Annals of Botany **133**: 851–869, 2024
https://doi.org/10.1093/aob/mcad084, available online at www.academic.oup.com/aob

ANNALS OF BOTANY
Nonprofit since 1887

SPECIAL ISSUE: AFRICAN FLORA IN A CHANGING WORLD

Diploid and tetraploid cytotypes of the flagship Cape species *Dicerothamnus rhinocerotis* (Asteraceae): variation in distribution, ecological niche, morphology and genetics

Zuzana Chumová[1,*], Zafar Monier[2], Kristýna Šemberová[1], Eliška Havlíčková[1,3], Douglas Euston-Brown[4], A. Muthama Muasya[2], Nicola G. Bergh[2,5] and Pavel Trávníček[1,*]

[1]Institute of Botany of the Czech Academy of Sciences, Zámek 1, Průhonice, 252 43, Czech Republic, [2]Bolus Herbarium, Department of Biological Sciences, University of Cape Town, Cape Town, 7707, South Africa, [3]Department of Botany, Faculty of Science, Charles University, Benátská 2, Prague, 120 00, Czech Republic, [4]Independent botanist, Scarborough, 7975, Cape Town, South Africa and [5]The Compton Herbarium, Kirstenbosch National Botanical Gardens, Cape Town, 7735, South Africa
[*]For correspondence. E-mail zuza.chumova@gmail.com or Pavel.Travnicek@ibot.cas.cz

Received: 29 April 2023 Returned for revision: 16 June 2023 Editorial decision: 26 June 2023 Accepted: 4 July 2023

• **Background and Aims** The Greater Cape Floristic Region is one of the world's biodiversity hotspots and is considered poor in polyploids. To test this assumption, ploidy variation was investigated in a widespread Cape shrub, *Dicerothamnus rhinocerotis* (renosterbos, Asteraceae). The aim was to elucidate the cytotype distribution and population composition across the species range, and to assess differences in morphology, environmental niches and genetics.
• **Methods** Ploidy level and genome size were determined via flow cytometry and cytotype assignment was confirmed by chromosome counting. Restriction site-associated DNA sequencing (RADseq) analyses were used to infer genetic relationships. Cytotype climatic and environmental niches were compared using a range of environmental layers and a soil model, while morphological differences were examined using multivariate methods.
• **Key Results** The survey of 171 populations and 2370 individuals showed that the species comprises diploid and tetraploid cytotypes, no intermediates and only 16.8 % of mixed populations. Mean 2C values were 1.80–2.06 pg for diploids and 3.48–3.80 pg for tetraploids, with very similar monoploid genome sizes. Intra-cytotype variation showed a significant positive correlation with altitude and longitude in both cytotypes and with latitude in diploids. Although niches of both cytotypes were highly equivalent and similar, their optima and breadth were shifted due to differences mainly in isothermality and available water capacity. Morphometric analyses showed significant differences in the leaves and corolla traits, the number of florets per capitulum, and cypsela dimensions between the two cytotypes. Genetic analyses revealed four groups, three of them including both cytotypes.
• **Conclusions** *Dicerothamnus rhinocerotis* includes two distinct cytotypes that are genetically similar. While tetraploids arise several times independently within different genetic groups, morphological and ecological differences are evident between cytotypes. Our results open up new avenues for questions regarding the importance of ploidy in the megadiverse Cape flora, and exemplify the need for population-based studies focused on ploidy variation.

Key words: Asteraceae, Compositae, *Elytropappus rhinocerotis*, flow cytometry, Gnaphalieae, ploidy level, RADseq, renosterbos, renosterveld, South Africa, *Stoebe* clade.

INTRODUCTION

Polyploidy, a state of bearing more than one complete set of chromosomes, is one of the key factors affecting and shaping plant diversity on Earth (Tank *et al*., 2015). It is estimated that almost all plant lineages in the world have undergone several polyploid events in their evolution, which are reflected in the construction of their genomes (e.g. Soltis and Soltis, 2016; Ren *et al*., 2018). A significant proportion of plants, already bearing the remnants of ancient polyploidization, are experiencing the effects of recent polyploidization, which is evidenced by the high ploidy-level variation in some families, genera and species. The geographic distribution of recent polyploidy on Earth is not uniform, but shows a distinct clinal pattern where the representation of polyploids increases from the equator to the poles (Rice *et al*., 2019).

One exception to this rule is considered to be the Cape flora, the megadiverse and highly endemic-rich flora of the south-western tip of the southern African subcontinent, which has been found to be more polyploid-poor than expected from its geographic location (Oberlander *et al*., 2016). The Greater Cape Floristic Region (GCFR), which is climatically characterized by winter rainfall and summer drought, is composed of two main biomes, Fynbos and Succulent Karoo (Bergh *et al*., 2014). The Succulent Karoo is associated with drier and warmer

conditions, and is found on plains and lower slopes with shale- and granite-derived soils and annual rainfall between 20 and 300 mm and summer temperatures reaching up to 44 °C. In contrast, Fynbos is found on sandy lowland coastal plains as well as mountains, but not in areas where annual rainfall is <200 mm, with vegetation types adapted to the different edaphic characteristics, including nutrient-poor quartzite-derived substrates and shale- or granite-derived loamy mesotrophic substrates (Low *et al.*, 1996; Mucina *et al.*, 2006; Rebelo *et al.*, 2006). The Fynbos biome is adapted to fires, which drive ecological processes such as regeneration, succession and vegetation dynamics (Bond and van Wilgen, 1996; Keeley *et al.*, 2012).

The Cape's polyploid paucity is usually explained as a consequence of its stable climatic and geological history (Dynesius and Jansson, 2000; Cowling *et al.*, 2015) and the fact that much of the local plant diversity is confined to a limited number of radiated plant lineages (Linder, 2003; Oberlander *et al.*, 2016). On the other hand, there is increasing evidence that cytotype variation is important in many plant groups that have diversified in the Cape, including some of the 'Cape clades' *sensu* Linder (2003): African Restionaceae (Linder *et al.*, 2017), *Oxalis* (Oxalidaceae; Krejčíková *et al.*, 2013*a*, *b*), *Heliophila* (Brassicaceae; Mandáková *et al.*, 2012; Dogan *et al.*, 2021), *Helichrysum* (Asteraceae tribe Gnaphalieae; Andrés-Sánchez *et al.*, 2019) and *Pteronia* (Asteraceae tribe Astereae; Chumová *et al.*, 2022). Only detailed and population-targeted screening can elucidate cytotype diversity within species and populations, which is necessary for estimating the frequency of recent polyploidy but often not detected when just a few individuals are surveyed. However, few such studies have been carried out in the Cape, although several plant families known for their frequent polyploidy in other parts of the world are significantly represented here. A notable example is the family Asteraceae, by far the largest family in the Cape (Manning and Goldblatt, 2012) and with many independently radiated Cape lineages (e.g. in tribes Arctotideae, Senecioneae and Gnaphalieae, amongst others). Within Asteraceae globally, ancient polyploidization is known to be associated with diversification (Huang *et al.*, 2016), and current species diversity harbours a high proportion of recent polyploids (Semple and Watanabe, 2009).

Dicerothamnus rhinocerotis (L.f.) Koek. (formerly *Elytropappus rhinocerotis* Less.; Koekemoer, 2019) is a member of one of Linder's (2003) 'Cape Floral clades' and part of the Cape '*Stoebe* clade' of Bergh and Linder (2009). The species is an endemic Cape member of the Asteraceae tribe Gnaphalieae (subtribe Gnaphaliinae) and is a shrub with minute, cupressoid, tightly adpressed leaves on slender wiry twigs, bearing masses of cryptic flowers in tiny brown flowerheads. Commonly known as 'renosterbos' ('rhinoceros bush'), this species is one of the dominant members of, and gives its name to, the Mediterranean-type shrubland known locally as 'renosterveld', one of the most highly threatened vegetation types (Kemper *et al.*, 1999; Cowling and Heijnis, 2001) within the megadiverse GCFR and one of the most threatened vegetation types globally. While renosterveld is part of the Fynbos biome, it differs from the typical Fynbos vegetation type in occurring on more nutrient-rich soils, so the formations are bordered by Fynbos on one side, in other places they grade into Succulent Karoo, especially at the lower end of the rainfall spectrum. Renosterveld has received notably less focus than other fragmented ecosystems, despite higher levels of habitat loss than other highly biodiverse regions (Topp and Loos, 2019). Similarly, the renosterbos, *D. rhinocerotis*, has received little taxonomic or ecological attention since the first half of the last century, despite being an important ecological indicator species with a strong palaeopollen record (e.g. Chase and Meadows, 2007) and documented ethnopharmacological uses (e.g. Dekker *et al.*, 1988; Maroyi, 2019) as well as being the flagship species of a major southern African vegetation type. The distribution of this species ranges from Cape Town in the south-western Cape, northwards, eastwards and inland outside of the GCFR into the higher-elevation cool growing-season grasslands of the Great Escarpment of South Africa. In addition to this apparently very broad environmental tolerance, renosterbos plants have a large reproductive output, producing >300 fruits on just one small twig 7.5 cm long (Levyns, 1926), and with disturbance (usually fire)-mediated recruitment from seedbanks (Levyns, 1929), giving it the ability to form monocultures in inappropriately managed veld. This has been noted for >100 years (e.g. Marloth, 1908) and can be a problem for farmers, as the plant is not edible by stock.

Initial population screening to estimate genome size variation within *D. rhinocerotis* demonstrated striking variation, resembling that found in mixed ploidy systems. Here, we further examine multiple individuals from numerous populations across the full geographic range of the species in order to elucidate the cytotype and genetic composition of populations and reveal the spatial distribution of variation in these factors. We also explore the nature and maintenance of cytotype variation by assessing differentiation between cytotypes in genetics, morphology and environmental niche.

MATERIALS AND METHODS

Plant material

A total of 171 *Dicerothamnus rhinocerotis* populations were sampled during 2014–22 (Fig. 1, Supplementary Data Table S1) over a substantial part of its distribution range in South Africa. At each locality, the following material was collected for each population, where possible: (1) well-developed intact leaf tissue from 1–20 individuals (depending on the population size) was stored at 4–8 °C in a plastic bag for flow cytometry estimation of ploidy level and nuclear genome size, flow cytometry analyses being performed within 14 days of collection, usually sooner; (2) from one to eight of the individuals sampled in (1) above, twigs were collected for morphometric investigations, including both flowers and fruit where available (sampled during the flowering season); and (3) mature achenes from several individuals for germination in the experimental garden of the Institute of Botany of the Czech Academy of Sciences in Průhonice were studied for karyology. Vouchers of the above material have been deposited in the PRC and NBG herbariums.

Flow cytometry

Flow cytometry was used to examine DNA ploidy level (*sensu* Suda *et al.*, 2006) for all collected leaf samples. The

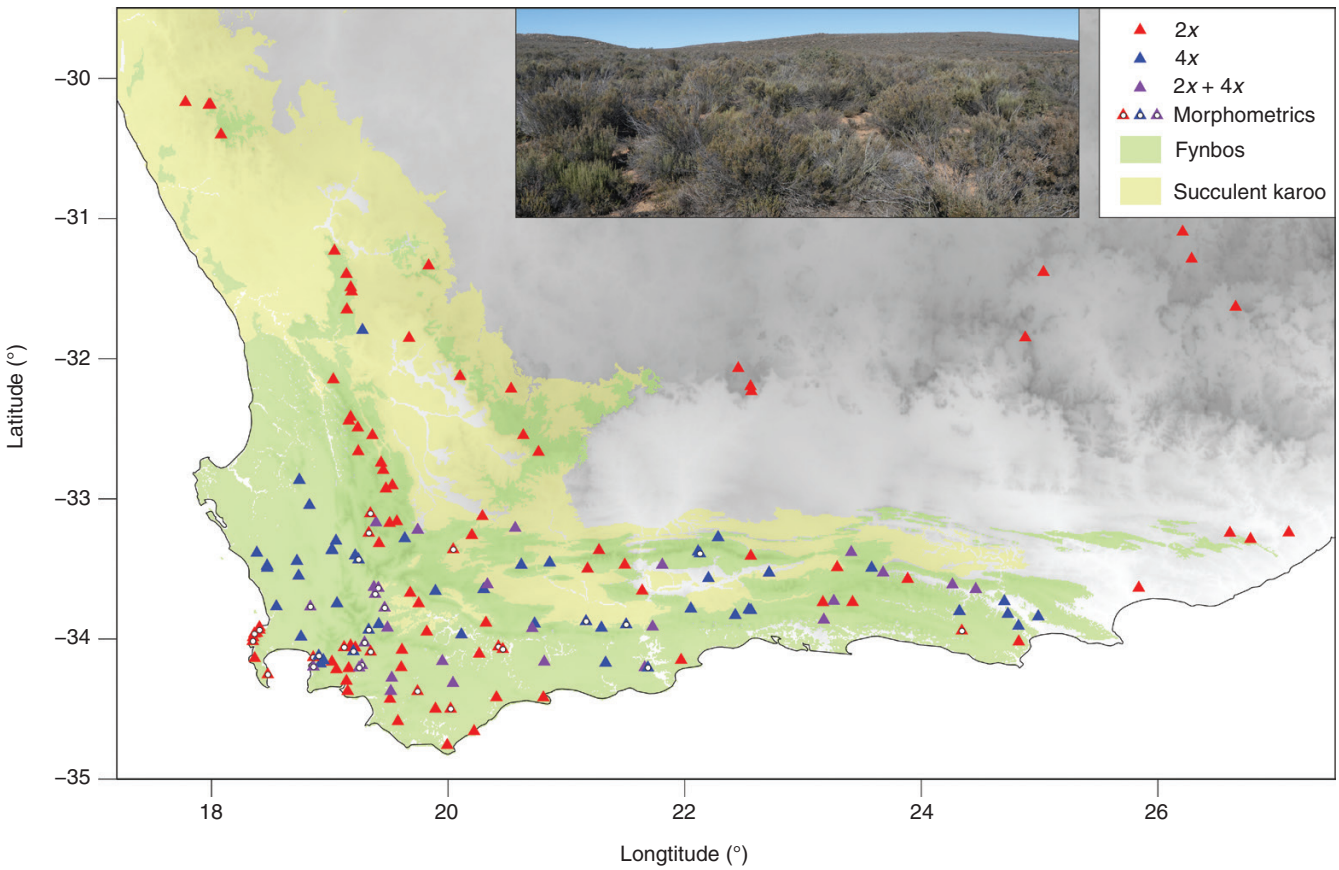

FIG. 1. Geographic distribution of cytotypes across the distribution range of *D. rhinocerotis* on an overlay showing elevation and the distribution of the Fynbos and Succulent Karoo biomes. Sites where only a single cytotype was are indicated in red for diploids and blue for tetraploids. Sites where diploids and tetraploids were found to co-occur are indicated in purple. Open triangles indicate sites sampled for morphometrics. The inset shows a typical image of the vegetation type renosterveld dominated by *D. rhinocerotis* (large mixed-ploidy population at the foothills of the Swartberg).

ploidy determination followed the simplified protocol of Otto (according to Doležel *et al.*, 2007) with the usage of DAPI (4′,6-diamidino-2-phenylindole at the final concentration of 4 µg · mL^{-1}) as a staining dye and *Solanum pseudocapsicum* as a primary internal standard (2C = 2.59 pg; Temsch *et al.*, 2010). Occasionally, one individual of *Hedera helix* served as a secondary internal standard (2C = 2.88 pg); its genome size was estimated in 15 independent measurements against the primary internal standard. A subset of individuals (64 plants from 29 populations plus 3 individuals of the closest related species *Dicerothamnus adpressus*) was also subjected to genome size estimation with the same sample preparation, but after incubation the suspension was stained with 1 mL of Otto II buffer (0.4 M $Na_2HPO_4 \cdot 12H_20$) supplemented with the intercalating fluorescent dye propidium iodide, RNAase IIA (both at the final concentration of 50 µg · mL^{-1}) and β-mercaptoethanol (2 µg · mL^{-1}).

