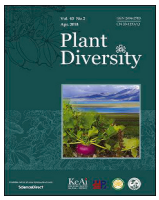
Plant Diversity xxx (xxxx) xxx

Contents lists available at ScienceDirect

Plant Diversity

journal homepage: http://www.keaipublishing.com/en/journals/plant-diversity/
http://journal.kib.ac.cn



Research paper

Polyploidy meets the Cape: Evolutionary dynamics of *Heliophila* (Brassicaceae)

Milan Pouch [a,b], Petr Šmarda [c], Pieter Winter [d], Fumi Tatsuzawa [e], Ihsan A. Al-Shehbaz [f], Klaus Mummenhoff [g], Ladislav Mucina [h,i], Martin A. Lysak [a,j], Terezie Mandáková [a,j,*]

[a] *Central European Institute of Technology, Masaryk University, Brno, Czechia*
[b] *National Centre for Biomolecular Research, Faculty of Science, Masaryk University, Brno, Czechia*
[c] *Department of Botany and Zoology, Faculty of Science, Masaryk University, Brno, Czechia*
[d] *South African National Biodiversity Institute, Kirstenbosch, Cape Town, South Africa*
[e] *Faculty of Agriculture, Iwate University, Iwate, Japan*
[f] *Missouri Botanical Garden, St. Louis, United States of America*
[g] *Biology Department, Osnabrück University, Osnabrück, Germany*
[h] *Harry Butler Institute, Murdoch University, Perth, Australia*
[i] *Department of Geography and Environmental Studies, Stellenbosch University, South Africa*
[j] *Department of Experimental Biology, Faculty of Science, Masaryk University, Brno, Czechia*

ARTICLE INFO

Article history:
Received 28 December 2025
Received in revised form
13 May 2026
Accepted 14 May 2026
Available online xxx

Keywords:
Brassicaceae
Chromosome evolution
Heliophila
Life-history evolution
Molecular phylogenetics
Polyploidy

ABSTRACT

The Cape Floristic Region (CFR), a global biodiversity hotspot, provides an exceptional setting for studying plant diversification. The South African genus *Heliophila* (Brassicaceae), notable for its species richness and morphological diversity, offers a unique system for examining the evolutionary role of polyploidy. Using an integrative framework combining phylogenomics, cytogenetics, and floral pigment biochemistry, we reconstructed the genus's evolutionary history based on ITS and 48 nuclear markers from 392 accessions representing over 120 taxa. Our analyses resolved four major clades originating during a rapid Miocene radiation (12–10 Mya) and revealed extensive chromosome number variation ($2n$ = 16–80) shaped by descending dysploidy and recurrent neopolyploidization. Ancestral state reconstruction indicates that the common ancestor was annual and that perenniality evolved independently at least 13 times, predominantly in lineages occupying cooler and wetter habitats. Environmental analyses show strong ecological differentiation among clades, suggesting adaptation to contrasting climatic and edaphic conditions across the CFR. Floral pigment profiling uncovered multiple species with delphinidin-based anthocyanins, revealing new ornamental potential. Together, these findings demonstrate that polyploidy, genome restructuring, and ecological heterogeneity have jointly shaped the diversification of *Heliophila*, underscoring its value as a model for studying polyploid evolution in biodiversity hotspots and highlighting the importance of wild genetic resources for conservation and crop improvement.

1. Introduction

The Cape Floristic Region (CFR) of southwestern South Africa is one of the world's most celebrated biodiversity hotspots, renowned for its exceptional vascular plant richness and endemism. Covering ~90,000 km^2, it harbors over 9000 vascular plant species, nearly 70% of which are endemic (Goldblatt and Manning, 2002). As Africa's only floristic region classified among the six global floral kingdoms, Capensis (Takhtajan, 1986), the CFR provides a unique natural laboratory for studying the processes that generate and maintain plant diversity. Ecologically, the CFR largely coincides with the Fynbos biome, a fire-prone, nutrient-poor shrubland of extraordinary species richness (Mucina and Rutherford, 2006). Adjacent lies the Succulent Karoo, a semi-arid biome with lower rainfall yet comparable floral diversity and

* Corresponding author. Central European Institute of Technology, Department of Experimental Biology, Faculty of Science, Masaryk University, Brno, Czechia.
E-mail address: terezie.mandakova@ceitec.muni.cz (T. Mandáková).
Peer review under the responsibility of Editorial Office of Plant Diversity.

https://doi.org/10.1016/j.pld.2026.05.005

Please cite this article as: M. Pouch, P. Šmarda, P. Winter *et al.*, Polyploidy meets the Cape: Evolutionary dynamics of *Heliophila* (Brassicaceae), Plant Diversity, https://doi.org/10.1016/j.pld.2026.05.005

specialization. Together, these biomes form a mosaic shaped by complex climatic gradients and geological history, fostering high ecological heterogeneity.

A striking feature of the Cape flora is its concentration of diversity within a few large evolutionary radiations: roughly half of all species belong to just 33 dominant lineages (Linder, 2003). Rapid speciation and ecological differentiation have produced species-rich genera such as *Erica*, *Pelargonium*, and *Protea* (Jones et al., 2009; Sauquet et al., 2009; Pirie et al., 2016). Although Brassicaceae is not among the most speciose families in the region, its genus *Heliophila* (tribe Heliophileae) stands out with > 100 species exhibiting remarkable morphological and ecological diversity (Mummenhoff et al., 2005; Mandáková et al., 2012; Al-Shehbaz et al., 2022; https://www.inaturalist.org/taxa/119379-Heliophila; Fig. 1B). Its distribution spans Fynbos, Succulent Karoo, Nama-Karoo, Gariep Desert, Namib Desert, and Drakensberg grasslands (Marais, 1970), highlighting exceptional ecological versatility.

Historically, the CFR has been considered relatively poor in polyploid taxa, with many major lineages lacking documented polyploid species (Goldblatt, 1978). However, recent studies have begun to challenge this view. For example, polyploidy has been documented in *Pteronia* (Asteraceae), with ploidy levels ranging from diploid to octoploid across 31% of studied species, and the different cytotypes exhibit distinct environmental preferences (Chumová et al., 2024). Substantial genome size variation and cytotypic diversity have also been reported in *Oxalis* (Krejčíková et al., 2013) and genome size shifts suggest polyploidy in families such as Restionaceae (Linder and Barker, 2014). These findings collectively indicate that whole-genome duplications may be more prevalent and evolutionarily significant within the Cape flora than previously recognized.

Among the polyploid lineages in the CFR, *Heliophila* emerges as a particularly informative system for studying genome evolution and diversification. Cytogenetic data reveal extensive chromosomal variation within the genus, with diploid to polyploid species showing chromosome numbers from $2n = 16$ to 80 (Mandáková et al., 2012). Recent whole-genome sequencing of *H. variabilis* ($2n = 22$) uncovered an allo-octoploid origin of the species, followed by extensive post-polyploid diploidization (Huang et al., 2023). The ancestral octoploid genome ($2n = 60$) likely arose through hybridization between two allotetraploid progenitors ($2n = 30$), each derived from distinct crucifer lineages. Over the past 12 million years, this genome has undergone significant reorganization, fractionation and rediploidization resulting in diploid-like genomes of the extant *Heliophila* species (Huang et al., 2023). These insights highlight polyploidy as a potential driver of adaptive radiation in *Heliophila* and challenge the view of a polyploid-poor CFR. Additionally, recent research on *Heliophila* has shown that lifespan transitions, driven primarily by water availability and seasonality, significantly impact wood anatomy, with perennials evolving in wetter and less seasonal environments (Baczyński et al., 2025). Despite these advances, much of *Heliophila*'s evolutionary history remains unresolved. Its extensive morphological disparity, including variation in growth forms, habitus and especially floral pigmentation, suggests complex evolutionary trajectories. Blue flowers are rare within the Brassicaceae and are known only from a few species of *Aphragmus* Andrz. ex DC., the monospecific *Dichasianthus* Ovcz. and Junussov, *Draba* L., *Eutrema* R.Br., and several *Solms-laubachia* Muschl. Notably, several species of *Heliophila* in its most species-rich clade also bear blue flowers (Mandáková et al., 2012). Recent biochemical and genomic studies further indicate promising, previously uncharacterized variation in *Heliophila* (Huang et al., 2023), opening potential for ornamental plant breeding.

Given the genomic complexity, morphological diversity and broad ecological range of the genus, a comprehensive phylogenetic and ecological analysis is essential for reconstructing its evolutionary history. This study aims to integrate molecular phylogenetics, cytogenetics, ecological reconstructions and floral pigment analyses to unravel the evolutionary dynamics of *Heliophila* and contribute to a broader understanding of speciation and genome evolution within the CFR.

2. Materials and Methods

2.1. Plant material

We analyzed 392 accessions representing 123 taxa of *Heliophila*, including 86 taxa that conform to the current classification, as well as five accessions of *Chamira circaeoides* (Table S1 and Data S1). The remaining taxa comprise names currently treated as synonyms that may require reinstatement, along with others likely to be described as new species in the future. Herbarium specimens associated with each accession were examined and re-identified or reclassified as documented in Table S1. The genus *Chamira*, recognized as the closest relative of *Heliophila* (Nikolov et al., 2019; Hendriks et al., 2023), was selected as the outgroup for phylogenetic analyses. Genomic DNA was extracted from freshly collected or silica-dried leaves using the NucleoSpin Plant II kit (Macherey–Nagel). For chromosome counts, young inflorescences were harvested from field specimens or cultivated plants, fixed in ethanol/acetic acid (3:1) for 24 h at 4 °C, and subsequently stored at −20 °C until further processing.

2.2. Chromosome counting

Selected inflorescences were rinsed in distilled water (twice for 5 min) and then in a citrate buffer containing 10 mM sodium citrate at pH 4.8 (again, twice for 5 min). Subsequently, the tissues were digested in citrate buffer containing a 0.3% enzyme mix (cellulase, cytohelicase, and pectolyase; all from Sigma–Aldrich) at 37 °C for 3 h. Following digestion, individual anthers were dissected on a microscope slide using 20 μl of acetic acid, which were then spread on the slide and placed on a metal hot plate at 50 °C for approximately 30 s. The prepared specimen was then fixed using a freshly prepared ethanol/acetic acid fixative in a 3:1 ratio. After fixation, the preparation was dried using a hair dryer and examined using a phase contrast microscope. For chromosome visualization, DAPI (4′,6-diamidino-2-phenylindole) was employed as a counterstain at a concentration of 2 μg/ml within Vectashield (Vector Laboratories). Chromosome spreads were then photographed using a Zeiss Z2 epifluorescence microscope equipped with a CoolCube CCD camera. Chromosomes were counted in 147 accessions of *Heliophila* (representing 47 distinct taxa) and in four accessions of *Chamira* (Table S1).

