South African Journal of Botany 172 (2024) 474–487

Contents lists available at ScienceDirect

South African Journal of Botany

journal homepage: www.elsevier.com/locate/sajb

Leafing through genetic barcodes: An assessment of 14 years of plant DNA barcoding in South Africa

Ryan D. Rattray[a,*], Ross D. Stewart[a,b], Hendrik J. Niemann[c], Oluwayemisi D. Olaniyan[a,d], Michelle van der Bank[a]

[a] The African Centre for DNA Barcoding, Department of Botany and Plant Biotechnology, University of Johannesburg, P. O. Box 524, Auckland Park, Johannesburg, 2006, South Africa
[b] Department of Biological & Agricultural Sciences, Sol Plaatje University, Kimberley, 8300, South Africa
[c] School of Animal, Plant and Environmental Sciences, University of the Witwatersrand, Private Bag X3, WITS, 2050, South Africa
[d] Department of Biological and Environmental Sciences, Kampala International University, P. O. Box 20000, Kampala, Uganda

ARTICLE INFO

Article History:
Received 15 March 2024
Revised 11 June 2024
Accepted 24 July 2024
Available online 6 August 2024

Edited by O.M. Grace

Keywords:
DNA barcoding
Biodiversity conservation
Plant diversity
Conservation policy
Species identification

ABSTRACT

South Africa, a global biodiversity hotspot, faces escalating threats to its rich plant diversity, including habitat loss, climate change, invasive species, and illegal harvesting. These threats are further exacerbated by the country's taxonomic impediment, which hinders both conservation and sustainable development efforts. This paper assesses the efficiency of DNA barcoding as a tool for species identification and biodiversity conservation within the South African context. While DNA barcoding offers promising applications in conservation, significant gaps and challenges persist. We provide a comprehensive overview of plant barcoding initiatives in South Africa – by querying the public data portal for plant barcoding records, 12,456 published specimen records encompassing 159 families and ca. 3,449 species were returned. These numbers highlight historical progress, database contributions, technological advancements, and taxonomic coverage. Despite South Africa contributing the third-highest number of Magnoliophyta records to the Barcode of Life Data System (BOLD), significant gaps in endemic families and geographic regions highlight the urgent need for targeted DNA barcoding initiatives and increased collaboration, as only ca. 16 % of the known flora has been barcoded. The underutilisation of BOLD and financial constraints pose significant barriers to expanding plant barcoding records. However, advancements in sequencing technologies offer cost-effective solutions. We advocate for concerted efforts to enhance DNA barcoding utilisation, harmonise databases, and prioritise sampling of underrepresented taxa to effectively preserve South Africa's diverse plant life.

1. Introduction

South Africa, renowned for its rich biodiversity, is one of the world's 17 "megadiverse" countries – encompassing approximately 21,466 plant species (ca. 6 % of the world's plant diversity) with an impressive 76.68 % endemism (Hoveka et al., 2020; SANBI, 2023). Furthermore, South Africa is the only country on Earth to have within its national borders an entire floral kingdom (also known as a phytochorion) – one of just six in the world (Good, 1947; Goldblatt, 1997; Carlin et al., 2020). The Cape Floral Kingdom supports a rich biodiversity, hosting over 9000 plant species, with 70 % restricted to approximately 90,000 km^2 (Pirie et al., 2016). South Africa's biodiversity extends to three hotspots – the Cape Floristic Region (CFR; confined to the Western and Eastern Cape provinces), the Maputaland-Pondoland-Albany Hotspot (extends from the extreme southern parts of Mozambique and Mpumalanga through Eswatini and KwaZulu-Natal to the Eastern Cape) and the Succulent Karoo (found along the west coast of the Western Cape and Northern Cape provinces and extending into neighbouring Namibia). Nine terrestrial biomes have been identified within the region, ranging from deserts to forests. In addition to the recognised biomes in South Africa, two unique biomes, namely the sub-Antarctic tundra situated in lowland areas and the sub-Antarctic polar desert exclusive to higher elevations (Smith and Mucina, 2006), have been identified on the subantarctic islands within the South African territory. Marion Island and Prince Edward Island hosts 22 and 21 indigenous vascular plant species, respectively.