Ploidy level was determined on a Partec CyFlow ML cytometer equipped with an LED UV chip, and genome size on a Partec CyFlow SL cytometer equipped with a 532-nm diode-pumped solid-state laser as a source of excitation light (both manufactured by Partec, Münster, Germany). Both approaches were used in accordance with best practices in plant flow cytometry (Sliwinska *et al.*, 2022).

Chromosome preparation and fluorescent in situ hybridization

Actively growing young roots were collected from seedlings germinated in Petri dishes. The root tips were pretreated with ice-cold water for 12 h, fixed in freshly prepared fixative (ethanol:acetic acid, 3:1) for 24 h at 4 °C and stored at −20 °C until further use. Chromosome spreads were prepared for 12 individuals as described by Mandáková and Lysak (2016). Preparation of the slides for fluorescent *in situ* hybridization (FISH) and the procedure itself was done according to Garcia *et al.* (2010).

Slides were examined and photographed on a Zeiss Axio Imager Z2 microscope system equipped with an ApoTome 2 Zen (Zeiss, Jena, Germany) and Adobe Photoshop software (Adobe Systems, San Jose, CA, USA) was used for processing the images.

DNA extraction

Total genomic DNA was extracted from ~0.5 g of leaf material taken from the herbarium (PRC) specimens by the sorbitol method (Štorchová *et al.*, 2000). The samples were cleaned with SPRI beads (Beckman Coulter Genomics, Danvers, MA, USA; ratio 1:0.4 DNA:beads) and the quality of the DNA was checked on 1 % agarose gel.

Generation of restriction site-associated DNA sequence data

Altogether 35 populations (91 individuals) of *D. rhinocerotis* and 3 populations (5 individuals) of *D. adpressus* as an outgroup (Supplementary Data Table S1) were used for library preparation. Genomic DNA from each sample was quantified using a Qubit 2.0 fluorometer (Invitrogen, Carlsbad, CA, USA) and diluted to the same concentration for all individuals. An initial amount of 500 ng DNA was used in a total volume of 30 µL. Construction of a ddRADseq library (following Peterson *et al.,* 2012) consisted of the following steps: digesting genomic DNA with two restriction enzymes, SphI and MluCI (New England BioLabs); ligating two different barcoded adaptors, P1 and P2 (Peterson *et al.*, 2012), onto the ends of 100 ng of the purified digestion products; size-selecting from the ligation products; and PCR-amplifying the remaining subset of fragments. After ligation of the adaptors, 48 ligation products differing in adapter barcode were pooled together, and cleaned with SPRI again (ratio 1:1.2 DNA:beads, two times), and size-selected fragments by Pippin Prep (Sage Science, Beverly, MA, USA) for the size-selection range 500–600 bp were amplified via multiple amplification reactions for each size-selected sample (Peterson *et al.*, 2012), purified twice with SPRI beads (ratio 1:1.5 DNA:beads), quantified on a Qubit fluorometer and pooled in equimolar concentrations. A 1.8 % agarose quality check gel was performed and final concentration was measured on a Bioanalyzer (Agilent, Santa Clara, CA, USA).

Sequencing was performed on an Illumina HiSeqX instrument at Macrogene (Seoul, Republic of Korea) using a 300-cycle kit (v.5, Illumina, San Diego, CA, USA) to obtain 150-bp paired-end reads. Raw sequence data are available in the NCBI BioProjects repository under accession number PRJNA962930.

Analysis of restriction site-associated DNA sequence data

The raw reads were quality-filtered and demultiplexed according to individual barcodes using the process_radtags.pl script implemented in Stacks (Catchen *et al.*, 2011). Despite our sampling aimed at diploid and tetraploid samples, we used the standard data analysis for diploid species using Stacks v.2.64 (Catchen *et al.*, 2011). This procedure has been shown to be efficient if closely related diploids and polyploids (presumably autopolyploids) are compared (e.g. Edgeloe *et al.*, 2022). Restriction site-associated DNA (RAD) loci were assembled and single nucleotide-polymorphisms (SNPs) were called *de novo*, using the denovo_map.pl pipeline. Settings parameters for the denovo_map.pl script were optimized on a pilot dataset including ten randomly chosen accessions according to Paris *et al.* (2017), and were set as follows: a minimum distance to identify a stack (-m 3), a maximum distance between stacks (M 3) and a maximum number of mismatches between loci of different individuals (-n 3). The 'populations' routine was used to extract and export the selected loci, filtered based on the number of heterozygous genotypes per locus; the maximum observed heterozygosity parameter (--max_obs_het) was set to 0.65 to avoid combining paralogues within the same RAD locus. Usage of only a single independent SNP per locus was guaranteed by implementing --write-random-snp flag. Data from the populations procedure was output to a vcf file using --vcf flag.

The vcf files were further processed using VCFtools (Danecek *et al.*, 2011), where the --minDP flag was used to select only genotypes with a given minimum depth, the --remove-indels flag to exclude sites with indels, and the --max-missing flag to exclude sites with a rate of missing data greater than the given threshold. For the Structure-like analyses, additional filtering for shared SNPs among a minimum number of individuals (--mac flag) was used. Subsequently, vcf files were converted to nexus format by the python script vcf2pyhilip.py (Ortiz, 2019), and/or read and further processed in R via the vcfR package (Knaus and Grünwald, 2017). Nexus files were used to explore possible reticulation via NeighborNet analysis implemented in SplitsTree4 (Huson and Bryant, 2006). The analysis with outgroup specimens was supplemented with bootstrap support inferred from 500 replicates. Resulting vcf files were processed in R by their transfer to standard genlight objects (vcfR2genlight function in vcfR) and further analysed via the adegenet R package (Jombart and Ahmed, 2011) which was used to conduct discriminant analysis of principal components (DAPC; Jombart *et al.*, 2010) and principal component analysis (PCA; via function glPca). For Structure-like analyses, the vcf file was transformed in PGDSpider v2.1.1.5 (Lischer and Excoffier, 2012) to Structure format. Structure analysis was conducted in Structure v.2.3.4 (Pritchard *et al.*, 2000) using 1 000 000 Markov chain Monte Carlo generations and a burn-in of 100 000, in ten replicates for each ancestral number of populations K from $K = 1$ to 15. Resulting data were analysed for optimal K using the approach of Evanno *et al.* (2005), and all ten replicates were merged into a single Q-matrix using the R package pophelper (Francis, 2017). The Structure input file was also used in sparse non-negative matrix factorization (sNMF) analysis algorithms (Frichot *et al.*, 2014) conducted in the R package LEA (Frichot and François, 2015), where the best-supported number of ancestral populations was estimated based on the cross-entropy criterion.

Individual assignment to genetic clusters was conducted with a user script in R via the barplot function based on Q matrices provided by the software (Structure and sNMF) or transformation of their outputs (DAPC). To visualize the genetic grouping assignment of individuals on NeighborNet, a user script was run in the phangorn R package (Schliep, 2011). To visualize genetic groupings at population level, a user script was run in the terra (Hijmans *et al.*, 2022) and marmap R packages (Pante and Simon-Bouhet, 2013). The GMTED 2010 topographic model of terrain (Danielson and Gesch, 2011) was used as a map base.

Niche modelling

To estimate potential differences in cytotype niche identity and breadth we modelled environmental features of the collection localities. Climate layers were obtained from the CHELSA database (Karger *et al.*, 2017) with a suite of Bioclim and derived variables in highest resolution (30 arc sec). In the first run, the full set of 19 Bioclim variables and additional climatic aggregations derived from monthly temperatures and precipitation characteristics were used, as well as topographic variables for the GCFR from Wüest *et al.* (2019). For the sake of simplicity, we have used a suite of GeoTIFF layers devoted to South African territory that are publicly available at the DRYAD

digital repository (https://doi.org/10.5061/dryad.1cs77qn). Soil characteristics were obtained from regionally modelled soil layers for the GCFR (Cramer *et al.*, 2019). Due to the availability of a complete set of environmental layers only for the GCFR and the fact that our collection was primarily focused on this area, niche modelling was performed only within this region. An overview of the environment variables used is available in Supplementary Data Table S2. In a first step, the georeferenced occurrence data for both cytotypes were spatially stratified to avoid uneven sampling across the distribution area. A 10-km radius was used for both cytotypes separately using the R package spThin (Aiello-Lammens *et al.*, 2015). Data from all raster layers were extracted using the extract function in the R package terra (Hijmans *et al.*, 2022) for georeferenced collection sites after transformation to the same coordinate system and spatial stratification. Pearson's correlation coefficient ($r < 0.7$) was used as a threshold to select uncorrelated layers for subsequent analyses (according to Dormann *et al.*, 2013; Supplementary Data Table S3). Our choice of niche modelling was based on comparison with background environmental data, so the next step required pseudo-absence data to set the global environment. Following recommendations (e.g. McCormack *et al.*, 2010; Warren *et al.*, 2010), the background area was taken from buffer zones around known occurrences or from distribution maps. Therefore, we used a 10-km buffer zone around the occurrence points of each cytotype to define the background area (using the R package geobuffer; Valentin, 2022). From the pseudo-distribution area thus defined, 100 times the number of occurrence points of each cytotype were randomly selected as background points and the same raster layer data were obtained. The background point datasets for both cytotypes were merged, cleaned of any duplicates, and then used in subsequent analyses.

Two different approaches were used to investigate the association of cytotype with environmental conditions. In the first approach, both cytotypes were subjected to quantification of niche overlap, equivalence and similarity, using an ordination technique that applies kernel smoothing to the presence of cytotypes in the environmental space using the R package ecospat (Broennimann *et al.*, 2012; Di Cola *et al.*, 2017). To interpret the niche characteristics, a PCA partitioning of the environmental space (first two axes) into a 500×500-cell grid was specified. The PCA output rasterized in this way was used for calculations of niche overlap rates (derived from Schoener's *D* statistic; Schoener, 1968), simulation tests of niche similarity and equivalence (Warren *et al.*, 2008) and niche optimum and niche width estimation (Theodoridis *et al.*, 2013; Kirchheimer *et al.*, 2016; Duchoslav *et al.*, 2021). A test of niche similarity was calculated bidirectionally to determine whether cytotypes differ from each other; equivalency was tested in order to test whether the niche overlap was more equivalent than would be expected by chance. Both tests were based on 1000 resamples simulating the null distribution, which was compared with observed Schoener's *D* as described in Molina-Henao and Hopkins (2019).

The second approach involved species distribution modelling (SDM) using Maxent v.3.4.4 (Phillips *et al.*, 2022) to define the potential spatial distribution of each cytotype, and their area of overlap. We used the defined background points from the first approach to achieve the same conditions for both approaches by a post-projection of the results onto a map. The cytotype distribution model was inferred from ten replicates of the SDM calculation for both cytotypes based on equal training sensitivity and specificity logistic threshold (ETSSLT; Liu *et al.* 2005). Cells of the average raster of suitable habitats for both cytotypes were treated in a binary manner, where all cells with a value below the ETSSLT for a given cytotype were replaced by 0, while all cells with a value above ETSSLT were replaced by 1. The potential co-occurrence of cytotypes was inferred by the overlay of these rasters.

All computations (except Maxent) and graphical outputs were conducted in R v.3.6.3 (R Core Team, 2021).

Multivariate morphometrics

Morphology was measured on a subset of flowering and fruiting diploid and tetraploid samples from both mixed- and single-cytotype populations mainly in the south-western Cape (Fig. 1, Supplementary Data Table S1). Macromorphological traits were measured using a calibrated Leica M50 stereomicroscope eyepiece micrometer. Each reported value represents the average of three measurements per specimen (e.g. three different leaves). An initial pilot study on a subset of specimens examined multiple traits that are generally known to correlate with ploidy (e.g. stomatal guard cell size; Levin, 2002; Beaulieu *et al.*, 2008; Hodgson *et al.*, 2010) or that we hypothesized, based on initial observations, to vary between ploidy levels. Micromorphological characters, i.e. stomatal guard cell dimensions, were examined using stomatal peels. Because the stomata-bearing (adaxial) leaf surface is densely covered with hairs in *D. rhinocerotis*, the peels were taken from the hairless adaxial surface of the middle involucral bracts, and imaged using a Leica DM300 compound microscope. Dimensions were obtained using ImageJ software (Schneider *et al.*, 2012). A final set of ten primary traits comprising stem, leaf, floret, fruit, and guard cell metrics was selected, from which additional ratios were calculated, including in particular pseudo-shape traits expressed as width/length ratios (Table 1, Fig. 2, Supplementary Data Table S4).

Prior to analyses, the traits were tested for normality (via base R function shapiro.test) and log-transformed if necessary. Pearson's correlation between traits was calculated using the corrplot R package (Murdoch and Chow, 1996; Friendly, 2002) as a decision criterion for retaining or excluding traits in multivariate analyses with sensitivity for trait collinearity ($|r| > 0.85$; i.e. discriminant analysis; e.g. Innangi *et al.*, 2020).

