



Exploring the phylogenomics of *Dipcadi* (Asparagaceae, Scilloideae) and assessing the utility of ITS as a DNA barcode

Abigael Kitur¹ · Brita Stedje¹ · Emily Wabuyele² · Solveig Bua Løken¹

Received: 28 November 2024 / Accepted: 16 March 2025 / Published online: 8 May 2025
© The Author(s) 2025

Abstract

Dipcadi (Asparagaceae, Scilloideae) has long presented phylogenetic and taxonomic challenges. A recent study by Chapano et al. (in press) based on ITS sequence data found several non-monophyletic species and an unresolved backbone topology of the genus. In this study, we present the first high-throughput sequencing (HTS) phylogeny of *Dipcadi*, using data from target capture of the angiosperms353 loci. The sampling is focused on African diversity. Additionally, we provide an updated ITS phylogeny representing 15 *Dipcadi* species, including 18 newly sequenced accessions. We expand taxon sampling, especially of *D. longifolium*, to further explore phylogenetic relationships and species limits in the genus. Although the angiosperms353 loci did not fully resolve the backbone topology of *Dipcadi*, both approaches generated similar topologies with six well-supported clades, but with incongruences regarding deep nodes. We corroborate monophyly of *D. marlothii* and *D. glaucum*. *Dipcadi longifolium* represents a species complex, possibly including *D. gracillimum*, but we refrain from taxonomic revisions based on the current sampling. *Dipcadi viride* and possibly *D. platyphyllum* are paraphyletic, warranting further attention. Our findings confirm that ITS can be effectively employed as a barcode marker in *Dipcadi*.

Keywords Asparagaceae · Angiosperms353 · Barcoding · *Dipcadi* · Internal transcribed spacer · Target capture

Introduction

Dipcadi Medik (Asparagaceae: Scilloideae) (APG IV 2016) consists of about forty species distributed in Africa, Madagascar, the Arabian Peninsula, the Mediterranean region and India, with centers of diversity in Southern Africa and India (Manning et al. 2009; Martínez-Azorín et al. 2011; Deshpande et al. 2016; Stedje and Kativu 2023). These bulbous plants have linear to lanceolate leaves that are arranged in a basal rosette, racemose inflorescences, bell-shaped or tubular flowers that are whitish, yellowish, green or brownish, with or without filiform caudate appendages on outer tepals, trilobed capsules and black, flat seeds (Stedje and Thulin

1995; Stedje 1996; Manning et al. 2004; Martínez-Azorín et al. 2011).

Most phylogenetic works that include *Dipcadi* have been of the Scilloideae and have included a limited number of *Dipcadi* accessions. These works are either based on a few plastid markers (Pfosser and Speta 1999; Manning et al. 2004, 2009) or plastid in combination with the nuclear ribosomal internal transcribed spacer (ITS) (Martínez-Azorín et al. 2011; Buerki et al. 2012). These studies retrieve a monophyletic *Dipcadi*. Recent molecular phylogenetic studies have treated both African (Chapano et al. in press) and Indian (Shelke et al. 2024) *Dipcadi* species, however, with unresolved backbone topologies and several inter- and intraspecific relationships remaining unclear. Chapano et al. (in press) presented an ITS phylogeny comprising 50 *Dipcadi* accessions representing 14 species and found several non-monophyletic morphospecies. One of these is the *D. longifolium* species complex, which based on the ITS region seems to represent at least three different evolutionary lineages.

The ITS region is often easy to amplify and is well-known for resolving phylogenetic relationships at shallow taxonomic levels (Álvarez and Wendel 2003). Although there

Handling Editor: Joerg Fuchs.

✉ Solveig Bua Løken
s.b.loken@nhm.uio.no

¹ Integrative Systematics of Plants and Fungi (ISOP) Group, Natural History Museum, University of Oslo, Oslo, Norway

² Department of Plant Sciences, Kenyatta University, Nairobi, Kenya

is no universal barcode for plants, the ITS region is widely used as one (Hollingsworth et al. 2011; Antil et al. 2023). The use of ITS in phylogenetics is, however, not without its challenges. The region is present in multiple copies within a genome, which can, if not properly homogenized through concerted evolution, lead to inaccurate sequencing and alignment problems following the comparison of nonhomologous sequences (Álvarez and Wendel 2003; Hollingsworth et al. 2016). Studies on the *Ornithogalum umbellatum* group by Herrmann (2002) and on *Ledebouria* by Lebatha et al. (2006) have reported problems in amplifying the ITS region, possibly due to the presence of multiple alleles, while in closely related genera, the region proves easily amplified (Chapano et al. in press).

Single-copy nuclear markers are considered the most reliable for phylogenetic inference, yet in plant systematics, issues regarding amplification of such regions have limited their use (Small et al. 2004; Zimmer and Wen 2012). As a result, most Sanger-based plant phylogenies are still based on plastid loci and/or the ITS region. However, advancements in high-throughput sequencing (HTS) methods have made plant nuclear DNA more accessible, introducing various new techniques to generate large genetic datasets for phylogenetics (Lemmon and Lemmon 2013; Dodsworth et al. 2019). Among these techniques, target capture has gained popularity (Dodsworth et al. 2019; Andermann et al. 2020). This method uses probes to selectively capture specific genomic regions, often targeting hundreds of loci across the genome (Albert et al. 2007; Gnirke et al. 2009; Weitemier et al. 2014). In plant systematics, the use of target capture has intensified after the development of the angiosperms353 probe set (Johnson et al. 2019). These probes are designed to target 353 putatively single-copy nuclear markers across all angiosperms, and the loci have proven useful in producing well-resolved and reliable phylogenies from deep to shallow taxonomic levels (e.g., (Villaverde et al. 2018; Larridon et al. 2020; Murphy et al. 2020; Bagley et al. 2020; Antonelli et al. 2021; Maurin et al. 2021; Ottenlips et al. 2021; Thomas et al. 2021a; b). Additionally, target capture is suitable for herbarium material with somewhat degraded DNA, making it a valuable tool for including old collections in phylogenetic studies (Hart et al. 2016; Brewer et al. 2019; Shee et al. 2020).

While target capture with the angiosperms353 probe set represents a powerful tool in plant phylogenetics, the method is challenging to perform in many research environments. Successful implementation needs technical expertise and necessary infrastructure, and processing and storage of large genomic datasets demands time and substantial bioinformatic skills. Analyses are most efficiently carried out on computer clusters, which are not available to all research institutions. Despite HTS decreasing the per base pair cost of data production, large-scale projects

involving numerous accessions remain costly (McKain et al. 2018). Therefore, traditional and cost-efficient methods like Sanger sequencing of a few loci remains practical and relevant for many plant systematists, especially in studies with large sample sizes. Sanger-based methods are also important for barcoding purposes. This procedure requires standard laboratory equipment, is relatively quick and data analysis is straightforward (McKain et al. 2018).

Our study builds on the work of Chapano et al. (in press), who found that the ITS region alone could not resolve the backbone phylogeny of *Dipcadi*. We therefore apply a target capture approach using the angiosperms353 probe set to increase our understanding of phylogenetic relationships within the genus. Additionally, we expand the taxon sampling of the *D. longifolium* species complex with the aim of delimiting the species within this group. Since the ITS region is readily amplified in *Dipcadi* (Chapano et al. in press), we want to assess its utility as a barcode marker in the genus by comparing its resolving power at the intragenetic level to that of the angiosperms353 loci. Our primary focus is on the delimitation of African *Dipcadi* species; Indian taxa will not be discussed in detail in this study.

Materials and methods

Material and sampling

In this study, we used 21 samples for generating genomic sequence data, originating from field-collected silica-dried leaves or herbarium material (Online resource 1). This sampling represented seven *Dipcadi* species hypotheses. Some of the silica-dried samples were collected from eight localities across Kilifi, Mombasa and Kwale Counties of Kenya in January 2022. Habitat characteristics were recorded for each plant population, from which the samples were collected. Vouchers are deposited at the National Museums of Kenya, East African Herbarium (EA). We also sampled from herbarium material both from the National History Museum, University of Oslo (O) and the National Herbarium and Botanic Gardens, Zimbabwe (SRGH). Sequence data for the out-group (*Albuca* sp.) was obtained from a separate study (Løken et al., in prep). For the ITS dataset, we also used 46 ITS sequences representing 14 species from Chapano et al. (in press). Eighteen of these sequences were originally retrieved from GenBank. Three accessions sequenced for this study were also ITS-sequenced by Chapano et al. (in press). All samples are identified based on a morphological species concept. Samples denoted “sp.” in the final phylogeny could not be referred to any described *Dipcadi* species.