2.3. ITS phylogeny

We newly amplified and sequenced the internal transcribed spacers ITS1 and ITS2 in a total of 154 *Heliophila* accessions representing 65 taxa and 4 accessions of *Chamira circaeoides* (Table S1). The resulting sequences have been deposited in GenBank under accession numbers OP196354–OP196492 for ITS1 and OP196493–OP196631 for ITS2. For PCR amplification, we employed a primer combination consisting of ITS1 and ITS4, as described by White et al. (1990). Subsequently, the PCR products were purified using the GenElute PCR Clean-Up Kit

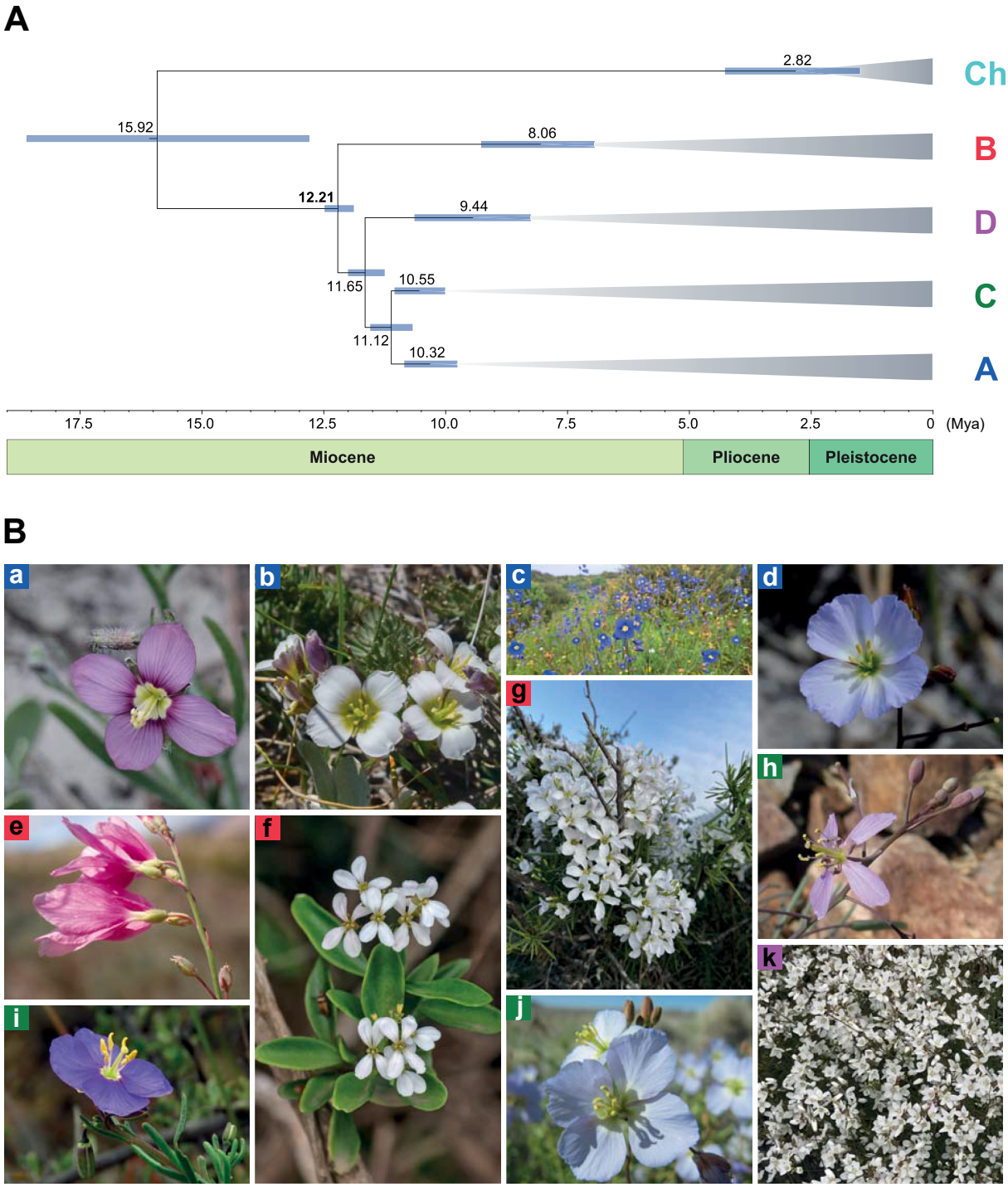


Fig. 1. Molecular dating and morphological diversity of the genus *Heliophila*. **(A)** Scheme of divergence time estimation for the genus *Heliophila* using a maximum-likelihood phylogeny inferred from concatenated single-copy nuclear genes. Node labels represent node ages (Mya) and node bars show 95% HPD intervals. Major clades are labeled A–D; Ch, *Chamira*. This schematic representation is derived from the full divergence time analysis shown in Fig. S5. **(B)** Selected representatives of the genus *Heliophila*: **(a)** *H. africana* (photo: A. McLeod), **(b)** *H. alpina* (photo: B. Cole), **(c)** *H. coronopifolia* (photo: J. Stephenson), **(d)** *H. promontorii* (photo: G. Laidler), **(e)** *H. juncea* (photo: R. Hoekstra), **(f)** *H. scandens* (photo: R. Boon), **(g)** *H. macrosperma* (photo: M. Tessendorf), **(h)** *H. deserticola* var. *micrantha* (photo: G. Laidler), **(i)** *H. suavissima* (photo: G. Laidler), **(j)** *H. deserticola* var. *deserticola* (photo: G. Laidler), and **(k)** *H. meyeri* (photo: C. Merry). Colored squares below the panel labels correspond to the major *Heliophila* clades A–D. All photographs are reproduced with the permission of their authors.

(Macherey–Nagel) and then subjected to Sanger sequencing services provided by Macrogen. In addition to these new sequences, available ITS sequences of *Heliophila* (231 accessions) and *Chamira* (1 accession) were included from our previous study (Dogan et al., 2021). A multiple sequence alignment was generated using MAFFT v.7.453 (Katoh and Standley, 2013) and was reviewed and trimmed using Geneious Prime 2022.1.1 (https://www.geneious.com). The construction of Maximum Likelihood (ML) phylogenetic trees was accomplished using IQ-TREE v.1.6.9 (Nguyen et al., 2015; Hoang et al., 2018) with the application of 1000 ultrafast bootstrap (UF bootstrap) replicates.

2.4. Nuclear single-copy markers

A subset of 141 *Heliophila* accessions representing 85 taxa, encompassing all major clades and subclades identified in the ITS tree, alongside two *Chamira* accessions, were chosen for an in-depth phylogenetic investigation (Table S1). We targeted 48 exonic markers from single-copy nuclear genes, with theoretical amplicons ranging from 440 to 561 bp, as described by Stockenhuber et al. (2015) (Table S2). PCR amplicons were generated using the Access Array System by Fluidigm, following the manufacturer's instructions. Amplicons were pooled per accession and sequenced on an Illumina MiSeq platform, generating 300-bp paired-end reads. Raw Illumina reads have been submitted to the NCBI Sequence Read Archive and will be publicly available upon acceptance (BioProject accession: PRJNA1429314). Sequence quality was assessed using FastQC (Andrews, 2010). Illumina adaptors and low-quality reads were removed using Trimmomatic v.0.39 (Bolger et al., 2014) with the following parameters: LEADING:3, TRAILING:3, SLIDINGWINDOW:5:20, MINLEN:60. Sequence assembly followed the HybPiper v.1.3.1 pipeline (Johnson et al., 2016) using the BWA mapping option. Reads were mapped to target sequences with reads_first.py under default parameters. Gene recovery statistics and sequence length summaries were obtained using get_seq_lengths.py and hybpiper_stats.py. Assembled target sequences were extracted using retrieve_sequences.py. Potential paralogs were assessed using paralog_investigator.py, and flagged loci were examined prior to downstream analyses. Multiple alignments were generated using MAFFT v.7.453 (Katoh and Standley, 2013) and reviewed in Geneious Prime 2022.1.1. Sequences shorter than 250 bp were excluded from the analysis. *Heliophila* accessions whose sequences appeared in less than 15% of alignments were removed from the analysis. The single-gene alignments were concatenated using Geneious Prime. The ML tree was constructed in IQ-TREE v.1.6.9 using a partitioning scheme with separate models for each gene, automatic substitution-model selection, and 1000 ultrafast bootstrap replicates.

To assess the topological congruence of phylogenetic trees and the species constituents of each major clade (ITS tree vs. tree inferred from 48 nuclear single-copy genes), branches with UF bootstrap values below 50% were first collapsed in both trees using Newick Utilities v.1.6.0 (Junier and Zdobnov, 2010). Subsequently, the ultimate tanglegrams were produced using Dendroscope v.3.8.5 (Huson and Scornavacca, 2012).

2.5. Combined/large-scale phylogeny of the genus Heliophila – concatenated ITS and nuclear single-copy markers

To obtain a phylogeny with full taxon sampling of all 392 *Heliophila* accessions plus five *Chamira circaeoides* accessions as an outgroup, two previously generated alignments (ITS1/ITS2 and 48 single-copy nuclear genes) were concatenated. The final ML tree was inferred using IQ-TREE v1.6.9 under a partitioning scheme with gene-specific models and with 1000 ultrafast bootstrap replicates.

2.6. Divergence time estimation

For molecular dating analyses, the aligned matrix of 48 single-copy nuclear markers and the ML species tree, inferred from this alignment, were used as inputs. These analyses were conducted using MCMCtree, which is implemented in PAML v.4.9j (Yang, 2007). To establish a secondary calibration for the *Heliophila* crown group, with an estimated age range of 12.3–10.7 million years ago (Mya), we relied on data derived from the comprehensive study by Hendriks et al. (2023), which focused on the phylogeny of the entire family Brassicaceae. In our analyses, two independent MCMC runs were executed, each consisting of 2.2 million iterations, with samples taken every 100 iterations. The initial 200,000 iterations were discarded as a burn-in. To assess the convergence of the MCMC runs, we conducted an analysis in Tracer v.1.7.1 (Rambaut et al., 2018) and ensured that Effective Sample Size (ESS) values for the tested parameters exceeded 200.