South Africa's commitment to conservation dates to the establishment of Africa's first protected areas in the 1890s with the priority of conserving wild-life or "game". This led to the first proclaimed protected areas being Sabi (est. 1895) and Shingwedzi (est. 1903) –

---

* Corresponding author.
E-mail addresses: ryanr@uj.ac.za (R.D. Rattray), ross.d.stewart@gmail.com (R.D. Stewart), hennieniemann92@gmail.com (H.J. Niemann), oluwayemisi.olaniyan.oo@gmail.com (O.D. Olaniyan), mvdbank@uj.ac.za (M. van der Bank).

https://doi.org/10.1016/j.sajb.2024.07.055

which later became the Kruger National Park (the first national park proclaimed in 1926; Van Wilgen et al., 2020). The Department of Forestry, with the aim of protecting forest and water resources, established protected areas in mountain water catchment areas (Langberg, est. 1886; Cederberg, est. 1895), coastal areas (Walkerbay, est. 1895), and indigenous forests (Knysna forest, est. 1894) (Van Wilgen et al., 2020). A 2021 report compiled through the collaborative efforts of Statistics South Africa (Stats SA), the South African National Biodiversity Institute (SANBI), and the Department of Forestry, Fisheries and the Environment (DFFE) states that in 2020, protected areas accounted for 112,807 km^2 (9.2 %) of the total land area of the country (Statistics South Africa, 2021).

In 2019, a status update on South Africa's biodiversity reported that 20,401 plant species had been assessed and listed on SANBI's Red List – 2,804 of which were threatened with extinction. Unfortunately, the situation seems to be on a downward trajectory with 17 % of revised assessments in the most recent iteration of the list (SANBI, 2024) being for species that have become even more threatened due to either intensifying existing threats or emerging threats. Furthermore, South Africa is reported as having the second-highest number of plant extinctions per country worldwide, with 36 species confirmed extinct and a further 70 possibly extinct (DEA, 2019). This worrying loss of species diversity is a direct consequence of factors such as changes in climate conditions, habitat loss, invasive species, and the illegal wildlife trade.

2. Threats to biodiversity

2.1. Plant poaching and illegal trade

The illegal wildlife trade, specifically illegal plant collections, gravely threatens biodiversity globally. Moreover, succulents have become critically threatened through illegal over-collecting for the ornamental trade (Margulies et al., 2022). In South Africa, a drastic increase in the illegal harvesting of succulent plants across the country's Arid Zone, particularly the Succulent Karoo Biome, has become of great concern (DFFE, 2022). Between 2020 and 2022, government agencies noted a sharp increase of 250 % per annum in illegal harvesting of endemic species to supply the global horticultural trade, with ca. 1.5 million plants removed from the wild since 2020 (Smith et al., 2023). The impact of this is evident in the uplisting of 85 *Conophytum* N.E.Br. species in the most recent iteration of SANBI's Red list (SANBI, 2024). Originally designated as Least Concern, Critically Rare, and Rare in 2020, all of these species are currently listed under a category of threat. Currently, 96 % of all *Conophytum* species are categorised as Vulnerable or higher. Numerous other caudiform and geophytic species are facing challenges as well, with patterns shifting quickly as more species are targeted by the expanding illicit wildlife trade.

Another notable threat to South African biodiversity is what has been described as the "South African cycad extinction crisis", whereby the country is at risk of losing 50 % of its cycads within the next 10 years. Of the 37 *Encephalartos* Lehm. species in the region, four are extinct in the wild, and 32 species have a conservation status ranging from vulnerable to critically endangered as a direct result of anthropogenic activities such as habitat loss and illegal harvesting (Williamson et al., 2016; Stewart et al., 2023). The resulting impacts have had dire consequences on several endemic species, such as *Conophytum herreanthus* subsp. *herreanthus* S.A.Hammer (Azioaceae), *Encephalartos brevifoliolatus* Vorster and *E. nubimontanus* P.J.H. Hurter (Zamiaceae), which have all been reported as extinct in the wild between 2009 and 2015 (DFFE, 2022; Van der Colff et al., 2023).

