Systematics and Evolutionary Ecology of *Osteospermum* section *Polygalinae* DC.

Lauren M.C. James

Dissertation presented in fulfilment of the requirements for the degree of

Master of Science specialising in Biological Sciences

Project supervisors:

Prof. G. Anthony Verboom, Dr Nicola G. Bergh and Prof. Allan G. Ellis

Administrative supervisor:

Prof. Michael D. Cramer

Department of Biological Sciences, Faculty of Science, University of Cape Town

September 2024

University of Cape Town

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

University of Cape Town

Contents

List of tables and figures……………………………………………………………..…………2

Abstract………………………………………………………………………….……..……….5

Acknowledgements……………..…………………………………….…..…….……..………6

Chapter 1: Introduction……………………………………………..…………..…..…..……7

Chapter 2: Species delimitation and phylogenetic relationships in *Osteospermum* section *Polygalinae*………………………………………………………………………………..……22

Chapter 3: The evolution of myrmecochory and its impact on the spatial scale of gene flow in *Osteospermum* section *Polygalinae*…………………………………………………………………73

Chapter 4: Synthesis……………………..…………………………………….………………104

Appendices…………………………………………………………………..……..……..…114

List of tables and figures

Chapter 1

Figure 1: Map of South Africa showing the boundary of the Cape Floristic Region and the distribution of *Osteospermum* section *Polygalinae* based on herbarium records.

Figure 2: Examples of the habitat variation in which species in *Osteospermum* section *Polygalinae* occur.

Figure 3: The defining traits and examples of trait variation of species in *Osteospermum* section *Polygalinae*.

Figure 4: Examples of the variation in growth form of species in *Osteospermum* section *Polygalinae*.

Figure 5: The four fruit forms on which Norlindh (1943) based his taxonomy of *Osteospermum* section *Polygalinae*.

Chapter 2

Figure 1: Map of South Africa showing the distribution of species in *Osteospermum* section *Polygalinae* and the localities of the populations sampled in this study.

Figure 2: Bayesian 90% majority-rule consensus tree of concatenated ETS and ITS sequences of species in *Osteospermum.*

Figure 3: Maximum likelihood phylogeny of 79 *Osteospermum* section *Polygalinae* accessions from different populations inferred from concatenated GBS data filtered to < 30 % missing per site.

Figure 4: Principal component analysis ordination of genotyping-by-sequencing single-nucleotide polymorphism data from 79 *Osteospermum* section *Polygalinae* populations.

Figure 5: Species delimitation in Clade 1.

Figure 6: Species delimitation in Clade 2.

Figure 7: Species delimitation in Clade 3.

Figure 8: Species delimitation in Clade 4.

Table 1: Taxa currently recognised in *Osteospermum* section *Polygalinae.*

Table 2: Details of the morphological traits scored for each population in the study.

Table 3: Comparison of the GBS dataset for phylogenetic inference when filtered for different proportions of missing data per site.

Table 4: The percentage of sampled populations of each lineage occurring in the vegetation type groups classified by Mucina & Rutherford (2006) as determined from the Vegetation Map of

South Africa, Lesotho and Swaziland (South African National Biodiversity Institute, 2009) and informed by field observations.

Table 5: The alignment of the species delimitations determined in this study with the current taxonomy of *Osteospermum* section *Polygalinae* based on morphology matching the type specimens.

Chapter 3

Figure 1: Photographs of *O. corymbosum*, the fruit dispersal experiment treatments and ants transporting fruits.

Figure 2: Phylogeny of *Osteospermum* section *Polygalinae* from Chapter 2 showing fruit traits of species at the tips and reconstructions of ancestral traits at the nodes.

Figure 3: Survival curves showing the removal rates of *O. corymbosum* fruits with elaiosomes and without elaiosomes when offered to ants in cafeteria experiments.

Figure 4: Comparison of fruit removal between *O. corymbosum* and Species E in cafeteria experiments at transect markers.

Figure 5: Isolation-by-distance plots for the seven individual species analysed and for all seven species on a common set of axes.

Table 1: Pairwise contrasts showing the differences in treatment effects on fruit removal by ants as modelled by the Cox proportional hazards model.

Table 2: Pairwise contrasts showing the differences in the number of fruits transported into ant nests between the three treatments as modelled by a binomial model.

Table 3: Model parameters for the comparison of IBD between species produced by a linear model with genetic distance as the response variable, and geographic distance, species, and the interaction between geographic distance and species as fixed-effects.

Appendices

Appendix 1: List of collection details for each population sampled for this study and the datasets in which they were used.

Appendix 2: Samples used for Bayesian phylogenetic analysis to test for monophyly of *O*. sect. *Polygalinae*, including preliminary morphological form assignment for specimens putatively in *O*. sect. *Polygalinae*, collection details and the gene regions sampled.

Appendix 3: Bootstrap majority-rule consensus trees for parsimony analysis of ETS and ITS sequences of taxa in *Osteospermum.*

Appendix 4: The eight continuous traits and ratios and six qualitative binary traits scored for each population.

Appendix 5: Maximum likelihood phylogeny of 79 *Osteospermum* section *Polygalinae* populations inferred from concatenated GBS data filtered to < 20 % missing per site.

Appendix 6: Ancestral gene pool assignment of each population in the four major clades inferred using sparse nonnegative matrix factorization (sNMF).

Appendix 7: PCA ordination of genotyping-by-sequencing (GBS) single-nucleotide polymorphism data for Clade 4 showing PC5 and PC6.

Abstract

Osteospermum L. (Asteraceae) is a genus of ca. 80 species of mainly yellow-flowered annual and perennial daisies, having most of its diversity in the Greater Cape Floristic Region of South Africa. The genus is characterised by remarkable variation in fruit structure and dispersal biology, putatively including species that employ myrmecochory (dispersal by ants). Myrmecochory is thought to be exceptionally prevalent in the species-rich Fynbos Biome and has gained interest as a possible driver of speciation due to short dispersal distances limiting gene flow and possible fitness advantages provided by seed deposition in ant nests. Due to its diversity of fruit types, *Osteospermum* provides an excellent system for the study of dispersal biology and its effect on diversification. However, the species-level taxonomy of *Osteospermum* has received little attention since it was last revised by Norlindh in 1943, with the current species concepts being ambiguous and possibly not representing evolutionary independent units and they are therefore inappropriate for testing evolutionary hypotheses. This study therefore first uses an integrative taxonomic approach to reassess the species-level taxonomy of *Osteospermum* section *Polygalinae* DC, a fynbos-centred group that contains a putatively myrmecochorous species, but within which species boundaries and relationships are especially unclear. The monophyly of *O.* sect. *Polygalinae* is confirmed using phylogenetic analysis of ETS and ITS sequences, while analysis of genomic (genotyping-by-sequencing, GBS), morphological and ecological data reveals *O.* sect. *Polygalinae* to comprise between 14 and 16 species, with well-resolved relationships. In the context of this revised species taxonomy and inferred phylogenetic tree, ancestral reconstructions of fruit traits are used, in conjunction with data from field experiments, to test the hypothesis of adaptation to ant dispersal. Myrmecochory, which is associated with the appearance of an elaiosome-like structure on the fruit, is found to have originated once, in *O. corymbosum*. The remaining species do not appear specifically adapted for dispersal, suggesting that they are passively dispersed. Field experiments confirm that elaiosomes in *O. corymbosum* attract ants and facilitate rapid transport of fruits underground. While previous studies suggest that myrmecochory stimulates diversification by reducing dispersal-mediated gene flow, a comparison of isolation-by-distance (IBD) based on GBS data finds limited support for this hypothesis, the IBD slope of *O. corymbosum* exceeding that of only four of the six species to which it was compared. Thus, the importance of myrmecochory as a driver of diversification in the Cape, as a consequence of a lower spatial scale of gene flow, may have been overstated. More importantly, perhaps, ant dispersal presents an important adaptation for avoiding predation during the inter-fire period and fire itself, which is critical in environments such as the Cape fynbos and may be particularly important for species that flower early in the inter-fire period.

Acknowledgements

I would like to thank my supervisors Tony Verboom, Nicola Bergh and Allan Ellis for their support and guidance throughout this project, for all they have taught me and for providing timeous and detailed feedback on drafts, but especially for sharing their enthusiasm and interest in the Cape flora, and in so doing, further enhancing mine. I am also grateful to Michael Cramer for taking up the role as my UCT supervisor when Tony left UCT.

I am immensely grateful for the support I have received from my lab-mates during this time, especially Seth Musker, Thaabiet Parker and Robert Sadler. I particularly want to thank Seth for his substantial help with bioinformatics and many technical conversations about analyses, Thaabiet for his all-round advice and guidance in the laboratory and Robert for initially introducing me to *Osteospermum* section *Polygalinae* and the many discussions we have had about taxonomy and fruit dispersal since.

I also want to thank all those who went with me on field trips, especially Sandy Smuts, Zafar Monier, Tony Verboom, Nicola Bergh and Hugh Verboom, and also Sonja Stock, Thaabiet Parker and Seth Musker. I am grateful to all the landowners who let me collect specimens on their properties and wish to thank the citizen scientists who assisted me with finding collection sites, the staff of Grootvadersbosch Nature Reserve, and Ellie and Rudi Goossens who collected a specimen for me in the Groot Winterhoek Mountains in the Eastern Cape. I also want to thank the staff of the Compton Herbarium and UCT Department of Biological Sciences for their assistance during this project, especially Isabella Gongota, Dulcinea Paulse and Michelle Smith. I am extremely grateful to the South African National Biodiversity Institute (Joan Wrench Kirstenbosch Scholarship) and UCT for funding, without which this work would not have been possible.

Finally, I will always be deeply grateful to my parents, Nick James and Helen James, sisters, Katherine James and Jacqui James and granny, Margaret Barber for their love, support, encouragement and shared appreciation of the natural world, which has and will continue to inspire me.

Chapter 1: Introduction

The species is a fundamental unit of biodiversity, representing evolutionarily distinct entities at the population-level that are independent from other such units (de Queiroz, 2005; Freudenstein et al., 2017). As such, species are essential for communication about biodiversity and making conservation decisions (Myers et al., 2000; Coates, Byrne & Moritz, 2018). They are also a fundamental unit of analysis in many areas of biological research, including evolutionary biology, ecology, and much of ecophysiology. In evolutionary biology, for example, the species is central to studies of speciation and diversification (Barraclough & Nee, 2001; Wiens, 2004), while in community ecology it is the basic unit of community assembly and composition (Webb et al., 2002; Vellend, 2010). Given the importance of species in how they impact our perception of biodiversity, conservation and research, it is critical that they are defined in a manner that is evolutionarily and biologically meaningful.

Species conceptualization and delimitation is challenging, being limited both by our understanding of evolutionary processes and the data and tools available for practical species discovery. These limitations have resulted in considerable disagreement over the definition of species and culminated in the proposal of many different species concepts (reviewed by: Mayden, 1997; de Queiroz, 1998; Coyne & Orr, 2004). Recently, however, there has been some consensus, with a view of species as independently evolving metapopulation lineages (de Queiroz, 1998, 2005, 2007) gaining widespread acceptance (e.g., Shaffer & Thomson, 2007; Wiens, 2007; Padial et al., 2010; Schlick-Steiner et al., 2010; Yeates et al., 2011; Yi et al., 2023; Vences, Miralles & Dufresnes, 2024), and an acknowledgment that it is not so much the question of what a species is, but rather the criteria that are best used to recognize them, that is the cause of debate (de Queiroz, 1998, 2005, 2007). This "unified species concept" leaves the criteria used to recognise independent lineages up to the taxonomist, as is appropriate for the organism being studied and based on the data and tools available at the time (de Queiroz, 2005, 2007). It also shifts the concept of species from being a definite entity to a hypothesis that can be tested using the best available evidence. This is appropriate since the methods and data which can be used to test species hypotheses change over time (de Queiroz, 2005; Rouhan & Gaudeul, 2014). Currently, developments in genomics and their application to taxonomy are enabling lineage relationships to be studied at the level of individuals and populations, providing new resolution and allowing investigation of macro and microevolutionary processes (Shaffer & Thomson, 2007; Camargo & Sites, 2013). However, although molecular data are helpful for identifying independent lineages, criteria are still needed for determining species boundaries (Freudenstein et al., 2017). Integrative taxonomy (Dayrat, 2005; Will, Mishler & Wheeler, 2005), the practise of using multiple different sources of data to delimit species, has emerged as a promising tool for establishing evolutionary meaningful species and bringing stability to taxonomy (Padial et al., 2009; Cicero et al., 2021).

Speciation is the process by which biological lineages split to give rise to new species. While details of the process vary from one instance to another, speciation is ultimately a process of genetic divergence, leading eventually to the build-up of reproductive incompatibility and evolutionary independence (Coyne & Orr, 2004). Genetic divergence leading to speciation is generally thought to mostly occur in allopatry, being driven by genetic drift and local adaptation and facilitated by reduced gene flow between populations, although new species can sometimes also arise in sympatry (Dieckmann & Doebeli, 1999; Coyne & Orr, 2004; Rundle & Nosil, 2005; Nosil, 2008). Understanding the spatial scale of gene flow, and its underlying determinants, is thus important for understanding speciation (Lenormand, 2002). In mobile organisms such as most animals, the mechanism of gene flow is the movement of juvenile or adult individuals, while in sessile organisms such as plants it is modulated by the movement of propagules such as pollen and seeds (Slatkin, 1985). Seed dispersal and pollination biology are thus central to understanding speciation in plants. Together with extinction, the process of lineage death, speciation underpins global patterns of species richness, with biodiversity hotspots commonly being attributed to regionally high rates of speciation and/or low rates of extinction (e.g. Linder, 2008; Rolland et al., 2014; Igea & Tanentzap, 2019; Harvey et al., 2020).

The Cape Floristic Region (CFR) of South Africa has remarkable vascular plant species richness, accommodating an estimated 9383 vascular plant species in an area of ca. 90 000 km^2 (Manning & Goldblatt, 2012), its area-adjusted diversity is greater than that of most tropical regions (Grobler & Cowling, 2022). As such, it is recognized as a global biodiversity hotspot (Myers et al., 2000). The high vascular plant endemism, 68.3% at the species level (Manning & Goldblatt, 2012), indicates that much of this floristic diversity has originated within the CFR, which has raised much interest in the processes of speciation in the region (e.g. Linder, 2003; Ellis et al., 2014). The CFR is situated in the southwestern region of South Africa (Figure 1), which is characterized by a predominantly Mediterranean-type climate with winter-rainfall and topographical heterogeneity with lowlands interrupted by the rugged Cape Fold Mountains which run north-south in the western region and east-west in the central and eastern regions (Goldblatt, 1978; Mucina & Rutherford, 2006; Manning & Goldblatt, 2012). The high species richness of the region is thought to result from the combination of low extinction and high speciation rates, with low extinction rates being due to relative climatic stability since the late Miocene, since the region experienced no major, ecosystem-resetting disturbance such as glaciations (Goldblatt, 1978, 1997; Dynesius & Jansson, 2000; Latimer, Silander & Cowling, 2005; Verboom et al., 2009). Several suggested drivers of recent radiations have been put forward, including: i) extreme environmental heterogeneity (elevation, climate, soils, fire regime) causing vicariance (non-adaptive divergence) and forming a mosaic of microhabitats providing different niches and consequently facilitating ecological speciation (adaptive divergence) (Cowling, 1987; Cowling, Holmes & Rebelo, 1992; Cowling & Lombard, 2002; van der Niet & Johnson, 2009;

Schnitzler et al., 2011; Verboom et al., 2015); ii) low dispersal rates and distances facilitating divergence (Bond & Slingsby, 1983; Pierce & Cowling, 1991; Johnson, 1992; Cowling et al., 1994; Latimer, Silander & Cowling, 2005); and iii) high levels of pollinator specialization (Johnson, 1996; van der Niet & Johnson, 2009; Anderson et al., 2014; van der Niet, Peakall & Johnson, 2014).

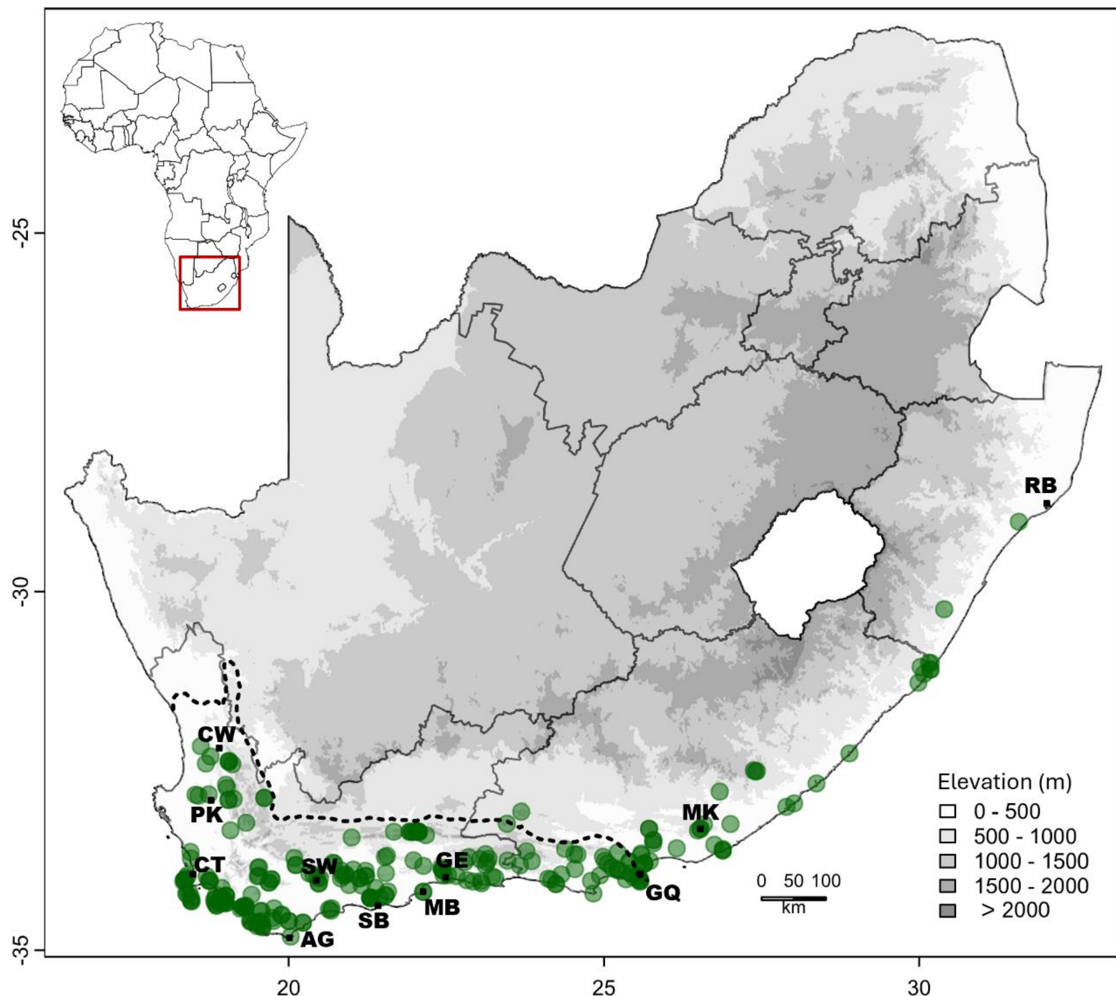


Figure 1. Map of South Africa showing the boundary of the Cape Floristic Region (dotted line) and the distribution of *Osteospermum* section *Polygalinae* based on herbarium records (green symbols). National and provincial borders are represented by solid black lines. Abbreviated town names: CW – Clanwilliam, PK – Piketberg, CT – Cape Town, AG – L'Agulhas, SW – Swellendam, SB – Still Bay, MB – Mossel Bay, GE – George, GQ – Gqeberha, MK – Makhanda, RB – Richards Bay.

Osteospermum L. (Compositae / Asteraceae) is the largest genus in the tribe Calenduleae Cass., currently comprising 91 recognised taxa (Sadler, 2024). While the genus is distributed throughout temperate Africa, most of its diversity is endemic to the Greater Cape Floristic Region (Nordenstam & Källersjö, 2009; Manning & Goldblatt, 2012; Snijman, 2013). Recently the genus *Calendula* L., which has its centre of diversity in the northern-hemisphere Mediterranean region, has been found to

be nested within *Osteospermum* (Nordenstam & Källersjö, 2009), and has consequently been synonymised with *Osteospermum*, within which it constitutes a section (Manning & Goldblatt, 2012). Most species in *Osteospermum* possess yellow flowers which lack a pappus and are aggregated into capitula that (with a single exception) are made up of female-sterile disc florets surrounded by a single row of male-sterile ray florets (Norlindh, 1943; Sadler et al., 2022). Species vary in life history and growth form, ranging from annual and perennial herbs, through low and sprawling perennial shrublets, to tall shrubs. The leaves are finely dissected to entire and in some species even succulent. Fruit (achene / cypsela) structure is also highly variable, suggesting a wide range of dispersal mechanisms, with fruit types including fleshy drupes, dry thin-walled fruits with three small or large wings, and hard-walled nutlet-like fruits with surfaces ranging from smooth to ridged or even spined (Norlindh, 1943; Nordenstam & Källersjö, 2009; Sadler, 2024). There is a large amount of homoplasy in these morphological features which has rendered the taxonomy of the group difficult. Historically, the taxonomy of the tribe has been based largely on fruit morphology, the latter being assumed to be taxonomically informative (Norlindh, 1943). However, the genus level taxonomy has been very unstable, with recent molecular phylogenetic studies showing that fruit types have evolved repeatedly (Sadler, 2024). The Core-*Osteospermum* clade, however, consists primarily of species having simple hard nutlet-like fruits (Norlindh, 1943; Sadler, 2024) and it is from these species that the genus gets its name, i.e., “bone seed”.

Osteospermum section *Polygalinae* DC. forms part of the Core-*Osteospermum* clade and is one of only three sections found to be monophyletic in recent work by Sadler (2024). Species in *O.* sect. *Polygalinae* mostly occur in the Fynbos Biome of the CFR (Mucina & Rutherford, 2006; Manning & Goldblatt, 2012), in which they are found in both fynbos vegetation, associated with nutrient-poor soils, and renosterveld vegetation, associated with more nutrient-rich and finer soils, as well as in grassier vegetation types in the eastern extent of their distribution range (Figure 1). Specimens have been collected in a range of habitats, occurring from near sea level to the peaks of the Cape Fold Mountains (Figure 2). The section is defined morphologically as being shrubs and subshrubs having entire sessile leaves, capitula with uniseriate or subuniseriate involucral bracts, very long yellow ray floret ligules, and homomorphic cylindrical fruits (Norlindh, 1943) (Figure 3). However, although these traits unite the section, the species-level taxonomy of the section has been historically unstable, primarily due to a vast spectrum of growth forms being present (Figure 4), which make species boundaries unclear. In his monograph of the Calenduleae, Norlindh (1943) separated the species in *O.* sect. *Polygalinae* based primarily on surface sculpturing of the fruits (Figure 5), recognising four species (*O. corymbosum* L., *O. imbricatum* L., *O. polygaloides* L. and *O. rotundifolium* (DC.) Norl.) and several subspecies and varieties, but also noting several morphological variants or forms that appeared different but were not separable based on the available evidence. Later he added another variety to *O. imbricatum* (Norlindh, 1960). In the 80 years subsequent to his revision of Calenduleae,

many more specimens have been collected and two new species, *O. australe* B.Nord. and *O. burttianum* B.Nord., have been described (Nordenstam, 2004). However, it remains difficult to assign specimens to species using the current taxonomy, which led Manning and Goldblatt (2012) to suggest that *O. imbricatum* (which comprises four subspecific taxa) be synonymised with *O. polygaloides* in the latest conspectus of Cape plants. Clearly, however, a comprehensive reassessment of species boundaries in *O.* sect. *Polygalinae* is required.

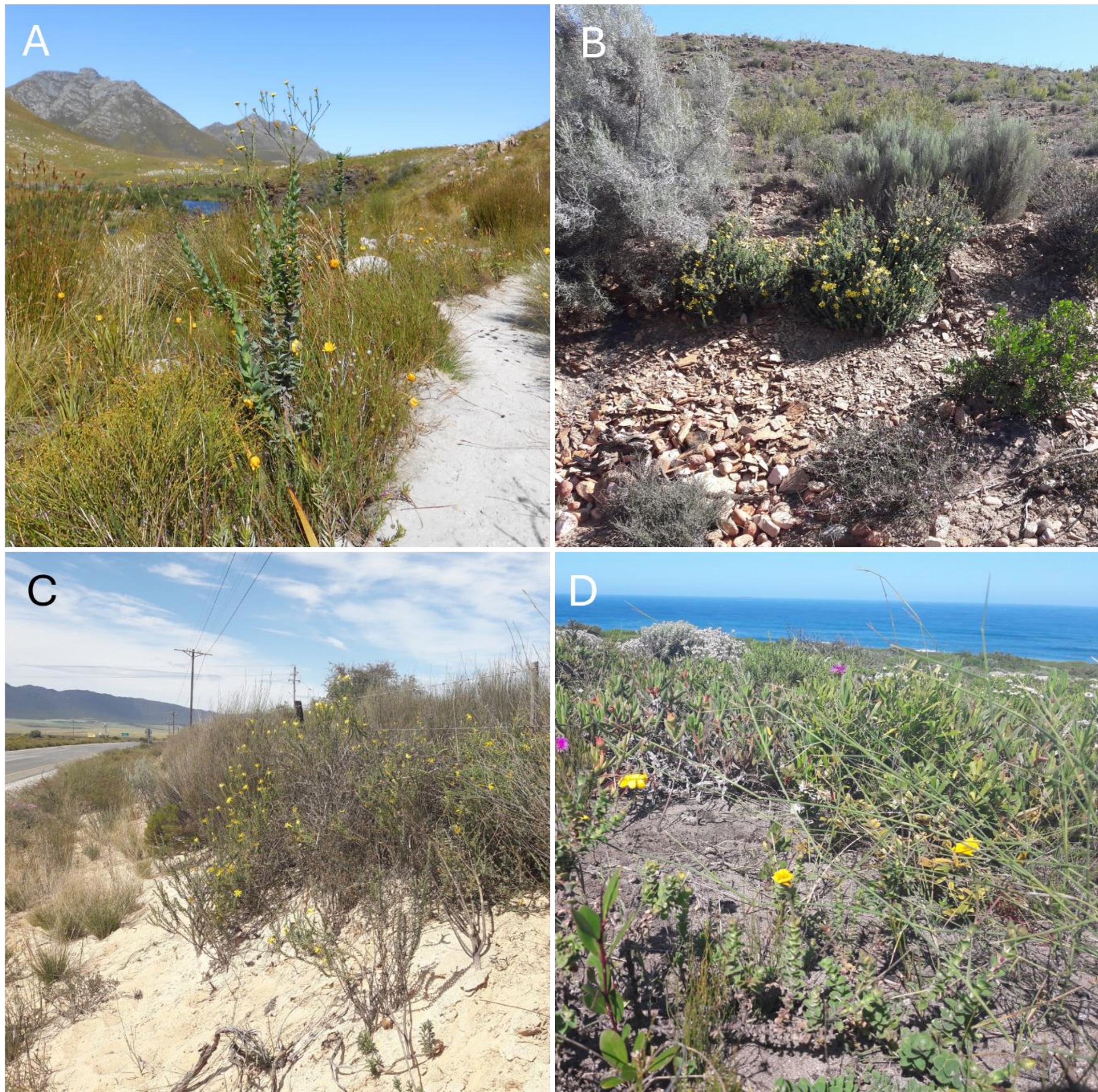


Figure 2. Examples of the habitat variation in which species in *Osteospermum* section *Polygalinae* occur. (A) *O. rotundifolium* in the Palmiet River Valley of the Kogelberg; (B) *O. imbricatum* at the foothills of the Kammanassie Mountains; (C) *O. polygaloides* growing in deep quartzitic sand near Sauer; (D) *O. imbricatum* in the coastal vegetation at Schoenmakerskop, Gqeberha. All photos were taken by L.M.C. James.

Figure 3. The defining traits and examples of trait variation of species in *Osteospermum* section *Polygalinae*. Uniseriate or subuniseriate involucral bracts in *O. imbricatum* (A) and *O. rotundifolium* (C); Long yellow ray floret ligules in *O. imbricatum* (B); Homomorphic cylindrical fruits of *O. imbricatum* (D) and *O. rotundifolium* (E); Entire sessile leaves in *O. imbricatum* subsp. *nervatum* (F), *O. corymbosum* (G) and *O. imbricatum* var. *microcephalum* (H). All photos were taken by L.M.C. James.

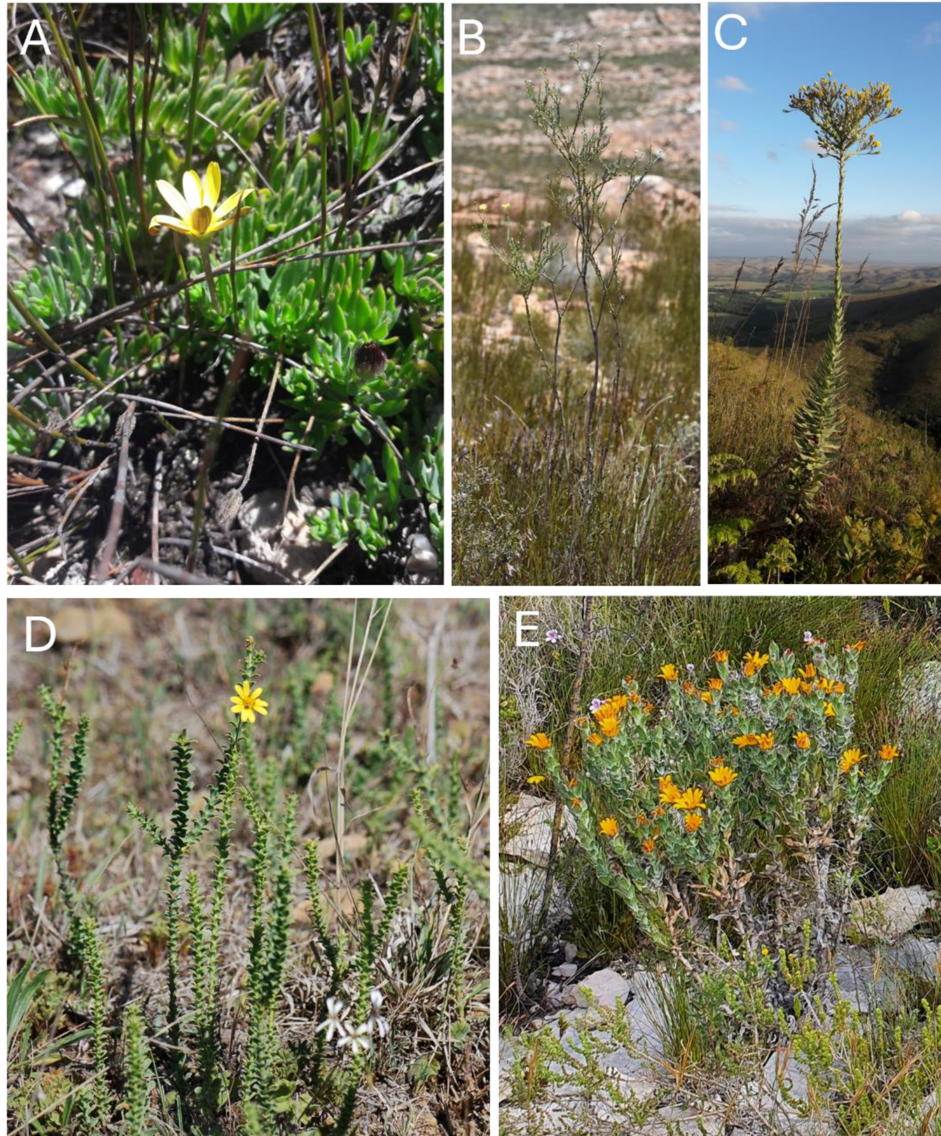


Figure 4. Examples of the variation in growth form of species in *Osteospermum* section *Polygalinae*. (A) A prostrate growth form in *O. polygaloides* at the summit of the Swartberg; (B) An erect, sparsely branching shrub in *O. imbricatum* on the northern slopes of the Swartberg; (C) A single stemmed growth form with branching only at the top of the main stem, forming a capitulescence in *O. corymbosum* in the Langeberg; (D) A small, sparsely branching shrublet growth form in *O. imbricatum* south of Riversdale; (E) A small shrub with erect branching stems in *O. australe* near Pearly Beach. Photos A to D were taken by L.M.C. James and E was taken by R. T. Sadler.

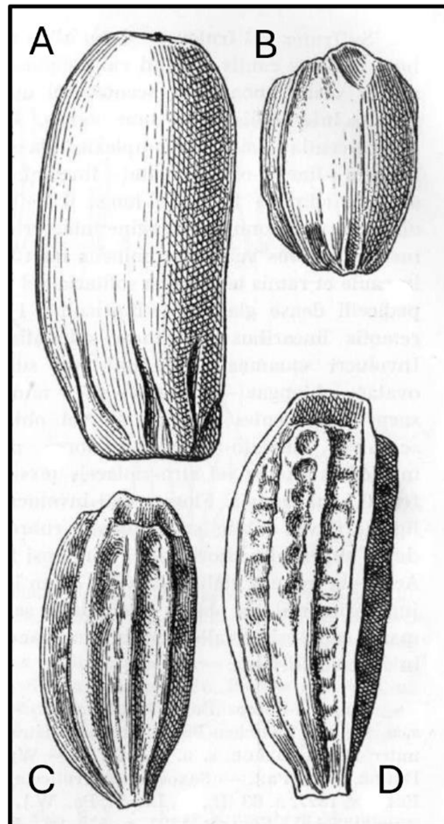


Figure 5. The four fruit forms on which Norlindh (1943) based his taxonomy of *Osteospermum* section *Polygalinae*. (A) Fruit terete, smooth in *O. corymbosum*, (B) Fruit with three slightly prominent ridges in *O. rotundifolium*; (C) Fruit with three obtuse primary ridges, six obtuse secondary ridges and pitted furrows between ridges in *O. polygaloides*; (D) Fruit with three acute primary ridges or often ridges minutely winged, six obtuse secondary ridges and pitted furrows between ridges *in O. imbricatum*. Diagram from Norlindh (1943).

The large variation in fruit morphology and associated dispersal modes in *Osteospermum* provides opportunities for studying the impacts of dispersal mode on diversification. The dispersal of seeds by ants (myrmecochory) has gained interest as a possible driver of speciation due to short dispersal distances limiting gene flow and the possible fitness advantages provided by seed deposition below-ground, in ant nests (Bond & Slingsby, 1983; Lengyel et al., 2009). Myrmecochory occurs in many habitats globally, and ant dispersed diaspores are characterized by having a lipid-rich appendage, called an elaiosome, that acts to reward ants (Sernander, 1906; Berg, 1975; van der Pijl, 1982; Lengyel et al., 2010). Myrmecochory was first reported in the Cape by Marloth (1913) and is thought to be fairly widespread in fynbos species based on the presence of structures inferred to be elaiosomes (Bond & Slingsby, 1983). Experimental tests of myrmecochory in the CFR are, however, few and have mostly focused on Proteaceae (e.g. Bond & Breytenbach, 1985; Slingsby & Bond, 1985; Christian & Stanton, 2004). *Osteospermum* has been reported as one of the genera in the CFR to contain myrmecochrous species (Bond & Slingsby, 1983), with Wood & Nordenstam, (2003) identifying *O. corymbosum* in *O.* sect. *Polygalinae* as being ant dispersed based on the presence of an

elaiosome-like structure on its fruits. It is possible that other species in *O*. sect *Polygalinae* may also possess elaiosome-like structures, but this has not been studied. While studies of myrmecochory at the macroevolutionary level suggest that myrmecochory may enhance diversification (Cowling et al., 1994; Lengyel et al., 2009), the gene flow consequences of myrmecochory remain poorly studied. A focused study on fruit dispersal biology and interpopulation gene flow in species in *O.* sect *Polygalinae* may therefore provide insights into the microevolutionary processes leading to population divergence. However, this is currently hampered by the unclear species boundaries in the section, with the individual species not necessarily representing single, evolutionarily independent units.

The recent development of next-generation sequencing methodologies provides access to large-scale genomic data with resolution appropriate for resolving patterns at both the species and population levels, and which can be used to study species limits and microevolutionary processes. This thesis makes use of these technologies, along with analysis of morphological characters, ecological data and field-based experiments, to achieve two objectives. The first is to reassess species boundaries in *O.* sect. *Polygalinae*, this being the focus of Chapter 2. For this I use molecular phylogenetics to infer species relationships and integrative taxonomy, using population genetic, morphological and ecological data, to delimit species. The second aim of this thesis is to investigate fruit dispersal in *O.* sect. *Polygalinae* and test whether the evolution of myrmecochory enhances population differentiation through its impact on the spatial scale of gene flow. This is the focus of Chapter 3, in which I first characterise the fruits of the newly delimited species in the section, then use field-based experiments to determine the functional significance of the elaiosome-like structure for dispersal in *O. corymbosum* and investigate the evolution of traits associated with myrmecochory. I then test whether the spatial scale of gene flow differs between myrmecochrous and non-myrmecochorous species.

References

Anderson, B., Allsopp, N., Ellis, A.G., Johnson, S.D., Midgley, J.J., Pauw, A. & Rodger, J. 2014. Biotic interactions. In *Fynbos: Ecology, Evolution, and Conservation of a Megadiverse Region*. Oxford: Oxford University Press. 224–247.

Barraclough, T.G. & Nee, S. 2001. Phylogenetics and speciation. *Trends in Ecology & Evolution*. 16(7):391–399. DOI: 10.1016/S0169-5347(01)02161-9.

Berg, R. 1975. Myrmecochorous plants in Australia and their dispersal by ants. *Australian Journal of Botany*. 23(3):475–508. DOI: 10.1071/BT9750475.

Bond, W.J. & Breytenbach, G.J. 1985. Ants, rodents and seed predation in Proteaceae. *South African Journal of Zoology*. 20(3):150–154. DOI: 10.1080/02541858.1985.11447928.

Bond, W.J. & Slingsby, P. 1983. Seed dispersal by ants in shrublands of the Cape Province and its evolutionary implications. *South African Journal of Science*. 79(6):231–233.

Camargo, A. & Sites, J. Jr. 2013. Species delimitation: a decade after the renaissance. In *The Species Problem - Ongoing Issues*. I.Y. Pavlinov, Ed. InTech. 225–247. DOI: 10.5772/52664.

Christian, C.E. & Stanton, M.L. 2004. Cryptic consequences of a dispersal mutualism: seed burial, elaiosome removal, and seed-bank dynamics. *Ecology*. 85(4):1101–1110. DOI: 10.1890/03-0059.

Cicero, C., Mason, N.A., Jiménez, R.A., Wait, D.R., Wang-Claypool, C.Y. & Bowie, R.C.K. 2021. Integrative taxonomy and geographic sampling underlie successful species delimitation. *Ornithology*. 138(2):1–15. DOI: 10.1093/ornithology/ukab009.

Coates, D.J., Byrne, M. & Moritz, C. 2018. Genetic diversity and conservation units: dealing with the species-population continuum in the age of genomics. *Frontiers in Ecology and Evolution*. 6:165. DOI: 10.3389/fevo.2018.00165.

Cowling, R.M. 1987. Fire and its role in coexistence and speciation in Gondwanan shrublands. *South African Journal of Science*. 83:106–112.

Cowling, R.M. & Lombard, A.T. 2002. Heterogeneity, speciation/extinction history and climate: explaining regional plant diversity patterns in the Cape Floristic Region. *Diversity and Distributions*. 8(3):163–179. DOI: 10.1046/j.1472-4642.2002.00143.x.

Cowling, R.M., Holmes, P.M. & Rebelo, A.G. 1992. Plant diversity and endemism. In *The Ecology of Fynbos*. R.M. Cowling, Ed. Cape Town: Oxford University Press. 62–112.

Cowling, R.M., Pierce, S.M., Stock, W.D. & Cocks, M. 1994. Why are there so many myrmecochorous species in the Cape fynbos? In *Plant-Animal Interactions in Mediterranean-type Ecosystems*. M. Arianoutsou & R.H. Groves, Eds. (Tasks for Vegetation Science no. 31). Dordrecht: Kluwer Academic Publishers. 159–168.

Coyne, J.A. & Orr, H.A. 2004. *Speciation*. Sunderland, Massachusetts: Sinauer Associates, Inc.

Dayrat, B. 2005. Towards integrative taxonomy. *Biological Journal of the Linnean Society*. 85(3):407–415. DOI: 10.1111/j.1095-8312.2005.00503.x.

Dieckmann, U. & Doebeli, M. 1999. On the origin of species by sympatric speciation. *Nature*. 400(6742):354–357. DOI: 10.1038/22521.

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*. 97(16):9115–9120. DOI: 10.1073/pnas.97.16.9115.

Ellis, A.G., Verboom, G.A., van der Niet, T., Johnson, S.D. & Linder, H.P. 2014. Speciation and extinction in the Greater Cape Floristic Region. In *Fynbos: Ecology, Evolution, and Conservation of a*

Megadiverse Region. N. Allsopp, J.F. Colville, & G.A. Verboom, Eds. Oxford: Oxford University Press. 119–141.

Freudenstein, J.V., Broe, M.B., Folk, R.A. & Sinn, B.T. 2017. Biodiversity and the species concept - lineages are not enough. *Systematic Biology*. 66(4):644–656. DOI: 10.1093/sysbio/syw098.

Goldblatt, P. 1978. An analysis of the flora of southern Africa: its characteristics, relationships, and origins. *Annals of the Missouri Botanical Garden*. 65(2):369–436.

Goldblatt, P. 1997. Floristic diversity in the Cape Flora of South Africa. *Biodiversity and Conservation*. 6:359–377.

Grobler, B.A. & Cowling, R.M. 2022. Which is the richest of them all? Comparing area-adjusted plant diversities of Mediterranean- and tropical-climate regions. *Frontiers of Biogeography*. 14(3):e56241. DOI: 10.21425/F5FBG56241.

Harvey, M.G., Bravo, G.A., Claramunt, S., Cuervo, A.M., Derryberry, G.E., Battilana, J., Seeholzer, G.F., McKay, J.S., et al. 2020. The evolution of a tropical biodiversity hotspot. *Science*. 370(6522):1343–1348. DOI: 10.1126/science.aaz6970.

Igea, J. & Tanentzap, A.J. 2019. Multiple macroevolutionary routes to becoming a biodiversity hotspot. *Science Advances*. 5:eaau8067. DOI: 10.1126/sciadv.aau8067.

Johnson, S.D. 1992. Plant-animal relationships. In *The Ecology of Fynbos*. R.M. Cowling, Ed. Cape Town: Oxford University Press. 175–205.

Johnson, S.D. 1996. Pollination, adaptation and speciation models in the Cape flora of South Africa. *Taxon*. 45(1):59–66. DOI: 10.2307/1222585.

Latimer, A.M., Silander, J.A. & Cowling, R.M. 2005. Neutral ecological theory reveals isolation and rapid speciation in a biodiversity hot spot. *Science*. 309(5741):1722–1725. DOI: 10.1126/science.1115576.

Lengyel, S., Gove, A.D., Latimer, A.M., Majer, J.D. & Dunn, R.R. 2009. Ants sow the seeds of global diversification in flowering plants. *PLoS ONE*. 4(5):e5480. DOI: 10.1371/journal.pone.0005480.

Lengyel, S., Gove, A.D., Latimer, A.M., Majer, J.D. & Dunn, R.R. 2010. Convergent evolution of seed dispersal by ants, and phylogeny and biogeography in flowering plants: a global survey. *Perspectives in Plant Ecology, Evolution and Systematics*. 12(1):43–55. DOI: 10.1016/j.ppees.2009.08.001.

Lenormand, T. 2002. Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*. 17(4):183–189. DOI: 10.1016/S0169-5347(02)02497-7.

Linder, H.P. 2003. The radiation of the Cape flora, southern Africa. *Biological Reviews of the Cambridge Philosophical Society*. 78(4):597–638. DOI: 10.1017/S1464793103006171.

