

Donaldson, J.S., (2003). *Cycads: status survey and conservation action plan*, IUCN-- the World Conservation Union.

794

795 Dosmann, M.S., (2006). Research in the Garden: Averting the Collections Crisis. *The*
796 *Botanical Review*, 72(3), 207–234.

797

798 Drury, C., Schopmeyer, S., Goergen, E., Bartels, E., Nedimyer, K., Johnson, M., & Lirman,
799 D., (2017). Genomic patterns in *Acropora cervicornis* show extensive population structure
800 and variable genetic diversity. *Ecology and evolution*, 7(16), 6188-6200.

801

802 Durel, C. E., Bertin, P., & Kremer, A., (1996). Relationship between inbreeding depression
803 and inbreeding coefficient in maritime pine (*Pinus pinaster*). *Theoretical and Applied*
804 *Genetics*, 92(3-4), 347-356.

805

806 Eaton, D. A., & Overcast, I. (2020). ipyrad: Interactive assembly and analysis of RADseq
807 datasets. *Bioinformatics*, 36(8), 2592-2594.

808

809 Edwards, R. D., Crisp, M. D., Cook, D. H., & Cook, L. G., (2017). Congruent
810 biogeographical disjunctions at a continent-wide scale: Quantifying and clarifying the role
811 of biogeographic barriers in the Australian tropics. *PLoS One*, 12(4).

812

813 Ekué, M. R., Gailing, O., Hölscher, D., Sinsin, B., & Finkeldey, R., (2008). Population
814 genetics of the cycad *Encephalartos barteri* ssp. *barteri* (Zamiaceae) in Benin with notes
815 on leaflet morphology and implications for conservation. *Belgian Journal of Botany*,
816 141(1), 78-94.

817

818 Ellstrand, N.C., (1993). Population Genetic Consequences of Small Population Size:
819 Implications for Plant Conservation. *Annual review of ecology and systematics*, 24(1),
820 217–242.

Ellstrand, N.C., Ornduff, R. & Clegg, J.M., (1990). Genetic Structure of the Australian Cycad, *Macrozamia Communis* (Zamiaceae). (5), 677.

Enßlin, A., Sandner, T. M., & Matthies, D. (2011). Consequences of ex situ cultivation of plants: Genetic diversity, fitness and adaptation of the monocarpic *Cynoglossum officinale* L. in botanic gardens. *Biological Conservation*, 144(1), 272-278.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14(8), 2611-2620.

Fant, J. B., Havens, K., Kramer, A. T., Walsh, S. K., Callicrate, T., Lacy, R. C., & Smith, P. P., (2016). What to do when we can't bank on seeds: What botanic gardens can learn from the zoo community about conserving plants in living collections. *Am. J. Bot*, 103, 1541-1543.

Fay, M.F., (2018). Orchid conservation: how can we meet the challenges in the twenty-first century? *Botanical studies*, 59(1), 16.

Feng, X., Wang, Y. & Gong, X., (2014). Genetic diversity, genetic structure and demographic history of *Cycas simplicipinna* (Cycadaceae) assessed by DNA sequences and SSR markers. *BMC Plant Biology*, 14(1), 187–187.

Finlay, C. M., Bradley, C. R., Preston, S. J., & Provan, J., (2017). Low genetic diversity and potential inbreeding in an isolated population of alder buckthorn (*Frangula alnus*) following a founder effect. *Scientific Reports*, 7(1), 1-8.

Fischer, M. C., Rellstab, C., Leuzinger, M., Roumet, M., Gugerli, F., Shimizu, K. K., & Widmer, A., (2017). Estimating genomic diversity and population differentiation–an empirical comparison of microsatellite and SNP variation in *Arabidopsis halleri*. *BMC Genomics*, 18(1), 69.

Fischer, M., Hock, M. & Paschke, M., (2003). Low genetic variation reduces cross-compatibility and offspring fitness in populations of a narrow endemic plant with a self-incompatibility system. *Conservation Genetics*, 4(3), 325–336.

Frankham, R., (2003). Genetics and conservation biology. *Comptes Rendus Biologies*, 326, 22–29.

Frankham, R., Ballou, J.D. & Briscoe, D.A., (2004). A primer of conservation genetics., *Comptes Rendus Biologies*, 326, 22–29.

Furlan, E. et al., (2012). Small population size and extremely low levels of genetic diversity in island populations of the platypus, *Ornithorhynchus anatinus*. *Ecology and evolution*, 2(4), 844–857.

Gargiulo, R., Saubin, M., Rizzuto, G., West, B., Fay, M. F., Kallow, S., & Trivedi, C., (2019). Genetic diversity in British populations of *Taxus baccata* L.: Is the seedbank collection representative of the genetic variation in the wild? *Biological conservation*, 233, 289-297.