Genomic DNA extraction, library prep, target capture and Illumina sequencing

We extracted total genomic DNA using the E.Z.N.A SP Plant DNA Kit (Omega Bio-tek Inc., Norcross, Georgia, USA) following the manufacturer's protocol with the following modifications: Prior to DNA extraction, we ground the dried samples with 2 tungsten-carbide beads in a TissueLyser Adapter at 25 Hz for 2 min until the samples were crushed into a fine powder. We extended the lysis time to 1 h instead of 10 min (when the first extraction did not yield enough DNA). To increase yield, we eluted the samples twice using 50 µl elution buffer and incubated them at 65 °C for 20 min during each elution. The quantity and quality of the extracted DNA was assessed using gel electrophoresis with 1% agarose gel, Qubit 2.0 fluorometer with the dsDNA HS assay kit (Life Technologies, Carlsbad, CA, USA), and NanoDrop spectrophotometer (Thermo Fisher Scientific, USA). When necessary, we pooled multiple extracts from the same sample to increase the total amount of DNA. We sent dried DNA extracts to Arbor Biosciences (Ann Arbor, Michigan, USA) for library preparation, target capture using the angiosperms353 probe set and sequencing. At Arbor, eight libraries were pooled in each target capture reaction, and enriched libraries were sequenced on an Illumina NovaSeq 6000 S4 platform using 150 bp paired-end reads and to a depth of 1 Gbp per sample.

Phylogenomic analyses of the target capture dataset

All analyses were performed on the Saga cluster provided by Sigma2—the National Infrastructure for High-Performance Computing and Data Storage in Norway. The raw reads were cleaned and trimmed using Trimmomatic (Bolger et al. 2014) with the following parameters: LEADING:30, TRAILING:30, SLIDING WINDOW:4:30 and MINLEN:36 parameters. We checked the quality of both raw and cleaned reads using FastQC v0.11.9 (Andrews 2010). Angiosperms353 exons were recovered using HybPiper v2.1.2 (Johnson et al. 2016) with default settings. Through HybPiper, the reads were mapped to the “mega353” reference file (McLay et al. 2021) using BWA v0.7.17 (Li and Durbin 2010), de novo assembly of each gene was done using SPAdes v3.15.3 (Prjibelski et al. 2020) and exons were extracted using exonerate (Slater and Birney 2005). Potential paralog sequences were recovered using HybPiper, aligned separately using MAFFT v7.407 (Katoh and Standley 2013) with default settings, and approximately-maximum-likelihood phylogenetic trees were generated using FastTree (Price et al. 2009, 2010). The resulting gene trees were inspected manually, and all genes that showed evidence of an ancient duplication event were exempted from later analysis.

All remaining exons were aligned separately using MAFFT with default settings. Each of the sequence alignments were trimmed with ClipKIT (Steenwyk et al. 2020) to remove all sites that were above 0.9 threshold of gappy-ness (mode=gappy). We calculated alignment statistics using AMAS (Borowiec 2016). Maximum likelihood gene trees were constructed using IQ-TREE v2.0-rc1 (Minh et al. 2020) with 1000 ultrafast bootstrap replicates (Hoang et al. 2018). The best-fitted substitution model for each alignment was selected using the ModelFinder option (Kalyaanamoorthy et al. 2017). We used the gene trees as input to ASTRAL-III version 5.7.8 (Zhang et al. 2018) for estimation of a species tree. We also conducted an analysis based on the supercontig sequences retrieved via HybPiper, following the same procedure as outlined above.

Phylogenetic analysis of the ITS datasets

Using HybPiper, we recovered the ITS region from our raw reads. We prepared a reference file consisting of 46 ITS sequences obtained from GenBank and from Christopher Chapano (The National Herbarium and Botanical Gardens, Harare, Zimbabwe; Online resource 1), and ran HybPiper with the default settings. The final ITS dataset consisted of 67 in-group taxa and *Albuca* sp. as out-group, including both the sequences used in the target file and the bioinformatically retrieved sequences, hereafter called ‘ITS-full’ (Online resource 1). To aid in comparison between the two methods and for inference of the reliability of the ITS phylogeny, we also created a reduced ITS dataset. This dataset included only the 22 bioinformatically retrieved ITS sequences, matching the accessions in the target capture dataset, hereafter called “ITS-red.”

Both datasets were aligned separately using MAFFT as implemented in Geneious Prime v.2023.0.3 (Kearse et al. 2012). The alignments were inspected manually, ends were trimmed and ambiguous nucleotides were edited where appropriate. For both ITS datasets, we performed a Bayesian phylogenetic analysis using MrBayes (Ronquist et al. 2012). Alignments were converted into MrBayes Nexus format using ALTER (Glez-Pen˜a et al. 2010). We used jModelTest v 2.1.10 (Darriba et al. 2012) and the Akaike Information Criterion (Akaike 1974) to find the best-fitted nucleotide substitution model. The Markov chain Monte Carlo (MCMC) chains were run for 10 000 000 generations, using one cold and three heated chains, and sampled every 1000 generations. The distribution of marginal likelihood values was visualized in Tracer v1.7.1 (Rambaut et al. 2018) to ensure convergence. The first 25% sampled trees were discarded as “burn-in,” and the remaining trees were used to compile a posterior probability (PP) distribution using a 50% majority rule.

Tree visualization

We visualized all phylogenetic trees in iTOL v5 (Letunic and Bork 2021) and did final figure edits in Adobe Photoshop (Adobe Inc., 2024). Taxon names are based on morphological identification of the species in the field and/or based on their vouchers.

Results

Target capture results

We produced genomic data for 21 samples and seven species of *Dipcadi*: *D. glaucum*, *D. gracillimum*, *D. longifolium*, *D. marlothii*, *D. platyphyllum*, *D. rigidifolium* and *D. viride*. *Albucca* sp. was included as out-group. On average, 6,767,285 read pairs were obtained per sample (ranging from 1,507,974 to 10,847,764), and an average of 56.6% reads per sample were on target. The recovery ranged from 349 to 352 out of the 353 genes targeted by the angiosperms353 probe set (Online resource 2).

Twenty-three paralog genes out of 49 suspected paralogs were excluded from downstream analysis, as their gene trees suggested an ancient duplication event. One gene was also excluded because it included sequences for only 2 accessions. The final analysis included 328 exons. Only three of these lacked sequences for in-group taxa (exon 6270: 4 missing, exon 6430: 10 missing, exon 7135: 2 missing). Across the alignments, the proportion of variable sites (PVS) was 23.8% on average (8.5–45.8%), while the average proportion of parsimony informative sites (PIS) was 9.0% (2.2–27.8%). Average alignment length was 799 bp (81–3498 bp), and the average percentage of missing characters was 9.9% (0.0 to 66.8%). The species tree produced with ASTRAL had a final quartet score of 1,517,018 out of 2,316,468 quartet trees, meaning 65.5% of the quartet trees in the gene trees are present in the species tree (Fig. 1, left).

The target capture phylogeny supports six distinct clades, A, B, C, D, E and F, all with maximum support (LPP = 1; Fig. 1, left). We have three pairs of sister clades, each strongly supported: clades A + B (LPP = 1), clades C + D (LPP = 1) and clades E + F (LPP = 1). The relationship between these pairs of sister clades is not resolved. One accession, *D. longifolium* (CC-2064) falls out as a sister to all other accessions of the in-group (LPP = 1). The whole in-group is also strongly supported (LPP = 1). Clade A consists of the East African, and most accessions of Zimbabwean *D. longifolium*. Within this clade, the East African accessions (AK-4, AK-5, AK-6, AK-7, IN-1970) form a monophyletic group with strong support (LPP = 1), while the Zimbabwean accessions (CC-3510, CC-3506, CC-3517) are paraphyletic with respect to the East African clade. Clade B consists of *D.*

rigidifolium and *D. viride*, but the relationship between these two taxa is unresolved. In clade C, we find two accessions of *D. platyphyllum*, and in clade D, two accessions of *D. glaucum*. Clade E comprises *D. longifolium* BS-458 and *D. gracillimum* CC-3515. Clade F consists of three accessions of *D. marlothii*. The analysis of the supercontig sequences resulted in an identical topology and no significantly higher support values.