Linder, H.P. 2008. Plant species radiations: where, when, why? *Philosophical Transactions of the Royal Society B: Biological Sciences*. 363(1506):3097–3105. DOI: 10.1098/rstb.2008.0075.

Manning, J.C. & Goldblatt, P. 2012. *Plants of the Greater Cape Floristic Region: the core Cape flora*. (Strelitzia no. 29). Pretoria: South African National Biodiversity Institute.

Marloth, R. 1913. *The flora of South Africa*. Cape Town: Darter Bros. & Co.

Mayden, R.L. 1997. A hierarchy of species concepts: the denouement in the saga of the species problem. In *Species: The units of diversity*. Chapman & Hall. 381–423.

Mucina, L. & Rutherford, M.C. 2006. *The vegetation of South Africa, Lesotho and Swaziland*. (Strelitzia no. 19). Pretoria: South African National Biodiversity Institute.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B. & Kent, J. 2000. Biodiversity hotspots for conservation priorities. *Nature*. 403(6772):853–858.

van der Niet, T. & Johnson, S.D. 2009. Patterns of plant speciation in the Cape floristic region. *Molecular Phylogenetics and Evolution*. 51:85–93.

van der Niet, T., Peakall, R. & Johnson, S.D. 2014. Pollinator-driven ecological speciation in plants: new evidence and future perspectives. *Annals of Botany*. 113(2):199–212. DOI: 10.1093/aob/mct290.

Nordenstam, B. 2004. Two new species of *Osteospermum* (Compositae–Calenduleae) from southwestern Cape Province, South Africa. *Edinburgh Journal of Botany*. 60(3):259–265. DOI: 10.1017/S0960428603000234.

Nordenstam, B. & Källersjö, M. 2009. Calenduleae. In *Systematics, Evolution and Biogeography of Compositae.* V.A. Funk, A. Susanna, T.F. Stuessy, & R.J. Bayer, Eds. Vienna: International Association for Plant Taxonomy. 527–538.

Norlindh, T. 1943. *Studies in the Calenduleae I: Monograph of the genera Dimorphotheca, Castalis, Osteospermum, Gibbaria and Chrysanthemoides*. Lund: C.W.K. Gleerup.

Norlindh, T. 1960. Additions to the monograph on *Osteospermum*. *Botaniska Notiser*. 4(113):385–399.

Nosil, P. 2008. Speciation with gene flow could be common. *Molecular Ecology*. 17(9):2103–2106. DOI: 10.1111/j.1365-294X.2008.03715.x.

Padial, J.M., Castroviejo-Fisher, S., Köhler, J., Vilà, C., Chaparro, J.C. & De La Riva, I. 2009. Deciphering the products of evolution at the species level: the need for an integrative taxonomy. *Zoologica Scripta*. 38(4):431–447. DOI: 10.1111/j.1463-6409.2008.00381.x.

Padial, J.M., Miralles, A., De La Riva, I. & Vences, M. 2010. The integrative future of taxonomy. *Frontiers in Zoology*. 7(16). DOI: 10.1186/1742-9994-7-16.

Pierce, S.M. & Cowling, R.M. 1991. Dynamics of soil-stored seed banks of six shrubs in fire-prone dune fynbos. *The Journal of Ecology*. 79(3):731–747. DOI: 10.2307/2260664.

van der Pijl, L. 1982. *Principles of dispersal in higher plants*. 3rd ed. Berlin: Springer-Verlag.

de Queiroz, K. 1998. The general lineage concept of species, species criteria, and the process of speciation: a conceptual unification and terminological recommendations. In *Endless Forms: Species and Speciation*. D.J. Howard & S.H. Berlocher, Eds. New York: Oxford University Press. 57–75.

de Queiroz, K. 2005. A unified concept of species and its consequences for the future of taxonomy. *Proceedings of the California Academy of Sciences*. 56(1):196–215.

de Queiroz, K. 2007. Species concepts and species delimitation. *Systematic Biology*. 56(6):879–886. DOI: 10.1080/10635150701701083.

Rolland, J., Condamine, F.L., Jiguet, F. & Morlon, H. 2014. Faster speciation and reduced extinction in the tropics contribute to the mammalian latitudinal diversity gradient. *PLoS Biology*. 12(1):e1001775. DOI: 10.1371/journal.pbio.1001775.

Rouhan, G. & Gaudeul, M. 2014. Plant taxonomy: a historical perspective, current challenges, and perspectives. In *Molecular Plant taxonomy: Methods and Protocols*. P. Besse, Ed. (Methods in Molecular Biology no. 1115). New York: Springer. 1–37.

Rundle, H.D. & Nosil, P. 2005. Ecological speciation. *Ecology Letters*. 8(3):336–352. DOI: 10.1111/j.1461-0248.2004.00715.x.

Sadler, R.T. 2024. Phylogeny, generic limits and evolution of dispersal traits in the Asteraceae tribe Calenduleae. MSc thesis. University of Cape Town.

Sadler, R., Parker, T., Verboom, G.A., Ellis, A.G., Jackson, M., Van Zyl, J., Manning, J. & Bergh, N.G. 2022. A molecular phylogeny of Calenduleae (Asteraceae) supports the transfer of *Dimorphotheca polyptera* DC. to *Osteospermum* L. *South African Journal of Botany*. 151:234–245. DOI: 10.1016/j.sajb.2022.09.001.

Schlick-Steiner, B.C., Steiner, F.M., Seifert, B., Stauffer, C., Christian, E. & Crozier, R.H. 2010. Integrative taxonomy: a multisource approach to exploring biodiversity. *Annual Review of Entomology*. 55(1):421–438. DOI: 10.1146/annurev-ento-112408-085432.

Schnitzler, J., Barraclough, T.G., Boatwright, J.S., Goldblatt, P., Manning, J.C., Powell, M.P., Rebelo, T. & Savolainen, V. 2011. Causes of plant diversification in the Cape biodiversity hotspot of South Africa. *Systematic Biology*. 60(3):343–357. DOI: 10.1093/sysbio/syr006.

Sernander, R. 1906. *Entwurf einer Monographie der europaischen Myrmekochoren*. V. 41. (no. 7). Uppsala: Kungliga Svenska Vetenskapsakademien Handlingar.

Shaffer, H.B. & Thomson, R.C. 2007. Delimiting species in recent radiations. *Systematic Biology*. 56(6):896–906. DOI: 10.1080/10635150701772563.

Slatkin, M. 1985. Gene flow in natural populations. *Annual Review of Ecology and Systematics*. 16:393–430.

Slingsby, P. & Bond, W.J. 1985. The influence of ants on the dispersal distance and seedling recruitment of *Leucospermum conocarpodendron* (L.) Buek (Proteaceae). *South African Journal of Botany*. 51(1):30–34. DOI: 10.1016/S0254-6299(16)31698-2.

Snijman, D.A. Ed. 2013. *Plants of the Greater Cape Floristic Region: the extra Cape flora*. (Strelitzia no. 30). Pretoria: South African National Biodiversity Institute.

Vellend, M. 2010. Conceptual synthesis in community ecology. *The Quarterly Review of Biology*. 85(2):183–206. DOI: 10.1086/652373.

Vences, M., Miralles, A. & Dufresnes, C. 2024. Next-generation species delimitation and taxonomy: Implications for biogeography. *Journal of Biogeography*. 00:1–14. DOI: 10.1111/jbi.14807.

Verboom, G.A., Archibald, J.K., Bakker, F.T., Bellstedt, D.U., Conrad, F., Dreyer, L.L., Forest, F., Galley, C., et al. 2009. Origin and diversification of the Greater Cape flora: Ancient species repository, hot-bed of recent radiation, or both? *Molecular Phylogenetics and Evolution*. 51:44–53.

Verboom, G.A., Bergh, N.G., Haiden, S.A., Hoffmann, V. & Britton, M.N. 2015. Topography as a driver of diversification in the Cape Floristic Region of South Africa. *New Phytologist*. 207(2):368–376. DOI: 10.1111/nph.13342.

Webb, C.O., Ackerly, D.D., McPeek, M.A. & Donoghue, M.J. 2002. Phylogenies and community ecology. *Annual Review of Ecology and Systematics*. 33(1):475–505. DOI: 10.1146/annurev.ecolsys.33.010802.150448.

Wiens, J.J. 2004. What is speciation and how should we study it? *The American Naturalist*. 163(6):914–923. DOI: 10.1086/386552.

Wiens, J.J. 2007. Species delimitation: new approaches for discovering diversity. *Systematic Biology*. 56(6):875–878. DOI: 10.1080/10635150701748506.

Will, K.W., Mishler, B.D. & Wheeler, Q.D. 2005. The perils of DNA barcoding and the need for integrative taxonomy. *Systematic Biology*. 54(5):844–851. DOI: 10.1080/10635150500354878.

Wood, A.R. & Nordenstam, B. 2003. An interesting new species of *Osteospermum* (Asteraceae-Calenduleae) from the Western Cape Province, South Africa, providing a link to the genus *Chrysanthemoides*. *South African Journal of Botany*. 69(4):572–578. DOI: 10.1016/S0254-6299(15)30297-0.

Yeates, D.K., Seago, A., Nelson, L., Cameron, S.L., Joseph, L. & Trueman, J.W.H. 2011. Integrative taxonomy, or iterative taxonomy? *Systematic Entomology*. 36(2):209–217. DOI: 10.1111/j.1365-3113.2010.00558.x.

Yi, H., Dong, S., Yang, L., Wang, J., Kidner, C. & Kang, M. 2023. Genome-wide data reveal cryptic diversity and hybridization in a group of tree ferns. *Molecular Phylogenetics and Evolution*. 184:107801. DOI: 10.1016/j.ympev.2023.107801.

Chapter 2: Species delimitation and phylogenetic relationships in *Osteospermum* section *Polygalinae*

Introduction

As the primary unit of biological classification, the species is fundamental to our perception, management, and conservation of biological diversity. Following a century of heated debate, the question of what a species is has recently reached some consensus (e.g., Shaffer & Thomson, 2007; Wiens, 2007; Padial et al., 2010; Schlick-Steiner et al., 2010; Yeates et al., 2011; Yi et al., 2023; Vences, Miralles & Dufresnes, 2024), with the broad adoption of a concept of species as distinct evolutionary lineages (Simpson, 1961; Wiley, 1978; de Queiroz, 1998). Although most species concepts have a similar ontological view of species (de Queiroz, 2005), the operational criteria used to recognise species vary and remain a source of disagreement (Freudenstein et al., 2017; Burbrink & Ruane, 2021). In addition, the purpose and utility of species has also changed over time: where species were originally defined purely as units of classification, they now routinely function as units of study/analysis in evolutionary research (Wiens, 2004; Rouhan & Gaudeul, 2014) and as units of conservation (Myers et al., 2000; Coates, Byrne & Moritz, 2018). This enhances the importance of defining species as meaningful, evolutionary independent entities.

The ability of taxonomists to delimit species as evolutionarily independent lineages has been limited by the types of data available at the time. Traditionally species discovery and delimitation relied largely on morphological evidence (e.g., de Candolle & de Candolle, 1837; Harvey & Sonder, 1865; Cronquist, 1981). However, widespread availability of DNA sequence data from the early 1990s (e.g., Chase et al. 1993) paired with a recognition that morphological character variation can be homoplasious (e.g., Givnish & Sytsma, 1997), has resulted in an increased reliance on sequence data, which has provided a novel and potentially more reliable basis for identifying independent lineages, and has thus revolutionised systematics (The Angiosperm Phylogeny Group, 2016). While Sanger sequencing of a small number of candidate loci has been widely used to infer phylogenetic relationships, these can be insufficient to resolve relationships between closely related taxa, such as at the population level which is relevant to the question of species delimitation, due to offering insufficient variation or yielding too few independent gene tree histories to allow gene tree variance to be accounted for (Doyle, 1992, 1997; Edwards, 2009). Thus, inference of population relationships based on Sanger sequences may be confounded by incomplete lineage sorting (Blanco-Pastor, Vargas & Pfeil, 2012). Next-generation sequencing overcomes these limitations by providing more resolution due to a larger dataset and by sampling many independently assorting loci from across the genome (Edwards, 2009; Knowles & Carstens, 2007; Parks, Cronn & Liston, 2009; Carstens et al., 2013; McCormack et al., 2013).One such approach involves sampling thousands of single nucleotide

polymorphisms (SNPs) from across the genome, which has been shown to be effective in revealing population relationships (Leaché & Oaks, 2017). This approach can thus allow recognition of independent lineages and provide a framework in which to delimit species based on phylogenetic groupings. However, since metapopulations can exhibit genetic structure, resulting in recovery of sister lineages that are phenotypically identical, other criteria are necessary to rank phylogenetic groups (e.g., phenotypical/ecological criteria) in order to delimit species that are meaningful biological entities (Freudenstein et al., 2017).

One of the main disagreements between taxonomists is on the criteria that indicate the boundaries of a species (e.g., reproductive isolation, ecological distinction, morphological distinction, monophyly, genetic distinction) that each operate at a different level along the speciation continuum (Mayden, 1997, 1999; de Queiroz, 1998, 2005). Integrative taxonomy is an approach that uses multiple different sources of data (e.g., molecular, morphological and ecological) to delimit species (Dayrat, 2005; Will, Mishler & Wheeler, 2005). Recently the availability of fine scale genomic data from next-generation sequencing has allowed the integration of population genetics methods into taxonomy and has provided new quantitative ways in which species boundaries can be tested. Integrative taxonomy including population genetics and thorough population-level sampling across geographic distributions is proving an effective approach for delimiting species, uncovering cryptic diversity and bringing stability to taxonomy (Quattrini et al., 2019; Cicero et al., 2021; Yi et al., 2023). This approach is semi-flexible, with the choice of suitable data to test species hypotheses being directed by the expertise of the taxonomist based on what they deemed appropriate for the particular group of organisms and their ecology (Sites & Marshall, 2004; Schlick-Steiner et al., 2010).

In addition to advances in the methods available for species delimitation, there has also been an enormous increase in the availability of specimens (Daru et al., 2018) and a growing appreciation of the importance of field-based sampling. For example, in the field of botany, taxonomy was often done in the context of a herbarium in the 19th and early 20th century, often relying on very few specimens with little knowledge of the habitats from which they were collected. Field based sampling provides insights into species ecology and allows for the discovery of characters which are not easily represented on herbarium specimens.

Although the methods and resources available for species delimitation have advanced substantially in the last half century, for many groups we still use the taxonomy established prior to this period. Where this taxonomy insufficiently captures true diversity and species relationships, it is critical to revise these groups so that they represent evolutionary meaningful independent units of biodiversity, as they are interpreted to be in the modern use of species. One such group is the genus *Osteospermum* L., a large genus of daisies (Asteraceae) that makes up the bulk of the species-poor tribe Calenduleae.

Apart from the publication of several newly discovered species (Wood & Nordenstam, 2003; Nordenstam, 2004; Manning, Goldblatt & Helme, 2012; Swanepoel, Cauwer & Wyk, 2021), the species-level taxonomy of *Osteospermum* has received little attention over the last 70 years, being last revised in 1943 by the Swedish taxonomist Tycho Norlindh in his monograph "Studies in the Calenduleae", and the genus has therefore been listed as a priority for taxonomic revision (Victor, Hamer & Smith, 2013; le Roux & Victor, 2017).

Osteospermum species are distributed throughout temperate Africa, but most of the diversity is situated in the Greater Cape Floristic Region of South Africa. Species in *Osteospermum* are mainly yellow-flowered annual and perennial herbs and shrubs, which show remarkable variation in fruit structure. Norlindh (1943) recognised ca. 80 species under *Osteospermum*, which he organised into two subgenera and 15 subgeneric sections distinguished largely by characters of the fruits. The monophyly of Norlindh's (1943) sections of *Osteospermum* has only recently been evaluated in the context of a phylogeny in which most currently recognised species are represented (Sadler, 2024). While the majority of the sections were not recovered as monophyletic, Sadler (2024) found support from nuclear ribosomal and plastid DNA sequences for the monophyly of *Osteospermum* section *Polygalinae* DC.

Osteospermum section *Polygalinae* is a fynbos-centred group of perennial shrubs and subshrubs that is morphologically distinct from other sections. However, the species-level taxonomy of the section has been unstable. This is largely due to traits that can be assessed from herbarium specimens (e.g., floral structure and fruit morphology) being fairly uniform, while traits that are difficult to determine from herbarium specimens (e.g., growth form) are very variable, resulting in much undocumented variation. The section is defined by having entire sessile leaves, capitula with uniseriate or subuniseriate involucral bracts, very long yellow ray floret ligules, and homomorphic cylindrical fruits (Norlindh, 1943). Sadler (2024) has also recently found the section to be defined by a number of synapomorphies, the most notable being large multiseriate trichomes on peduncles and involucral bracts, compared to the species in the rest of *Osteospermum* which have biseriate trichomes. Species in *O.* sect. *Polygalinae* vary greatly in height, growth form, and leaf shape. While the basal leaves of many species are much longer, and sometimes differ substantially in shape, from the medial or upper cauline leaves, they are frequently missing from herbarium specimens (Norlindh, 1943 and personal obs.). Growth form can be affected by plant age and by individual growth history, e.g., whether the individual has been damaged (Norlindh, 1943 and personal obs.), while leaf and fruit traits vary depending on maturity, which cannot always be detected in dried material. This has caused confusion and disagreement between taxonomists (e.g. de Candolle & de Candolle, 1837; Harvey & Sonder, 1865; Norlindh, 1943) and the erection and synonymization of a number of taxa over time. These

problems highlight the importance of new sources of data and of field-based sampling for assessing species circumscriptions.

Norlindh (1943) recognised four species in *O.* sect. *Polygalinae*: *O. rotundifolium* (DC.) Norl., *O. corymbosum* L., *O. polygaloides* L. (with two varieties), and *O. imbricatum* L. (with two subspecies, one of which contains two varieties). He separated these species primarily using the presence, prominence and number of ridges and pits on the fruits. Additionally, leaf shape separated *O. rotundifolium* (DC.) Norl. and *O. corymbosum* from other species, and *O. polygaloides* was separated from *O. imbricatum* based on the density of involucral bract trichomes. Leaf traits were used primarily to separate subspecies and varieties. Later, Norlindh (1960) described another variety of *O. imbricatum* L., namely *O. imbricatum* var. *microcephalum* Norl. which was primarily distinguished on the basis of its small capitula and thin flowering branches. Norlindh (1943) noted a number of specimens with unusual morphology, particularly in *O. imbricatum,* and lacking sufficient evidence to recognise these as distinct taxa, recognized these as forms (e.g., *O. imbricatum* mountain form and *O. imbricatum* coastal form). This was mainly due to him working with few specimens and having had limited field time, resulting in him relying heavily on traits whose variability could be assessed from herbarium collections, particularly fruit morphology. Subsequently many more specimens have been collected, resulting in two further species being described in *O*. sect. *Polygalinae*, *O. australe* B.Nord. and *O. burttianum* B.Nord. (Nordenstam, 2004) Both these species have restricted distributions and are represented by few collections in herbaria. In the latest conspectus of Cape plants, Manning & Goldblatt (2012) included *O. imbricatum* within *O. polygaloides* due to ambiguity of diagnostic traits in the taxonomic literature and existing keys making specimens of these species difficult to separate. However, this was not formalised and *O*. sect. *Polygalinae* therefore currently comprises ten taxa (Table 1).

Table 1. Taxa currently recognised in *Osteospermum* section *Polygalinae*

	Species		Subspecies		Varieties
1.	*O. australe* B.Nord.				
2.	*O. burttianum* B.Nord.				
3.	*O. corymbosum* L.				
4.	*O. imbricatum* L.	a.	subsp. *imbricatum*	I.	var. *imbricatum*
				II.	var. *microcephalum* Norl.
		b.	subsp. *nervatum* (DC.) Norl.	I.	var. *nervatum*
				II.	var. *helichrysoides* (DC.) Norl.
5.	*O. polygaloides* L.			I.	var. *polygaloides*
				II.	var. *latifolium* Norl.
6.	*O. rotundifolium* (DC.) Norl.				

It is evident that *O.* sect. *Polygalinae* is in dire need of a comprehensive reassessment of species boundaries to bring taxonomic stability and delimit biologically and evolutionary meaningful species. Recent studies have demonstrated the success of combining genomic population-level data with morphological, ecological and geographic distribution data in delimiting taxonomically difficult groups in the Cape flora (e.g., Shaik et al., 2023, 2024; Wootton, Forest & Verboom, 2023). In this chapter I follow a similar approach, using integrative taxonomy to assess trait variation and relationships at the level of populations, sampled from across the full morphological and geographic distribution range of *O*. sect. *Polygalinae* and representing geographic regions and putative taxa with multiple samples as far as possible. This comprehensive approach allows for a fine-scale assessment of species identity, independent of historical taxon concepts of the group. I follow the species concept of species being independent evolutionary lineages (de Queiroz, 1998, 2005). To that end, I first use Sanger-sequenced nuclear ribosomal DNA sequences to confirm monophyly of *O.* sect. *Polygalinae*, including newly collected and morphologically variable specimens for which I generate novel DNA sequence data. I then investigate population relationships, and assess species delimitation hypotheses using phylogenetic, population genetic, and multivariate analyses on a large dataset of SNP genotypes generated from genotyping-by-sequencing (Elshire et al., 2011), a set of morphological traits, and ecological data (vegetation type). Species are recognised as independent genetic lineages (de Queiroz, 1998, 2005) that are ecologically distinct from their sister species, as indicated by distinct morphology or occurrence in a different habitat (Freudenstein et al., 2017).

Methods

Population sampling

To guide field sampling, ca. 700 specimens from the South African herbaria NBG, SAM, BOL, PRE, and GRA (Theirs, 2024) were examined to obtain an initial overview of morphological variation, flowering phenology and geographic patterns in *O*. sect. *Polygalinae*. Specimens were first sorted into preliminary morphological forms, independent of current taxonomic placement. Traits used in this sorting included leaf size, shape and arrangement; peduncle branching and length; number of capitula; and growth form as inferred from degree of branching, stem thickness and notes about plant height. Selected specimens (n = 252) representing all of the preliminary forms were then georeferenced, following the National Science Collections Facility of South Africa georeferencing protocol V1.1 (Engelbrecht, 2021), to assess the geographic distribution of this morphological variability and determine the distribution range of *O*. sect. *Polygalinae*. Observation records on iNaturalist were used to identify additional sampling sites when there were few herbarium specimens of a particular morphological form, or when there were noticeable spatial or temporal sampling gaps. Where habitat appeared to be continuous but there were gaps in distribution records, fieldwork also targeted these areas to look for unrecorded populations.

Most field work was conducted between February 2022 and July 2023, but a few specimens were collected in the preceding year. The genomic sequencing methods used in this study required DNA of high quality (see below), making it necessary to collect fresh samples that were rapidly dehydrated rather than use herbarium material. Field work aimed to collect samples from as many populations as possible while capturing the variation in morphology and distribution, attempting to collect samples from at least three populations per morphological form. A total of 80 populations were sampled from across the known distribution of *O.* sect. *Polygalinae* (Figure 1; Appendix 1), with a few additional specimens of *Osteospermum* species from outside *O.* sect. *Polygalinae* sampled as outgroups (Appendix 1).

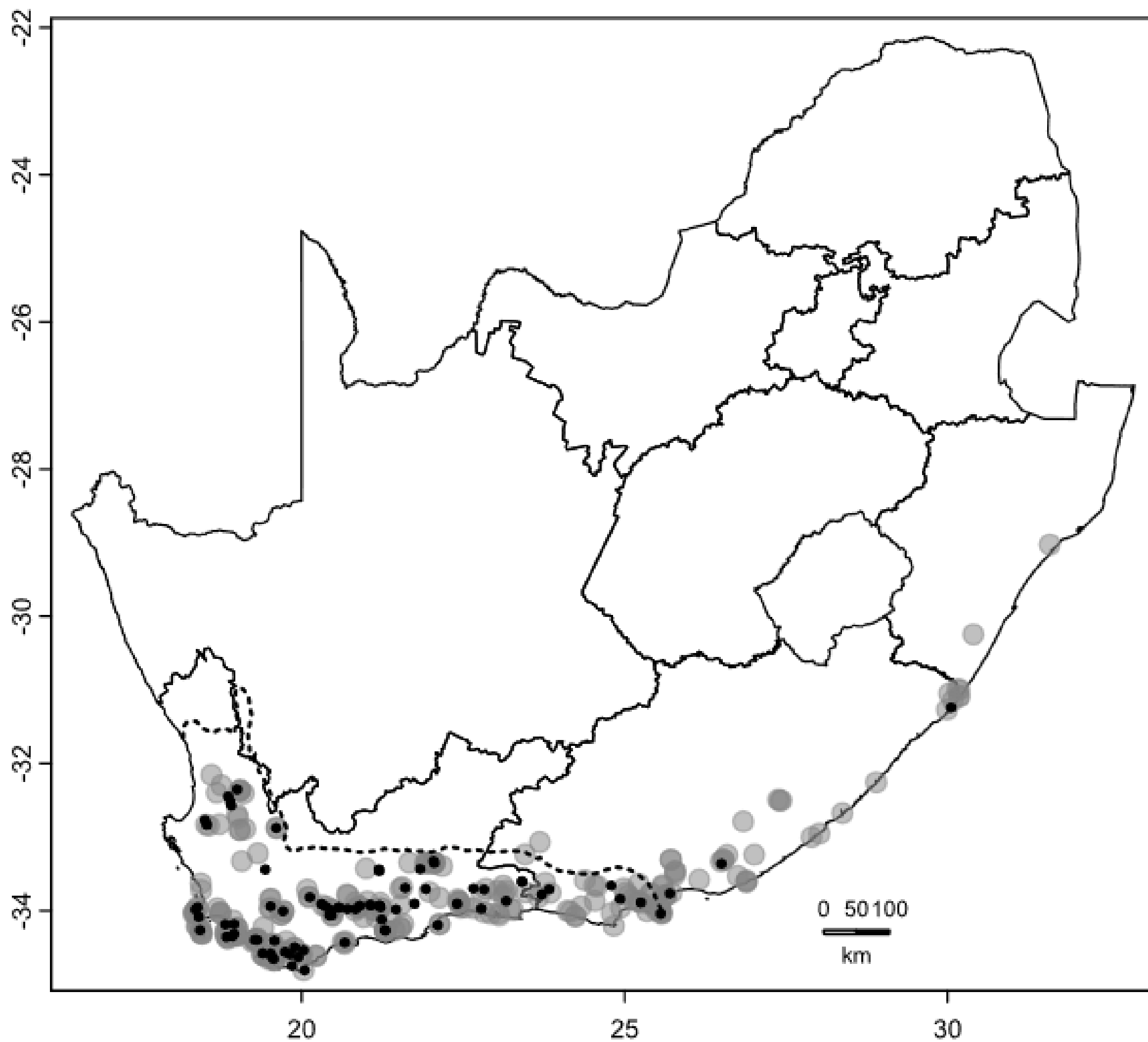


Figure 1. Map of South Africa showing the distribution of species in *Osteospermum* section *Polygalinae* and the localities of the populations sampled in this study. Georeferences of herbarium specimen records (n = 252) from NBG, SAM, BOL, PRE, and GRA are represented by grey symbols and populations sampled in this study are represented by black symbols. National and provincial borders are represented by solid black lines. The boundary of the CFR is indicated by a dotted line (after Goldblatt and Manning 2000).

In most cases the geographical boundaries of populations were easy to determine because individuals were found in distinct stands, often with no other individuals found within several hundred to thousand meters around them. To promote genetic independence of sampled populations, I attempted to maintain a distance of at least 5 km between sampled populations of morphologically similar individuals, though a few sampled populations were ± 3 km apart. With the exception of three populations, represented by NGB2603, SM623 and LMCJ204, I personally sampled all populations, allowing comparison of habitats, within-population phenotypic variation and different life history stages. Field work led to the discovery of two single-population forms (represented by LMCJ180 and LMCJ298 that are morphologically distinct from all previous collections seen in herbaria.

Three individuals were sampled where possible at each population, with additional individuals being sampled where populations exhibited large morphological or developmental variability. Samples were collected both for morphological assessment and DNA extraction for molecular analyses. Samples for molecular work were collected for two purposes: (i) to test for monophyly of *O.* sect. *Polygalinae* using data from Sanger sequencing, and (ii) to explore the relationships between populations and determine species boundaries using a large genomic dataset produced from genotyping-by-sequencing (GBS). For (ii), one individual was used from every population, while for (i) a subset of samples was selected to represent the range of morphological and geographical variability. Previous studies examining species delimitation in the Cape flora using GBS data have found little within-population genetic diversity (e.g., Parker, 2019; Shaik et al., 2023; Wootton, Forest & Verboom, 2023), justifying the use of one sample per population for the GBS dataset, and thus making it feasible to sample a greater number of populations.

Samples of leaf material were stored in silica-gel for quick dehydration in preparation for DNA extraction. Samples of capitula, fruits and leaves were fixed in FAA (60% ethanol, 25% distilled water, 10% formalin, 5% glacial acetic acid) for later morphological examination. Additionally, fruits were collected into brown paper packets. Pressed voucher specimens will be deposited in the Compton Herbarium (NBG). All specimens were identified as fully as possible based on the current taxonomy (Norlindh, 1943; Nordenstam, 2004) and compared to images of the type specimens so that the current taxon concepts could be assessed against the species delimitation criteria applied here. Each population was given a unique code consisting of a prefix from the first three letters of the taxon name followed by a number corresponding to the latitude of its collection locality relative to other populations given the same prefix, for example the western-most population identified as *O. corymbosum* is given the population code "COR1".

DNA extraction

To obtain sufficient DNA of a quality appropriate for GBS (requirements: purity OD 260/280 = 1.8-2.0; concentration ≥ 20 ng/µL; amount ≥ 0.6 µg), DNA extraction protocols were adapted for samples at different stages of leaf development and/or having secondary metabolite concentrations. Between 30 and 100 mg of silica-dried leaf material per sample was ground into fine powder in a Retsch MM400 mixer mill (Retsch, Haan, Germany) using stainless steel grinding beads. Most samples were then extracted using a modified hexadecyltrimethylammonium bromide (CTAB) extraction protocol based on that of Doyle & Doyle (1987), following specifications given by Inglis et al. (2018) for sample lysis and extraction using a high salt CTAB lysis buffer. Further modifications included using 0.4% β-mercaptoethanol and adding RNase A (2 µl of 100 mg/ml RNase or 10 µl of 10 mg/ml RNase) to the lysis buffer. Samples were placed in a -20°C freezer for 45 to 150 minutes or overnight to allow

maximum precipitation of DNA. DNA pellets were suspended in 60 - 75 µl of TE buffer and stored at 4°C.

For some samples sodium dodecyl sulphate (SDS) was used as the surfactant for cell lysis and polysaccharide removal instead of CTAB. In such cases 800 µl of a lysis buffer (500 mM NaCl, 100 mM Tris-HCl, 50 mM EDTA, adjusted to pH 8.0, and 1% 2-mercaptoethanol) was added to each sample after grinding, followed by 80 µl of 20% SDS. Following the incubation step, 80 µl of 4°C 5M potassium acetate was added to each sample and samples were cooled for 10 minutes in a -20°C freezer to precipitate out SDS, proteins and polysaccharides, before centrifuging and pipetting the supernatant into new tubes. Before DNA precipitation with isopropanol, 150 µl 5 M NaCl was added in addition to sodium acetate.

To decrease secondary metabolite concentrations, about half the samples were pre-washed after grinding with a sorbitol buffer prepared following Inglis et al. (2018) but with EDTA concentration increased to 10 mM, before continuing with the standard extraction protocol. Alternatively for some samples, 3.75 mg of activated charcoal per sample was added to the CTAB extraction buffer as suggested by Križman et al. (2006). This however had limited success.

DNA extract quality was assessed for purity using a spectrophotometer (NanoDrop ND-1000; Thermo Fisher Scientific, U.S.A.), for fragmentation and qualitative concentration using 1% agarose gel electrophoresis of DNA and visualisation with ethidium bromide under UV light, and for quantitative concentration using a fluorometer and fluorescent dye (Qubit™ dsDNA BR Assay Kit; Thermo Fisher Scientific, U.S.A.). Where necessary, samples were purified using magnetic beads (AMPure XP reagent; Beckman Coulter, Inc., Brea, CA, U.S.A.).

Species-level phylogenetic analysis to test for monophyly of Osteospermum sect. Polygalinae

To investigate the monophyly and membership of *O*. sect. *Polygalinae,* I selected a subset of samples (39 populations) representing species and putative species whose morphology identified them as potentially belonging to *O*. sect. *Polygalinae* (Appendix 2), but many of which had not been included in the phylogeny of Sadler (2024). The phylogenetic placement of these samples was assessed with outgroup *Osteospermum* species (n = 18) representing all major clades and most subclades of the *Osteospermum* phylogeny inferred by Sadler (2024) (Appendix 2). Nuclear ribosomal DNA sequences of the internal and external transcribed spacer regions (ITS and ETS) were obtained from R. Sadler (n = 34) or were newly generated (n = 23). ITS and ETS regions were chosen because they were found to be more informative than plastid spacer regions in the study by Sadler (2024). Trees were rooted on *O. ilicifolium* L., found to be sister to all remaining species of *Osteospermum* (Sadler, 2024).

DNA extracts were amplified by polymerase chain rection (PCR) using the ITS4 and ITS5 primers of White et al. (1990) for the ITS region, and AST-1 (Markos & Baldwin, 2001) and ETS-18S (Baldwin & Markos, 1998) primers for the ETS region. PCR solutions contained either (i) Supertherm taq (0.25 µl), with H_2O (15.25 µl), 10x buffer (2.5 µl), 25mM $MgCl_2$ (2.5 µl), 10 mM dNTP mix (1 µl) and 100% DMSO (0.5 µl), or (ii) 5 U/µl KAPA taq (0.1 µl) with H_2O (17.4 µl), 10x KAPA taq buffer A (2.5 µl), 25 mM $MgCl_2$ (1 µl), 10 mM dNTP mix (0.5 µl) and 100% DMSO (0.5 µl). One microlitre of each of the respective forward and reverse primers at 10 µM and 1 µl of DNA extract were added to each tube, to give a total reaction volume of 25 µl. PCR reactions were performed in an Applied Biosystems 2720 Thermal Cycler. The thermal profile for ITS was as follows: an initial 94°C denaturation step for 2 min; followed by 30 cycles of 94°C for 1 min, 50°C for 1 min and 72°C for 1 min; and ending with 72°C for 9 min, and then held at 4°C. For ETS the thermoprofile comprised: an initial 94°C denaturation step for 2 min; followed by 30 cycles of 94°C for 1 min, an annealing temperature between 54.2 and 62°C for 1 min, and 72°C for 1 min; and ending with 72°C for 7 or 9 min, and then held at 4°C. The quality and concentration of amplified DNA was checked using gel electrophoresis as described previously. Post-PCR clean-up and Sanger sequencing in both directions using the PCR primers was performed by the Central Analytical Facility of the University of Stellenbosch. Sequence chromatograms were checked, edited, and assembled in ChromasPro version 2.1.10 (Technelysium Pty. Ltd., 2003-2021). All sequences were compiled in a matrix and aligned automatically using the MUSCLE algorithm (Edgar, 2004) through the EMBL-EBI online tools website (EMBL-EBI, 2023) and alignments were manually checked and trimmed in BioEdit version 7.2.5 (Hall, 1999).

To assess gene tree congruence, parsimony bootstrap analysis of separate ETS and ITS alignments was performed using the accessions for which both ETS and ITS regions had successfully been sequenced (n = 53). Parsimony bootstrap analyses were performed in PAUP* version 4.0a (Swofford, 2002) using the NNI swapping algorithm, with 1000 bootstrap replicates, saving a maximum of 500 trees per replicate, using 10 random stepwise addition replicates with one tree held at each step, and excluding parsimony-uninformative characters. Bootstrap majority-rule consensus trees were compared visually for conflicting nodes with high support. Two accessions (LMCJ95 and LMCJ217) showed strongly supported incongruence (reciprocal bootstrap support, BS, >70 %) between ETS and ITS, and LMCJ190 showed weak incongruence (ETS BS = 98 %; ITS BS = 54 %) (Appendix 3). Taxa involved in incongruence were incorporated into concatenated analysis of ETS and ITS using a taxon duplication/deconstruction approach (Pirie et al., 2008, 2009) in which their ETS and ITS sequences were entered separately into the concatenated matrix, with the sites representing the other gene region coded as missing. These taxa were thus represented twice in the analysis in order to visualise both their ITS and ETS placements simultaneously and to prevent the conflict from affecting

tree topology. For concatenated analysis, four additional accessions for which only ITS sequences were available were included, increasing the number of included taxa to 57.

The best fitting nucleotide substitution models for ETS and ITS alignments were TN+F+R2 and TNe+G4 respectively, as identified using the BIC in ModelFinder (Kalyaanamoorthy et al., 2017) implemented in IQ-TREE version 1.6.12 (Tamura & Nei, 1993; Nguyen et al., 2015; Chernomor, Von Haeseler & Minh, 2016). Bayesian inference of the concatenated alignment was performed using the BEAST v2.7.5 software platform (Bouckaert et al., 2019). Due to limited model options for BEAST analyses, the TNe+G4 model from the SSM package (Bouckaert & Dong, 2017) was selected as the best available model for both ETS and ITS. Parameters were set up in BEAUti with the site model partitioned so that parameters of the TNe+G4 model were estimated separately for the ETS and ITS sequences. An optimised relaxed clock (Drummond et al., 2006; Zhang & Drummond, 2020; Douglas, Zhang & Bouckaert, 2021) and a birth-death tree prior (Gernhard, 2008) were selected, with remaining priors set to their default. Four independent MCMC chains were run, each starting from a different randomly generated starting tree and running for 10^7 generations, sampled every 1000th generation. Convergence and estimated sample sizes (ESS) of parameters were assessed in Tracer v1.7.2 (Rambaut et al., 2018), with a burn-in of 10% being determined as appropriate. The three best runs (all with ESS values >200 for all parameters) were combined using Logcombiner, yielding a total of 27,003 trees. A 90% majority rule consensus true was computed from all trees in PAUP*, and visualised in Figtree version 1.4.4 (Rambaut, 2018).

GBS sequencing and phylogenetic analysis to infer population relationships

DNA extracts from one individual per *O*. sect. *Polygalinae* population (Appendix 1) and an outgroup (*O. junceum* P.J.Bergius, LMCJ271) (n = 80) were sent to Novogene Co., Ltd., Beijing, China for GBS library construction and sequencing. For library construction, DNA was digested using restriction enzymes MseI and NlaIII, and barcode adapters were ligated to the ends of the resulting fragments. DNA fragments were amplified by PCR and then size-selected and purified. Sequencing was performed using an Illumina NovaSeq PE150 High-throughput Sequencing platform producing paired-end reads 144 bp in length.

Genotyping-by-sequencing data were processed using the computation facilities provided by the University of Cape Town's ICTS High Performance Computing team (https://ucthpc.uct.ac.za/). Data were first pre-processed using fastp version 0.23.4 (Chen, 2023). Quality filtering was performed in two steps. Firstly, reads containing greater than 40% of sites with a Phred score less than 20 were discarded. Secondly, a sliding window of four bases width (default) was used to cut all data in and to the right of the window when mean Phred quality in the window was less than 20. Reads were discarded if their length was less than 75 bp or complexity less than 20%. Adapters were

automatically detected and trimmed. In overlapped regions of paired reads, correction was performed to improve quality of bases when a base of high quality was present at a site to which bases of low quality could be compared and corrected. PolyG tails were trimmed and overrepresented sequences were detected, both using default settings. After filtering this yielded a mean of 3.7 million reads per sample (min: 2.0 million, max: 8.8 million).

For read mapping, a reference genome of *Calendula officinalis* L. was used (GCA_029618835) (Porghahreman et al., 2023). The genus *Calendula* is nested within *Osteospermum* (Nordenstam & Källersjö, 2009; Manning & Goldblatt, 2012; Sadler et al., 2022; Sadler, 2024) and so is appropriately closely-related. Reads were mapped to the reference genome using BBMap version 39.06 (Bushnell, 2023) with k = 11. Genotype calling and filtering was performed using BCFTools version 1.19-1-g0172073. (Danecek et al., 2021). Maximum depth was set to 5000. SNPs were set as missing if their depth was less than four or if they passed a test for excess heterozygosity at $p > 0.05$. Sites with 30% or more missing data across all samples were excluded. Different filtering parameters were used depending on the downstream data analysis. For population-level phylogenetic analysis, invariant sites and biallelic SNPs were retained. For other analyses only biallelic SNPs were retained.

The VCF file was converted to PHYLIP format using vcf2phylip version 2.9 (Ortiz, 2019). Model selection and maximum likelihood tree inference were performed using default settings in IQ-TREE 2.2.2.7 (Minh et al., 2020). Model selection was performed on the concatenated alignment file as described previously, with TPM3u+F+I+R4 selected as the best-fit model. Branch support was assessed using the ultrafast bootstrap (Hoang et al., 2018) and the SH-aLRT test (Guindon et al., 2010), each with 1000 replicates, as well as site concordance factors (sCF) calculated using the --scfl option to assess the proportion of informative sites supporting each branch (Mo et al., 2023). A branch was considered to be well supported if ultrafast bootstrap (UFB) ≥ 95%, SH-aLRT ≥ 80% and sCF > 33%. *Osteospermum junceum* was used as an outgroup to root the phylogeny. Phylogenetic tree inference was performed using the dataset filtered for both 30% and 20% missing data per site and the tree topologies and branch supports were then compared.

Population genetic analyses

For the purpose of identifying genetic clusters, Principal component analysis (PCA) and STRUCTURE-like methods were applied to the SNP dataset using R version 4.3.2 (R Core Team, 2023). STRUCTURE-like methods use population genetic models to infer population structure by calculating the proportions of genomic data per sample that originate from different inferred ancestral gene pools (Pritchard, Stephens & Donnelly, 2000; Frichot et al., 2014). Principal component analysis is a multivariate analysis based on Euclidean distances and is therefore not as sophisticated as STRUCTURE-like methods, not being based on genetic models, but is mathematically appropriate for

analysing biallelic SNP data when assessing intraspecific genetic variation and variation between closely related species at a population level (Patterson, Price & Reich, 2006; Georges et al., 2023). For biallelic data in geno format, 0 and 2 represent the homozygote states for the two alleles, while 1 represents the heterozygous state and is therefore a meaningful intermediate. The Euclidean distance between two individuals is calculated using the Pythagorean theorem, calculating the square root of the sum of the squares of the distances between individuals at each locus (Georges et al., 2023). Both STRUCTURE-like methods and PCA are commonly used for inferring population level genetic patterns. However each has different assumptions and biases, and it is therefore recommended to use both methods and assess their agreement when interpreting genetic groupings to determine whether patterns are robust to the different assumptions of the two approaches (Patterson, Price & Reich, 2006; Frichot et al., 2014). Some key differences between these methods include their assumptions, how population structure is determined and how admixture is represented. Firstly, while STRUCTURE-like methods are more sophisticated than PCA, being based on population genetic models, they can be vulnerable to bias if their model assumptions are violated (Patterson, Price & Reich, 2006; Frichot et al., 2014). Secondly, for PCA the number of significant eigenvectors needed to represent the genetic structure in the population can be determined using Tracy-Widom statistics without assigning individuals to groups *a priori* (Patterson, Price & Reich, 2006). This is an advantage over STRUCTURE-like analyses which require the number of ancestral gene pools (*K* value) to be specified before assignment. Patterson, Price & Reich (2006) and Frichot & François (2015) recommend comparing a range of *K* values for STRUCTURE-like analyses, which can be guided by PCA results, using *K* values up to one more than the number of significant PCA eigenvalues. Lastly, admixture is modelled in STRUCTURE-like methods by inferring portions of the ancestral gene pool from specific genetic groups (Pritchard, Stephens & Donnelly, 2000), while admixture in PCAs is represented by intermediate positions between parent clusters, which makes admixture harder to detect and can make relationships between clusters difficult to interpret in PCAs (Patterson, Price & Reich, 2006; Frichot et al., 2014; Yi & Latch, 2022). In addition to these differences, in both STRUCTURE-like methods and PCA, the identification of genetic groups can be impacted when there are unequal sample numbers between groups, with there being a bias against the detection of groups with small sample numbers as distinct (Patterson, Price & Reich, 2006; McVean, 2009; Puechmaille, 2016). This is a serious drawback when being used for species delimitation when some species are only known from a small number of populations (often just one) and the use of multiple methods to allow cross-validation is therefore recommended.