Morphometric data were analysed using PCA and canonical discriminant analysis using the R package morphotools (Koutecký, 2015). The main focus was on deterministic features associated with cytotype delineation. All data were analysed using R v.3.6.3 (R Core Team, 2021).

RESULTS

Ploidy variation and cytogeography

Two distinct genome size peaks, corresponding to two DNA ploidy levels, 2*x* and 4*x*, were detected among the 2370 analysed individuals from 171 populations (Fig. 1, Supplementary

TABLE 1. *Overview of recovered genetic groupings (colours as in Figs 4 and 5), their relationship to biomes, and their cytotype composition.*

	Genetic cluster				Level
	Group 1 (bright yellow)	**Group 2 (pale yellow)**	**Group 3 (orange)**	**Group 4 (brown)**	
Biome					
Albany Thicket	0	3	0	0	Populations
Fynbos	6	12	6	6	
Grassland	0	2	0	0	
Nama-Karoo	0	1	0	0	
Succulent Karoo	1	2	1	0	
Cytotype					
2*x*	4	25	9	9	Individuals
4*x*	14	24	0	2	

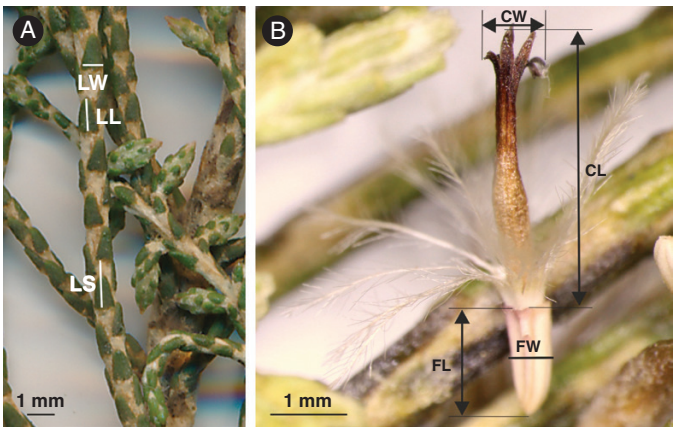

FIG. 2. (A) Twig segment and (B) floret of *D. rhinocerotis* showing how selected morphological traits were measured. CW, corolla width; CL, corolla length; FL, fruit length; FW, fruit width; LS, leaf spacing; LW, leaf width; LL, leaf length.

Data Fig. S1, Supplementary Data Table S1). No intermediate genome sizes were detected, although a low level of genome size variation is evident within each ploidy level (see below). The number of chromosomes counted for diploids was $2n = 2x = 16$, with one visible pair of linked 35S–5S rDNA loci on slides, and $2n = 4x = 32$ for tetraploids, with two visible pairs of 35S rDNA (Supplementary Data Table S1, Supplementary Data Fig. S1). Most of the investigated populations were ploidy uniform (93 diploid, 50 tetraploid), although we detected 29 mixed populations that comprised individuals of both cytotypes.

The more geographically widespread cytotype is the 2*x* *D. rhinocerotis*, which occurs throughout the species' range, and is the only cytotype detected outside of the GCFR in the Great Escarpment and in the northern GCFR (Fig. 1). The 2*x* cytotype is also the only one detected on the Cape Peninsula, in coastal habitats, and in the far south-east of the species range in the Eastern Cape Province. The tetraploid cytotype is much more geographically restricted, being confined to the southern parts of the range within the GCFR, and with only a few specimens detected north of the −33° latitude line. However, the 4*x* cytotype is extremely common from the south-western Cape coastal forelands eastwards throughout the Overberg and Little Karoo, in these areas co-occurring with the 2*x* cytotype at many localities.

Within-cytotype variation in nuclear genome size

A degree of variation in genome size was observed within each cytotype (Supplementary Data Fig. S1, Supplementary Data Table S1). Mean 2C values varied from 1.80 to 2.06 pg in diploids and from 3.48 to 3.80 pg in tetraploids. Diploid and tetraploid cytotypes shared very similar monoploid genome sizes (1Cx values) ranging from 0.44 to 0.52 pg, corroborated by non-significance of the ANOVA test: $F_{(1,32)} = 1.24$, $P = 0.273$. A very similar genome size was found in the closely related species *D. adpressus*, 2C = 1.97 ± 0.03 pg, corresponding to the range of diploids.

Intra-cytotype variation in genome size was compared with environmental variability in each environmental layer. Of the 45 raster layers, 18 showed significant linear correlations with genome size of just one of the cytotypes, while 15 showed significant correlation with genome size in both cytotypes (Supplementary Data Figs S2 and S3 for diploids and tetraploids, respectively; Supplementary Data Table S2). These significant correlations were almost always in the same direction for both cytotypes (e.g. potential evaporation is positively correlated with genome size in both cytotypes, while mean temperature of the driest quarter, bio9, is negatively correlated with genome size in both cytotypes; Supplementary Data Table S2). Intra-cytotype variation also showed significant positive correlation with altitude, latitude and longitude in diploids, and with altitude and longitude in tetraploids (Fig. 3).

Restriction site-associated DNA sequence assessment of population structure

After demultiplexing and filtering for raw reads, RADseq data averaged 3.47 million (s.d. 1.25 million) reads per sample. However, four samples were excluded due to a lower number of reads (below the threshold set at 1.5 million; Supplementary

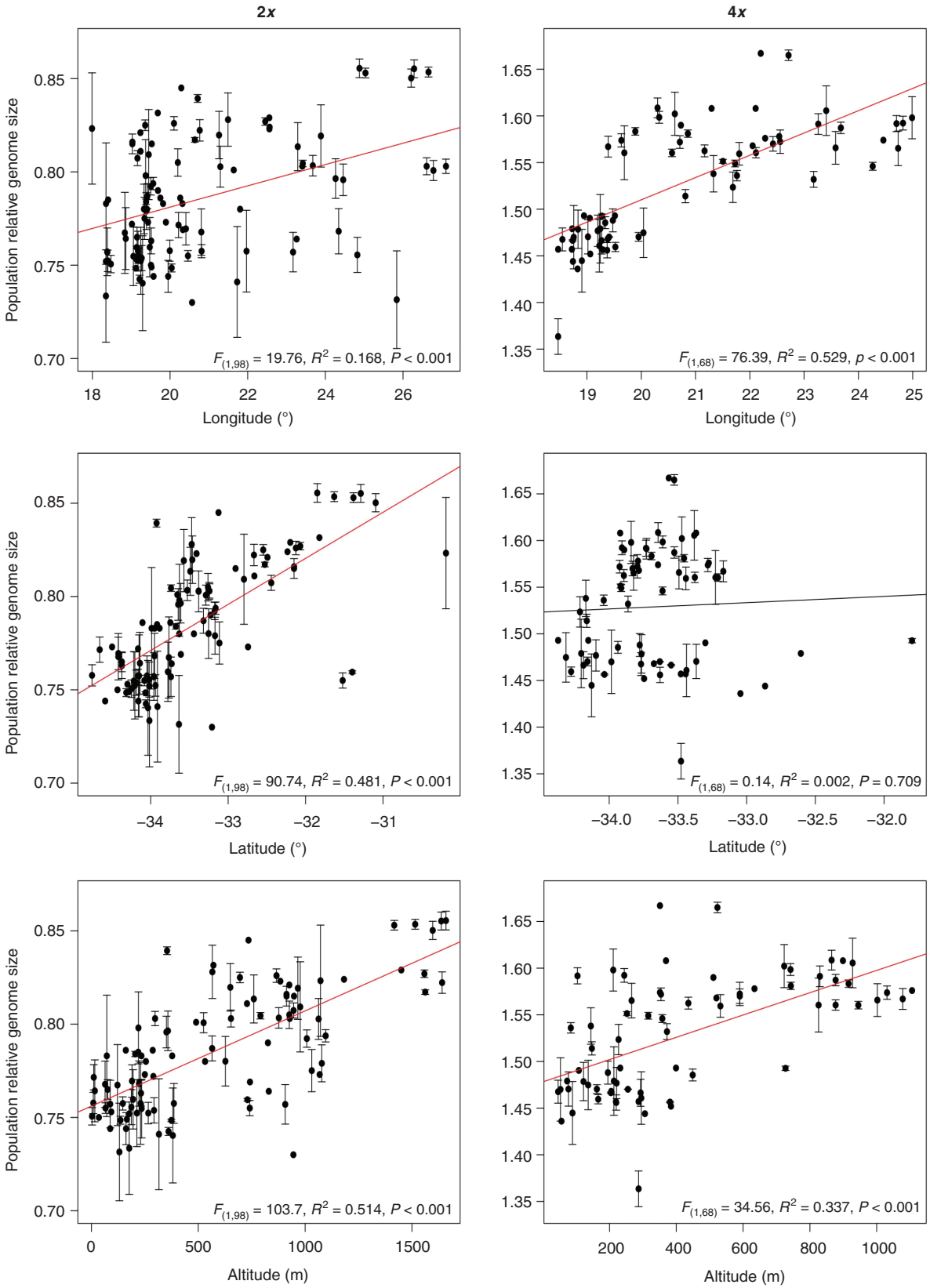

FIG. 3. Relationship between relative genome size and geographic data for populations of both cytotypes (left column for diploids, right column for tetraploids) of *D. rhinocerotis*. Mean values of relative genome size of the population (plotted with standard deviation) are compared with longitude (top row), latitude (middle row) and altitude (bottom row) and fitted with linear models. Statistically significant fits are shown by the red line.

Data Table S5). Thereafter, the mean number of reads increased to 3.58 ± 1.15 million (range 1.68–6.82 million) for all 92 included samples. An initial run of the *de novo* map STACKS pipeline was performed on all samples (including five *D. adpressus* individuals as an outgroup), yielding 5316 loci with 5308 variant SNPs that were shared by at least 80 % of individuals (-R flag of 0.8 in populations). Further filtering in vcftools for minimum depth of coverage (five reads per SNP) and maximum proportion of missing data per sample (20 %) left 4090 SNPs, which were used for determining the position of the outgroup and to identify genetic clusters within *D. rhinocerotis* (Fig. 4).

For the second set of data analyses, the outgroup specimens were removed (by simply removing specimen names from the popmap) before proceeding to the population-level analyses. This yielded 5431 loci with 5422 SNP variants. Filtering with vcftools with minimum depth of coverage (seven reads per SNP) and maximum missing data per locus (20 %) left 3489 SNPs. This dataset was used in the NeighborNet and PCA analyses. Due to the sensitivity of Structure-like analyses to non-shared (private) alleles, additional filtering for minimal SNP sharing between at least three individuals (--mac 3) resulted in 1047 SNPs. This third dataset was used in all three clustering analyses (DAPC, Structure, sNMF).

Initial analysis with outgroup specimens of *D. adpressus* showed a clear separation of this species from all samples of *D. rhinocerotis* (bootstrap support 100). Within the latter, there is closer relative proximity of clusters, with high mutual connectivity, evidenced by the NeighborNet network and only moderate bootstrap support (<100) for clusters of individuals (Fig. 4).

All of the analyses performed to reveal genetic structure at the individual/population level within *D. rhinocerotis* yielded very similar results (Fig. 5). There are at least three genetic clusters identified by Structure-like analyses. The fourth cluster clearly distinguished in the DAPC and sNMF analyses comprises intermediate individuals in the Structure analysis (Fig. 5A). Both the simple ordination using PCA (Fig. 5B) and the NeighborNet analysis (Fig. 5C) identify the same groupings, indicating that these are supported by multiple independent approaches. In addition, these clusters show a strong correlation with geography. To gain insight into the relationship between genetic groups and habitat, we extracted the presence of genetic clusters in populations and summarized them by biome (based on the vegetation map of Mucina and Rutherford, 2006; Table 1). The largest and most widespread cluster (group 2, coloured pale yellow in Figs 4 and 5) comprises a roughly equal number of diploids and tetraploids (Table 1, Figs 4 and 5). Group 2 is distributed longitudinally along the east–west axis of the GCFR and beyond (from the west coast to the easternmost populations) and occupies all five biomes (Table 1). Group 2 is also most closely associated with the sister species *D. adpressus* (Fig. 4). The remaining clusters are geographically more localized. Samples assigned to group 1 (coloured bright yellow) occur only in the Fynbos and Succulent Karoo biomes and are largely confined to the coastal forelands at the

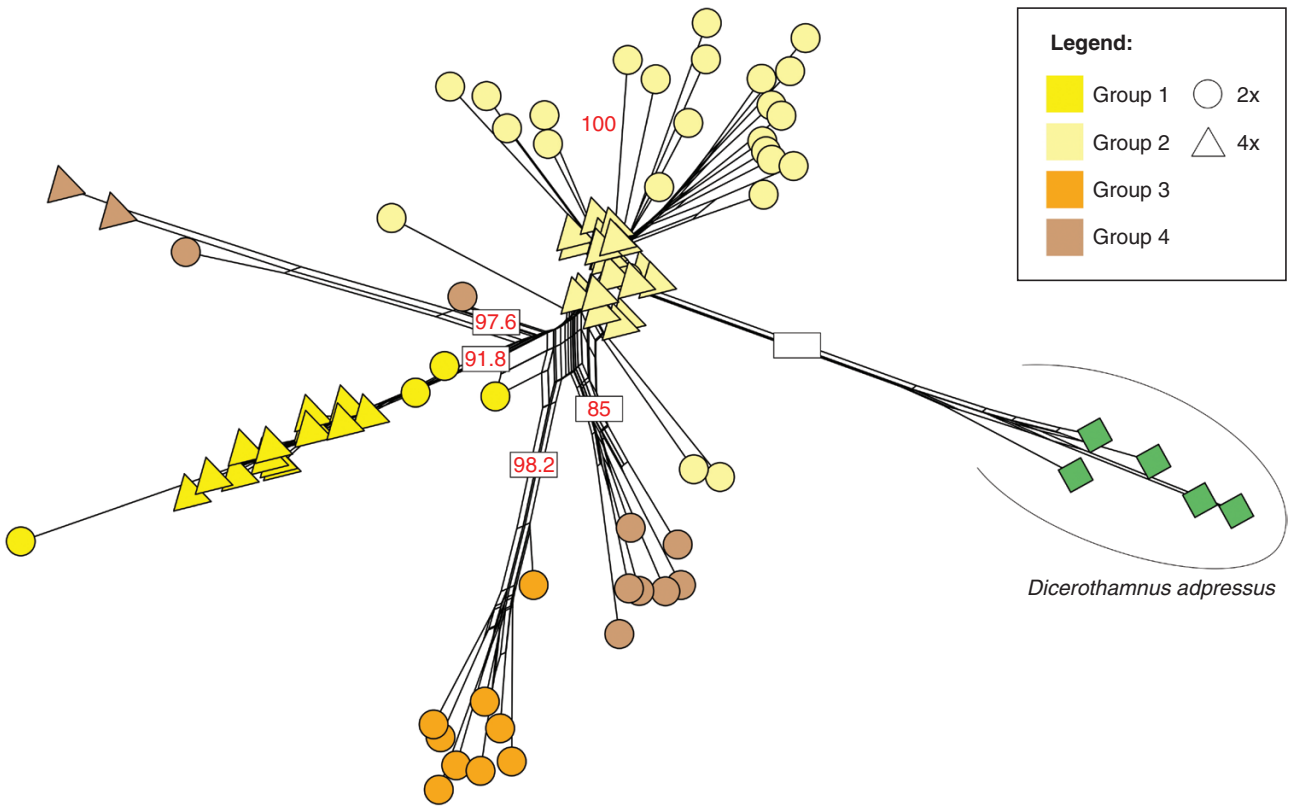

FIG. 4. NeighborNet network showing relationships between samples of *D. rhinocerotis* and of its closest relative, *D. adpressus*, based on RADseq data (4947 SNPs). Bootstrap support of major clades is shown.