2.2. Climate change

Over the second half of the 20th century, the world's ecosystems have changed more rapidly than ever. Rising levels in both CO$_2$ and temperature associated with climate change have led to changes in hydrological cycles and extreme weather patterns (droughts and floods; Habibullah et al., 2022). The impact of climate change in South Africa is highlighted by the higher temperature increase and extreme rainfall occurrences experienced in the country over the past 50 years, with the mean annual temperature increasing 1.5 times the global average and mean precipitation increasing by 2 % – both of which have serious ramifications for the distribution and survival of plant species, especially endemics, within the region (Hoveka et al., 2022).

Greater mean precipitation can also negatively impact plant biodiversity by leading to significant reductions in species richness, promoting grass dominance at the expense of other plant types, and increasing physiological stress and mortality rates among plants due to more extreme wet and dry periods (Reynaert et al., 2021). These alterations disrupt ecosystem balance and diminish overall biodiversity (Reynaert et al., 2021). Therefore, climate change has now been recognised as a serious concern, with data predicting the growth or decline of many suitable habitats, shifts in phenological patterns and growth, and the possibility of local extinctions under various climate-controlled scenarios, such as present-day and future climate conditions. These scenarios have predicted that South Africa's threatened endemic species may experience changes in their range and distributions by 2050. Based on such data, improvements must be made to protected areas and incorporated into species extinction risk assessments (Hoveka et al., 2020). Moreover, climate change significantly intensifies wildfires, increasing their frequency and severity. Higher global temperatures result in drier conditions and extended fire seasons, facilitating ignition and spread. More frequent and severe droughts dry out vegetation, heightening fire risk. Altered precipitation patterns impact fuel moisture and vegetation growth, changing wildfire dynamics. Extreme weather events like heatwaves, strong winds, and lightning storms are becoming more common, further promoting wildfire conditions (Kraaij et al., 2018). Climate change also influences vegetation distribution and abundance, affecting fuel availability. As human populations expand into wildland areas, the vulnerability of the wildland-urban interface to wildfires increases. Additionally, wildfires release greenhouse gases, creating feedback loops that exacerbate global warming and future fire behaviour. Understanding these complex interactions is vital for developing effective wildfire management and adaptation strategies (Kraaij et al., 2018).

The impact of climate change on biodiversity within the South African context is evident in the most recent iteration of SANBI's Red List (SANBI, 2024). The findings indicate that climate change had an impact on 12 % of the recently updated assessments. The Namaqualand, in particular, saw substantial population declines as a result of the prolonged drought that began in 2012 and continued until 2021. The tree aloes *Aloidendron pillansii* (L.Guthrie) Klopper & Gideon F. Sm. and *Aloidendron dichotomum* (Masson) Klopper & Gideon F.Sm., as well as other iconic species of the Richtersveld, such as *Pachypodium namaquanum* (Wyley ex Harv.) Welw. (the '*halfmens*') and *Aloe pearsonii* Schönland, are just a few disturbing examples of South African species declining as a result of climate change.

2.3. Land degradation, fragmentation, and habitat loss

Land degradation refers to the adverse alteration of land conditions caused by human activities, including climate change (IPCC, 2019). This decrease in biological productivity, ecological integrity, and overall land value undermines the multifaceted significance of the land, encompassing its ecological, social, and economic dimensions. Approximately 3.2 billion people, or over 10 % of the global annual gross product, are affected by land degradation, resulting in biodiversity loss and diminished ecosystems. The extent of degraded land ranges from under 1 billion ha to around 6 billion ha, with observed degradation levels of 15 % to 63 % between 1977 and 2003

(Mani et al., 2021). Roughly 25 % of Earth's land area is acknowledged as degraded – impacting 1.3 to 3.2 billion people in poverty. The drivers of land degradation include biophysical, institutional (land governance, land tenure systems, policy frameworks, economic factors and resource management practises), and socio-economic factors (Mani et al., 2021).