2.7. Ancestral state reconstruction of chromosome numbers

The full taxon-sampling ML tree (combined dataset of ITS and single-copy nuclear genes) was used as the backbone for the ancestral state reconstruction of chromosome numbers. We utilized the R package *ape* v.5.6-2 (Paradis and Schliep, 2019) and the "drop.tip" function to remove accessions/species from the original ML tree with unknown chromosome numbers. To reconstruct the evolution of chromosome numbers, we used the program ChromEvol v.2.0 (Glick and Mayrose, 2014). For accessions with ambiguous chromosome counts (e.g., estimated cytological ranges such as $2n \approx 40$–44), we encoded these as multiple equally probable chromosome-number states, following the ChromEvol2 procedure for ambiguous observations. Each integer value within the observed range was supplied to ChromEvol with identical prior probability.

To evaluate the effect of root-state specification, we first conducted analyses without constraining the ancestral chromosome number. Because empirical observations indicate that several early-diverging lineages within the genus possess $n = 13$ chromosome pairs, and because Brassicaceae are characterized predominantly by descending dysploidy following polyploidization events (Mandáková and Lysak, 2018), we additionally performed analyses with the root fixed at $n = 13$ as a biologically informed scenario. Comparing constrained and unconstrained analyses allowed us to assess the robustness of inferred evolutionary patterns to alternative root-state assumptions.

A total of eight standard models implemented in ChromEvol2 were tested, and the best-fitting model was selected using the Akaike Information Criterion (AIC).

2.8. Life form ancestral state reconstruction

As in the case of chromosome number evolution, the ML tree inferred from the combined dataset was used for the ancestral state reconstruction of life forms. We used the R package *ape* v.5.6-2 (Paradis and Schliep, 2019) and the "drop.tip" function to selectively remove accessions with unclear life form from the tree. R package *phytools* v.2.3.0 (Revell, 2024) was employed in the analysis of life form evolution of the genus *Heliophila*. Firstly, four possible models for a two-state character (annual vs. perennial) were tested using the "fitMk" function implemented in phytools, selecting the best model based on AIC. Marginal ancestral state reconstruction on the given phylogeny and predicted model of evolution was done using the function "ancr" of the phytools. Estimated posterior probabilities of each state in each node were plotted on the tree.

2.9. Environmental variables

We obtained geographical data for our georeferenced samples by downloading (i) Chelsea v.2.1 bioclimatic variables (Karger et al., 2017) from the Chelsea v.2.1 repository, which provides data at a resolution of approximately 1 km. This data is accessible at https://chelsa-climate.org/downloads/ (Tables S3 and S4). Additionally, we acquired (ii) predicted nutrient content data for South Africa (Cramer et al., 2019) at a resolution of about 250 m,

and (iii) altitude data from the digital elevation model supplied by the Shuttle Radar Topography Mission through WorldClim (Hijmans et al., 2005), which has horizontal and vertical resolutions of about 30 m and 16 m, respectively (Table S4). Data downloads and extractions were performed using the raster R package (Hijmans, 2022). We refined the dataset by excluding variables with a significant amount of missing data and those that were less pertinent or remained constant for our study group. To ensure data suitability for statistical analysis, we examined data distributions within the four major *Heliophila* clades. We made necessary transformations to allow for standard parametric statistics. Variable values were compared among the clades using ANOVA and Tukey post hoc tests in R. To obtain an overall perspective on the potential environmental preferences of individual *Heliophila* clades, we conducted a linear discriminant analysis (DA) involving both samples and their environmental variables. To simplify the resulting diagrams, we excluded bioclimatic variables (Tables S3 and S4) that exhibited smaller differences among the compared clades and variables with high intercorrelations (Pearson correlation coefficient > 0.9). Samples with missing soil data (21 out of 378) were also omitted from the DA analysis. The DA was performed using the R package *MorphoTools* v.1.1 (Koutecký, 2015). Differences in environmental preferences of annuals and perennials within the clades were tested with linear mixed effect models using *nlme* package (Pinheiro et al., 2025) in R, treating lifeform as a categorical predictor and clade membership (A, B handled together with D, and C) as a random effect.

2.10. High-performance liquid chromatography (HPLC)

Flavonoids in *Heliophila* species were analyzed building upon previous knowledge of flavonoids in *H. coronopifolia* (Saito et al., 2011). Analytical HPLC of flavonoids was conducted using an LC 10A system manufactured by Shimadzu Corporation in Kyoto, Japan. A C18 column (4.6 mm × 250 mm) from Waters in Milford, Massachusetts, USA was employed, and the column temperature was maintained at 40 °C. Elution was performed at a flow rate of 1 ml/min, and detection was carried out using a photodiode array detector. For the analysis of flavonoids, approximately 0.3–0.5 mg of dried petals were extracted with 200 μl of a solvent mixture called MAW (MeOH:Acetic acid:H_2O = 4:1:5, v/v/v) at room temperature for 2 h. The HPLC elution process followed a 40-min linear gradient, transitioning from 20% to 85% of solvent B (consisting of 1.5% H_3PO_4, 20% HOAc, 25% MeCN in H_2O) in solvent A (1.5% H_3PO_4 in H_2O). Subsequently, a 5-min re-equilibration at 20% solvent B was conducted, specifically for the separation of flavonoids.

3. Results

3.1. Four major Heliophila clades revealed by the ITS phylogeny

We inferred a maximum-likelihood (ML) phylogeny from ITS1/ITS2 sequences using 385 *Heliophila* accessions, with five accessions of the sister species *Chamira circaeoides* included as an outgroup (Fig. S1). The analysis recovered four major *Heliophila* clades, here designated A–D. Clade A contained 192 accessions (49.9% of the dataset), clade B 66 accessions (17.1%), clade C 116 accessions (30.1%), and clade D 11 accessions (2.9%). Clades B and D formed a weakly supported superclade, whereas clades A and C grouped together in a second superclade. Node support for the four major clades was high, with Ultrafast Bootstrap (UF bootstrap) values ranging from 86% to 100%. In contrast, support for the B–D and A–C superclades was relatively low, at 52% and 82%, respectively.

3.2. Consistency and robustness of major Heliophila clades in both ITS and single-copy nuclear gene phylogenies

Using 48 single-copy nuclear markers, we reconstructed a species tree based on a final set of 141 *Heliophila* accessions representing the main subclades of the ITS phylogeny. Initially, we amplified and sequenced these markers in 149 *Heliophila* and two *Chamira* accessions. Eight *Heliophila* accessions with < 15% marker recovery were excluded prior to analysis. The mean marker coverage across the remaining accessions was 77.1%. Concatenation of the 48 markers yielded an alignment of 24,126 base pairs.

The inferred ML species tree (Fig. S2) recovered four major *Heliophila* clades (A–D), fully consistent with the ITS phylogeny (Fig. S1). Clade A comprised 71 accessions (50.4%), clade B 21 (14.9%), clade C 40 (28.4%), and clade D 9 accessions (6.4%). The topology indicated that clade B represents the earliest-diverging lineage within the *Heliophila* crown group, followed by the divergence of clade D from the sister superclade formed by clades A and C. All major clades (A–D) received maximal support (UF bootstrap = 100%), confirming the stable and well-resolved placement of clade B within the crown group.

Both phylogenies – based on rDNA ITS1/ITS2 and on 48 single-copy nuclear markers– consistently recovered the same four infrageneric evolutionary clades (A–D), despite the smaller taxon sampling in the nuclear dataset (Fig. S3). This congruence underscores the robustness of the major clades for subsequent phylogenetic and ecological analyses. Examination of within–clade relationships revealed largely congruent topologies, with only minor discrepancies in clades B and D. In clade C, the placement of two basal subclades differed between datasets, and the internal topology of clade A showed substantial incongruence, likely influenced by several poorly supported nodes in the ITS tree (Fig. S1; UF bootstrap < 50%).

3.3. Concatenated phylogeny

The final combined ML phylogeny (Fig. S4) was inferred from a concatenated dataset comprising the ITS regions and 48 nuclear single-copy genes. The analysis included all 392 *Heliophila* accessions analyzed in this study, along with 5 accessions of *Chamira circaeoides* used as the outgroup. Four major clades were recovered: clade A consisted of 194 accessions (49.5%), clade B of 67 accessions (17.1%), clade C of 120 accessions (30.6%), and clade D of 11 accessions (2.8%). As expected, the topology of the four clades was congruent with that of the nuclear-concatenated tree, with clade B in the sister (ancestral) position closest to *Chamira*, followed by clade D and the sister clades A and C.

3.4. Diversification of major Heliophila clades occurred between 12 to 7 Mya

Divergence time estimates for the genus *Heliophila*, based on the ML species tree inferred from the nuclear-marker dataset (Figs. 1A and S5), indicate that the most recent common ancestor (MRCA) of *Heliophila* originated approximately 12.2 Mya (95% HPD: 12.49–11.88 Mya) in the middle Miocene. The four major evolutionary clades diverged in relatively rapid succession over the following ~5 million years. Clade B represents the earliest diverging lineage, followed by the separation of clade D from the AC superclade around 11.6 Mya, and the subsequent split between clades A and C at approximately 11.1 Mya.

The crown ages of the major clades fall within a relatively narrow temporal window: clades A and C originated around 10–11 Mya, clade D around 9–10 Mya, and clade B slightly later, around 8 Mya. Together, these estimates indicate a rapid Miocene radiation,

with most deep divergences within *Heliophila* occurring between 12 and 7 Mya.