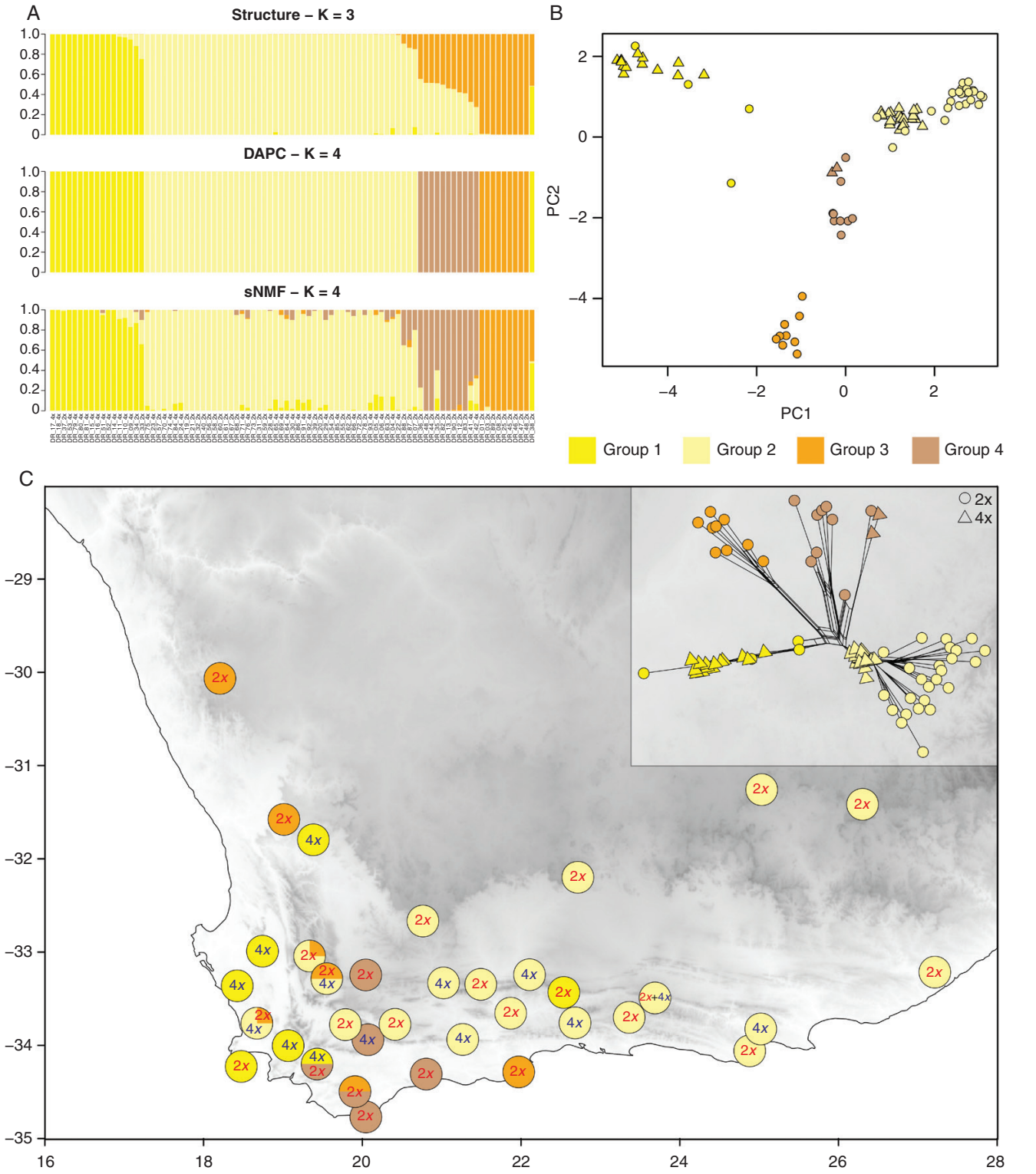

FIG. 5. Genetic structure in *D. rhinocerotis* based on RADseq data. (A) Q-matrices calculated by three different approaches based on 1047 SNPs yielded three (Structure) or four (DAPC, sNMF) distinct groups. (B) The first two axes of PCA based on 3489 SNPs. (C) Population overview with spatial distribution of genetic groups inferred by DAPC, supplemented with NeighborNet analysis from SplitsTree, coloured according to DAPC.

extreme south-west of the GCFR (Table 1, Figs 4 and 5). The next cluster, group 3 (coloured orange), comprises only diploid samples geographically mostly confined to the north–south axis of the GCFR and extending further north than any other samples. As for group 1, these samples occur mostly in the Fynbos but also in the Succulent Karoo biome (Table 1, Figs 4 and 5).

Group 4 (coloured brown) occurs mainly in the Fynbos biome and is localized around the Agulhas plain region (Table 1, Figs 4 and 5).

Niche modelling

PCA data analysis was carried out using the 23 raster layers that provide non-correlated environmental data of background pseudo-absence points serving as an ordination space for passive projection of presence data of both cytotypes (Fig. 6A). The most highly-contributing environmental variables for the first axis include hydrology (vertical distance to the channel network, valley depth), temperature (mean temperature of the warmest quarter), evapotranspiration (topographic moisture index) and soil pH (Fig. 6B). The second axis is primarily governed by temperature and precipitation features of the cool season (isothermality and minimum temperature of the coldest month, precipitation of the wettest month) and soil nutrients (particularly extractable potassium and phosphorus) (Fig. 6B).

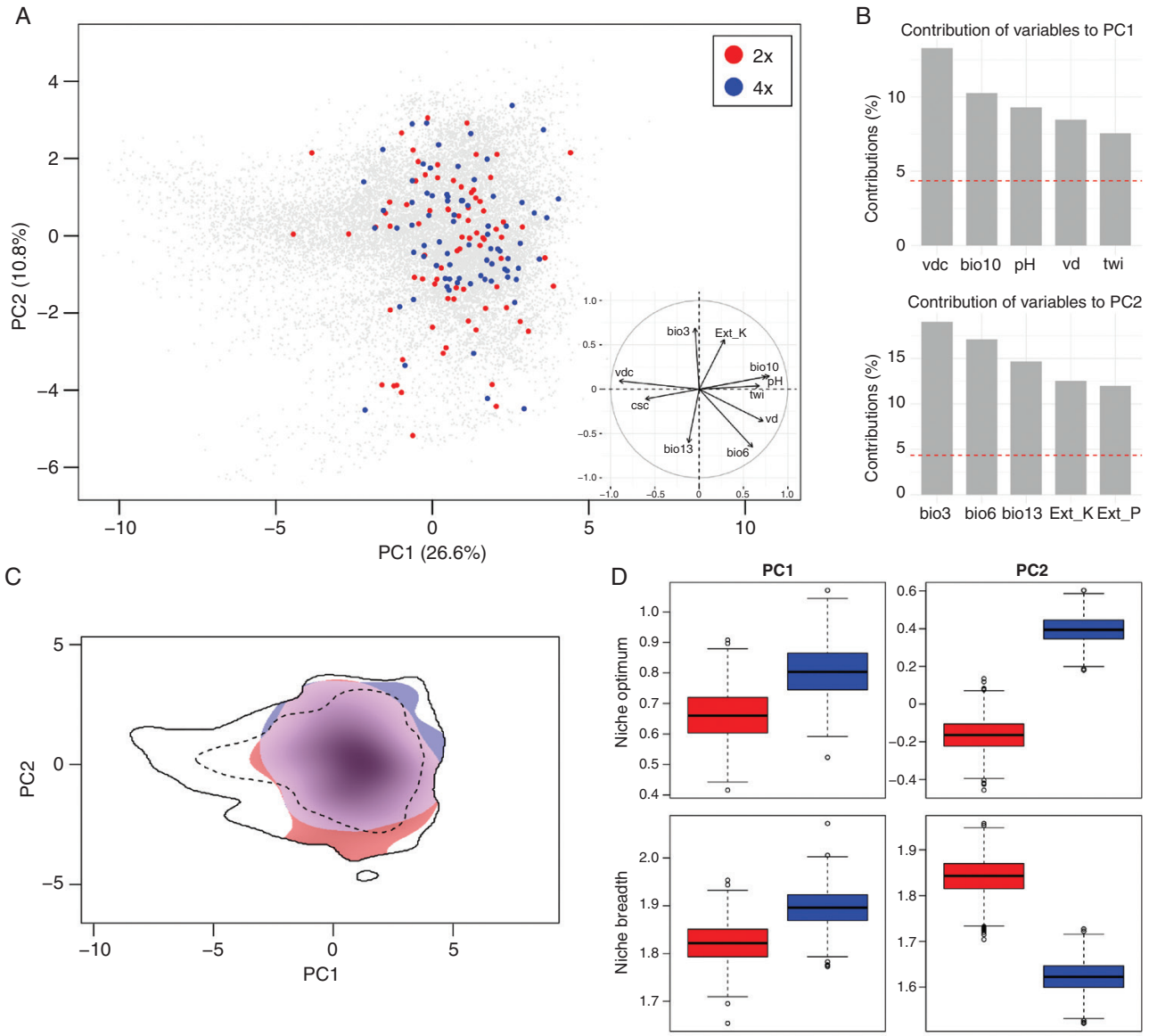

FIG. 6. Environmental niche modelling of *D. rhinocerotis* using the ecospat R package-based publicly available datasets that include CHELSA Bioclim data, additional extra-CHELSA and topographic data taken from Wüest *et al.* (2019) and soil data from Cramer *et al.* (2019). (A) PCA ordination (first two axes with explained variability in parentheses) based on background points (grey) and the position of sampling points of both cytotypes: diploid (red) and tetraploid (blue). Inset shows the contribution of main environmental characters to the first two PCA axes. (B) Percentage contributions of the five most important environmental characters for the first two PCA axes. (C) Comparison of niches occupied by two cytotypes. Shaded colours follow the common pattern for cytotypes and their putative niche overlap is in violet. Full and dashed contour lines illustrate 100 and 50 %, respectively, of available environments delimited by a 10-km buffer zone around the occurrence points of each cytotype. (D) Niche optima (top two panels) and niche breadths (bottom two panels) for both cytotypes along the first (left) and second (right) PCA axes. Abbreviations: vdc, vertical distance to channel network; bio10, mean temperature of warmest quarter; pH, soil pH; vd, valley depth; twi, topographic wetness index; bio3, isothermality; bio6, minimal temperature of coldest month; bio13, precipitation of wettest month; Ext_K, extractable K from soil; Ext_P, extractable P from soil; csc, cross-sectional curvature.

Despite this, the PCA shows poor separation of cytotypes by environmental variables, except for a slight separation along the axis suggested by extractable potassium and bio13 (precipitation of the wettest month), with diploids occurring in areas with slightly higher average precipitation in the wettest month and in areas with somewhat lower average extractable K. However, *t*-tests indicate no significant difference between the cytotypes for either of these variables (Supplementary Data Fig. S4). In contrast, significant differences between cytotypes (in tests without background point interference; Supplementary Data Fig. S4) were revealed for isothermality (bio3), available water capacity (awc), precipitation of the warmest quarter (bio18), valley depth (vd), soil pH (pH) and mean temperature of the driest quarter (bio9).

The highly significant variables for the first two PCA axes contribute most to the hypervolume of the niches of both cytotypes, and cause their slight differentiation in kernel density in the ordination space, although the degree of sharing of the common niche space is remarkable (Fig. 6C). Schoener's *D* statistic ($D = 0.723$) indicates a high overlap of niches. The niche equivalence test showed a marginally insignificant shift towards non-equivalent niches ($P = 0.055$; Supplementary Data Fig. S5A); therefore, niche equivalency between cytotypes can not be excluded. The test of niche similarity between cytotypes in both directions proved that the niche overlap was more similar than random ($2x \Rightarrow 4x$, $P = 0.049$; $4x \Rightarrow 2x$, $P = 0.045$; Supplementary Data Fig. S5B, C). On the other hand, niche optima and breadth showed displacement along the first two PCA axes, indicating niche shift between cytotypes (Fig. 6D). In particular, there is a notable shift along the second axis that is related to gradients in temperature, precipitation and soil nutrients (see above).