For both methods, VCF files were first converted to genlight format using the function “vcf2genlight” from the package vcfR version 1.15.0 (Knaus & Grünwald, 2017), and then to geno format using the function “gl2geno” from the package dartR version 2.9.7 (Mijangos et al., 2022). PCA was performed using the “pca” function from LEA package version 3.14.0 (Frichot & François, 2015). PCA of SNP

data can be severely biased by the presence of even a small percentage of missing data (Yi & Latch, 2022). The SNP data were therefore filtered for a range of data missingness values: < 30% (22701 SNPs), < 5% (5337 SNPs), < 3% (3929 SNPs) and < 1% (1186 SNPs) missing data and compared. This approach was preferred over the alternative commonly used approach of data imputation, which can introduce bias due to averaging (Yi & Latch, 2022; Georges et al., 2023). All PCAs showed the same general pattern so the PCA based on the dataset with < 5% missing data is presented as it represents the best trade-off between bias due to missing data and having sufficient data to allow good resolution.

Since both PCA and phylogenetic analysis resolved four well-supported deeply diverged major clades or well-distinguished clusters within *O.* sect. *Polygalinae* (see results), further population genetic analyses, which aimed to resolve distinct species-level entities, were conducted separately for each of these lineages. This allowed increased population-level resolution and avoided model assumption violations. For each major clade, PCAs were performed as above, but samples in each major clade were first filtered separately to remove all sites with missing data. This reduced the number of SNPs to 3828 in Clade 1, 7917 in Clade 2, 11642 in Clade 3, and 5471 in Clade 4. PCAs were followed by Tracy-Widom tests to determine the significant number of eigenvalues using the “tracy.widom” function from the LEA package.

Gene pool ancestry was estimated using sNMF which uses a nonparametric algorithm based on sparse nonnegative matrix factorization that yields results comparable to those generated by STRUCTURE (Pritchard, Stephens & Donnelly, 2000) and ADMIXTURE (Alexander, Novembre & Lange, 2009), but under much faster computation run times and free from the assumption of Hardy–Weinberg equilibrium (Frichot et al., 2014). The “snmf” function from the LEA package was run on the data for each clade with *K* ranging from one to ten, with 10% of genotypes masked for cross-entropy calculation, and alpha ranging from 51 to 150, totalling 100 repetitions per *K* value. The “snmf” function estimates an entropy criterion, the minimum value of which can be used to assist *K* value selection. However, the entropy criterion was uninformative for this dataset, so a *K* value range from two to the number of significant eigenvalues of the PCA plus two was selected for each of the four clades. The sNMF results were then further analysed using the “clumpak” function from package starmie version 0.1.3 (Tonkin-Hill & Lee, 2024), using the “stephens” method (Stephens, 2000), using a parallelized version of the function created by Seth Musker (https://github.com/SethMusker/starmie) to speed up run time. CLUMPK is an algorithm used for combining Q-matrices (admixture coefficients) per *K* produced from different runs of STRUCTURE-like analyses to assess cluster assignment between runs while avoiding switching of labels assigned to individuals per run (Kopelman et al., 2015). The percentage assignment to major and minor clusters was assessed and structure plots were used to represent the major clusters identified by the clumpak algorithm for the range of *K* values. I tested the impact of missing data on gene pool assignment by sNMF, using SNP

data filtered for < 30% and 0% missing data per site. These produced qualitatively similar results, so I present the plots produced using the < 30% missing data which contained 22701 sites.

Morphological data collection and analysis

Morphological data were collected for every population for eight continuous traits and ratios and six qualitative binary traits (Table 2). Traits were measured on pressed voucher specimens except plant height, which was measured or estimated in the field, and fruit traits which were scored from fruit preserved in FAA or collected directly and dried in brown paper packets to prevent crushing and deformation associated with pressing. Measurements were taken using a tape measure, a ruler, or digital vernier calipers. Leaf traits and peduncle length were measured on one individual per population with the mean being determined, as far as possible, from three replicate measures (some specimens had fewer than three peduncles supporting capitula or were immature). Peduncle branching and number of capitula per branch stem were counted on all mature flowering or fruiting individuals sampled (mean of 3.1 individuals per population) and the mean calculated for each population. Due to there being no standard terminology in place for the stem-like structures supporting the capitula and determining their arrangement, I here use the term "peduncle" for the entire stem-like structure between the true stem and the capitula. Stems and peduncles are distinguished based on: (i) the presence of leaves on stems but not on peduncles, (ii) leaves on stems being consistently spaced while bracteoles on peduncles are not, and (iii) peduncles having a greater density of glandular trichomes than stems. Fruits from each population were examined using a 10x hand lens and anther appendage colour and involucral bract trichome traits were examined under a dissecting microscope. Morphological similarity of populations in each major clade was assessed by PCA of continuous traits using the "prcomp" function from the stats package version 4.3.2 in R (R Core Team, 2023). The full morphological dataset is given in Appendix 4.

Table 2. Details of the morphological traits scored for each population in the study. An abbreviated code is given for each continuous trait used in the PCA.

Trait	Code	Trait type and units	Trait description
Plant height	HGT	Continuous (cm)	The distance from the base of the stem to the topmost tip of a flowering plant. Measured using a tape measure (precise to 1 cm), ruler (precise to 1 cm) or estimated (precise to 30 cm if > 1 m, precise to 10 cm if < 1 m).

Leaf width	LWD	Continuous (mm)	The maximum width of a fully developed medial cauline leaf. Measured using vernier calipers (precise to 0.01 mm).
Leaf length	LLN	Continuous (mm)	The length of a fully developed medial cauline leaf. Measured using vernier calipers (precise to 0.01 mm).
Leaf width:length	LWL	Continuous (ratio)	The width divided by the length of a leaf. Small values indicate more linear leaves.
Relative position of widest point of leaf	LWP	Continuous (ratio)	The position of the widest point of the leaf relative to the length of the leaf, giving an idea of leaf shape. Calculated as leaf length - length from tip to widest point)/ leaf length. Smaller values indicate that leaves are widest near the base while larger values indicate leaves that are wider near the tip (e.g., ovate vs obovate).
Peduncle length	PLN	Continuous (mm)	The measurement of the longest mature peduncles supporting fully developed capitula and taking curvature into account. In cases where the peduncles branch, this was the length of a terminal branch supporting a single capitulum. Measured using vernier calipers (precise to 0.01 mm).
Peduncle branching	PBR	Continuous (count)	The maximum number of branching events along the peduncle, between its point of attachment to the stem and the terminal capitulum.
Number of capitula per branch stem	CPT	Continuous (count)	The maximum number of capitula at the end of a stem. The number will be higher if there are either multiple unbranched peduncles at the end of a stem or if peduncles branch multiple times and therefore support multiple capitula.
Apical anther appendage colour		Binary	The colour of extruded apical anther appendages, scored as 0 = yellow/gold/white/transparent, or 1= black/grey. Examined under a dissecting microscope.
Involucral bract trichome presence		Binary	The presence of trichomes on the involucral bracts, scored as 0 = trichomes absent or 1 = trichomes present. Examined under a dissecting microscope.

Involucral bract trichome pigmentation	Binary	The pigmentation of trichomes (not including the apical gland, if present) on involucral bracts, scored as 0 = none pigmented (white or absent) or 1 = some pigmented. Examined under a dissecting microscope.
Involucral bract hairs	Binary	The presence and abundance of elongated filamentous non-glandular involucral bract trichomes, here referred to as hairs, scored as 0 = absent or sparse or 1 = abundant and forming a dense layer over the bract. Examined under a dissecting microscope.
Fruit surface texture	Binary	The surface sculpturing of a fruit, scored as 0 = mostly smooth (no pitting or ridges), or 1 = conspicuously pitted and ridged. Examined using a 10x hand lens.
Fruit primary ridge edges	Binary	The cross-sectional shape of the three primary longitudinal ridges on the fruit, scored as 0 = ridges absent or obtuse and not membranous, or 1= sharply acute and often membranous. Examined using a 10x hand lens. Norlind (1943) distinguished *O. imbricatum* from *O. polygaloides* based on the former having ridges that are sharply acute with the edges drawn into a thin membranous structure like a tiny wing.

Ecological data

To enable a consideration of ecology in the grouping of populations into species-level taxa, each population was characterized in terms of the vegetation type group in which it was located. For this purpose, I queried the vegetation type group from the Vegetation Map of South Africa, Lesotho and Swaziland (South African National Biodiversity Institute, 2009) at the coordinates of each population using the R package sf version 1.0.16 (Pebesma, 2018). The vegetation types in South Africa are mapped based on a model using plant distribution records, topography, geology, soils and climate data (Mucina & Rutherford, 2006), and so represent an integrative measure of ecological association. Since vegetation type can change at small spatial scales and therefore not always be accurately captured by the map when populations are close to the edge of a vegetation type polygon, vegetation type assignment was also informed by field observations. The percentage of populations per vegetation type group was calculated for the population groups identified from the molecular and morphological data to determine whether they are ecologically differentiated.

Results

Monophyly of Osteospermum section Polygalinae

The final ETS alignment comprised 413 bp, with 143 distinct patterns, 60 parsimony informative sites, 49 singleton sites and 304 constant sites. The ITS alignment comprised 633 bp, 217 distinct patterns, 85 parsimony-informative sites, 71 singleton sites and 477 constant sites. The concatenated Bayesian 90% majority-rule consensus tree (Figure 2) recovered the same main clades B, C, D, E and I, within *Osteospermum* as the more broadly sampled phylogeny of Sadler (2024). Clade I (Bayesian PP, expressed as a percentage = 100) contains all accessions that match the type species of *O*. sect. *Polygalinae*, *O*. *polygaloides* L. as well as the accessions that match the other taxa currently included in this section (*O. australe, O. burttianum, O. corymbosum, O. imbricatum,* and *O. rotundifolium*). All newly sampled morphologically distinct accessions sampled in this study are also resolved within Clade I, with one exception, confirming their inclusion in *O*. sect. *Polygalinae*.

The one morphologically distinctive newly sampled accession not recovered in Clade I, LMCJ 204 (marked by an asterisk in Figure 2), was placed within Clade E (Figure 2, Bayesian PP = 99.95) and is strongly supported as sister to both accessions of *O. junceum* within Clade H (Figure 2, Bayesian PP = 100). Accession LMCJ 204 is from the Groot Winterhoek mountains in the Eastern Cape Province. It is a large proteoid shrub with robust, woody stems, large, entire leaves, dense capitulescences and fleshy involucral bracts. Of the known species of *Osteospermum*, it is morphologically most similar to both *O. junceum* (with which it shares large, slightly fleshy fruits and involucral bracts in multiple series) and *O. corymbosum* (with which it shares large, fleshy leaves and densely branched capitulescences). Its placement indicates that it is most closely related to the former, and that it is not a member of *O*. sect. *Polygalinae*. This species is undescribed and has been listed as "*Osteospermum* sp. 3" by Manning and Goldblatt (2012). It has not previously been included in a molecular analysis. Since this specimen is not a member of *O*. sect. *Polygalinae*, it was excluded from further analyses and will not be discussed further, but its phylogenetic placement confirms the morphological indications that it is a novel species in need of description.

Six well-supported clades (J through O) were recovered within *O.* sect. *Polygalinae* (Clade I) (Figure 2), but the relationships amongst these clades and two unplaced accessions were unresolved (Bayesian PP < 90). The ITS and ETS accessions of each of the three deconstructed incongruent samples were recovered in different clades, with the ETS of LMCJ190 and ITS of LMCJ95 being resolved in Clade L, while the ITS of LMCJ190 and ETS of LMCJ95 are resolved in Clade M. The ETS and ITS accessions of LMCJ217 were resolved in Clades N and O respectively. With the exception of Clade J, which houses both accessions of *O. imbricatum* subsp. *nervatum*, none of the clades within *O.* sect. *Polygalinae* (Clade I) correspond to a currently recognised taxon. Samples identified as *O. polygaloides* were recovered in clades K, N and O and as an unplaced accession; while those

identified as *O. imbricatum* were recovered in clades J, L, M and O and as the other unplaced accession.

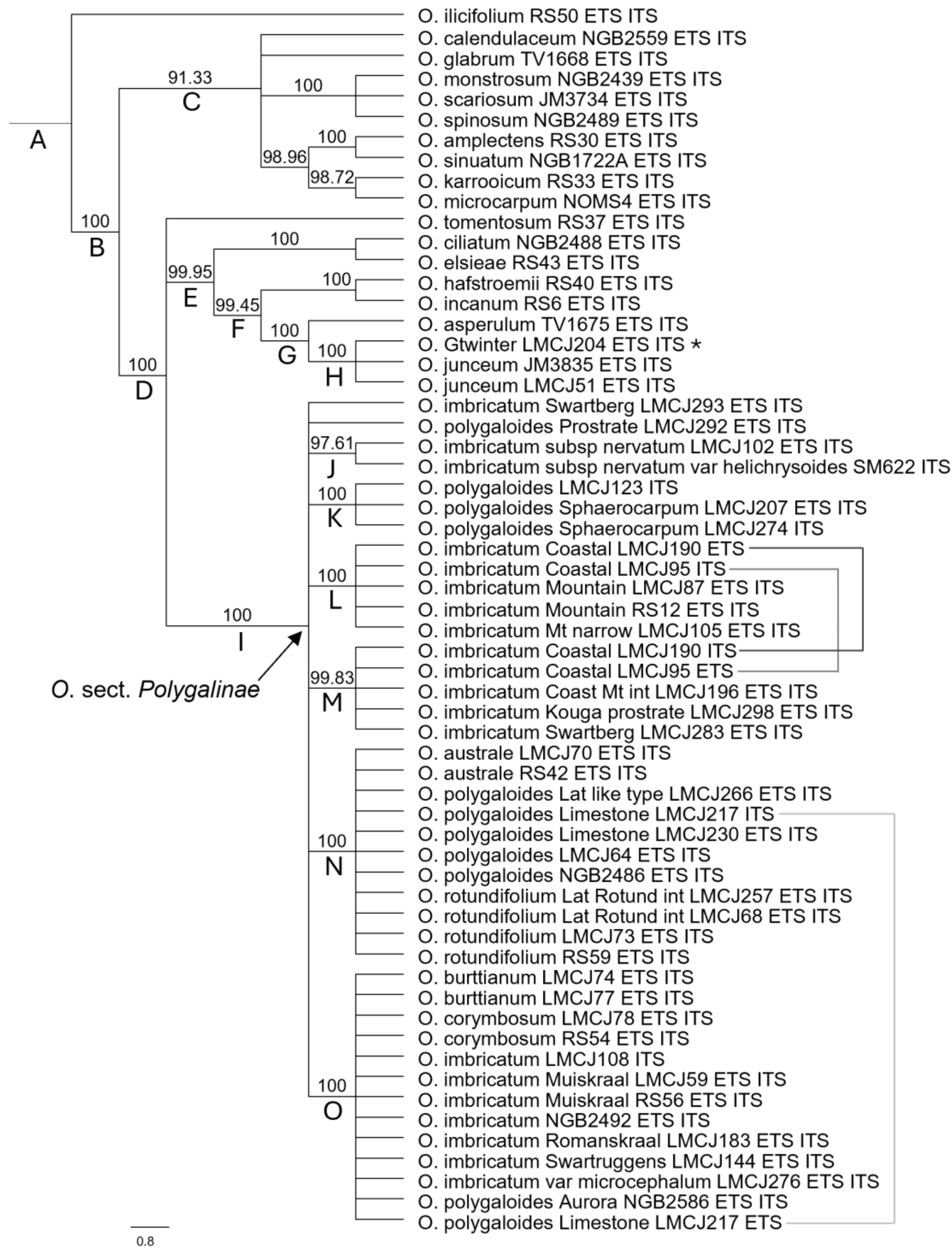


Figure 2. Bayesian 90% majority-rule consensus tree of concatenated ETS and ITS sequences of species in *Osteospermum*. Numbers above branches indicate Bayesian Posterior Probabilities (PP) expressed as percentages. Accessions are named according to their assignment using the currently accepted taxonomy and are given morphological form names when their morphology does not match that of the type. In most cases each tip is represented by both ETS and ITS sequences, but three sequences (LMCJ95, 217 and 190) have been deconstructed into separate ETS and ITS accessions due to topological incongruence. Also, only ITS sequences were available for SM622 (*O. imbricatum* subsp. *nervatum* var. *helichrysoides*), LMCJ123 (*O. polygaloides*) and LMCJ108 (*O. imbricatum*).

Relationships within Osteospermum section Polygalinae

Filtering the concatenated alignment from the GBS dataset to include only sites having < 20% missing data, relative to < 30% missing data, reduced the total number of sites by ca. 30% and the number of parsimony-informative sites by ca. 31% (Table 3). However, the phylogenetic trees inferred from these two data sets were very similar, both recovering the same four major clades with high support (SH-aLRT support > 91% and UFB support > 98%) and the same generally well supported subclades (Figure 3 and Appendix 5). Since branch support was generally higher when using the larger dataset, this phylogeny was preferred as a framework for downstream interpretation and analysis (Figure 3).

Table 3. Comparison of the GBS dataset for phylogenetic inference when filtered for different proportions of missing data per site.

Proportion of missing data per site	Total number of sites	Constant sites	Parsimony-informative sites	Singleton sites
<30%	444479	425730	11080	7669
<20%	309340	296127	7624	5589

The general topology of the GBS phylogeny (Figure 3) is congruent with the phylogeny produced from ETS and ITS sequences (Figure 2) but has much greater resolution. The relationships between the six subclades resolved from ETS and ITS sequences (Figure 2) were recovered in the GBS phylogeny (Figure 3) as follows: Clades J, L and M and the two unplaced lineages were all resolved in Clade 1; accessions in Clade O were split between Clade 2 and 3 with the exception of the ETS accession of LMCJ217, which was placed in Clade 4; all accessions in Clades K and N were recovered in Clade 4.

In the GBS phylogeny Clades 3 and 4 are well supported as sisters (SH-aLRT = 99.9%, UFB = 100%, sCF = 36.1%), which are together sister to Clade 2 (SH-aLRT = 99.7%, UFB = 100%, sCF = 39.3%), while Clade 1 is sister to all other clades (Figure 3). Further well supported relationships were revealed within the four major clades (e.g., subclades and lineages A to P), some of which align with currently recognized taxa, for example, Subclade C contains all accessions identified as *O. imbricatum* subsp. *nervatum* and Subclade N contains all accessions identified as *O. australe* (Figure 3).

The four major clades inferred in the GBS phylogeny were independently resolved as distinct genetic clusters in the PCA of all samples (Figure 4). Accessions POL12, POL14, POL15 and POL17

(Subclade K) from Clade 4 form a central cluster in the PCA separate from the four main clusters (Figure 4). Together with the long branch subtending Clade K (Figure 3), this suggests that these samples could be sufficiently genetically distinct to be treated as a fifth major clade. However, due to its small sample size relative to the other major clades, it was instead included in Clade 4 for comparative analyses, which was appropriate due to being sister to the rest of Clade 4 (Figure 3). Since each of the four major clades represent a well-supported and independent genetic group, the internal structure of each clade and the genetic and morphological relationships of its component samples were investigated separately for each clade to determine species circumscriptions. This also allowed species-level resolution within the molecular PCAs and sNMF analyses. The sNMF structure plots for the full range of *K* values investigated are given in Appendix 6 and only those relevant for species delimitation corresponding to the number of genetic clusters in the PCAs are presented here.

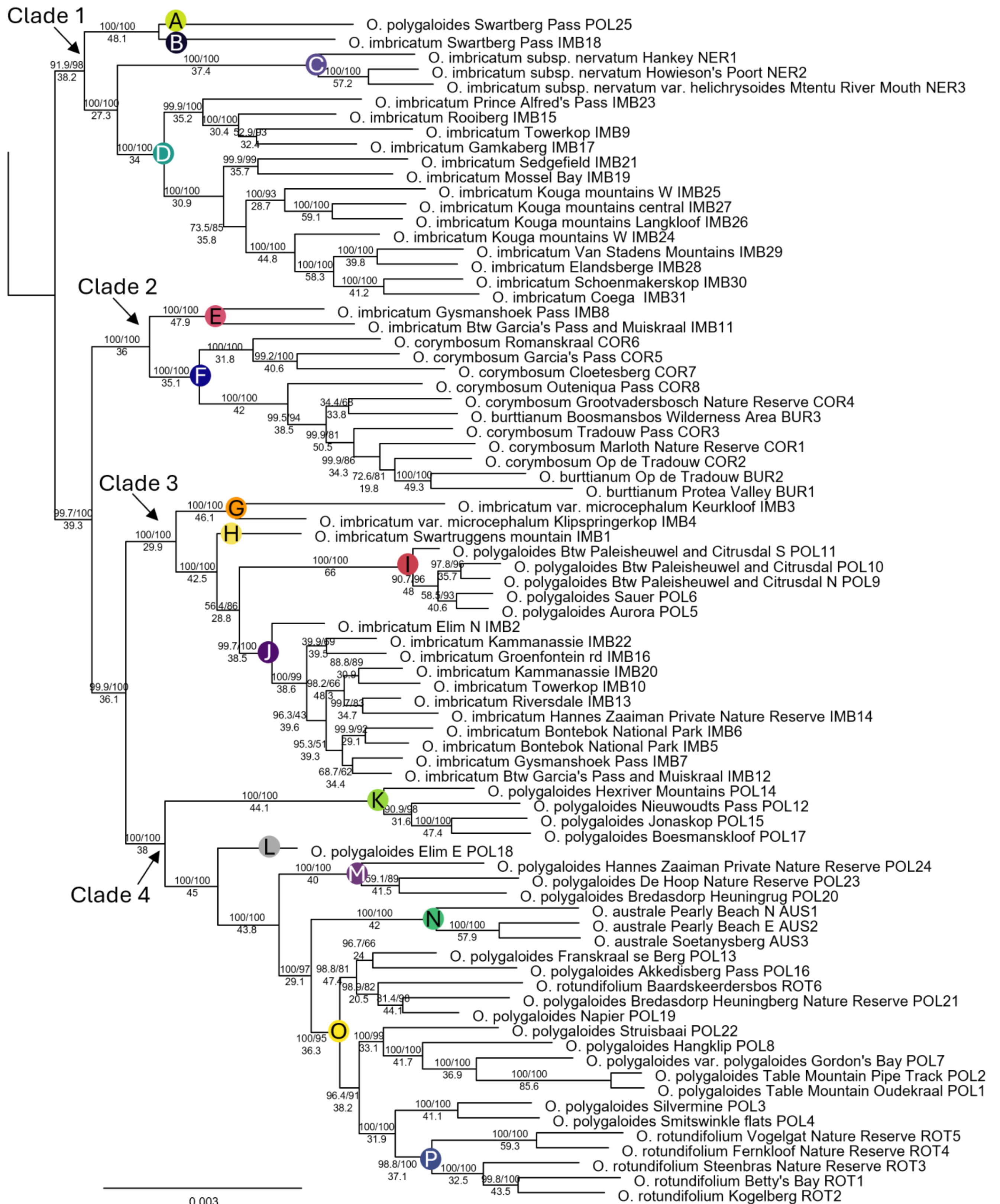


Figure 3. Maximum likelihood phylogeny of 79 *Osteospermum* section *Polygalinae* accessions from different populations inferred from concatenated GBS data filtered to < 30 % missing per site. The tree was rooted on an outgroup (*O. junceum*, not shown). Major clades are indicated by arrows and subclades / lineages by letters in coloured circles. Tip labels give the best name under the current taxonomy, collection locality and population code. Branch supports are given as SH-aLRT support (%) / ultrafast bootstrap support (%), with sCF % below.

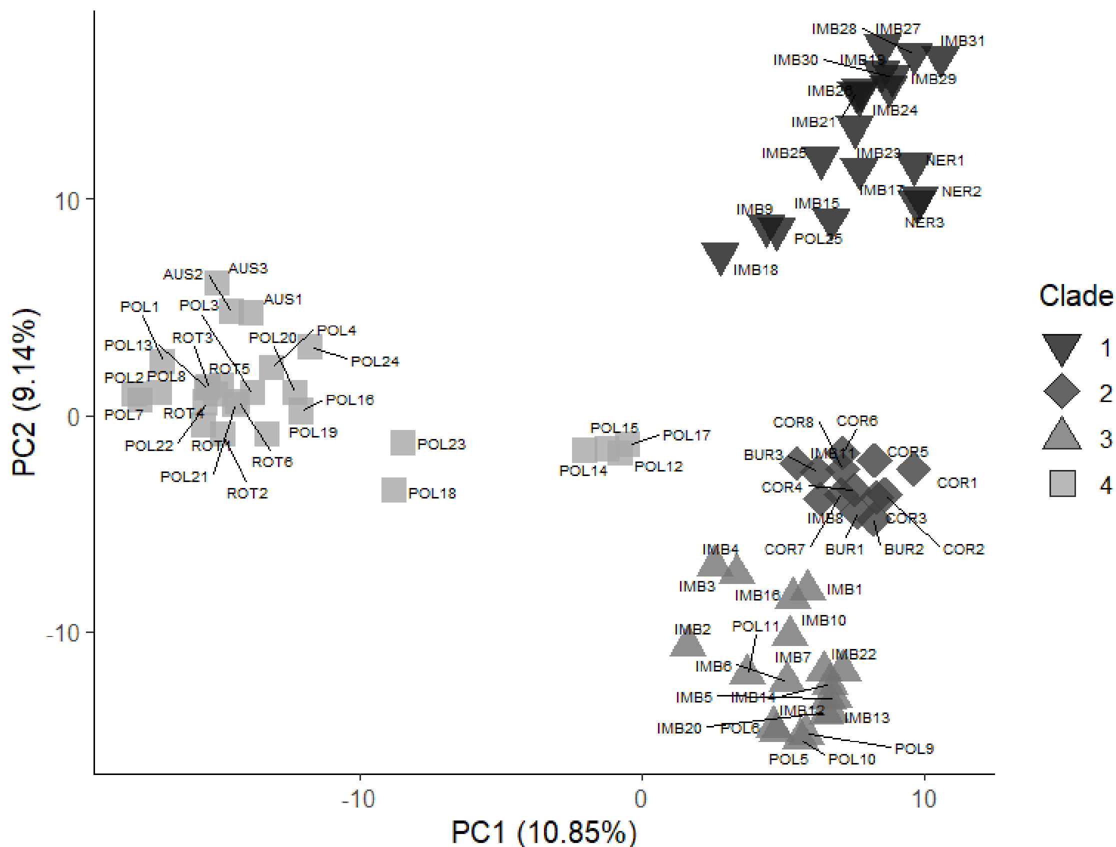


Figure 4. Principal component analysis ordination of genotyping-by-sequencing single-nucleotide polymorphism data from 79 *Osteospermum* section *Polygalinae* populations. Symbols correspond to the major clades resolved in the maximum likelihood phylogeny (Figure 3).

Delimitation of species-level entities within major clades

Clade 1

Clade 1 comprises lineages A to D, all of which consist of samples collected from localities east of Mossel Bay near the coast, or from mountains in the Little Karoo (all east of 21° longitude) (Figure 5a). Within Clade 1, Subclade C is a well-supported lineage (Figure 5b) that is genetically (Figure 5d i) and morphologically (Figure 5c) distinct. It also has a unique ancestral gene pool assignment (Figure 5e). Subclade D, the sister of Subclade C, was also resolved with good support (Figure 5b) and, despite being very morphologically and genetically heterogeneous (Figures 5c, d, e), it is clearly distinct from Subclade C. Although Subclade D shows some well supported internal phylogenetic structure, its component lineages are not morphologically differentiated. Despite evidence for some gene pool sharing with Subclade D at the western end of the latter's range (Figure 5a), lineages A and B, each represented by a single population, are genetically distinct from each other and from the rest

of Clade 1 (Figure 5d ii). Morphologically they are distinct, both from each other and from subclades C and D, although both lie close to samples from Subclade D in PCA morphospace (Figure 5c). Lineage B is morphologically unique, being the only sample in which glandular trichomes on the involucral bract are absent (Appendix 4). The Lineage A sample is also unique, its fruits having obtuse rounded ridge edges, contrasting to the acute membranous ridge edges on fruits of all other members of Clade 1. The samples from lineages A and B also differ from each other in anther appendage colour and growth form, Lineage A having a distinct low-growing prostrate habit with plants being only about 10 cm high, while the Lineage B samples have erect, near-vertical stems, growing to a height of 1.5 m. Where populations in Subclades A and B occur in Sandstone Fynbos, those in subclades C and D were collected from a range of vegetation types (Table 4).

Clade 2

Clade 2 comprises two well-supported subclades, E and F (Figure 6b), with all populations restricted to the Cape Fold Mountains, stretching from the western Langeberg near Swellendam eastwards to the western Outeniqua Mountains near George (Figure 6a; between 20.3° and 22.5° E). All populations in Clade 2 were found in Sandstone Fynbos (Table 4). Subclade E is genetically distinct (Figure 6d) and has a unique ancestral gene pool assignment (Figure 6e). Unfortunately, only one of the two populations of Subclade E could be included in the PCA of morphological data due to all individuals in population IMB8 being immature when sampled. Apart from COR6, which occupies an intermediate position, IMB11 is well separated from Subclade F accessions in the PCA morphospace (Figure 6c) and based on the vegetative traits that could be measured for IMB8 and IMB11, they are morphologically highly similar, particularly in terms of their distinctive leaf shape. Moreover, where the populations in Subclade F are characterised by fruits with elaiosomes and smooth surfaces, Subclade E fruits lack elaiosomes and have ridged and pitted surfaces. Subclade F is well-supported as sister to Subclade E. Subclade F comprises two well-supported smaller clades that are also recovered in the PCA of the GBS data (Figure 6d) and ancestral gene pool assignment (Figure 6e), however, this genetic division lacks morphological corroboration (Figure 6c).

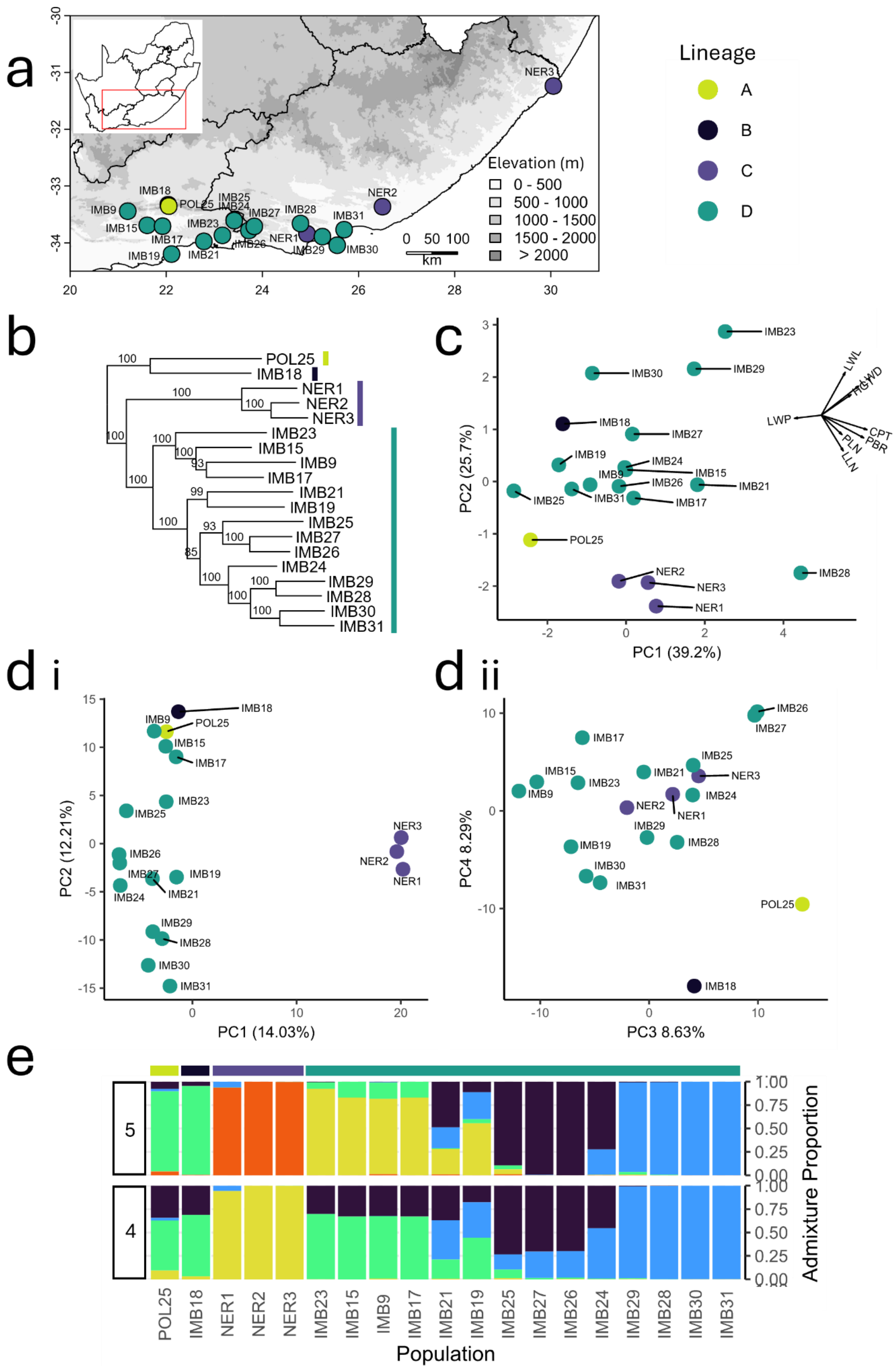

a
Lineage
A
B
C
D
Elevation (m)
0 - 500
500 - 1000
1000 - 1500
1500 - 2000
> 2000
km
b
c
PC1 (39.2%)
PC2 (25.7%)
d i
PC1 (14.03%)
PC2 (12.21%)
d ii
PC3 8.63%
PC4 8.29%
e
Admixture Proportion
Population

Figure 5. Species delimitation in Clade 1. (a) Map showing the collection localities of populations in Clade 1. Note that symbols for POL25 (Lineage A) and IMB18 (Lineage B) are superimposed, as are symbols representing IMB24 & 25 (Subclade D). (b) Maximum likelihood phylogeny, with branch labels describing ultra-fast bootstrap support. Subclades / lineages are indicated by coloured bars. (c) Principal component analysis (PCA) ordination of continuous morphological traits with inset showing the morphological trait vectors. Traits are abbreviated as follows: HGT = plant height; LWD = leaf width; LLN = leaf length; LWL = leaf width to length ratio; LWP = relative position of widest point of leaf; PLN = peduncle length; PBR = peduncle branching; CPT = number of capitula per branch stem. (d) PCA ordination of genotyping-by-sequencing (GBS) single-nucleotide polymorphism data. Two biplots are presented, showing the significant axes of variation, (i) PC1 and PC2 and (ii) PC4, as determined by Tracy-Widom tests. (e) Ancestral gene pool assignment of each population inferred using sparse nonnegative matrix factorization (sNMF), for K = 4 and K = 5. The colours of the vertical bars represent different gene pools. The species delimited in this study are indicated by the colours of the circles in a, c and d; the vertical bars in b; and the horizontal bars above the plot in e.

Table 4. The percentage of sampled populations of each lineage occurring in the vegetation type groups classified by Mucina & Rutherford (2006) as determined from the Vegetation Map of South Africa, Lesotho and Swaziland (South African National Biodiversity Institute, 2009) and informed by field observations. The number of populations sampled in each lineage are given in parentheses.

Lineage	Sandstone Fynbos	Sand Fynbos	Limestone Fynbos	Quartzite Fynbos	Silcrete Fynbos	Granite Fynbos	Shale Fynbos	Shale Band Vegetation	Ferricrete Fynbos	Conglomerate Fynbos	Eastern Strandveld	Limestone Renosterveld	Shale Renosterveld	Albany Thicket	Indian Ocean Coastal Belt
A (1)	100														
B (1)	100														
C (3)							33			33					33
D (14)	64	7	7					7			7			7	
E (2)	100														
F (11)	100														
G (2)	50							50							
H (1)				100											
I (5)	60	40													
J (11)	9				18		27		9			9	18	9	
K (4)	100														
L (1)									100						
M (3)			100												
N (3)			100												
O (12)	50		8			17	8	8					8		
P (5)	100														

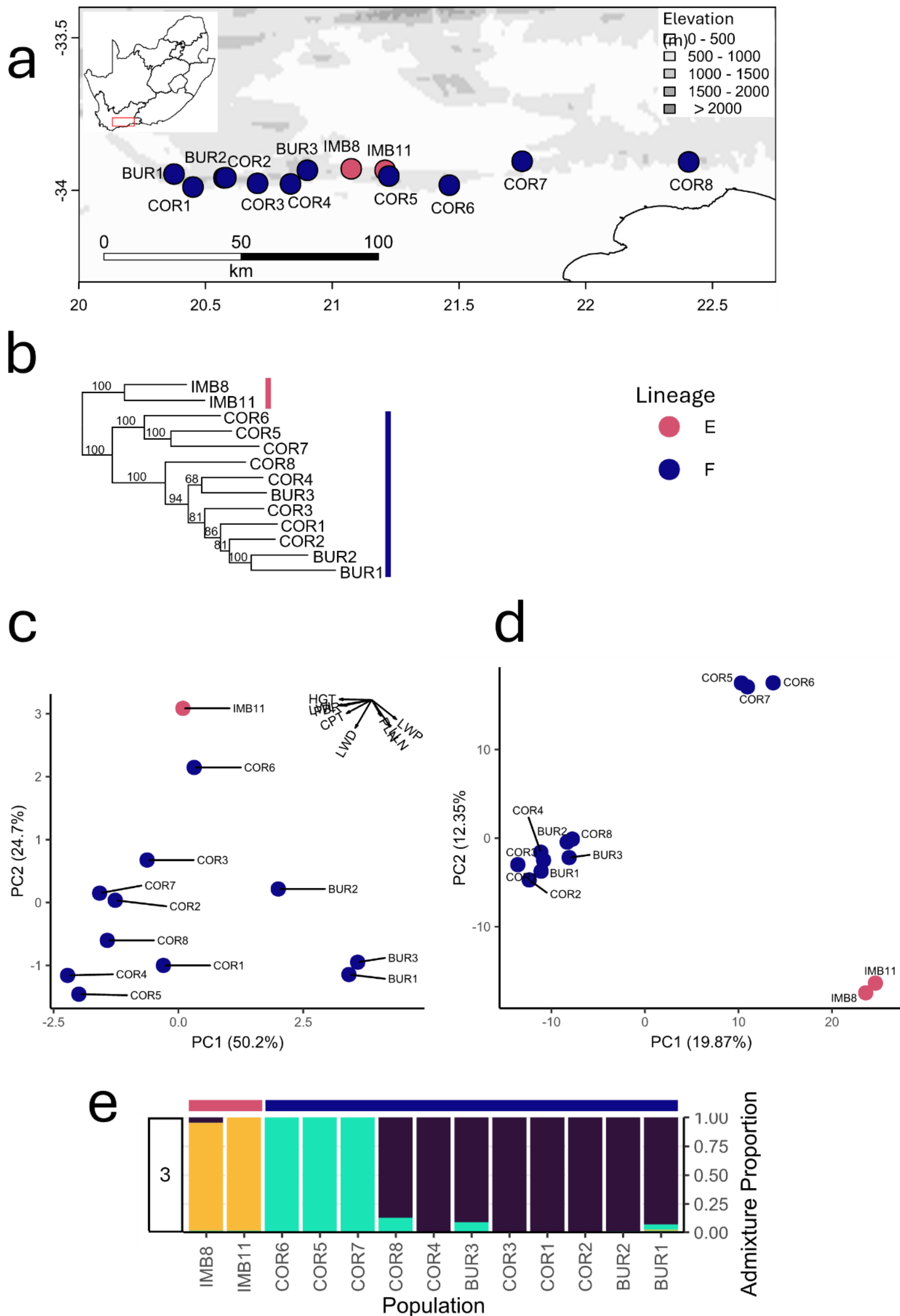


Figure 6. Species delimitation in Clade 2. (a) Map showing the collection localities of populations in Clade 2. Note that symbols for BUR2 and COR2 (both Subclade F) overlap. (b) Maximum likelihood phylogeny, with branch labels describing ultra-fast bootstrap support and subclades indicated by coloured bars. (c) Principal component analysis (PCA) ordination of continuous morphological traits with inset showing the morphological trait vectors. Traits are abbreviated as follows: HGT = plant height; LWD = leaf width; LLN = leaf length; LWL = leaf width to length ratio; LWP = relative

position of widest point of leaf; PLN = peduncle length; PBR = peduncle branching; CPT = number of capitula per branch stem. Note that IMB8 is not included in this plot. (d) PCA ordination of genotyping-by-sequencing (GBS) single-nucleotide polymorphism data showing the significant axes of variation, PC1 and PC2, determined by Tracy-Widom tests. (e) Ancestral gene pool assignment of each population inferred using sparse nonnegative matrix factorization (sNMF), for K = 3. The colours of the vertical bars represent different gene pools. The species delimited in this study are indicated by the colours of the circles in a, c and d; the vertical bars in b; and the horizontal bars above the plot in e.

Clade 3

Clade 3 comprises lineages G to J, with all populations from the northwestern and central CFR (Figure 7a). The westernmost of these is Subclade I, which is a well-supported lineage (Figure 7b), that is genetically (Figure 7d) and morphologically (Figure 7c) distinct and has a unique ancestral gene pool assignment (Figure 7e). Populations in Subclade I occupy sandy habitats, primarily in Sandstone Fynbos (Table 4). While Subclade G is also well supported and genetically and morphologically distinct (Figure 7b, c, d), its two member populations (IMB3 and IMB4) are widely separated from each other in both the genetic and morphological PCA ordinations. In the case of the genetic ordination the separation appears to arise because IMB3 has accumulated some unique molecular variation, as indicated by its very long subtending branch in the phylogeny (Figure 7b). In the morphological ordination, however, it appears due to individuals in population IMB3 being older, taller, and more structurally complex, resulting in more capitula per branch stem. Both populations, however, share a distinctive leaf shape, having leaves with a high width to length ratio, and being widest at the base. The sNMF analysis inferred a largely distinct, highly shared gene pool for these two populations (Figure 7e), although this is also partly shared with Lineage H, which comprises a single isolated population (IMB1) from the Swartruggens Mountains (Figure 7a). Subclade J is also well supported (Figure 7b) and displays remarkable genetic and morphological cohesion (Figure 7c, d, e), despite being widespread (Figure 7a) and occurring in a range of vegetation types (Table 4). The morphological distinctiveness and distinctive genetic ancestry of Subclade J are only compromised by Lineage H (Figure 7a). Although Lineage H is not morphologically distinct from members of Subclade J based on the ordination, it can be separated from them based on anther appendage colour, and it is also distinct in the genetic ordination (Figure 7d). Lineage H also has a unique ecology, being the only population out of all *O*. sect. *Polygalinae* populations sampled occurring in Quartzite Fynbos (Table 4). Although subclades I and J are supported as distinct clades, the phylogenetic relationships between lineages H, I and J are uncertain due to low support (UFB = 86%) on the node indicating a sister relationship between subclades I and J (Figure 7b).

Figure 7. Species delimitation in Clade 3. (a) Map showing the collection localities of populations in Clade 3. Note that symbols for IMB5 and IMB6 (both Subclade J) overlap. (b) Maximum likelihood phylogeny, with branch labels describing ultra-fast bootstrap support. Subclades are indicated by coloured bars. (c) Principal component analysis (PCA) ordination of continuous morphological traits with inset showing the morphological trait vectors. Traits are abbreviated as follows: HGT = plant height; LWD = leaf width; LLN = leaf length; LWL = leaf width to length ratio; LWP = relative position of widest point of leaf; PLN = peduncle length; PBR = peduncle branching; CPT = number of capitula per branch stem. (d) PCA ordination of genotyping-by-sequencing (GBS) single-nucleotide polymorphism data showing the significant axes of variation, PC1 and PC2, determined by Tracy-Widom tests. (e) Ancestral gene pool assignment of each population inferred using sparse nonnegative matrix factorization (sNMF), for K = 3 and K = 4. The colours of the vertical bars represent different gene pools. The species delimited in this study are indicated by the colours of the circles in a, c and d; the vertical bars in b; and the horizontal bars above the plot in e.

