From the Neotropics to the Namib: evidence for rapid ecological divergence following extreme long-distance dispersal

ARTICLE *in* BOTANICAL JOURNAL OF THE LINNEAN SOCIETY · OCTOBER 2015

Impact Factor: 2.53 · DOI: 10.1111/boj.12334

4 AUTHORS:

Enelge Gildenhuys

Stellenbosch University

5 PUBLICATIONS **6** CITATIONS

SEE PROFILE

Allan G. Ellis

Stellenbosch University

44 PUBLICATIONS **674** CITATIONS

SEE PROFILE

Scott P. Carroll

University of California, Davis/Institute for …

76 PUBLICATIONS **2,945** CITATIONS

SEE PROFILE

Johannes J Le Roux

Stellenbosch University

51 PUBLICATIONS **777** CITATIONS

SEE PROFILE

The LINNEAN SOCIETY of London BOTANICAL Journal of the Linnean Society

Botanical Journal of the Linnean Society, 2015, ••, ••–••. With figures

From the Neotropics to the Namib: evidence for rapid ecological divergence following extreme long-distance dispersal

ENELGE GILDENHUYS[1], ALLAN G ELLIS[2], SCOTT CARROLL[3,4] and JOHANNES J LE ROUX[1]*

[1]*Centre for Invasion Biology, Department of Botany and Zoology, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa*
[2]*Department of Botany and Zoology, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa*
[3]*Department of Entomology, University of California, Davis, CA, USA*
[4]*Institute for Contemporary Evolution, Davis, CA 95616, USA*

Received 26 January 2015; revised 29 July 2015; accepted for publication 5 August 2015

Extreme long-distance dispersal is an important process in plant biogeography. Such events can lead to rapid diversification due to founder effects, genetic drift and novel selection in recipient environments. Balloon vines (*Cardiospermum* spp.) are mainly Neotropical, but include two native southern African species, the endemic desert-adapted *C. pechuelii* and the moist subtropical *C. corindum* (which also occurs in the Neotropics). We used phylogenetic approaches (internal transcribed spacer (ITS), *rpl32* and *trnL-trnF* DNA sequencing data) and population genetics (amplified fragment length polymorphism (AFLP) analyses) to confirm the long-distance dispersal of *C. corindum* to southern Africa and to reveal the subsequent divergence of the morphologically and ecologically extreme but genetically close *C. pechuelii*. We could not judge whether incongruences between ecological requirements and morphology and gene trees for the African species resulted from ongoing gene flow or incomplete lineage sorting, but our findings do support recent divergence of *C. pechuelii* from *C. corindum* in Africa following transoceanic dispersal of the lineage. © 2015 The Linnean Society of London, *Botanical Journal of the Linnean Society*, 2015, ••, ••–••.

ADDITIONAL KEYWORDS: balloon vine – *Cardiospermum* – endemic – speciation.

INTRODUCTION

Extreme long-distance dispersal is more prevalent in plants than previously thought, with numerous recently described examples of distant transoceanic exchanges (e.g. Givnish *et al*., 2004; Renner, 2004; Tremetsberger *et al*., 2005; Nettel & Dodd, 2007; Christenhusz & Chase, 2013; Takayama *et al*., 2013; Le Roux *et al*., 2014). Such dispersal events likely result in reduced genetic variation due to strong founder effects (Austerlitz *et al*., 1997) and strong local selection in novel environments (Carson & Templeton, 1984; Templeton, 2008). Radiations following single long-distance colonization events have been well-documented on islands (e.g. Wagner & Funk, 1995; Baldwin & Sanderson, 1998; Carlquist, Baldwin & Carr, 2003; Price & Wagner, 2004; Knope *et al*., 2012; García-Verdugo & Fay, 2014, and references therein), often leading to the evolution of floras that are distinct from those found in the areas from which they were derived (Whittaker, Jones & Partomihardjo, 1997). Here we investigate the possibility of rapid divergence in *Cardiospermum* L. (Sapindaceae) following extreme long-distance dispersal between the South American and African continents.

The genus *Cardiospermum* (17 species), commonly known as balloon vines, are mostly restricted to the Neotropics with a few exceptions (Gildenhuys *et al*., 2013). Numerous species have also been introduced globally to areas where they are now considered invasive (Gildenhuys *et al*., 2013, 2014). For example, in

*Corresponding author. E-mail: jleroux@sun.ac.za; jacoleroux01@gmail.com

southern Africa *C. grandiflorum* Sw. and *C. halicacabum* L. are thought to be introduced from South America and invasive, whereas *C. corindum* L. is regarded as native to both southern Africa and South America (Gildenhuys *et al.*, 2014). A phylogenetic treatment of the group suggests that the latter represents an extreme long-distance dispersal event from South America to southern Africa, which occurred between 5.9 and 15.1 Ma (Gildenhuys *et al.*, 2014).