ITS results

Our ITS-red phylogeny included 22 successfully sequenced accessions representing seven *Dipcadi* species and one out-group, while the ITS-full phylogeny consisted of 68 accessions representing 15 *Dipcadi* species and one outgroup (Fig. 1, right and Fig. 2). The final alignment length for both ITS datasets was 731 bp, and average sequence length was 670 bp (321–726 bp). In the Bayesian analyses, for both the reduced and the full dataset, the two runs converged with an average standard deviation of split frequencies (SDSF) < 0.01. The resulting phylogenies are congruent, but with varying support for a few nodes (Fig. 1, right and Fig. 2).

The analyses of the target capture and ITS-red datasets produced similar topologies, with some notable incongruences (Fig. 1). Clades A–F are all strongly supported based on both ITS (BPP = 1) and target capture (LPP = 1); however, the relationships between them differ. Based on ITS, clade A (most accessions of *D. longifolium*) is sister to clade C + D (*D. platyphyllum* and *D. glaucum*), while in the target capture tree, clade A is sister to clade B (*D. viride*). Based on ITS, clade B is moderately supported as sister to A + C + D (BPP = 0.85). Clade E + F is not supported based on ITS, but clade E is rather a weakly supported sister to A + C + D + B (BPP = 0.52, but LPP = 0.83 in ITS-full). Clade F is strongly supported as sister to the remaining clades (BPP = 1). Like in the target capture phylogeny, *D. longifolium* (CC-2064) is sister to the remaining in-group taxa (BPP = 1). In addition, the East African accessions of *D. longifolium* within clade A do not form a monophyletic group, as one of the Zimbabwean accessions is nested within this clade.

In ITS-full, *Dipcadi* forms two major clades representing an African-Malagasy-European group (clades A–F) and a largely Indian group (clade G) (Fig. 2). Clades A–F in ITS-full (Fig. 2) are all highly supported (BPP = 0.95–1) and comparable to the corresponding named clades in Fig. 1. Clade A is moderately supported, consisting of most accessions of *D. longifolium*, three accessions of *D. gracillimum* (Manning-2682, CC-2071, CC-2072), one accession of *D. serotinum* (Pedersen-1971) and one accession of *D. sp* (Kuchar-18019) from Ethiopia (BPP = 0.85). The *D. serotinum* accessions is strongly supported as sister to the *D. longifolium* and *D. gracillimum* accessions

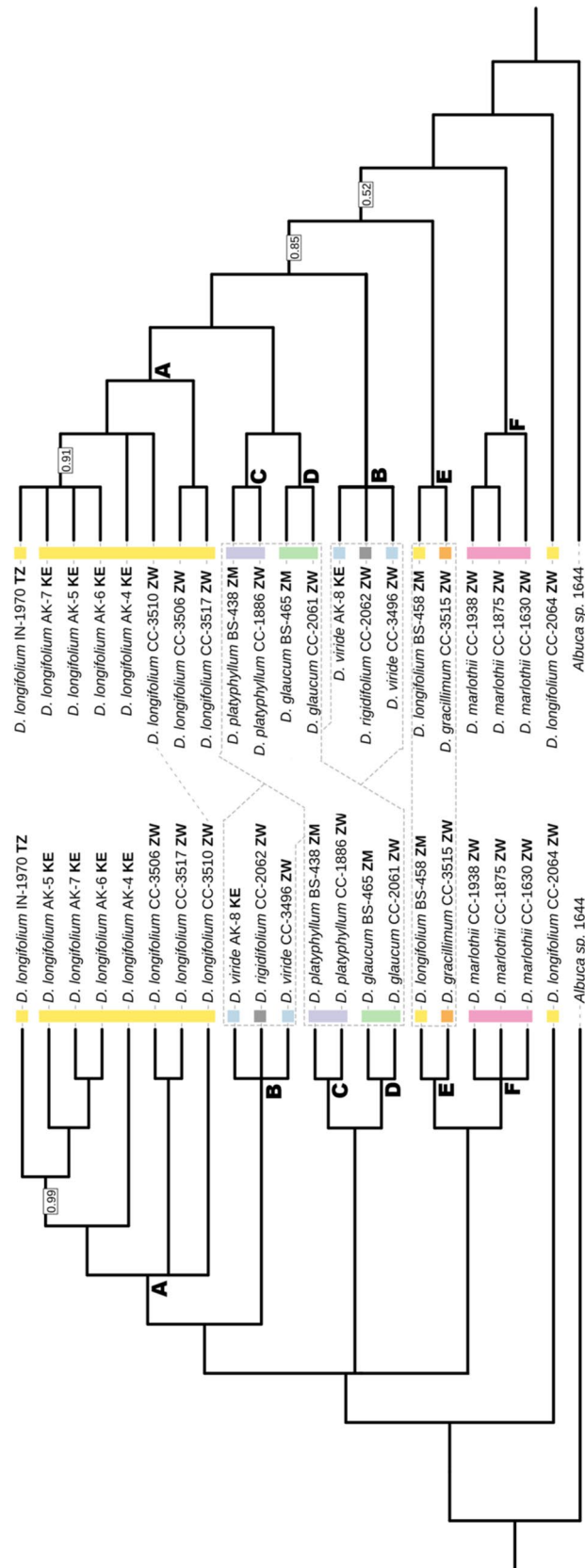


Fig. 1 Comparison of phylogenetic trees based on the target capture (ITS-red, left) and the reduced ITS dataset (ITS-full, right), showing 21 *Dipcadi* accessions and *Albuca* sp. as outgroup. The target capture phylogeny is based on 328 nuclear loci from the angiosperms353 probe set and inferred using ASTRAL-III. The ITS phylogeny is a 50% majority-rule consensus tree based on the nuclear internal transcribed spacer (ITS) region and inferred using MrBayes. Support values are local (left) and Bayesian (right) posterior probabilities. All nodes with support values below 0.8 are collapsed, except one node in ITS-red because it received higher support in the full ITS dataset (ITS-full, Fig. 2). Support values between 0.8 and 0.99 are displayed on branches, and remaining nodes have maximum support. Major clades are named A-F and species color-coded are colored (yellow *D. longifolium*, light blue *D. viride*, gray *D. rigidifolium*, purple *D. platyphyllum*, light green *D. glaucum*, pink *D. marlothii*, orange *D. marlothii*, orange *D. marlothii*). Topological incongruences between the two methods are highlighted by dashed boxes or lines connecting clades or accessions. Taxon names are based on morphological identifications and confirmed by inspecting vouchers. Details on each accession are found in Online Resource 1. Abbreviations: DDipcadi, KE Kenya, TZ Tanzania, ZM Zambia, ZW Zimbabwe