Clade 4

Clade 4 comprises lineages K to P, which are all restricted to the western region of the CFR. All members are concentrated in the extreme southwest except for members of Subclade K, which occur northwards from the Riviersonderend Mountains to the Cederberg (Figure 8a). Subclade K is a very distinct group that is well supported as sister to the rest of Clade 4 (Figure 8b). Not only is it genetically distinct from the rest of Clade 4 (Figure 8d i), but it is also particularly genetically distinct from the rest of *O*. sect. *Polygalinae* (Figure 4, central cluster). It is also derived from a unique ancestral gene pool (Figure 8e) and all populations sampled occur in Sandstone Fynbos (Table 4). Although it is morphologically similar to other members of Clade 4 for the continuous traits measured (Figure 8c), it has distinctive, near-spherical fruits and is one of only two subclades in Clade 4 that lack purple pigmented trichomes on their involucral bracts. Subclade N is the other subclade which lacks pigmented involucral bract trichomes, but rather than short glandular trichomes it has white elongated filamentous trichomes that form a thick arachnoid indumentum. Subclade N is also a well-supported lineage (Figure 8b), that is genetically distinct (Figure 8d i), has a unique gene pool assignment (Figure 8e) and is morphologically distinct, despite lying close to ROT3 in PCA morphospace (Figure 8c). All populations constituting Subclade N were found in Limestone Fynbos (Table 4).

Lineages L and M, comprising one and three populations respectively, are each well supported and separated in the phylogeny (Figure 8b) but are indistinguishable from each other and from members of the subclades grouped as O (Grade O) based on the morphological traits measured (Figure 8c). These three lineages also have overlapping distributions (Figure 8a). Where Subclade M is genetically distinct (Figure 8d ii) and assigned to a unique ancestral gene pool (Figure 8e), the genetic distinctness of Lineage L (POL18, grey symbol) is marginal in the genetic PCA produced for Clade 4 (Figure 8d i). Interestingly, the genetic PCA performed across all of *O*. sect. *Polygalina* separated Lineage L (POL18) from the rest of Clade 4, placing it closer to Clade 3 (Figure 4). This might suggest that it has some genetic differences to members of Clade 4 that were lost when the data were

filtered to remove missing sites for the Clade 4 PCA. The data used for ancestral gene pool assignment were not filtered as strictly, and Lineage L (POL18) was inferred to have a mixed gene pool, largely shared with populations POL13, POL16, ROT6, POL21, POL19 and POL22 from Grade O (Figure 8e). Lineage L is represented by a population occurring in Ferricrete Fynbos (Table 4), a vegetation type in which only one other population was found, being IMB2 from Subclade J of Clade 3. All populations constituting Subclade M were in Limestone Fynbos, while populations in Grade O were from a range of vegetation types, although 50% were found in Sandstone Fynbos (Table 4).

Finally, Grade O and Subclade P, together form a supported monophyletic group, with Subclade P being embedded within Grade O and rendering the latter paraphyletic (Figure 8b). Where Grade O is genetically and morphologically heterogeneous (Figure 8c and d), with populations assigned to highly mixed gene pools (Figure 8e), Subclade P is genetically distinct (Figure 8d ii) and morphologically distinguishable, except from population ROT6 (Subclade O). One population from Subclade P, ROT3, appears isolated from the rest of Subclade P in PCA morphospace (Figure 8c), however this is due to its individuals being younger and shorter and therefore having shorter peduncles, with fewer peduncle branches and capitula, and it was otherwise morphologically similar. Subclade P shows some ancestral gene pool overlap with populations POL3 and POL4 of Grade O (Figure 8e). All populations of Subclade P occurred in Sandstone Fynbos (Table 4).

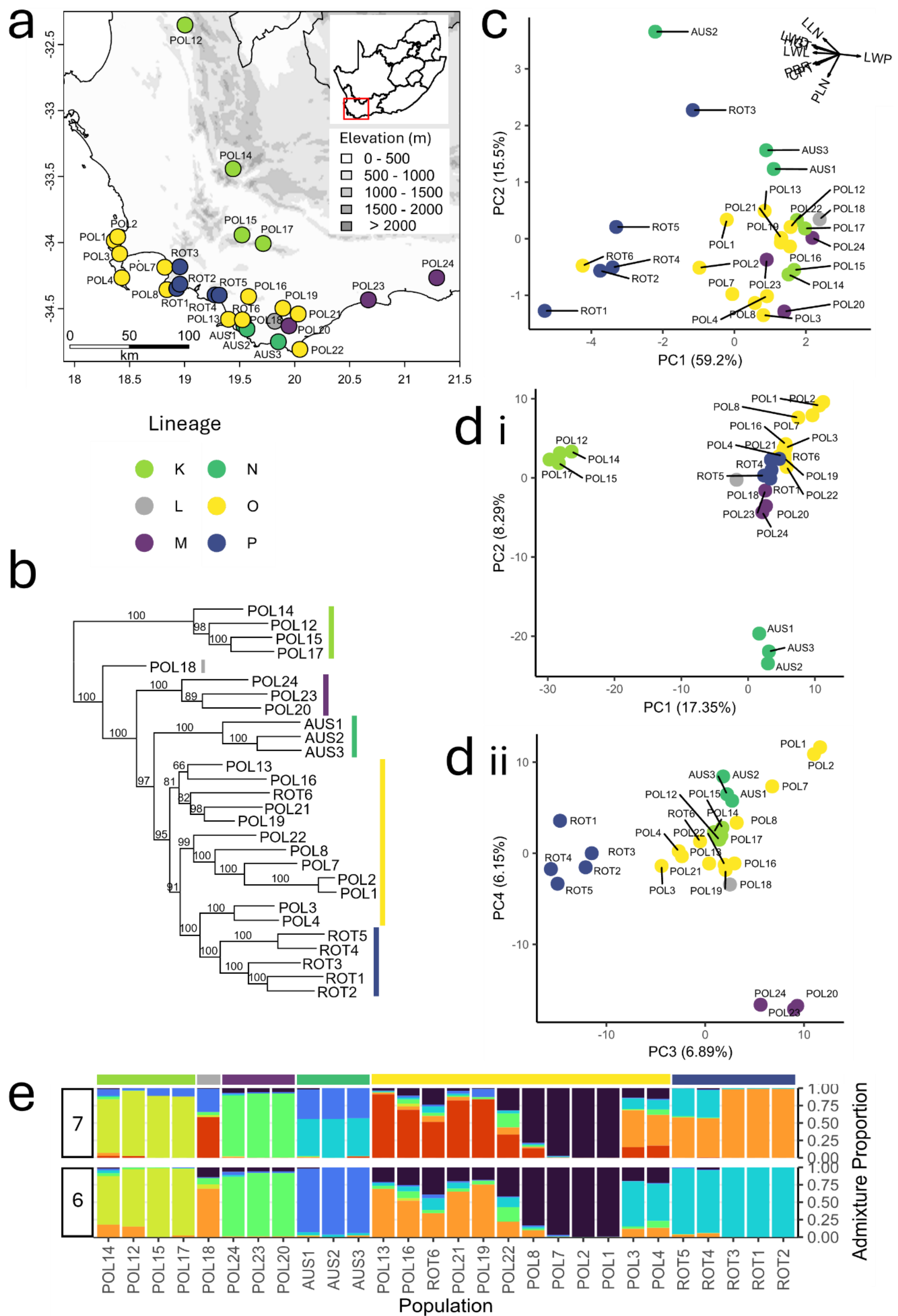


Figure 8. Species delimitation in Clade 4. (a) Map showing the collection localities of populations in Clade 4. Note that the symbol for AUS1 (Subclade N) is partly obscured by ROT6 (Grade O) and AUS2 (Subclade N). (b) Maximum likelihood phylogeny, with branch labels describing ultra-fast bootstrap support. Subclades are indicated by coloured bars. (c) Principal component analysis (PCA) ordination of continuous morphological traits with inset showing the morphological trait vectors. Traits are abbreviated as follows: HGT = plant height; LWD = leaf width; LLN = leaf length; LWL = leaf width to length ratio; LWP = relative position of widest point of leaf; PLN = peduncle length; PBR = peduncle branching; CPT = number of capitula per branch stem. (d) PCA ordination of genotyping-by-sequencing (GBS) single-nucleotide polymorphism data. Two biplots are presented, showing four of the significant axes of variation, (i) PC1 and PC2 and (ii) PC3 and PC4, as determined by Tracy-Widom tests. PC5 and PC6 were also significant, but uninformative for species delimitation, so that plot is included in Appendix 7 rather than here. (e) Ancestral gene pool assignment of each population inferred using sparse nonnegative matrix factorization (sNMF), for K = 6 and K = 7. The colours of the vertical bars represent different gene pools. The species delimited in this study are indicated by the colours of the circles in a, c and d; the vertical bars in b; and the horizontal bars above the plot in e.

Discussion

Monophyly of Osteospermum section Polygalinae

Currently, species are recognised as belonging to *O.* sect. *Polygalinae* by being shrubs or subshrubs with alternate, sessile, leathery leaves with entire margins that are usually elongated on lower stems, capitula with uniseriate or subuniseriate involucral bracts, long yellow ray floret ligules, and homomorphic, cylindrical fruits (Norlindh, 1943). In this, the first comprehensive test of monophyly, *O.* sect. *Polygalinae* is shown to be a monophyletic group (Figure 2) consistent with these defining traits. In particular, the feature of having uniseriate or subuniseriate involucral bracts separates these species from other similar *Osteospermum* species, which provides a morphological means to exclude accession LMCJ204 from the Groot Winterhoek Mountains in the Eastern Cape.

There is strong evidence both from the phylogenetic analysis of the GBS data and the PCA of SNP data for four independent and geographically structured lineages within *O.* sect. *Polygalinae* (Figures 3 and 4). Populations in Clade 1 are distributed in the eastern CFR, populations in Clade 2 are restricted to the Cape Fold Belt running east-west from the Langeberg near Swellendam eastward to the western portion of the Outeniqua Mountains near George, populations in Clade 3 occur in the lowlands in the central longitudes of the CFR, in the westernmost region of the Langeberg and the smaller mountains in western portion of the CFR, but notably not the main north-south Cape Fold Belt, where some members of Clade 4 are found, while most populations in Clade 4 are restricted to the southwestern region . Despite clear genetic separation between Clades 1 to 4 (Figure 3), these four major clades lack unique diagnostic synapomorphies (Figure 3 and Appendix 4), each instead displaying a range of attributes which have arisen repeatedly in other clades. It is for this reason that a taxonomy based on morphology alone has been unable to resolve entities corresponding to natural evolutionary units in this group. Norlindh (1943) primarily used fruit traits to define his species

concepts, particularly the fruit shape and surface texture and the number and prominence of the primary longitudinal ridges. In light of my data, it is clear that these features mostly do not represent synapomorphies, particularly within *O. imbricatum* and *O. polygaloides* which are here shown to be highly polymorphic. Accessions consistent with Norlindh's concept of *O. imbricatum* are placed in Clades 1, 2, and 3, while those consistent with his concept of *O. polygalodes* occur in Clades 1, 3 and 4 (Figure 3). Clearly, similar fruit traits have evolved independently in multiple lineages and cannot therefore be used to define species representing independent evolutionary lineages (de Queiroz, 1998, 2007). I therefore assess species delimitation independently of previous taxonomic concepts, relying rather on evolutionary relationships as revealed by a synthetic analysis of phylogenetic, genetic, and morphological data.

Species conceptualisation and operational delimitation criteria

The segregation of *O.* sect. *Polygalinae* into four well-supported major clades, each comprising multiple subclades that are in many cases morphologically and/or ecologically differentiated, suggests the presence of many more species in *O.* sect. *Polygalinae* than are currently recognised. In formalizing this taxonomic diversity, I employ a species concept which views species as evolutionary independent lineages (de Queiroz, 1998, 2005) whilst also recognizing the importance of ecological role as a feature of species (Freudenstein et al., 2017). The continuous nature of the speciation process, however, as well as variation in the order in which indicator criteria emerge (de Queiroz 2007), argues against a strict reliance on any single criterion as a basis for species hypothesis testing. Hence, the approach employed here is semi-flexible, consistent with the principles of integrative taxonomy (Sites & Marshall, 2004; Dayrat, 2005; Schlick-Steiner et al., 2010). My key criteria for the recognition of species in this study are as follows:

1. Although recent speciation by budding is expected to yield parent species that are initially paraphyletic, species are generally expected to be monophyletic. Shallow species paraphyly is permitted, however, where speciation is assumed to have proceeded via a budding process.
2. Species are generally expected to have distinct ecological roles, as indicated by habitat and/or morphological differentiation. Thus, species which show internal phylogenetic and genetic structure that is not accompanied by ecological divergence are not recognised as distinct at the species level.
3. When there is no evidence of ecological differentiation between lineages, but they are deeply diverged and if grouped together would form deeply paraphyletic entities, they are

considered as distinct, cryptic species. In such cases species are generally diagnosable based on their distributions.

Species in Osteospermum section Polygalinae

Clade 1

I recognise four species in Clade 1 (Table 5). Three of these, corresponding to lineages A, B and C (hereafter Species A, B, and C), are morphologically and genetically clearly defined, while the fourth is a widespread and morphologically variable species complex which requires further study (Subclade D, hereafter Species D). The most distinct of these is Species C, which has the easternmost distribution of all species in *O.* sect. *Polygalinae*. Herbarium specimens corresponding morphologically to this species indicate a distribution extending from near George (Western Cape Province) in the CFR, through the Eastern Cape to just south of Richards Bay (Kwazulu-Natal Province). It occurs in coastal or near-coastal subtropical grassy habitats, being collected in Conglomerate Fynbos, Shale Fynbos and Indian Ocean Coastal Belt vegetation in this study. This species, which is morphologically diagnosable by its exceptionally long, narrow leaves (e.g., Chapter 1; Figure 3F), was first recognized taxonomically by de Candolle (1837) who described it as *Osteospermum nervatum*. Norlindh (1943) considered it to be conspecific with *O. imbricatum* because its fruits matched his concept of that species and due to the apparent existence of intermediate forms, and reduced it to *O. imbricatum* subsp. *nervatum*. Apart from its genetic distinctiveness, the morphological distinctiveness of Species C, even in the western part of its range where it overlaps with Species D, supports its recognition as a distinct species.

Species D is a widespread species and occurs across a range of contrasting vegetation types and habitats which include moist coastal lowlands (e.g., Chapter 1; Figure 2D) and moist mountain slopes, and the relatively hotter and drier mountains of the Little Karoo and Kouga Mountains. This species comprises specimens placed by Norlindh (1943) within his broad concept of *O. imbricatum* based on their fruit traits, but which do not otherwise match the type of that species. Specimens such as these have contributed to confusion regarding the concept of *O. imbricatum* and its separation from *O. polygaloides* (Manning & Goldblatt, 2012). While Species D shows supported internal phylogenetic structure, a lack of clear correspondence to morphology currently prevents a formal recognition of this structure as taxonomic entities within Species D. Accordingly, I suggest that this complex be provisionally recognized as a single species, pending further work, involving denser population sampling and more detailed evaluation of morphology.

Species A and B are both known only from single localities in the Swartberg. Although the morphological distinctness of these entities from Species D is not entirely clear from the PCA based

on the set of quantitative traits measured (Figure 5c), both species are readily distinguishable. While Species B has a similar growth form to some members of Species D, being a sparsely branching erect shrub about 1.5 m tall (e.g., Chapter 1; Figure 4B), it uniquely lacks glandular trichomes on its involucral bracts, has very widely spaced leaves, and its leaves and stems are glaucous and concolourous. This contrasts with all other *O.* sect. *Polygalinae* species which all have involucral bract trichomes and either imbricate leaves or brighter colouration with greater contrast between leaves and stems and frequently have maroon leaf margins or apices. Species B is the only *O.* sect. *Polygalinae* species occurring on the drier northern slopes of the Swartberg, where it grows in a rocky habitat at around 1414 m elevation. In contrast, Species A has a unique, low-growing prostrate habit, being only about 10 cm tall with secund leaves (e.g., Chapter 1; Figure 4A). It grows along the southern summit ridge of the Swartberg (around 1614 m elevation) and is associated with cool, damp conditions. Neither Species A nor Species B have been recognised by previous taxonomists.

Clade 2

Clade 2 contains two distinct species, corresponding to Subclades E and F (hereafter Species E and F). Species E is previously undescribed and known only from two populations, both collected on the dry, lower northern slopes of the eastern Langeberg in Sandstone Fynbos. Superficially this species most closely resembles Species C (*O. imbricatum* subsp. *nervatum*) from Clade 1, due to its similar erect growth form with sparse branching in the lower stems and long, narrow lower cauline leaves. However, Species C and E are only distantly related (Figure 3) and differ greatly in other morphological traits such as height, anther appendage colour, involucral bract trichome pigmentation, and fruit morphology, so the vegetative similarities can be assumed to be the result of convergence.

In contrast to Species E, Species F occurs in damper habitats, growing mostly on the moister southern slopes (e.g., Chapter 1; Figure 4C) and only being found on northern slopes at high elevations (> 890 m). Species F consists of populations representing three distinct morphologies (Figure 6c): (i) a low-growing cushion forming shrublet (plant height < 0.6 m) with long, narrow oblanceolate leaves, which was previously described as *O. burttianum* B.Nord; (ii) a tall (> 1 m, often > 2 m, tall), robust, and thick-stemmed shrub that branches only near the top of the plant to form a dense corymbose capitulescence and has broad leaves that taper towards the apex, previously described as *O. corymbosum* L.; and (iii) a tall (ca. 1.5 m tall) shrub, corresponding to accession COR6 from the farm "Romanskraal" in the eastern Langberg, which resembles *O. corymbosum* but is much more delicate, with narrower stems, smaller leaves, less peduncle branching and fewer capitula per branch stem. To my knowledge, specimens with the Romanskraal morphology have not been previously collected or described. While these three morphologies are clearly distinguishable, they are not corroborated by the genetic data which separates three of the more eastern population (COR5, COR6 and COR7), but notably not the easternmost population, from the rest, one of these corresponding to the Romanskraal

morphology and two to the *O. corymbosum* morphology, the latter being morphologically indistinguishable from the other *O. corymbosum* populations (Figure 6d). Further, populations corresponding to the *O. burttianum* morphology (BUR1, BUR2 and BUR3) do not form a clade, with population BUR3 being phylogenetically isolated from the other two. Only three *O. burttianum* populations are known, all occurring on high elevation mountain ridges (> 1030 m) in the central Langeberg, where they have never been observed growing in direct sympatry with individuals having the *O. corymbosum* morphology. Other than this elevation difference, however, there are no notable ecological differences between the populations of Subclade F, all being found in moist Sandstone Fynbos. In this context, I find little evidence to suggest that Subclade F consists of more than one species, with the observed morphological variation instead being interpretable as ecotypic or plastic. I therefore recognise Subclade F as a single species (Species F) comprising three morphological varieties (Table 5). Further studies are needed to determine the basis for the morphological differences.

Clade 3

Clade 3 contains four species, corresponding to lineages G, H, I and J (hereafter Species G, H, I and J). Species I is a highly distinct but previously undescribed species that occurs in sandy habitats in the northwestern portion of the CFR (e.g., Chapter 1; Figure 2C). It is a shrub growing to ca. 1 m tall, with unique very large fruits that are smooth apart from three slight ridges, and has the longest ray ligules of all species in *O.* sect. *Polygalinae*. Species J matches the morphology of the type specimen of *O. imbricatum* L., being a shrublet (< 80 cm tall), with small, generally rounded, highly imbricate leaves, and unbranched peduncles with capitula almost always solitary (rarely two capitula per branch stem). It is a lowland species, with all populations found at less than 670 m elevation in this study. It occurs in either renosterveld or fynbos, but tends to occur on soils derived from relatively nutrient-rich geologies such as ferricrete, silcrete, shale, and limestone, as opposed to sandstone (e.g., Chapter 1; Figure 2B). Species G, previously described as a variety of Species J, i.e., *O. imbricatum* subsp. *imbricatum* var. *microcephalum* Norl. (Norlindh 1943), is genetically and morphologically distinct from Species J and has very different ecology, occurring in the Langberg in fynbos either on sandstone or on the high elevation shale band near the summit ridge (> 800 m elevation). Species G has a unique cordate-amplexicaul leaf shape (e.g., Chapter 1; Figure 3H) and thin peduncles, which distinguish it from all other species in *O.* sect. *Polygalinae*. In contrast to Species G, I and J, Species H is not as clearly morphologically distinct. For the set of traits examined, it differs from Species J only by anther appendage colour, having white or yellow anther appendages compared to the black or grey appendages of Species J (e.g., Chapter 1; Figure 3B). While the admixed gene pool of this species is suggestive of hybridization between Species G and Species J, its geographical isolation (> 100 km) from both of these species suggests that this gene sharing is more likely a consequence of ancient introgression. Alternatively, it may be a result of bias in the gene pool assignment due to this

being a single population. Species H is nevertheless ecologically unique, occurring at over 1100 m elevation in the Swartruggens Mountains, where it inhabits a dry habitat in Quartzite Fynbos. This, paired with its genetic distinctness support it as a unique entity, and I recognise it here as a distinct, albeit morphologically cryptic, species.

Clade 4

Clade 4 contains four species corresponding to lineages K, M, N and P (hereafter Species K, M, N and P), a species complex corresponding to Grade O (hereafter Species O), and an entity of uncertain status (Lineage L). Species N is the most distinctive and charismatic of these, having a showy capitulum with many ray ligules, each with a prominent red streak on the abaxial surface, long involucral bracts, with red tips and white arachnoid hairs covering the involucral bracts and young leaves (e.g., Chapter 1; Figure 4E). Within Species N, plant height, leaf size, and capitulum size vary between populations, possibly as a result of growth conditions, leading to a wide spread of points in the PCA of morphological characters. This species is nevertheless easily distinguished from all others and was described as *Osteospermum australe* B.Nord. It is restricted to limestone ridges in the Cape Agulhas region, a habitat known for edaphic endemics (Cowling & Holmes, 1992; Willis, Cowling & Lombard, 1996).

The distinctness of Species K, particularly its unique near-spherical fruits, was also noticed by Bertil Nordenstam, but despite writing "*Osteospermum sphaerocarpum*" on several herbarium sheets, he never formally described this species. Species K is vegetatively very similar to Species M and O, all these species being shrubs with oblong to linear leaves. However, Species K has strictly solitary capitula and unpigmented involucral bract trichomes, while Species M and O possess peduncles that sometimes branch (once in M, up to three times in O), capitula which are not strictly solitary, and involucral bracts with pigmented trichomes. Species K also has a distinct distribution and ecology, occurring on the drier northern slopes of the Riviersonderend Mountains and northwards to the Cederberg, always in Sandstone Fynbos at high elevation (between ca. 500 and 1000 m). In contrast, Species M grows strictly on limestone ridges along the coast of the Agulhas Plain and eastwards to Still Bay, while Species O grows on a range of substrates mostly at low elevations (< 450 m), having a distribution range stretching from the Cape Peninsula, eastwards to the Agulhas Plain. Despite Species M having an overlapping distribution with Species O and being morphologically similar to some of its members, it is genetically and edaphically distinct. For this reason, and because its inclusion in Species O would render the latter polyphyletic, I recognise Species M as a distinct, albeit morphologically semi-cryptic, species.

Species O is widespread and morphologically variable, but this variability is not correlated with phylogenetic relationship, and I accordingly recommend that it be provisionally treated as a single species. Further work, involving denser sampling and a more detailed assessment of morphological

character variation, may, however, reveal it to be a complex of multiple species. Notwithstanding its variable morphology, Species O contains specimens that best match the type specimen of *O. polygaloides* L. in both morphology and distribution, being shrublets with narrow leaves that are linear to oblanceolate, peduncles with sparse branching and fruits which lack membranous ridge edges, and so it is most appropriately accorded this name.

Species P corresponds to the type morphology of *O. rotundifolium* (DC.) Norl. and is both genetically and morphologically distinct, being a sparsely branching shrub which stands > 1 m tall and has large, broad leaves, dark, blue-purple pigmented trichomes, and branching peduncles supporting multiple capitula per branch stem (e.g., Chapter 1; Figure 2A). The one exception is population ROT6 which is placed in Species O based on genetic data but would be morphologically identified as *O. rotundifolium*. Species P is distributed between the Kogelberg Mountains in the west and the Kleinrivierberge in the east. Within this range it associates with damp habitats, being most frequently found in valleys and furrows near streams, strictly in Sandstone Fynbos. Norlindh (1943) described *O. polygaloides* var. *latifolium* Norl. as a variety of *O. polygaloides* that had broadly elliptic, ovate or suborbiculate leaves. However, this designation was based exclusively on fruit traits, *O. polygaloides* var. *latifolium* otherwise being indistinguishable from *O. rotundifolium* and having the same distribution. From my observations, subtleties of fruit shape and surface texture can vary within a species and are highly dependent on fruit maturity, which is difficult to determine on dry herbarium specimens. I therefore do not recognise O. *polygaloides* var. *latifolium* as distinct from *O. rotundifolium*. While the recognition of Species P renders Species O shallowly paraphyletic, this does not compromise the status of the latter as an evolutionarily independent lineage, given that transient paraphyly is an expectation of widespread parent species giving rise to more localized daughter species via a budding speciation mechanism (Rieseberg & Brouillet, 1994; Hörandl & Stuessy, 2010).

Lastly, Lineage L, is a perplexing entity. In the genetic PCA of all samples, Lineage L (POL18) is placed away from the core Clade 4 cluster (Figure 4), a position that corresponds with its placement in the phylogeny, being intermediate between Species K and the rest of Clade 4 (Figure 8b). However, in the genetic PCA of just Clade 4 samples (Figure 8d), Lineage L only appears marginally genetically differentiated. This contrasts to Species K and POL23 of Species M, which are also not placed in the core cluster in the overall PCA (Figure 4) but are distinctly differentiated in the PCA of Clade 4 samples (Figure 8d). Additionally, Lineage L does not have a unique ancestral gene pool assignment but rather a mixed gene pool most similar to members of Species O. Apart from Lineage L comprising individuals of very small stature (12.5 cm tall) and having slightly more oblanceolate leaves, both fairly plastic traits, they are morphologically indistinct. A lack of genetic and morphological differentiation does not support the recognition of Lineage L as a distinct species, even if its species identity remains unclear. One possibility is that Lineage L represents a hybrid between members of Clade 3 and Clade 4, thus accounting for it intermediate placement in the PCA, its

somewhat basal placement in Clade 4, and its admixed gene pool profile (Patterson, Price & Reich, 2006; Smith et al., 2017; Linan et al., 2021; Yi et al., 2023). A hybrid interpretation of Lineage L is also geographically consistent, this lineage inhabiting Ferricrete Fynbos in an area of distributional overlap between Species O (Clade 4) and Species J (Clade 3).

Table 5. The alignment of the species delimitations determined in this study with the current taxonomy of *Osteospermum* section *Polygalinae* based on morphology matching the type specimens.

Clade	Lineage/ Species	Species status	Alignment with current taxonomy
1	A	Species	NA
1	B	Species	NA
1	C	Species	*O. imbricatum* subsp. *nervatum* (DC.) Norl. and *O. imbricatum* subsp. *nervatum* var. *helichrysoides* (DC.) Norl.
1	D	Species complex	NA
2	E	Species	NA
2	F	Species with three varieties: "corymbosum", "romanskraal" and "burttianum"	*O. corymbosum* L. ("corymbosum" and "romanskraal" varieties) and *O. burttianum* B.Nord ("burttianum" variety)
3	G	Species	*O. imbricatum* var. *microcephalum* Norl.
3	H	Species (cryptic to Species J)	NA
3	I	Species	NA
3	J	Species	*O. imbricatum* L.
4	K	Species	NA
4	L	Suspected hybrid	NA
4	M	Species (cryptic to Species O)	NA
4	N	Species	*O. australe* B.Nord
4	O	Species complex	*O. polygaloides* L.
4	P	Species	*O. rotundifolium* (DC.) Norl. and *O. polygaloides var. latifolium* Norl.

Conclusion

This study draws on evidence from genetic, morphological, and ecological data in an integrative taxonomic approach to refine species delimitation in *O.* sect. *Polygalinae*. Although there are difficulties in determining the species status of two entities (Lineages H and L), I propose that at least 14 and up to 16 species be recognised within *O.* section *Polygalinae* (Table 5), an increase from the ten taxa (six species) currently recognised. This work represents a significant contribution to our understanding of species boundaries and relationships within *O.* section *Polygalinae,* a clade of Cape Asteraceae with a history of unstable alpha-taxonomy. A key challenge in this group is the high levels of homoplasy displayed by most morphological characters. As a consequence, the choice of which characters best reflect phylogenetic relationships has been unavoidably subjective, with Norlindh (1943) choosing to emphasize fruit characters, but these being an unreliable basis on which to identify taxa, resulting in Manning and Goldblatt (2012) suggesting that all elements historically included in *O. imbricatum* and *O. polygaloides* be lumped into a single, heterogeneous taxon. The relationships inferred here using genomic data are based on thousands of independently assorting loci and may therefore be less homoplasious than morphological characters (Givnish & Sytsma, 1997). As such, the patterns revealed by genomic data provide a valuable framework for reassessing morphological variation, and a more objective basis for identifying which morphological characters best reflect phylogenetic relationships. Additionally, using the approach of integrative taxonomy (Dayrat, 2005), including data linked to the ecological role of taxa (Freudenstein et al., 2017), enables the circumscription of biologically meaningful species. This study again demonstrates the utility of genomic data in conjunction with integrative taxonomy for resolving species relationships in taxonomically difficult groups in the Cape flora (e.g. Shaik et al., 2023, 2024; Wootton, Forest & Verboom, 2023). This approach allows the recognition of species that represent evolutionary meaningful units of biodiversity, which should bring stability to their taxonomy and allow them to be appropriately utilised for answering evolutionary questions.

References

Alexander, D.H., Novembre, J. & Lange, K. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*. 19(9):1655–1664. DOI: 10.1101/gr.094052.109.

Baldwin, B.G. & Markos, S. 1998. Phylogenetic utility of the external transcribed spacer (ETS) of 18S–26S rDNA: congruence of ETS and ITS trees of *Calycadenia* (Compositae). *Molecular Phylogenetics and Evolution*. 10(3):449–463. DOI: 10.1006/mpev.1998.0545.

Blanco-Pastor, J.L., Vargas, P. & Pfeil, B.E. 2012. Coalescent simulations reveal hybridization and incomplete lineage sorting in Mediterranean *Linaria*. *PLoS ONE*. 7(6):e39089. DOI: 10.1371/journal.pone.0039089.

Bouckaert, R. & Dong, X. 2017. BEAST2-Dev/substmodels: standard nucleotide substitution models. Zenodo. Available: http://doi.org/10.5281/zenodo.995740.

Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., et al. 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLOS Computational Biology*. 15(4):e1006650. DOI: 10.1371/journal.pcbi.1006650.

Burbrink, F.T. & Ruane, S. 2021. Contemporary philosophy and methods for studying speciation and delimiting species. *Ichthyology & Herpetology*. 3(109):874–894.

Bushnell, B. 2023. BBMap short read aligner, and other bioinformatic tools. Sourceforge. Available: https://sourceforge.net/projects/bbmap/.

de Candolle, A.P. & de Candolle, A. 1837. *Prodromus Systematis Naturalis Regni Vegetabilis*. V. 6. Parisii: Sumptibus Sociorum Treuttel et Würtz, 1824-73.

Carstens, B.C., Pelletier, T.A., Reid, N.M. & Satler, J.D. 2013. How to fail at species delimitation. *Molecular Ecology*. 22(17):4369–4383. DOI: 10.1111/mec.12413.

Chase, M.W., Soltis, D.E., Olmstead, R.G., Morgan, D., Les, D.H., Mishler, B.D., Duvall, M.R., Price, R.A., et al. 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene rbcL. *Annals of the Missouri Botanical Garden*. 80(3):528–580.

Chen, S. 2023. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *iMeta*. 2(e107). Available: https://doi.org/10.1002/imt2.107.

Chernomor, O., Von Haeseler, A. & Minh, B.Q. 2016. Terrace aware data structure for phylogenomic inference from supermatrices. *Systematic Biology*. 65(6):997–1008. DOI: 10.1093/sysbio/syw037.

Cicero, C., Mason, N.A., Jiménez, R.A., Wait, D.R., Wang-Claypool, C.Y. & Bowie, R.C.K. 2021. Integrative taxonomy and geographic sampling underlie successful species delimitation. *Ornithology*. 138(2):1–15. DOI: 10.1093/ornithology/ukab009.

Coates, D.J., Byrne, M. & Moritz, C. 2018. Genetic diversity and conservation units: dealing with the species-population continuum in the age of genomics. *Frontiers in Ecology and Evolution*. 6:165. DOI: 10.3389/fevo.2018.00165.

Cowling, R.M. & Holmes, P.M. 1992. Endemism and speciation in a lowland flora from the Cape Floristic Region. *Biological Journal of the Linnean Society*. 47(4):367–383. DOI: 10.1111/j.1095-8312.1992.tb00675.x.

Cronquist, A. 1981. *An integrated system of classification of flowering plants*. New York: Columbia University Press.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., et al. 2021. Twelve years of SAMtools and BCFtools. *GigaScience*. 10(2):giab008. DOI: 10.1093/gigascience/giab008.

Daru, B.H., Park, D.S., Primack, R.B., Willis, C.G., Barrington, D.S., Whitfeld, T.J.S., Seidler, T.G., Sweeney, P.W., et al. 2018. Widespread sampling biases in herbaria revealed from large-scale digitization. *New Phytologist*. 217:939–955. DOI: 10.1111/nph.14855.

Dayrat, B. 2005. Towards integrative taxonomy. *Biological Journal of the Linnean Society*. 85(3):407–415. DOI: 10.1111/j.1095-8312.2005.00503.x.

Douglas, J., Zhang, R. & Bouckaert, R. 2021. Adaptive dating and fast proposals: Revisiting the phylogenetic relaxed clock model. *PLOS Computational Biology*. 17(2):e1008322. DOI: 10.1371/journal.pcbi.1008322.

Doyle, J.J. 1992. Gene trees and species trees: molecular systematics as one-character taxonomy. *Systematic Botany*. 17(1):144–163.

Doyle, J.J. 1997. Trees within trees: genes and species, molecules and morphology. *Systematic Biology*. 46(3):537–553.

Doyle, J.J. & Doyle, J.L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*. 19(1):11–15.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J. & Rambaut, A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology*. 4(5):e88. DOI: 10.1371/journal.pbio.0040088.

Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*. 32(5):1792–1797. DOI: 10.1093/nar/gkh340.

Edwards, S.V. 2009. Is a new and general theory of molecular systematics emerging? *Evolution*. 63(1):1–19. DOI: 10.1111/j.1558-5646.2008.00549.x.

Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S. & Mitchell, S.E. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*. 6(5):e19379. DOI: 10.1371/journal.pone.0019379.

EMBL-EBI. 2023. *MUSCLE Multiple Sequence Alignment (MSA)*. Available: https://www.ebi.ac.uk/jdispatcher/msa/muscle [2023, November 13].

Engelbrecht, I. 2021. NSCF Georeferencing protocol. Version 1.1. National Science Collections Facility South Africa.

Freudenstein, J.V., Broe, M.B., Folk, R.A. & Sinn, B.T. 2017. Biodiversity and the species concept - lineages are not enough. *Systematic Biology*. 66(4):644–656. DOI: 10.1093/sysbio/syw098.

Frichot, E. & François, O. 2015. LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*. 6(8):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(4):973–983. DOI: 10.1534/genetics.113.160572.

Georges, A., Mijangos, L., Kilian, A., Patel, H., Aitkens, M. & Gruber, B. 2023. Distances and their visualization in studies of spatial-temporal genetic variation using single nucleotide polymorphisms (SNPs). *bioRxiv*. DOI: 10.1101/2023.03.22.533737.

Gernhard, T. 2008. The conditioned reconstructed process. *Journal of Theoretical Biology*. 253(4):769–778. DOI: 10.1016/j.jtbi.2008.04.005.

Givnish, T.J. & Sytsma, K.J. 1997. Homoplasy in molecular vs. morphological data: the likelihood of correct phylogenetic inference. In *Molecular evolution and adaptive radiation*. T.J. Givnish & K.J. Sytsma, Eds. Cambridge: Cambridge University Press. 55–101.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W. & Gascuel, O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*. 59(3):307–321. DOI: 10.1093/sysbio/syq010.

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

Harvey, W.H. & Sonder, O.W. 1865. *Flora capensis*. V. 3. Dublin: Hodges, Smith, and Co.

Hoang, D.T., Chernomor, O., Von Haeseler, A., Minh, B.Q. & Vinh, L.S. 2018. UFBoot2: improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*. 35(2):518–522. DOI: 10.1093/molbev/msx281.

Hörandl, E. & Stuessy, T.F. 2010. Paraphyletic groups as natural units of biological classification. *Taxon*. 59(6):1641–1653. DOI: 10.1002/tax.596001.

Inglis, P.W., Marilia de Castro, R.P., Resende, L. V. & Grattapaglia, D. 2018. Fast and inexpensive protocols for consistent extraction of high quality DNA and RNA from challenging plant and fungal samples for high-throughput SNP genotyping and sequencing applications. *PLoS ONE*. 13(10):1–14. DOI: 10.1371/journal.pone.0206085.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., Von Haeseler, A. & Jermiin, L.S. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*. 14(6):587–589. DOI: 10.1038/nmeth.4285.

Knaus, B.J. & Grünwald, N.J. 2017. VCFR: a package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources*. 17(1):44–53. DOI: 10.1111/1755-0998.12549.

Knowles, L.L. & Carstens, B.C. 2007. Delimiting species without monophyletic gene trees. *Systematic Biology*. 56(6):887–895.

Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A. & Mayrose, I. 2015. Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*. 15(5):1179–1191. DOI: 10.1111/1755-0998.12387.

Križman, M., Jakše, J., Baričevič, D., Javornik, B. & Prošek, M. 2006. Robust CTAB-activated charcoal protocol for plant DNA extraction. *Acta agriculturae Slovenica*. 87(2):427–433. DOI: 10.14720/aas.2006.87.2.15122.

Leaché, A.D. & Oaks, J.R. 2017. The utility of single nucleotide polymorphism (SNP) data in phylogenetics. *Annual Review of Ecology, Evolution, and Systematics*. 48(1):69–84. DOI: 10.1146/annurev-ecolsys-110316-022645.

Linan, A.G., Lowry, P.P., Miller, A.J., Schatz, G.E., Sevathian, J. & Edwards, C.E. 2021. RAD-sequencing reveals patterns of diversification and hybridization, and the accumulation of reproductive isolation in a clade of partially sympatric, tropical island trees. *Molecular Ecology*. 30(18):4520–4537. DOI: 10.1111/mec.15736.

Manning, J.C. & Goldblatt, P. 2012. *Plants of the Greater Cape Floristic Region: the core Cape flora*. (Strelitzia no. 29). Pretoria: South African National Biodiversity Institute.

Manning, J.C., Goldblatt, P. & Helme, N.A. 2012. *Osteospermum norlindhianum* and *O. nordenstamii,* two new species of *Osteospermum* sect. Trefenestratae (Calenduleae) from the Greater Cape Floristic Region. *Bothalia*. 41(1):59–66.

Markos, S. & Baldwin, B.G. 2001. Higher-level relationships and major lineages of *Lessingia* (Compositae, Astereae) based on nuclear rDNA internal and external transcribed spacer (ITS and ETS) sequences. *Systematic Botany*. 26(1):168–183.

Mayden, R.L. 1997. A hierarchy of species concepts: the denouement in the saga of the species problem. In *Species: The units of diversity*. Chapman & Hall. 381–423.

Mayden, R.L. 1999. Consilience and a hierarchy of species concepts: advances toward closure on the species puzzle. *Journal of Nematology*. 31(2):95–116.

McCormack, J.E., Hird, S.M., Zellmer, A.J., Carstens, B.C. & Brumfield, R.T. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution*. 66:526–538. DOI: 10.1016/j.ympev.2011.12.007.

McVean, G. 2009. A genealogical interpretation of principal components analysis. *PLoS Genetics*. 5(10):e1000686. DOI: 10.1371/journal.pgen.1000686.

Mijangos, J.L., Gruber, B., Berry, O., Pacioni, C. & Georges, A. 2022. *DARTR* v2: An accessible genetic analysis platform for conservation, ecology and agriculture. *Methods in Ecology and Evolution*. 13(10):2150–2158. DOI: 10.1111/2041-210X.13918.

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., Von Haeseler, A. & Lanfear, R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*. 37(5):1530–1534. DOI: 10.1093/molbev/msaa015.

Mo, Y.K., Lanfear, R., Hahn, M.W. & Minh, B.Q. 2023. Updated site concordance factors minimize effects of homoplasy and taxon sampling. *Bioinformatics*. 39(1):btac741. DOI: 10.1093/bioinformatics/btac741.

Mucina, L. & Rutherford, M.C. 2006. *The vegetation of South Africa, Lesotho and Swaziland*. (Strelitzia no. 19). Pretoria: South African National Biodiversity Institute.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B. & Kent, J. 2000. Biodiversity hotspots for conservation priorities. *Nature*. 403(6772):853–858.

Nguyen, L.-T., Schmidt, H.A., Von Haeseler, A. & Minh, B.Q. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*. 32(1):268–274. DOI: 10.1093/molbev/msu300.

Nordenstam, B. 2004. Two new species of *Osteospermum* (Compositae–Calenduleae) from southwestern Cape Province, South Africa. *Edinburgh Journal of Botany*. 60(3):259–265. DOI: 10.1017/S0960428603000234.

Nordenstam, B. & Källersjö, M. 2009. Calenduleae. In *Systematics, Evolution and Biogeography of Compositae.* V.A. Funk, A. Susanna, T.F. Stuessy, & R.J. Bayer, Eds. Vienna: International Association for Plant Taxonomy. 527–538.

Norlindh, T. 1943. *Studies in the Calenduleae I: Monograph of the genera Dimorphotheca, Castalis, Osteospermum, Gibbaria and Chrysanthemoides*. Lund: C.W.K. Gleerup.

Norlindh, T. 1960. Additions to the monograph on *Osteospermum*. *Botaniska Notiser*. 4(113):385–399.

Ortiz, E.M. 2019. vcf2phylip v2.0: convert a VCF matrix into several matrix formats for phylogenetic analysis. v2.0. Zenodo. Available: DOI:10.5281/zenodo.2540861.

Padial, J.M., Miralles, A., De La Riva, I. & Vences, M. 2010. The integrative future of taxonomy. *Frontiers in Zoology*. 7(16). DOI: 10.1186/1742-9994-7-16.

Parker, T. 2019. An integrative taxonomic approach using genome-wide SNP data resolves variation in the *Dimorphotheca pluvialis-sinuata* species complex (Calenduleae: Asteraceae) in South Africa. Honours thesis. University of Cape Town.

Parks, M., Cronn, R. & Liston, A. 2009. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. *BMC Biology*. 7(1):84. DOI: 10.1186/1741-7007-7-84.

Patterson, N., Price, A.L. & Reich, D. 2006. Population structure and eigenanalysis. *PLoS Genetics*. 2(12):2074–2093. DOI: 10.1371/journal.pgen.0020190.

Pebesma, E. 2018. Simple features for R: standardized support for spatial vector data. *The R Journal*. 10(1):439–446. DOI: 10.32614/RJ-2018-009.

Pirie, M.D., Humphreys, A.M., Galley, C., Barker, N.P., Verboom, G.A., Orlovich, D., Draffin, S.J., Lloyd, K., et al. 2008. A novel supermatrix approach improves resolution of phylogenetic relationships in a comprehensive sample of danthonioid grasses. *Molecular Phylogenetics and Evolution*. 48(3):1106–1119. DOI: 10.1016/j.ympev.2008.05.030.

Pirie, M.D., Humphreys, A.M., Barker, N.P. & Linder, H.P. 2009. Reticulation, data combination, and inferring evolutionary history: an example from Danthonioideae (Poaceae). *Systematic Biology*. 58(6):612–628. DOI: 10.1093/sysbio/syp068.