3.5. Chromosome evolution of the genus Heliophila was driven by descending dysploidy and neopolyploidization

Descriptive statistics of chromosome numbers were calculated only from accessions with unambiguous counts; however, accessions with chromosome-number ranges were incorporated into the ancestral state reconstruction following the procedure outlined in the Methods. We counted chromosomes in 147 *Heliophila* accessions (Table S1; Fig. S4 and S6), identifying 11 different chromosome numbers ($2n$ = 16–80). Clade A exhibited five chromosome numbers ($2n$ = 16, 18, 20, 36, and 80), with $2n$ = 20 being the most common (42 of 74 counts). Seven chromosome numbers were observed among just 22 accessions of clade B ($2n$ = 16, 20, 22, 24, 26, 28, and 32). In clade C, two chromosome numbers were detected, with $2n$ = 22 predominant (38 of 46 counts), whereas clade D showed two chromosome numbers among five accessions ($2n$ = 26 and 44). Across the entire genus, $2n$ = 20 and 22 were the most frequent chromosome numbers (29.9% and 30.6%, respectively).

Ancestral chromosome number reconstruction was first conducted without constraining the root state. Under this unconstrained analysis, the CONST_RATE_DEMI_EST model was identified as the best-fitting (AIC = 387.2). The highest root probability was inferred for n = 11 (F = 0.655), followed by n = 10 (F = 0.294), whereas n = 12 (F = 0.044) and n = 13 (F = 0.003) received only minor support. This reconstruction implied ancestral ascending dysploidy preceding the establishment of higher chromosome numbers in several early-diverging lineages. Because such extensive ancestral ascending dysploidy is inconsistent with the predominance of descending dysploidy documented in Brassicaceae (Mandáková and Lysak, 2018), we additionally performed an analysis with the root constrained to n = 13 as a biologically informed alternative scenario. Under this constraint, the same model (CONST_RATE_DEMI_EST) was recovered as the best-fitting (AIC = 394.6) (Fig. S7 and Fig. 2). All subsequent interpretations of chromosome evolution patterns presented below are based on this root-constrained analysis.

The presence of minimally diploidized genomes with $2n$ = 26 across clades B, C, and D suggests that this cytotype represented an intermediate stage in post-polyploid diploidization prior to the diversification of the major *Heliophila* clades. Chromosome numbers equal to or exceeding *ca.* n = 16–30 likely reflect neopolyploidization, a phenomenon found in all four clades.

For clade A, the inferred ancestral chromosome number was $2n$ = 20 (Fig. S7 and 2). Descending dysploidy occurred repeatedly, with reductions from 20 → 18 → 16, and the transition to 18 evolving independently multiple times. Several neopolyploid cytotypes were detected, including $2n$ = 36 (tetraploidy from $2n$ = 18) in *Heliophila elata* and *H. cornuta*, and $2n$ = 80 (octoploidy from $2n$ = 20) in *H. sp. aff. macowaniana*.

In clade B, the ancestral chromosome number was inferred as $2n$ = 26. Species within this clade underwent progressive descending dysploidy (26 → 16), followed by further neopolyploidization and subsequent dysploid reductions. A striking case of rapid intraspecific chromosome evolution was observed in the *H. juncea* subclade, which exhibited five cytotypes: $2n$ = 16, 28, 32, ~44–46, and ~60–70.

In clade C, chromosome numbers were reduced from $2n$ = 26 to the dominant $2n$ = 22. Two *Heliophila carnosa* accessions represented probable neopolyploids: one with 44–46 chromosomes (tetraploidy from $2n$ = 22 or 23) and one with $2n$ = 80 (octoploidy from $2n$ = 20).

Chromosome numbers were available for only five accessions in the less speciose clade D. One accession possessed 13 chromosome pairs ($2n$ = 26), whereas four others were likely neopolyploids with $2n$ = 40–44 (tetraploidy from $2n$ = 20 or 22).

3.6. Life form ancestral state reconstruction

A total of 390 *Heliophila* accessions were included in the analysis of life-form evolution, comprising 248 annuals (63.6%) and 142 perennials (36.4%). Only two accessions that could not be assigned to a taxon identity were excluded from the original dataset (Table S1). We tested four evolutionary models for a two-state character in the phytools package. The ARD model (all rates different), which had the lowest AIC value (159.689), was selected for marginal ancestral state reconstruction of life-form evolution (Fig. 3).

The most recent common ancestor of *Chamira* + *Heliophila*, as well as the crown ancestor of *Heliophila*, was inferred to have been predominantly annual, with marginal probabilities of 99.8% and 92.2%, respectively. Interestingly, all 67 accessions from the basal clade B were perennial, whereas all 10 examined accessions from the small clade D were annual. Annuals also dominated the large clades A and C: clade A comprised 143 annuals (73.7%) and 51 perennials (26.3%), while clade C comprised 95 annuals (79.8%) and 24 perennials (20.2%).

Several transitions from the ancestral annual state to the perennial state were recovered within the predominantly annual A + C + D superclade: two in clade C (the *Heliophila suavissima* and *H. carnosa* clades) and eleven in clade A, including lineages such as *H. alpina*, *H. subulata*, *H. cornuta*, *H. lactea*, and a complex of accessions related to *H. linearis* and *H. elata*, distributed across multiple parts of clade A. Conversely, only two transitions from perennial to annual life forms were inferred in clade A, represented by accessions of the annual *H. stenocarpa* and *H. arenaria* nested within the predominantly perennial *H. elata* clade. Both annual and perennial life forms were observed among accessions of *H. lactea*.

In the comparative visual analysis of major chromosomal alterations or neopolyploid occurrences (Fig. S7 and 2) in conjunction with lifeform data (Fig. 3), the transitions between annual and perennial lifeforms do not show an evident association with changes in chromosome numbers or neopolyploidy events. This observation holds, with potential exception in the case of *Heliophila elata* subsp. *pillansii*, where neopolyploidy appears to coincide with the evolutionary shift towards perennial life form. However, substantiating this correlation necessitates an examination of ploidy levels in other taxonomically related perennial taxa for a comprehensive understanding.

3.7. Environmental variables

The four main *Heliophila* clades (A–D) exhibit significant variations in nearly all tested environmental variables (Data S1 and S2). These differences highlight the distinct environmental requirements and preferences of each *Heliophila* clade, shaping their ability to inhabit specific types of ecosystems, which are effectively summarized in the discriminant analysis (Fig. S8).

The basal clade B, consisting of perennial species, tends to inhabit higher elevations with an overall colder climate, higher rainfall and humidity. These areas are also characterized by more acidic soils, as well as a higher content of total nitrogen and organic compounds (Figs. 4 and S8; Data S2).

The related annual clade D contains only a few species, making it difficult to evaluate its ecological preferences based on the characteristics of the few sampled localities. Clade D seems to

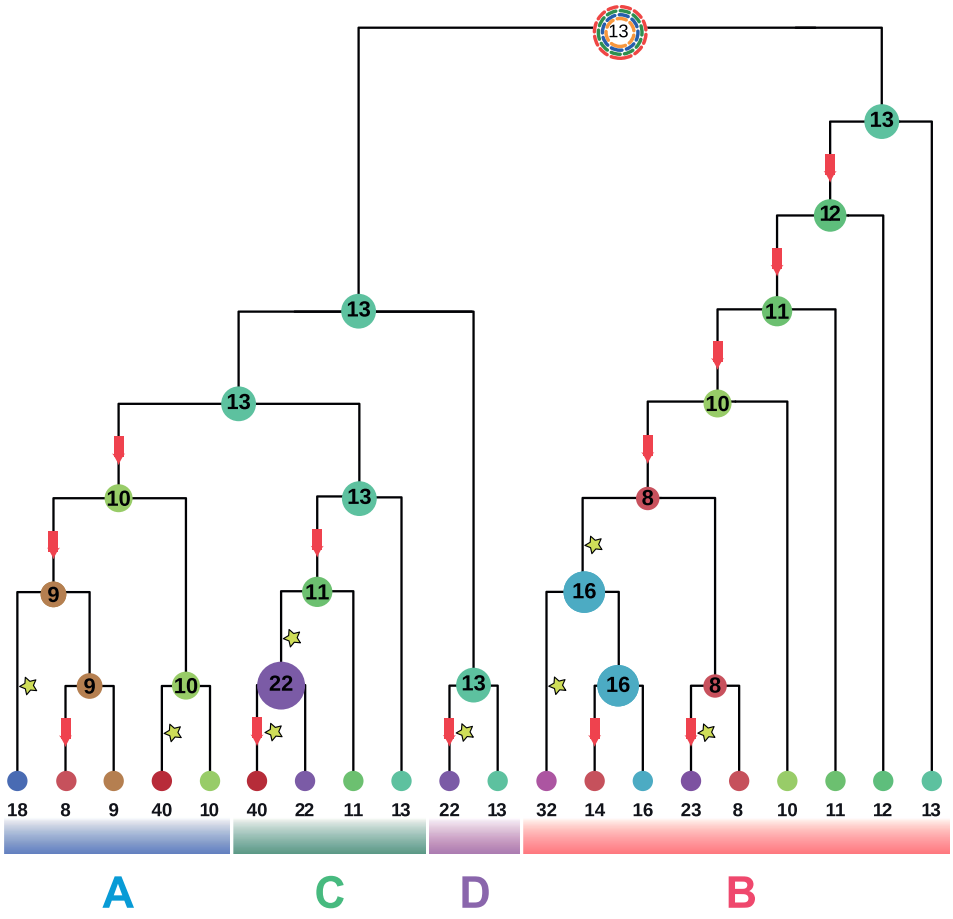


Fig. 2. Chromosome number evolution in the genus *Heliophila*. The phylogenetic framework is subdivided into four major clades (A–D). Extant species exhibit a wide range of haploid chromosome numbers ($n = 8, 9, 10, 11, 12, 13, 14, 16, 18, 22, 32$, and 40). Ancestral state reconstruction indicates an ancestral haploid number of $n = 13$. Red arrows denote descending dysploidy events, whereas yellow stars indicate whole-genome duplication events. This schematic summary is based on the full ChromEvol ancestral state reconstruction presented in Fig. S7.

share many characteristics with its sister clade B, such as a tendency to grow in areas with more rainfall and humidity, characterized by low total nitrogen and carbon. However, contrary to perennial clade B, annuals in clade D tend to occur at lower elevations with an overall warmer climate (Fig. 4). In any of the tested soil or bioclimatic variables, this clade does not significantly differ from all three remaining *Heliophila* clades (Data S2 and Fig. S8).