The key factors driving land degradation in South Africa are historical inequity in land access, unsustainable land management practices, overgrazing, high population-to-land ratio, and woody plant encroachment. These factors are interconnected with issues of justice and climate change, where historical injustices in land distribution (particularly due to apartheid policies) have led to high population densities and livestock numbers, contributing to land degradation. Unsustainable land management practices, influenced by complex land tenure systems, worsen degradation. Woody plant encroachment, influenced by climate change, overgrazing, and unsustainable land management further exacerbates degradation – about 60 % of the country's land is degraded, and 91 % is susceptible to desertification (Mani et al., 2021). Continued mining activities in the Northern Cape have also been identified as a primary contributor to the remarkable increase in the number of threatened species that have been added to the SANBI Red List from this province (SANBI, 2024).

Habitat loss has been the biggest driver of the extinction of South African plant species between 1750 and 2021. During this period, 26 species became extinct primarily due to urban and agricultural expansion, afforestation, and dam construction (Van der Colff et al., 2023).

2.4. Invasive alien species

South Africa displays the imprint of European colonisation on its ecosystems. Alien plant species were minimal before the seventeenth century, but introductions worldwide began in the 1600s for purposes such as agriculture, timber, and ornamental gardens. Forestry initiatives in the nineteenth and mid-twentieth centuries, largely influenced by government efforts, introduced numerous tree species like *Acacia* Mill. (Australian wattles), *Eucalyptus* L'Hér., and *Pinus* L.; with World War I, afforestation intensified with the cultivation of over 400 tree species (Richardson et al., 2020). Furthermore, despite a diverse native flora, the persistent demand for ornamental alien plants, dating back to 1652, has led to widespread invasions around human settlements (Richardson et al., 2020).

Invasive alien species pose various significant threats to native plants and ecosystems. These species tend to outcompete native plants for resources, disrupt ecosystem processes, reduce biodiversity, alter soil properties, increase the risk of wildfires, deplete water resources, and lead to secondary invasions by other invasive species. The impact of alien species on South African flora is a complex issue that necessitates integrated management strategies to mitigate negative effects and safeguard the region's unique biodiversity (Erckie et al., 2022; Zengeya and Wilson, 2023). For instance, the presence of *Lantana camara* L. (commonly known as lantana) induces alterations in the physical structure of the ecosystem, precipitating shifts in community composition. This, in turn, results in a decrease in both invertebrate abundance and species richness – concomitant with a decline in plant species diversity within Groenkloof Nature Reserve, South Africa (Samways et al., 1996; Raphela and Duffy, 2022).