African and South American *C. corindum* are similar morphologically, but *C. corindum* in Namibia is found in savannah regions (Fig. 1C), different from the more forested areas typical of the species in South America. *Cardiospermum corindum* is also widespread in both subtropical and arid areas in Africa and has been recorded in countries like Angola (J. Le Roux, personal observation), Tanzania (S Carroll, personal observation, Uganda (A. Witt, personal communication), Zimbabwe, Botswana etc. (Global Biodiversity Information Facility, GBIF). Some New World *C. corindum* populations are also found in arid regions such as the Sonoran Desert (S Carroll, personal observation). African *C. corindum* from southern Africa appears phylogenetically more closely related to the Namib Desert endemic, *C. pechuelii* Kuntze, than to South American *C. corindum*, despite its morphological resemblance to the latter (Gildenhuys *et al.*, 2014). This paraphyletic grouping of *C. corindum* is peculiar given that *C. pechuelii* is distinct from African *C. corindum* based on morphology and habitat requirements (Fig. 1). *Cardiospermum pechuelii* has smaller leaf surface areas, a more compact shrub-like growth form and reduced fruit capsules compared to *C. corindum*, potentially xerophytic adaptations that may conserve water in the extremely arid Namib Desert habitats where it occurs (Fig. 1B, C). Morphological differences between *C. pechuelii* and *C. corindum* from South America and southern Africa persist when

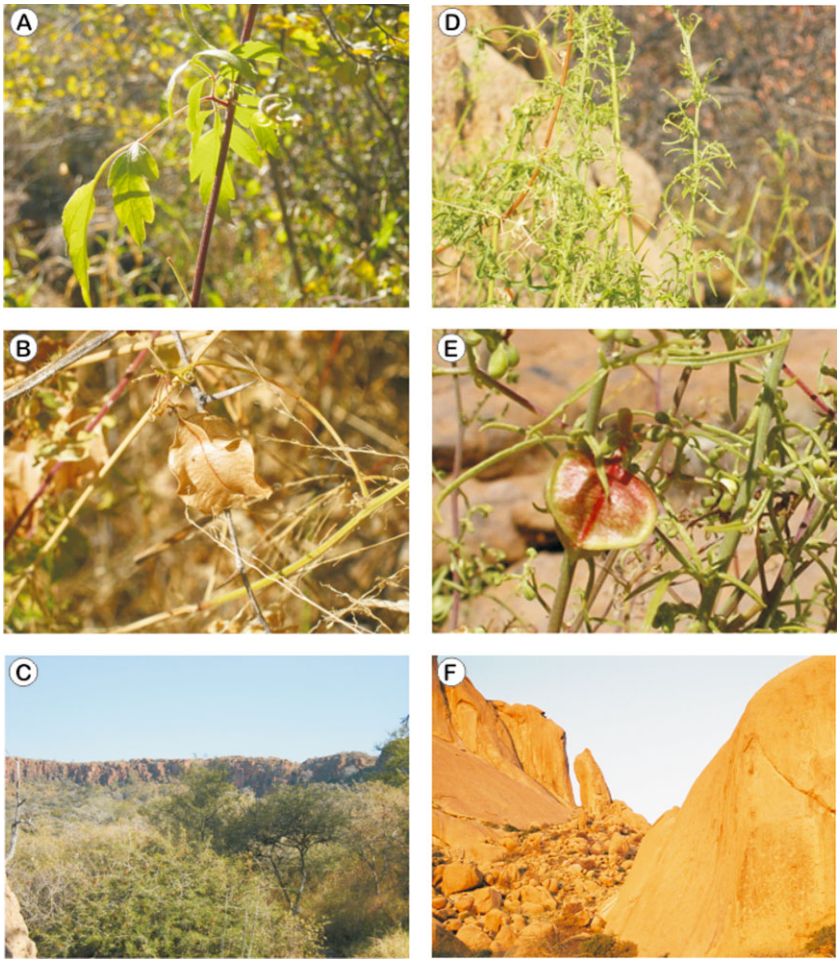

Figure 1. Typical morphological and environmental characteristics of *Cardiospermum corindum* (A–C) and *C. pechuelii* (D–F) in their native ranges in Namibia: leaf morphology (A, D); fruit morphology (B, E); and habitats occupied: the savannah Waterberg region (C) and rocky outcrops of the Namib Desert (F).

grown under common garden conditions, indicating that phenotypic plasticity is an unlikely cause of these differences (E. Gildenhuys, unpubl. data). The latter, in conjunction with the paraphyletic nature of *C. corindum*, suggests that African *C. corindum* constitutes a separate taxon rather than being a conspecific variant of *C. pechuelii*. The strikingly different morphological adaptations in African *C. pechuelii* and *C. corindum* and their peculiar phylogenetic relationship provide an interesting case to investigate how adaptation and divergence can occur following extreme long-distance dispersal.

Here, expanding on a previous study on the biogeography of selected *Cardiospermum* spp. (Gildenhuys *et al*., 2014), we aim to better understand the historic relationships between *C. pechuelii* and *C. corindum* using phylogenetic and population genetic approaches. This will allow us to elucidate the dynamics of diversification following intercontinental long-distance dispersal. First, using DNA sequencing data from multiple gene regions we hope to shed light on the number and order of colonization events and the patterns of diversification of African taxa. Second, a population genetic approach will allow us to compare genetic diversity estimates between taxa and geographical regions and will be informative concerning the occurrence of genetic bottlenecks during colonization and the extent of reproductive isolation following the diversification of African taxa. Finally, this study is relevant to ongoing biological control strategies against invasive balloon vine species in southern Africa.

MATERIALS AND METHODS

Material collection

Leaf material was collected for 25 and nine *C. corindum* plants from South America and southern Africa, respectively, and 32 *C. pechuelii* plants from Namibia (Fig. 2) (Supporting Information Table S1). Voucher specimens for each species have been deposited in the Geo-Potts Herbarium, Bloemfontein, South Africa (voucher numbers BLFU1336-1340). Whole genomic DNA was extracted from silica dried leaf material using the cetyltrimethyl ammonium bromide method as described by Doyle & Doyle (1990), modified by adding 1% PVP-40T. DNA quality and quantity were checked using a NanoDrop spectrophotometer (NanoDrop ND-1000, Inqaba Biotec, South Africa) and samples diluted to a final concentration of *c.* 50 ng/μL.