Porghahreman, F., Torabi Giglou, M., Farhadpour, M., Ghorbani, A. & Hajieghrari, B. 2023. *Calendula officinalis* isolate TN-82-766, whole genome shotgun sequencing project. National Center for Biotechnology Information. Available: https://www.ncbi.nlm.nih.gov/nuccore/JARQZI000000000.1.

Pritchard, J.K., Stephens, M. & Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics*. 155(2):945–959. DOI: 10.1093/genetics/155.2.945.

Puechmaille, S.J. 2016. The program STRUCTURE does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. *Molecular Ecology Resources*. 16(3):608–627. DOI: 10.1111/1755-0998.12512.

Quattrini, A.M., Wu, T., Soong, K., Jeng, M.-S., Benayahu, Y. & McFadden, C.S. 2019. A next generation approach to species delimitation reveals the role of hybridization in a cryptic species complex of corals. *BMC Evolutionary Biology*. 19(1):116. DOI: 10.1186/s12862-019-1427-y.

de Queiroz, K. 1998. The general lineage concept of species, species criteria, and the process of speciation: a conceptual unification and terminological recommendations. In *Endless Forms: Species and Speciation*. D.J. Howard & S.H. Berlocher, Eds. New York: Oxford University Press. 57–75.

de Queiroz, K. 2005. A unified concept of species and its consequences for the future of taxonomy. *Proceedings of the California Academy of Sciences*. 56(1):196–215.

de Queiroz, K. 2007. Species concepts and species delimitation. *Systematic Biology*. 56(6):879–886. DOI: 10.1080/10635150701701083.

R Core Team. 2023. Version 4.3.2. Available: https://www.R-project.org/.

Rambaut, A. 2018. FigTree. Version 1.4.4. Available: http://github.com/rambaut/figtree/.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G. & Suchard, M.A. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*. 67(5):901–904. DOI: 10.1093/sysbio/syy032.

Rieseberg, L.H. & Brouillet, L. 1994. Are many plant species paraphyletic? *Taxon*. 43(1):21–32. DOI: 10.2307/1223457.

Rouhan, G. & Gaudeul, M. 2014. Plant taxonomy: a historical perspective, current challenges, and perspectives. In *Molecular Plant taxonomy: Methods and Protocols*. P. Besse, Ed. (Methods in Molecular Biology no. 1115). New York: Springer. 1–37.

le Roux, M.M. & Victor, J.E. 2017. A priority list of genera that are in need of taxonomic revision in South Africa. Available: https://www.sanbi.org/documents/a-priority-list-of-genera-that-are-in-need-of-taxonomic-revision-in-south-africa/ [2021, September 28].

Sadler, R.T. 2024. Phylogeny, generic limits and evolution of dispersal traits in the Asteraceae tribe Calenduleae. MSc thesis. University of Cape Town.

Sadler, R., Parker, T., Verboom, G.A., Ellis, A.G., Jackson, M., Van Zyl, J., Manning, J. & Bergh, N.G. 2022. A molecular phylogeny of Calenduleae (Asteraceae) supports the transfer of *Dimorphotheca polyptera* DC. to *Osteospermum* L. *South African Journal of Botany*. 151:234–245. DOI: 10.1016/j.sajb.2022.09.001.

Schlick-Steiner, B.C., Steiner, F.M., Seifert, B., Stauffer, C., Christian, E. & Crozier, R.H. 2010. Integrative taxonomy: a multisource approach to exploring biodiversity. *Annual Review of Entomology*. 55(1):421–438. DOI: 10.1146/annurev-ento-112408-085432.

Shaffer, H.B. & Thomson, R.C. 2007. Delimiting species in recent radiations. *Systematic Biology*. 56(6):896–906. DOI: 10.1080/10635150701772563.

Shaik, Z., Bergh, N.G., Oxelman, B. & Verboom, G.A. 2023. Genomic, phenotypic, and ecological data provide decisive support for multiple species in the *Seriphium plumosum* complex (Gnaphalieae: Asteraceae). *Bulletin of the Society of Systematic Biologists*. 2(2):1–23. DOI: 10.18061/bssb.v2i2.9000.

Shaik, Z., Verboom, G.A., Oxelman, B. & Bergh, N.G. 2024. Revision of *Seriphium plumosum* (Gnaphalieae: Asteraceae) in southern Africa with description of new species in *Seriphium*. *South African Journal of Botany*. 165:367–383. DOI: 10.1016/j.sajb.2023.12.046.

Simpson, G.G. 1961. *Principles of animal taxonomy*. New York: Columbia University Press.

Sites, J.W. & Marshall, J.C. 2004. Operational criteria for delimiting species. *Annual Review of Ecology, Evolution, and Systematics*. 35(1):199–227. DOI: 10.1146/annurev.ecolsys.35.112202.130128.

Smith, J.F., Clark, J.L., Amaya-Márquez, M. & Marín-Gómez, O.H. 2017. Resolving incongruence: Species of hybrid origin in *Columnea* (Gesneriaceae). *Molecular Phylogenetics and Evolution*. 106:228–240. DOI: 10.1016/j.ympev.2016.10.001.

South African National Biodiversity Institute. 2009. The Vegetation Map of South Africa, Lesotho and Swaziland [vector geospatial dataset]. Available: http://bgis.sanbi.org/SpatialDataset/Detail/331.

Stephens, M. 2000. Dealing with label switching in mixture models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*. 62(4):795–809. DOI: 10.1111/1467-9868.00265.

Swanepoel, W., Cauwer, V.D. & Wyk, A.E.V. 2021. A new species of *Osteospermum* subgen. *Tripteris* (Asteraceae: Calenduleae) from the Namib Desert, Namibia. *Phytotaxa*. 487(3):185–194. DOI: 10.11646/phytotaxa.487.3.1.

Swofford, D.L. 2002. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Tamura, K. & Nei, M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*. 10(3):512–526. DOI: 10.1093/oxfordjournals.molbev.a040023.

The Angiosperm Phylogeny Group. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society*. 181(1):1–20. DOI: 10.1111/boj.12385.

Theirs, B.M. 2024. *Index Herbariorum*. Available: https://sweetgum.nybg.org/science/ih/ [2024, July 25].

Tonkin-Hill, G. & Lee, S. 2024. starmie: population structure model inference and visualisation. Available: https://github.com/sa-lee/starmie.

Vences, M., Miralles, A. & Dufresnes, C. 2024. Next-generation species delimitation and taxonomy: Implications for biogeography. *Journal of Biogeography*. 00:1–14. DOI: 10.1111/jbi.14807.

Victor, J.E., Hamer, M. & Smith, G.F. 2013. A biosystematics research strategy for the algae, animals, bacteria and archaea, fungi and plants of South Africa 2013–2018. In *SANBI Biodiversity Series 23*. Pretoria: South African National Biodiversity Institute.

White, T.J., Bruns, T., Lee, S. & Taylor, J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In *PCR protocols: a guide to methods and applications*.

M.A. Innis, D.H. Gelfand, J.J. Sninsky, & T.J. White, Eds. Academic Press. 315–322. DOI: 10.1016/B978-0-12-372180-8.50042-1.

Wiens, J.J. 2004. What is speciation and how should we study it? *The American Naturalist*. 163(6):914–923. DOI: 10.1086/386552.

Wiens, J.J. 2007. Species delimitation: new approaches for discovering diversity. *Systematic Biology*. 56(6):875–878. DOI: 10.1080/10635150701748506.

Wiley, E.O. 1978. The evolutionary species concept reconsidered. *Society of Systematic Biologists*. 27(1):17–26.

Will, K.W., Mishler, B.D. & Wheeler, Q.D. 2005. The perils of DNA barcoding and the need for integrative taxonomy. *Systematic Biology*. 54(5):844–851. DOI: 10.1080/10635150500354878.

Willis, C.K., Cowling, R.M. & Lombard, A.T. 1996. Patterns of endemism in the limestone flora of South African lowland fynbos. *Biodiversity and Conservation*. 5(1):55–73. DOI: 10.1007/BF00056292.

Wood, A.R. & Nordenstam, B. 2003. An interesting new species of *Osteospermum* (Asteraceae-Calenduleae) from the Western Cape Province, South Africa, providing a link to the genus *Chrysanthemoides*. *South African Journal of Botany*. 69(4):572–578. DOI: 10.1016/S0254-6299(15)30297-0.

Wootton, L.M., Forest, F. & Verboom, G.A. 2023. Consilience across multiple, independent genomic data sets reveals species in a complex with limited phenotypic variation. *Systematic Biology*. 72(4):753–766. DOI: 10.1093/sysbio/syad024.

Yeates, D.K., Seago, A., Nelson, L., Cameron, S.L., Joseph, L. & Trueman, J.W.H. 2011. Integrative taxonomy, or iterative taxonomy? *Systematic Entomology*. 36(2):209–217. DOI: 10.1111/j.1365-3113.2010.00558.x.

Yi, X. & Latch, E.K. 2022. Nonrandom missing data can bias Principal Component Analysis inference of population genetic structure. *Molecular Ecology Resources*. 22(2):602–611. DOI: 10.1111/1755-0998.13498.

Yi, H., Dong, S., Yang, L., Wang, J., Kidner, C. & Kang, M. 2023. Genome-wide data reveal cryptic diversity and hybridization in a group of tree ferns. *Molecular Phylogenetics and Evolution*. 184:107801. DOI: 10.1016/j.ympev.2023.107801.

Zhang, R. & Drummond, A. 2020. Improving the performance of Bayesian phylogenetic inference under relaxed clock models. *BMC Evolutionary Biology*. 20(1):54. DOI: 10.1186/s12862-020-01609-4.

Chapter 3: The evolution of myrmecochory and its impact on the spatial scale of gene flow in *Osteospermum* section *Polygalinae*

Introduction

Myrmecochory, the transport of seeds or fruits by ants, is a short-distance dispersal mechanism associated with the directed movement of diaspores to ant nests. This mutualism has evolved repeatedly, occurring in more than 11 000 species (4.5%), 334 genera (2.5%), and 77 families (17%) of angiosperms (Lengyel et al., 2010). Although cosmopolitan in distribution, myrmecochory is most prevalent in summer-dry sclerophyll shrublands growing on nutrient-poor soils, particularly soils that are low in phosphorous (Berg, 1975; Milewski & Bond, 1982; Westoby et al., 1982). As such, it is especially prevalent in the kwongan of Western Australian (ca. 1500 species) and the fynbos of the South African Cape (ca. 1300 species), but less common in the Californian chaparral and Mediterranean shrublands, which have similar climates and fire regimes but more fertile soils (Berg, 1975; Milewski & Bond, 1982; Westoby et al., 1982; Bond & Slingsby, 1983; Lengyel et al., 2010; Stock & Verboom, 2012). In the sclerophyll shrublands of the Cape, myrmecochory is most prevalent in fynbos vegetation, which is associated with leached, quartzite-derived sands, compared to renosterveld or succulent karoo vegetation which occurs on more fertile, shale or granite-derived clays and loams (Milewski & Bond, 1982; Cowling et al., 1994). While it has been suggested that ant dispersal enhances seedling establishment in nutrient-poor environments by transporting seeds to nutrient-enriched ant nests (Milewski & Bond, 1982; Bond & Slingsby, 1983), experiments have shown that this advantage is situation-specific, and thus not alone able to explain the evolution of myrmecochory (Rice & Westoby, 1986; Manzaneda & Rey, 2012). Instead, the selective advantage of myrmecochory may be the protection from fire, pathogens and seed predators, such as rodents and granivorous birds, provided by underground deposition of diaspores in ant nests, as threats to diaspores on the soil surface are cosmopolitan (O'Dowd & Hay, 1980; van der Pijl, 1982; Bond & Slingsby, 1983; Bond & Breytenbach, 1985; Slingsby & Bond, 1985; Manzaneda, Fedriani & Rey, 2005).

In sclerophyll shrublands in Australia and the Cape, most recruitment is restricted to a post-fire window period and seeds must therefore be able to survive the interval between successive fires, which is typically 10 - 30 years in fynbos and 10 - 40 years in kwongan (le Maitre & Midgley, 1992; Holmes & Newton, 2004; Keeley et al., 2012; Kraaij & van Wilgen, 2014; Miller & Dixon, 2014), so refuge in ant nests may be important for seed survival. The strong post-fire recruitment pattern observed in fynbos and kwongan is partly a function of nutrient limitation in these systems, which experience a short period of increased nutrient availability directly after fire (le Maitre & Midgley, 1992; Keeley et al., 2012; Miller & Dixon, 2014; Verboom, Slingsby & Cramer, 2024). Furthermore,

mechanisms to enhance seed survival may be particularly important in low nutrient habitats where seed production is costly, with greater seed survival allowing plants to produce fewer seeds (Pierce & Cowling, 1991; Cowling et al., 1994). In the Cape, experiments have shown high seed losses due to rodent predation, impacting both obligate reseeders, which make up the majority of myrmecochores, and resprouters such as resprouting Restionaceae and Cyperaceae (Bond, 1984; Bond & Slingsby, 1984; Bond & Breytenbach, 1985; Cowling et al., 1994; Christian & Stanton, 2004; van Blerk, West & Midgley, 2017). Although rodents are largely viewed as seed predators, there is evidence that some rodents scatter-hoard seeds in underground caches and may successfully disperse seeds and increase their survival (Midgley et al., 2002; Weighill, Huysamer & Anderson, 2017; White, Bronner & Midgley, 2017), but the frequency and species-specificity of this interaction requires further study.

The possession of an elaiosome, a structure that is of nutritional value to ants but not of use for seed germination or seedling growth, has been identified as the key functional trait of myrmecochorous diaspores (Sernander, 1906; Westoby et al., 1982; Lengyel et al., 2010). Elaiosomes are most commonly present as lipid-rich appendages, which attract ants that then carry the seeds to their nests before consuming the elaiosomes and discarding the seeds (Berg, 1975; van der Pijl, 1982). Myrmecochorous diaspores tend to be hard, heavy and have a smooth surface (Berg, 1975; Slingsby & Bond, 1981). This makes them difficult for ants to grip once the elaiosome has been removed, ensuring that the majority of seeds remain underground in the ant nest and few are redeposited on the surface in discard piles (Berg, 1975; Slingsby & Bond, 1981). Elaiosomes have evolved independently in at least 101 lineages, representing one of the most extensive examples of convergent evolution relating to a mutualism (Lengyel et al., 2009, 2010). Elaiosomes have developed from a variety of tissues and may even have different origins in closely related species (Speta, 1972; van der Pijl, 1982). Ants respond very quickly (often within a few seconds) to the attractants of a fallen seed, leading to quick burial (Berg, 1975; Slingsby & Bond, 1981; van der Pijl, 1982). This quick response is triggered by detection of unsaturated fatty acids released from the elaiosome (Marshall, Beattie & Bollenbacher, 1979; Brew, O'Dowd & Rae, 1989; Midgley & Bond, 1995; Pfeiffer, Huttenlocher & Ayasse, 2010).

The ant species recruited by plants to disperse seeds are usually wide-spread and generalist scavengers, with a few ant species serving many plant species, so myrmecochory is not a co-evolved mutualism (O'Dowd & Hay, 1980; Slingsby & Bond, 1981; Westoby et al., 1982; Johnson, 1992; Levine et al., 2019). For myrmecochory to be successful, the recruited ants must be species that nest underground and carry the seed into the ant nest before consuming the elaiosome (Christian & Stanton, 2004). Unlike plant species richness, ant species richness is not exceptionally high in the Fynbos Biome of the Cape (Braschler, Chown & Gaston, 2012). In the Cape the most reliable ant species for myrmecochory are *Anoplolepis custodiens* (F. Smith) and *A. steingroeveri* (Forel), with *Pheidole capensis* (Mayr), *Camponotus niveosetosus* (Mayr) and *Tetramorium sericeiventre* (Emery)

(formerly T. *quadrispinosum*) also frequently reported (Slingsby & Bond, 1981, 1985; Bond & Slingsby, 1983, 1984; Botes et al., 2006). All these ant species have widespread distributions in southern Africa, occurring in multiple biomes and are not restricted to fynbos (Slingsby, 2017). In some areas of the Cape, native ants have been replaced by the invasive Argentine ant, *Linepithema humile* (Mayr) (formerly *Iridomyrmex humilis*) (Slingsby, 2017), and experiments have shown reduced dispersal and recruitment of myrmecochorous species in invaded areas (Bond & Slingsby, 1984; Christian, 2001). This is a conservation concern, especially when myrmecochorous species are local endemics, which is the case for many species in the Cape flora (Bond & Slingsby, 1984; Cowling et al., 1994; McDonald et al., 1995; Goldblatt & Manning, 2002).

The evolution of myrmecochory is believed to influence the rate of speciation and diversification of lineages since ants typically move seeds short distances, which is thought to limit gene flow between populations more than other dispersal strategies (Bond & Slingsby, 1983; Slingsby & Bond, 1985; Johnson, 1992; Lengyel et al., 2009). The global mean dispersal distance of ant dispersed seeds is 0.96 m (range 0.01-77 m, n = 2524) with a dispersal curve typically showing a peak at short distances and a long tail (Gómez & Espadaler, 1998). In contrast, seeds dispersed by wind can be carried hundreds of meters (Johnson, 1992) and bird dispersed seeds even further, peaking at around 1 km but having the potential to be dispersed up to 400 km (Mokotjomela, Musil & Esler, 2013). In plants, gene flow is controlled by both the dispersal of seeds and pollen, and dispersal distance therefore controls the connectivity of populations and the spatial genetic cohesion of a species by determining whether there is panmictic mating across the species (Wright, 1943, 1946; Johnson, 1992; Ronce, 2007; Orsini et al., 2013; Cruzan & Hendrickson, 2020). Limited gene flow between populations can promote local adaptation and genetic drift which can lead to speciation through divergent selection and geographic speciation by isolation (Lenormand, 2002; Linder, 2003; Givnish, 2010; Orsini et al., 2013). When pollen dispersal is limited more than seed dispersal, for example due to short flight distances of insect pollinators, the spatial scale of gene flow will be predominantly determined by seed dispersal distances (Johnson, 1992; Cruzan & Hendrickson, 2020). Seed dispersal distance also dictates a species ability to colonise new areas or reestablish after local extinction, impacting geographic distribution and demographics (van der Pijl, 1982; Ronce, 2007). Unlike other dispersal modes, myrmecochory provides the safe storage of seeds in ant nests which may increase seed survival and possibly germination success and lead to increased fitness and reduced rates of extinction in myrmecochrous lineages (Bond & Slingsby, 1983; Bond & Breytenbach, 1985; Christian & Stanton, 2004; Lengyel et al., 2009).

Both the South African Cape and Western Australian floras are extraordinarily species rich with very high endemism and since these are also the areas where myrmecochory is most prevalent, it has been hypothesised that myrmecochory may be associated with increased rates of diversification in these regions (Bond & Slingsby, 1983; Cowling, Holmes & Rebelo, 1992; Givnish, 2010; Mucina & Majer,

2012). In a worldwide comparison of myrmecochorous and non-myrmecochrous (passive, water, wind, exozoochory or endozoochory) sister lineages, Lengyel et al (2009) found that on average, myrmecochorous lineages contained over twice the number of species compared to their non-myrmecochorous sister lineages. This pattern was particularly strong in the Palearctic, Australian, and Holarctic regions. In the Paleotropical zone, however, which was largely represented by the South Africa Cape, myrmecochorous lineages did not contain significantly more species than their sister clades, which may suggest that other factors and processes are influencing rates of speciation in this region. In contrast, Cowling et al. (1994) found myrmecochorous genera to be more species rich in a comparison of 31 myrmecochorous and 76 non-myrmecochorous genera in the Cape fynbos. Cowling et al. (1994) suggested that diversification in myrmecochorous lineages in the fynbos was due to them being obligate reseeders producing few, large seeds and having weakly persistent seedbanks (as a consequence of growing in low-nutrient soils), which makes them more vulnerable to fire-induced population reduction and local extinction. They argue that this results in myrmecochorous species occurring in small, isolated populations with gene flow reduced by short dispersal distances, which would make populations more likely to diverge, especially in the very heterogeneous environments of the Cape which promote ecological diversification and favour local endemism (Linder, 1985; Cowling & Holmes, 1992). However, a factor which has not been considered in explaining diversification in this system is that the most common dispersal mode is passive dispersal (Moll & McKenzie, 1994) which is perhaps expected to have similar short dispersal distances. Nonetheless, in a study examining the most important factors predicting local endemism in the southern Langeberg (a mountain range in the CFR) using a logistic model based on growth form, post-fire regeneration strategy and dispersal mode, dispersal mode was found to be the most important factor determining endemism, with the likelihood of endemism of an ant-dispersed species being 1.72 times that of passively dispersed species and 4 times that of wind dispersed species (McDonald et al., 1995). It therefore remains unclear whether myrmecochory is truly driving diversification in the Cape and more empirical studies are needed to test how different dispersal modes impact gene flow and speciation.

The genus *Osteospermum* (Asteraceae; Calenduleae) provides an excellent system for the study of dispersal biology and its effect on diversification because species within the genus exhibit a large range of fruit forms, including fleshy drupes; thin-walled fruits with large papery wings; and hard nutlet-like fruits with surface corrugations, spines, or ridges and some bearing fatty appendages, suggesting a large variety of dispersal mechanisms (Norlindh, 1943; Wood & Nordenstam, 2003; Sadler, 2024). The Core-*Osteospermum* clade is comprised primarily of species with nutlet-like fruits (Sadler, 2024), some of which are putatively myrmecochorous, although the only species specifically mentioned by Bond & Slingsby (1983) is *O. asperulum* (DC.) Norl. Wood & Nordenstam (2003) have, however, suggested that *O. corymbosum* L. (a member of *Osteospermum* section *Polygalinae*), may be myrmecochorous based on the presence of a pale coloured, fleshy, basal appendage derived

from the pericarp on the fruits, which they interpreted as an elaiosome (Figure 1). Studies on the adaptive significance of elaiosome-like structures in the Cape flora are few and have mainly focused on species in the Proteaceae (e.g., Bond & Breytenbach, 1985; Slingsby & Bond, 1985; Christian & Stanton, 2004). No known studies have investigated any of the six Asteraceae genera in the Cape inferred to have myrmecochorous species (Bond & Slingsby, 1983). The effectiveness of the elaiosome-like structure in *O. corymbosum* as an adaptation for myrmecochory, and whether the other species in *O.* section *Polygalinae* possess elaiosome-like structures is therefore unknown.

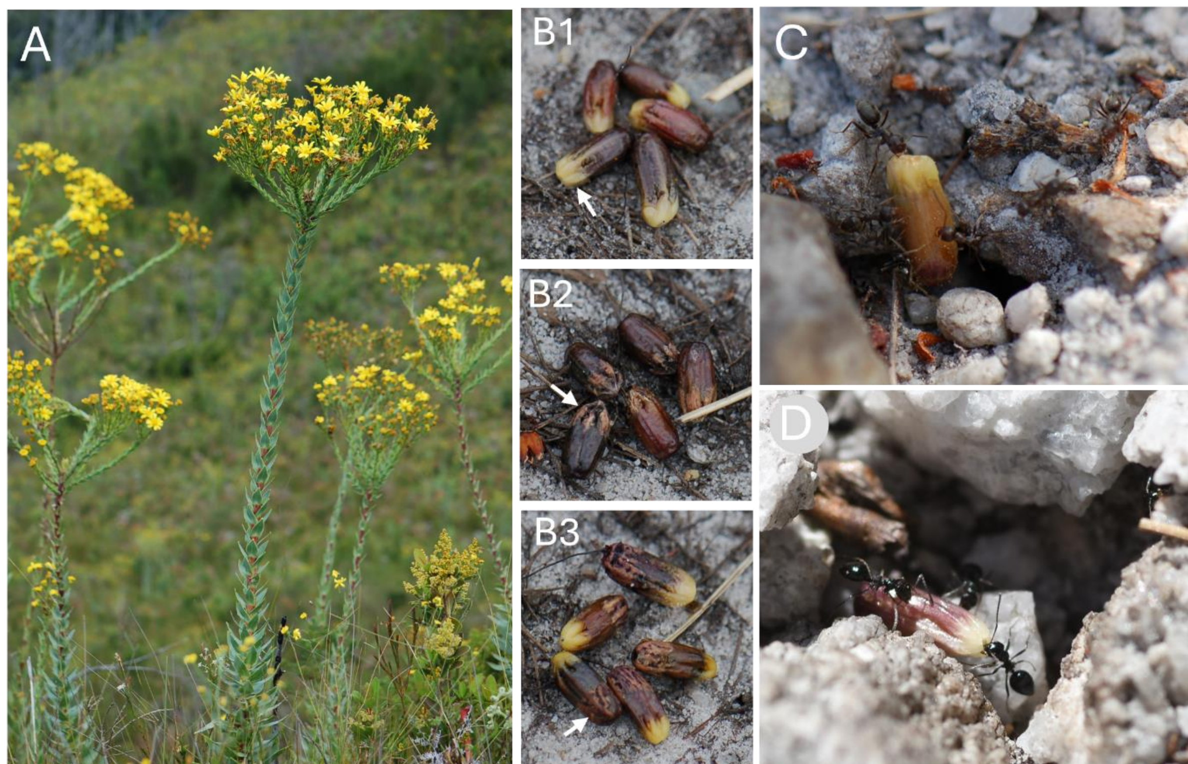


Figure 1. (A) *Osteospermum corymbosum* stand at Grootvadersbosch Nature Reserve, Langeberg, South Africa. (B) The three fruit treatments used in cafeteria experiments to test the functional significance of elaiosomes for dispersal by ants. B1: Unmanipulated fruit, arrow indicating elaiosome; B2: Fruit with elaiosome removed, arrow indicating area of elaiosome removal; B3: Control, fruit with elaiosome but damaged at the opposite end, arrow indicating damage. (C) Ants transporting a fruit of *O. corymbosym* into the entrance of an ant nest. (D) Ants transporting a fruit of *O. burttianum*. All photos were taken by L.M.C. James.

In this chapter I first test the hypothesis that the fruits of *O. corymbosum* are adapted for dispersal by ants and investigate whether other species in *O.* sect. *Polygalinae* share these traits. To do so, I employ the classic protocol of Baum & Larson (1991) using ancestral character reconstructions to trace the evolution of key fruit characters, including the presence of elaiosome-like structures and a smooth, unridged surface texture, in *O.* section *Polygalinae*, and comparative field experiments to test

the efficacy of these structures in promoting fruit dispersal by ants. I then test the hypothesis that the evolution of myrmecochory enhances population differentiation through its effect on the spatial scale of gene flow. Interpopulation gene flow can be studied by assessing the correlation of genetic distance with geographic distance (isolation-by-distance, IBD), with no relationship expected when there is frequent gene flow between populations resulting in panmictic mating within the species, and a strong relationship of IBD expected when gene flow is present but limited (Wright, 1943, 1946; Orsini et al., 2013). I determine IBD patterns in seven species in *O.* section *Polygalinae* using single nucleotide polymorphism (SNP) data to calculate intraspecific genetic distances between populations, and the geographic distances between them. Interspecific IBD patterns are then compared to test whether the spatial scale of gene flow is lower in myrmecochorous species, as indicated by a steeper IBD slope.

Methods

Evolution of fruit traits in Osteospermum section Polygalinae

Fruits were collected in the field from 50 populations representing 13 out of 15 putative species in *O.* sect. *Polygalinae* (see Chapter 2) over the spring-autumn periods of 2021-2023. This represented all the species except Species H and M, for which viable, mature fruits were unavailable. Fruit morphology for each species was characterised based on three discrete traits (presence of an elaiosome, surface texture, presence of ridges with membranous edges) and two continuous traits (mass and width) that could influence dispersal. Discrete traits were assessed for all populations while continuous traits were measured from one representative population per species, except for two morphologically variable species complexes (Species D and O; Chapter 2), for which measurements were taken from one population of each morphologically distinct population group for which fruits were available, and the overall mean calculated across these (i.e., Species D:, IMB19, IMB21, IMB26, IMB28; Species O: POL21, POL2, POL4, POL21).

Fruits were examined in the field while fresh for the presence of an appendage that could be an elaiosome. Fruits from *O. australe* (Species N) were already dry upon collection, so could not be assessed for elaiosome presence. Fruits were then either preserved in FAA (60% ethanol, 25% distilled water, 10% formalin, 5% glacial acetic acid) solution or allowed to dry in brown paper packets before assessment of other morphological traits in the laboratory. To characterise sculpturing of the fruit surfaces, fruits were examined using a 10x hand lens and two characters were scored: firstly, whether the fruit surface was smooth or pitted between longitudinal ridges; and secondly, whether or not the primary ridges (if present) were sharply acute, with membranous edges. Fruit characters within each species were similar for these discrete traits (see Chapter 2), so a single score was assigned to each species. Fruit mass and width was measured for three mature dried fruits per representative population and the mean calculated. The mass of each fruit was measured using an

electronic balance precise to 0.01 mg (Shimadzu Corporation Japan, AUW220D). Fruit width at the midpoint of the length was measured using digital callipers precise to 0.01 mm.

Ancestral character state estimation was performed using the “ace” function from the ape package version 5.7.1 (Paradis & Schliep, 2019) in R (R Core Team, 2023) for the three discrete and two continuous traits. For this purpose, the maximum likelihood tree inferred from concatenated GBS data in Chapter 2 was pruned to contain only one tip per species. This tree was then made ultrametric using the “chronos” function in the ape package using default parameters and with no calibration points specified. For the discrete traits ancestral characters were estimated by maximum likelihood, with default settings and a binary state model with equal transition rates (Mk1). For the continuous characters a Brownian motion model was used and fitted by residual maximum likelihood (REML), which are the recommended default settings (Paradis & Schliep, 2019). Only taxa with known trait states were included in the ancestral character state estimation.

Experimental tests of myrmecochory

To determine whether *O. corymbosum* (Species F) fruits are dispersed by ants and possess a true elaiosome, two comparative field experiments were performed between 20 and 25 October 2022 at Grootvadersbosch Nature Reserve on the southern slopes of the central Langeberg mountains in South Africa, where large stands of *O. corymbosum* occur at elevations above 400 m (Figure 1A). While all morphological varieties of Species F possess elaiosome-like structures (see Chapter 2), my experiments focussed on the “corymbosum” variety which corresponds to *O. corymbosum* as currently circumscribed.

Experiment 1 aimed to test whether ants disperse fruits of *O. corymbosum* and which component of the fruit acts to attract ants. Freshly collected fruits were subjected to three treatments: (i) unmanipulated fruit with the pale fleshy appendage intact (unmanipulated); (ii) fruit with the appendage scraped off (elaiosome removed); and (iii) fruit with the pericarp scraped off at the opposite end of the fruit to the fleshy appendage, leaving the appendage intact but mimicking the damage inflicted on the fruit with removal of the appendage (control) (Figure 1B).

Following treatment application, five fruits representing each treatment (15 fruits in total) were placed in piles situated within 70 cm of active ant nests in a cafeteria-style experiment. This experiment was repeated once at each of six sites situated along a mountain track where vegetation was low, permitting access to ant nests and allowing unhindered observations. At each site the time from placing of fruits in piles (start of the experiment) to the time at which each fruit was moved from its original position was recorded. Where possible, fruits were tracked after removal and a record made of whether they were transported into the entrance of an ant nest. For every fruit that could be

recovered at the end of the experiment, the straight-line distance the fruit had been moved was recorded. For fruits that had been taken into ant nests, the distance from the original position to the entrance of the ant nest was measured. The experiment was run for 32 to 45 minutes, except at site 5 where the experiment was run for 73 minutes to allow for comparable dispersal activity due to ants taking longer to initially find the fruits. This may have been due to the site 5 cafeteria experiment being conducted latest in the day (late afternoon – early evening), while ant activity was always greatest during the hottest time of the day (early afternoon). The effects of differences between sites were accounted for in analyses.

The time to removal of fruits between treatments was compared using survival analysis implemented using the "Surv" function from the survival package version 3.5.7 (Therneau, 2023) in R. Survival analysis considers the length of time until occurrence of an event, which here is removal of a fruit, and compares between treatments using a hazards ratio. Survival models also account for the possibility that an event may not be observed within the observation period, resulting in observations for the individuals for which the event has not yet occurred being censored, i.e. the fruit was not removed because the experiment ended. This approach also controls for time length differences between runs of the experiment. A Cox proportional hazards mixed effects model was fitted in which time to removal was the response variable, treatment the fixed effect, and site a random effect using the "coxme" function from the coxme package version 2.2.20 (Therneau, 2024). Estimated marginal means were calculated to enable pairwise comparison of the treatment effects, and contrasts compared, using functions from the eemeans package version 1.10.0 (Lenth, 2024). Due to the complexity of plotting models with random effects, a simple Kaplan-Meier survival model was used to plot the data for visual comparison using the "survfit" function from the survival package. To compensate for between-site variation in the time taken by ants to discover the fruits, the data were adjusted such that the first fruit was removed at time zero. Differences between treatment curves were tested using the "survdiff" function.

The proportion of tracked fruits that were taken into the ant nest was compared between the three treatments using a binomial generalized linear mixed-effects model with treatment as the fixed effect and site as a random effect using the "glmer" function from package lme4 version 1.1.35.1 (Bates et al., 2015). The distances fruits were moved was compared between treatments using a linear mixed-effects model with treatment as the fixed effect and site as a random effect using the "lmer" function from lme4. For both models, functions from the emmeans package were used to calculate pairwise contrasts between treatments. For fruits that were transported from their original positions, the mean distance moved was calculated for each treatment at each site.

Ants at each site were collected in 80% ethanol and identified under a compound microscope using keys in Slingsby (2017). Identifications were confirmed by Peter Slingsby from photographs.

Experiment 2 sought to test whether the removal of *O. corymbosum* fruits by ants differs from that of fruits of its sister species (Species E) which lack elaiosomes. Both Species E and *O. corymbosum* occur in the Langeberg mountains, the former being known only from two populations on the drier northern slopes, and the latter being found mostly on the wetter southern slopes or at high elevations on the northern slopes. For Species E, fruits were collected from the population situated between Garcia's Pass and the farm "Muiskraal" (Appendix 1; population IMB11), and for *O. corymbosum* (Species F), fruits were collected from the population at Grootvadersbosch Nature Reserve, where the experiment was conducted (Appendix 1; population COR4). Mature fruits collected from both species were stored separately in zip-sealed plastic bags in a cold box to maintain freshness until setup of the experiment. Ten markers were placed at two-meter intervals to make a transect through a stand of *O. corymbosum* plants. The ground around the markers was examined and any naturally fallen fruit was removed. Five fruits of each species (10 in total) were then placed at the base of each marker and left for several hours (between 166 and 391 minutes) to allow natural dispersal to take place. The number of fruits of each species that had been moved away from the marker were then recorded. This set-up was repeated four times (40 replicates) in three separate stands of *O. corymbosum* over a period of three days. By using pairwise species comparisons, the setup of the experiment accounted for differences in the spatial and temporal conditions between markers. Differences in the number of fruits removed between species were compared using a Wilcoxon signed rank test for paired samples, implemented using the "wilcox.test" function from the stats package version 4.3.2 in R (R Core Team, 2023).

Between-species variation in the spatial scale of gene flow

The spatial scale of gene flow was quantified using IBD, with Mantel tests (Mantel, 1967; Diniz-Filho et al., 2013) being used to assess significance of both single species IBD patterns and between-species comparison of IBD slope differences. For this purpose, pairwise genetic distances were first calculated between populations (each represented by one individual) for all species for which genetic (SNP) data were available for four or more populations (n = 7) (Chapter 2). Generation of the SNP dataset followed the same steps as in Chapter 2, except that filtering was done separately for each species to remove intraspecific invariant sites, and up to 50% missing data was allowed per site. The number of SNPs for each species were as follows: Species D, 11281; Species F, 12200; Species I, 7356; Species J, 16383; Species K, 4460; Species O, 12199 and Species P, 8064. VCF files were converted to genind objects using the "vcfR2genind" function from the vcfR package version 1.15.0 (Knaus & Grünwald, 2017). Pairwise between-population genetic distances were calculated using the genetic distance measure of Kosman & Leonard (2005), which is appropriate for diploid organisms and biallelic SNP data, using the "gd.kosman" function from the PopGenReport package version 3.1

(Adamack & Gruber, 2014). Pairwise geographic distances between populations were calculated using coordinates on the WGS ellipsoid using the "pointDistance" function from the raster package version 3.6.26 (Hijmans, 2023). For each species, IBD was fitted using a linear model with pairwise genetic distance as the response variable and pairwise geographic distance as the effect variable using the "lm" function from the stats package. Since pairwise comparisons are non-independent, significance was assessed using one-tailed Mantel tests, based on 999 matrix permutations, in which the genetic distance matrix was shuffled using the "shuf.dist" function from the BKlibR package version 0.0.0 (Byungju, 2021), and the t-value associated with the slope coefficient was used as the test statistic.

An otherwise strong pattern of IBD in *O. corymbosum* was contradicted by a single population (Outeniqua Pass; COR8) whose genetic profile is suggestive of recent, long-distance dispersal from the central Langeberg. This population was therefore omitted from IBD inference for this species. Also, since the Species O samples included two populations which are situated < 5 km apart and which are genetically very similar (POL1 and POL2), one of these was selected randomly (POL2) and removed from the analysis on the basis of being genetically non-independent.

IBD patterns were compared between species by fitting a linear model with genetic distance as the response variable, and geographic distance, species, and the interaction between geographic distance and species as fixed-effect predictors. To facilitate assessment of the effect of dispersal mode on gene flow, *O. corymbosum*, the only known myrmecochorous species, was specified as the reference. Significance was assessed using a two-tailed Mantel test with 9999 permutations in which the genetic distance matrices were shuffled within each species and the t-values associated with each predictor used as test statistics.

Results

Evolution of fruit traits in Osteospermum section Polygalinae

Fruits of all examined species were confirmed to be hard-walled nutlet-like fruits (Figure 2). An elaiosome-like appendage was only present on fruits of Species F (*O. corymbosum*) and ancestral character state reconstruction therefore suggested a single gain of this trait on the branch leading to this species (Figure 2). Fruit surface sculpturing traits, i.e., surface texture and presence of ridges with membranous edges, were both shown to be homoplasious, with the ancestral conditions uncertain, although membranous ridge edges are more likely a derived trait with at least two independent origins, in Clades 1 and 3 (Figure 2). Fruits in *O.* sect. *Polygalinae* were estimated to have an ancestral mass of 0.028 g, with heavier fruits evolving in species F, I and K. The ancestral fruit width was inferred to be 3 mm and to remain fairly constant with little variability between species, except

for in Species I and K, which have noticeably wider fruits (Figure 2). Elaiosomes in Species F (*O. corymbosum*) appear to have evolved in association with greater than ancestral mass, a smooth surface lacking ridges, but an intermediate width and a similar narrowly cylindrical shape to most other species in *O.* sect. *Polygalinae* (Figure 2; Clade 2). This contrasts with Species I and K which have also evolved greater mass and have smooth surface texture but have much greater widths, making their shape more rounded.

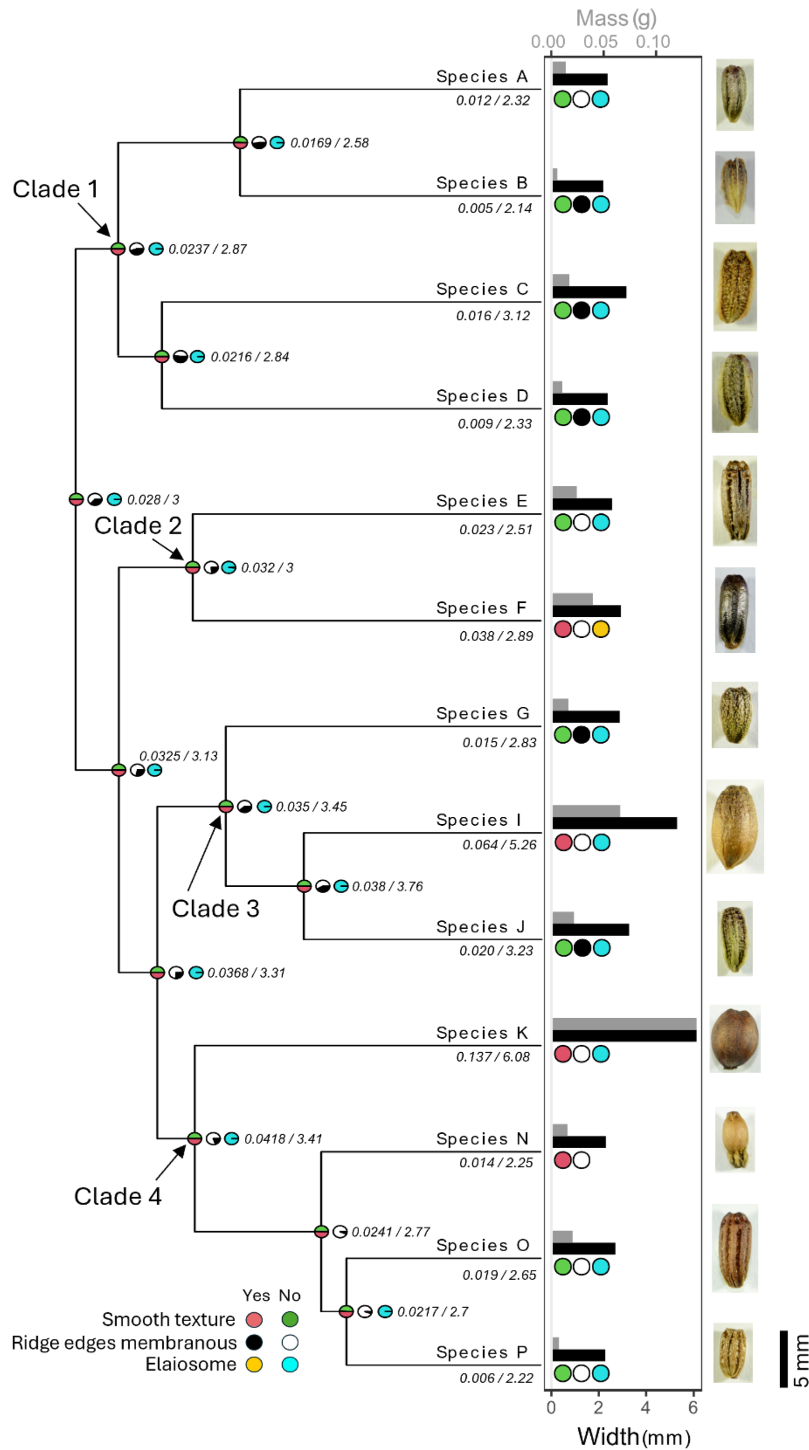


Figure 2. Phylogeny of *Osteospermum* section *Polygalinae* from Chapter 2 showing fruit traits of species at the tips (Species H and M omitted) and reconstructions of ancestral traits at the nodes. Discrete traits (indicated by coloured circles: fruit surface texture, ridge edge structure, and presence / absence of elaiosome) were reconstructed using maximum likelihood; trait likelihoods are shown as pie charts at the nodes. Continuous traits are shown as bar graphs (mass in grey and width in black) and numbers (mass (g) / width (mm)) at the tips and were reconstructed using a Brownian motion model and residual maximum likelihood and are indicated as number at the nodes. Fruit images on the right are scaled to comparable size.

Experimental tests of myrmecochory

In Experiment 1, ants were observed to remove fruits of *O. corymbosum* within seconds of fruits being placed near ant nests. All unmanipulated (elaiosome-bearing) fruits were removed at all sites within 1844 seconds (30.73 mins) of the fruits being discovered by ants (Figure 3). In contrast only 63% of fruits without elaiosomes were removed and were removed at a slower rate (Figure 3). The ants observed in my experiments were identified as *Anoplolepsis steingroeveri* (Forel).