3. Navigating taxonomic impediments in biodiversity

Finding practical solutions to the rising threats to biodiversity is of great importance. However, accurate taxonomic data is crucial for implementing any measure undertaken. Taxonomists are responsible for generating and curating the baseline data critical to successful conservation and sustainable development. Nonetheless, factors such as large gaps in taxonomic knowledge, a decline in the number of experts, poor funding, and inadequate infrastructure hinder taxonomic work generally in many regions worldwide (Coleman, 2015). These factors, which have come to be known as 'taxonomic impediments', may hinder the progress of cataloguing and understanding the full extent of a country's biological diversity. South Africa is not left out in facing these taxonomic impediments. Despite having a rich biodiversity, the country also struggles with gaps in taxonomic knowledge and a shortage of experts, which can impact conservation efforts and sustainable development initiatives. In a comprehensive assessment of the state of plant taxonomy in South Africa, Victor and Smith (2011) and Victor et al. (2015) concluded that there is a critical shortage of human capacity in South Africa to conduct plant taxonomic research to benefit its biodiversity. These studies also highlighted the lack of funding and research material as major obstacles hindering the advancement of plant taxonomy in South Africa. Consequently, the progress with taxonomic treatments of South African plants, such as the Flora of Southern Africa project, has been slower than expected, with only a fraction of species treated after many years – impacting the availability of up-to-date information for conservation assessments. Buys and Smith (2006) estimated the discovery and description of new plant species in South Africa to 60 per year between 1995 and 2005, contrasting with the global rate of 2350 new species described annually. This suggests a comparatively slower pace of species documentation within the region. The rates of new species descriptions, when expressed as percentages, highlight the significance of undiscovered plant species in South Africa compared to the global context. In South Africa, the estimated 1400–1575 yet-to-be-described species represent approximately 12.5 % to 14 % of the total described species (11,208) (Mamathaba et al., 2022). This indicates a significant opportunity for new discoveries within the country's rich biodiversity. Globally, these unknown species constitute less than 1 % (about 0.36 % to 0.40 %) of the estimated 390,000 known plant species, showing that while important nationally, South Africa's undiscovered species form a minor part of the global botanical landscape (Mamathaba et al., 2022). These percentages underscore the need for continued taxonomic research in South Africa to fully document its flora within the broader context of global plant diversity. The richness and complexity of the South African flora also make it difficult for taxonomists to identify the plants' biodiversity in the region. In a study that sought to identify the biodiversity knowledge gaps for conserving South Africa's endemic flora, Hoveka et al. (2020) attributed the challenge of fully documenting and researching all plant species in South Africa to the extensive species diversity in the region. According to them, South Africa is home to an estimated 250,000 to 1000,000 plant and animal species collectively, including over 13,000 plant species that are endemic to the region, thereby making many species remain poorly understood and threatened with extinction. Additionally, the high level of endemism in the region further complicates taxonomic efforts. Regions such as the CFR, the Succulent Karoo, and the Maputaland-Pondoland-Albany are recognised as global biodiversity hotspots with significant levels of endemism. Identifying and studying these endemic species requires focused sampling endeavours and specialised expertise already in short supply.

The existence of "cryptic" plant species adds additional complexity to the challenge of taxonomic impediments in South Africa. Cryptic plant species are difficult to differentiate based on traditional morphological characteristics because they appear similar but are genetically distinct. This is evident in the *Salicornia meyeriana* Moss (Slenzka et al., 2013) complex, where two subspecies, '*overbergensis*' and '*soetendaliana*,' were described despite their morphological similarities. These subspecies were identified using diagnostic mutations in the External Transcribed Spacer (ETS) region, and their unique distribution and ecological preferences (Slenzka et al., 2013). The '*overbergensis*' subspecies is associated with estuarine habitats, while the '*soetendaliana*' subspecies is found in inland areas of the Overberg region in South Africa. Despite the challenge in morphological differentiation, the genetic, geographical, and ecological distinctiveness of these subspecies allowed for their recognition and description (Slenzka et al., 2013). The difficulty in distinguishing cryptic species can lead to misidentification or underestimation of species diversity, impacting conservation strategies and management decisions.

These taxonomic impediments, compounded by the complexities of South African plant diversity, highlight the importance of finding innovative solutions. Incorporating genetic data, such as that obtained through DNA barcoding, has emerged as a promising approach to mitigate these challenges and foster collaboration between molecular approaches and traditional taxonomy.

4. DNA barcoding as a tool for species identification and biodiversity conservation

DNA barcoding is a molecular technique developed by Paul Hebert (Hebert et al., 2003) to identify a wide range of taxa based on short, standardised DNA sequences from a specific genome region. His core idea behind DNA barcoding was to assign a unique genetic barcode to each species, similar to how product barcodes are used for identification in stores.

The protocols for DNA barcoding typically involve several key steps, such as DNA extraction, PCR amplification, sequencing, and data analysis. First, DNA is extracted from the tissue or specimen of the organism of interest using various extraction methods, such as silica-based extraction, CTAB-based extraction, or commercial DNA extraction kits. Then, a specific DNA region, such as the mitochondrial region cytochrome c oxidase subunit 1 (*CO1*) in animals or the plastid regions ribulose bisphosphate carboxylase large subunit (*rbcL*) and maturase K (*matK*) in plants, can be amplified using polymerase chain reaction (PCR) with primers designed to target the barcode region. This step generates multiple copies of the DNA fragment for sequencing. Next, the amplified DNA fragments are sequenced using Sanger sequencing or high-throughput sequencing technologies like Illumina, PacBio, or Oxford Nanopore. The resulting sequences represent the DNA barcode of the organism. The obtained DNA barcode sequences are compared to reference databases like BOLD (Barcode of Life Data Systems) or the National Center for Biotechnology Information's (NCBI) GenBank to identify the species. Bioinformatics tools are used to analyse the sequences, perform alignments, and assess genetic distances for species identification.