Gonzalez-Astorga, J., Vergara-Silva, F., Vovides, A. P., Nicolalde-Morejon, F., Cabrera-Toledo, D., & Perez-Farrera, M. A., (2008). Diversity and genetic structure of three

875 species of *Dioon* Lindl. (Zamiaceae, Cycadales) from the Pacific seaboard of Mexico.
876 *Biological Journal of the Linnean Society*, 94(4), 765-776.
877
878 González-Astorga, J., Vovides, A.P., et al., (2008). Diversity and genetic structure of the
879 endangered cycad *Dioon sonorense* (Zamiaceae) from Sonora, Mexico: Evolutionary and
880 conservation implications. *Biochemical Systematics and Ecology*, 36(12), 891–899.
881
882 González-Varo, J. P., Albaladejo, R. G., Aparicio, A., & Arroyo, J., (2010). Linking genetic
883 diversity, mating patterns and progeny performance in fragmented populations of a
884 Mediterranean shrub. *Journal of Applied Ecology*, 47(6), 1242-1252.
885
886 Gutiérrez-Ortega, J. S., Salinas-Rodríguez, M. M., Ito, T., Pérez-Farrera, M. A., Vovides,
887 A. P., Martínez, J. F., .& Murakami, M. (2020). Niche conservatism promotes speciation in
888 cycads: the case of *Dioon merolae* (Zamiaceae) in Mexico. *New Phytologist*, 227(6),
889 1872-1884.
890
891 Griffith, M. P., Calonje, M., Meerow, A. W., Tut, F., Kramer, A. T., Hird, A., & Husby, C. E.,
892 (2015). Can a botanic garden cycad collection capture the genetic diversity in a wild
893 population?. *International Journal of Plant Sciences*, 176(1), 1-10.
894
895 Griffith, M.P., Magellan, T.M. & Tomlinson, P.B., (2014). Variation in Leaflet Structure in
896 *Cycas* (Cycadales: Cycadaceae): Does Anatomy Follow Phylogeny and Geography?
897 *International Journal of Plant Sciences*, 175(2), 241–255.
898
899 Griffiths, K. E., Balding, S. T., Dickie, J. B., Lewis, G. P., Pearce, T. R., & Grenyer, R.,
900 (2015). Maximizing the phylogenetic diversity of seed banks. *Conservation Biology*, 29(2),
901 370-381.

902

903 Griffith, M. P., Calonje, M. A., Stevenson, D. W., Husby, C. E., & Little, D. P., (2012). Time,
904 place, and relationships: cycad phenology in a phylogenetic and biogeographic context.
905 *Memoirs of the New York Botanical Garden*, 106, 59-81.
906
907 Hall, J.A. & Walter, G.H., (2014). Relative Seed and Fruit Toxicity of the Australian
908 Cycads *Macrozamia miquelii* and *Cycas ophiolitica*: Further Evidence for a Megafaunal
909 Seed Dispersal Syndrome in Cycads, and Its Possible Antiquity. *Journal of Chemical*
910 *Ecology*, 40(8), 860–868.
911
912 Hall, J. A., & Walter, G. H., (2013). Seed dispersal of the Australian cycad *Macrozamia*
913 *miquelii* (Zamiaceae): Are cycads megafauna-dispersed “grove forming” plants?.
914 *American Journal of Botany*, 100(6), 1127-1136.
915
916 Hall, P., Walker, S., & Bawa, K., (1996). Effect of forest fragmentation on genetic
917 diversity and mating system in a tropical tree, *Pithecellobium elegans*. *Conservation*
918 *Biology*, 10(3), 757-768.
919
920 Hall, W.T. & McGavin, M.D., (1968). Clinical and neuropathological changes in cattle
921 eating the leaves of *Macrozamia lucida* or *Bowenia serrulata* (Family Zamiaceae).
922 *Pathologia veterinaria*, 5(1), 26–34.
923
924 Hamilton, M.B., (1994). Ex Situ Conservation of Wild Plant Species: Time to Reassess the
925 Genetic Assumptions and Implications of Seed Banks. *Conservation Biology*, 8(1), 39–49.
926
927 Hedrick, P.W. & Miller, P.S., (1992). Conservation Genetics: Techniques and

928 Fundamentals. *Ecological Applications*, 2(1), 30–46.

929

930 Hefley, T. J., Hooten, M. B., Drake, J. M., Russell, R. E., & Walsh, D. P., (2016). When
931 can the cause of a population decline be determined?. *Ecology letters*, 19(11), 1353-1362.

932

933 Hensen, I., & Oberprieler, C., (2005). Effects of population size on genetic diversity and
934 seed production in the rare *Dictamnus albus* (Rutaceae) in central Germany. *Conservation*
935 *Genetics*, 6(1), 63-73.