Many species of clade C are confined to arid plateaus characterizing the N and E margins of the distribution range of the Cape *Heliophila*. This makes clade C ecologically the most distinct of the *Heliophila* clades, differing in many soil and bioclimatic characteristics from the rest of the annual or perennial *Heliophila* groups (Figs. 4 and S8; Data S2). Notably, many species from this clade thrive in arid environments characterized by reduced precipitation, higher temperature maxima, greater diurnal and annual temperature fluctuations, lower near-surface humidity, and low net primary productivity. These habitats are often associated with slightly saline soils (having higher pH, conductance, and concentrations of Na and K) and a low content of organic C and total N.

Compared to clade C, the species of sister clade A mainly occur in Cape lowlands. They are characterized by a preference for lower altitudes and higher temperatures, and a longer and warmer growing season (defined as sum of temperatures and the number of days with temperatures above a given temperature threshold). The soil properties and other bioclimatic characteristics do not differ much from those of the other *Heliophila* clades (Figs. 4 and S8; Data S3). The notable exception is *H. alpina*, occupying a disjunct position on the high Drakensberg escarpment and isolated from the core clade members.

Perennial species exhibited pronounced distinctions from their annual counterparts within the clade across a spectrum of soil and bioclimatic parameters (Table S5 and Data S3). Notably, perennials consistently manifested preferences divergent from annuals, displaying an inclination towards regions characterized by elevated rainfall, particularly during the driest and warmest quarters of the year. These regions were typically situated in colder climates at higher altitude, characterized by a shorter vegetation season, indicative of fewer growing degree days. Regarding soil attributes, perennials displayed a proclivity for environments with elevated levels of nitrogen and organic carbon.

3.8. Chemical characterization of flavonoids in Heliophila species

A total of 14 individuals from seven *Heliophila* species were analyzed for flavonoid composition using HPLC (Table S6 and Fig. S9). Known anthocyanins in the genus *Heliophila* (Pigments, Pigs. 1–6 (Saito et al., 2011)), newly identified anthocyanins (DM, Pigs. 7 and 8) (Tatsuzawa et al., 2007)) and a flavonol (Pig. 9) were identified in *H. juncea*, *H. deserticola*, *H. variabilis*, *H. arenaria* and *H.* sp. cf. *coronopifolia*. Interestingly, *H. seselifolia* and *H. digitata* exhibited unknown anthocyanins as major components.

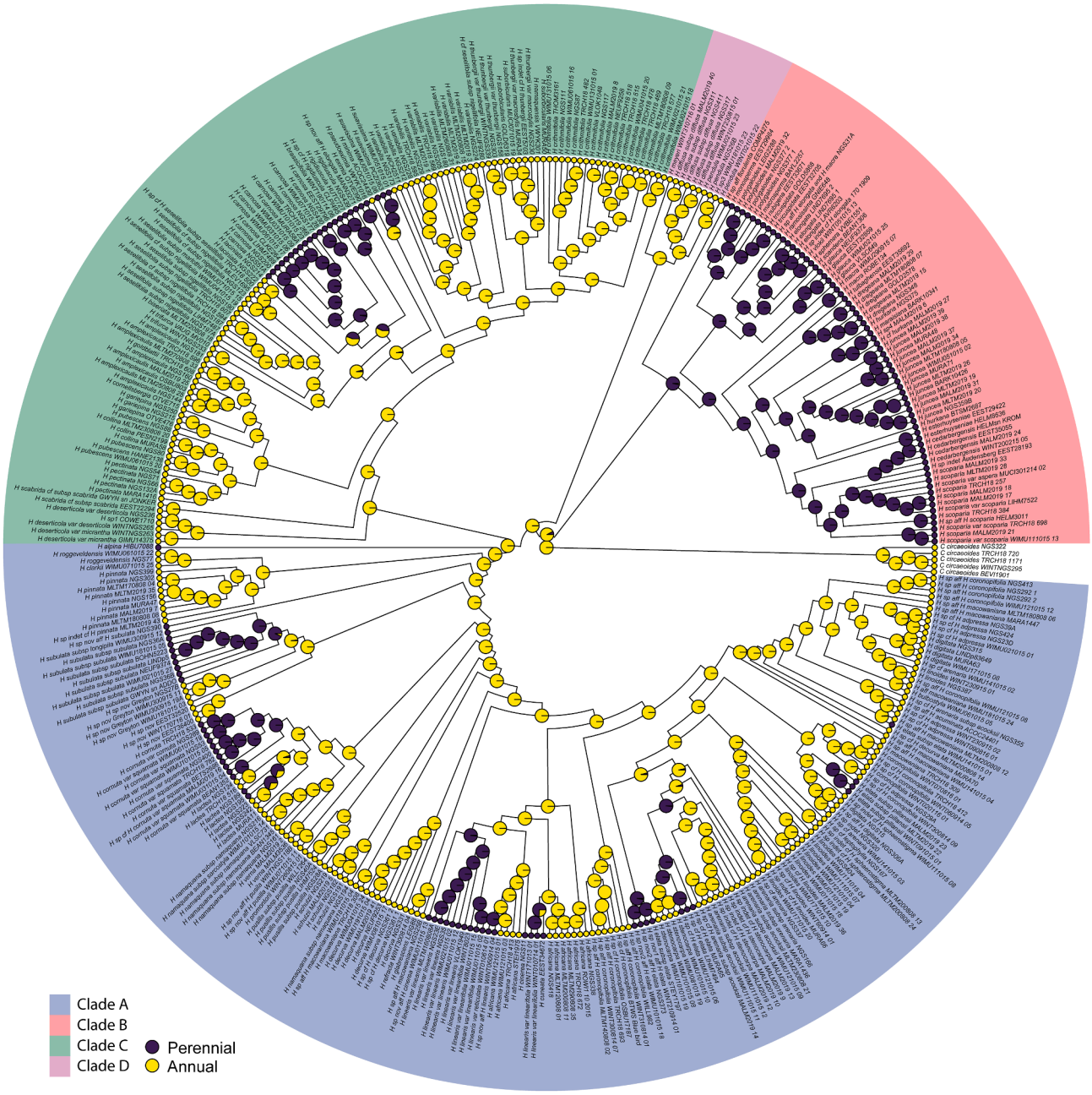


Fig. 3. Marginal ancestral state reconstruction of life form using concatenated ITS1/ITS2 and 48 nuclear single-copy genes. The tree with predicted character states evolution of annual (yellow) and perennial (black) lifeforms. Posterior probability of each state is shown using pie charts on the respective nodes (ancestors) together with actual lifeform of extant taxa on terminal branches. The major clades are labeled A–D. Branch lengths do not reflect evolutionary distances.

Additionally, acylated cyanidin glycosides were found in *H. juncea* (clade B), *H. deserticola* (clade C), and *H. variabilis* (clade C) as major components. Acylated delphinidin glycosides were observed in *H. arenaria* (clade A) and *H.* sp. cf. *coronopifolia* (clade A), along with trace amounts of acylated cyanidin glycosides. Additionally, an acylated kaempferol glycoside (Pig. 9) was detected, known for its intermolecular co-pigment effect that imparts a blue hue to the flower.

Intriguingly, unidentified anthocyanins were identified as major components in *Heliophila deserticola* and *H. variabilis*, with Pigs. 7 and 8. This suggests subtle differences in the chemical structures of acylated anthocyanins in these two species. Further insights into the chemical structures of these unidentified anthocyanins could shed light on these distinctions. The presence of flavonoid 3′,5′-hydroxylase in clade A (blue) was notably inhibited, leaving only the flavonoid 3′-hydroxylase active, resulting in the red color of clade B and the white color of clade C, based on the chemical structure of anthocyanins in *H. juncea*, *H. deserticola*, *H. variabilis*, *H. arenaria* and sp. cf. *H. longifolia*. Moreover, the blue flowers of clade A coexist with flavonol (Pig. 9), intensifying the bluish flower color of *Heliophila*, which may indicate the presence of pollinators with distinct color preferences compared to those of clades B and C. *H. digitata* and *H. seselifolia* could be subject to further investigation once the chemical structures of the unidentified anthocyanins are elucidated.