The differences detected in niche hypervolume between cytotypes are strongly reflected in the reconstruction of their potential range of distribution within the GCFR only, using Maxent species distribution modelling (Fig. 7). Suitable habitat is inferred for diploids on the Cape Peninsula, the coastline, and further inland on the high-lying regions of the Great Escarpment, and the inner margins of the Cederberg and Swartruggens mountains (at the inner angle of the 'L' formed

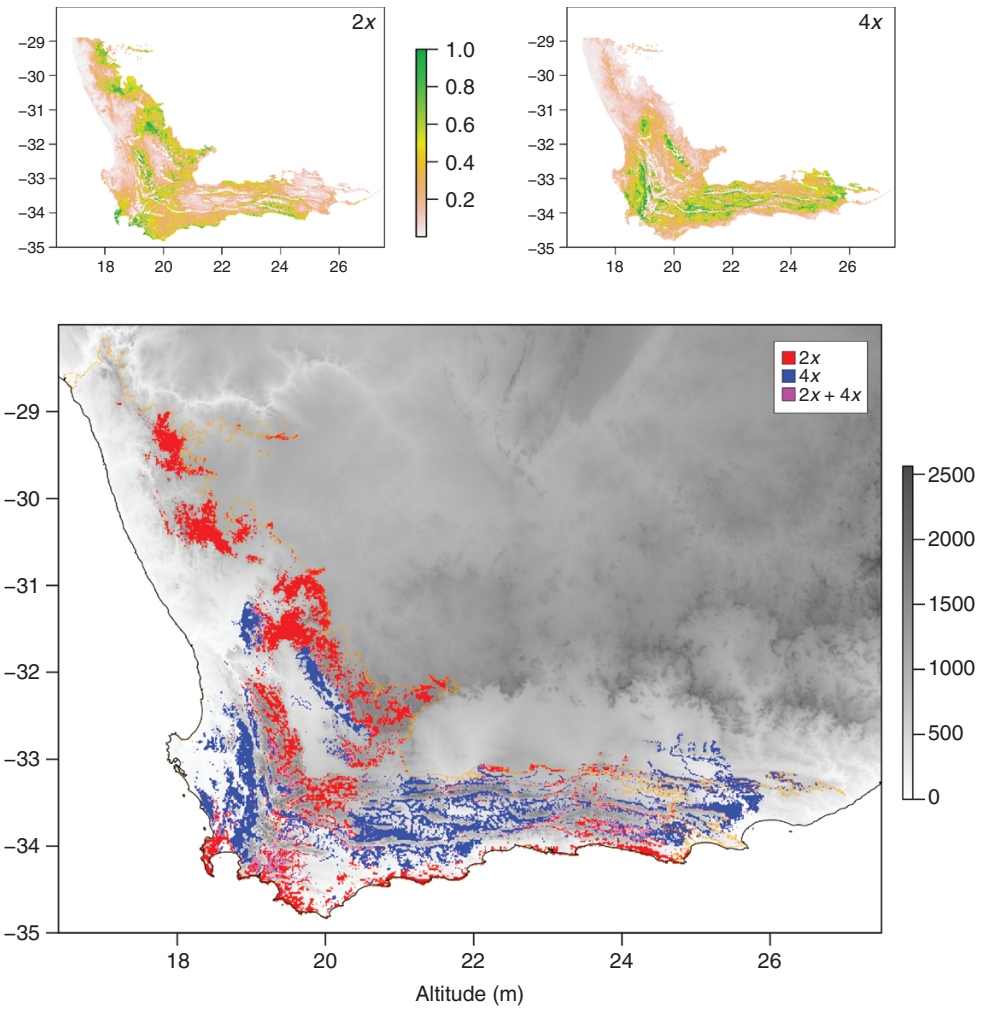

FIG. 7. Species distribution models (SDMs) calculated in Maxent using user-defined background points (see Materials and methods section for details) for both cytotypes (top panels). Potential spatial distributions are presented as the average probability taken from ten independent replicates and show the suitability of areas on the likelihood scale shown in the colour scale (0–1). The lower panel shows the overlap of the two cytotypes (their presence was inferred from an SDM based on equal training sensitivity and specificity logistic threshold: mean value for $2x = 0.407$, $4x = 0.449$) with the likely region of co-occurrence. The orange line shows the border of the GCFR.

by the Cape Fold Belt mountains). Tetraploids, in contrast, are modelled as preferring the lowlands of the western coastal platform, the Little Karoo, and the seaward-facing escarpments of the Roggeveld and northern Bokkeveld mountains, as well as the far eastern regions of the GCFR in the grassy fynbos of the Albany region. Regions where both cytotypes are modelled as co-occurring cover smaller areas but are widespread from near Cape Town through to the Overberg region, and many mountain fringes along the east–west arm of the GCFR.

Multivariate morphometrics

Before analyses, all variables were tested for normality and log-transformed if required. Correlation analysis between all traits revealed low to relatively high Pearson correlation coefficients ($0.0 < |r| \leq 0.85$), all of which are below the $|r| > 0.85$ threshold for very high correlation (Supplementary Data Table S3, Supplementary Data Fig. S6). Several traits show significant differences among cytotypes, as indicated by *t*-tests (box and whisker plots in Supplementary Data Fig. S7). Highly significant differences were detected in the spacing, length and shape (length/width ratio) of the leaves, in corolla dimensions and shape, in the number of florets per capitulum, and in cypsela dimensions between the two cytotypes.

Due to the correlated nature of the features, PCA was performed on the entire dataset and demonstrates a degree of separation between the two cytotypes, although some outlier tetraploid samples occur well within the space otherwise occupied by diploids (Fig. 8A). The contribution of the five most important features for the first and second axes of the PCA is plotted in Fig. 8B. A summary of the strength of the contribution of each trait to the first two PCA axes is given in Table 2.

To gain greater insight into the morphological delineation of the two cytotypes, we performed a canonical discriminant analysis on the entire dataset (Fig. 9A), and with stepwise selection of traits with highest discriminatory power (Fig. 9B). Stepwise selection resulted in retention of only three traits whose linear combination included significant discriminatory power for the entire dataset. Both approaches showed good separation of cytotypes and revealed the importance of individual traits for cytotype discrimination (Table 2), with the two most discriminatory traits being corolla width and leaf shape (Fig. 7C).

DISCUSSION

The Cape flora is extremely diverse, making it one of the hotspots of biodiversity on Earth (Myers *et al.*, 2000). At the same time, the local flora is one of the best-studied of a particular area because its uniqueness has attracted the attention of biologists for centuries (elaborated in detail by Glen and Germishuizen, 2010). We therefore currently have reasonably good assumptions about the number of species that may

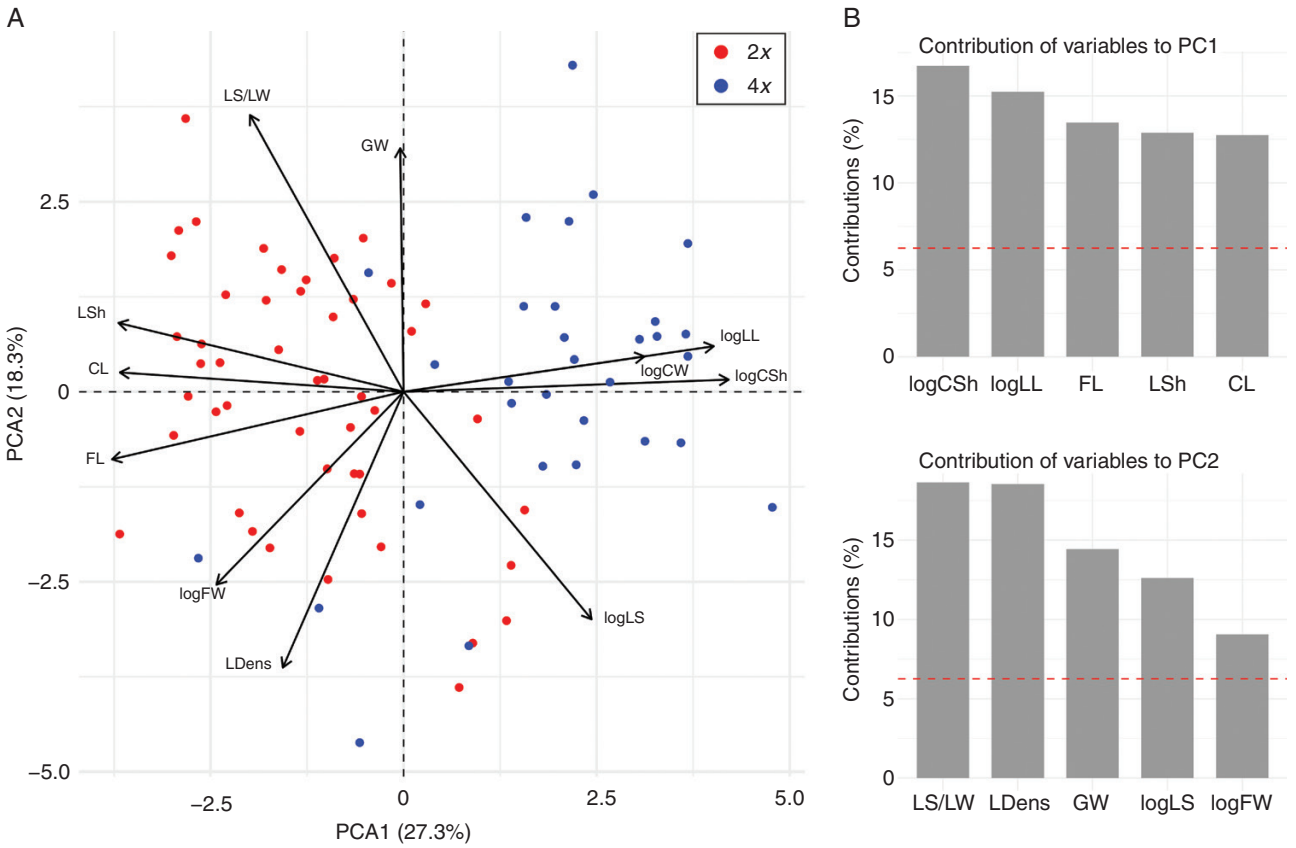

FIG. 8. Ordination biplot of morphometric traits of both cytotypes (2*x*, red; 4*x*, blue) of *D. rhinocerotis* analysed by PCA (A). Two first PCA axes are plotted with their explained variability in parentheses and a subset of traits with the best contribution to them. Abbreviations: CSh, corolla shape; LL, leaf length; FL, fruit length; LSh, leaf shape; CL, corolla length; LS, leaf spacing; LW, leaf width; LDens, leaf density; GW, guard cell width; FW, fruit width.

TABLE 2. *Description of morphological traits and their significance in PCA (PC1 and PC2) and canonical discriminant analysis (CDA; for full data and stepwise selection) for* D. rhinocerotis*. Traits that were log-transformed prior the analyses due to normality violence are noted with the prefix 'log'.*

Trait	Character description	PCA		CDA	
		PC1	PC2	Full	Stepwise
Leaf length (logLL)*	Length of leaves	0.390	0.071	0.777	
Leaf width (LW)*	Width of leaves at widest point	0.014	0.220	−0.112	
Leaf spacing (logLS)*	Distance between bases of vertically adjacent leaves	0.237	−0.355	0.424	
Corolla length (CL)*	Length of corolla tubes	−0.357	0.030	−0.509	
Corolla width (logCW)*	Width of corolla tubes at widest point	0.303	0.055	0.593	0.637
Fruit length (FL)*	Length of mature fruit	−0.367	−0.106	−0.416	
Fruit width (logFW)*	Width of fruit at widest point	−0.236	−0.301	−0.412	
Number of florets (logFlor)	Number of florets per capitulum	−0.120	−0.099	−0.454	−0.488
Guard cell length (GL)	Length of stomatal guard cell	0.009	0.257	0.085	
Guard cell width (GW)	Width of stomatal guard cell at widest point	−0.004	0.380	0.104	
Leaf shape (LSh)	LW/LL	−0.359	0.107	−0.821	−0.882
Leaf density (LDens)	LL/LS	−0.152	−0.431	−0.309	
LS/LW		−0.193	0.432	−0.408	
Corolla shape (logCSh)	CW/CL	0.409	0.019	0.692	
Fruit shape (logFSh)	FW/FL	0.104	−0.231	−0.016	
Guard cell shape (GSh)	GW/GL	−0.008	0.241	0.088	

*How the character was measured is shown in Fig. 2.

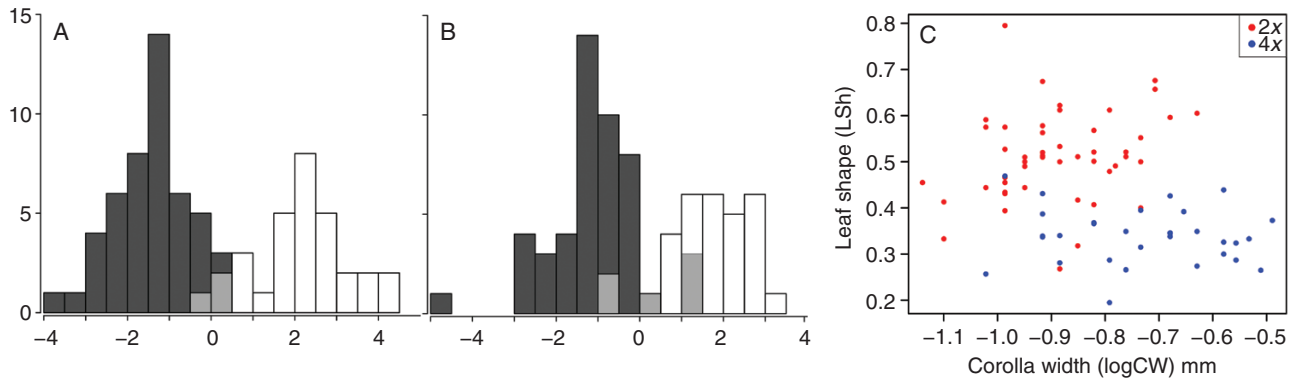

FIG. 9. Discriminant analyses of morphometric traits for both cytotypes (2*x*, dark grey; 4*x*, white) using (A) complete set of traits and (B) bidirectional stepwise selection of traits with highest discrimination power. (C) Two traits with the best contribution to differentiation of both cytotypes (corolla width and leaf shape) are plotted against each other. Corolla width was log-transformed prior to analysis due to normality violation.

occur here (Manning and Goldblatt, 2012; Snijman, 2013), although new taxa are regularly described, especially from remote areas. On the other hand, population-level variation, including cytotype diversity, remains remarkably overlooked, likely due to the long-standing hypothesis of polyploid paucity in the region (nicely documented in Oberlander *et al.*, 2016). This lack of knowledge is only rarely mitigated by studies that use population-targeted approaches accounting for cytotype diversity (e.g. Krejčíková *et al.*, 2013*a*, *b*; Glennon *et al.*, 2019). Here we present such a study aimed at elucidating aspects of ploidy variability in one of the dominant species of renosterveld vegetation, and the most well-known since it is the namesake for the vegetation type, *D. rhinocerotis*.