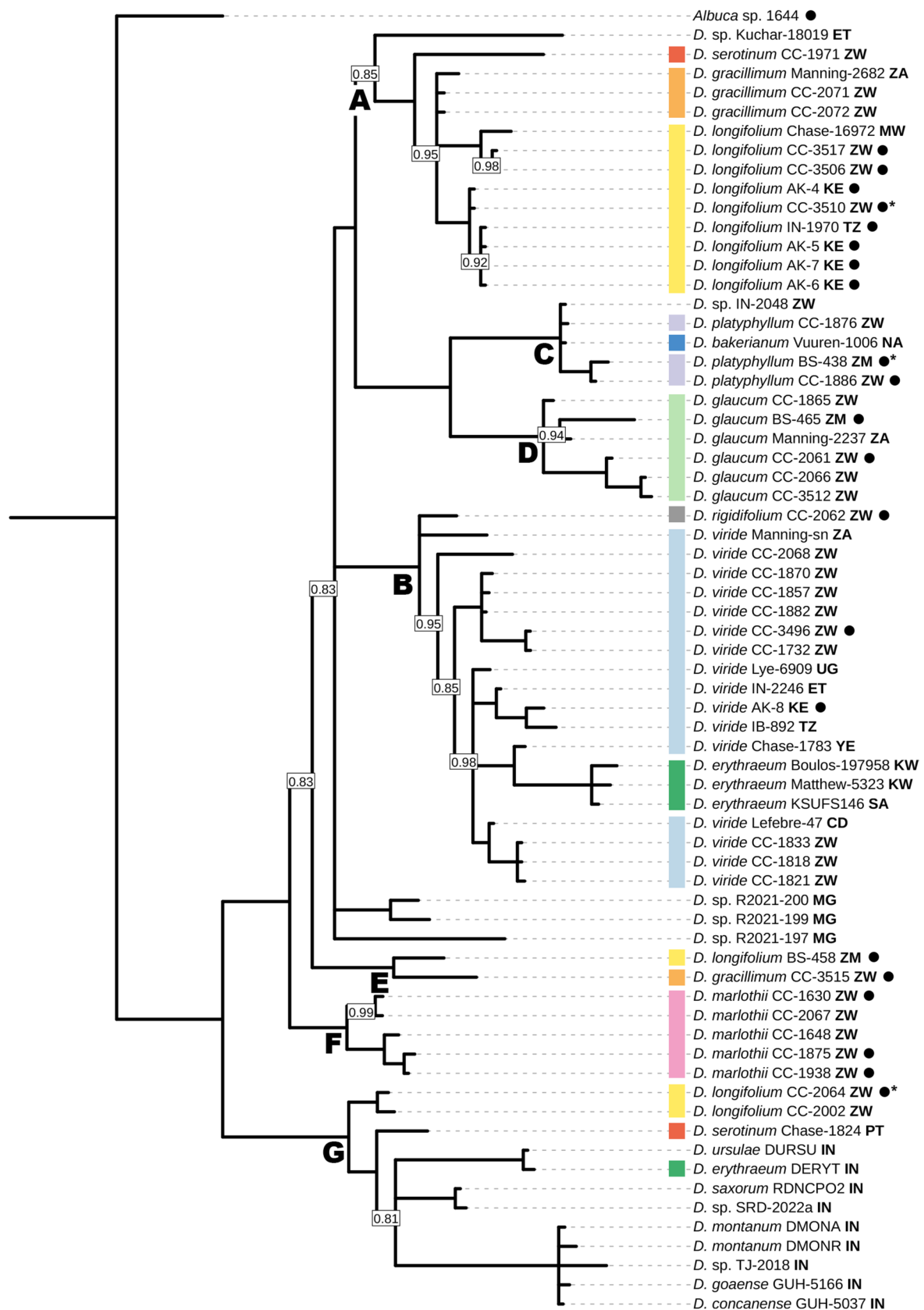


Fig. 2 The 50% majority-rule consensus tree based on the full ITS dataset: 67 *Dipcadi* accessions and *Albuca* sp. as outgroup. The phylogeny is based on the nuclear internal transcribed spacer (ITS) region and inferred using MrBayes. All nodes with Bayesian posterior probability (BPP) values below 0.8 are collapsed, BPP between 0.8 and 0.99 are displayed on branches and remaining nodes have maximum support. Major clades are named A–G, and species names are colored (red *D. serotinum*, orange *D. gracillimum*, yellow *D. longifolium*, purple *D. platyphyllum*, dark blue *D. bakerianum*, light green *D. glaucum*, light green *D. rigidifolium*, light green *D. viride*, dark green *D. erythraeum*, pink *D. marlothii*). Accessions sequenced for this study are marked with a black circle. Accessions that are sequenced both for this study and for Chapano et al. 2024 are marked with a black circle and an asterisk. Details on each accession are found in Online Resource 1. Taxon names are based on morphological identifications and confirmed by inspecting vouchers. Abbreviations: CD Democratic Republic of Congo, DD *Dipcadi*, ESS Spain, ET Ethiopia, IN India, KE Kenya, KW Kuwait, MG Madagascar, MW Malawi, NA Namibia, PT Portugal, SA Saudi Arabia, TZ Tanzania, UG Uganda, YE Yemen, ZA South Africa, ZM Zambia, ZW Zimbabwe

(BPP = 1). The accessions of *D. gracillimum* form a polytomy with two *D. longifolium* clades. Clade B consists of two subclades and three accessions falling outside of these. The uppermost subclade (BPP = 1) is composed of accessions of *D. viride* from Zimbabwe only. The lowermost subclade (BPP = 0.98) consists of *D. erythraeum* as well as accessions of *D. viride* from various parts of the species' distribution area. These subclades are moderately supported as sisters (BPP = 0.85). An accession of *D. viride* from Zimbabwe (CC-2068) comes out as sister to these two subclades (BPP = 0.95). One accession of *D. rigidifolium* (Chapano-2062) and one of *D. viride* (Manning-sn) fall outside the above-mentioned subclades. Clade C includes *D. platyphyllum*, *D. bakerianum* and one accession of *D. sp.* (IN-2048), but the relationships between the accessions are partly unresolved. Clade D consists only of *D. glaucum*, separated into two subclades and one single sample. In a polytomy with clades A–D, there are three accessions of an undetermined *Dipcadi* from Madagascar. Clade E includes one accession of *D. longifolium* (BS-458) and one accession of *D. gracillimum* (CC-3515). Clade F recovered accessions of *D. marlothii* forming two subclades. As sister to clades A–F, we find a strongly supported clade G that includes mostly Indian species. The first branch in this clade includes two African accessions of *D. longifolium* (CC-2064 and CC-2002) (BPP = 1), followed by one accession of *D. serotinum* (Chase-1824). Subsequently we find a moderately supported (BPP = 0.81) clade containing the Indian species; *D. ursulae*, *D. erythraeum*, *D. saxorum*, *D. montanum*, *D. concanense* and *D. goanense*.

Geographic affinities

Indian samples form a clade with a sample of *D. serotinum* from Portugal and two samples of *D. longifolium* (clade G, Fig. 2), and those from the Arabian Peninsula form a clade inside the larger clade of African, European and Malagasy species (i.e., *D. erythraeum* in clade B, Fig. 2). Clades C–F comprise accessions from southern parts of Africa, while clades A and B include more geographically diverse accessions. Accessions from Madagascar and the Arabian Peninsula are phylogenetically positioned with mainland African accessions and distant from the more proximal clade of Indian species. The Mediterranean accessions of *D. serotinum* were not clustered together.

Discussion

In this study, we provide the first HTS phylogeny of *Dipcadi*, based on 328 nuclear loci obtained through target capture and the angiosperms353 probe set. We compare the topology of the HTS dataset to one obtained through a more traditional and accessible method, namely Sanger sequencing of the ITS region. Although we do not provide formal taxonomic changes, we provide several suggestions for future systematics of *Dipcadi*.

Species delimitation in *Dipcadi*

In our ITS phylogeny, we have expanded the sampling in Chapano et al. (in press) with 18 accessions: Eight *Dipcadi longifolium*, one *D. gracillimum*, one *D. platyphyllum*, two *D. glaucum*, three *D. marlothii*, one *D. rigidifolium* and two *D. viride*. Since the target capture and ITS analyses showed a high degree of congruence regarding main clades, and because the full ITS tree includes a larger number of accessions, the discussion on species delimitation will be based on the full ITS phylogeny (Fig. 2) unless otherwise stated.