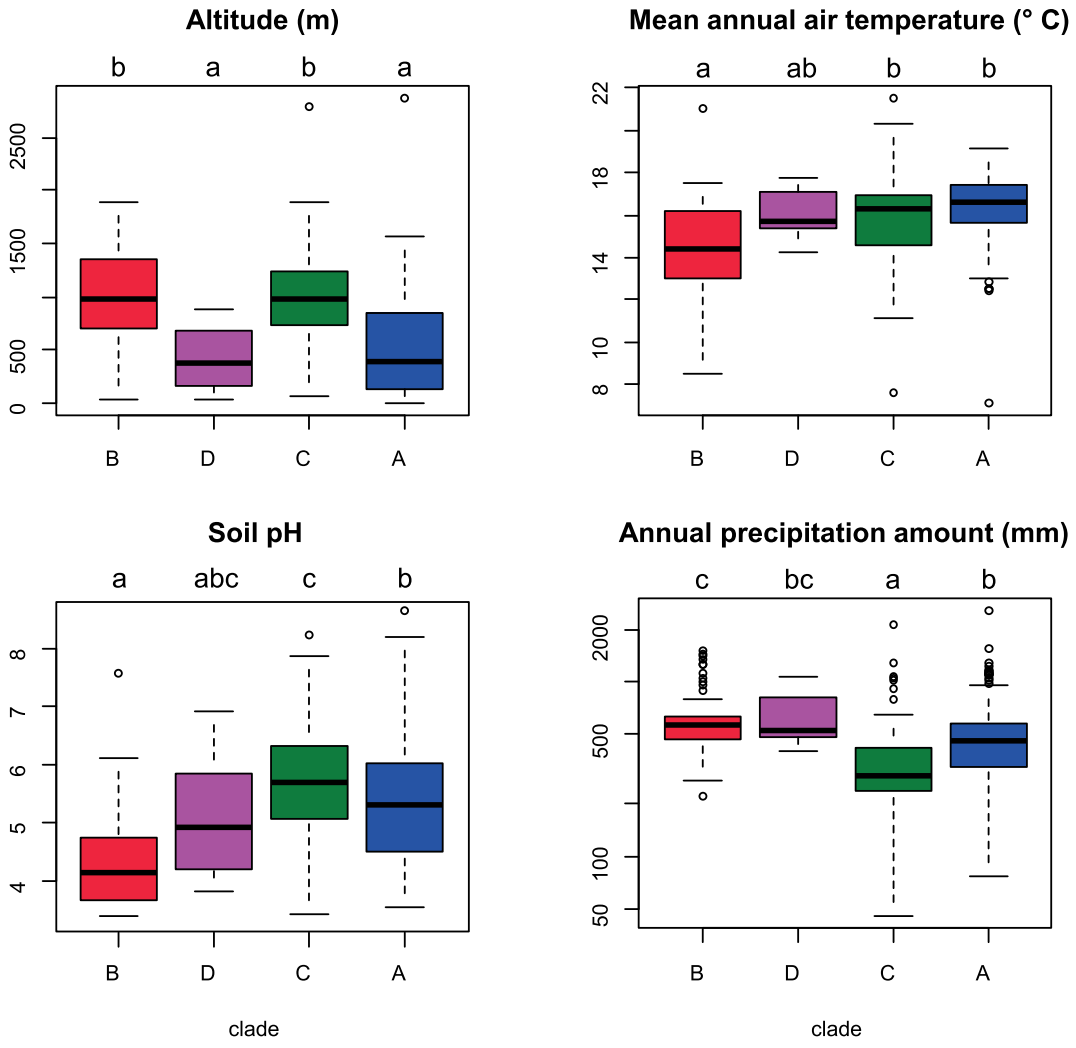


Fig. 4. Four selected environmental variables showing major differences in environments occupied by species from individual *Heliophila* clades: warmer lowlands (clade A), colder and wetter highlands on acidic soils (clade B), dry and warm highlands with less acidic soils (clade C), and slightly wetter lowlands (clade D). Boxplots show median (horizontal line), interquartile range (box), non-outlier range (whiskers), and outliers (circles). Non-significant pairwise comparisons ($p > 0.05$, Tukey post-hoc test) are indicated by shared letters above the boxplots.

4. Discussion

4.1. Miocene aridification, edaphic mosaics, and life-history shifts: Environmental drivers of Heliophila divergence

Our study advances previous phylogenetic work on *Heliophila* by expanding taxon sampling and employing a robust low-copy nuclear, multi-locus framework. Earlier analyses (Mandáková et al., 2012; Mummenhoff et al., 2005) relied largely on ITS, which, while useful, can be confounded by concerted evolution and hybridization (Álvarez and Wendel, 2003). By leveraging a nuclear multilocus dataset, we recover improved resolution of relationships within the genus and identify the early divergence of clade B, which retains several ancestral traits and is most closely related to the sister genus *Chamira*.

The timing of the four major *Heliophila* clades (*ca.* 12–10 Mya) coincides with the Late Miocene aridification of southern Africa (Verboom et al., 2009), a climatic transition that facilitated the rise of modern Fynbos and Succulent Karoo biomes, systems marked by exceptional diversity and turnover (Linder, 2003; Schnitzler et al., 2012). Escalating aridity combined with fine-scale soil heterogeneity likely imposed strong selective filters, fostering ecological differentiation. Consistent with this, clades exhibit distinct environmental preferences across nearly all tested parameters: clade B (dominantly perennial) occurs mainly in high-altitude, high-rainfall sites with acidic soils, whereas annual-dominated clades A and C occupy warmer, drier regions with nutrient-poor, sandy substrates. These patterns indicate tight coupling between life-history strategy and habitat specialization, echoing broader CFR trends of climatic and edaphic association.

Clade D provides a complementary perspective: most collections originate from marginal (non-core) Cape vegetation, specifically mesic “Fynbos Thickets” (Rebelo et al., 2006), riparian thickets, and ravine-forest edges. Notably, *Chamira circaeoides*, the lone species of Chamireae endemic to the Cape, is likewise restricted to remnant Fynbos Thickets (L. Mucina, unpublished). These mesic habitats are hypothesized to predate the dominance of Fynbos/renosterveld shrublands (Mucina, 2023), suggesting ancient ecological links at the base of Heliophileae and offering insight into the tribe's origins within the Cape flora.

Within the CFR's compact geography, the interplay of climate gradients, diverse substrates, and fire regimes has generated exceptional habitat heterogeneity and specialization (Cowling et al., 2015). Our results indicate that *Heliophila* diversified along

this environmental mosaic via contrasting adaptive strategies: drought escape in annuals (rapid flowering and seed set) supported their persistence in arid, nutrient-poor contexts (clades A and C), while longer-lived forms in cooler, wetter uplands (clade B) optimized resource retention and reproductive success through perennial growth. These trajectories align with biogeographic dynamics reported in other Cape lineages - repeated arid-adaptation in *Pelargonium* (Jones et al., 2009), biome-linked evolution in *Protea* (Valente et al., 2010), and montane–lowland niche partitioning in Restionaceae (Linder and Bouchenak-Khelladi, 2015). The elevational stratification observed among *Heliophila* clades fits these broader patterns: older lineages occupy uplands, while younger, more rapidly diversifying lineages radiate into lower, environmentally dynamic zones (Hughes and Atchison, 2015).

Finally, the high species richness of *Heliophila* in arid biomes points to the combined roles of drought tolerance (Mucina et al., 2024) and edaphic specialization in shaping evolutionary success, paralleling geophytic diversification across the Cape (Manning and Goldblatt, 2012). Soils likely impose selection on root architecture, nutrient acquisition, and water-use efficiency, structuring strategies across clades. While niche conservatism is common in many CFR groups (Verboom et al., 2009), the pronounced differentiation among clades A, B, and C in *Heliophila* indicates adaptive shifts in both life history and environmental tolerance under climatic and edaphic pressures.

4.2. Shifting strategies: annual–perennial transitions and their evolutionary drivers

Our analyses indicate that the ancestral life form in *Heliophila* was most likely annual, yet the consistently perennial nature of clade B suggests a more nuanced trajectory (Mucina et al., 2024). If clade B reflects ancestral traits, annuality may have evolved later within the predominantly annual superclade A + C + D, whose common ancestor is strongly inferred to be annual (Fig. 3). This timing coincides with Late Miocene aridification, when selective pressures favored short life cycles, rapid reproduction, and drought escape - hallmarks of annual species (Boyko et al., 2023). Similar transitions linked to climatic unpredictability have been documented in other Cape lineages, including *Pelargonium* and Restionaceae (Jones et al., 2009; Linder and Barker, 2014).

Following the split from clade B, annual-dominated lineages underwent multiple independent reversals to perenniality, typically associated with shifts to higher elevations, cooler climates, and edaphic niches offering stable water availability. These reversals align with broader angiosperm trends where perenniality evolves under conditions of reduced seasonality and greater resource stability (Friedman, 2020). The presence of perennials in basal clade B further complicates the picture, raising the possibility that perenniality was ancestral and annuality emerged as an adaptation to increasing drought stress. This evolutionary lability underscores the dynamic interplay between life-history traits and environmental pressures.

Recent work on wood anatomy provides additional insight: lifespan transitions strongly influence wood functional traits in *Heliophila*, with perennials in wetter, less seasonal environments exhibiting distinct anatomical adaptations (Baczyński et al., 2025). Secondary woodiness may occur more frequently in warmer niches, though current evidence remains limited. These findings highlight how life-history evolution intersects with structural and physiological traits, reinforcing the role of ecological context in shaping diversification.

In sum, annual–perennial transitions in *Heliophila* reflect a complex evolutionary response to climatic shifts and habitat heterogeneity, mediated by genomic and developmental mechanisms. Future research integrating phylogenomics, ecological modeling, and functional studies of flowering-time and woodiness genes will be critical for disentangling these drivers. Understanding how life-history strategies adapt to environmental change is essential for predicting plant responses under ongoing climate shifts and for conservation planning in biodiversity hotspots.

4.3. Genomic remodeling as an engine of diversification: Polyploidy and dysploidy in Heliophila

Polyploidy is a pervasive evolutionary force in flowering plants, generating genomic complexity, phenotypic novelty, and opportunities for diversification (Soltis et al., 2015; Van de Peer et al., 2017). In Brassicaceae, whole-genome duplications (WGDs) followed by extensive diploidization have repeatedly shaped lineage evolution (Mandáková and Lysak, 2018). Post-polyploid restructuring often involves descending dysploidy - a reduction in chromosome number - which stabilizes karyotypes, promotes reproductive isolation, and facilitates ecological expansion (Schranz et al., 2006; Mandáková et al., 2012).

Our results reveal similar dynamics in Heliophileae. Chromosome numbers ranging from $2n = 16$ to 80 reflect a history of ancient polyploidy, recurrent neopolyploidy, and descending dysploidy, processes that underpin the diversification of major infrageneric clades. The recent chromosome-scale assembly of *Heliophila variabilis* (a meso-octoploid, $2n = 22$) provides a window into these events: its genome likely originated from an allo-octoploid ancestor ($2n = 60$) formed by hybridization between two allotetraploid progenitors (Huang et al., 2023). Subsequent diploidization, genome downsizing, and chromosomal reduction produced the extant karyotypic diversity, with additional WGDs in some lineages amplifying complexity. These patterns mirror broader Brassicaceae trends, where ancient polyploidy coupled with chromosomal restructuring has driven radiation in multiple clades (Hohmann et al., 2015; Mandáková et al., 2017). In *Heliophila*, descending dysploidy appears particularly influential, enabling genome stabilization and possibly facilitating adaptive radiation. Polyploidy itself may act as a catalyst by generating novel genetic variation, expanding ecological amplitude, and fostering reproductive isolation through chromosomal incompatibilities (Van de Peer et al., 2017). Such genomic flexibility likely supported *Heliophila*'s colonization of contrasting environments - from arid Karoo plains to moist montane systems - echoing similar ploidy-linked niche differentiation in Cape radiations such as *Pelargonium* and Restionaceae (Bouchenak-Khelladi and Linder, 2017; Jones et al., 2009).

Beyond structural changes, polyploidy may have influenced functional traits, including flowering time, stress tolerance, and floral morphology - features known to enhance ecological success in Brassicaceae (Lukens et al., 2004; Schranz and Osborn, 2000). These links between genome architecture and phenotype underscore the role of polyploidy and dysploidy as engines of diversification in *Heliophila*.