Cytogeography and intra-cytotype variation

Our data clearly show that *D. rhinocerotis* consists of diploids and tetraploids, which usually form cytotypically uniform populations, since <20 % of examined populations comprise mixes of cytotypes. The estimated chromosome number for the diploid $2n = 16$ corresponds well with the known data for the most closely related species within the South African Gnaphalieae, where even *Disparago* and *Stoebe* have a base chromosome number $x = 8$ (Bayer *et al.*, 2000). The position of 5S and 35S rDNA is in congruence with Garcia *et al.* (2010), who presented the linked location of rDNA probes in the tribe Gnaphalieae.

Despite extensive sampling over the species range and quite common co-occurrence of both cytotypes, no other minor ploidy levels have been found, which is indicative of reproductive isolation between cytotypes (e.g. Husband and Sabara, 2004; Trávníček *et al*., 2010). This may be due to, or at least reinforced by, a 2-week shift in flowering time between cytotypes, at least at some co-occurring sites (authors, pers. obs.), where the overlap in flowering is actually very short (similar, for example, to Nuismer and Cunningham, 2005; Pegoraro *et al*., 2016). Bergh *et al.* (2007) surveyed genetic variation across the range of *D. rhinocerotis* but found no indication of any genetic discontinuities, even though their sampling strategy would have resulted in the inclusion of both diploids and tetraploids. Instead, they recovered large intrapopulation variation and some signal of isolation by distance. However, the genetic marker system utilized [multilocus dominant inter simple sequence repeat (ISSR) markers] may be poor at detecting variation based on ploidy differences, and Bergh *et al.* (2007) did not distinguish cytotypes. Accordingly, our RADseq data revealed the existence of at least four relatively well-separated clusters, none of which were exclusively tetraploid, and the genetic distances between cytotypes do not exceed the usual population-level distance (see below).

The nature of our data, which reflect sampling across nearly the entire range, allows us to compare populations in terms of intra-cytotype variation in relative genome size and to determine the degree to which this correlates with geographic and environmental data (*sensu* e.g. Chumová *et al*., 2015, 2021; Vít *et al*., 2016). Our findings, the first to our knowledge for the Cape flora, show gradual changes in genome size as a function of latitude, longitude and altitude in diploids as well as longitude and altitude in tetraploids (Fig. 3). These three environmental axes are strongly correlated with climatic clines in the GCFR, such as precipitation amount and seasonality (Cowling and Holmes, 1992; Procheş *et al*., 2005), temperature, and to a lesser degree, edaphic characteristics, since the Cape mountains largely comprise coarse-grained, acidic nutrient-poor quartzites while the lowlands harbour a more diverse range of soil types. This assumption is underpinned by multiple environmental layers that show a correlation with the genome size variability found in both cytotypes (Supplementary Data Figs S2 and S3, Supplementary Data Table S2). The GCFR is also one of the most environmentally heterogeneous regions in southern Africa, which correlates highly with species diversity (Cramer and Verboom, 2017) and may create opportunities for co-existence of entities with very similar niche requirements. Correspondence of within-cytotype variation in genome size and environmental gradients has been found in similar studies conducted mainly in the northern hemisphere (e.g. Slovák *et al*., 2009; Cires *et al*., 2010; Díez *et al*., 2013; Duchoslav *et al*., 2013; Štubňová *et al*., 2017; Rejlová *et al*., 2019). Most of these studies show linear changes along gradients, but there is no uniform response to geographic variables and each species/cytotype responds differently. Interestingly, all significant changes in average population genome sizes in *D. rhinocerotis* show a positive response to all gradients examined. An exception is the absence of a correlated response of tetraploids to latitude, but this is probably due to the relatively narrow distribution of tetraploids along latitude (most populations are clustered between 33° and 34.5 °S; Fig. 3).

Genetic differentiation of individuals, populations and cytotypes

Using the RADseq procedure for non-model organisms, we were able to unambiguously separate two closely related species, *D. adpressus* and *D. rhinocerotis* (Fig. 4), and distinguish four relatively well-separated genetic groups among individuals of the latter (Fig. 5A, B). The known ploidy level at the individual level of *D. rhinocerotis* shows that tetraploids are present in three of these groups. Only the fourth genetic group (orange in Fig. 5) is composed exclusively of diploids. We encountered a predominance of tetraploids in the first group (dark yellow in Fig. 5) comprising >70 % of the sampled individuals (Table 1). A well-balanced proportion of cytotypes was found in group 2 (light yellow in Fig. 5), which also represents the largest group in terms of the number of individuals sampled (Table 1). However, our data show no clustering of individual genotypes based on ploidy level, suggesting a high degree of SNP sharing between the two cytotypes and the very likely emergence of tetraploids independently in each genetic cluster containing them.

The affiliation of individuals to populations provides an opportunity to examine the spatial distribution of these groups in the GCFR (Fig. 5C). It clearly shows that group 2 (pale yellow) is also the most spatially distributed, extending almost from the west coast in a longitudinal direction to the easternmost populations included in the study. This group is also the only one that is able to colonize all the biomes that were assigned to the populations we sampled (Table 1). The other three genetic groups are more spatially restricted but sometimes somewhat dispersed. The first group (bright yellow) is almost exclusively restricted to the south-western part of the GCFR, with the exception of a single spatially distant population in the pass through the Swartberg Mountains between De Rust and Klaastrom. The third group (orange), although only nine individuals have been assigned to it, is widely dispersed and extends from Mossel Bay on the south coast to the northernmost and westernmost populations near Kamieskroon. The fourth group (brown) is spatially compact from the southernmost tip of Agulhas to Touwsrivier. The overlapping area, where all groups meet to form genetically mixed populations, is the south-western part of the GCFR. This fits well with the centre of species diversity (Born *et al*., 2006) and the regionalization of the GCFR, where four distinct regions have been found based on phylogenetic comparative analyses, with three of them interconnecting in this region (Myburg and Daniels, 2022).

Environmental differentiation of cytotypes

The majority of populations are ploidy uniform, and there are virtually distinct niche differences between the two cytotypes, despite both cytotypes possessing a high degree of overlap in niche hypervolume (Fig. 6C). However, derived niche optima and niche breadths show notable shifts along the first two PCA axes (Fig. 6D). This corresponds to the relatively high value of Schoener's *D* statistic alongside a proven test for higher niche overlap similarity than random. Contrarily, the test for niche equivalency showed that both cytotypes are likely equivalent in their niches. This means, in other words, that both cytotypes are similar enough to be

able to share the same niche; on the other hand, their niche preferences are little bit shifted, which is clearly mirrored in their real as well as potential distribution (Figs 1 and 7, respectively).

Diploids have slightly lower isothermality, reflected in their modelled preference for coastal or, alternatively, high-altitude habitats. Diploids also occur in areas with slightly greater available water capacity, likely reflecting either higher water tables or greater clay content in soils, in keeping with the largely shale bedrock underpinning large parts of the interiors of the Cape Provinces of South Africa. Diploids also appear to occur in habitats in which the driest season is both hotter and drier than the driest season experienced by tetraploids, based on differences in bio18 (precipitation of the warmest quarter) and bio9 (mean temperature of the driest quarter). We interpret this as an indication of the diploid cytotype occurring largely in areas that experience more seasonal (winter-only) precipitation, with warm dry summers, while the tetraploid cytotype, with its large distribution in the eastern part of the GCFR, experiences a more aseasonal precipitation regime, with more rainfall during the very hot summers. These findings differ from the pattern observed in the Cape *Schoenus* (Elliott *et al.*, 2023), a taxon occurring in both sandstone and shale substrates and at a wider elevation range, where different ploidies show overlapping distribution in the GCFR but polyploids are prevalent in drier locations that have more variation in precipitation between dry and wet months. In addition, there is a significant difference in the soil pH ranges of the cytotypes, with diploids occurring in more acidic soils across the GCFR. This is surprising given the modelled association of diploids with coastal habitats, but may be accounted for by the requirement for richer soils by *D. rhinocerotis*, the coastal shales and granites being more leached by higher rainfall.

Although we do not know the origin of *Dicerothamnus* tetraploids with certainty, our genetic data rather point to their autopolyploid origin, which is also indirectly supported by the fact that niche differentiation resulted only in niche shifting, not in much greater niche breadth or occupation of completely different habitats, which is common in allopolyploids (e.g. Wu *et al.*, 2016; Dai *et al.*, 2020; Akiyama *et al.*, 2021). A similar scenario with a high level of niche similarity and equivalence, but accompanied by non-identity of niches, was observed in the diploid–autopolyploid complex of *Arabidopsis arenosa* (Molina-Henao and Hopkins, 2019). Similarly, assessment of niche differentiation between diploids and their auto- and allopolyploid descendants in the *Alyssum montanum* species complex shows only marginal shifts in the niches of autopolyploids compared with substantial shifts in allopolyploids (Arrigo *et al.*, 2016).

Morphological differentiation of cytotypes

Dicerothamnus rhinocerotis is known for its morphological variability (e.g. Levyns, 1926, 1935*a*, *b*), but only with the ability to distinguish cytotypes can some of the underlying causes be assessed. Our data show that the overall morphological variability is at least partially ploidy dependent (Figs 8 and 9); in particular, variability in leaf and corolla traits is highly associated with ploidy level (Supplementary Data Fig. S7). Interestingly, morphological studies conducted by Levyns in the first half of the 20th century distinguished several morphological types. She noted three forms of the species, one very tall and confined to inland watercourses, and the remaining two forms being more common and widespread, and frequently co-occurring (Levyns, 1926, 1935*a*, *b*). She described these two main forms as being easily distinguished when in flower due to differences in overall arrangement of the flowerheads, but harder to distinguish when in the purely vegetative condition. Although we are not able to directly associate Levyns' main forms with ploidy levels, some coincidences are remarkable. The more common of the two main forms, first discovered growing abundantly on Signal Hill in the city of Cape Town, Levyns called the 'Signal Hill' form, and also noted that it is very common throughout the range of the species. This form is somewhat greener in colour due to contracted internodes and smaller, more crowded leaves, and bears the capitula closer to the older, woody portion of each branch; it corresponds well in both morphology and distribution with the diploid cytotype. The other form, the 'Stellenbosch' form, has slightly larger leaves and more elongated axes, showing more of the white-woolly stem and so having a greyer appearance overall. In this form, small groups of capitula are borne near the tips of the young branches. Levyns (1929) described the Stellenbosch form as covering enormous tracts of country in the Stellenbosch, Malmesbury, Ceres, Worcester, Swellendam and Riversdale districts of the south-western Cape, and also co-occurring with the Signal Hill form on the Cape Peninsula. The Stellenbosch form corresponds with the tetraploid cytotype.

Taxonomic treatment

The potential shift in flowering time, the absence of triploids even within mixed populations, the different niche preferences and, last but not least, a high level of morphological discrimination of cytotypes, suggest that the two cytotypes of *D. rhinocerotis* can be considered as separate taxa. On the other hand, despite extensive examination of variation, Levyns (1935*a*) did not consider this variation worthy of taxonomic recognition. She found analogous variability in other close relatives in the former genus *Elytropappus* (*E. gnaphaloides* and *E. scaber*), interpreting this predisposition to variability as a genus-level character (Levyns, 1935*a*). Supportive evidence provided by our RADseq data and ISSR data of Bergh *et al.* (2007) points to the absence of any geographic structuring of genetic variation that could be associated with cytotype differentiation. Berg *et al.* (2007) examined spatial patterns of genetic variation across the range of the species using anonymous multilocus dominant ISSR markers, and recovered relatively high measures of genetic variation. No two individuals shared the same ISSR banding pattern, and proportions of polymorphic loci were very high, while nearly 80 % of the ISSR variation resided among individuals within localities rather than among localities. In contradiction to our results they also found little geographic structuring of genetic variation, although the north-western-most regions (Richersveld and northern Namaqualand) were distinct, and harboured the lowest variation. However, this difference may simply be due to the sensitivity of the two methods, where RADseq is at the highest

level among population-based methods (comparable to microsatellites; e.g. Brandrud *et al.*, 2017). Berg *et al.* (2007) also recovered signals of isolation-by-distance across fairly large spatial scales. These findings are consistent with high levels of outcrossing and gene flow. On the other hand, the data presented here show the opposite scenario of genetic clustering without the usual transitions from one to the other. The level of admixture at the individual level is relatively low (if we consider four groups, i.e. we take the single transition group in the Structure analysis as a separate cluster, which other methods clearly indicate), suggesting a low level of admixture between different genetic clusters. In addition, no clustering towards cytotype separation is present, and even tetraploids do not show higher levels of admixture (Fig. 5).

All of the above evidence points to the fact that the morphological differentiation and niche shifts demonstrated in *D. rhinocerotis* cytotypes are not reflected in the genetic structure and probably only demonstrate adaptation to local ecological conditions, which also causes shifts in phenotypic plasticity towards apparent differences in phenology. However, the origin of tetraploids is still unknown and deserves further study.

Conclusions

Although our study deals with only a tiny fraction of the immense plant variation of the Cape flora, it highlights some aspects of hidden or hitherto overlooked diversity. In particular, the research approach advocates population-targeted screening that may provide new insights into intraspecific variation accompanied by ploidy differentiation. Such studies are still very rare in the Cape, although some plant groups (specifically Asteraceae) represent ideal models. Our data on the common and well-known flagship Cape species *D. rhinocerotis* show a geographically structured distribution of cytotypes that is accompanied by quite distinguished but ploidy-independent genetic structure, and niche and morphological differentiation. This opens a number of questions that could be answered in future research on Cape systems and elsewhere, ranging from cytotype formation and maintenance in mixed populations to tetraploid origin, and to the adaptability of different cytotypes in the face of climate change and other questions.

SUPPLEMENTARY DATA

Supplementary data are available online at https://academic.oup.com/aob and consist of the following.

Table S1: overview table of sampled populations, number of analysed plants and results provided by flow cytometry. Table S2: list and brief description of all environmental data and their correlation with within-cytotype genome size variability at population level. Table S3: list and brief description of uncorrelated environmental data that were used for niche modelling. Table S4: raw morphometric data measured in individuals included in the morphology analysis. Table S5: raw RADseq data inferred from Stacks *de novo* procedure. Figure S1: examples of histograms documenting flow cytometry analyses in *D. rhinocerotis*. Figure S2: genome size variation correlated with environmental data in diploid populations of *D. rhinocerotis*. Figure S3: genome size variation correlated with environmental data in tetraploid populations of *D. rhinocerotis*. Figure S4: box and whisker plots across all environmental variables tested between cytotypes of *D. rhinocerotis*. Figure S5: results of niche equivalency and similarity tests between cytotypes. Figure S6: correlogram of all morphological traits based on Pearson correlation coefficients. Figure S7: box and whisker plots across all morphological traits estimated for cytotypes of *D. rhinocerotis*.