Based on the results from Chapano et al. (in press), we sampled *Dipcadi longifolium* more extensively with the aim of resolving this species complex. Their results indicated that *D. longifolium* represents at least three genetically distinct lineages based on ITS sequence data. Our results support this, as target capture data also reveals a polyphyletic *D. longifolium* representing three evolutionary lineages with samples identified as such present in clade A, clade E and also sister to clades A–F (a subclade of clade G in ITS-full, Fig. 2) (Fig. 1). *Dipcadi longifolium* as currently circumscribed appears to be comprised of three distinct species that are found in disparate lineages in the phylogeny. East African samples of *D. longifolium*, occurring in sandy beaches are all found in clade A. The type specimen of *D. longifolium* was collected from Mafamale Island, a small

island off the northern coast of Mozambique, growing in loose sand (Obermeyer 1964), and we believe that our East African accessions found in this habitat represents the concept of this species. Hence, the two other lineages of *D. longifolium* (clade E and the subclade of clade G, Fig. 2) should be investigated further and described as new species. By increasing sampling of *D. longifolium* in this study, we aimed to understand these two clades better. However, since most newly sampled accessions of *D. longifolium* fall into clade A, more evidence on the distribution and morphology of the other two clades is needed before we can provide any formal description of new species. At this point, the diversity in the morphospecies *D. longifolium* is cryptic in the sense that we cannot distinguish the genetic lineages based on their morphology. We have inspected all vouchers, and they are confirmed as *D. longifolium* based on their morphology. The clades are not distinguished by geography either, since all clades include accessions from Zimbabwe. It might be worthwhile investigating ecological preference further: *D. longifolium* accessions in clade G are sampled in open, nutrient-rich habitats (fallow field/grassland), accessions in clade E are sampled from sandy soils, while accessions in clade A are sampled from various habitats (mainly on sand or clay).

Dipcadi gracillimum also appears polyphyletic. Interestingly, both accessions of *D. gracillimum* form a sister relationship to a *D. longifolium* (clades A and E, Fig. 2). This suggests that the two names may refer to the same morphospecies, as also pointed out by Chapano et al. (in press). While the two species are morphologically similar, *D. gracillimum* has narrower leaves, a potentially plastic trait influenced by the environment, which could lead to misidentifications. Further investigations are needed to determine whether the name *D. gracillimum* could be applied to clade E, or if it is a synonym of *D. longifolium*.

The newly sequenced accession of *D. platyphyllum* falls into clade C, together with the other accessions of *D. platyphyllum*, *D. bakerianum* and one accession of *D. sp.* The relationships within this clade are unresolved, making *D. platyphyllum* appear paraphyletic, as also found in Chapano et al. (in press). *Dipcadi bakerianum* and *D. platyphyllum* are clearly defined morphologically. The distinguishing feature of *D. platyphyllum* is its elliptic lanceolate leaves with wavy margins. *Dipcadi bakerianum* is described as having a knobbed ovary and capsule, swollen scape and lack of caudate appendages on the tepals (Obermeyer 1964). However, the ITS region does not separate these two species, indicating a close relationship between them. As we do not have target capture data from *D. bakerianum*, we cannot conclude whether the two species represent the same genetic lineage or if there is too little variation in their ITS sequences to delimit them.

Newly sequenced accessions of both *D. glaucum* (clade D) and *D. marlothii* (clade F) cluster as expected and corroborate monophyly of these two species, respectively. Both species were also recovered as monophyletic in Chapano et al. (in press). In addition to being genetically distinct, they are both well-defined morphologically, as pointed out by Chapano et al. (in press). *Dipcadi glaucum* has wide strap-shaped leaves that are glaucous (blue green with fine waxy substance on their surfaces) and many flowers that are arranged spirally in the inflorescence (Hyde et al. 2023). *Dipcadi marlothii*, on the other hand, is distinguished by having many, hairy, narrow and often spiralized leaves (Stedje 1996). Both these species exhibit some level of genetic variation within.

Both accessions of *D. viride* sequenced for this study fall out in clade B. As in Chapano et al. (in press), *D. viride* appears to be paraphyletic with a long-branched *D. erythraeum* clade nested within. The two species are similar morphologically, and their distribution overlaps in the Arabian Peninsula (POWO 2024). The appearance of *D. viride* as a paraphyletic group is consistent with the results of Manning et al. (2009), which was based on a plastid phylogeny and a more limited sampling than ours. Chapano et al. (in press) investigated if any morphological characters support the subclades of *D. viride* but found no correlation between morphological and genetic differentiation. The Indian accession of *D. erythraeum* likely represents a taxon other than *D. erythraeum* from the Arabian Peninsula. It seems likely that this accession is a misidentified *D. ursulae*, but as we have no voucher for the Indian specimen, we cannot conclude on this matter. However, Shelke et al. (2024) also reported an apparent polyphyly of *D. erythraeum*.

One accession of *D. rigidifolium* is retrieved in a polytomy with the large and heterogeneous *D. viride/erythraeum* clade B. In the target capture phylogeny, it forms a polytomy with two *D. viride* accessions, suggesting the two species are not genetically separated. *Dipcadi rigidifolium* resembles *D. viride* in the presence of caudate appendages on tepals but differs by having distinctive leaves that are rigid in texture with ribs and thickened margins. This morphological distinction supports *D. rigidifolium*'s status at species level, but a more extensive sampling is necessary to confirm this. As both *D. rigidifolium* and *D. erythraeum* are distinct morphologically, it seems likely that *D. viride* represents a species complex that could be resolved with additional sampling.

We provide no new data for the species *D. serotinum*, but as in Chapano et al. (in press), it appears polyphyletic. The Indian species *D. goaense*, *D. montanum*, *D. saxorum*, *D. concanense* and *D. ursulae* were included for a comprehensive overview of the genus, though a discussion of these taxa is beyond the scope of this study but is recently discussed in Shelke et al. (2024). We find no incongruences regarding the placement of these taxa relative to their phylogeny.

Interestingly, Shelke et al. (2024) retrieved a clade comprising *D. longifolium* and *D. gracillimum* (two specimens also included in our clade A) as sister to *D. viride*; the same relationship as in our target capture phylogeny.

Performance of angiosperms353 loci

Our phylogeny, based on 328 nuclear orthologs from the angiosperms353 probe set, does not give full resolution in the genus *Dipcadi* (Fig. 1, left). The alignments showed a low proportion of variable sites (23.8% on average), suggesting a recent divergence of the species within the genus. As a result, deep-level relationships can be harder to detect, as shown also by others (e.g., Murphy et al. 2020; Thomas et al. 2021a). Additionally, there is substantial gene tree discordance in our dataset, as indicated by the final quartet score from ASTRAL. While large datasets provide the framework for robust phylogenetic inference, they also contain conflicting phylogenetic signals. This is due to factors such as variable evolutionary rates across, incomplete lineage sorting (ILS), introgression, gene duplication and loss, horizontal gene transfer, hybridization and stochastic error (Soltis et al. 2009; Liu et al. 2015). We expect incomplete lineage sorting (ILS) in recent radiations like these (Shee et al. 2020; Murphy et al. 2020; Thomas et al. 2021a), but this result might also suggest some level of reticulate evolution in *Dipcadi*. Results on chromosome numbers in *Dipcadi* may also provide indications of reticulation (e.g., Samanta et al. 2023). Chromosome numbers are known for 13 species of the genus (Goldblatt and Manning 2011). Many are diploid with $2n = 12$, but $2n = 10$ is also known. Some occurrences of higher ploidy levels are reported both for $x = 5$ and $x = 6$. In addition, there are two reports of $2n = 22$. This number can be interpreted as the result of an allopolyploidization involving parents with $2n = 10$ and $2n = 12$, thus indicating reticulation in the genus.

To improve resolution, we analyzed supercontig sequences (i.e., both exons and flanking intronic regions) as inclusion of faster-evolving intronic regions have been beneficial in other cases involving rapid radiations (Larridon et al. 2020; Murphy et al. 2020; Bagley et al. 2020; Thomas et al. 2021a). However, in *Dipcadi*, the supercontigs did not improve resolution relative to the exon-only dataset. This might be attributed to alignment uncertainties following higher levels of missing data due to less sequencing coverage of intronic regions. Other authors have reported similar results from target capture projects (Johnson et al. 2016; Villaverde et al. 2018).

Our sampling is limited in terms of number of species, with 37% of currently described *Dipcadi* species represented. Additional sampling, including accessions spanning the whole diversity of the genus might improve resolution.

Comparison of target capture and ITS datasets

In this study, the two analyses resulted in similar tree topologies, despite the differences in dataset size. There were incongruences among deep nodes in the tree, yet we retrieve the same main clades containing the same terminal taxa (Fig. 1). We compared the two methods to test the reliability of ITS in species-level identification in *Dipcadi* and thus its utility as a DNA barcode marker.