936

937 Hill, K.D., (1996). A taxonomic revision of the genus *Cycas* (Cycadaceae) in Australia.
938 *Telopea*, 7(1), 1–64.

939

940 Hill, K.D., (1994). The *Cycas* rumphii complex (Cycadaceae) in New Guinea and the
941 western Pacific. *Australian Systematic Botany*, 7(6), 543–567.

942

943 Hughes, A. R., Inouye, B. D., Johnson, M. T., Underwood, N., & Vellend, M., (2008).
944 Ecological consequences of genetic diversity. *Ecology letters*, 11(6), 609-623.

945

946 Hoban, S., Callicrate, T., Clark, J., Deans, S., Dosmann, M., Fant, J., Gailing, O., Havens,
947 K., Hipp, A.L., Kadav, P. and Kramer, A.T., (2020). Taxonomic similarity does not predict
948 necessary sample size for ex situ conservation: a comparison among five genera.
949 *Proceedings of the Royal Society B*, 287(1926), 20200102.

950

951 Hodel, R. G., Chen, S., Payton, A. C., McDaniel, S. F., Soltis, P., & Soltis, D. E., (2017).
952 Adding loci improves phylogeographic resolution in red mangroves despite increased

953 missing data: comparing microsatellites and RAD-Seq and investigating loci filtering.
954 *Scientific Reports*, 7(1), 1-14.
955
956 Hou, L., Cui, Y., Li, X., Chen, W., Zhang, Z., Pang, X., & Li, Y., (2018). Genetic evaluation
957 of natural populations of the endangered conifer *Thuja koraiensis* using microsatellite
958 markers by restriction-associated DNA sequencing. *Genes*, 9(4), 218.
959
960 Huang, S., Hsieh, H. T., Fang, K., & Chiang, Y. C., (2004). Patterns of genetic variation
961 and demography of *Cycas taitungensis* in Taiwan. *The Botanical Review*, 70(1), 86-92.
962
963 Hurka, H., (1994). Conservation genetics and the role of botanical gardens. In
964 *Conservation Genetics*. Basel*: Birkhäuser Basel*, 371–380.
965
966 Ingham, J. A., Forster, P. I., Crisp, M. D., & Cook, L. G., (2013). Ancient relicts or recent
967 dispersal: how long have cycads been in central Australia? *Diversity and distributions*,
968 19(3), 307-316.
969
970 James, H. E., Forster, P. I., Lamont, R. W., & Shapcott, A., (2018). Conservation
971 genetics and demographic analysis of the endangered cycad species *Cycas megacarpa*
972 and the impacts of past habitat fragmentation. *Australian Journal of Botany*, 66(2), 173-
973 189.
974
975 Johnson, J. S., Gaddis, K. D., Cairns, D. M., Konganti, K., & Krutovsky, K. V. (2017).
976 Landscape genomic insights into the historic migration of mountain hemlock in response
977 to Holocene climate change. *American Journal of Botany*, 104(3), 439-450.

978
979 Jones, D.L., (2002). *Cycads of the world* 2nd ed., Washington, D.C.: Smithsonian
980 Institution Press.
981
982 Kärkkäinen, K., & Savolainen, O. (1993). The degree of early inbreeding depression
983 determines the selfing rate at the seed stage: model and results from *Pinus sylvestris*
984 (Scots pine). *Heredity*, 71(2), 160-166.
985
986 Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., & Prodöhl, P. A., (2013).
987 diveRsity: An R package for the estimation and exploration of population genetics
988 parameters and their associated errors. *Methods in Ecology and Evolution*, 4(8), 782-
989 788.
990
991 Keppel, G., Hodgskiss, P.D. & Plunkett, G.M., (2008). Cycads in the insular South- west
992 Pacific: dispersal or vicariance? *Journal of Biogeography*, 35(6), 1004– 1015.
993
994 Keppel, G., Lee, S. W., & Hodgskiss, P. D., (2002). Evidence for long isolation among
995 populations of a Pacific cycad: genetic diversity and differentiation in *Cycas seemannii* A.
996 Br. (Cycadaceae). *Journal of Heredity*, 93(2), 133-139.
997
998 Keppel, G., (2002). Low genetic variation in a Pacific cycad: conservation concerns for
999 *Cycas seemannii* (Cycadaceae). *Oryx*, 36(01), 41–49.
1000
1001 Kettle, C. J., Hollingsworth, P. M., Jaffré, T., Moran, B., & Ennos, R. A. (2007). Identifying
1002 the early genetic consequences of habitat degradation in a highly threatened tropical
1003 conifer, *Araucaria nemorosa* Laubenfels. *Molecular Ecology*, 16(17), 3581-3591.
1004