Gene and AFLP amplification

One nuclear internal transcribed spacer (ITS) and two plastid (*rpl32* and *trnL-trnF*) loci were used in this study. Briefly, the nuclear (nDNA) ITS region was amplified using the universal primers ITS1 (5′-TCC GTA GGT GAA CCT GCG G-3′) and ITS4 (5′-TCC TCC GCT TAT TGA TAT GC-3′) and in certain cases ITS5 (5′-GGA AGG AGA AGT CGT AAC AAG G-3′) (White *et al*., 1990). These primers amplify the ribosomal transcribed spacer regions 1 and 2, and the 5.8S region. Some PCR products for the ITS region were cloned due to sequence ambiguity caused by polymerase drop-off (see Gildenhuys *et al*. (2014) for further details of the PCR and cloning methods). For plastid DNA regions, the *trnL* intron and the *trnL-trnF* intergenic spacer (hereafter *trnL-F*) were amplified using the primers c (5′-CGA AAT CGG TAG ACG CTA CG-3′) and f (5′-ATT TGA ACT GGT GAC ACG AG-3′) (Taberlet *et al*., 1991) and rpl32-trnL (5′-CTG CTT CCT AAG AGC AGC GT-3′) and rpl32-F (5′-CAG TTC CAA AAA AAC GTA CTT-3′) (Shaw *et al*., 2007) were used for amplification of the *rpl32* region. Full details of amplification conditions used here are described in Gildenhuys *et al*. (2014). Forty-four individuals were sequenced for both plastid DNA regions (*rpl32* and *trnL-F*), of which a subset of 37 individuals were sequenced for ITS.

PCR amplification of AFLP fragments was performed using the 'universal' protocol described by Blignaut, Ellis & Le Roux (2013). Briefly, following digestion and pre-selective PCR, selective PCR amplification with one primer pair (*Eco*R1-CAT and *Mse*1-CTT) resulted in good quality and repeatable profiles. Forty-six samples were amplified successfully for AFLP markers. To estimate error rates and repeatability of AFLP loci, all samples were duplicated either by PCR or re-extraction to account for an additional 49 AFLP profiles included in the repeatability analyses.

Phylogenetic analysis

DNA sequence data were edited and aligned in BioEdit v 7.0.5.3 (Hall, 1999) and MAFFT (Katoh & Standley, 2013). Two data matrices were created, a combined plastid matrix (*rpl32* and *trnL-F*) for 44 samples and a combined matrix including all three loci (ITS, *rpl32* and *trnL-F*) for 37 samples. Sequence data for all three loci for accessions of *C. halicacabum*, the sister taxon to *C. corindum* and *C. pechuelii* (Gildenhuys *et al*., 2014), were included as outgroup data to root our phylogenetic trees.

A plastid DNA haplotype network was constructed from the *rpl32* and *trnL-F* dataset using statistical with 95% connection limit in TCS v 1.21 (Clement, Posada & Crandall, 2000). We chose network analysis in addition to phylogenetic tree building approaches to visualize and explore the data as networks are better suited to explore relationships between gene regions sampled within a species which are often not hierarchical as assumed by traditional phylogenetic

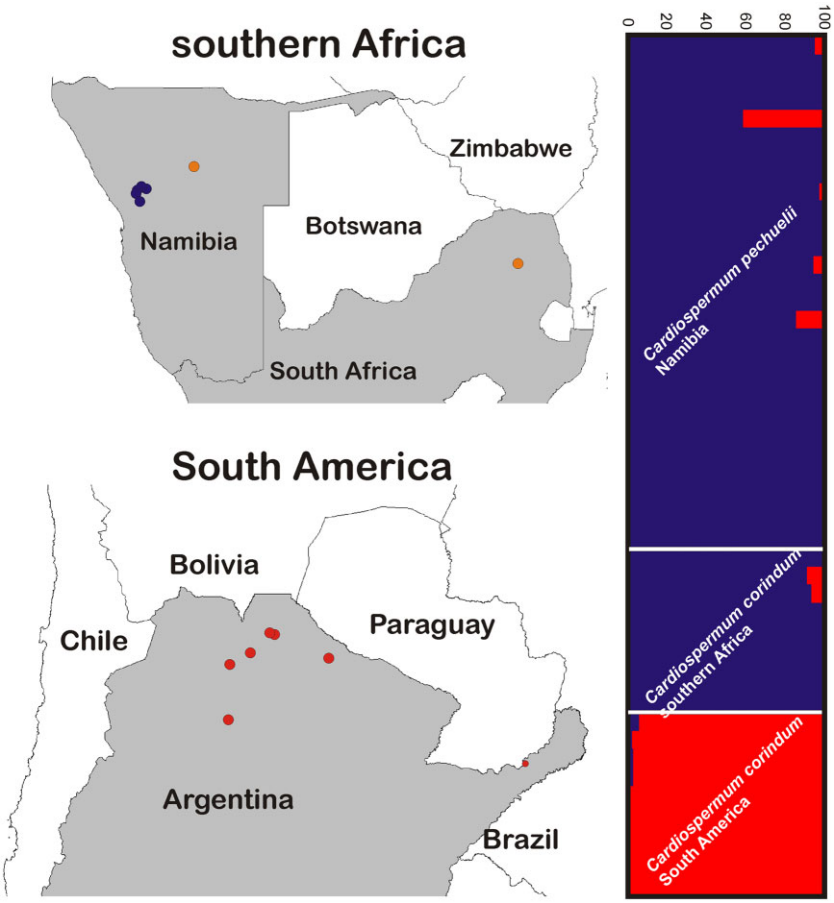

Figure 2. Maps indicating sampled localities of populations of *C. corindum* in southern Africa (orange), *C. pechuelii* (blue) in Namibia and *C. corindum* in South America (red). The bar plot on the right shows genetic structure across these populations determined using Bayesian assignment tests. For $K = 2$ (most likely number of genetic clusters) each horizontal bar illustrates a different individual and each colour the proportion of its genome assigned to each genetic cluster.