Similarly to the target capture phylogeny, the ITS region did not fully resolve the topology of *Dipcadi* (Fig. 1, right and Fig. 2). Poor backbone resolution in the ITS phylogeny is likely due to sequence saturation. Due to its noncoding, rapidly evolving nature, the ITS region accumulates numerous mutations that can obscure deeper phylogenetic signals. As a result, the ITS region is not suited to infer higher-level phylogenetic relationships, particularly when using a single marker as the basis for phylogenetic reconstruction (Gee 2003; Felsenstein 2004). An important factor to consider when using ITS in phylogenetic inference is its presence as heterogeneous copies within a genome (Álvarez and Wendel 2003; Uncu et al. 2015; Hollingsworth et al. 2016). If the different copies of the ITS region within a genome are not homogenized as expected through concerted evolution, heterogeneous copies might persist (Álvarez and Wendel 2003). This phenomenon is particularly prevalent in plants (Krabberød et al. 2025). Chapano et al. (in press) mentioned that retrieval of heterogeneous versions of the ITS region might explain the seemingly distinct genetic lineages of *D. longifolium*. As we retrieve similarly divergent clades using the angiosperms353 loci, we believe that in our case, the ITS accurately reflects the genetic variation in *D. longifolium*.

Even though we find incongruences between the target capture and ITS phylogenies regarding deep relationships, ITS seems to separate *Dipcadi* species similarly to the angiosperms353 loci (Fig. 1). Although we cannot readily recommend the ITS region as a reliable marker for inferring phylogenetic relationships alone, in the case of *Dipcadi*, the region seems effective for identifying accessions to species level and for inferring the nearest sister species relations. The ITS region is well-known for its usability as a barcode marker (Hollingsworth et al. 2011; Antil et al. 2023), and our results support its barcoding power for the genus *Dipcadi*. This is a relevant finding for research projects with limited finances, infrastructure and/or bioinformatic skills. Moreover, amplifying the ITS region for *Dipcadi* is straightforward, although this is not the case for even closely related genera (Herrmann 2002; Lebatha et al. 2006). For phylogenetic reconstruction using more available methods that are cost-effective for large sample sizes, we also call for the detection of phylogenetically informative, single-copy nuclear markers that can be amplified and Sanger sequenced.

Conclusions

Our study provides the first HTS phylogeny of *Dipcadi* as well as an ITS phylogeny based on a large-scale sampling of the genus, highlighting the complementary strengths of targeted loci and ITS data. Both datasets support *D. marlothii* and *D. glaucum* as monophyletic species. Our data confirms that *D. longifolium* is polyphyletic, representing several distinct evolutionary lineages. One of the lineages is not genetically distinct from *D. gracillimum*, suggesting these names are used to describe the same entity. Our focused sampling of *D. longifolium* was not sufficient to resolve its taxonomy and describe the unnamed lineages. The habitat and geographic origin of the type specimen suggests that the name *D. longifolium* should be reserved for clade A, with the other lineages requiring different names. Our results corroborate that *D. viride* is paraphyletic, possibly representing a species complex in close association to *D. rigidifolium* and *D. erythraeum*. Both *D. rigidifolium* and *D. erythraeum* are morphologically distinct, and additional sampling is needed to draw clear species boundaries and define taxa within *D. viride*. While *D. platyphyllum* is distinct based on angiosperms353 loci, it is indistinguishable from *D. bakerianum* based on ITS, yet *D. bakerianum* was not included in the target capture dataset. Our study provides a robust framework for future taxonomic revisions of the genus *Dipcadi*.

Despite the large molecular dataset provided by angiosperms353 loci, the phylogenetic backbone of *Dipcadi* remains unresolved, likely due to conflicts in the dataset arising from incomplete lineage sorting (ILS) and potential reticulate evolution. Additional sampling of more *Dipcadi* species might also improve resolution. These challenges underscore the need for further investigations to clarify species boundaries and evolutionary relationships within *Dipcadi*. Comprehensive sampling of *Dipcadi*, targeted sampling of species, and an integrative approach is needed to achieve accurate delimitation of species. Understanding the alpha taxonomy of *Dipcadi* and the distributions of the species will help inform conservation strategies that preserve the diversity of *Dipcadi* species.

Our study also confirms that the ITS region can serve as a barcode marker in *Dipcadi* that can place samples within major clades and identify many species. This represents a valuable tool for research groups with limited resources.

Information on Electronic Supplementary Material

Online Resource 1. Material list with taxon name, voucher information (collector's name, abbreviation where used in figures, collection number, herbarium catalogue number or garden accession number where available); year of collection; country, specific locality of collection, habitat where available, coordinates; origin of molecular data; GenBank accession number.

Online Resource 2. Statistics retrieved from HybPiper.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00606-025-01947-0>.

Acknowledgements This work was funded by the project NOR-PART-2016/10013 and the Natural History Museum, University of Oslo, Norway. We thank Mika Bendiksby, Jesus Adrian Chimal Bal-leros and Ann Marie Evankow for comments on the manuscript. We acknowledge all those who have collected samples used in this study. We thank Lisbeth Thorbek for assistance in the laboratory.

Author contributions B. S. obtained funding for the study. B.S., E.W. and S.B.L. contributed to conception of the study. A. K., B.S. and E.W. collected plant material in the field. B.S. identified and provided additional material. A.K. performed laboratory work. A.K. and S.B.L. performed bioinformatic analyses. A.K. and S.B.L. wrote the main manuscript text. S.B.L. prepared figures. All authors reviewed and edited the manuscript. All authors agree to be accountable for all aspects of the work and approved the final version to be published.

Funding Open access funding provided by University of Oslo (incl Oslo University Hospital).

Data availability Raw Illumina sequencing reads are deposited in the European Nucleotide Archive. Alignments, gene trees and output files are available on Dryad. ITS sequences are uploaded to GenBank.

Declarations

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adobe Inc. 2024. Adobe Photoshop, version 26.0.0. Available at: <https://adobe.com/products/photoshop.html>
- Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19:716–723. <https://doi.org/10.1109/TAC.1974.1100705>
- Albert TJ, Molla MN, Muzny DM, Nazareth L, Wheeler D, Song X, Richmond TA, Middle CM, Rodesch MJ, Packard CJ, Weinstock GM, Gibbs RA (2007) Direct selection of human genomic loci by microarray hybridization. Nature Meth 4:903–905. <https://doi.org/10.1038/nmeth1111>
- Álvarez I, Wendel JF (2003) Ribosomal ITS sequences and plant phylogenetic inference. Molec Phylogen Evol 29:417–434. [https://doi.org/10.1016/S1055-7903\(03\)00208-2](https://doi.org/10.1016/S1055-7903(03)00208-2)
- Andermann T, Torres Jiménez MF, Matos-Maraví P, Batista R, Blanco-Pastor JL, Gustafsson ALS, Kistler L, Liberal IM, Oxelman B,