1005 Kono, M. & Tobe, H., (2007). Is *Cycas revoluta* (Cycadaceae) Wind- or Insect-Pollinated?
1006 *American Journal of Botany* (5), 847.
1007
1008 Kramer, A.T. & Havens, K., (2009). Plant conservation genetics in a changing world.
1009 *Trends in Plant Science*, 14(11), 599–607.
1010
1011 Lemopoulos, A., Prokkola, J. M., Uusi-Heikkilä, S., Vasemägi, A., Huusko, A., Hyvärinen,
1012 P., & Vainikka, A. (2019). Comparing RADseq and microsatellites for estimating genetic
1013 diversity and relatedness—Implications for brown trout conservation. *Ecology and*
1014 *Evolution,* 9(4), 2106-2120.
1015
1016 Liddle, D.R., (2009). Management Program for Cycads in the Northern Territory of
1017 Australia. *Northern Territory Department of Natural Resources, Environment, the Arts*
1018 *and Sport, Darwin.*
1019
1020 Li, Q., Xu, Z., & He, T., (2002). Ex situ genetic conservation of endangered *Vatica*
1021 *guangxiensis* (Dipterocarpaceae) in China. *Biological Conservation*, 106(2), 151-156.
1022
1023 Long-Qian, X. & Xun, G., (2006). Genetic differentiation and relationships of populations in
1024 the *Cycas balansae* complex (Cycadaceae) and its conservation complications. *Annals of*
1025 *Botany*, 97(5), 807–812.
1026
1027 Long-Qian, X. et al., (2004). ISSR Variation in the Endemic and Endangered Plant *Cycas*
1028 *guizhouensis* (Cycadaceae). *Annals of Botany*, 94(1), 133–138.
1029
1030 Lönn, M. & Prentice, H.C., (2002). Gene diversity and demographic turnover in central and

peripheral populations of the perennial herb *Gypsophila fastigiata*. *Oikos*, 99(3), 489–498.

Lowe, P.R., (2008). Species and Subspecies. *Ibis*, 64(1), 179–185.

Mable, B. K., (2019). Conservation of adaptive potential and functional diversity: integrating old and new approaches. *Conservation Genetics*, 20(1), 89-100.

Mahy, G., & Jacquemart, A. L., (1999). Early inbreeding depression and pollen competition in *Calluna vulgaris* (L.) Hull. *Annals of Botany*, 83(6), 697-704.

Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P., (2003). Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, 18(4), 189-197.

Mankga, L. T., & Yessoufou, K., (2017). Factors driving the global decline of cycad diversity. *AoB Plants*, 9(4), plx022.

Marler, T. E., (2010). Time-size trade-offs in responses of cycads to male cone herbivory. *Communicative & Integrative Biology*, 3(6), 602-603.

Meerow, A.W. & Nakamura, K., (2007). Ten microsatellite loci from *Zamia integrifolia* (Zamiaceae). *Molecular Ecology Notes*, 7(5), 824–826.

Meerow, A. W., Francisco-Ortega, J., Calonje, M., Griffith, M. P., Ayala-Silva, T., Stevenson, D. W., & Nakamura, K., (2012). *Zamia* (Cycadales: Zamiaceae) on Puerto Rico:

asymmetric genetic differentiation and the hypothesis of multiple introductions. *American Journal of Botany*, 99(11), 1828-1839.

Mondini, L., Noorani, A., & Pagnotta, M. A. (2009). Assessing plant genetic diversity by molecular tools. *Diversity*, 1(1), 19-35.

Mondoni, A., Probert, R. J., Rossi, G., Vegini, E., & Hay, F. R., (2011). Seeds of alpine plants are short lived: implications for long-term conservation. *Annals of Botany*, 107(1), 171-179.

Muriira, N. G., Muchugi, A., Yu, A., Xu, J., & Liu, A., (2018). Genetic diversity analysis reveals genetic differentiation and strong population structure in Calotropis plants. *Scientific reports*, 8(1), 1-10.

Nadarajan, J., Benson, E. E., Xaba, P., Harding, K., Lindstrom, A., Donaldson, J., & King, E., (2018). Comparative Biology of Cycad Pollen, Seed and Tissue-A Plant Conservation Perspective. *The Botanical Review*, 84(3), 295-314.

Nagalingum, N.S., Marshall, C.R., Quental, T.B., Rai, H.S., Little, D.P. and Mathews, S., (2011). Recent synchronous radiation of a living fossil. *Science,* 334(6057), 796–799.

Namoff, S., Husby, C. E., Francisco-Ortega, J., Noblick, L. R., Lewis, C. E., & Griffith, M. P., (2010). How well does a botanical garden collection of a rare palm capture the genetic variation in a wild population? *Biological Conservation*, 143(5), 1110-1117.

Namroud, M. C., Beaulieu, J., Juge, N., Laroche, J., & Bousquet, J., (2008). Scanning the genome for gene single nucleotide polymorphisms involved in adaptive population

differentiation in white spruce. *Molecular Ecology*, 17(16), 3599-3613.