4.4. Color innovation and its evolutionary significance: Floral pigments in Heliophila

Floral pigmentation is a key trait influencing pollinator interactions, reproductive isolation, and adaptation (Grotewold, 2006). Although based on a limited taxon sampling, our exploratory pigment analysis reveals a broader diversity of anthocyanins in *Heliophila* than previously recognized. Several species beyond *H. coronopifolia* show evidence of delphinidin-based pigmentation,

suggesting that blue flowers may have evolved multiple times in the genus (Davies et al., 2012; Grotewold, 2006). These preliminary results indicate possible F3′5′H activity in parts of clade A, but functional verification is still required. Flower color transitions often correlate with pollination syndromes: bee-pollinated species favor blue/violet, while bird or butterfly-pollinated flowers tend toward red or pink (Bradshaw and Schemske, 2003; Wessinger and Rausher, 2012). Given the predominance of generalist insect pollinators in *Heliophila*, blue pigmentation may reflect bee preferences for short wavelengths (Chittka and Menzel, 1992; Schiestl and Johnson, 2013), although any link between color shifts and pollination remains speculative.

As blue flowers are rare in Brassicaceae due to the typical absence of F3′5′H (He et al., 2020), these findings point to potentially novel modifications of the flavonoid pathway as drivers of floral diversification, but the evolutionary timing and frequency of such innovations remain uncertain. Because anthocyanins can contribute both to pollinator interactions and abiotic stress tolerance, including UV protection, drought resilience, and pathogen defense (Landi et al., 2015), floral color variation in *Heliophila* may reflect combined ecological pressures. Post-polyploid genome restructuring could also play a role in pigment diversification (Edger et al., 2017; Landis et al., 2018). Overall, these results highlight promising avenues for future work, including functional assays of F3′5′H, broader metabolomic sampling, and targeted pollinator studies, to clarify the genetic and ecological drivers of color evolution in *Heliophila*.

5. Conclusion

Our comprehensive phylogenetic analysis of nearly 400 accessions of *Heliophila* reveals four well-supported intrageneric clades that diverged during the Late Miocene (*ca.* 12–10 Mya). Chromosome number variation ($2n = 16–80$) reflects a complex history of descending dysploidy and recurrent neopolyploidization, occurring independently across major clades. Ancestral state reconstruction indicates that the common ancestors of *Chamira* and *Heliophila* were predominantly annual, yet life-history evolution within *Heliophila* has been highly dynamic: at least 13 independent shifts from annuality to perenniality occurred within the largely annual A + C + D superclade, contrasted by only two reversals to annuality. Ecological analyses demonstrate distinct environmental preferences among clades, suggesting that chromosomal changes and life-history transitions have contributed to niche differentiation and habitat specialization. Together, these findings underscore the intricate interplay between genome evolution, life-history strategies, and ecological adaptation in shaping the remarkable diversity of *Heliophila*.

CRediT authorship contribution statement

Milan Pouch: Investigation, Visualization, Writing – original draft, Writing – review & editing. **Petr Šmarda**: Investigation, Visualization, Writing – review & editing. **Pieter Winter**: Plant material collection, Investigation, Writing – review & editing. **Fumi Tatsuzawa**: Investigation, Writing – review & editing. **Ihsan A. Al-Shehbaz**: Plant material collection, Investigation, Writing – review & editing. **Klaus Mummenhoff**: Plant material collection, Writing – review & editing. **Ladislav Mucina**: Plant material collection, Conceptualization, Data curation, Writing – review & editing. **Martin A. Lysak**: Plant material collection, Conceptualization, Funding acquisition, Writing – review & editing. **Terezie Mandáková**: Plant material collection, Conceptualization, Funding acquisition, Investigation, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by the Czech Science Foundation (project no. 26-20931S) and project TowArdsNextGENeration Crops (no. 17 CZ.02.01.01/00/22_008/0004581) of the ERDF Programme Johannes Amos Comenius. Core Facility Genomics of CEITEC Masaryk University is gratefully acknowledged for the obtaining of the scientific data presented in this paper. Computational resources were provided by the e-INFRA CZ project (ID:90254), supported by the Ministry of Education, Youth and Sports of the Czech Republic. LM acknowledges logistic support of the Iluka Chair in Vegetation Science and Biogeography at the Murdoch University, Perth, Australia. The South African National Biodiversity Institute is thanked for providing a vehicle for a collecting trip.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.pld.2026.05.005.

References

Al-Shehbaz, I.A., Mandáková, T., Mummenhoff, K., et al., 2022. *Heliophila verna* (Brassicaceae), a new species from the Northern Cape of South Africa. Phytotaxa 555, 195–199. https://doi.org/10.11646/phytotaxa.555.2.8.

Álvarez, I., Wendel, J.F., 2003. Ribosomal ITS sequences and plant phylogenetic inference. Mol. Phylogenet. Evol. 29, 417–434. https://doi.org/10.1016/S1055-7903(03)00208-2.

Andrews, S., 2010. FastQC: a Quality Control Tool for High Throughput Sequence Data [software]. http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Baczyński, J., Oskolski, A.A., Winter, P.J.D., et al., 2025. Lifespan outperforms climate as a predictor of wood functional traits, but secondary woodiness shows no clear climatic pattern in *Heliophila*, a diverse clade from the cape floristic region. Ann. Bot. mcaf046. https://doi.org/10.1093/aob/mcaf046.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Bouchenak-Khelladi, Y., Linder, H.P., 2017. Frequent and parallel habitat transitions as driver of unbounded radiations in the Cape flora. Evolution 71, 2548–2561. https://doi.org/10.1111/evo.13364.

Boyko, J.D., Hagen, E.R., Beaulieu, J.M., et al., 2023. The evolutionary responses of life-history strategies to climatic variability in flowering plants. New Phytol. 240, 1587–1600. https://doi.org/10.1111/nph.18971.

Bradshaw, H.D., Schemske, D.W., 2003. Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. Nature 426, 176–178. https://doi.org/10.1038/nature02106.

Chittka, L., Menzel, R., 1992. The evolutionary adaptation of flower colours and the insect pollinators' colour vision. J. Comp. Physiol. 171, 171–181. https://doi.org/10.1007/BF00188925.

Chumová, Z., Havlíčková, E., Zeisek, V., et al., 2024. Deciphering *Pteronia*'s evolution in the cape floristic region: a comprehensive study disputes polyploid deficiency and affirms diploid radiation. Plant J. 119, 2236–2254. https://doi.org/10.1111/tpj.16914.

Cowling, R.M., Potts, A.J., Bradshaw, P.L., et al., 2015. Variation in plant diversity in mediterranean-climate ecosystems: the role of climatic and topographical stability. J. Biogeogr. 42, 552–564. https://doi.org/10.1111/jbi.12429.

Cramer, M.D., Wootton, L.M., Van Mazijk, R., et al., 2019. New regionally modelled soil layers improve prediction of vegetation type relative to that based on global soil models. Divers. Distrib. 25, 1736–1750. https://doi.org/10.1111/ddi.12973.

Davies, K.M., Albert, N.W., Schwinn, K.E., 2012. From landing lights to mimicry: the molecular regulation of flower colouration and mechanisms for pigmentation patterning. Funct. Plant Biol. 39, 619–638. https://doi.org/10.1071/FP12195.

Dogan, M., Pouch, M., Mandáková, T., et al., 2021. Evolution of tandem repeats is mirroring post-polyploid cladogenesis in *Heliophila* (Brassicaceae). Front. Plant Sci. 11, 607893. https://doi.org/10.3389/fpls.2020.607893.

Edger, P.P., Smith, R., McKain, M.R., et al., 2017. Subgenome dominance in an interspecific hybrid, synthetic allopolyploid, and a 140-year-old naturally