tree reconstruction methods (Posada & Crandall, 2001). A Bayesian inference phylogenetic tree was reconstructed from separate plastid DNA and nDNA partitions using Mr Bayes v 3.2 (Ronquist & Huelsenbeck, 2003). Tree topologies for the two datasets were congruent, but with lower nodal support for the plastid DNA data (results not shown). We therefore re-ran the analysis for all three loci. jModelTest (Posada, 2008) and the Akaike information criterion (Akaike, 1973) was used to determine the best-fit model for our data. The Bayesian model was run for 4 million generations sampling every 1000th generation and a consensus tree was built, discarding the first 25% of trees as burn-in. Posterior probabilities (PP) were calculated using a majority rule consensus method to assess tree topology support.

AFLP GENOTYPING

Genemarker Version 2.2.0 (SoftGenetics, LLC, CA, USA) was used to visually investigate data quality and remove failed AFLP profiles. AFLP loci scoring followed the methods proposed by Ley & Hardy (2013). Briefly, cleaned data were imported into PEAKSCANNER© software v 1.0 (Applied Biosystems, Foster City CA, USA) where peaks were automatically scored using default settings. The fragment sizing table was exported from PEAKSCANNER© and imported into tinyFLP v1.22 (Arthofer, 2010) to optimize peak selection. Settings for tinyFLP followed those described by Ley & Hardy (2013) with the following adjustments: minimum peak height, 1000; minimum fragment size, 80 bp; maximum fragment size, 400 bp; and maximum frequency, 100%. The modified dataset was transformed into a binary matrix using tinyCAT (Arthofer, 2010).

AFLP REPEATABILITY

Duplicate samples (re-extracted and/or duplicated PCRs) were used to estimate the broad sense herit-

ability (H^2) as a measure of repeatability for each locus in SPAGeDi v 1.4 (Hardy & Vekemans, 2002), where H^2 = Variance $(V)_{\text{among individuals}}/(V_{\text{among individuals}} + V_{\text{within individuals}})$. We used 1000 permutations and coded data as haploid (H^2 neglects dominance effects). The output data gives an individual F_{ST} output for each locus which is a direct measurement of H^2 because, similar to H^2, F_{ST} measures the amount of genetic variance due to the population effect, i.e. $F_{ST} = V_{\text{among populations}}/(V_{\text{among populations}} + V_{\text{within populations}}$ (Ley & Hardy, 2013). All loci with low heritability ($H^2 < 0.7$), and thus low repeatability, were omitted from further analyses.

We used Structure v 2.3.4 (Falush, Stephens & Pritchard, 2007) to assign individual AFLP genotypes to genetic clusters using a Bayesian approach. Simulations were run for one to seven genetic clusters (K) using the admixture model and correlated allele frequencies. Ten runs of 500 000 iterations and burn-in of 20 000 were conducted for each K value. Structure harvester (Earl & von Holdt, 2012) was used to determine the optimal number of K, using the Δ K method of Evanno, Regnaut & Goudet (2005).

All genetic diversity analyses were performed in GenAlEx6 v 6.50 (Peakall & Smouse, 2006). We estimated genetic diversity indices: (1) treating southern African *C. corindum* and *C. pechuelii* as two separate populations; and (2) treating all African accessions as a single population. Population differentiation and diversity indices included analysis of molecular variance (AMOVA), number of effective alleles (*Ne*), number of private alleles (P_A), Nei's gene diversity (h) and Shannon's diversity index (I). Pairwise population divergences were estimated using Nei's genetic distance (Nei, 1978). Genetic diversity indices can be informative concerning the evolutionary history of African balloon vines in two ways. First, comparable levels of unique alleles in African *C. pechuelii* and *C. corindum* populations may indicate that these two species have diversified shortly after the arrival of the ancestral taxon in Africa from South America and therefore would have experienced the same amount of time for new mutations to accrue. On the other hand, low numbers of unique alleles and reduced gene diversity in *C. pechuelii* compared with African *C. corindum* populations may indicate a more recent divergence and that the onset of diversification between these two species did not directly follow the arrival of *C. corindum* in southern Africa.

RESULTS

Sequence variation

The aligned combined matrix of all three loci (*rpl32*, *trnL-F* and ITS) contained 2039 base pairs (bp) and the matrix including only plastid loci contained 1460 bp. All DNA sequences have been deposited in GenBank (http://www.ncbi.nlm.nih.gov/genbank/). The *rpl32* and *trnL-F* matrix (no outgroups) required 27 gaps (indels) for alignment, ranging from 1 to 6 bp in size. Gaps > 2 bp, were shortened to 1 bp to avoid treating a single mutation (inversion, deletion or translocation) as multiple mutations. The combined matrix, with the two additional *C. halicacabum* outgroup sequences obtained from Gildenhuys *et al*. (2014), required 29 gaps for alignment.