Nazareno, A. G., Bemmels, J. B., Dick, C. W., & Lohmann, L. G., (2017). Minimum sample sizes for population genomics: an empirical study from an Amazonian plant species. *Molecular Ecology Resources*, 17(6), 1136-1147.

Nei, M., (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89(3), 583-590.

Newbold, T., Hudson, L. N., Arnell, A. P., Contu, S., De Palma, A., Ferrier, S., & Burton, V. J., (2016). Has land use pushed terrestrial biodiversity beyond the planetary boundary? *A global assessment. Science*, 353(6296), 288-291.

Nikitsky Botanical Gardens – National Scientific Centre of RAS et al., (2017). The Role of Botanical Gardens in the Conservation of the Family Cactaceae juss. Species. *Tambov University Reports. Series: Natural and Technical Sciences*, 22(5-1), 828–832.

Norstog, K. & Nicholls, T.J., (1997). *The biology of the cycads*, Ithaca, N.Y.: Comstock Pub. Associates.

Norstog, K.J. & Fawcett, P.K.S., (1989). Insect-Cycad Symbiosis and its Relation to the Pollination of *Zamia furfuracea* (Zamiaceae) by *Rhopalotria mollis* (Curculionidae). *American Journal of Botany*, 76(9), 1380.

Oakley, C. G., Lundemo, S., Ågren, J., & Schemske, D. W., (2019). Heterosis is common and inbreeding depression absent in natural populations of *Arabidopsis thaliana*. *Journal*

1109 *of Evolutionary Biology*, 32(6), 592-603.

1110

1111 Octavio-Aguilar, P., González-Astorga, J. & Vovides, A.P., (2009). Genetic diversity
1112 through life history of *Dioon edule* Lindley (Zamiaceae, Cycadales). *Plant Biology*, 11(4),
1113 525–536.

1114

1115 Ornduff, R., (1992). Features of Coning and Foliar Phenology, Size Classes, and Insect
1116 Associates of *Cycas armstrongii* (Cycadaceae) in the Northern Territory, Australia.
1117 *Bulletin of the Torrey Botanical Club*, 119(1), 39.

1118

1119 Pan, Y., Wang, X., Sun, G., Li, F., & Gong, X., (2016). Application of RAD sequencing for
1120 evaluating the genetic diversity of domesticated *Panax notoginseng* (Araliaceae). *PloS*
1121 *One*, 11(11).

1122

1123 Pauls, S. U., Nowak, C., Bálint, M., & Pfenninger, M., (2013). The impact of global
1124 climate change on genetic diversity within populations and species. *Molecular*
1125 *Ecology*, 22(4), 925-946.

1126

1127 Paz-Vinas, I., Loot, G., Hermoso, V., Veyssiere, C., Poulet, N., Grenouillet, G., & Blanchet,
1128 S., (2018). Systematic conservation planning for intraspecific genetic diversity. *Proceedings*
1129 *of the Royal Society B: Biological Sciences*, 285(1877), 20172746.

1130

1131 Peakall, R., Ebert, D., Scott, L. J., Meagher, P. F., & Offord, C. A., (2003). Comparative
1132 genetic study confirms exceptionally low genetic variation in the ancient and endangered
1133 relictual conifer, *Wollemia nobilis* (Araucariaceae). *Molecular Ecology*, 12(9), 2331-2343.

1134

1135 Peakall, R. & Smouse, P.E., (2012). GenAlEx 6.5: genetic analysis in Excel. Population
1136 genetic software for teaching and research--an update. *Bioinformatics*, 28(19), 2537–2539.
1137
1138 Pérez-Farrera, M. A., Vovides, A. P., Octavio-Aguilar, P., González-Astorga, J., De La
1139 Cruz-rodríguez, J., Hernández-Jonapá, R., & Villalobos-Méndez, S. M., (2006).
1140 Demography of the cycad *Ceratozamia mirandae* (Zamiaceae) under disturbed and
1141 undisturbed conditions in a biosphere reserve of Mexico. *Plant Ecology,* 187(1), 97-108.
1142
1143 Pinares, A., Gonzalez-Astorga, J., Vovides, A. P., Lazcano, J., & Vendrame, W. A.,
1144 (2009). Genetic diversity of the endangered endemic *Microcycas calocoma* (Miq.) A. DC
1145 (Zamiaceae, Cycadales): Implications for conservation. *Biochemical Systematics and*
1146 *Ecology*, 37(4), 385-394.
1147
1148 Plenk, K., Bardy, K., Höhn, M., & Kropf, M., (2019). Long-term survival and successful
1149 conservation? Low genetic diversity but no evidence for reduced reproductive success at
1150 the north-westernmost range edge of *Poa badensis* (Poaceae) in Central
1151 Europe. *Biodiversity and conservation*, 28(5), 1245-1265.
1152
1153 Plomion, C., Bartholomé, J., Lesur, I., Boury, C., Rodríguez-Quilón, I., Lagraulet, H., & de
1154 Miguel, M., (2016). High-density SNP assay development for genetic analysis in maritime
1155 pine (*Pinus pinaster*). *Molecular Ecology Resources*, 16(2), 574-587.
1156
1157 Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure
1158 using multilocus genotype data. *Genetics*, 155(2), 945-959.
1159
1160 Pujol, B., Archambeau, J., Bontemps, A., Lascoste, M., Marin, S., & Meunier, A., (2017).
1161 Mountain landscape connectivity and subspecies appurtenance shape genetic

differentiation in natural plant populations of the snapdragon (*Antirrhinum majus* L.). *Botany Letters*, 164(2), 111-119.