In many animal groups, the "Folmer region" of the *CO1* gene (a 650 bp fragment at the 5′end of the gene) is often used as a single barcode marker to identify all animal species. However, it has been difficult to find a single barcode marker for land plants as most candidate loci did not meet the criteria for selecting a DNA barcoding marker *viz*: universal applicability of primers; genetic variability between species; conservation of markers; amplifiability of primers using PCR; high resolution of the marker; ease of primer design, and compatibility with sequencing technologies and bioinformatics tools (Letsiou et al., 2024). These criteria ensure the markers across diverse organisms, their ability to differentiate species accurately, and their compatibility with laboratory techniques for efficient and reliable species identification in plant DNA barcoding studies. Fazekas et al. (2008) attributed the difficulty of finding a universal plant barcode marker to factors such as higher levels of genetic diversity, complex evolutionary histories, hybridisation, and polyploidy – making species boundaries difficult to define.

To address these limitations, the Consortium for the Barcode of Life (CBOL) Plant Working Group (2009) was tasked with testing the seven leading candidate chloroplast barcoding regions at the time (*atpF–atpH, matK, psbK–psbI, rbcL, rpoB, rpoC1* and *trnH–psbA*) and finding a suitable DNA barcode for land plants among them. This required the collection, pooling, and collaborative analysis of data to evaluate the candidate markers in relation to three criteria: sequence quality, discriminatory power, and universality (ease of amplification and sequencing). Four of the seven barcodes could be ruled out based on the aforementioned criteria: *atpF–atpH* due to limited recovery of high-quality bidirectional sequences and falling below the median for species resolution in both single and multilocus barcodes, *psbK–psbI* for having the lowest overall sequencing success during the trials and for also exhibiting significant issues producing bidirectional reads, and both *rpoB* and *rpoC1* as a result of their poor discriminating power.

Selecting a plant barcode among the three plastid loci that remained was challenging because *matK, rbcL*, and *trnH–psbA* each had distinct advantages and disadvantages. The majority preference of the CBOL Plant Working group (2009) was to suggest a 2-locus core barcode for terrestrial plants made up of portions of *rbcL+matK* that would be supplemented with further markers as needed. The decision was mostly based on the simple recovery of *rbcL* and the strong discriminatory ability of *matK*, which together offset each other's weaknesses and produced a higher level of species resolution than the majority of the multilocus barcodes investigated as well as all seven single locus barcodes. Additionally, because both markers are coding regions, errors in editing or assembly, the existence of pseudogenes, and proper sequence orientation can all be automatically checked by electronically translating DNA sequences to amino acids.

Supplementary loci often include the plastid *trnH-psbA* intergenic spacer because of its higher species discrimination success and/or the nuclear ribosomal internal transcribed spacers (nrITS or nrITS2) because it generally provides improved species resolution (Okuyama and Kato, 2009; Hollingsworth et al., 2011). Nonetheless, universal agreement on a single combination is still lacking, but exploring diverse marker combinations continues to improve the effectiveness of plant DNA barcoding for precise species identification.