Haplotype network

The haplotype network consisted of 14 haplotypes, illustrating a close relationship between samples from southern African *C. corindum* and *C. pechuelii*, with these two taxa even sharing identical haplotypes (Fig. 3). Moreover, *C. corindum* from southern Africa and *C. pechuelii* were more closely related to each other than to any *C. corindum* haplotypes identified from South America, with the latter isolated by several mutational steps (Fig. 3).

Phylogenetic tree

The best-fit model for the combined matrix was identified as the general time reversible model (GTR + G). *Cardiospermum pechuelii* rendered *C. corindum* paraphyletic, with African *C. corindum* being more closely related to *C. pechuelii* than to any South American accessions (Fig. 3). African and South American accessions did not form reciprocally monophyletic groups, with one South American sample (JLR291) sister to the remainder on the tree. Southern African accessions formed a single monophyletic group nested in South American *C. corindum*. This well supported southern African clade (PP = 1) included *C. pechuelii* and *C. corindum*, and further structure in this clade was only moderately supported (Fig. 3).

Genetic diversity and structure analysis

Forty-one AFLP loci were considered reproducible (H^2 scores > 0.7). Bayesian assignment analysis identified two genetic clusters within the data (Fig. 2). The first cluster corresponded to all *C. pechuelii* and *C. corindum* individuals from southern Africa (indicated in blue in Fig. 2) and the second cluster to South American *C. corindum* (indicated in red in Fig. 2).

AMOVA results showed that 37% of the AFLP genetic variation resided within populations and 63% of the variation resided among populations (Table 1). Even though sample sizes were higher for *C. pechuelii* than for *C. corindum*, all genetic diversity indices

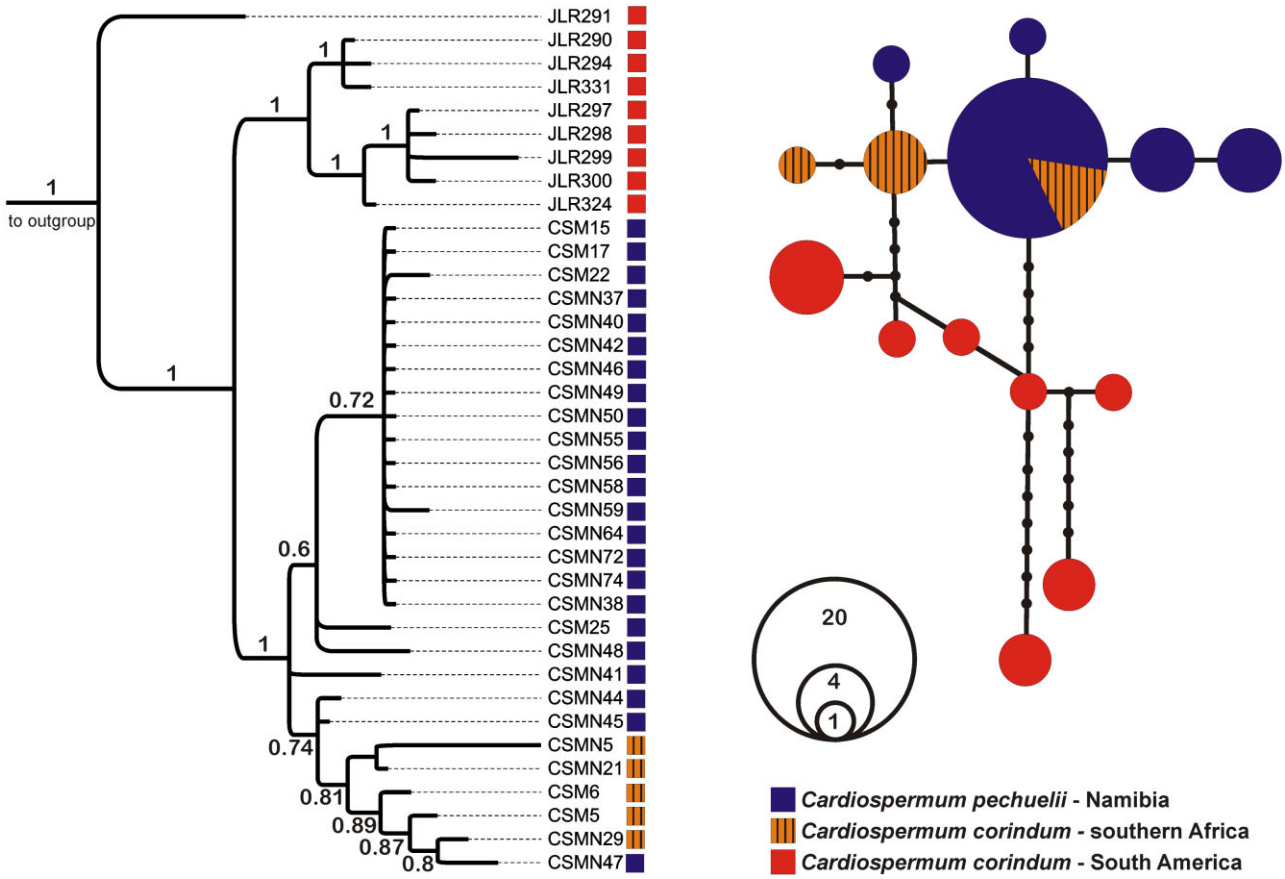

Figure 3. Bayesian phylogeny of combined nDNA and plastid DNA sequences indicating the relationships among *Cardiospermum corindum* from South America and southern Africa and *C. pechuelii* from Namibia (left hand side). Branch support is given as posterior probabilities and accessions colour coded according to species and geographic affiliations. A haplotype network (right hand side) shows relationships for these species based on only the plastid DNA. The sizes of the circles in the network are proportional to the number of individuals sharing each haplotype.