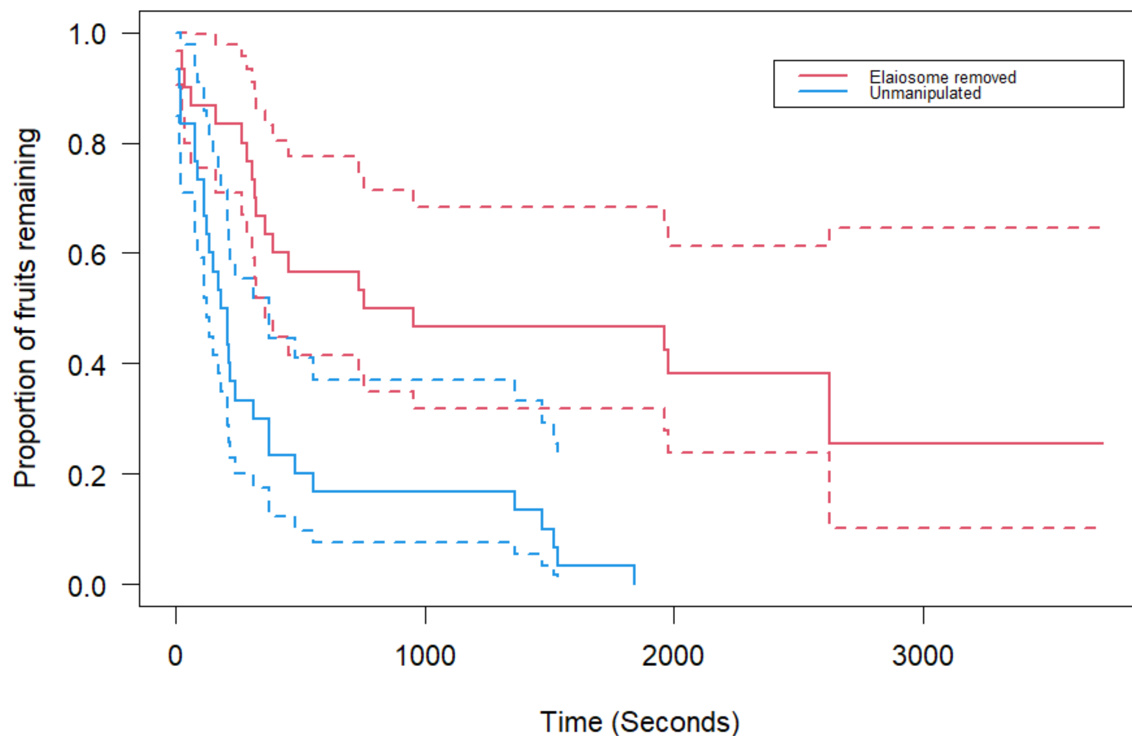


Figure 3. Survival curves showing the removal rates of *O. corymbosum* fruits with elaiosomes (unmanipulated; blue) and without elaiosomes (elaiosome removed; red) when offered to ants in cafeteria experiments. Plots were produced from a Kaplan-Meier survival model with data time adjusted so that t = 0 corresponds to the time the first fruit was removed at a site. Solid lines show fruit removal rates and dashed lines indicated 95% confidence intervals. The survival rates (i.e., proportion of unmoved fruit) of fruits with and without elaiosomes were significantly different (Chisq = 18.8, df = 1, p-value < 0.001, n = 30).

The Cox proportional hazards mixed effect model indicated significant differences in removal rates of fruits with and without elaiosomes, the removal of unmanipulated fruit being 5.2 times as likely as fruits with elaiosomes removed after controlling for the site effect (Table 1). However, there was no significant difference in the rate of removal of damaged, elaiosome-bearing fruit (control) relative to unmanipulated fruit (Table 1), indicating that differences in the removal rate of fruits with and

without elaiosomes is due to elaiosome presence or absence rather than being an artifact of damage associated with elaiosome removal. Moreover, although a large proportion of fruits with elaiosomes removed were transported from their original position, a significantly higher proportion of unmanipulated fruits were taken into the entrance of an ant nest (mean ± SE; 0.78 ± 0.11) than were fruits with elaiosomes removed (0.03 ± 0.03), as indicated by the binomial model (Table 2). Although a slightly smaller proportion of damaged, elaiosome-bearing (i.e., control) fruits (0.55 ± 0.10) were taken into the ant nest than were unmanipulated fruits, this difference was not significant, but the proportion of control fruits taken into the ant nest was significantly greater than that of fruits with elaiosomes removed (Table 2). This suggests that it is the presence of an elaiosome that increases the probability of a fruit being taken into the ant nest.

Unmanipulated fruits were also transported a significantly greater distance (mean ± SE; 29.55 ± 6.40 cm) than fruits with elaiosomes removed (8.36 ± 2.42 cm) as indicated by the linear model (elaiosome removed – unmanipulated; estimate = -21.30, SE = 5.43, df = 54.1, t-ratio = -3.926, p-value = 0.001). The distance fruits were moved did not differ significantly between the two treatments in which elaiosomes were retained (unmanipulated – control; estimate = 8.65, SE = 5.59, df = 52.5, t-ratio = 1.549, p-value = 0.277) with control fruits being transported only a slightly shorter distance (23.52 ± 5.69 cm) than unmanipulated fruits. Although the model did not show a significant difference in the distance moved between control fruits and those with elaiosomes removed (elaiosome removed – control: estimate = -12.65, SE = 6.27, df = 54.8, t-ratio = -2.018, p-value = 0.118), this is likely a result of there being fewer data points for these two treatments. This was because many fruits with elaiosomes removed were not moved from their original positions, so could not be included in the dataset, and many control fruits were transported away by ants so rapidly that they could not be tracked for their distances to be measured.

Table 1. Pairwise contrasts showing the differences in treatment effects on fruit removal by ants as modelled by the Cox proportional hazards model. The estimate gives the relative probability of removal of fruits between treatments. Note that the model produces results on the log scale, so model output has been converted to the response scale. The treatments are coded as follows: Unmanipulated fruits bearing elaiosomes (unmanipulated), fruits with damage on the opposite end to the elaiosome (control) and fruits with the elaiosome removed (elaiosome removed).

Contrast	Estimate	SE	z-ratio	p-value
Unmanipulated - elaiosome removed	5.197	0.715	134.963	< 0.001
Control - elaiosome removed	3.691	0.708	43.380	< 0.001
Control - unmanipulated	0.710	0.723	0.348	0.542

Table 2. Pairwise contrasts showing the differences in the number of fruits transported into ant nests between the three treatments as modelled by a binomial model. Note that the model results presented are on the log odds scale. The treatments are coded as follows: Unmanipulated fruits bearing elaiosomes (unmanipulated), fruits with damage on the opposite end to the elaiosome (control) and fruits with the elaiosome removed (elaiosome removed).

Contrast	Estimate	SE	z-ratio	p-value
Elaiosome removed - Unmanipulated	-5.26	1.374	-3.828	0.0004
Elaiosome removed - Control	-4.02	1.310	-3.065	0.0062
Unmanipulated - Control	1.24	0.787	1.580	0.2541

In Experiment 2, significantly more fruits of *O. corymbosum* were removed than fruits of Species E (n = 40, V = 63, p-value = 0.008). Dispersal did not readily take place during the course of this experiment, with fruits of either species only being removed at 11 of the 40 markers. Since Experiment 1 demonstrated that ants enthusiastically disperse fruits of *O. corymbosum* when they find them, it is possible that the many cases of no removal are due to non-detection. I attribute this to differences in weather between the days on which Experiments 1 and 2 were conducted, the latter taking place on days which were overcast with cool temperatures, while the weather was hot with bright sunshine during Experiment 1. Despite the low removal rates in Experiment 2, the two species still showed a strong and highly significant difference in the rate of fruit removal. When assessed across all 40 markers, the mean (±SE) number of fruits removed (out of five offered) was 0.55 ± 0.19 for *O. corymbosum* and 0.08 ± 0.04 for Species E (Figure 4A), while when assessed only across the 11 markers at which the removal of at least one fruit was recorded, the corresponding removal rates were 2.00 ± 0.45 and 0.27 ± 0.14 respectively (Figure 4B).

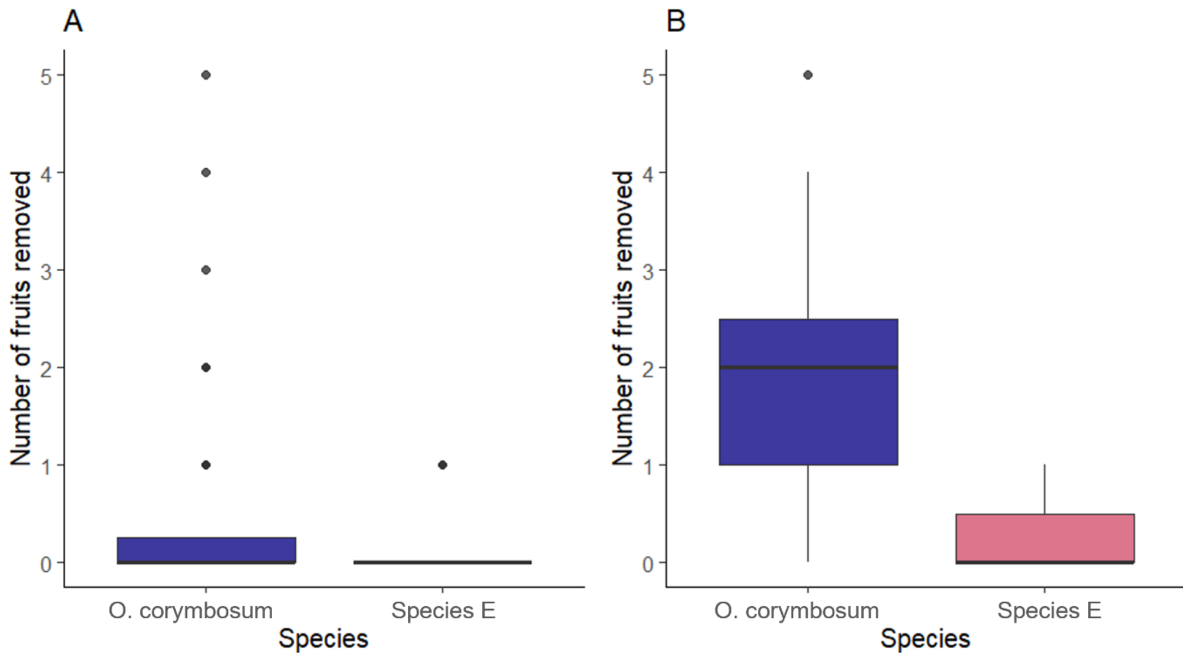


Figure 4. Comparison of fruit removal between *O. corymbosum* (blue; with elaiosome) and Species E (pink; no elaiosome) in cafeteria experiments at transect markers. Ten fruits (five of each species) were offered at each marker. (A) number of fruits removed per marker across all 40 replicates. Note that very few fruits were removed overall in this experiment. (B) comparison of fruit removal only in the eleven replicates where at least one fruit (of either species) was removed.

Between-species variation in the spatial scale of gene flow

Four of the seven species investigated showed significant IBD patterns (Figure 5), with geographic distance explaining a high proportion of the between-population genetic distance in Species F (*O. corymbosum*) and Species P, which had the steepest IBD slopes. Since the power of Mantel tests is limited by the number of pairwise comparisons, the non-significance of IBD in Species I (n = 5) and K (n = 4), and the weak significance of IBD in Species P (n = 5) may be a consequence of small sample size. All other species had 10 or more populations sampled. Species J exhibits no significant IBD despite the number of populations sampled being large (n = 11).

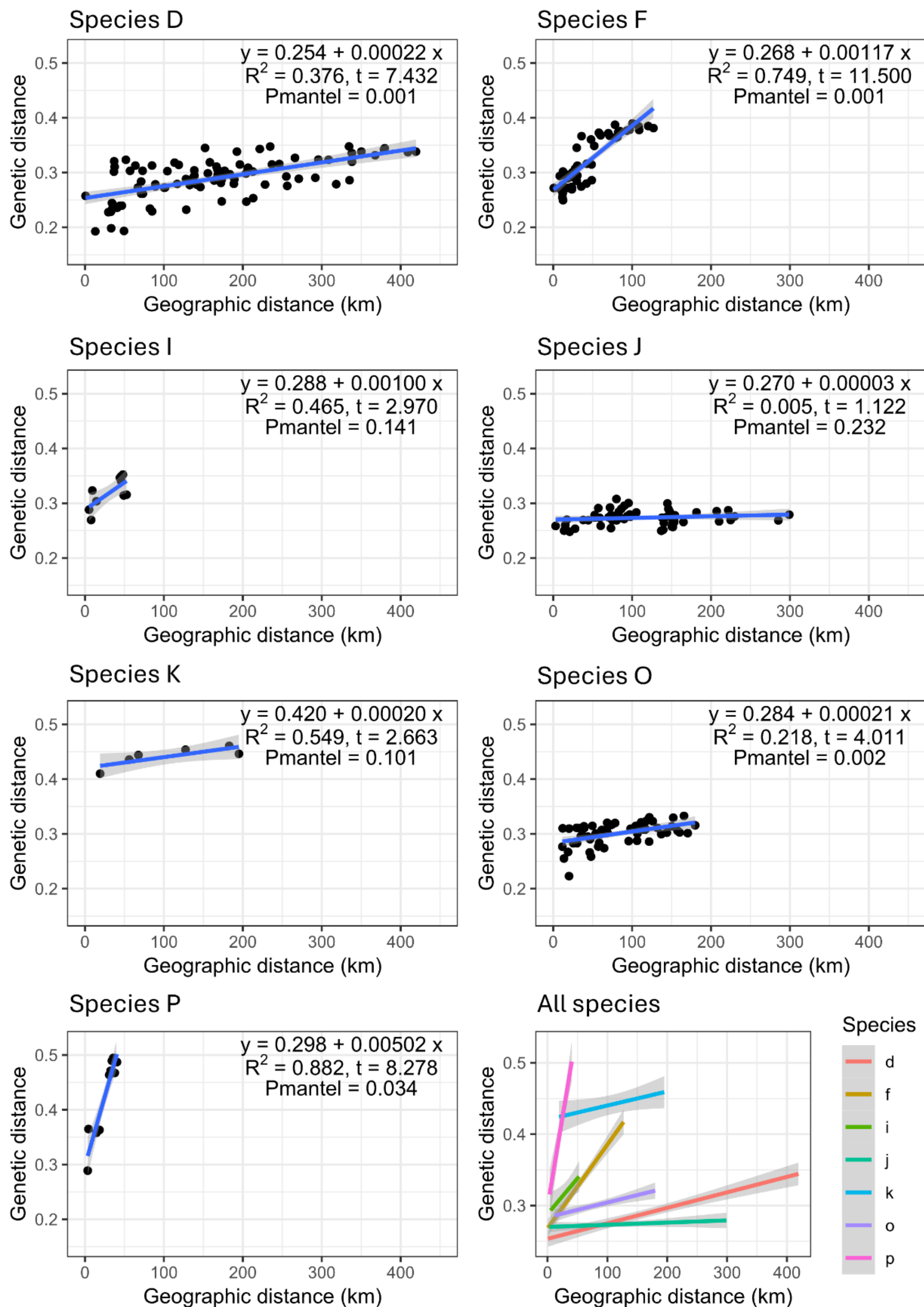


Figure 5. Isolation-by-distance plots for the seven individual species analysed (D, F, I, J, K, O and P) and for all seven species on a common set of axes (All species). Trend lines show predictions of linear models, grey shaded areas indicate confidence intervals and model parameters are given above the plots. Significance of IBD relationships were tested using Mantel tests and are indicated by Pmantel.

Comparing the IBD slopes between species (Figure 5; All species), Species D, J, K and O all had a significantly weaker IBD pattern (i.e., shallower slope) than the myrmecochorous Species F (*O. corymbosum*), implying that the former have higher levels of between-population gene flow over comparable geographical distances (Table 3; Figure 5). Species P, however, had a steeper IBD slope than Species F, while the slopes of Species F and I were statistically indistinguishable (Table 3; Figure 5). This suggests that between-population gene flow is most limited in Species P, followed by Species F (myrmecochorous) and Species I.

Table 3. Model parameters for the comparison of IBD between species produced by a linear model with genetic distance as the response variable, and geographic distance, species, and the interaction between geographic distance and species as fixed-effects. Since pairwise comparisons of genetic and geographic distances are non-independent, significance was assessed using Mantel tests (P_{mantel}) in addition to the normal test of significance given by the linear model (P_{param}). The myrmecochorous Species F (*O. corymbosum*), was used as the reference.

	Estimate	Std. Error	t-value	P_{param}	P_{mantel}
(Intercept)	0.2680	0.0059	45.126	< 0.001	< 0.001
Geodistance	0.0012	0.0001	11.789	< 0.001	< 0.001
Species D	-0.0145	0.0074	-1.952	0.052	0.875
Species I	0.0200	0.0153	1.311	0.191	0.040
Species J	0.0021	0.0085	0.243	0.808	< 0.001
Species K	0.1524	0.0188	8.110	< 0.001	< 0.001
Species O	0.0156	0.0085	1.849	0.066	0.007
Species P	0.0302	0.0166	1.823	0.070	0.859
Geodistance:Species D	-0.0010	0.0001	-9.362	< 0.001	< 0.001
Geodistance:Species I	-0.0002	0.0004	-0.451	0.653	0.382
Geodistance:Species J	-0.0011	0.0001	-10.427	< 0.001	< 0.001
Geodistance:Species K	-0.0010	0.0002	-5.654	< 0.001	< 0.001
Geodistance:Species O	-0.0010	0.0001	-8.172	< 0.001	< 0.001
Geodistance:Species P	0.0038	0.0006	6.948	< 0.001	0.007

Discussion

The functional significance of elaiosomes for dispersal

The results of this study provide clear support for the hypothesis that *Osteospermum corymbosum* is ant-dispersed, with the elaiosomes on the fruits of this species functioning as an adaptation for ant-dispersal. Fruits of *O. corymbosum* in natural settings were dispersed at a significantly higher rate than the fruits of its elaiosome-lacking sister species (Species E), with removal rate of *O. corymbosum* fruits being nearly ten-fold higher than those of Species E when there was evidence of fruit removal (Figure 4). Moreover, these patterns are apparent and significant despite the experiment being conducted under sub-optimal weather conditions, since ants (including *Anoplolepsis spp.*) are generally most active under sunny conditions with warm temperatures (De Bie & Hewitt, 1990; Cros, Cerdá & Retana, 1997; Botes et al., 2006; Slingsby, 2017).

My experiments testing the functional role in dispersal of the pale appendage inferred to be an elaiosome on the fruits of *O. corymbosum* (Wood & Nordenstam, 2003) showed that the presence of the elaiosome was necessary for successful dispersal by ants, with ants showing preference for fruits with elaiosomes (100% removal and removal 5.2 times more likely than fruits without elaiosomes) and removing them at a faster rate than fruits without elaiosomes (Figure 3). Most importantly, even though a large proportion of fruits without elaiosomes were moved from their original position (63 %), few of these were taken into the ant nest (3% compared to 78% of fruits with elaiosomes), most being transported only a short distance compared to fruits with elaiosomes. These results support the theory that diaspore burial is an important part of myrmecochory (Bond & Slingsby, 1984; Bond & Breytenbach, 1985; Christian & Stanton, 2004) and show that an elaiosome is important for completion of this process, and not only for initial ant attraction and seed transport. Furthermore, it is possible that initial transport of fruits without elaiosomes was elevated in experimental settings due to ants being attracted by the presence of fruits with elaiosomes, since ants were observed to lose interest in fruits without elaiosomes after investigation and initial movement. In addition to my experiments, fruits of Species F were also observed to be dispersed by ants and transported into ant nests at other localities, including both *O. corymbosum* (Tradouw Pass and Outeniqua Pass) and *O. burttianum* (Op de Tradouw) (Figure 1D). It can be concluded that the appendage on the base of fruits of *O. corymbosum* truly is an elaiosome that plays an important role in fruit dispersal. Although ant-dispersal and elaiosomes are postulated to be frequent in the Cape flora (Bond & Slingsby, 1983), this is one of few studies to test this hypothesis experimentally, and the first confirming the functional role of an elaiosome in myrmecochorous Asteraceae in the Cape.

The evolution of fruit traits as adaptations for dispersal

Consistent with the situation that *O. corymbosum* and *O. burttianum* are not separate species (see Chapter 2), a single species (Species F) in *O.* sect. *Polygalinae* was found to have elaiosomes. The ancestral character state reconstruction therefore shows a single origin of elaiosomes on the branch leading to Species F. Following Baum and Larson (1991), the adaptive significance of this trait, and ant-dispersal, is best understood with reference to the antecedent fruit morphology and dispersal mode, as well as the selective environment occupied by this species. The ancestral fruit is inferred to have been a narrowly cylindrical, hard nutlet-like fruit. It is uncertain whether the surface was smooth or pitted and ridged, but it most probably possessed obtuse ridges lacking membranous ridge edges (Figure 2). As such it does not appear to be adapted for a particular active dispersal mode and was most likely passively dispersed. Fruit morphology in *O.* sect. *Polygalinae* seems somewhat labile, however, and the ancestral state possibly has all the prerequisites necessary to adapt to different dispersal mechanisms given selection for them and sufficient time (e.g., membranous ridge edges could become wings for wind dispersal and smooth nutlets could gain elaiosomes). However, fruit shape remains fairly invariant, with contemporary species in *O.* sect. *Polygalinae* not appearing particularly adapted for dispersal other than *O. corymbosum*, suggesting that most are passively dispersed. The inferred ancestral width was about 3mm, and *O. corymbosum* (Species F) remains consistent with this, having fruits that are only slightly narrower. Species I and K, however, have evolved fruits which are larger and more rounded (Figure 2). Fruit width may be constrained in myrmecochorous lineages because fruits need to be narrow enough to fit into ant nest entrances and tunnels, so being ancestrally narrow may have been an important prerequisite in *O.* sect. *Polygalinae* for adaptation for ant dispersal. From field observation, the fruits of *O. corymbosum* appear to be a very similar width to the entrance holes of ant nests (Figure 1C), with one observation being made where the ants had to widen the entrance slightly before a fruit would fit through. *Osteospermum corymbosum* has a smooth fruit surface texture, a trait it shares only with Species I, K and N. Fruits of *O. corymbosum* are thus consistent with the syndrome of myrmecocohrous diaspores being typically hard and smooth, an adaptation which ensures that ants struggle to grip them once the elaiosome has been removed, with the consequence that the fruit remains in the ant nest rather than being discarded on surface middens (Berg, 1975; Slingsby & Bond, 1981; Gómez, Espadaler & Bas, 2005). Ants transporting fruits in my experiments were primarily observed to grip the soft elaiosome with their mandibles while pulling fruit, although they were also capable of moving fruits lacking elaiosomes. In species with ridges and pitted surfaces ants may be able to grip the fruit itself more easily. Furthermore, the fruits of *O. corymbosum* have above average mass, being heavier than the inferred ancestral mass, but not nearly as heavy as Species I and K. Fruit mass may also be a constrained adaptation (but see below) since ants may struggle to carry very heavy fruits and may instead remove the elaiosome upon discovery rather than transport the entire fruit back to the nest. This behaviour

was observed by Berg (1975) and Westoby et al. (1982) who noticed smaller ant species that were attracted to elaiosomes but unable to carry the fruit.

Osteospermum australe (Species N) also appears to possess many of the fruit traits which reflect adaptation for myrmecochory. While fruits of *O. australe* have both smaller mass and width than *O. corymbosum*, they are also smooth and narrow. However, since the fruits of *O. australe* sampled in this study had already dried at the time of collection, the presence of an elaiosome could not be determined since elaiosomes on other *Osteospermum* species have been observed to dry out very quickly, shrink, turn brown and become undetectable (personal observation). I also note that none of the species examined possessed any structure that could be interpreted as an intermediate in terms of elaiosome development. Since fruits without elaiosomes were also observed to be transported by ants in my experiments, it is possible that once an elaiosome precursor first appears the fitness advantage provided by ant dispersal strongly selects for the evolution of full elaiosomes.

The ecological role of myrmecochory

The adaptive significance of ant-mediated fruit dispersal is believed to lie in its effectiveness as a seed burial strategy, this being particularly beneficial in environments in which surface-deposited seeds are at risk of predation or being destroyed by fire (Berg, 1975; Bond & Breytenbach, 1985; Slingsby & Bond, 1985; Manzaneda, Fedriani & Rey, 2005). Fire plays an important role in the ecology of the fynbos vegetation (Kraaij & van Wilgen, 2014) in which *O. corymbosum* occurs and as such the adaptations of this species are inextricably linked to surviving fire. A noteworthy feature of *O. corymbosum* stands is that all individuals are of a similar age, probably reflecting the behaviour of this species to recruit rapidly post-fire. Based on the evidence of local anecdotal reports of the dates of fires, personal field observations (Outeniqua Pass area burnt in 2018, visited 2022; Tradouw Pass area burnt 2017, visited 2021, 2022 and 2023; Marloth Nature Reserve burnt 2012, visited 2022), and herbarium records of prior plant presence in sites in which I found few or no individuals, *O. corymbosum* is an obligate reseeder with an early post-fire successional ecology. The plants become tall and dominate communities at around four years post-fire, with all individuals then flowering and fruiting prolifically. Thereafter, it experiences die-back around five to six years after fire and is almost completely absent after ten years. Since postfire regeneration is entirely from seed, with seed being produced soon after fire, it is critical that seeds are able to persist for several years before the next fire occurs. Consequently, the removal of seeds from the soil surface to underground ant nests, where they enjoy protection from predators and fire, may be an important selective advantage.

Rodent predation has been shown to be a major threat to myrmecochrous seeds when ants are excluded (Bond & Breytenbach, 1985). O'Dowd & Hay (1980) found that even short-distance surface movement of seeds is beneficial since rodents focus their foraging around mother plants. At my

Experiment 2 sites, fruits left out at the markers overnight after the results had been recorded appeared to have been eaten by seed predators by the next morning, with only shards of the outer pericarp remaining. This indicates that seed predation by nocturnal rodents is an active threat in *O. corymbosum* stands and highlights the importance of rapid dispersal of fruits to safe sites the same day that they are released. Since my results showed low natural dispersal rates in cool overcast weather, it is possible that there would be selection for myrmecochorous seeds to only mature and drop when the weather is favourable for ant activity. Since the fruits of *O. corymbosum* appear to be heavier than that inferred for its ancestor, it is possible that an increase in seed mass became possible once the fruit was being taken underground and was released from predation pressure, since rodents may prefer relatively larger seeds (Midgley et al., 2002; Rusch, Midgley & Anderson, 2013). Increased seed mass is also important for seed dormancy and would have been beneficial for a fast-growing post fire ephemeral since the seed would be stocked with more resources, giving the seedling a competitive advantage (le Maitre & Midgley, 1992; Leishman et al., 2000; Pausas & Lamont, 2022). However, since seed size is also a constraint of myrmecochory, as discussed above, there may be an optimum seed mass and width. The evolution of larger fruits in Species I and K presents an interesting comparison. Since they are not adapted for ant dispersal, they would appear to be particularly vulnerable to predation if only passively dispersed, although their larger mass and rounded shape may facilitate rolling away from the mother plant. Another possibility is that they may be dispersed by rodents and buried in caches, as has been found for a few *Leucadendron* (Proteaceae) species in the Cape (Midgley et al., 2002; Midgley & Anderson, 2005; White, Bronner & Midgley, 2017).

The impact of myrmecochory on the spatial scale of gene flow and implications for diversification

While myrmecochory has been suggested to stimulate diversification through its effect in limiting between-population gene flow (Bond & Slingsby, 1983; Johnson, 1992; Cowling et al., 1994; Lengyel et al., 2009), this study finds limited support for this hypothesis. One constraint is the presence of just a single origin of myrmecochory in the *O.* sect. *Polygalinae* clade, which limits the power to test the hypothesis. Nevertheless, although the slope of the IBD curve of *O. corymbosum* (Species F) is comparatively steep, it is not the steepest of all species in this clade, with Species I displaying a similar IBD slope and Species P having a steeper slope (Figure 5). Since the species in this study are inferred to be predominantly passively dispersed, dispersal distance is likely not the factor limiting gene flow since there is little reason to expect that ants will disperse seeds smaller distances than passive dispersal, which might explain the lack of correspondence between dispersal mode and the spatial scale of gene flow in this study. Following the findings of Lengyel et al. (2009), who found myrmecochorous lineages to be generally more species rich than their non-myrmecochrous sister lineages globally (including lineages with passive dispersal), and Cowling et al. (1994) who found

myrmecochorous genera in the Cape fynbos to be significantly more speciose than non-myrmecochrous genera, it might be suggested that the influence of myrmecochory on diversification compared to passive dispersal may be due to reduced extinction of myrmecochorous lineages rather than a lower spatial scale of gene flow. However, the results of both these studies are limited by lack of knowledge of dispersal modes at a species level since the only work detailing the occurrence of myrmecochory in the Cape, on which both these studies relied, is that of Bond & Slingsby (1983), who characterised the occurrence of myrmecochroy at a genus level. Consequently, a genus containing a single myrmecochorous species might be categorised as myrmecochorous by these studies, overlooking the dispersal modes of perhaps the majority of species in the genus. To highlight this point, *Osteospermum* contains 91 taxa (Sadler, 2024), but only two have ever been specifically reported to possess elaiosome-like structures (Bond & Slingsby, 1983; Wood & Nordenstam, 2003). The influence of passive dispersal in promoting diversification in the Cape as a result of low spatial scale of gene flow compared to other dispersal mechanisms may therefore have been overlooked in studies focussing on myrmecochory. There is therefore no firm evidence that myrmecochory is driving diversification in the Cape more than passive dispersal as a consequence of short-dispersal distances resulting in a lower spatial scale of gene flow and further studies are needed that confirm the dispersal mechanisms of species.

There are factors apart from dispersal mode that may also influence the IBD patterns of the species studied here and could potentially overwhelm the signature of dispersal mode. Firstly, since gene flow is impacted by pollen dispersal as well as seed dispersal, differences in pollen movement between species would also impact IBD patterns. However, floral structure is very similar across all species in *O*. sect. *Polygalinae* (Norlindh, 1943; personal observation), so it is unlikely that they would have different pollinators. In the field I frequently noticed pollen-covered monkey beetles (Scarabaeidae: Hopliini) in the capitula, which are one of the most important insect pollinators of Asteraceae in the Cape Floristic Region, and reported to have fairly short flight distances (Picker & Midgley, 1996; Goldblatt & Manning, 2011; Karolyi et al., 2016). Gene flow differences as a consequence of pollination mode are therefore not expected in the species studied here. Secondly, range continuity as a consequence of topography or habitat continuity may also influence the spatial scale of gene flow within a species and impact IBD patterns. The heterogeneous environment in the Cape may result in species having patchy distributions as a consequence of being restricted to certain substrates and microhabitats, which may reduce gene flow between patches (Ellis et al., 2014). The Cape Fold Mountains are thought to play a significant role in acting as a gene flow barrier (Verboom et al., 2015). This may play a role in limiting gene flow in the species studied here, perhaps apart from in Species J, which is restricted to lowland areas and did not show significant IBD, implying substantial gene flow and panmictic mating across its distribution. Notably, Species P (*O. rotundifolium*) which has the strongest IBD, always occurs in moist areas in valleys, so its strong IBD pattern may be due to

limited gene flow between drainage basins, as has explained restricted gene flow in *Argyroderma* N.E.Br. (Aizoaceae) (Ellis, Weis & Gaut, 2007). The transient nature of suitable habitats due to the fire cycle may also impact range continuity and thus gene flow (Cowling, 1987). The impact of this may be particularly strong in early post-fire successional species with ephemeral life histories, such as *O. corymbosum*, since there may be no pollen dispersal between adjacent populations if their flowering times are not in sync, perhaps preventing gene flow for many decades or even centuries. Most other species in *O.* sect. *Polygalinae* were not noticed to have such a strongly ephemeral life-history, with some appearing to have a persistence strategy, often being found in more exposed areas of their habitats where there was less competition, such as near rocky outcrops (*O. polygaloides*, Species O), on mountain summits where vegetation was less dense (Species A and Species K) or in drier habitats where vegetation density is also lower (*O. imbricatum*, Species J; Species B; and Species D). So post-fire flowering may particularly constrain gene flow in *O. corymbosum*, possibly contributing to its stronger IBD pattern. Lastly, IBD slopes do not exclusively indicate gene flow patterns and are also impacted by differences in lineage age, mutation rates, patterns of colonisation and the strength of clinal selection between populations in a species (Nei, 1972; Coyne & Orr, 2004; Rundle & Nosil, 2005; Nosil, 2008). The extent to which this influences IBD patterns in species in this study is unknown, however, based on the *O.* sect. *Polygalinae* phylogeny (see Chapter 2; Figure 3), it seems unlikely that lineage age would explain the strong IBD pattern in Species P, since Species P appears to be a relatively young species.

In conclusion, there is little evidence to indicate a consistent negative effect of ant dispersal on the spatial scale of gene flow relative to the passive dispersal mode which is prevalent in *O.* sect. *Polygalinae* and which dominates the Cape flora (Moll & McKenzie, 1994). Thus, suggestions that ant dispersal is a major driver of diversification in the Cape remain open to question. More significantly, ant dispersal presents an important adaptation for avoiding predation during the inter-fire period, and the destruction of seeds by fire itself, by directed dispersal into underground ant nests. This is, perhaps, especially important for fynbos species that flower early in the inter-fire period and whose seeds are large enough to be of interest to rodents. This leads to a prediction of the kinds of environments and fruit/seed morphologies in which myrmecochory is most likely to arise.

References

Adamack, A.T. & Gruber, B. 2014. POPGENREPORT: simplifying basic population genetic analyses in R. *Methods in Ecology and Evolution*. 5(4):384–387. DOI: 10.1111/2041-210X.12158.

Bates, D., Mächler, M., Bolker, B. & Walker, S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*. 67(1):1–48. DOI: 10.18637/jss.v067.i01.

Baum, D.A. & Larson, A. 1991. Adaptation review: A phylogenetic methodology for studying character macroevolution. *Systematic Zoology*. 40(1):1–18.

Berg, R.Y. 1975. Myrmecochorous plants in Australia and their dispersal by ants. *Australian Journal of Botany*. 23(3):475–508. DOI: 10.1071/BT9750475.

van Blerk, J.J., West, A.G. & Midgley, J.J. 2017. Restio culm felling is a consequence of pre-dispersal seed predation by the rodent *Rhabdomys pumilio* in the Fynbos. *South African Journal of Botany*. 112:123–127. DOI: 10.1016/j.sajb.2017.05.017.

Bond, W.J. 1984. Fire survival of Cape Proteaceae-influence of fire season and seed predators. *Vegetatio*. 56(2):65–74. DOI: 10.1007/BF00033048.

Bond, A.W. & Slingsby, P. 1984. Collapse of an ant-plant mutualism: the Argentine ant (*Iridomyrmex humilis*) and myrmecochorous Proteaceae. *Ecology*. 65(4):1031–1037.

Bond, W.J. & Breytenbach, G.J. 1985. Ants, rodents and seed predation in Proteaceae. *South African Journal of Zoology*. 20(3):150–154. DOI: 10.1080/02541858.1985.11447928.

Bond, W.J. & Slingsby, P. 1983. Seed dispersal by ants in shrublands of the Cape Province and its evolutionary implications. *South African Journal of Science*. 79(6):231–233.

Botes, A., McGeoch, M.A., Robertson, H.G., van Niekerk, A., Davids, H.P. & Chown, S.L. 2006. Ants, altitude and change in the northern Cape Floristic Region. *Journal of Biogeography*. 33(1):71–90. DOI: 10.1111/j.1365-2699.2005.01336.x.

Braschler, B., Chown, S.L. & Gaston, K.J. 2012. The Fynbos and Succulent Karoo Biomes do not have exceptional local ant richness. *PLoS ONE*. 7(3):e31463. DOI: 10.1371/journal.pone.0031463.

Brew, C.R., O'Dowd, D.J. & Rae, I.D. 1989. Seed dispersal by ants: behaviour-releasing compounds in elaiosomes. *Oecologia*. 80(4):490–497. DOI: 10.1007/BF00380071.

Byungju, K. 2021. BKlibR. Available: https://github.com/airbj31/BKlibR.

Christian, C.E. 2001. Consequences of a biological invasion reveal the importance of mutualism for plant communities. *Nature*. 413(6856):635–639. DOI: 10.1038/35098093.

Christian, C.E. & Stanton, M.L. 2004. Cryptic consequences of a dispersal mutualism: seed burial, elaiosome removal, and seed-bank dynamics. *Ecology*. 85(4):1101–1110. DOI: 10.1890/03-0059.

Cowling, R.M. 1987. Fire and its role in coexistence and speciation in Gondwanan shrublands. *South African Journal of Science*. 83:106–112.

Cowling, R.M. & Holmes, P.M. 1992. Endemism and speciation in a lowland flora from the Cape Floristic Region. *Biological Journal of the Linnean Society*. 47(4):367–383. DOI: 10.1111/j.1095-8312.1992.tb00675.x.

Cowling, R.M., Holmes, P.M. & Rebelo, A.G. 1992. Plant diversity and endemism. In *The Ecology of Fynbos*. R.M. Cowling, Ed. Cape Town: Oxford University Press. 62–112.

Cowling, R.M., Pierce, S.M., Stock, W.D. & Cocks, M. 1994. Why are there so many myrmecochorous species in the Cape fynbos? In *Plant-Animal Interactions in Mediterranean-type Ecosystems*. M. Arianoutsou & R.H. Groves, Eds. (Tasks for Vegetation Science no. 31). Dordrecht: Kluwer Academic Publishers. 159–168.

Coyne, J.A. & Orr, H.A. 2004. *Speciation*. Sunderland, Massachusetts: Sinauer Associates, Inc.

Cros, S., Cerdá, X. & Retana, J. 1997. Spatial and temporal variations in the activity patterns of Mediterranean ant communities. *Écoscience*. 4(3):269–278. DOI: 10.1080/11956860.1997.11682405.

Cruzan, M.B. & Hendrickson, E.C. 2020. Landscape genetics of plants: challenges and opportunities. *Plant Communications*. 1(6):100100. DOI: 10.1016/j.xplc.2020.100100.

De Bie, G. & Hewitt, P.H. 1990. Thermal responses of the semi-arid zone ants *Ocymyrmex weitzeckeri* (Emery) and *Anoplolepis custodiens* (Smith). *Journal of the Entomological Society of Southern Africa*. 53(1):65–73.

Diniz-Filho, J.A.F., Soares, T.N., Lima, J.S., Dobrovolski, R., Landeiro, V.L., Telles, M.P.D.C., Rangel, T.F. & Bini, L.M. 2013. Mantel test in population genetics. *Genetics and Molecular Biology*. 36(4):475–485. DOI: 10.1590/S1415-47572013000400002.

Ellis, A.G., Weis, A.E. & Gaut, B.S. 2007. Spatial scale of local adaptation and population genetic structure in a miniature succulent, *Argyroderma pearsonii*. *New Phytologist*. 174(4):904–914. DOI: 10.1111/j.1469-8137.2007.02043.x.

Ellis, A.G., Verboom, G.A., van der Niet, T., Johnson, S.D. & Linder, H.P. 2014. Speciation and extinction in the Greater Cape Floristic Region. In *Fynbos: Ecology, Evolution, and Conservation of a Megadiverse Region*. N. Allsopp, J.F. Colville, & G.A. Verboom, Eds. Oxford: Oxford University Press. 119–141.

Givnish, T.J. 2010. Ecology of plant speciation. *Taxon*. 59(5):1326–1366.

Goldblatt, P. & Manning, J.C. 2002. Plant diversity of the Cape region of southern Africa. *Annals of the Missouri Botanical Garden*. 89(2):281–302. DOI: 10.2307/3298566.

Goldblatt, P. & Manning, J.C. 2011. Hopliine beetles (Scarabaeidae: Rutelinae: Hopliini), specialized pollinators of the southern African flora. *Curtis's Botanical Magazine*. 28(4):238–259. DOI: 10.1111/j.1467-8748.2011.01752.x.

Gómez, C. & Espadaler, X. 1998. Myrmecochorous dispersal distances: a world survey. *Journal of Biogeography*. 25(3):573–580. DOI: 10.1046/j.1365-2699.1998.2530573.x.

Gómez, C., Espadaler, X. & Bas, J.M. 2005. Ant behaviour and seed morphology: a missing link of myrmecochory. *Oecologia*. 146(2):244–246. DOI: 10.1007/s00442-005-0200-7.

Hijmans, R. 2023. raster: geographic data analysis and modeling. Version 3.6-26. Available: https://CRAN.R-project.org/package=raster.

Holmes, P.M. & Newton, R.J. 2004. Patterns of seed persistence in South African fynbos. *Plant Ecology*. 172(1):143–158. DOI: 10.1023/B:VEGE.0000026035.73496.34.

Johnson, S.D. 1992. Plant-animal relationships. In *The Ecology of Fynbos*. R.M. Cowling, Ed. Cape Town: Oxford University Press. 175–205.

Karolyi, F., Hansal, T., Krenn, H.W. & Colville, J.F. 2016. Comparative morphology of the mouthparts of the megadiverse South African monkey beetles (Scarabaeidae: Hopliini): feeding adaptations and guild structure. *PeerJ*. 4:e1597. DOI: 10.7717/peerj.1597.

Keeley, J.E., Bond, W.J., Bradstock, R.A., Pausas, J.G. & Rundel, P.W. 2012. *Fire in Mediterranean ecosystems ecology, evolution and management*. New York: Cambridge University Press.

Knaus, B.J. & Grünwald, N.J. 2017. VCFR: a package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources*. 17(1):44–53. DOI: 10.1111/1755-0998.12549.

Kosman, E. & Leonard, K.J. 2005. Similarity coefficients for molecular markers in studies of genetic relationships between individuals for haploid, diploid, and polyploid species. *Molecular Ecology*. 14(2):415–424. DOI: 10.1111/j.1365-294X.2005.02416.x.

Kraaij, T. & van Wilgen, Brian.W. 2014. Drivers, ecology, and management of fire in fynbos. In *Fynbos: Ecology, Evolution, and Conservation of a Megadiverse Region*. N. Allsopp, J.F. Coville, & G.A. Verboom, Eds. Oxford: Oxford University Press. 47–72.

Leishman, M.R., Wright, I.J., Moles, A.T. & Westoby, M. 2000. The evolutionary ecology of seed size. In *Seeds: the ecology of regeneration in plant communities*. 2nd ed. M. Fenner, Ed. UK: CABI Publishing. 31–57. DOI: 10.1079/9780851994321.0031.

Lengyel, S., Gove, A.D., Latimer, A.M., Majer, J.D. & Dunn, R.R. 2009. Ants sow the seeds of global diversification in flowering plants. *PLoS ONE*. 4(5):e5480. DOI: 10.1371/journal.pone.0005480.

Lengyel, S., Gove, A.D., Latimer, A.M., Majer, J.D. & Dunn, R.R. 2010. Convergent evolution of seed dispersal by ants, and phylogeny and biogeography in flowering plants: a global survey. *Perspectives in Plant Ecology, Evolution and Systematics*. 12(1):43–55. DOI: 10.1016/j.ppees.2009.08.001.

Lenormand, T. 2002. Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*. 17(4):183–189. DOI: 10.1016/S0169-5347(02)02497-7.

Lenth, R. 2024. emmeans: estimated marginal means, aka least-squares means. Version 1.10.0. Available: https://CRAN.R-project.org/package=emmeans.

Levine, N., Ben-Zvi, G., Seifan, M. & Giladi, I. 2019. Investment in reward by ant-dispersed plants consistently selects for better partners along a geographic gradient. *AoB PLANTS*. 11(3):plz027. DOI: 10.1093/aobpla/plz027.

Linder, H.P. 1985. Gene flow, speciation, and species diversity patterns in a species rich area: the Cape flora. In *Species and Speciation*. V. 4. E.S. Vrba, Ed. (Transvaal Museum Monograph). Pretoria: Transvaal Museum. 53–57.

Linder, H.P. 2003. The radiation of the Cape flora, southern Africa. *Biological Reviews of the Cambridge Philosophical Society*. 78(4):597–638. DOI: 10.1017/S1464793103006171.

le Maitre, D.C. & Midgley, J.J. 1992. Plant reproductive ecology. In *The Ecology of Fynbos*. R.M. Cowling, Ed. Cape Town: Oxford University Press. 135–174.

Mantel, N. 1967. The detection of disease clustering and generalized regression approach. *Cancer Research*. 27(2):209–220.