FUNDING

The study was financed by the Ministry of Education, Youth and Sports of the Czech Republic, Operational Programme Research, Development and Education, The European Structural and Investment Funds, EU (project no. CZ.02.2.26/0.0/0.0/16_027/0007852) to Z.C. in 2018 and a Joan Wrench/SANBI Honours student bursary to Z.M. in 2021. Computational resources were provided by the e-INFRA CZ project (ID 90140), supported by the Ministry of Education, Youth and Sports of the Czech Republic. Additional support was provided as part of a long-term research project of the Czech Academy of Sciences, Institute of Botany (RVO 67985939).

ACKNOWLEDGEMENTS

We thank Jan Ptáček, Dora Čertnerová and Martin Čertner for their help in the field, Soňa Píšová, Michaela Jandová, Ladislava Paštová and Terezie Malík Mandáková for their help in the laboratories, and Léanne Dreyer for all her help and especially for providing space for a flow cytometer in her laboratory at Stellenbosch University during a 6-month internship of Z.C. Cape Nature is acknowledged for providing collecting permits.

LITERATURE CITED

Aiello-Lammens ME, **Boria RA**, **Radosavljevic A**, **Vilela B**, **Anderson RP. 2015**. spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography* **38**: 541–545. doi:10.1111/ecog.01132.

Akiyama R, **Sun J**, **Hatakeyama M**, ***et al*. 2021**. Fine-scale empirical data on niche divergence and homeolog expression patterns in an allopolyploid and its diploid progenitor species. *New Phytologist* **229**: 3587–3601.

Andrés-Sánchez S, **Verboom GA**, **Galbany-Casals M**, **Bergh NG. 2019**. Evolutionary history of the arid climate-adapted *Helichrysum* (Asteraceae: Gnaphalieae): Cape origin and association between annual life-history and low chromosome numbers. *Journal of Systematics and Evolution* **57**: 468–487.

Arrigo N, **de La Harpe M**, **Litsios G**, ***et al*. 2016**. Is hybridization driving the evolution of climatic niche in *Alyssum montanum*? *American Journal of Botany* **103**: 1348–1357. doi:10.3732/ajb.1500368.

Bayer RJ, **Puttock CF**, **Kelchner SA. 2000**. Phylogeny of South African Gnaphalieae (Asteraceae) based on two noncoding chloroplast sequences. *American Journal of Botany* **87**: 259–272.

Beaulieu JM, **Leitch IJ**, **Patel S**, **Pendharkar A**, **Knight CA. 2008**. Genome size is a strong predictor of cell size and stomatal density in angiosperms. *New Phytologist* **179**: 975–986. doi:10.1111/j.1469-8137.2008.02528.x.

Bergh NG, **Linder HP. 2009**. Cape diversification and repeated out-of-southern-Africa dispersal in paper daisies (Asteraceae–Gnaphalieae). *Molecular Phylogenetics and Evolution* **51**: 5–18. doi:10.1016/j.ympev.2008.09.001.

Bergh NG, **Hedderson TA**, **Linder HP**, **Bond WJ. 2007**. Palaeoclimate-induced range shifts may explain current patterns of spatial genetic

variation in renosterbos (*Elytropappus rhinocerotis*, Asteraceae). *Taxon* **56**: 393–408. doi:10.1002/tax.562011.

Bergh NG, **Verboom GA**, **Rouget M**, **Cowling RM. 2014**. Vegetation types of the Greater Cape Floristic Region. In: **Allsopp N**, **Colville JF**, **Verboom GA**, eds. *Fynbos: ecology, evolution, and conservation of a megadiverse region*. Oxford; New York: Oxford University Press, 26–46.

Bond WJ, **van Wilgen BW. 1996**. *Fire and plants*. London: Chapman and Hall.

Born J, **Linder H**, **Desmet P. 2006**. The Greater Cape Floristic Region. *Journal of Biogeography* **34**: 147–162.

Brandrud MK, **Paun O**, **Lorenzo MT**, **Nordal I**, **Brysting AK. 2017**. RADseq provides evidence for parallel ecotypic divergence in the autotetraploid *Cochlearia officinalis* in Northern Norway. *Scientific Reports* **7**: 5573. doi:10.1038/s41598-017-05794-z.

Broennimann O, **Fitzpatrick MC**, **Pearman PB**, ***et al***. **2012**. Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography* **21**: 481–497.

Catchen JM, **Amores A**, **Hohenlohe P**, **Cresko W**, **Postlethwait JH. 2011**. Stacks: building and genotyping loci de novo from short-read sequences. *G3 Genes Genomes Genetics* **1**: 171–182. doi:10.1534/g3.111.000240.

Chase BM, **Meadows ME. 2007**. Late Quaternary dynamics of southern Africa's winter rainfall zone. *Earth-Science Reviews* **84**: 103–138. doi:10.1016/j.earscirev.2007.06.002.

Chumová Z, **Krejčíková J**, **Mandáková T**, **Suda J**, **Trávníček P. 2015**. Evolutionary and taxonomic implications of variation in nuclear genome size: lesson from the grass genus *Anthoxanthum* (Poaceae). *PLoS One* **10**: e0133748. doi:10.1371/journal.pone.0133748.

Chumová Z, **Mandáková T**, **Trávníček P. 2021**. On the origin of tetraploid vernal grasses (*Anthoxanthum*) in Europe. *Genes* **12**: 966. doi:10.3390/genes12070966.

Chumová Z, **Belyayev A**, **Mandáková T**, ***et al***. **2022**. The relationship between transposable elements and ecological niches in the Greater Cape Floristic Region: a study on the genus *Pteronia* (Asteraceae). *Frontiers in Plant Science* **13**: 982852. doi:10.3389/fpls.2022.982852.

Cires E, **Cuesta C**, **Revilla MA**, **Fernández Prieto JA. 2010**. Intraspecific genome size variation and morphological differentiation of *Ranunculus parnassifolius* (Ranunculaceae), an Alpine–Pyrenean–Cantabrian polyploid group. *Biological Journal of the Linnean Society* **101**: 251–271. doi:10.1111/j.1095-8312.2010.01517.x.

Di Cola V, **Broennimann O**, **Petitpierre B**, ***et al***. **2017**. ecospat: an R package to support spatial analyses and modeling of species niches and distributions. *Ecography* **40**: 774–787. doi:10.1111/ecog.02671.

Cowling RM, **Heijnis CE. 2001**. The identification of broad habitat units as biodiversity entities for systematic conservation planning in the Cape Floristic Region. *South African Journal of Botany* **67**: 15–38. doi:10.1016/s0254-6299(15)31087-5.

Cowling RM, **Holmes PM. 1992**. Flora and vegetation. In: **Cowling RM**, ed. *The ecology of fynbos: nutrients, fire and diversity*. Cape Town: Oxford University Press, 23–61.

Cowling RM, **Potts AJ**, **Bradshaw PL**, ***et al***. **2015**. Variation in plant diversity in mediterranean-climate ecosystems: the role of climatic and topographical stability. *Journal of Biogeography* **42**: 552–564.

Cramer MD, **Verboom GA. 2017**. Measures of biologically relevant environmental heterogeneity improve prediction of regional plant species richness. *Journal of Biogeography* **44**: 579–591.

Cramer MD, **Wootton LM**, **van Mazijk R**, **Verboom GA**. **2019**. New regionally modelled soil layers improve prediction of vegetation type relative to that based on global soil models. *Diversity and Distributions* **25**: 1736–1750.

Dai X, **Li X**, **Huang Y**, **Liu X. 2020**. The speciation and adaptation of the polyploids: a case study of the Chinese *Isoetes* L. diploid-polyploid complex. *BMC Evolutionary Biology* **20**: 118. doi:10.1186/s12862-020-01687-4.

Danecek P, **Auton A**, **Abecasis G**, ***et al***.; **1000 Genomes Project Analysis Group**. **2011**. The variant call format and VCFtools. *Bioinformatics* **27**: 2156–2158. doi:10.1093/bioinformatics/btr330.

Danielson JJ, **Gesch DB. 2011**. *Global multi-resolution terrain elevation data 2010 (GMTED2010)*. Washington, DC: US Geological Survey.

Dekker TG, **Fourie TG**, **Matthee E**, ***et al***. **1988**. Studies of South African medicinal plants. Part 7. Rhinocerotinoic acid: a labdane diterpene with anti-inflammatory properties from *Elytropappus rhinocerotis*. *South African Journal of Chemistry* **41**: 33–35.

Díez CM, **Gaut BS**, **Meca E**, ***et al***. **2013**. Genome size variation in wild and cultivated maize along altitudinal gradients. *New Phytologist* **199**: 264–276. doi:10.1111/nph.12247.

Dogan M, **Pouch M**, **Mandáková T**, ***et al***. **2021**. Evolution of tandem repeats is mirroring post-polyploid cladogenesis in *Heliophila* (Brassicaceae). *Frontiers in Plant Science* **11**: 607893.

Doležel J, **Greilhuber J**, **Suda J. 2007**. Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols* **2**: 2233–2244. doi:10.1038/nprot.2007.310.

Dormann CF, **Elith J**, **Bacher S**, ***et al***. **2013**. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **36**: 27–46.

Duchoslav M, **Šafářová L**, **Jandová M. 2013**. Role of adaptive and non-adaptive mechanisms forming complex patterns of genome size variation in six cytotypes of polyploid *Allium oleraceum* (Amaryllidaceae) on a continental scale. *Annals of Botany* **111**: 419–431. doi:10.1093/aob/mcs297.

Duchoslav M, **Jandová M**, **Kobrlová L**, **Šafářová L**, **Brus J**, **Vojtěchová K. 2021**. Intricate distribution patterns of six cytotypes of *Allium oleraceum* at a continental scale: niche expansion and innovation followed by niche contraction with increasing ploidy level. *Frontiers in Plant Science* **11**: 591137.

Dynesius M, **Jansson R. 2000**. Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. *Proceedings of the National Academy of Sciences of the USA* **97**: 9115–9120. doi:10.1073/pnas.97.16.9115.

Edgeloe JM, **Severn-Ellis AA**, **Bayer PE**, ***et al***. **2022**. Extensive polyploid clonality was a successful strategy for seagrass to expand into a newly submerged environment. *Proceedings of the Royal Society B: Biological Sciences* **289**: 20220538.

Elliott TL, **Muasya AM**, **Bureš P. 2023**. Complex patterns of ploidy in a holocentric plant clade (*Schoenus*, Cyperaceae) in the Cape biodiversity hotspot. *Annals of Botany* **131**: 143–156. doi:10.1093/aob/mcac027.

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. doi:10.1111/j.1365-294X.2005.02553.x.

Francis RM. 2017. pophelper: an R package and web app to analyse and visualize population structure. *Molecular Ecology Resources* **17**: 27–32. doi:10.1111/1755-0998.12509.

Frichot E, **François O. 2015**. LEA: an R package for landscape and ecological association studies. *Methods in Ecology and Evolution* **6**: 925–929. doi:10.1111/2041-210x.12382.

Frichot E, **Mathieu F**, **Trouillon T**, **Bouchard G**, **François O. 2014**. Fast and efficient estimation of individual ancestry coefficients. *Genetics* **196**: 973–983. doi:10.1534/genetics.113.160572.

Friendly M. 2002. Corrgrams: exploratory displays for correlation matrices. *American Statistician* **56**: 316–324. doi:10.1198/000313002533.

Garcia S, **Panero JL**, **Široký J**, **Kovařík A. 2010**. Repeated reunions and splits feature the highly dynamic evolution of 5S and 35S ribosomal RNA genes (rDNA) in the Asteraceae family. *BMC Plant Biology* **10**: 176. doi:10.1186/1471-2229-10-176.

Glen HF, Germishuizen G, eds. **2010**. *Botanical exploration of southern Africa*. Pretoria: South African National Biodiversity Institute.

Glennon KL, **Suda J**, **Cron GV. 2019**. DNA ploidy variation and population structure of the morphologically variable *Helichrysum odoratissimum* (L.) Sweet (Asteraceae) in South Africa. *International Journal of Plant Sciences* **180**: 755–767. doi:10.1086/704355.

Hijmans A, **Phillips S**, **Leathwick J**, **Elith J. 2022**. *terra: spatial data handling with terra. R package version 1.5.29*. https://CRAN.R-project.org/package=terra (10 November 2022, date last accessed).

Hodgson JG, **Sharafi M**, **Jalili A**, ***et al***. **2010**. Stomatal vs. genome size in angiosperms: the somatic tail wagging the genomic dog? *Annals of Botany* **105**: 573–584. doi:10.1093/aob/mcq011.

Huang C-H, **Zhang C**, **Liu M**, ***et al***. **2016**. Multiple polyploidization events across Asteraceae with two nested events in the early history revealed by nuclear phylogenomics. *Molecular Biology and Evolution* **33**: 2820–2835. doi:10.1093/molbev/msw157.

Husband BC, **Sabara HA. 2004**. Reproductive isolation between autotetraploids and their diploid progenitors in fireweed, *Chamerion angustifolium* (Onagraceae). *New Phytologist* **161**: 703–713. doi:10.1046/j.1469-8137.2004.00998.x.

Huson DH, **Bryant D. 2006**. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* **23**: 254–267. doi:10.1093/molbev/msj030.

Innangi M, **Fioretto A**, **Fondón CL**, ***et al***. **2020**. *Tuber aestivum* is associated with changes in soil chemistry and reduced biological quality in a *Quercus pubescens* stand in northern Italy. *Pedobiologia* **80**: 150648. doi:10.1016/j.pedobi.2020.150648.

Jombart T, **Ahmed I. 2011**. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* **27**: 3070–3071. doi:10.1093/bioinformatics/btr521.

Jombart T, **Devillard S**, **Balloux F. 2010**. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* **11**: 94. doi:10.1186/1471-2156-11-94.

Karger DN, **Conrad O**, **Böhner J**, ***et al***. **2017**. Climatologies at high resolution for the earth's land surface areas. *Scientific Data* **4**: 170122. doi:10.1038/sdata.2017.122.

Keeley JE, **Bond WJ**, **Bradstock RA**, **Pausas JG**, **Rundel PW. 2012**. *Fire in Mediterranean ecosystems: ecology, evolution and management*. Cambridge: Cambridge University Press.

Kemper J, **Cowling RM**, **Richardson DM. 1999**. Fragmentation of South African renosterveld shrublands: effects on plant community structure and conservation implications. *Biological Conservation* **90**: 103–111. doi:10.1016/s0006-3207(99)00021-x.