Table 1. AMOVA results indicating the distribution of genetic variation within and among populations of *Cardiospermum corindum* from southern Africa and South America and *C. pechuelii* from Namibia

	SS	MS	Est. Var.	%
Among populations	70.2	35.1	2.6	63%
Within populations	65.1	1.5	1.5	37%

(number of effective alleles, Shannon's information index, expected heterozygosity, number of PA and percentage polymorphic loci) indicated the highest level of diversity for *C. corindum* from South America and lowest for African samples (Table 2). Pooled African samples (*C. corindum* and *C. pechuelii*) still harboured substantially lower genetic diversity than South American samples (Table 3). Nei's pairwise genetic distances (Nei, 1978) not only show that genetic divergence is higher between South American and southern African *C. corindum* than between southern African *C. corindum* and *C. pechuelii*, but also that differentiation is lower between *C. pechuelii* and South American *C. corindum* than between *C. corindum* from southern Africa and South America (Table 4).

DISCUSSION

The genetic relationships within the *pechuelii-corindum* group are consistent with a model of speciation in response to novel environmental conditions following a single long-distance transoceanic dispersal. *Cardiospermum pechuelii*, a desert endemic from Namibia, arose after *C. corindum* dispersed from South America to southern Africa between 5.9 and 15.1 Ma (Gildenhuys *et al.*, 2014). The drier conditions of the Namib Desert probably led to strong directional selection and, in the presence of genetic

Table 2. Genetic diversity indices based on AFLP analyses for *Cardiospermum corindum* from southern Africa and South America and *C. pechuelii* from Namibia**.**

	Ne	*I*	*He*	P_A	*PP*
C. pechuelii	1.030	0.035	0.019	3	21.95%
C. corindum southern Africa	1.074	0.061	0.041	1	12.20%
C. corindum South America	1.290	0.285	0.182	21	65.85%

Effective number of alleles (*Ne*), Shannon's information index (*I*), expected heterozygosity (*He*), number of private alleles (P_A) and % polymorphic loci (*PP*) are shown.

Table 3. Genetic diversity indices based on AFLP analyses for *Cardiospermum corindum* and *C. pechuelii*, grouping all African samples as a single population and South American *Cardiospermum corindum* samples as another**.**

	Ne	*I*	*He*	P_A	*PP*
Southern African samples	1.056	0.069	0.038	6	34.15%
C. corindum South America	1.290	0.285	0.182	21	65.85%

Effective number of alleles (*Ne*), Shannon's information index (*I*), expected heterozygosity (*He*), number of private alleles (P_A) and % polymorphic loci (*PP*) are shown.

Table 4. Pairwise population matrix of Nei's genetic distance based on AFLP analyses for *Cardiospermum corindum* from southern Africa and South America and *C. pechuelii* from Namibia

	C. pechuelii	*C. corindum* southern Africa	*C. corindum* South America
C. pechuelii	0.000		
C. corindum southern Africa	0.050	0.000	
C. corindum South America	0.141	0.153	0.000

drift (as evidenced by reduced genetic diversity in African taxa Tables 2 and 3), probably facilitated diversification of a morphologically and ecologically unique taxon. Although our data cannot distinguish the action of selection from drift, the low genetic differentiation between African *C. corindum* and *C. pechuelii* observed here (Tables 2 and 3) favours selection over isolation and genetic drift as the causal mechanism for the strikingly different morphologies and habitat of these two taxa (Fig. 1). Moreover, African *C. corindum* is genetically more closely related to the morphologically divergent *C. pechuelii* than to morphologically similar South American *C. corindum* (Figs 2, 3).

The mixed clade including southern African balloon vines is well supported (PP = 1) with two further clades that are moderately supported (Fig. 3). Again, for this African clade, *C. pechuelii* renders African *C. corindum* paraphyletic, i.e. some *C. pechuelii* accessions are phylogenetically more closely related to African *C. corindum* than some African *C. corindum* accessions are to each other. These relationships, along with population genetic structure based on AFLP data, possibly indicate that reproductive isolation is incomplete between these taxa and that gene flow is still ongoing. However, given their recent divergence, coupled with geographical isolation and distinct morphologies, we cannot dismiss the possibility of incomplete lineage sorting (the failure of gene copies to coalesce within the duration of the species) as an equally parsimonious explanation for the genetic patterns observed here. The placement of the *C. pechuelii* accession, CSMN47, within the African *C. corindum* plastid DNA clade (Fig. 3) also suggests that past hybridization between African *C. corindum* and *C. pechuelii* may have led to introgression and possible plastid capture. Taken together these results indicate that the evolution of *C. pechuelii* in Namibia probably occurred very recently and it is therefore conceivable that the extreme long-distance dispersal event of *C. corindum* from South America to Africa occurred towards the lower end of the 5.9–15.1 Ma estimation of Gildenhuys *et al*. (2014). Our phylogenetic analysis also illustrates the need for the revision of alpha taxonomy of the *pechuelii-corindum* group. Specifically, *C. pechuelii* should be treated as an

African desert ecotype/variant or lineage of *C. corindum* if species delimitations are based on phylogenetic monophyly.