Manzaneda, A.J. & Rey, P.J. 2012. Geographical and interspecific variation and the nutrient-enrichment hypothesis as an adaptive advantage of myrmecochory. *Ecography*. 35(4):322–332. DOI: 10.1111/j.1600-0587.2011.06923.x.

Manzaneda, A.J., Fedriani, J.M. & Rey, P.J. 2005. Adaptive advantages of myrmecochory: the predator-avoidance hypothesis tested over a wide geographic range. *Ecography*. 28(5):583–592. DOI: 10.1111/j.2005.0906-7590.04309.x.

Marshall, D.L., Beattie, A.J. & Bollenbacher, W.E. 1979. Evidence for diglycerides as attractants in an ant-seed interaction. *Journal of Chemical Ecology*. 5(3):335–344. DOI: 10.1007/BF00987919.

McDonald, D.J., Juritz, J.M., Cowling, R.M. & Knottenbelt, W.J. 1995. Modelling the biological aspects of local endemism in South African Fynbos. *Plant Systematics and Evolution*. 195(3–4):137–147. DOI: 10.1007/BF00989290.

Midgley, J.J. & Anderson, B. 2005. Scatterhoarding in Mediterranean shrublands of the SW Cape, South Africa. In *Seed fate: predation, dispersal and seedling establishment*. P.-M. Forget, J.E. Lambert, P.E. Hulme, & S.B. Vander Wall, Eds. Wallingford UK: CAB International. 197–204.

Midgley, J.J. & Bond, W.J. 1995. Relative attractiveness of seeds of myrmecochorous Australian and South African plants to ants, and the chemical basis of this attraction. *South African Journal of Botany*. 61(4):230–232. DOI: 10.1016/S0254-6299(15)30520-2.

Midgley, J., Anderson, B., Bok, A. & Fleming, T. 2002. Scatter-hoarding of Cape Proteaceae nuts by rodents. *Evolutionary Ecology Research*. 4:623–626.

Milewski, A.V. & Bond, W.J. 1982. Convergence of myrmecochory in mediterranean Australia and South Africa. In *Ant-plant interactions in Australia*. R.C. Buckley, Ed. The Hague: Dr W. Junk Publishers. 89–98. DOI: 10.1007/978-94-009-7994-9_9.

Miller, B.P. & Dixon, K.W. 2014. Chapter 6: Plants and fire in kwongan vegetation. In *Plant Life on the Sandplains in Southwest Australia, a Global Biodiversity Hotspot*. H. Lambers, Ed. Crawley, Australia: University of Western Australia Publishing.

Mokotjomela, T.M., Musil, C.F. & Esler, K.J. 2013. Potential seed dispersal distances of native and non-native fleshy fruiting shrubs in the South African Mediterranean climate region. *Plant Ecology*. 214(9):1127–1137. DOI: 10.1007/s11258-013-0237-3.

Moll, E.J. & McKenzie, B. 1994. Modes of dispersal of seeds in the Cape fynbos. In *Plant-Animal Interactions in Mediterranean-type Ecosystems*. M. Arianoutsou & R.H. Groves, Eds. (Tasks for Vegetation Science no. 31). Dordrecht: Kluwer Academic Publishers. 151–157.

Mucina, L. & Majer, J.D. 2012. Ants and the origins of plant diversity in old, climatically stable landscapes: A great role for tiny players. *South African Journal of Botany*. 83:44–46. DOI: 10.1016/j.sajb.2012.07.003.

Nei, M. 1972. Genetic distance between populations. *The American Naturalist*. 106(949):283–292.

Norlindh, T. 1943. *Studies in the Calenduleae I: Monograph of the genera Dimorphotheca, Castalis, Osteospermum, Gibbaria and Chrysanthemoides*. Lund: C.W.K. Gleerup.

Nosil, P. 2008. Speciation with gene flow could be common. *Molecular Ecology*. 17(9):2103–2106. DOI: 10.1111/j.1365-294X.2008.03715.x.

O'Dowd, D.J. & Hay, M.E. 1980. Mutualism between harvester ants and a desert ephemeral: seed escape from rodents. *Ecology*. 61(3):531–540. DOI: 10.2307/1937419.

Orsini, L., Vanoverbeke, J., Swillen, I., Mergeay, J. & De Meester, L. 2013. Drivers of population genetic differentiation in the wild: isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Molecular Ecology*. 22(24):5983–5999. DOI: 10.1111/mec.12561.

Paradis, E. & Schliep, K. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*. 35(3):526–528. DOI: 10.1093/bioinformatics/bty633.

Pausas, J.G. & Lamont, B.B. 2022. Fire-released seed dormancy - a global synthesis. *Biological Reviews*. 97(4):1612–1639. DOI: 10.1111/brv.12855.

Pfeiffer, M., Huttenlocher, H. & Ayasse, M. 2010. Myrmecochorous plants use chemical mimicry to cheat seed-dispersing ants. *Functional Ecology*. 24(3):545–555. DOI: 10.1111/j.1365-2435.2009.01661.x.

Picker, M.D. & Midgley, J.J. 1996. Pollination by monkey beetles (Coleoptera: Scarabaeidae: Hopliini): flower and colour preferences. *African Entomology*. 4(1):7–14.

Pierce, S.M. & Cowling, R.M. 1991. Dynamics of soil-stored seed banks of six shrubs in fire-prone dune fynbos. *The Journal of Ecology*. 79(3):731–747. DOI: 10.2307/2260664.

van der Pijl, L. 1982. *Principles of dispersal in higher plants*. 3rd ed. Berlin: Springer-Verlag.

R Core Team. 2023. Version 4.3.2. Available: https://www.R-project.org/.

Rice, B. & Westoby, M. 1986. Evidence against the hypothesis that ant-dispersed seeds reach nutrient-enriched microsites. *Ecology*. 67(5):1270–1274. DOI: 10.2307/1938682.

Ronce, O. 2007. How does it feel to be like a rolling stone? Ten questions about dispersal evolution. *Annual Review of Ecology, Evolution, and Systematics*. 38(1):231–253. DOI: 10.1146/annurev.ecolsys.38.091206.095611.

Rundle, H.D. & Nosil, P. 2005. Ecological speciation. *Ecology Letters*. 8(3):336–352. DOI: 10.1111/j.1461-0248.2004.00715.x.

Rusch, U.D., Midgley, J.J. & Anderson, B. 2013. Rodent consumption and caching behaviour selects for specific seed traits. *South African Journal of Botany*. 84:83–87. DOI: 10.1016/j.sajb.2012.09.007.

Sadler, R.T. 2024. Phylogeny, generic limits and evolution of dispersal traits in the Asteraceae tribe Calenduleae. MSc thesis. University of Cape Town.

Sernander, R. 1906. *Entwurf einer Monographie der europaischen Myrmekochoren*. V. 41. (no. 7). Uppsala: Kungliga Svenska Vetenskapsakademien Handlingar.

Slingsby, P. 2017. *ANTS of Southern Africa*. Muizenberg: Slingsby Maps.

Slingsby, P. & Bond, W.J. 1981. Ants — Friends of the Fynbos. *Veld and Flora*. 67:39–45.

Slingsby, P. & Bond, W.J. 1985. The influence of ants on the dispersal distance and seedling recruitment of *Leucospermum conocarpodendron* (L.) Buek (Proteaceae). *South African Journal of Botany*. 51(1):30–34. DOI: 10.1016/S0254-6299(16)31698-2.

Speta, F. 1972. Entwicklungsgeschichte und karyologie von elaiosomen an samen und früchten. In *Naturkundliches Jahrbuch der Stadt Linz*. V. 18. Linz. 9–65.

Stock, W.D. & Verboom, G.A. 2012. Phylogenetic ecology of foliar N and P concentrations and N:P ratios across mediterranean-type ecosystems. *Global Ecology and Biogeography*. 21(12):1147–1156. DOI: 10.1111/j.1466-8238.2011.00752.x.

Therneau, T.M. 2023. A package for survival analysis in R. Version 3.5-7. Available: https://CRAN.R-project.org/package=survival.

Therneau, T.M. 2024. coxme: mixed effects cox models. Version 2.2-20. Available: https://CRAN.R-project.org/package=coxme.

Verboom, G.A., Bergh, N.G., Haiden, S.A., Hoffmann, V. & Britton, M.N. 2015. Topography as a driver of diversification in the Cape Floristic Region of South Africa. *New Phytologist*. 207(2):368–376. DOI: 10.1111/nph.13342.

Verboom, G.A., Slingsby, J.A. & Cramer, M.D. 2024. Fire-modulated fluctuations in nutrient availability stimulate biome-scale floristic turnover in time, and elevated species richness, in low-nutrient fynbos heathland. *Annals of Botany*. 133(5–6):819–832. DOI: 10.1093/aob/mcad199.

Weighill, B., Huysamer, A. & Anderson, B. 2017. The nightshift: Seed dispersal and consumption differences by rodents before and after dark. *South African Journal of Botany*. 108:267–271. DOI: 10.1016/j.sajb.2016.11.005.

Westoby, M., Rice, B., Shelley, J., Haig, D. & Kohen, J. 1982. Plant's use of ants for dispersal at West Head, NSW. In *Ant-plant interactions in Australia*. R. Buckley, Ed. The Hague: Dr W. Junk Publishers. 75–87.

White, J.D., Bronner, G.N. & Midgley, J.J. 2017. Camera-trapping and seed-labelling reveals widespread granivory and scatter-hoarding of nuts by rodents in the Fynbos Biome. *African Zoology*. 52(1):31–41. DOI: 10.1080/15627020.2017.1292861.

Wood, A.R. & Nordenstam, B. 2003. An interesting new species of *Osteospermum* (Asteraceae-Calenduleae) from the Western Cape Province, South Africa, providing a link to the genus *Chrysanthemoides*. *South African Journal of Botany*. 69(4):572–578. DOI: 10.1016/S0254-6299(15)30297-0.

Wright, S. 1943. Isolation by distance. *Genetics*. 28(114): 114–138.

Wright, S. 1946. Isolation by distance under diverse systems of mating. *Genetics*. 31(1):39–59. DOI: 10.1093/genetics/31.1.39.

Chapter 4: Synthesis

The availability of genomic data with population-level resolution allows the delimitation of species that are independent evolutionary units and thus realistic representations of units of biodiversity. Whereas previous taxonomic work on *Osteospermum* section *Polygalinae* resulted in the recognition of six species (Norlindh, 1943; Nordenstam, 2004), application of an integrative taxonomic approach based on a genomic framework (Chapter 2) reveals the existence of 14 to 16 species in this clade. Historically, morphology was the only source of information available for inferring species relationships (Rouhan & Gaudeul, 2014). However, since most morphological traits are also functional and under selection, they often show strong evolutionary divergence and convergence (Johnson, 1992; Anderson et al., 2014; Cramer et al., 2014), this erasing the phylogenetic signal they contain, and compromising their phylogenetic and taxonomic utility (Givnish & Sytsma, 1997). Thus, where Norlindh (1943) relied heavily on fruit traits to separate taxa in the Calenduleae, recent studies at the generic and sectional level (Sadler et al., 2022; Sadler, 2024), and now also at the species level in *O*. sect. *Polygalinae* (this study), reveal these traits to be highly homoplasious. Largely due to the greater number of traits, molecular data appears less susceptible to homoplasy (Givnish & Sytsma, 1997), especially when generated by next-generation sequencing (NGS) methods that allow thousands of independently assorting loci to be sampled from across the genome (Kusy et al., 2019; Van Damme et al., 2022). This study demonstrates how a large NGS molecular dataset combined with extensive population sampling allows lineages to be resolved with strong support at very fine spatial and phylogenetic (population level) scales, thereby providing an evolutionary framework within which species can be delimited. An integrative taxonomic approach, which draws on morphological, ecological, and spatial data, can then be used to delimit species that are meaningful evolutionary units (de Queiroz, 2007), and which reflect their ecological roles (Freudenstein et al., 2017). While most of the species identified with this approach show concordance across multiple types of data (e.g. morphology, genetics), further sampling is required to resolve relationships within Species D and *O. polygaloides* (Species O) and to determine whether these entities comprise more than one species. This, however, reflects the iterative nature of hypothesis-driven taxonomy, in which species are delimited as well as possible given the data available (Yeates et al., 2011). The two morphologically cryptic species, Species H and M, also need further study to determine whether they possess phenotypic differences that can be used to distinguish them from closely related *O. imbricatum* (Species J) and *O. polygaloides* (Species O) respectively. Notwithstanding these uncertainties, according to my genetic data, all species in *O.* sect. *Polygalinae* resolved in this study reflect evolutionarily independent units whose interrelationships are now known. Beyond providing the basis for a natural taxonomy, this allows novel insights into the evolution, history and future of this group. This is well exemplified by Chapter 3, which uses the newly defined species and their phylogenetic relationships to study fruit evolution in *O.* sect. *Polygalinae* and to test a hypothesis relating to the

microevolutionary consequences of fruit traits and their relation to dispersal. In the paragraphs that follow, I highlight some further insights that can be gained from the findings in the previous two chapters, discuss their relevance in a broader context, and offer suggestions for future research.

My proposed *O.* sect. *Polygalinae* taxonomy indicates that the widespread *O. imbricatum* and *O. polygaloides*, as conceptualized by Norlindh (1943), both comprise a series of species that are more geographically restricted than was previously thought, with some of these species being confined to a particular substrate. Additionally, extensive fieldwork reveals that populations of most species are generally small, comprising few individuals aggregated in a tight stand, rather than being distributed throughout the environment. This has conservation implications since species with small ranges, habitat specificity, and small, localized populations are more vulnerable to extinction (Matthies et al., 2004; Işik, 2011). The two species from the Swartberg (Species A and B), for example, are at present known from just one locality each. Further exploration is, however, needed to determine the full extent of their distributions. Both Species M and *O. australe* (Species N) are limestone endemics, being restricted to limestone ridges in the Agulhas Plain and eastwards to Still Bay. This is a region with high local endemism in which, it is suggested, specialization to patchily distributed calcareous substrates has promoted plant speciation (Cowling & Holmes, 1992; Willis, Cowling & Lombard, 1996; Grobler & Cowling, 2021). The vegetation of this area has, however, been fragmented by agricultural land transformation, the spread of alien invasive plants, and human settlement, all of these threatening the endemic flora (Willis, Cowling & Lombard, 1996; Lombard et al., 1997). Although portions of this vegetation are now protected in the De Hoop Nature Reserve, Agulhas National Park, Nuwejaars Wetlands Special Management Area (https://nuwejaars.com), and other private reserves, many populations and taxa remain threatened. *Osteospermum australe*, for example, has not been found again at the type locality in the limestone hills near Cape Agulhas where it was collected in 1962 (Nordenstam, 2004), owing to this area being massively transformed by the expansions of the towns L'Agulhas and Struisbaai. While most species in *O.* sect. *Polygalinae* occur in fynbos vegetation on nutrient-poor soils, *O. imbricatum* (Species J) is frequently associated with renosterveld, a vegetation type which is critically endangered on account of agricultural land transformation. Presently, less than 10% of the historical area occupied by renosterveld remains intact (Kemper, Cowling & Richardson, 1999; Topp & Loos, 2019) and further areas continue to be lost every year (Moncrieff, 2021). While many renosterveld species have very localized distributions and are therefore endangered (Newton & Knight, 2010), this fortunately does not seem to be the case for *O. imbricatum* which has a relatively wide distribution in lowland areas of the central CFR. A lack of significant IBD between populations of *O. imbricatum* (Chapter 3) also suggests panmictic gene flow across its range, though it may well be the case that this pattern is historical, with the effects of recent habitat fragmentation not yet being apparent. As a result of the present study, all species in *O.* sect.

Polygalinae will now be able to have their proper conservation status accurately assessed once they have been formally described.

The four main clades within *O*. sect. *Polygalinae* are geographically displaced, showing limited range overlap. Moreover, where their ranges overlap, their members seldom grow sympatrically, usually being found at different elevations or in different vegetation types in that area. For example, in areas of range overlap between Species J (Clade 3) and Species F (Clade 2) in the central CFR, the former is restricted to lowlands and foothill habitats and the latter to mountain tops. The restricted distribution and dominance of each main clade in a different broad geographic region, suggests that each clade has diversified locally within its range. Additionally, all clades have species occurring across a diversity of habitats and elevations, and have species of a variety of growth forms. The repeated evolution of growth forms in similar habitats in different clades suggests convergent evolution of growth forms. This provides the opportunity to test hypotheses, in an evolutionarily replicated manner, about the role of environment in stimulating speciation and growth form evolution. Of particular interest, perhaps, is the repeated evolution in *O.* sect. *Polygalinae* of a tall, virgate growth form which is generally associated with rugged mountainous areas. Plants with this growth form, commonly referred to as "wand plants", occur in multiple families in the fynbos and while some hypotheses have been suggested to explain the functional significance of this growth form (Bailey et al., 2019), evolutionary drivers of this convergence are unknown. The species taxonomy and phylogenetic data presented in this study, could provide a valuable framework for exploring this question further.

Knowledge of dispersal ecology allows us to investigate evolutionary questions about the variables that select for different fruit morphologies and dispersal strategies. *Osteospermum* species occur across a range of vegetation types associated with *inter alia* different substrates, habitat openness and aridity (Mucina & Rutherford, 2006; Manning & Goldblatt, 2012). The genus also contains multiple different fruit types adapted for different dispersal modes, including winged fruits adapted for wind dispersal (Stock, 2021; Sadler, 2024), fleshy drupes adapted for dispersal by vertebrates, primarily birds (Norlindh, 1943; Gosper, 2004; Mokotjomela, Musil & Esler, 2013), and nutlet-like fruits with elaiosomes adapted for ant dispersal, as shown in this study. Many of these dispersal-related fruit types have evolved more than once (Sadler, 2024). Hard nutlet-like *Osteospermum* fruits without any apparent adaptations for dispersal are presumed to be passively dispersed by gravity, but there is a need for further work to determine whether scatter-hoarding by rodents plays a role in their dispersal (Midgley et al., 2002; Midgley & Anderson, 2005; White, Bronner & Midgley, 2017). This could perhaps explain the greater size and mass of fruits of Species I and Species K compared to other species in *O*. sect. *Polygalinae*, which fall into the size category of fruits preferred by rodents for scatter-hoarding (Rusch, Midgley & Anderson, 2013). There is evidence that wind dispersal in *Osteospermum* and other genera with winged fruits in the Cape is more prevalent in dry, open habitats

(Sadler, 2024) and globally there is a trend of myrmecochory being most prevalent in shrublands with phosphorous-poor soils (Milewski & Bond, 1982; Westoby et al., 1982). *Osteospermum* therefore provides an excellent study system for testing the selective role of environmental variables for different dispersal strategies.

Dispersal ecology is also important for predicting the invasive potential and conservation threats of species. Long distance dispersed species are more likely to spread rapidly through a novel environment and become invasive (Travis & Dytham, 2002; Wilson et al., 2009). This has been seen in *O. moniliferum* L. (known for a time as *Chrysanthemoides monilifera* (L.) Norl.), which is native to southern Africa but has become a rampant invader in Australia (Gosper, 2004; Scott & Batchelor, 2014; Emmett et al., 2023). Its fleshy drupes are apparently adapted for dispersal by vertebrates, particularly birds, which have been easily co-opted in its non-native range and which transport fruits long distances, facilitating rapid invasion (Gosper, 2004; Mokotjomela, Musil & Esler, 2013; Scott & Batchelor, 2014). In contrast, short-distance dispersed species are more vulnerable to extinction due to their inability to recolonise areas in which they have become locally extinct (Cowling et al., 1994; Cain, Milligan & Strand, 2000). When these species are also specialised for a mutualism, as in myrmecochorous species, they may be at even greater risk since they depend on the presence of another species for survival (Bond, 1994). Since myrmecochorous seeds can be dispersed by multiple different ant species, there may be some buffering against this threat (Slingsby & Bond, 1981; Milewski & Bond, 1982; Westoby et al., 1982; Levine et al., 2019). However, when entire ant communities are displaced, myrmecochrous plant species are at great risk of extinction. This has been witnessed in areas invaded by the Argentine ant (*Linepithema humile* Mayr) in Mediterranean shrublands in both South Africa and Spain (Bond & Slingsby, 1984; Christian, 2001; Gómez & Oliveras, 2003). Argentine ants displace native ant species and are ineffective seed dispersers, leaving seeds on the soil surface where they are vulnerable to predation, resulting in reduced seedling emergence of myrmecochrous species and consequent shifts in the plant community composition in invaded-areas (Bond & Slingsby, 1984; Christian, 2001; Gómez & Oliveras, 2003; Christian & Stanton, 2004). From these examples, understanding the dispersal ecology of plant species is clearly crucial for predicting both their invasive potential and extinction vulnerability.

Lastly, having species defined so as to correspond to evolutionary independent lineages, and knowing how they are related, are powerful tools for studying the processes of speciation (Barraclough & Nee, 2001). Until recently many of the hypothesized drivers of speciation in the Cape could not be rigorously tested due to a lack of appropriate phylogenetic context, but this has changed as molecular phylogenies have become gradually available over the past 30 years (Barraclough, 2006; Ellis et al., 2014). Most studies have looked at broad phylogenetic patterns or used sister-species comparisons (e.g., van der Niet & Johnson, 2009), which is a macroevolutionary approach that allows correlations to be drawn between diversification patterns and variables hypothesized to be driving speciation (Ellis

et al., 2014). The microevolutionary processes at the population level that lead up to these patterns have received less attention, probably due to the challenge of collecting the fine-scale datasets needed to investigate them (Ellis et al., 2014). While it is widely suggested that myrmecochory has contributed to diversification in the Cape flora through its effect on gene flow (e.g., Bond & Slingsby, 1983; Linder, 1985; Cowling, 1987; Goldblatt, 1997; Latimer, Silander & Cowling, 2005), this has only been studied at the macroevolutionary scale (Cowling et al., 1994; Lengyel et al., 2009). Population genetics and inter-population gene flow of myrmecochrous plant species have been investigated in other parts of the world (Maeyama, 2000; Nakagawa, 2010; Pascov et al., 2015; Zhang et al., 2017; Lee, Kim & Kim, 2018), but there are no studies which test, in a comparative manner, the role of ant-modulated dispersal in limiting gene flow in the context of a process that could lead to population differentiation and speciation. The present study is therefore novel in testing the microevolutionary consequences of myrmecochory. Although only one species was found to be myrmecochorous, so statistical power is limited to the extent that it is not possible to draw firm conclusions about the impact of ant dispersal relative to passive dispersal on the spatial scale of gene flow in *O*. sect. *Polygalinae*, the present study nonetheless demonstrates the kind of fine-level work that is needed to truly untangle the mechanism by which myrmecochory might enhance speciation. Further fine-scale studies are needed, first to confirm the dispersal mode of well-defined plant species and then to test whether this is truly driving speciation as a product of reduced gene flow between populations. Studies of pollination biology of the Cape flora have gradually accumulated the type of species-level data needed to test hypotheses of how pollinators are driving speciation (van der Niet & Johnson, 2009; van der Niet, Peakall & Johnson, 2014), so this is possible for seed dispersal ecology too. My hope is that the findings presented in this thesis can contribute to this immense endeavour of discovering how seed dispersal is involved in the generation of the incredible floral diversity of the Cape, and aid in the conservation of its species.

References

Anderson, B., Allsopp, N., Ellis, A.G., Johnson, S.D., Midgley, J.J., Pauw, A. & Rodger, J. 2014. Biotic interactions. In *Fynbos: Ecology, Evolution, and Conservation of a Megadiverse Region*. Oxford: Oxford University Press. 224–247.

Bailey, L.A., Potts, A.J., Cowling, R.M., Whitfield, M.C. & Smit, B. 2019. Wand plant architecture in the Fynbos: Testing the rodent herbivory hypothesis. *South African Journal of Botany*. 124:564–572. DOI: 10.1016/j.sajb.2019.03.035.

Barraclough, T.G. 2006. What can phylogenetics tell us about speciation in the Cape flora? *Diversity and Distributions*. 12(1):21–26. DOI: 10.1111/j.1366-9516.2006.00208.x.

Barraclough, T.G. & Nee, S. 2001. Phylogenetics and speciation. *Trends in Ecology & Evolution*. 16(7):391–399. DOI: 10.1016/S0169-5347(01)02161-9.

Bond, W.J. 1994. Do mutualisms matter? Assessing the impact of pollinator and disperser disruption on plant extinction. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*. 344(1307):83–90. DOI: 10.1098/rstb.1994.0055.

Bond, A.W. & Slingsby, P. 1984. Collapse of an ant-plant mutualism: the Argentine ant (*Iridomyrmex humilis*) and myrmecochorous Proteaceae. *Ecology*. 65(4):1031–1037.

Bond, W.J. & Slingsby, P. 1983. Seed dispersal by ants in shrublands of the Cape Province and its evolutionary implications. *South African Journal of Science*. 79(6):231–233.

Cain, M.L., Milligan, B.G. & Strand, A.E. 2000. Long-distance seed dispersal in plant populations. *American Journal of Botany*. 87(9):1217–1227. DOI: 10.2307/2656714.

Christian, C.E. 2001. Consequences of a biological invasion reveal the importance of mutualism for plant communities. *Nature*. 413(6856):635–639. DOI: 10.1038/35098093.

Christian, C.E. & Stanton, M.L. 2004. Cryptic consequences of a dispersal mutualism: seed burial, elaiosome removal, and seed-bank dynamics. *Ecology*. 85(4):1101–1110. DOI: 10.1890/03-0059.

Cowling, R.M. 1987. Fire and its role in coexistence and speciation in Gondwanan shrublands. *South African Journal of Science*. 83:106–112.

Cowling, R.M. & Holmes, P.M. 1992. Endemism and speciation in a lowland flora from the Cape Floristic Region. *Biological Journal of the Linnean Society*. 47(4):367–383. DOI: 10.1111/j.1095-8312.1992.tb00675.x.

Cowling, R.M., Pierce, S.M., Stock, W.D. & Cocks, M. 1994. Why are there so many myrmecochorous species in the Cape fynbos? In *Plant-Animal Interactions in Mediterranean-type Ecosystems*. M. Arianoutsou & R.H. Groves, Eds. (Tasks for Vegetation Science no. 31). Dordrecht: Kluwer Academic Publishers. 159–168.

Cramer, M.D., West, A.G., Power, S.C., Skelton, R. & Stock, W.D. 2014. Plant ecophysiological diversity. In *Fynbos: Ecology, Evolution, and Conservation of a Megadiverse Region*. Oxford: Oxford University Press. 248–272.

Ellis, A.G., Verboom, G.A., van der Niet, T., Johnson, S.D. & Linder, H.P. 2014. Speciation and extinction in the Greater Cape Floristic Region. In *Fynbos: Ecology, Evolution, and Conservation of a Megadiverse Region*. N. Allsopp, J.F. Colville, & G.A. Verboom, Eds. Oxford: Oxford University Press. 119–141.

Emmett, A.M., Scott, J.K., Webber, B.L., Severn-Ellis, A.A. & Bell, K.L. 2023. Combining multiple lines of evidence to elucidate the origin and introduction pathway of bitou bush (*Chrysanthemoides*

monilifera subsp. *rotundata*) in Australia. *Biological Invasions*. 25(6):1881–1905. DOI: 10.1007/s10530-023-03017-3.

Freudenstein, J.V., Broe, M.B., Folk, R.A. & Sinn, B.T. 2017. Biodiversity and the species concept - lineages are not enough. *Systematic Biology*. 66(4):644–656. DOI: 10.1093/sysbio/syw098.

Givnish, T.J. & Sytsma, K.J. 1997. Homoplasy in molecular vs. morphological data: the likelihood of correct phylogenetic inference. In *Molecular evolution and adaptive radiation*. T.J. Givnish & K.J. Sytsma, Eds. Cambridge: Cambridge University Press. 55–101.

Goldblatt, P. 1997. Floristic diversity in the Cape flora of South Africa. *Biodiversity and Conservation*. 6:359–377.

Gómez, C. & Oliveras, J. 2003. Can the Argentine ant (*Linepithema humile* Mayr) replace native ants in myrmecochory? *Acta Oecologica*. 24(1):47–53. DOI: 10.1016/S1146-609X(03)00042-0.

Gosper, C.R. 2004. Fruit characteristics of invasive bitou bush, *Chrysanthemoides monilifera* (Asteraceae), and a comparison with co-occurring native plant species. *Australian Journal of Botany*. 52(2):223–230. DOI: 10.1071/BT03046.

Grobler, B.A. & Cowling, R.M. 2021. The composition, geography, biology and assembly of the coastal flora of the Cape Floristic Region. *PeerJ*. 9:e11916. DOI: 10.7717/peerj.11916.

Işik, K. 2011. Rare and endemic species: why are they prone to extinction? *Turkish Journal of Botany*. 35(4):411–417. DOI: 10.3906/bot-1012-90.

Johnson, S.D. 1992. Plant-animal relationships. In *The Ecology of Fynbos*. R.M. Cowling, Ed. Cape Town: Oxford University Press. 175–205.

Kemper, J., Cowling, R.M. & Richardson, D.M. 1999. Fragmentation of South African renosterveld shrublands: effects on plant community structure and conservation implications. *Biological Conservation*. 90:103–111.

Kusy, D., Motyka, M., Bocek, M., Masek, M. & Bocak, L. 2019. Phylogenomic analysis resolves the relationships among net-winged beetles (Coleoptera: Lycidae) and reveals the parallel evolution of morphological traits. *Systematic Entomology*. 44(4):911–925. DOI: 10.1111/syen.12363.

Latimer, A.M., Silander, J.A. & Cowling, R.M. 2005. Neutral ecological theory reveals isolation and rapid speciation in a biodiversity hot spot. *Science*. 309(5741):1722–1725. DOI: 10.1126/science.1115576.

Lee, S., Kim, B. & Kim, Y. 2018. Genetic diagnosis of a rare myrmecochorous species, *Plagiorhegma dubium* (Berberidaceae): Historical genetic bottlenecks and strong spatial structures among populations. *Ecology and Evolution*. 8(17):8791–8802. DOI: 10.1002/ece3.4362.

Lengyel, S., Gove, A.D., Latimer, A.M., Majer, J.D. & Dunn, R.R. 2009. Ants sow the seeds of global diversification in flowering plants. *PLoS ONE*. 4(5):e5480. DOI: 10.1371/journal.pone.0005480.

Levine, N., Ben-Zvi, G., Seifan, M. & Giladi, I. 2019. Investment in reward by ant-dispersed plants consistently selects for better partners along a geographic gradient. *AoB PLANTS*. 11(3):plz027. DOI: 10.1093/aobpla/plz027.

Linder, H.P. 1985. Gene flow, speciation, and species diversity patterns in a species rich area: the Cape flora. In *Species and Speciation*. V. 4. E.S. Vrba, Ed. (Transvaal Museum Monograph). Pretoria: Transvaal Museum. 53–57.

Lombard, A.T., Cowling, R.M., Pressey, R.L. & Mustart, P.J. 1997. Reserve selection in a species-rich and fragmented landscape on the Agulhas Plain, South Africa. *Conservation Biology*. 11(5):1101–1116. DOI: 10.1046/j.1523-1739.1997.96043.x.

Maeyama, T. 2000. Genetic relationship of myrmecophyte (*Anthorrhiza caerulea*) individuals within and among territories of the arboreal ant (*Dolichoderus* sp.) detected using random amplified polymorphic DNA markers. *Austral Ecology*. 25(3):273–282. DOI: 10.1046/j.1442-9993.2000.01034.x.

Manning, J.C. & Goldblatt, P. 2012. *Plants of the Greater Cape Floristic Region: the core Cape flora*. (Strelitzia no. 29). Pretoria: South African National Biodiversity Institute.

Matthies, D., Bräuer, I., Maibom, W. & Tscharntke, T. 2004. Population size and the risk of local extinction: empirical evidence from rare plants. *Oikos*. 105(3):481–488. DOI: 10.1111/j.0030-1299.2004.12800.x.

Midgley, J.J. & Anderson, B. 2005. Scatterhoarding in Mediterranean shrublands of the SW Cape, South Africa. In *Seed fate: predation, dispersal and seedling establishment*. P.-M. Forget, J.E. Lambert, P.E. Hulme, & S.B. Vander Wall, Eds. Wallingford UK: CAB International. 197–204.

Midgley, J., Anderson, B., Bok, A. & Fleming, T. 2002. Scatter-hoarding of Cape Proteaceae nuts by rodents. *Evolutionary Ecology Research*. 4:623–626.

Milewski, A.V. & Bond, W.J. 1982. Convergence of myrmecochory in mediterranean Australia and South Africa. In *Ant-plant interactions in Australia*. R.C. Buckley, Ed. The Hague: Dr W. Junk Publishers. 89–98. DOI: 10.1007/978-94-009-7994-9_9.

Mokotjomela, T.M., Musil, C.F. & Esler, K.J. 2013. Potential seed dispersal distances of native and non-native fleshy fruiting shrubs in the South African Mediterranean climate region. *Plant Ecology*. 214(9):1127–1137. DOI: 10.1007/s11258-013-0237-3.

Moncrieff, G.R. 2021. Locating and dating land cover change events in the renosterveld, a critically endangered shrubland ecosystem. *Remote Sensing*. 13(5):834. DOI: 10.3390/rs13050834.

Mucina, L. & Rutherford, M.C. 2006. *The vegetation of South Africa, Lesotho and Swaziland*. (Strelitzia no. 19). Pretoria: South African National Biodiversity Institute.

Nakagawa, M. 2010. Fine-scale genetic structure within plots of *Polygala reinii* (Polygalaceae) having an ant-dispersal seed. *Journal of Plant Research*. 123(3):355–362. DOI: 10.1007/s10265-009-0276-y.

Newton, I.P. & Knight, R.S. 2010. How homogeneous is West Coast Renosterveld? Implications for conservation. *Bothalia*. 40(2):219–226. DOI: 10.4102/abc.v40i2.223.

van der Niet, T. & Johnson, S.D. 2009. Patterns of plant speciation in the Cape floristic region. *Molecular Phylogenetics and Evolution*. 51:85–93.

van der Niet, T., Peakall, R. & Johnson, S.D. 2014. Pollinator-driven ecological speciation in plants: new evidence and future perspectives. *Annals of Botany*. 113(2):199–212. DOI: 10.1093/aob/mct290.

Nordenstam, B. 2004. Two new species of *Osteospermum* (Compositae–Calenduleae) from southwestern Cape Province, South Africa. *Edinburgh Journal of Botany*. 60(3):259–265. DOI: 10.1017/S0960428603000234.

Norlindh, T. 1943. *Studies in the Calenduleae I: Monograph of the genera Dimorphotheca, Castalis, Osteospermum, Gibbaria and Chrysanthemoides*. Lund: C.W.K. Gleerup.

Pascov, C.M., Nevill, P.G., Elliott, C.P., Majer, J.D., Anthony, J.M. & Krauss, S.L. 2015. The critical role of ants in the extensive dispersal of *Acacia* seeds revealed by genetic parentage assignment. *Oecologia*. 179(4):1123–1134. DOI: 10.1007/s00442-015-3400-9.

de Queiroz, K. 2007. Species concepts and species delimitation. *Systematic Biology*. 56(6):879–886. DOI: 10.1080/10635150701701083.

Rouhan, G. & Gaudeul, M. 2014. Plant taxonomy: a historical perspective, current challenges, and perspectives. In *Molecular Plant taxonomy: Methods and Protocols*. P. Besse, Ed. (Methods in Molecular Biology no. 1115). New York: Springer. 1–37.

Rusch, U.D., Midgley, J.J. & Anderson, B. 2013. Rodent consumption and caching behaviour selects for specific seed traits. *South African Journal of Botany*. 84:83–87. DOI: 10.1016/j.sajb.2012.09.007.

Sadler, R.T. 2024. Phylogeny, generic limits and evolution of dispersal traits in the Asteraceae tribe Calenduleae. MSc thesis. University of Cape Town.

Sadler, R., Parker, T., Verboom, G.A., Ellis, A.G., Jackson, M., Van Zyl, J., Manning, J. & Bergh, N.G. 2022. A molecular phylogeny of Calenduleae (Asteraceae) supports the transfer of *Dimorphotheca polyptera* DC. to *Osteospermum* L. *South African Journal of Botany*. 151:234–245. DOI: 10.1016/j.sajb.2022.09.001.

Scott, J.K. & Batchelor, K.L. 2014. Management of *Chrysanthemoides monilifera* subsp. *rotundata* in Western Australia. *Invasive Plant Science and Management*. 7(1):190–196. DOI: 10.1614/IPSM-D-13-00052.1.

Slingsby, P. & Bond, W.J. 1981. Ants — Friends of the Fynbos. *Veld and Flora*. 67:39–45.

Stock, S. 2021. Secondary dispersal of winged fruit in the daisy tribe Calenduleae: the effect of wing number and wing-loading. Honours thesis. University of Cape Town.

Topp, E.N. & Loos, J. 2019. Fragmented landscape, fragmented knowledge: a synthesis of renosterveld ecology and conservation. *Environmental Conservation*. 46(2):171–179. DOI: 10.1017/S0376892918000498.

Travis, J.M.J. & Dytham, C. 2002. Dispersal evolution during invasions. *Evolutionary Ecology Research*. 4:1119–1129.

Van Damme, K., Cornetti, L., Fields, P.D. & Ebert, D. 2022. Whole-genome phylogenetic reconstruction as a powerful tool to reveal homoplasy and ancient rapid radiation in waterflea evolution. *Systematic Biology*. 71(4):777–787. DOI: 10.1093/sysbio/syab094.

Westoby, M., Rice, B., Shelley, J., Haig, D. & Kohen, J. 1982. Plant's use of ants for dispersal at West Head, New South Wales. In *Ant-plant interactions in Australia*. R. Buckley, Ed. The Hague: Dr W. Junk Publishers. 75–87.

White, J.D., Bronner, G.N. & Midgley, J.J. 2017. Camera-trapping and seed-labelling reveals widespread granivory and scatter-hoarding of nuts by rodents in the Fynbos Biome. *African Zoology*. 52(1):31–41. DOI: 10.1080/15627020.2017.1292861.

Willis, C.K., Cowling, R.M. & Lombard, A.T. 1996. Patterns of endemism in the limestone flora of South African lowland fynbos. *Biodiversity and Conservation*. 5(1):55–73. DOI: 10.1007/BF00056292.

Wilson, J.R.U., Dormontt, E.E., Prentis, P.J., Lowe, A.J. & Richardson, D.M. 2009. Something in the way you move: dispersal pathways affect invasion success. *Trends in Ecology & Evolution*. 24(3):136–144. DOI: 10.1016/j.tree.2008.10.007.

Yeates, D.K., Seago, A., Nelson, L., Cameron, S.L., Joseph, L. & Trueman, J.W.H. 2011. Integrative taxonomy, or iterative taxonomy? *Systematic Entomology*. 36(2):209–217. DOI: 10.1111/j.1365-3113.2010.00558.x.

Zhang, X., Ge, J., Yang, J., Dunn, B. & Chen, G. 2017. Genetic diversity of *Stemona parviflora*: A threatened myrmecochorous medicinal plant in China. *Biochemical Systematics and Ecology*. 71:193–199. DOI: 10.1016/j.bse.2017.02.012.

Appendices

Appendix 1. List of collection details for each population sampled for this study and the datasets in which they were used. Note that for the datasets indicated "GBS" refers to both GBS phylogenetic and population genetic datasets used for species delimitation, while "Sanger" refers to phylogenetic inference using sequences from the ITS and ETS gene regions to test for monophyly. In the collection localities, "Nature Reserve" is abbreviated as "N.R.".