- Bacon CD, Antonelli A (2020) A guide to carrying out a phylogenomic target sequence capture project. *Frontiers Genet* 10:1407. <https://doi.org/10.3389/fgene.2019.01407>
- Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. Available at: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Antil S, Abraham JS, Sripoorna S, Maurya S, Dagar J, Makhija S, Bhagat P, Gupta R, Sood U, Lal R, Toteja R (2023) DNA barcoding, an effective tool for species identification: a review. *Molec Biol Rep* 50:761–775. <https://doi.org/10.1007/s11033-022-08015-7>
- Antonelli A, Clarkson JJ, Kainulainen K, Maurin O, Brewer GE, Davis AP, Epitawalage N, Goyder DJ, Livshultz T, Persson C, Pokorný L, Straub SCK, Struwe L, Zuntini AR, Forest F, Baker WJ (2021) Settling a family feud: a high-level phylogenomic framework for the Gentianales based on 353 nuclear genes and partial plastomes. *Amer J Bot* 108:1143–1165. <https://doi.org/10.1002/ajb2.1697>
- APG IV TAP (2016) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Bot J Linn Soc* 181:1–20. <https://doi.org/10.1111/boj.12385>
- Bagley JC, Uribe-Convers S, Carlsen MM, Muchhala N (2020) Utility of targeted sequence capture for phylogenomics in rapid, recent angiosperm radiations: Neotropical *Burmeistera* bellflowers as a case study. *Molec Phylogen Evol* 152:106769. <https://doi.org/10.1016/j.ympev.2020.106769>
- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Borowiec ML (2016) AMAS: a fast tool for alignment manipulation and computing of summary statistics. *PeerJ* 4:e1660. <https://doi.org/10.7717/peerj.1660>
- Brewer GE, Clarkson JJ, Maurin O, Zuntini AR, Barber V, Bellot S, Biggs N, Cowan RS, Davies NMJ, Dodsworth S, Edwards SL, Eiserhardt WL, Epitawalage N, Frisby S, Grall A, Kersey PJ, Pokorný L, Leitch IJ, Forest F, Baker WJ (2019) Factors affecting targeted sequencing of 353 nuclear genes from herbarium specimens spanning the diversity of Angiosperms. *Frontiers Pl Sci* 10:1102
- Buerki S, Jose S, Yadav SR, Goldblatt P, Manning JC, Forest F (2012) Contrasting biogeographic and diversification patterns in two Mediterranean-type ecosystems. *PLoS ONE* 7:e39377. <https://doi.org/10.1371/journal.pone.0039377>
- Chapano C, Løken SB, Kativu S, Zimudzi C, Bjørå CS, Stedje B (in press) Circumscription of African *Dipcadi* (Asparagaceae, Scilloideae) species based on ITS sequence data and morphology. *Phytotaxa* (in press)
- Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: More models, new heuristics and parallel computing. *Nature Meth* 9:772–772. <https://doi.org/10.1038/nmeth.2109>
- Deshpande AS, Mfirgafl A, Krfishnan S, Narkhede S, Janarthan MK (2016) Extended distribution of *Dipcadi concanense* (Dalzell) Baker - A highly threatened plant taxon of the family Asparagaceae. *J Threat Taxa* 8:9586–9588. <https://doi.org/10.11609/jott.2403.8.13.9586-9588>
- Dodsworth S, Pokorný L, Johnson MG, Kim JT, Maurin O, Wickett NJ, Forest F, Baker WJ (2019) Hyb-seq for flowering plant systematics. *Trends Pl Sci* 24:887–891. <https://doi.org/10.1016/j.tplants.2019.07.011>
- Felsenstein J (2004) *Inferring Phylogenies*. Sinauer Associates, Inc. Publishers, Sunderland
- Gee H (2003) Ending incongruence. *Nature* 425:782–782. <https://doi.org/10.1038/425782a>
- Glez-Peña D, Gómez-Blanco D, Reboiro-Jato M, Fdez-Riverola F, Posada D (2010) ALTER: program-oriented conversion of DNA and protein alignments. *Nucl Acids Res* 38:W14–W18. <https://doi.org/10.1093/nar/gkq321>
- Gnirke A, Melnikov A, Maguire J, Rogov P, LeProust EM, Brockman W, Fennell T, Giannoukos G, Fisher S, Russ C, Gabriel S, Jaffe DB, Lander ES, Nusbaum C (2009) Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat Biotechnol* 27:182–189. <https://doi.org/10.1038/nbt.1523>
- Goldblatt P, Manning JC (2011) A review of chromosome cytology in Hyacinthaceae subfamily Ornithogaloideae (*Albuca*, *Dipcadi*, *Ornithogalum* and *Pseudogaltonia*) in sub-Saharan Africa. *S African J Bot* 3:581–591
- Hart ML, Forrest LL, Nicholls JA, Kidner CA (2016) Retrieval of hundreds of nuclear loci from herbarium specimens. *Taxon* 65:1081–1092
- Herrmann N (2002) Biological Flora of Central Europe: *Ornithogalum angustifolium* nom. prov., Syn. p.p. *O. orthophyllum* ssp. *kochii* = *O. kochii* Parl., *O. gussonei* Ten. *Flora* 197:409–428. <https://doi.org/10.1078/0367-2530-00059>
- Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS (2018) UFBoot2: Improving the ultrafast bootstrap approximation. *Molec Biol Evol* 35:518–522. <https://doi.org/10.1093/molbev/msx281>
- Hollingsworth PM, Graham SW, Little DP (2011) Choosing and using a plant DNA barcode. *PLoS ONE* 6:e19254. <https://doi.org/10.1371/journal.pone.0019254>
- Hollingsworth PM, Li D-Z, van der Bank M, Twyford AD (2016) Telling plant species apart with DNA: from barcodes to genomes. *Philos Trans Roy Soc London B Biol Sci* 371:1–9
- Hyde M, Wursten, B. T., Ballings, P., Coates Palgrave, M. (2023) *Flora of Zimbabwe*. Available at: <https://www.zimbabweflora.co.zw/>. Accessed 3 Sep 2024
- Johnson MG, Gardner EM, Liu Y, Medina R, Goffinet B, Shaw AJ, Zerega NJC, Wickett NJ (2016) HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Appl Pl Sci* 4:1600016. <https://doi.org/10.3732/apps.1600016>
- Johnson MG, Pokorný L, Dodsworth S, Botigué LR, Cowan RS, DeVault A, Eiserhardt WL, Epitawalage N, Forest F, Kim JT, Leebens-Mack JH, Leitch IJ, Maurin O, Soltis DE, Soltis PS, Wong GKS, Baker WJ, Wickett NJ (2019) A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using k-medoids clustering. *Syst Biol* 16:594–606
- Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Meth* 14:587–589. <https://doi.org/10.1038/nmeth.4285>
- Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molec Biol Evol* 30:772–780. <https://doi.org/10.1093/molbev/mst010>
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>
- Krabberød AK, Stokke E, Thoen E, Skrede I, Kauserud H (2025) The ribosomal operon database: A full-length rDNA operon database derived from genome assemblies. *Molec Ecol Resour* 25:e14031. <https://doi.org/10.1111/1755-0998.14031>
- Larridon I, Villaverde T, Zuntini AR, Pokorný L, Brewer GE, Epitawalage N, Fairlie I, Hahn M, Kim J, Maguilla E, Maurin O, Xanthos M, Hipp AL, Forest F, Baker WJ (2020) Tackling rapid radiations with targeted sequencing. *Frontiers Pl Sci* 10:1655. <https://doi.org/10.3389/fpls.2019.01655>
- Lebatha P, Buys MH, Stedje B (2006) *Ledebouria*, *Resnova* and *Drimiopsis*: A tale of three genera. *Taxon* 55:643–652. <https://doi.org/10.2307/25065640>