Kirchheimer B, **Schinkel CCF**, **Dellinger AS**, ***et al***. **2016**. A matter of scale: apparent niche differentiation of diploid and tetraploid plants may depend on extent and grain of analysis. *Journal of Biogeography* **43**: 716–726. doi:10.1111/jbi.12663.

Knaus BJ, **Grünwald NJ. 2017**. vcfr: a package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources* **17**: 44–53. doi:10.1111/1755-0998.12549.

Koekemoer M. 2019. Taxonomy and reclassification of South African Asteraceae genus *Elytropappus* (Gnaphalieae, Asteraceae), the description of two new genera and two new species. *Phytotaxa* **403**: 248–284. doi:10.11646/phytotaxa.403.4.1.

Koutecký P. 2015. MorphoTools: a set of R functions for morphometric analysis. *Plant Systematics and Evolution* **301**: 1115–1121.

Krejčíková J, **Sudová R**, **Lučanová M**, ***et al***. **2013***a*. High ploidy diversity and distinct patterns of cytotype distribution in a widespread species of *Oxalis* in the Greater Cape Floristic Region. *Annals of Botany* **111**: 641–649. doi:10.1093/aob/mct030.

Krejčíková J, **Sudová R**, **Oberlander KC**, **Dreyer LL**, **Suda J. 2013***b*. Cytogeography of *Oxalis pes-caprae* in its native range: where are the pentaploids? *Biological Invasions* **15**: 1189–1194.

Levin DA. 2002. *The role of chromosomal change in plant evolution*. New York: Oxford University Press.

Levyns MR. 1926. A preliminary note on the rhenoster bush (*Elytropappus rhinocerotis*) and the germination of its seed. *Transactions of the Royal Society of South Africa* **14**: 383–388. doi:10.1080/00359192609519647.

Levyns MR. 1929. The problem of the rhenoster bush. *South African Journal of Science* **26**: 166–169.

Levyns MR. 1935*a*. A revision of *Elytropappus* Cass. *Journal of South African Botany* **1**: 89–103.

Levyns MR. 1935*b*. Veld-burning experiments at Oakdale, Riversdale. *Transactions of the Royal Society of South Africa* **23**: 231–243. doi:10.1080/00359193509518893.

Linder HP. 2003. The radiation of the Cape flora, southern Africa. *Biological Reviews* **78**: 597–638. doi:10.1017/s1464793103006171.

Linder HP, **Suda J**, **Weiss-Schneeweiss H**, **Trávníček P**, **Bouchenak-Khelladi Y. 2017**. Patterns, causes and consequences of genome size variation in Restionaceae of the Cape flora. *Botanical Journal of the Linnean Society* **183**: 515–531. doi:10.1093/botlinnean/box005.

Lischer HEL, **Excoffier L. 2012**. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* **28**: 298–299. doi:10.1093/bioinformatics/btr642.

Liu CR, **Berry PM**, **Dawson TP**, **Pearson RG. 2005**. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* **28**: 385–393. doi:10.1111/j.0906-7590.2005.03957.x.

Low AB, **Rebelo AG**, **Bredenkamp GJ. 1996**. *Vegetation of South Africa, Lesotho and Swaziland: a companion to the vegetation map of South Africa, Lesotho and Swaziland*. Pretoria: Department of Environmental Affairs & Tourism.

Mandáková T, **Lysak MA. 2016**. Chromosome preparation for cytogenetic analyses in *Arabidopsis* In: **Stacey G**, **Birchler J**, **Ecker J**, **Martin CR**, **Stitt M**, **Zhou J-M**, eds. *Current protocols in plant biology*. Hoboken: John Wiley, 43–51.

Mandáková T, **Mummenhoff K**, **Al-Shehbaz IA**, **Mucina L**, **Muehlhausee A**, **Lysak MA. 2012**. Whole-genome triplication and species radiation in the southern African tribe Heliophileae (Brassicaceae). *Taxon* **61**: 989–1000.

Manning J, Goldblatt P, eds. 2012. *Plants of the Greater Cape Floristic Region*. Pretoria: SANBI.

Marloth R. 1908. *Das Kapland*. Jena: Gustav Fischer.

Maroyi A. 2019. Ethnomedicinal uses, phytochemistry and pharmacological properties of *Elytropappus rhinocerotis*. *Journal of Pharmaceutical Sciences and Research* **11**: 3508–3513.

McCormack JE, **Zellmer AJ**, **Knowles LL. 2010**. Does niche divergence accompany allopatric divergence in *Aphelocoma* Jays as predicted under ecological speciation?: Insights from tests with niche models. *Evolution* **64**: 1231–1244. doi:10.1111/j.1558-5646.2009.00900.x.

Molina-Henao YF, **Hopkins R. 2019**. Autopolyploid lineage shows climatic niche expansion but not divergence in *Arabidopsis arenosa*. *American Journal of Botany* **106**: 61–70. doi:10.1002/ajb2.1212.

Mucina L, **Roux A**, **Rutherford MC**, ***et al.*** **2006**. Succulent Karoo biome. In: **Mucina L**, **Rutherford M**, eds. *The vegetation of South Africa, Lesotho and Swaziland*. Pretoria: South African National Biodiversity Institute, 220–299.

Mucina L, Rutherford M, eds. **2006**. *The vegetation of South Africa, Lesotho and Swaziland*. Pretoria: South African National Biodiversity Institute.

Murdoch DJ, **Chow ED. 1996**. A graphical display of large correlation matrices. *American Statistician* **50**: 178–180. doi:10.2307/2684435.

Myburgh AM, **Daniels SR. 2022**. Between the Cape Fold Mountains and the deep blue sea: comparative phylogeography of selected codistributed ectotherms reveals asynchronous cladogenesis. *Evolutionary Applications* **15**: 1967–1987. doi:10.1111/eva.13493.

Myers N, **Mittermeier RA**, **Mittermeier CG**, **da Fonseca GAB**, **Kent J. 2000**. Biodiversity hotspots for conservation priorities. *Nature* **403**: 853–858. doi:10.1038/35002501.

Nuismer SL, **Cunningham BM. 2005**. Selection for phenotypic divergence between diploid and autotetraploid *Heuchera grossulariifolia*. *Evolution* **59**: 1928–1935.

Oberlander KC, **Dreyer LL**, **Goldblatt P**, **Suda J**, **Linder HP. 2016**. Species-rich and polyploid-poor: insights into the evolutionary role of whole-genome duplication from the Cape flora biodiversity hotspot. *American Journal of Botany* **103**: 1336–1347. doi:10.3732/ajb.1500474.

Ortiz, EM. 2019. *vcf2phylip v2.0: convert a VCF matrix into several matrix formats for phylogenetic analysis*. Zenodo. https://zenodo.org/record/2540861#.YGYaJK8zaUk.

Pante E, **Simon-Bouhet B. 2013**. marmap: a package for importing, plotting and analyzing bathymetric and topographic data in R. *PLoS One* **8**: e73051. doi:10.1371/journal.pone.0073051.

Paris JR, **Stevens JR**, **Catchen JM. 2017**. Lost in parameter space: a road map for stacks. *Methods in Ecology and Evolution* **8**: 1360–1373. doi:10.1111/2041-210x.12775.

Pegoraro L, **Cafasso D**, **Rinaldi R**, **Cozzolino S**, **Scopece G. 2016**. Habitat preference and flowering-time variation contribute to reproductive isolation between diploid and autotetraploid *Anacamptis pyramidalis*. *Journal of Evolutionary Biology* **29**: 2070–2082. doi:10.1111/jeb.12930.

Peterson BK, **Weber JN**, **Kay EH**, **Fisher HS**, **Hoekstra HE. 2012**. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One* **7**: e37135. doi:10.1371/journal.pone.0037135.

Phillips SJ, **Dudík M**, **Schapire RE. 2022**. *Maxent software for modeling species niches and distributions (Version 3.4.1)*. http://biodiversityinformatics.amnh.org/open_source/maxent/ (9 April 2022, date last accessed).

Pritchard JK, **Stephens M**, **Donnelly P. 2000**. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959. doi:10.1093/genetics/155.2.945.

Procheş S, **Cowling RM**, **du Preez DR. 2005**. Patterns of geophyte diversity and storage organ size in the winter-rainfall region of southern Africa. *Diversity and Distributions* **11**: 101–109. doi:10.1111/j.1366-9516.2005.00132.x.

R Core Team. 2021. *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. https://www.R-project.org/.

Rebelo AG, **Boucher C**, **Helme N**, **Mucina L**, **Rutherford MC. 2006**. Fynbos biome In: **Mucina L**, **Rutherford MC**, eds. *The vegetation of South Africa, Lesotho and Swaziland*. Pretoria: South African Biodiversity Institute, 53–219.

Rejlová L, **Chrtek J**, **Trávníček P**, **Lučanová M**, **Vít P**, **Urfus T. 2019**. Polyploid evolution: the ultimate way to grasp the nettle. *PLoS One* **14**: e0218389. doi:10.1371/journal.pone.0218389.

Ren R, **Wang H**, **Guo C**, ***et al***. **2018**. Widespread whole genome duplications contribute to genome complexity and species diversity in angiosperms. *Molecular Plant* **11**: 414–428. doi:10.1016/j.molp.2018.01.002.

Rice A, **Šmarda P**, **Novosolov M**, ***et al***. **2019**. The global biogeography of polyploid plants. *Nature Ecology & Evolution* **3**: 265–273. doi:10.1038/s41559-018-0787-9.

Schliep KP. 2011. phangorn: phylogenetic analysis in R. *Bioinformatics* **27**: 592–593. doi:10.1093/bioinformatics/btq706.

Schneider CA, **Rasband WS**, **Eliceiri KW. 2012**. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**: 671–675. doi:10.1038/nmeth.2089.

Schoener TW. 1968. The *Anolis* lizards of Bimini: resource partitioning in a complex fauna. *Ecology* **49**: 704–726. doi:10.2307/1935534.

Semple JC, **Watanabe K. 2009**. A review of chromosome numbers in Asteraceae with hypotheses on chromosomal base number evolution. In: **Funk VA**, **Susanna A**, **Stuessy TF**, **Bayer RJ**, eds. *Systematics, evolution and biogeography of Compositae*. Vienna: International Association for Plant Taxonomy, 61–72.

Sliwinska E, **Loureiro J**, **Leitch IJ**, ***et al***. **2022**. Application-based guidelines for best practices in plant flow cytometry. *Cytometry. Part A* **101**: 749–781.

Slovák M, **Urfus T**, **Vít P**, **Marhold K. 2009**. The Balkan endemic *Picris hispidissima* (Compositae): morphology, nuclear DNA content and relationship to the polymorphic *P. hieracioides*. *Plant Systematics and Evolution* **278**: 187–201. doi:10.1007/s00606-008-0137-5.

Snijman D, ed. **2013**. *The extra Cape flora*. Pretoria: SANBI.

Soltis PS, **Soltis DE. 2016**. Ancient WGD events as drivers of key innovations in angiosperms. *Current Opinion in Plant Biology* **30**: 159–165. doi:10.1016/j.pbi.2016.03.015.

Štorchová H, **Hrdličková R**, **Chrtek J**, **Tetera M**, **Fitze D**, **Fehrer J. 2000**. An improved method of DNA isolation from plants collected in the field and conserved in saturated NaCl/CTAB solution. *Taxon* **49**: 79–84. doi:10.2307/1223934.

Štubňová E, **Hodálová I**, **Kučera J**, **Mártonfiová L**, **Svitok M**, **Slovák M. 2017**. Karyological patterns in the European endemic genus *Soldanella* L.: absolute genome size variation uncorrelated with cytotype chromosome numbers. *American Journal of Botany* **104**: 1241–1253. doi:10.3732/ajb.1700153.

Suda J, **Krahulcová A**, **Trávníček P**, **Krahulec F. 2006**. Ploidy level versus DNA ploidy level: an appeal for consistent terminology. *Taxon* **55**: 447–450.

Tank DC, **Eastman JM**, **Pennell MW**, ***et al***. **2015**. Nested radiations and the pulse of angiosperm diversification: increased diversification rates often follow whole genome duplications. *New Phytologist* **207**: 454–467. doi:10.1111/nph.13491.

Temsch EM, **Greilhuber J**, **Krisai R. 2010**. Genome size in liveworts. *Preslia* **82**: 63–80.

Theodoridis S, **Randin C**, **Broennimann O**, **Patsiou T**, **Conti E. 2013**. Divergent and narrower climatic niches characterize polyploid species of European primroses in *Primula* sect. *Aleuritia*. *Journal of Biogeography* **40**: 1278–1289. doi:10.1111/jbi.12085.

Topp EN, **Loos J. 2019**. Local and landscape level variables influence butterfly diversity in critically endangered South African renosterveld. *Journal of Insect Conservation* **23**: 225–237.

Trávníček P, **Eliášová A**, **Suda J. 2010**. The distribution of cytotypes of *Vicia cracca* in Central Europe: the changes that have occurred over the last four decades. *Preslia* **82**: 149–163.

Valentin S. 2022. *geobuffer: geodesic buffer around points (long, lat) using metric radius. R package version 0.0.0.9000*. https://github.com/valentinitnelav/geobuffer (date last accessed 4 April 2022).

Vít P, **Krak K**, **Trávníček P**, **Douda J**, **Lomonosova MN**, **Mandák B. 2016**. Genome size stability across Eurasian *Chenopodium* species (Amaranthaceae). *Botanical Journal of the Linnean Society* **182**: 637–649. doi:10.1111/boj.12474.

Warren DL, **Glor RE**, **Turelli M. 2008**. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution* **62**: 2868–2883. doi:10.1111/j.1558-5646.2008.00482.x.

Warren DL, **Glor RE**, **Turelli M. 2010**. ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography* **33**: 607–611.

Wu H, **Ma Z**, **Wang M-M**, **Qin A-L**, **Ran J-H**, **Wang X-Q. 2016**. A high frequency of allopolyploid speciation in the gymnospermous genus *Ephedra* and its possible association with some biological and ecological features. *Molecular Ecology* **25**: 1192–1210. doi:10.1111/mec.13538.

Wüest RO, **Boucher FC**, **Bouchenak-Khelladi Y**, **Karger DN**, **Linder HP. 2019**. Dissecting biodiversity in a global hotspot: uneven dynamics of immigration and diversification within the Cape Floristic Region of South Africa. *Journal of Biogeography* **46**: 1936–1947.