Qu, W. M., Liang, N., Wu, Z. K., Zhao, Y. G., & Chu, D., (2019). Minimum sample sizes for invasion genomics: Empirical investigation in an invasive whitefly. *Ecology and Evolution*. 10(1), 38–49.

Rae, D., (2011). Fit for purpose: the importance of quality standards in the cultivation and use of live plant collections for conservation. *Biodiversity and Conservation*, 20(2), 241-258.

R Core Team, (2019). R: A Language and Environment for Statistical Computing, version 3.4.4. Vienna, Austria: R Foundation for Statistical Computing.

Reed, T.E., Jenouvrier, S. & Visser, M.E., (2013). Phenological mismatch strongly affects individual fitness but not population demography in a woodland passerine A. Roulin, ed. *The Journal of Animal Ecology*, 82(1), 131–144.

Rodríguez-Rodríguez, P., Pérez de Paz, P.L. & Sosa, P.A., (2018). Species delimitation and conservation genetics of the Canarian endemic *Bethencourtia* (Asteraceae). *Genetica*, 146(2), 199–210.

Roodt, D., Lohaus, R., Sterck, L., Swanepoel, R.L., Van de Peer, Y. and Mizrachi, E. (2017). Evidence for an ancient whole genome duplication in the cycad lineage. *PLoS ONE*, 12(9), p. e0184454.

Rosche, C., Schrieber, K., Lachmuth, S., Durka, W., Hirsch, H., Wagner, V., & Hensen,

I., (2018). Sex ratio rather than population size affects genetic diversity in *Antennaria dioica*. *Plant Biology*, 20(4), 789-796.

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, 145(4), 1219-1228.

Ruhsam, M., Hollingsworth, P. M., Squirrell, J., & Ennos, R. A., (2010). Significant differences in outcrossing rate, self-incompatibility, and inbreeding depression between two widely hybridizing species of Geum. *Biological Journal of the Linnean Society*, 101(4), 977-990.

Salas-Leiva, D. E., Meerow, A. W., Calonje, M., Francisco-Ortega, J., Griffith, M. P., Nakamura, K., & Knowles, D., (2017). Shifting Quaternary migration patterns in the Bahamian archipelago: Evidence from the *Zamia pumila* complex at the northern limits of the Caribbean island biodiversity hotspot. *American Journal of Botany*, 104(5), 757-771.

Sampson, J. F., Byrne, M., Gibson, N., & Yates, C., (2016). Limiting inbreeding in disjunct and isolated populations of a woody shrub. *Ecology and evolution*, 6(16), 5867-5880.

Schoen, D. J., & Brown, A. H. (2001). The conservation of wild plant species in seed banks: attention to both taxonomic coverage and population biology will improve the role of seed banks as conservation tools. *BioScience*, 51(11), 960-966.

Shafer, A. B., Peart, C. R., Tusso, S., Maayan, I., Brelsford, A., Wheat, C. W., & Wolf, J. B., (2017). Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods in Ecology and Evolution*, 8(8), 907-917.

Sharma, I.K., Jones, D.L. & Forster, P.I., (2004). Genetic differentiation and phenetic relatedness among seven species of the *Macrozamia plurinervia* complex (Zamiaceae). *Biochemical Systematics and Ecology*, 32(3), 313–327.

Sharma, I.K., Jones, D.L. & Forster, P.I., (1999). Contribution of isozymic analysis in differentiating *Macrozamia moorei* and KD Hill & *M. johnsonii* F. Muell (Zamiaceae). *Austrobaileya*, 5(2), 363–365.

Sharma, I. K., Jones, D. L., Forster, P. I., & Young, A. G., (1998). The extent and structure of genetic variation in the *Macrozamia pauli-guilielmi* complex (Zamiaceae). *Biochemical Systematics and Ecology*, 26(1), 45-54.

Shuguang, J., Yang, Z., Nian, L., Zezheng, G., Qiang, W., Zhenhua, X. and Hal, R. (2006). Genetic variation in the endangered endemic species *Cycas fairylakea* (Cycadaceae) in China and implications for conservation. *Biodiversity & Conservation*, 15(5), 1681–1694.