The recent divergence of African *C. pechuelii* and *C. corindum* and their relationship to South American *C. corindum* have important implications for ongoing biological control against invasive balloon vine species in southern Africa (Gildenhuys *et al*., 2014). First, our study highlights the need for sound taxonomy prior to implementing biological control programmes. Host-specificity trials of biological control agents currently considered for release in South Africa used mostly South American provenances of *C. corindum* (Mc Kay *et al*., 2010; Simelane *et al*., 2014). Given its genetic distinctiveness, it is evident that African *C. corindum* should be included as a separate taxon in host-specificity trials and that trials on South American provenances may not accurately predict host preferences in Africa. The complex provenances of balloon vines identified here also pose a unique opportunity to shift focus from classical to indigenous avenues of biological control, i.e. seeking and testing fauna associated with indigenous balloon vines as potential control agents against introduced balloon vines. Promoting native natural enemies to control non-native taxa probably has much untapped potential, but has hitherto been largely unexplored. Indigenous biological control may open the door to better understanding of how eco-evolutionary management may influence the structure of local and regional communities in general. For example, Andres *et al*. (2013) recently identified a host shift of hybrid lineages of native Australian *Leptocoris* soapberry bugs onto invasive balloon vine (*C. halicacabum*) in Australia. Parental, non-hybrid lineages of these soapberry bugs usually prefer distantly related native plant hosts (Andres *et al*., 2013).

The peculiar distribution and evolutionary history of African balloon vines also presents a system for testing and exploring numerous evolutionary hypotheses such as species concepts. Species concepts have been debated for decades and are still disputed with *c.* 26 different species concept definitions (Wilkins, 2011). The lack of agreement between scientists on species concepts often poses problems for identifying species. For example, a quick glance at *C. corindum* from South America and southern Africa is unlikely to reveal any immediate morphological differences, but to tease this relationship apart properly in-depth morphological and genetic analyses will have to be incorporated. Similarly, despite marked morphological differences between *C. pechuelii* and *C. corindum* our genetic results indicate recent diversification and therefore outcrossing experiments could add valuable information for understanding how complete reproductive isolation is between these two taxa.

The evolutionary diversification of *C. pechuelii* from *C. corindum* is also a remarkable example of the extent to which diversification can occur, from humid subtropical to desert environments. This may not be unexpected, as others have found species of *Cardiospermum* and *Paullinia* L. to have significantly higher evolutionary rates compared to other genera in Sapindaceae (Harrington, 2008). Our study also adds to a growing body of evidence highlighting the importance of extreme long-distance dispersal in shaping regional biodiversity and how evolution can be triggered given novel environmental conditions.

ACKNOWLEDGEMENTS

We thank David Simelane and Andries Fourie who kindly provided us with *Cardiospermum* DNA samples from South America and Kate Le Roux, Hannes Wiese, Mandelise van Loggerenberg, Ralph Riley and Jenella Loye for field assistance. We are grateful for sampling permits from the New South Wales Department of Environment and Conservation and Queensland Government Department of Environment and Resource Management (Australia) and the Ministry of Environment and Tourism (Namibia). We thank the Australian Tropical Herbarium and Australian National Herbarium for herbarium specimens. Financial support was provided by the DST-NRF Centre of Excellence for Invasion Biology and the Working for Water Programme through their collaborative project on 'Research for Integrated Management of Invasive Alien Species'. E. G. acknowledges the South African National Research Foundation's (NRF) Scarce Skills scholarship programme. J. L. R. also acknowledges Stellenbosch University's Subcommittee B and the NRF Thuthuka Programme for research funding. S. C. acknowledges support from the School of Life Sciences at the University of Queensland, St. Lucia.

REFERENCES

Akaike H. 1973. Akaike Information theory as an extension of the maximum likelihood principle. In: Petrov BN, Csaki F, eds. *Second international symposium on information theory*. Budapest: Akademiai Kiado, 267–281.

Andres JA, Thampy PR, Mathieson MT, Loye J, Zalucki MP, Dingle H, Carroll SP. 2013. Hybridization and adaptation to introduced balloon vines in an Australian soapberry bug. *Molecular Ecology* **22:** 6116–6130.

Arthofer W. 2010. tinyFLP and tinyCAT: software for automatic peak selection and scoring of AFLP data tables. *Molecular Ecology Resources* **10:** 385–388.

Austerlitz F, Jung-Muller B, Godelle B, Gouyon P-H. 1997. Evolution of coalescence times, genetic diversity and structure during colonization. *Theoretical Population Biology* **51:** 148–164.

Baldwin BG, Sanderson MJ. 1998. Age and rate of diversification of the Hawaiian silversword alliance (Compositae). *Proceedings of the National Academy of Sciences of the United States of America* **95:** 9402–9406.

Blignaut M, Ellis AG, Le Roux JJ. 2013. Towards a transferable and cost-effective plant AFLP protocol. *PLoS ONE* **8:** e61704.

Carlquist S, Baldwin BG, Carr GD. 2003. *Tarweeds and silverswords: evolution of the Madiinae (Asteraceae)*. Saint Louis, MO: Missouri Botanical Garden Press.

Carson HL, Templeton AR. 1984. Genetic revolutions in relation to speciation phenomena: the founding of new populations. *Annual Review of Ecology, Evolution, and Systematics* **15:** 97–131.

Christenhusz MJM, Chase MW. 2013. Biogeographical patterns of plants in the Neotropics – dispersal rather than plate tectonics is most explanatory. *Botanical Journal of the Linnean Society* **171:** 277–286.

Clement M, Posada D, Crandall KA. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9:** 1657–1660.

Doyle J, Doyle J. 1990. Isolation of plant DNA from fresh tissue. *Focus (San Francisco, Calif.)* **12:** 13–15.

Earl DA, von Holdt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4:** 359–361.

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

Falush D, Stephens M, Pritchard JK. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes* **7:** 574–578.