- Lemmon EM, Lemmon AR (2013) High-throughput genomic data in systematics and phylogenetics. *Annual Rev Ecol Syst* 44:99–121. <https://doi.org/10.1146/annurev-ecolsys-110512-135822>
- Letunic I, Bork P (2021) Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucl Acids Res* 49:W293–W296. <https://doi.org/10.1093/nar/gkab301>
- Li H, Durbin R (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26:589–595. <https://doi.org/10.1093/bioinformatics/btp698>
- Liu L, Xi Z, Davis CC (2015) Coalescent methods are robust to the simultaneous effects of long branches and incomplete lineage sorting. *Molec Biol Evol* 32:791–805. <https://doi.org/10.1093/molbev/msu331>
- Manning J, Goldblatt P, Fay M (2004) A revised generic synopsis of Hyacinthaceae in Sub-Saharan Africa, based on molecular evidence, including new combinations and the new tribe Pseudoprosperaeae. *Edinburgh J Bot* 60:533–568. <https://doi.org/10.1017/S0960428603000404>
- Manning JC, Forest F, Devey DS, Fay MF, Goldblatt P (2009) A molecular phylogeny and a revised classification of Ornithogaloideae (Hyacinthaceae) based on an analysis of four plastid DNA regions. *Taxon* 58:77–107. <https://doi.org/10.1002/tax.581011>
- Martínez-Azorín M, Crespo MB, Juan A, Fay MF (2011) Molecular phylogenetics of subfamily Ornithogaloideae (Hyacinthaceae) based on nuclear and plastid DNA regions, including a new taxonomic arrangement. *Ann Bot (Oxford)* 107:1–37. <https://doi.org/10.1093/aob/mcq207>
- Maurin O, Anest A, Bellot S, Biffin E, Brewer G, Charles-Dominique T, Cowan RS, Dodsworth S, Epitawalage N, Gallego B, Giaretta A, Goldenberg R, Gonçalves DJP, Graham S, Hoch P, Mazine F, Low YW, McGinnic C, Michelangeli FA, Morris S, Penneys DS, Pérez Escobar OA, Pillon Y, Pokorný L, Shimizu G, Staggemeier VG, Thornhill AH, Tomlinson KW, Turner IM, Vasconcelos T, Wilson PG, Zuntini AR, Baker WJ, Forest F, Lucas E (2021) A nuclear phylogenomic study of the angiosperm order Myrtales, exploring the potential and limitations of the universal Angiosperms353 probe set. *Amer J Bot* 108:1087–1111. <https://doi.org/10.1002/ajb2.1699>
- McKain MR, Johnson MG, Uribe-Convers S, Eaton D, Yang Y (2018) Practical considerations for plant phylogenomics. *Appl Plant Sci* 6:e1038. <https://doi.org/10.1002/aps3.1038>
- McLay TGB, Birch JL, Gunn BF, Ning W, Tate JA, Nauheimer L, Joyce EM, Simpson L, Schmidt-Lebuhn AN, Baker WJ, Forest F, Jackson CJ (2021) New targets acquired: Improving locus recovery from the Angiosperms353 probe set. *Appl Plant Sci* 9. <https://doi.org/10.1002/aps3.11420>
- Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R (2020) IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molec Biol Evol* 37:1530–1534. <https://doi.org/10.1093/molbev/msaa015>
- Murphy B, Forest F, Barraclough T, Rosindell J, Bellot S, Cowan R, Golos M, Jebb M, Cheek M (2020) A phylogenomic analysis of *Nepenthes* (Nepenthaceae). *Molec Phylog Evol* 144:106668. <https://doi.org/10.1016/j.ympev.2019.106668>
- Obermeyer AA (1964) The South African species of *Dipcadi*. *Bothalia* 8:117–136. <https://doi.org/10.4102/abc.v8i2.1615>
- Ottenlips MV, Mansfield DH, Buerki S, Feist MAE, Downie SR, Dodsworth S, Forest F, Plunkett GM, Smith JFI, (2021) Resolving species boundaries in a recent radiation with the Angiosperms353 probe set: the *Lomatium packardiae*/L. *anomalum* clade of the *L. triternatum* (Apiaceae) complex. *Amer J Bot* 108:1217–1233. <https://doi.org/10.1002/ajb2.1676>
- Pfossor M, Speta F (1999) Phylogenetics of Hyacinthaceae based on plastid DNA sequences. *Ann Missouri Bot Gard* 86:852–875. <https://doi.org/10.2307/2666172>
- POWO (2024) Plants of the World Online. Royal Botanic Gardens, Kew. Available at: <https://powo.science.kew.org>. Accessed 9 May 2023
- Price MN, Dehal PS, Arkin AP (2009) FastTree: Computing large minimum evolution trees with profiles instead of a distance matrix. *Molec Biol Evol* 26:1641–1650. <https://doi.org/10.1093/molbev/msp077>
- Price MN, Dehal PS, Arkin AP (2010) FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS ONE* 5:e9490. <https://doi.org/10.1371/journal.pone.0009490>
- Prijbelski A, Antipov D, Meleshko D, Lapidus A, Al K (2020) Using SPAdes de novo assembler. *Curr Protoc Bioinf* 70:e102. <https://doi.org/10.1002/cpbi.102>
- Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA (2018) Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst Biol* 67:901–904. <https://doi.org/10.1093/sysbio/syy032>
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61:539–542. <https://doi.org/10.1093/sysbio/sys029>
- Samanta T, Jha TB, Ray S, Jha S (2023) Comparative cytogenetics and fluorescent chromosome banding in five Indian species of *Dipcadi* Medik. *Plants* 12:2534. <https://doi.org/10.3390/plants12132534>
- Shee ZQ, Frodin DG, Cámara-Leret R, Pokorný L (2020) Reconstructing the complex evolutionary history of the Papuanian *Schefflera* radiation through herbariomics. *Frontiers Plant Sci* 11:258. <https://doi.org/10.3389/fpls.2020.00258>
- Shelke PE, Tamboli AS, Surveswaran S, Yadav SR, Choo YS, Pak JH, Lekhak MM (2024) Molecular phylogenetic reconstruction improves the taxonomic understanding of Indian *Dipcadi* (Asparagaceae) and reveals a new species from the bank of Hiranyakeshi River, Maharashtra, India. *J Pl Res* 137:829–845. <https://doi.org/10.1007/s10265-024-01558-9>
- Slater G, Birney E (2005) Automated generation of heuristics for biological sequence comparison. *BMC Bioinform* 6:31. <https://doi.org/10.1186/1471-2105-6-31>
- Small RL, Cronn RC, Wendel JF (2004) Use of nuclear genes for phylogeny reconstruction in plants. *Austral Syst Bot* 17:145. <https://doi.org/10.1071/SB03015>
- Soltis DE, Albert VA, Leebens-Mack J, Bell CD, Paterson AH, Zheng C, Sankoff D, dePamphilis CW, Wall PK, Soltis PS (2009) Polyploidy and angiosperm diversification. *Amer J Bot* 96:336–348. <https://doi.org/10.3732/ajb.0800079>
- Stedje B, Kativu S (2023) Flora Zambesiaca: Hyacinthaceae. Royal Botanic Gardens, Kew
- Stedje B, Thulin M (1995) Synopsis of Hyacinthaceae in tropical East and North-East Africa. *Nordic J Bot* 15:591–601. <https://doi.org/10.1111/j.1756-1051.1995.tb02127.x>
- Stedje B (ed) (1996) Flora of Tropical East Africa: Hyacinthaceae. AA Balkema Publishers, Rotterdam
- Steenwyk JL, Buida TJ III, Li Y, Shen XX, Rokas A (2020) ClipKIT: A multiple sequence alignment trimming software for accurate phylogenomic inference. *PLoS Biol* 18:e3001007. <https://doi.org/10.1371/journal.pbio.3001007>
- Thiers B (2024) Index Herbariorum: A global directory of public herbaria and associated staff. New York Botanical Garden, Bronx. Available at: <http://sweetgum.nybg.org/science/ih/>. Accessed 14 Nov 2024
- Thomas AE, Igea J, Meudt HM, Albach DC, Lee WG, Tanentzap AJ (2021a) Using target sequence capture to improve the phylogenetic resolution of a rapid radiation in New Zealand *Veronica*. *Amer J Bot* 108:1289–1306. <https://doi.org/10.1002/ajb2.1678>
- Thomas SK, Liu X, Du ZY, Dong Y, Cummings A, Pokorný L, Xiang QY, Leebens-Mack JH (2021b) Comprehending Cornales: phylogenetic reconstruction of the order using the Angiosperms353

- probe set. *Amer J Bot* 108:1112–1121. <https://doi.org/10.1002/ajb2.1696>
- Uncu AO, Uncu AT, Celik I, Doganlar S, Frary A (2015) A primer to molecular phylogenetic analysis in plants. *Crit Rev Pl Sci* 34:454–468. <https://doi.org/10.1080/07352689.2015.1047712>
- Villaverde T, Pokorny L, Olsson S, Rincón-Barrado M, Johnson MG, Gardner EM, Wickett NJ, Molero J, Riina R, Sanmartín I (2018) Bridging the micro- and macroevolutionary levels in phylogenomics: Hyb-Seq solves relationships from populations to species and above. *New Phytol* 220:636–650. <https://doi.org/10.1111/nph.15312>
- Weitemier K, Straub SCK, Cronn RC, Fishbein M, Schmickl R, McDonnell A, Liston A (2014) Hyb-Seq: Combining target enrichment and genome skimming for plant phylogenomics. *Appl Pl Sci* 2:1400042. <https://doi.org/10.3732/apps.1400042>
- Zhang C, Rabiee M, Sayyari E, Mirarab S (2018) ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* 19:153. <https://doi.org/10.1186/s12859-018-2129-y>
- Zimmer EA, Wen J (2012) Using nuclear gene data for plant phylogenetics: Progress and prospects. *Molec Phylogen Evol* 65:774–785. <https://doi.org/10.1016/j.ympev.2012.07.015>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.