Slatkin, M. (1981). Estimating levels of gene flow in natural populations. *Genetics*, 99(2), 323-335.

Sletvold, N., Mousset, M., Hagenblad, J., Hansson, B., & Ågren, J. (2013). Strong inbreeding depression in two Scandinavian populations of the self-incompatible perennial herb *Arabidopsis lyrata*. *Evolution*, 67(10), 2876-2888.

Sullivan, E. R., Barker, C., Powell, I., & Ashton, P. A. (2019). Genetic diversity and connectivity in fragmented populations of *Rhinanthus minor* in two regions with

contrasting land-use. *Biodiversity and Conservation*, 28(12), 3159-3181.

Schmidt, L., Fischer, M., & Oja, T. (2018). Two closely related species differ in their regional genetic differentiation despite admixing. *AoB Plants*, 10(1), ply007.

Sork, V. L., Nason, J., Campbell, D. R., & Fernandez, J. F. (1999). Landscape approaches to historical and contemporary gene flow in plants. *Trends in Ecology & Evolution*, 14(6), 219-224.

Stojanova, B., Šurinová, M., Zeisek, V., Münzbergová, Z., & Pánková, H. (2020). Low genetic differentiation despite high fragmentation in the endemic serpentinophyte *Minuartia smejkalii* (M. verna agg., Caryophyllaceae) revealed by RADSeq SNP markers. *Conservation Genetics*, 1-12.

Storfer, A. (1999). Gene flow and endangered species translocations: a topic revisited. *Biological Conservation*, 87(2), 173-180.

Suinyuy, T.N., Donaldson, J.S. & Johnson, S.D., (2009). Insect pollination in the African cycad *Encephalartos friderici-guilielmii* Lehm. *South African Journal of Botany*, 75(4), 682–688.

Sunde, J., Yıldırım, Y., Tibblin, P., & Forsman, A. (2020). Comparing the Performance of Microsatellites and RADseq in Population Genetic Studies: Analysis of Data for Pike (*Esox lucius*) and a Synthesis of Previous Studies. *Frontiers in genetics*, 11, 218.

Swart, C., Donaldson, J. & Barker, N., (2018). Predicting the distribution of *Encephalartos latifrons*, a critically endangered cycad in South Africa.

Biodiversity and Conservation, 27(8), 1961–1980.

Szczecińska, M., Sramko, G., Wołosz, K., & Sawicki, J., (2016). Genetic diversity and population structure of the rare and endangered plant species *Pulsatilla patens* (L.) Mill in East Central Europe. *PLoS One*, 11(3).

Tang, W., (1990). Reproduction in the cycad *Zamia pumila* in a fire-climax habitat: an eight-year study. *Bulletin of the Torrey Botanical Club*, 368-374.

Terry, I., Tang, W., & Marler, T. E., (2012). Pollination systems of Island Cycads: predictions based on island biogeography. In Proceedings of Cycad 2008. The 8th International Conference on Cycad Biology, Panama City, Panama, 13-15 January 2008 (102-132). *New York Botanical Garden Press*.

Torgersen, J.S., (2017). Crime, culture and collecting: the illicit cycad market in South Africa.

Van Geert, A., Van Rossum, F., & Triest, L., (2008). Genetic diversity in adult and seedling populations of *Primula vulgaris* in a fragmented agricultural landscape. *Conservation Genetics*, 9(4), 845.

Vogl, C., Karhu, A., Moran, G., & Savolainen, O., (2002). High resolution analysis of mating systems: inbreeding in natural populations of Pinus radiata. *Journal of Evolutionary Biology*, 15(3), 433-439.

Vovides, A.P., (1990). Spatial Distribution, Survival, and Fecundity of *Dioon edule*

(Zamiaceae) in a Tropical Deciduous Forest in Veracruz, Mexico, with Notes on Its Habitat. *American Journal of Botany*, 77(12), 1532.

Vovides, A. P., Ogata, N., Sosa, V., & E. Peña-García (1997). Pollination of endangered Cuban cycad *Microcycas calocoma* (Miq.) A. DC. Botanical Journal of the Linnean Society, 125(3), 201-210.

Wang, J. (2019). A parsimony estimator of the number of populations from a STRUCTURE-like analysis. *Molecular ecology resources*, 19(4), 970-981.

Watkinson, A. R., & Powell, J. C. (1997). The life history and population structure of *Cycas armstrongii* in monsoonal northern Australia. *Oecologia*, 111(3), 341-349.

Whitlock, R., Hipperson, H., Thompson, D. B. A., Butlin, R. K., & Burke, T., (2016). Consequences of *in-situ* strategies for the conservation of plant genetic diversity. *Biological Conservation*, 203, 134-142.

Williams, C. G., & Savolainen, O., (1996). Inbreeding depression in conifers: implications for breeding strategy. *Forest Science*, 42(1), 102-117.