García-Verdugo C, Fay MF. 2014. Ecology and evolution on oceanic islands: broadening the botanical perspective. *Botanical Journal of the Linnean Society* **174:** 271–275.

Gildenhuys E, Ellis AG, Carrol SP, Le Roux JJ. 2014. Combining known native range distributions and phylogeny to resolve biogeographic uncertainties of balloon vines (*Cardiospermum*, Sapindaceae). *Diversity and Distributions* **21:** 163–174.

Gildenhuys E, Ellis AG, Carroll S, Le Roux JJ. 2013. The ecology, biogeography, history and future of two globally important weeds: *Cardiospermum halicacabum* Linn. and *C. grandiflorum* Sw. *NeoBiota* **19:** 45–65.

Givnish TJ, Millam KC, Evans TM, Hall JC, Pires JC, Berry PE, Sytsma KJ. 2004. Ancient vicariance or recent long-distance dispersal? Inferences about phylogeny and South American-African disjunctions in Rapateaceae and Bromeliaceae based on *ndhF* sequence data. *International Journal of Plant Sciences* **165:** S35–S54.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41:** 95–98.

Hardy OJ, Vekemans X. 2002. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* **2:** 618–620.

Harrington M. 2008. *Phylogeny and evolutionary history of Sapindaceae and* Dodonaea. PhD Thesis, James Cook University.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30:** 772–780.

Knope ML, Morden CW, Funk VA, Fukami T. 2012. Area and the rapid radiation of Hawaiian *Bidens* (Asteraceae). *Journal of Biogeography* **39:** 1206–1217.

Le Roux JJ, Strasberg D, Rouget M, Morden C, Koordom M, Richardson DM. 2014. Relatedness defies biogeography: the tale of two island endemics (*Acacia koa* and *A. heterophylla*). *New Phytologist* **204:** 230–242.

Ley AC, Hardy OJ. 2013. Improving AFLP analysis of large scale patterns of genetic variation – a case study with the Central African lianas *Haumania* spp (Marantaceae) showing interspecific gene flow. *Molecular Ecology* **22:** 1984–1997.

Mc Kay F, Oleiro M, Fourie A, Simelane D. 2010. Natural enemies of balloon vine *Cardiospermum grandiflorum* (Sapindaceae) in Argentina and their potential use as biological control agents in South Africa. *International Journal of Tropical Insect Science* **30:** 67–76.

Nei M. 1978. Estimation of average heterozygosity and genetic distance for small number of individuals. *Genetics* **89:** 583–590.

Nettel A, Dodd RS. 2007. Drifting propagules and receding swamps: genetic footprints of mangrove recolonization and dispersal along tropical coasts. *Evolution* **61:** 958–971.

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

Posada D. 2008. jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution* **25:** 1253–1256.

Posada D, Crandall KA. 2001. Intraspecific phylogenetics: trees grafting into networks. *Trends in Ecology and Evolution* **16:** 37–45.

Price JP, Wagner WL. 2004. Speciation in Hawaiian angiosperm lineages: cause, consequence, and mode. *Evolution* **58:** 2185–2200.

Renner S. 2004. Plant dispersal across the tropical Atlantic by wind and sea currents. *International Journal of Plant Sciences* **165:** S23–S33.

Ronquist F, Huelsenbeck JP. 2003. Mr Bayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics (Oxford, England)* **19:** 1572–1574.

Shaw J, Lickey EB, Schilling EE, Small RL. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare. *American Journal of Botany* **94:** 275–288.

Simelane DO, Mawela KV, Mc Kay F, Oleiro M. 2014. Field and laboratory studies to determine the suitability of *Cissoanthonomus tuberculipennis* (Curculionidae) for

release against *Cardiospermum grandiflorum* in South Africa. *Biocontrol Science and Technology* **24:** 734–750.

Taberlet P, Gielly L, Pautou G, Bouvet J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* **17:** 1105–1109.

Takayama K, Tamura M, Tateishi Y, Webb EL, Kajita T. 2013. Strong genetic structure over the American continents and transoceanic dispersal in the mangrove genus *Rhizophora* (Rhizophoraceae) revealed by broad-scale nuclear and chloroplast DNA analysis. *American Journal of Botany* **100:** 1191–1201.

Templeton AR. 2008. The reality and importance of founder speciation in evolution. *Bioessays: News and Reviews in Molecular, Cellular and Developmental Biology* **30:** 470–479.

Tremetsberger K, Weiss-Shneeweiss H, Stuessy T, Samuel R, Kadlec G, Orits MA, Talavera S. 2005. Nuclear ribosomal DNA and karyotypes indicate a NW African origin of South American *Hypochaeris* (Asteraceae, Cichorieae). *Molecular Phylogenetics and Evolution* **35:** 102–116.

Wagner WL, Funk VA. 1995. *Hawaiian biogeography: evolution on a hot spot archipelago*. Washington, DC: Smithsonian Institution Press.

White TJ, Bruns T, Lee S, Taylor J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ, eds. *PCR protocols – a guide to methods and applications*. San Diego, CA: Academic Press, 315–322.

Whittaker RTJ, Jones SH, Partomihardjo T. 1997. The rebuilding of an isolated rain forest assemblage: how disharmonic is the flora of Krakatau? *Biodiversity and Conservation* **6:** 1671–1696.

Wilkins JS. 2011. Philosophically speaking, how many species concepts are there? *Zootaxa* **1765:** 58–60.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Table S1. Species, sample ID, geographical origin, site, latitude, longitude and GenBank accession numbers of samples used in this study.