established neo-allopolyploid monkeyflower. Plant Cell 29, 2150–2167. https://doi.org/10.1105/tpc.17.00010.
Friedman, J., 2020. The evolution of annual and perennial plant life histories: ecological correlates and genetic mechanisms. Annu. Rev. Ecol. Evol. Syst. 51, 461–481. https://doi.org/10.1146/annurev-ecolsys-110218-024638.
Glick, L., Mayrose, I., 2014. ChromEvol: assessing the pattern of chromosome number evolution and the inference of polyploidy along a phylogeny. Mol. Biol. Evol. 31, 1914–1922. https://doi.org/10.1093/molbev/msu122.
Goldblatt, P., 1978. An analysis of the flora of Southern Africa: its characteristics, relationships, and orgins. Ann. Mo. Bot. Gard. 65, 369–436. https://doi.org/10.2307/2398858.
Goldblatt, P., Manning, J.C., 2002. Plant diversity of the Cape region of Southern Africa. Ann. Mo. Bot. Gard. 89, 281–302. https://doi.org/10.2307/3298566.
Grotewold, E., 2006. The genetics and biochemistry of floral pigments. Annu. Rev. Plant Biol. 57, 761–780. https://doi.org/10.1146/annurev.arplant.57.032905.105248.
He, Q., Lu, Q., He, Y., et al., 2020. Dynamic changes of the anthocyanin biosynthesis mechanism during the development of heading Chinese Cabbage (*Brassica rapa* L.) and *Arabidopsis* under the control of BrMYB2. Front. Plant Sci. 11, 593766. https://doi.org/10.3389/fpls.2020.593766.
Hendriks, K.P., Kiefer, C., Al-Shehbaz, I.A., et al., 2023. Global Brassicaceae phylogeny based on filtering of 1,000-gene dataset. Curr. Biol. 33, 4052–4068.e6. https://doi.org/10.1016/j.cub.2023.08.026.
Hijmans, R.J., 2022. raster: Geographic Data Analysis and Modeling [R package]. https://CRAN.R-project.org/package=raster.
Hijmans, R.J., Cameron, S.E., Parra, J.L., et al., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25, 1965–1978. https://doi.org/10.1002/joc.1276.
Hoang, D.T., Chernomor, O., von Haeseler, A., et al., 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522. https://doi.org/10.1093/molbev/msx281.
Hohmann, N., Wolf, E.M., Lysak, M.A., et al., 2015. A time-calibrated road map of Brassicaceae species radiation and evolutionary history. Plant Cell 27, 2770–2784. https://doi.org/10.1105/tpc.15.00482.
Huang, Y., Guo, X., Zhang, K., et al., 2023. The meso-octoploid Heliophila variabilis genome sheds a new light on the impact of polyploidization and diploidization on the diversity of the Cape flora. Plant J. 116, 446–466. https://doi.org/10.1111/tpj.16383.
Hughes, C.E., Atchison, G.W., 2015. The ubiquity of alpine plant radiations: from the Andes to the Hengduan Mountains. New Phytol. 207, 275–282. https://doi.org/10.1111/nph.13230.
Huson, D.H., Scornavacca, C., 2012. Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. Syst. Biol. 61, 1061–1067. https://doi.org/10.1093/sysbio/sys062.
Johnson, M.G., Gardner, E.M., Liu, Y., et al., 2016. HybPiper: extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. Appl. Plant Sci. 4, 1600016. https://doi.org/10.3732/apps.1600016.
Jones, C.S., Bakker, F.T., Schlichting, C.D., et al., 2009. Leaf shape evolution in the South African genus *Pelargonium* L'Her. (Geraniaceae). Evolution 63, 479–497. https://doi.org/10.1111/j.1558-5646.2008.00552.x.
Junier, T., Zdobnov, E.M., 2010. The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. Bioinformatics 26, 1669–1670. https://doi.org/10.1093/bioinformatics/btq243.
Karger, D.N., Conrad, O., Böhner, J., et al., 2017. Climatologies at high resolution for the earth's land surface areas. Sci. Data 4, 170122. https://doi.org/10.1038/sdata.2017.122.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780. https://doi.org/10.1093/molbev/mst010.
Koutecký, P., 2015. MorphoTools: a set of R functions for morphometric analysis. Plant Syst. Evol. 301, 1115–1121. https://doi.org/10.1007/s00606-014-1153-2.
Krejčíková, J., Sudová, R., Lučanová, M., et al., 2013. High ploidy diversity and distinct patterns of cytotype distribution in a widespread species of oxalis in the greater cape floristic region. Ann. Bot. 111, 641–649. https://doi.org/10.1093/aob/mct030.
Landi, M., Tattini, M., Gould, K.S., 2015. Multiple functional roles of anthocyanins in plant-environment interactions. Environ. Exp. Bot. 119, 4–17. https://doi.org/10.1016/j.envexpbot.2015.05.012.
Landis, J.B., Soltis, D.E., Li, Z., et al., 2018. Impact of whole-genome duplication events on diversification rates in angiosperms. Am. J. Bot. 105, 348–363. https://doi.org/10.1002/ajb2.1060.
Linder, H.P., 2003. The radiation of the Cape flora, Southern Africa. Biol. Rev. Camb. Phil. Soc. 78, 597–638. https://doi.org/10.1017/s1464793103006171.
Linder, H.P., Barker, N.P., 2014. Does polyploidy facilitate long-distance dispersal? Ann. Bot. 113, 1175–1183. https://doi.org/10.1093/aob/mcu047.
Linder, H.P., Bouchenak-Khelladi, Y., 2015. The causes of southern African spatial patterns in species richness: speciation, extinction and dispersal in the Danthonioideae (Poaceae). J. Biogeogr. 42, 914–924. https://doi.org/10.1111/jbi.12474.
Lukens, L.N., Quijada, P.A., Udall, J., et al., 2004. Genome redundancy and plasticity within ancient and recent Brassica crop species. Biol. J. Linn. Soc. 82, 665–674. https://doi.org/10.1111/j.1095-8312.2004.00352.x.
Mandáková, T., Li, Z., Barker, M.S., et al., 2017. Diverse genome organization following 13 independent mesopolyploid events in Brassicaceae contrasts with convergent patterns of gene retention. Plant J. 91, 3–21. https://doi.org/10.1111/tpj.13553.
Mandáková, T., Lysak, M.A., 2018. Post-polyploid diploidization and diversification through dysploid changes. Curr. Opin. Plant Biol. 42, 55–65. https://doi.org/10.1016/j.pbi.2018.03.001.
Mandáková, T., Mummenhoff, K., Al-Shehbaz, I.A., et al., 2012. Wholegenome triplication and species radiation in the southern African tribe Heliophileae (Brassicaceae). Taxon 61, 989–1000. https://doi.org/10.1002/tax.615006.
Manning, J., Goldblatt, P., 2012. Plants of the Greater Cape Floristic Region, Volume 1: the Core Cape Flora, Strelitzia. South African National Biodiversity Institute, Pretoria.
Marais, W., 1970. Cruciferae. In: Codd, L.E., De Winter, B., Killick, D.J., Rycroft, H.B. (Eds.), Flora of Southern Africa. Government Printer, Pretoria, pp. 1–118.
Mucina, L., 2023. Biomes of the Southern Hemisphere, Biome Ecology. Springer, Cham. https://doi.org/10.1007/978-3-031-26739-0.
Mucina, L., Mummenhoff, K., Winter, P., et al., 2024. Drought and life-history strategies in *Heliophila* (Brassicaceae). New Phytol. 241, 532–534. https://doi.org/10.1111/nph.19352.
Mucina, L., Rutherford, M.C. (Eds.), 2006. The Vegetation of South Africa, Lesotho and Swaziland, Strelitzia. South African National Biodiversity Institute, Pretoria.
Mummenhoff, K., Al-Shehbaz, I.A., Bakker, F.T., et al., 2005. Phylogeny, morphological evolution, and speciation of endemic Brassicaceae genera in the Cape Flora of southern Africa. Ann. Mo. Bot. Gard. 92, 400–424.
Nguyen, L.T., Schmidt, H.A., von Haeseler, A., et al., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274. https://doi.org/10.1093/molbev/msu300.
Nikolov, L.A., Shushkov, P., Nevado, B., et al., 2019. Resolving the backbone of the Brassicaceae phylogeny for investigating trait diversity. New Phytol. 222, 1638–1651. https://doi.org/10.1111/nph.15732.
Paradis, E., Schliep, K., 2019. Ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35, 526–528. https://doi.org/10.1093/bioinformatics/bty633.
Pinheiro, J., Bates, D., R Core Team, 2025. nlme: Linear and Nonlinear Mixed Effects Models [R package]. https://CRAN.R-project.org/package=nlme.
Pirie, M.D., Oliver, E.G.H., Mugrabi de Kuppler, A., et al., 2016. The biodiversity hotspot as evolutionary hot-bed: spectacular radiation of Erica in the cape floristic region. BMC Evol. Biol. 16, 190. https://doi.org/10.1186/s12862-016-0764-3.
Rambaut, A., Drummond, A.J., Xie, D., et al., 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67, 901–904. https://doi.org/10.1093/sysbio/syy032.
Rebelo, A.G., Boucher, C., Helme, N., et al., 2006. Fynbos biome. In: The Vegetation of South Africa, Lesotho and Swaziland, pp. 52–219.
Revell, L.J., 2024. Phytools 2.0: an updated R ecosystem for phylogenetic comparative methods (and other things). PeerJ 12, e16505. https://doi.org/10.7717/peerj.16505.
Saito, N., Tatsuzawa, F., Toki, K., et al., 2011. The blue anthocyanin pigments from the blue flowers of *Heliophila coronopifolia* L. (Brassicaceae). Phytochemistry 72, 2219–2229. https://doi.org/10.1016/j.phytochem.2011.07.020.
Sauquet, H., Weston, P.H., Anderson, C.L., et al., 2009. Contrasted patterns of hyperdiversification in Mediterranean hotspots. Proc. Natl. Acad. Sci. U.S.A. 106, 221–225. https://doi.org/10.1073/pnas.0805607106.
Schiestl, F.P., Johnson, S.D., 2013. Pollinator-mediated evolution of floral signals. Trends Ecol. Evol. 28, 307–315. https://doi.org/10.1016/j.tree.2013.01.019.
Schnitzler, J., Graham, C.H., Dormann, C.F., et al., 2012. Climatic niche evolution and species diversification in the Cape flora, South Africa. J. Biogeogr. 39, 2201–2211. https://doi.org/10.1111/jbi.12028.
Schranz, M., Osborn, T., 2000. Novel flowering time variation in the resynthesized polyploid Brassica napus. J. Hered. 91, 242–246. https://doi.org/10.1093/jhered/91.3.242.
Schranz, M.E., Lysak, M.A., Mitchell-Olds, T., 2006. The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. Trends Plant Sci. 11, 535–542. https://doi.org/10.1016/j.tplants.2006.09.002.
Soltis, P.S., Marchant, D.B., Van de Peer, Y., et al., 2015. Polyploidy and genome evolution in plants. Curr. Opin. Genet. Dev., Genomes Evol 35, 119–125. https://doi.org/10.1016/j.gde.2015.11.003.
Stockenhuber, R., Zoller, S., Shimizu-Inatsugi, R., et al., 2015. Efficient detection of novel nuclear markers for Brassicaceae by transcriptome sequencing. PLoS One 10, e0128181. https://doi.org/10.1371/journal.pone.0128181.
Takhtajan, A., 1986. Floristic Regions of the World. University of California Press, Berkeley.
Tatsuzawa, F., Toki, K., Saito, N., et al., 2007. Four acylated cyanidin 3-sambubioside-5-glucosides from the purple-violet flowers of Lobularia maritima. Heterocycles 71, 1117. https://doi.org/10.3987/COM-07-11034.
Valente, L.M., Reeves, G., Schnitzler, J., et al., 2010. Diversification of the African genus *Protea* (Proteaceae) in the cape biodiversity hotspot and beyond: equal rates in different biomes. Evolution 64, 745–760. https://doi.org/10.1111/j.1558-5646.2009.00856.x.
Van de Peer, Y., Mizrachi, E., Marchal, K., 2017. The evolutionary significance of polyploidy. Nat. Rev. Genet. 18, 411–424. https://doi.org/10.1038/nrg.2017.26.

Verboom, G.A., Archibald, J.K., Bakker, F.T., et al., 2009. Origin and diversification of the greater cape flora: ancient species repository, hot-bed of recent radiation, or both? Mol. Phylogenet. Evol. 51, 44–53. https://doi.org/10.1016/j.ympev.2008.01.037.

Wessinger, C.A., Rausher, M.D., 2012. Lessons from flower colour evolution on targets of selection. J. Exp. Bot. 63, 5741–5749. https://doi.org/10.1093/jxb/ers267.

White, T.J., Bruns, T.D., Lee, S.B., et al., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M.A., Gelfand, D.H., Sninsky, J.J., et al. (Eds.), PCR Protocols: a Guide to Methods and Applications. Academic Press, New York, pp. 315–322. https://doi.org/10.1016/B978-0-12-372180-8.50042-1.

Yang, Z., 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24, 1586–1591. https://doi.org/10.1093/molbev/msm088.