Willi, Y., Van Buskirk, J., & Hoffmann, A. A., (2006). Limits to the adaptive potential of small populations. *Annu. Rev. Ecol. Evol. Syst*., 37, 433-458.

Wright, L.I., Tregenza, T. & Hosken, D.J., (2007). Inbreeding, inbreeding depression and extinction. *Conservation Genetics*, 9(4), 833–843.

1322 Wu, X., Ruhsam, M., Wen, Y., Thomas, P.I., Worth, J.R., Lin, X., Wang, M., Li, X., Chen,
1323 L., Lamxay, V. and Le Canh, N., (2020). The last primary forests of the Tertiary relict
1324 *Glyptostrobus pensilis* contain the highest genetic diversity. *Forestry: An International*
1325 *Journal of Forest Research*, 93(3), 359-375.
1326
1327 Wu, C. S., & Chaw, S. M., (2015). Evolutionary stasis in cycad plastomes and the first
1328 case of plastome GC-biased gene conversion. *Genome biology and evolution*, *7*(7), 2000-
1329 2009.
1330
1331 Yang, S.-L. & Meerow, A.W., (1996). The *Cycas pectinata* (Cycadaceae) complex: genetic
1332 structure and gene flow. *International Journal of Plant Sciences*, 157(4), 468–483.
1333
1334 Yang, J., Gao, Z., Sun, W., & Zhang, C., (2016). High regional genetic differentiation of an
1335 endangered relict plant *Craigia yunnanensis* and implications for its conservation. *Plant*
1336 *diversity*, 38(5), 221-226.
1337
1338 Yessoufou, K., Daru, B. H., Tafirei, R., Elansary, H. O., & Rampedi, I., (2017). Integrating
1339 biogeography, threat and evolutionary data to explore extinction crisis in the taxonomic
1340 group of cycads. *Ecology and evolution*, 7(8), 2735-2746.
1341
1342 Yoder, A. D., Poelstra, J. W., Tiley, G. P., & Williams, R. C., (2018). Neutral theory is the
1343 foundation of conservation genetics. *Molecular biology and evolution*, 35(6), 1322-1326.
1344
1345 Young, A., Boyle, T. & Brown, T., (1996). The population genetic consequences of habitat
1346 fragmentation for plants. *Trends in Ecology & Evolution*, 11(10), 413– 418.
1347

Zhan, Q. Q., Wang, J. F., Gong, X., & Peng, H., (2011). Patterns of chloroplast DNA variation in *Cycas debaoensis* (Cycadaceae): conservation implications. *Conservation Genetics*, 12(4), 959-970.

Zhang, Y. Y., Shi, E., Yang, Z. P., Geng, Q. F., Qiu, Y. X., & Wang, Z. S., (2018). Development and application of genomic resources in an endangered palaeoendemic tree, *Parrotia subaequalis* (Hamamelidaceae) from eastern China. *Frontiers in plant science*, 9, 246.

Zimmerman, S. J., Aldridge, C. L., & Oyler-McCance, S. J. (2020). An empirical comparison of population genetic analyses using microsatellite and SNP data for a species of conservation concern. *BMC genomics*, 21, 1-16.

Figures and Tables

Figure 1 *Cycas calcicola* populations growing in the wild in the Katherine region and Litchfield National Park. (A) Large population of *C. calcicola* growing on sandstone in the Litchfield National Park, Northern Territory. (B) Small group of *C. calcicola* growing on limestone within the Katherine region.

Figure 2 Distribution of samples of *C. calcicola* in Northern Territory. Map of the northern part of the Northern Territory, Australia showing sampling sites of wild (Litchfield National Park and Katherine Region) and cultivated *ex-situ* conservation collections (Spirit Hills Conservation Site (Cul.) and Daly River (Cul.), representing the entire range of the species. Inset: sampling locations in Darwin region within Australia. Area of occurrence: representing the extent of occurrence for each species based in herbarium specimen records Australasian Virtual Herbarium (https://avh.chah.org.au).

Figure 3 STRUCTURE plot of *Cycas calcicola* populations. Plot representing 72 samples from nine populations. The most likely number of genetic groups for the species was *K*=2 (D*K* = 1883) indicating two clusters within the data. *Ex-situ* cultivated populations = Katherine CUL, Daly River and Spirit Hills.

Figure 4. DAPC graph of *Cycas calcicola* populations. Discriminant analysis of principal components (DAPC) of nine *C. calcicola* populations. DAPC is a summary of 22 principal components with three discriminate functions (*K* = 2) and a proportion of conserved variance of 0.527. Inset: indicates the first axis of the DAPC. *Ex-situ* cultivated populations = Katherine CUL, Daly River and Spirit Hills. Inset: shows the first axis of the DAPC, which helps to demonstrate the separation between genetic groups.