Clade	Lineage	Population code	GBS voucher collection number	Current taxon name	Collection locality	Elevation (m)	Latitude (degrees)	Longitude (degrees)	Date collected	Datasets
1	A	POL25	LMCJ292	*O. polygaloides*	Swartberg pass	1614	-33.3526	22.05305	08-Jan-23	GBS, Morphological, Vegetation type, Sanger
1	B	IMB18	LMCJ293	*O. imbricatum*	Swartberg pass	1414	-33.3288	22.04199	08-Jan-23	GBS, Morphological, Vegetation type, Sanger
1	C	NER1	LMCJ90	*O. imbricatum* subsp. *nervatum*	Hankey	145	-33.84	24.92876	02-Sep-22	GBS, Morphological, Vegetation type
1	C	NER2	LMCJ102	*O. imbricatum* subsp. *nervatum*	Howieson's Poort	479	-33.3623	26.49861	04-Sep-22	GBS, Morphological, Vegetation type, Sanger
1	C	NER3	SM623	*O. imbricatum* subsp. *nervatum* var. *helichrysoides*	Mtentu River Mouth	38	-31.2389	30.05327	12-Jul-22	GBS, Morphological, Vegetation type, Sanger
1	D	IMB9	LMCJ283	*O. imbricatum*	Towerkop	1197	-33.4414	21.19924	07-Jan-23	GBS, Morphological, Vegetation type, Sanger
1	D	IMB15	LMCJ310	*O. imbricatum*	Rooiberg	908	-33.6901	21.60449	12-Jan-23	GBS, Morphological, Vegetation type
1	D	IMB17	LMCJ307	*O. imbricatum*	Gamkaberg	891	-33.7053	21.92342	11-Jan-23	GBS, Morphological, Vegetation type
1	D	IMB19	LMCJ190	*O. imbricatum/polygaloides*	Mossel Bay	168	-34.196	22.11004	07-Oct-22	GBS, Morphological, Vegetation type, Sanger
1	D	IMB21	LMCJ81	*O. imbricatum*	Sedgefield	179	-33.9728	22.78332	30-Aug-22	GBS, Morphological, Vegetation type
1	D	IMB23	LMCJ196	*O. imbricatum*	Prince Alfred's Pass	454	-33.8666	23.16927	07-Oct-22	GBS, Morphological, Vegetation type, Sanger
1	D	IMB24	LMCJ301	*O. imbricatum*	Kouga mountains W	1119	-33.6016	23.41133	09-Jan-23	GBS, Morphological, Vegetation type
1	D	IMB25	LMCJ298	*O. imbricatum*	Kouga mountains W	1178	-33.6074	23.4147	09-Jan-23	GBS, Morphological, Vegetation type, Sanger
1	D	IMB26	LMCJ304	*O. imbricatum*	Kouga mountains Langkloof	792	-33.7808	23.7195	10-Jan-23	GBS, Morphological, Vegetation type

1	D	IMB27	LMCJ105	*O. imbricatum*	Kouga mountains central	592	-33.7092	23.83229	06-Sep-22	GBS, Morphological, Vegetation type, Sanger
1	D	IMB28	LMCJ91	*O. imbricatum*	Elandsberge	666	-33.6591	24.79287	02-Sep-22	GBS, Morphological, Vegetation type, Sanger
1	D	IMB29	LMCJ87	*O. imbricatum*	Van Stadens Mountains	470	-33.8894	25.24956	01-Sep-22	GBS, Morphological, Vegetation type, Sanger
1	D	IMB30	LMCJ95	*O. imbricatum*	Schoenmakerskop	40	-34.0409	25.55615	03-Sep-22	GBS, Morphological, Vegetation type, Sanger
1	D	IMB31	LMCJ99	*O. imbricatum*	Coega	48	-33.7672	25.70136	03-Sep-22	GBS, Morphological, Vegetation type
2	E	IMB8	LMCJ157	*O. imbricatum*	Gysmanshoek Pass	677	-33.9294	21.07485	03-Oct-22	GBS, Morphological, Vegetation type
2	E	IMB11	LMCJ200	*O. imbricatum*	Btw Garcia's Pass and Muiskraal	535	-33.9342	21.20864	19-Oct-22	GBS, Morphological, Vegetation type, Sanger
2	F	BUR1	LMCJ279	*O. burttianum*	Protea Valley	1047	-33.9465	20.37597	06-Jan-23	GBS, Morphological, Vegetation type
2	F	COR1	LMCJ151	*O. corymbosum*	Marloth N.R.	526	-33.9891	20.45081	01-Oct-22	GBS, Morphological, Vegetation type
2	F	BUR2	LMCJ76	*O. burttianum*	Op de Tradouw	1032	-33.9591	20.5725	12-Feb-22	GBS, Morphological, Vegetation type, Sanger
2	F	COR2	LMCJ78	*O. corymbosum*	Op de Tradouw	895	-33.9583	20.58047	12-Feb-22	GBS, Morphological, Vegetation type, Sanger
2	F	COR3	LMCJ153	*O. corymbosum*	Tradouw Pass	342	-33.977	20.70536	02-Oct-22	GBS, Morphological, Vegetation type
2	F	COR4	LMCJ202	*O. corymbosum*	Grootvadersbosch N.R.	428	-33.9789	20.83548	25-Oct-22	GBS, Morphological, Vegetation type
2	F	BUR3	LMCJ317	*O. burttianum*	Boosmansbos Wilderness Area	1330	-33.9341	20.90184	04-Mar-23	GBS, Morphological, Vegetation type
2	F	COR5	LMCJ169	*O. corymbosum*	Garcia's Pass	543	-33.9532	21.22329	03-Oct-22	GBS, Morphological, Vegetation type, Sanger
2	F	COR6	LMCJ180	*O. corymbosum*	Romanskraal	487	-33.983	21.4618	05-Oct-22	GBS, Morphological, Vegetation type, Sanger
2	F	COR7	LMCJ188	*O. corymbosum*	Cloetesberg	752	-33.9047	21.74899	06-Oct-22	GBS, Morphological, Vegetation type
2	F	COR8	LMCJ197	*O. corymbosum*	Outeniqua Pass	605	-33.9067	22.40621	08-Oct-22	GBS, Morphological, Vegetation type
3	G	IMB3	LMCJ276	*O. imbricatum* var. *microcephalum*	Keurkloof	828	-33.8181	20.1273	05-Jan-23	GBS, Morphological, Vegetation type, Sanger

3	G	IMB4	LMCJ325	*O. imbricatum* var. *microcephalum*	Klipspringerkop	506	-33.9102	20.3111	10-Jun-23	GBS, Morphological, Vegetation type
3	H	IMB1	LMCJ144	*O. imbricatum/polygaloides*	Swartruggens mountain	1191	-32.8769	19.61005	23-Sep-22	GBS, Morphological, Vegetation type, Sanger
3	I	POL5	LMCJ139	*O. polygaloides/imbricatum*	Aurora	80	-32.7683	18.50371	22-Sep-22	GBS, Morphological, Vegetation type, Sanger
3	I	POL6	LMCJ136	*O. polygaloides/imbricatum*	Sauer	54	-32.8328	18.53687	22-Sep-22	GBS, Morphological, Vegetation type
3	I	POL9	LMCJ134	*O. polygaloides/imbricatum*	Btw Paleisheuwel and Citrusdal N	686	-32.4474	18.86514	21-Sep-22	GBS, Morphological, Vegetation type
3	I	POL10	LMCJ129	*O. polygaloides/imbricatum*	Btw Paleisheuwel and Citrusdal	661	-32.5267	18.90011	21-Sep-22	GBS, Morphological, Vegetation type
3	I	POL11	LMCJ126	*O. polygaloides/imbricatum*	Btw Paleisheuwel and Citrusdal S	668	-32.5713	18.91533	21-Sep-22	GBS, Morphological, Vegetation type
3	J	IMB2	LMCJ226	*O. imbricatum*	Elim N	84	-34.5626	19.74758	07-Nov-22	GBS, Morphological, Vegetation type
3	J	IMB5	LMCJ150	*O. imbricatum*	Bontebok National Park	109	-34.0621	20.4309	30-Sep-22	GBS, Morphological, Vegetation type, Sanger
3	J	IMB6	LMCJ146	*O. imbricatum/polygaloides*	Bontebok National Park	92	-34.0607	20.46409	30-Sep-22	GBS, Morphological, Vegetation type
3	J	IMB7	LMCJ158	*O. imbricatum*	Gysmanshoek Pass	507	-33.9212	21.06006	03-Oct-22	GBS, Morphological, Vegetation type
3	J	IMB10	LMCJ286	*O. imbricatum*	Towerkop	626	-33.4607	21.20184	07-Jan-23	GBS, Morphological, Vegetation type
3	J	IMB12	LMCJ165	*O. imbricatum*	Btw Garcia's Pass and Muiskraal	536	-33.9347	21.20874	03-Oct-22	GBS, Morphological, Vegetation type
3	J	IMB13	LMCJ170	*O. imbricatum*	Riversdale	165	-34.1205	21.24196	04-Oct-22	GBS, Morphological, Vegetation type
3	J	IMB14	LMCJ179	*O. imbricatum*	Hannes Zaaiman Private N.R.	59	-34.2707	21.2962	04-Oct-22	GBS, Morphological, Vegetation type
3	J	IMB16	LMCJ288	*O. imbricatum*	Groenfontein rd	667	-33.4339	21.83507	08-Jan-23	GBS, Morphological, Vegetation type
3	J	IMB20	LMCJ113	*O. imbricatum*	Kammanassie	667	-33.6977	22.66084	07-Sep-22	GBS, Morphological, Vegetation type
3	J	IMB22	LMCJ108	*O. imbricatum*	Kammanassie	574	-33.7125	22.82494	07-Sep-22	GBS, Morphological, Vegetation type, Sanger
4	K	POL12	LMCJ123	*O. polygaloides*	Nieuwoudts Pass	498	-32.3524	19.00411	21-Sep-22	GBS, Morphological, Vegetation type, Sanger

4	K	POL14	NGB2603	*O. polygaloides*	Hexriver Mountains	1000	-33.441	19.44082	09-Jan-22	GBS, Morphological, Vegetation type
4	K	POL15	LMCJ207	*O. polygaloides*	Jonaskop	986	-33.9419	19.52157	03-Nov-22	GBS, Morphological, Vegetation type, Sanger
4	K	POL17	LMCJ272	*O. polygaloides*	Boesmanskloof	773	-34.007	19.71403	17-Dec-22	GBS, Morphological, Vegetation type, Sanger
4	L	POL18	LMCJ232	*O. polygaloides/imbricatum*	Elim E	36	-34.5954	19.8198	08-Nov-22	GBS, Morphological, Vegetation type
4	M	POL20	LMCJ230	*O. polygaloides*	Bredasdorp Heuningrug	68	-34.6305	19.94797	08-Nov-22	GBS, Morphological, Vegetation type, Sanger
4	M	POL23	LMCJ213	*O. polygaloides*	De Hoop N.R.	218	-34.4335	20.66943	05-Nov-22	GBS, Morphological, Vegetation type
4	M	POL24	LMCJ174	*O. polygaloides*	Hannes Zaaiman Private N.R.	117	-34.2661	21.29043	04-Oct-22	GBS, Morphological, Vegetation type
4	N	AUS1	LMCJ236	*O. australe*	Pearly Beach N	158	-34.6264	19.53732	08-Nov-22	GBS, Morphological, Vegetation type
4	N	AUS2	LMCJ70	*O. australe*	Pearly Beach E	115	-34.6546	19.56702	19-Nov-21	GBS, Morphological, Vegetation type, Sanger
4	N	AUS3	LMCJ222	*O. australe*	Soetanysberg	235	-34.751	19.85186	07-Nov-22	GBS, Morphological, Vegetation type
4	O	POL1	LMCJ320	*O. polygaloides*	Table Mountain Oudekraal	73	-33.9846	18.36111	16-Apr-23	GBS, Morphological, Vegetation type
4	O	POL2	LMCJ266	*O. polygaloides*	Table Mountain Pipe Track	296	-33.9561	18.39292	26-Nov-22	GBS, Morphological, Vegetation type, Sanger
4	O	POL3	LMCJ262	*O. polygaloides*	Silvermine	433	-34.085	18.41039	25-Nov-22	GBS, Morphological, Vegetation type
4	O	POL4	LMCJ268	*O. polygaloides*	Smitswinkle flats	83	-34.2655	18.43019	13-Dec-22	GBS, Morphological, Vegetation type
4	O	POL7	LMCJ64	*O. polygaloides*	Gordon's Bay	42	-34.1867	18.82024	01-Nov-21	GBS, Morphological, Vegetation type, Sanger
4	O	POL8	LMCJ255	*O. polygaloides*	Hangklip	52	-34.3561	18.83985	23-Nov-22	GBS, Morphological, Vegetation type
4	O	POL13	LMCJ246	*O. polygaloides*	Franskraal se Berg	263	-34.5808	19.39862	09-Nov-22	GBS, Morphological, Vegetation type
4	O	ROT6	LMCJ239	*O. rotundifolium*	Baardskeerdersbos	105	-34.5841	19.52538	09-Nov-22	GBS, Morphological, Vegetation type
4	O	POL16	LMCJ248	*O. polygaloides/imbricatum*	Akkedisberg Pass	141	-34.4086	19.58124	10-Nov-22	GBS, Morphological, Vegetation type

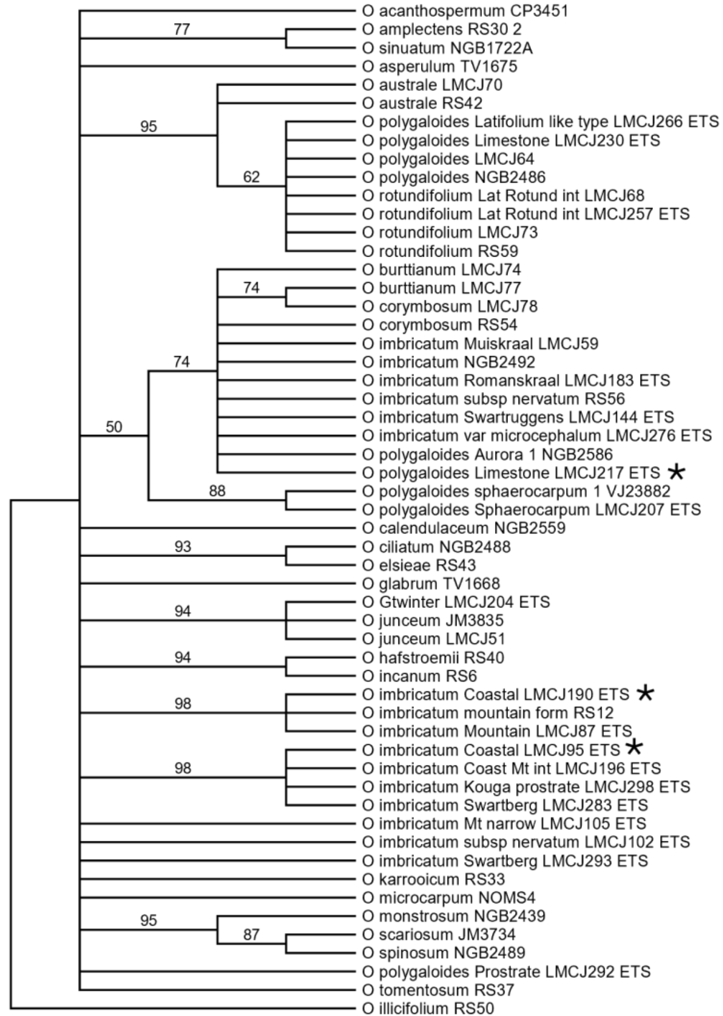
4	O	POL19	LMCJ225	*O. polygaloides*	Napier	280	-34.497	19.8931	07-Nov-22	GBS, Morphological, Vegetation type
4	O	POL21	LMCJ215	*O. polygaloides*	Bredasdorp Heuningberg N.R.	202	-34.5407	20.03212	06-Nov-22	GBS, Morphological, Vegetation type
4	O	POL22	LMCJ217	*O. polygaloides*	Struisbaai	28	-34.8084	20.04685	06-Nov-22	GBS, Morphological, Vegetation type, Sanger
4	P	ROT1	LMCJ61	*O. rotundifolium*	Betty's Bay	70	-34.3476	18.92931	01-Nov-21	GBS, Morphological, Vegetation type
4	P	ROT2	LMCJ257	*O. rotundifolium*	Kogelberg	42	-34.3147	18.95736	24-Nov-22	GBS, Morphological, Vegetation type, Sanger
4	P	ROT3	LMCJ332	*O. rotundifolium*	Steenbras N.R.	512	-34.1836	18.95744	13-Jul-23	GBS, Morphological, Vegetation type
4	P	ROT4	LMCJ250	*O. rotundifolium*	Fernkloof N.R.	99	-34.3963	19.27693	10-Nov-22	GBS, Morphological, Vegetation type
4	P	ROT5	LMCJ72	*O. rotundifolium*	Vogelgat N.R.	36	-34.3978	19.31456	20-Nov-21	GBS, Morphological, Vegetation type, Sanger
NA	NA	Outgroup	LMCJ271	*O. junceum*	Greyton-McGregor trail	601	-34.0151	19.71055	17-Dec-22	GBS
NA	NA		LMCJ204	*O. sp. nov.*	Groot Winterhoek Mountains (E.C.)		-33.6262	25.23435	19-Oct-22	Sanger

Appendix 2. Samples used for Bayesian phylogenetic analysis to test for monophyly of *O*. sect. *Polygalinae*, including preliminary morphological form assignment for specimens putatively in *O*. sect. *Polygalinae*, collection details and the gene regions sampled. Sequences were either sourced from R. Sadler (Unpublished MSc, 2024) or were newly generated in this study.

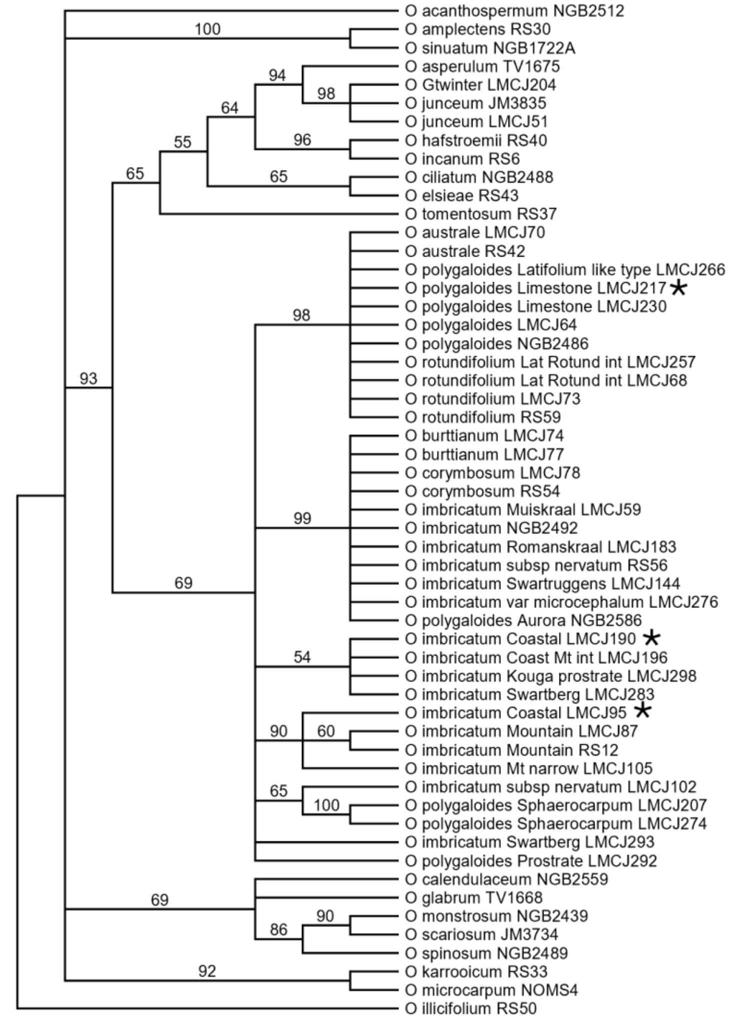
Species	Expected placement	Form	Voucher collection number	Population code	Collection locality	Gene region		Sequence source
O. amplectens	Outgroup		RS30			ETS	ITS	R. Sadler
O. asperulum	Outgroup		TV1675			ETS	ITS	R. Sadler
O. australe	*O*. sect. *Polygalinae*	Australe	LMCJ70	AUS2	Pearly Beach E	ETS	ITS	R. Sadler
O. australe	*O*. sect. *Polygalinae*	Australe	RS42	AUS2	Pearly Beach E	ETS	ITS	R. Sadler
O. burttianum	*O*. sect. *Polygalinae*	Burttianum	LMCJ74	BUR2	Op de Tradouw	ETS	ITS	R. Sadler
O. burttianum	*O*. sect. *Polygalinae*	Burttianum	LMCJ77	BUR2	Op de Tradouw	ETS	ITS	R. Sadler
O. calendulaceum	Outgroup		NGB2559			ETS	ITS	R. Sadler
O. ciliatum	Outgroup		NGB2488			ETS	ITS	R. Sadler
O. corymbosum	*O*. sect. *Polygalinae*	Corymbosum	LMCJ78	COR2	Op de Tradouw	ETS	ITS	R. Sadler
O. corymbosum	*O*. sect. *Polygalinae*	Corymbosum	RS54	COR5	Garcia's Pass	ETS	ITS	R. Sadler
O. elsieae	Outgroup		RS43			ETS	ITS	R. Sadler
O. glabrum	Outgroup		TV1668			ETS	ITS	R. Sadler
O. sp. nov.	*O*. sect. *Polygalinae*	Grootwinterhoek	LMCJ204	NA	Groot Winterhoek Mountains (E.C.)	ETS	ITS	This study
O. hafstroemii	Outgroup		RS40			ETS	ITS	R. Sadler
O. ilicifolium	Outgroup		RS50			ETS	ITS	R. Sadler
O. imbricatum	*O*. sect. *Polygalinae*	Classic imbricatum	NGB2492	IMB5	Bontebok National Park	ETS	ITS	R. Sadler
O. imbricatum	*O*. sect. *Polygalinae*	Classic imbricatum	LMCJ108	IMB22	Kammanassie	-	ITS	This study
O. imbricatum	*O*. sect. *Polygalinae*	Coastal	LMCJ190	IMB19	Mossel Bay	ETS	ITS	This study
O. imbricatum	*O*. sect. *Polygalinae*	Coastal	LMCJ95	IMB30	Schoenmakerskop	ETS	ITS	This study
O. imbricatum	*O*. sect. *Polygalinae*	Coastal-Mountain intermediate	LMCJ196	IMB23	Prince Alfred's Pass	ETS	ITS	This study
O. imbricatum	*O*. sect. *Polygalinae*	Helichrysoides	SM622	NER3	Mtentu River Mouth	-	ITS	This study
O. imbricatum	*O*. sect. *Polygalinae*	Kouga prostrate	LMCJ298	IMB25	Kouga mountains W	ETS	ITS	This study
O. imbricatum	*O*. sect. *Polygalinae*	Microcephalum	LMCJ276	IMB3	Keurkloof	ETS	ITS	This study
O. imbricatum	*O*. sect. *Polygalinae*	Mountain	LMCJ87	IMB29	Van Stadens Mountains	ETS	ITS	This study
O. imbricatum	*O*. sect. *Polygalinae*	Mountain	RS12	IMB28	Elandsberge	ETS	ITS	R. Sadler
O. imbricatum	*O*. sect. *Polygalinae*	Mt narrow	LMCJ105	IMB27	Kouga mountains central	ETS	ITS	This study
O. imbricatum	*O*. sect. *Polygalinae*	Muiskraal	RS56	IMB11	Btw Garcia's Pass and Muiskraal	ETS	ITS	R. Sadler
O. imbricatum	*O*. sect. *Polygalinae*	Muiskraal	LMCJ59	IMB11	Btw Garcia's Pass and Muiskraal	ETS	ITS	R. Sadler

O. imbricatum	*O.* sect. *Polygalinae*	Nervatum	LMCJ102	NER2	Howieson's Poort	ETS	ITS	This study
O. imbricatum	*O.* sect. *Polygalinae*	Romanskraal	LMCJ183	COR6	Romanskraal	ETS	ITS	This study
O. imbricatum	*O.* sect. *Polygalinae*	Swartberg	LMCJ283	IMB9	Towerkop	ETS	ITS	This study
O. imbricatum	*O.* sect. *Polygalinae*	Swartberg	LMCJ293	IMB18	Swartberg pass	ETS	ITS	This study
O. imbricatum	*O.* sect. *Polygalinae*	Swartruggens	LMCJ144	IMB1	Swartruggens mountain	ETS	ITS	This study
O. incanum	Outgroup		RS6			ETS	ITS	R. Sadler
O. junceum	Outgroup		JM3835			ETS	ITS	R. Sadler
O. junceum	Outgroup		LMCJ51			ETS	ITS	R. Sadler
O. karrooicum	Outgroup		RS33			ETS	ITS	R. Sadler
O. microcarpum	Outgroup		NOMS4			ETS	ITS	R. Sadler
O. monstrosum	Outgroup		NGB2439			ETS	ITS	R. Sadler
O. polygaloides	*O.* sect. *Polygalinae*	Aurora	NGB2586	POL5	Aurora	ETS	ITS	R. Sadler
O. polygaloides	*O.* sect. *Polygalinae*	Classic polygaloides	LMCJ64	POL7	Gordon's Bay	ETS	ITS	R. Sadler
O. polygaloides	*O.* sect. *Polygalinae*	Classic polygaloides	NGB2486	NA	Cape Peninsula	ETS	ITS	R. Sadler
O. polygaloides	*O.* sect. *Polygalinae*	Classic polygaloides/ Sphaerocarpum	LMCJ123	POL12	Nieuwoudts Pass	-	ITS	This study
O. polygaloides	*O.* sect. *Polygalinae*	Limestone	LMCJ217	POL22	Struisbaai	ETS	ITS	This study
O. polygaloides	*O.* sect. *Polygalinae*	Limestone	LMCJ230	POL20	Bredasdorp Heuningrug	ETS	ITS	This study
O. polygaloides	*O.* sect. *Polygalinae*	Prostrate	LMCJ292	POL25	Swartberg pass	ETS	ITS	This study
O. polygaloides	*O.* sect. *Polygalinae*	Sphaerocarpum	LMCJ207	POL15	Jonaskop	ETS	ITS	This study
O. polygaloides	*O.* sect. *Polygalinae*	Sphaerocarpum	LMCJ274	POL17	Boesmanskloof	-	ITS	This study
O. polygaloides	*O.* sect. *Polygalinae*	Table Mountain	LMCJ266	POL2	Table Mountain Pipe Track	ETS	ITS	This study
O. rotundifolium	*O.* sect. *Polygalinae*	Classic Rotundifolium	LMCJ68	ROT2	Kogelberg	ETS	ITS	R. Sadler
O. rotundifolium	*O.* sect. *Polygalinae*	Classic Rotundifolium	LMCJ73	ROT5	Vogelgat N.R.	ETS	ITS	R. Sadler
O. rotundifolium	*O.* sect. *Polygalinae*	Classic Rotundifolium	RS59	ROT5	Vogelgat N.R.	ETS	ITS	R. Sadler
O. rotundifolium	*O.* sect. *Polygalinae*	Latifolium-Rotundifolium intermediate	LMCJ257	ROT2	Kogelberg	ETS	ITS	This study
O. scariosum	Outgroup		JM3734			ETS	ITS	R. Sadler
O. sinuatum	Outgroup		NGB1722A			ETS	ITS	R. Sadler
O. spinosum	Outgroup		NGB2489			ETS	ITS	R. Sadler
O. tomentosum	Outgroup		RS37			ETS	ITS	R. Sadler

ETS



ITS



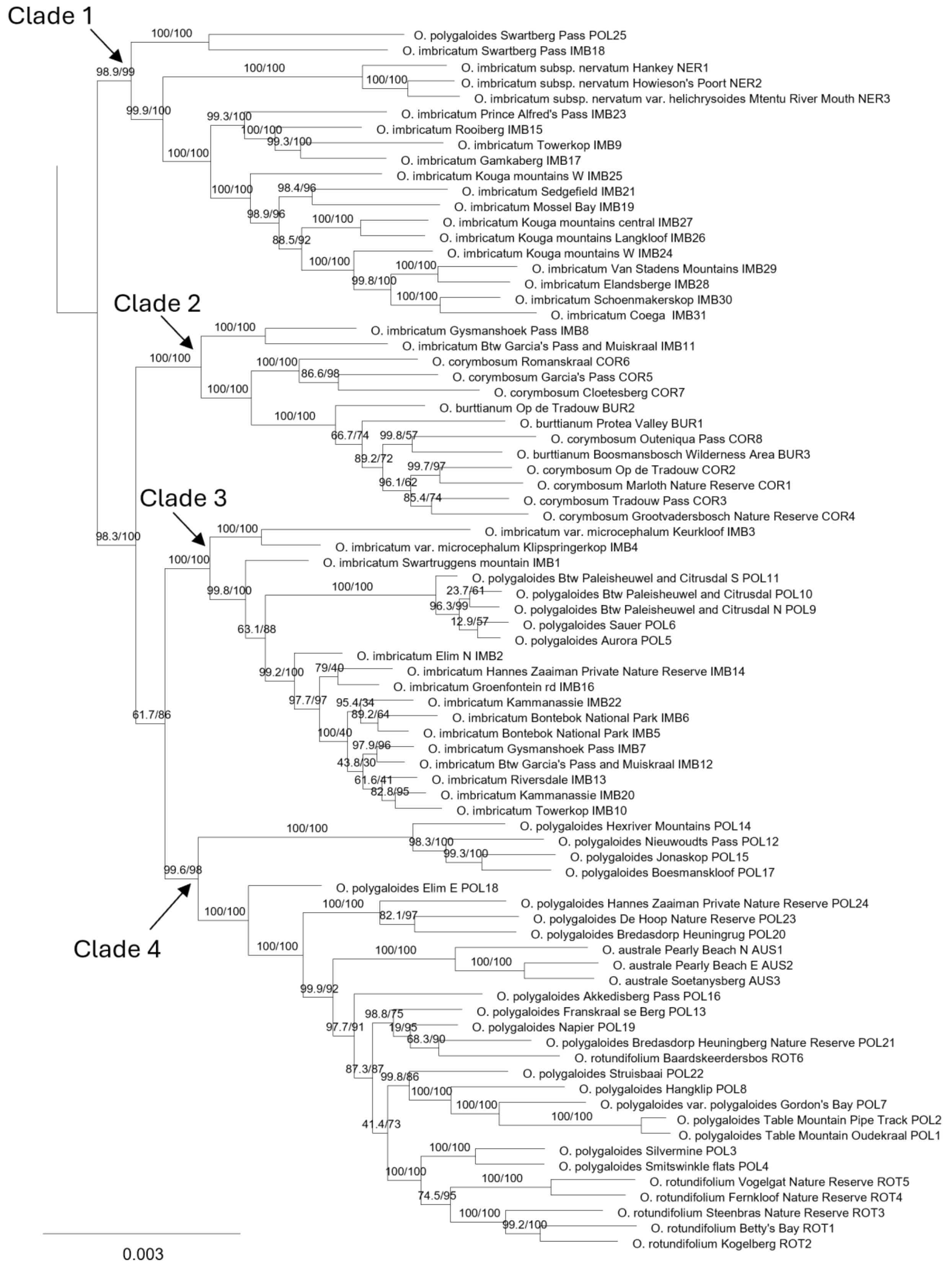
Appendix 3. Bootstrap majority-rule consensus trees from parsimony analysis of ETS (left) and ITS (right) sequences of taxa in *Osteospermum*. Numbers above branches indicate bootstrap support. Asterisks indicate accessions with incongruent placements.

Appendix 4. The eight continuous traits and ratios and six qualitative binary traits scored for each population. Leaf traits and peduncle length were measured on one individual per population and the mean determined from three replicate measurements. Peduncle branching and number of capitula per branch stem were counted on all mature flowering or fruiting individuals sampled, the number of specimens is indicated for each population. Measurement and scoring details are given in Chapter 2; Table 2.

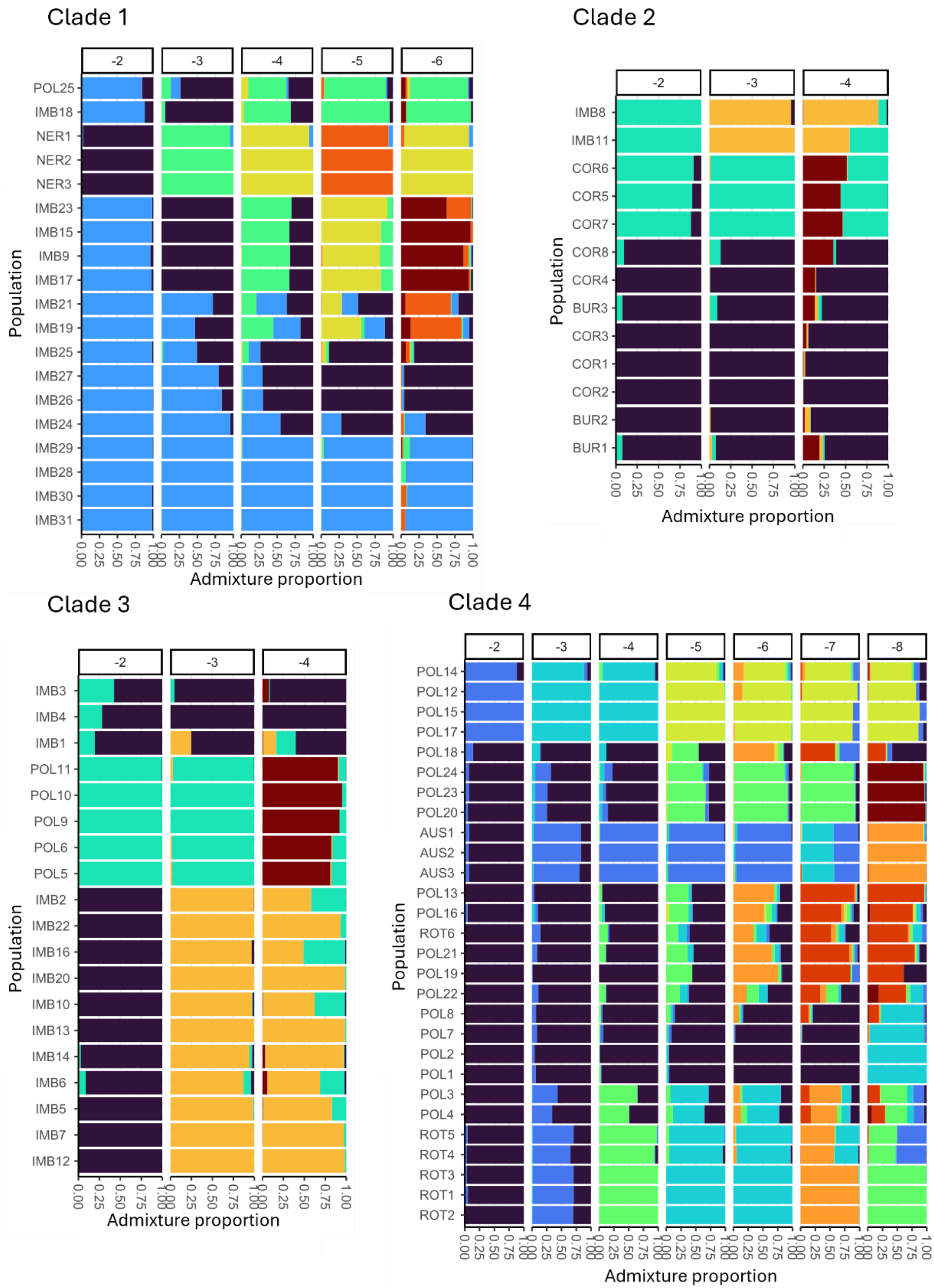
Clade	Lineage	Population code	Number of specimens	Height (cm)	Leaf width (mm)	Leaf length (mm)	Leaf width : length	Relative position of widest point of leaf	Peduncle length (mm)	Peduncle branching	Number of capitula per branch stem	Apical anther appendage colour	Involucral bract trichomes presence	Involucral bract trichome pigmentation	Involucral bract hairs	Fruit surface texture	Fruit ridge edges
1	A	POL25	3	10	2.87	20.55	0.14	0.69	61.45	0.00	1.33	0	1	1	0	1	0
1	B	IMB18	3	150	3.65	11.72	0.31	0.65	21.86	0.67	2.00	1	0	0	0	1	1
1	C	NER1	3	25	4.41	28.56	0.16	0.35	75.69	3.33	4.67	1	1	0	0	1	1
1	C	NER2	5	45	4.46	24.72	0.21	0.68	97.16	2.60	4.20	1	1	0	0	1	1
1	C	NER3	1	50	3.56	24.03	0.15	0.38	62.06	3.00	6.00	1	1	0	0	NA	NA
1	D	IMB9	1	80	3.67	13.47	0.27	0.60	44.80	2.00	4.00	1	1	0	0	NA	NA
1	D	IMB15	3	170	5.13	16.63	0.31	0.59	72.61	2.00	3.00	0	1	1	0	0	1
1	D	IMB17	4	100	5.07	19.02	0.27	0.44	44.09	2.25	4.50	0	1	0	0	1	1
1	D	IMB19	4	70	3.27	11.68	0.28	0.60	57.86	0.50	2.50	1	1	1	0	1	1
1	D	IMB21	3	140	8.38	17.20	0.49	0.61	156.57	3.00	5.33	1	1	1	0	1	1
1	D	IMB23	3	150	11.02	12.50	0.88	0.30	60.10	2.67	5.33	1	1	1	0	1	1
1	D	IMB24	3	200	4.80	16.79	0.29	0.63	77.16	1.33	3.67	0	1	0	0	NA	NA
1	D	IMB25	2	15	3.04	15.15	0.20	0.74	30.73	0.00	1.50	0	1	1	0	NA	NA
1	D	IMB26	3	120	6.86	22.46	0.31	0.58	91.46	0.00	3.00	1	1	0	0	1	1
1	D	IMB27	3	125	8.20	20.21	0.41	0.51	36.28	0.67	3.67	1	1	0	0	1	1
1	D	IMB28	3	169	8.97	24.97	0.36	0.59	99.06	6.67	13.33	0	1	0	0	1	1
1	D	IMB29	3	150	13.56	20.76	0.66	0.57	34.98	1.33	5.33	0	1	0	0	1	1
1	D	IMB30	3	40	7.97	12.00	0.67	0.58	52.72	0.00	2.33	0	1	1	0	1	1
1	D	IMB31	3	32	3.83	11.89	0.33	0.66	99.09	1.00	3.00	1	1	0	0	1	1
2	E	IMB8	3	100	6.32	31.81	0.20	0.11	NA	NA	NA	NA	NA	NA	NA	NA	NA
2	E	IMB11	5	110	2.87	10.57	0.27	0.00	29.39	2.80	4.80	0	1	1	0	1	0
2	F	BUR1	3	50	7.64	48.17	0.16	0.62	120.38	1.33	3.00	0	1	1	0	0	0
2	F	COR1	1	100	12.93	43.26	0.30	0.30	54.51	6.00	9.00	0	1	1	0	NA	NA

2	F	BUR2	4	30	6.14	32.15	0.19	0.62	51.34	4.00	6.50	0	1	1	0	0	0
2	F	COR2	2	200	11.14	32.82	0.34	0.21	40.90	5.00	9.50	0	1	1	0	0	0
2	F	COR3	4	167	9.35	28.22	0.33	0.28	27.44	4.25	7.25	0	1	0	0	0	0
2	F	COR4	3	215	11.71	27.53	0.42	0.42	56.24	4.67	24.33	0	1	0	0	0	0
2	F	BUR3	3	30	10.60	61.89	0.17	0.74	20.72	0.33	1.67	0	1	1	0	0	0
2	F	COR5	3	170	15.14	34.97	0.43	0.34	51.33	6.00	15.67	0	1	0	0	0	0
2	F	COR6	4	150	3.80	14.39	0.28	0.31	33.96	3.00	4.50	0	1	1	0	0	0
2	F	COR7	3	200	11.40	27.87	0.41	0.22	40.50	4.67	9.00	0	1	1	0	0	0
2	F	COR8	3	200	14.79	32.58	0.45	0.24	52.30	3.67	7.00	0	1	0	0	0	0
3	G	IMB3	3	150	13.38	9.91	1.35	0.09	60.20	4.33	9.00	1	1	0	0	1	1
3	G	IMB4	4	70	11.25	8.14	1.38	0.06	45.98	2.00	3.00	1	1	0	0	1	1
3	H	IMB1	3	54	4.68	13.03	0.36	0.48	35.19	0.00	1.00	0	1	0	0	NA	NA
3	I	POL5	3	100	2.99	14.85	0.20	0.69	15.95	0.00	1.00	0	1	0	0	NA	NA
3	I	POL6	3	120	4.12	13.62	0.30	0.64	23.81	0.00	1.00	0	1	0	0	0	0
3	I	POL9	3	100	5.10	19.98	0.26	0.70	12.41	0.00	1.00	0	1	0	0	NA	NA
3	I	POL10	5	150	3.72	12.45	0.30	0.66	20.94	0.00	1.00	0	1	0	0	NA	NA
3	I	POL11	3	100	5.53	23.03	0.24	0.57	20.19	0.00	1.00	0	1	0	0	0	0
3	J	IMB2	3	30	4.83	8.83	0.55	0.57	12.66	0.00	1.00	1	1	0	0	NA	NA
3	J	IMB5	3	50	3.68	7.52	0.49	0.63	18.51	0.00	1.00	1	1	0	0	1	1
3	J	IMB6	3	60	2.86	10.59	0.27	0.58	30.57	0.00	1.00	1	1	0	0	NA	NA
3	J	IMB7	3	15	6.52	10.74	0.61	0.64	43.70	0.00	1.00	1	1	0	0	NA	NA
3	J	IMB10	3	34.5	3.31	9.40	0.36	0.64	20.88	0.00	1.00	1	1	0	0	NA	NA
3	J	IMB12	3	50	5.44	12.29	0.44	0.58	27.65	0.00	1.00	1	1	0	0	NA	NA
3	J	IMB13	3	40	4.79	7.55	0.64	0.53	34.28	0.00	1.00	1	1	0	0	NA	NA
3	J	IMB14	3	50	5.86	11.45	0.52	0.46	17.96	0.00	1.00	1	1	0	0	1	1
3	J	IMB16	3	50	3.94	8.86	0.45	0.58	33.38	0.00	1.00	1	1	0	0	NA	NA
3	J	IMB20	3	74	4.73	12.29	0.39	0.62	53.80	0.00	1.33	1	1	0	0	1	1
3	J	IMB22	3	56	4.70	10.18	0.46	0.61	28.11	0.00	1.33	1	1	0	0	1	1
4	K	POL12	3	50	2.43	18.38	0.13	0.66	22.04	0.00	1.00	0	1	0	0	0	0
4	K	POL14	2	62	2.01	15.50	0.13	0.64	44.19	0.00	1.00	0	1	0	0	NA	NA
4	K	POL15	3	30	1.78	11.88	0.15	0.51	34.98	0.00	1.00	0	1	0	0	0	0

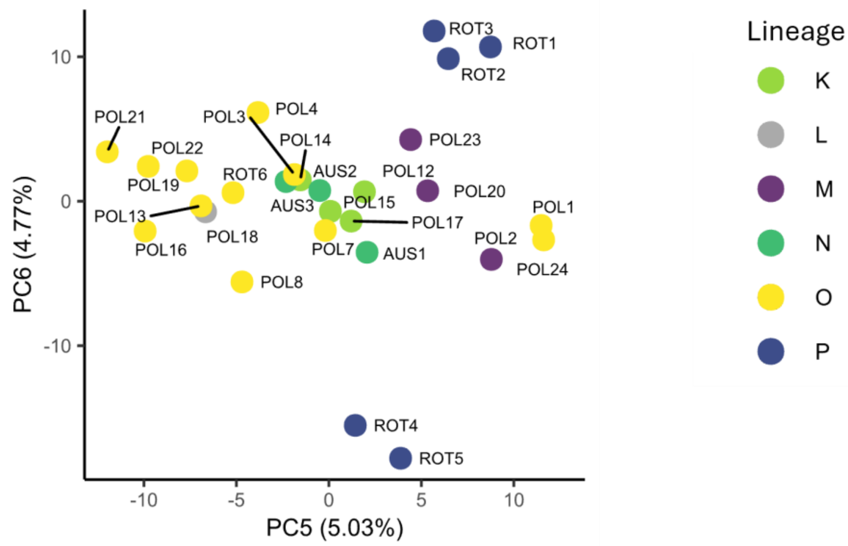
4	K	POL17	3	20	2.31	18.70	0.12	0.67	23.29	0.00	1.00	0	1	0	0	0	0
4	L	POL18	4	12.5	2.22	9.83	0.23	0.72	5.54	0.00	1.00	0	1	1	0	1	0
4	M	POL20	3	60	1.45	9.06	0.16	0.55	50.74	0.00	1.67	0	1	1	0	NA	NA
4	M	POL23	3	120	3.17	15.79	0.20	0.74	37.95	0.33	2.67	0	1	1	0	NA	
4	M	POL24	4	20	1.48	8.85	0.17	0.59	13.68	0.00	1.00	0	1	1	0	NA	NA
4	N	AUS1	3	45	6.95	19.40	0.36	0.58	6.19	0.00	1.00	0	1	0	1	NA	NA
4	N	AUS2	2	200	20.11	44.92	0.45	0.33	11.07	0.00	1.00	0	1	0	1	NA	NA
4	N	AUS3	3	50	7.33	26.01	0.28	0.54	7.84	0.00	1.00	0	1	0	1	0	0
4	O	POL1	3	90	5.55	20.67	0.27	0.31	26.73	1.33	2.33	0	1	1	0	1	0
4	O	POL2	5	80	6.60	39.58	0.18	0.55	65.10	2.40	3.60	0	1	1	0	1	0
4	O	POL3	4	40	3.31	13.78	0.24	0.42	61.53	0.00	1.50	0	1	1	0	1	0
4	O	POL4	3	20	3.62	20.21	0.18	0.67	46.37	1.33	3.33	0	1	1	0	1	0
4	O	POL7	1	50	3.50	20.38	0.17	0.34	49.12	2.00	3.00	0	1	1	0	NA	NA
4	O	POL8	3	50	2.95	14.57	0.20	0.30	57.27	0.00	2.00	NA	1	NA	0	1	0
4	O	POL13	3	30	3.09	22.34	0.14	0.22	26.74	0.00	1.00	0	1	1	0	NA	NA
4	O	ROT6	3	200	15.93	14.71	1.08	0.24	42.17	4.00	6.67	0	1	1	0	0	0
4	O	POL16	2	47	1.60	12.09	0.13	0.44	26.32	0.00	1.00	0	1	1	0	NA	NA
4	O	POL19	3	20	4.27	21.51	0.20	0.61	34.52	0.00	2.00	0	1	1	0	1	0
4	O	POL21	3	40	4.55	16.18	0.28	0.63	28.02	0.00	1.67	0	1	1	0	1	0
4	O	POL22	3	30	2.01	12.70	0.16	0.39	17.61	0.00	1.00	0	1	1	0	0	0
4	P	ROT1	1	130	21.68	31.17	0.70	0.34	52.44	5.00	17.00	0	1	1	0	NA	NA
4	P	ROT2	6	150	15.90	23.24	0.69	0.29	42.41	4.50	8.50	0	1	1	0	1	0
4	P	ROT3	3	100	15.65	32.81	0.49	0.30	13.86	0.33	1.33	0	1	1	0	NA	NA
4	P	ROT4	3	120	18.23	20.83	0.88	0.18	57.98	2.67	4.67	0	1	1	0	1	0
4	P	ROT5	4	200	15.90	26.34	0.61	0.27	49.20	3.00	4.25	0	1	1	0	NA	NA



Appendix 5. Maximum likelihood phylogeny of 79 *Osteospermum* section *Polygalinae* populations inferred from concatenated GBS data filtered to < 20 % missing per site. The tree was rooted on an outgroup (*O. junceum,* not shown). Major clades are indicated by arrows. Tip labels give the species name under the current taxonomy, collection locality and population code. Branch supports are given as follows: SH-aLRT support (%) / ultrafast bootstrap support (%).



Appendix 6. Ancestral gene pool assignment of each population in the four major clades inferred using sparse nonnegative matrix factorization (sNMF). For each major clade ancestral gene pool assignment was investigated for K values ranging from two to the number of significant eigenvalues of the PCA (as determined by Tracy-Widom tests) plus two. Colours indicate gene pool assignment.



Appendix 7. PCA ordination of genotyping-by-sequencing (GBS) single-nucleotide polymorphism data for Clade 4 showing PC5 and PC6. These axes showed significant variation as determined by Tracy-Widom tests but were not informative for species delimitation. Colours indicate lineage membership.