Ideally, an appropriate DNA barcode should display a clear "barcoding gap" (Meyer and Paulay, 2005); i.e., the genetic distances between species (interspecific variability) are greater than the genetic distances within a species (intraspecific variability). It was identified at an early stage as being essential to the success of DNA barcoding. Indeed, the presence of a species' barcoding gap has frequently been cited as proof that DNA barcoding "works" in real-world scenarios (Stoeckle and Thaler, 2014). Using a taxonomically understudied group like cowries (marine gastropods in the Cypraeidae family) as an example, Meyer & Pauley (2005) pointed out that there was substantial overlap between intra- and interspecific variation, which usually compromises species identification. In such cases, thresholds are often used to separate variations, which might increase error rates. Previous research (Phillips et al., 2024) shows that the existence of a gap depends on sampling effort for target taxa and their close relatives and can, therefore, be improved upon with better-optimized sampling strategies. In cases where a DNA barcode might not be present/sufficient, the application of a character-based DNA barcode approach can also be a solution. It is an approach that converts sequence data into diagnostic characters and can be used to distinguish between different species, particularly in cases when there is little to no interspecific variation or a "barcoding gap." (DeSalle 2006, 2007; Wiemers and Fiedler 2007; Rach et al. 2008; Waugh et al. 2008; Gibbs, 2018). Many problems associated with a lack of a DNA barcode gap can also be solved by employing next-generation sequencing approaches.

The future of plant DNA barcoding lies in next-generation sequencing (NGS) technologies, through the generation of very long sequences often referred to as super barcodes. Super barcodes, like the complete chloroplast genome, have demonstrated potential in differentiating closely related plant species (Li et al., 2015; Wu et al., 2021). The chloroplast genome offers more sequence diversity and can differentiate between closely related plants, providing a comprehensive approach to plant DNA barcoding. For example, this approach has been successfully applied to identify closely related species of *Aloidendron* (A.Berger) Klopper & Gideon F.Sm. (Malakasi et al., 2019), *Araucaria* Juss. (Ruhsam et al., 2015), *Camellia* L. (Yang et al., 2013), *Echinacea* Moench (Zhang et al., 2017), *Epimedium* L. (Zhang et al., 2016), *Fritillaria* Tourn. ex L. (Wu et al., 2021), *Taxus* L. (Fu et al., 2019), and *Theobroma* L. (Kane et al., 2012). Complete plastid sequencing, however, does not solve the fundamental problem that plastid genomes do not always reflect species boundaries (Rieseberg and Soltis, 1991; Nichols, 2001; Hollingsworth et al., 2011; Hollingsworth et al., 2016). Thus, although the continued use of complete plastid genomes as barcodes is expected, substantial gains in resolving power are expected to come from NGS technologies utilising the nuclear genome (on its own or in conjunction with plastid and/or mitochondrial sequence data).

Super barcodes in plants could be produced in various ways utilising next-generation techniques; the optimal approach would take several factors into account. Annotated chromosome-scale whole-genome sequences (WGS) are thought to be the best data set for super barcodes (Pezzini et al., 2023); however, this application is rarely possible due to project budget constraints (especially in the context of developing countries like South Africa) and the quality of plant material required. For most studies, DNA is extracted from dried plant material obtained from silica collections or herbaria – with DNA from the latter often being degraded with fragments shorter than 100 bp (Forrest et al., 2019). The use of WGS is challenging in this situation. Hence, short-read sequencing technologies must be employed instead. Numerous strategies have been devised and implemented by researchers to generate thousands of markers from the nuclear, plastid, and mitochondrial genomes – enhancing the resolution and accuracy of plant species identification. These approaches are expected to become more accessible as NGS costs continue to decrease with continued technological advancements, market competition, economic scalability, and new competitors with innovative approaches. While genotyping by sequencing (GBS) and restriction site-associated genotyping (RAD) are primarily useful for plant collection management as well as diversity and evolutionary investigations (Villano et al., 2023), genome skimming and target capture are considered to be more suitable for super barcoding (Guo et al., 2023).

Genome skim data has successfully been used as a super barcode to separate taxonomically complicated taxa like *Rhododendron* L. (Fu et al., 2022) and *Schima* Reinw. ex Blume (Yu et al., 2022). Full skim data has been investigated for its potential wide-spread application as a plant DNA barcode (Coissac et al., 2016); however, the more loci used, the more problems with hybridisation, introgression, and general incongruence are exposed, and the more complicated it is to distinguish between species in many clades (Bohmann et al., 2020). While high-copy-number areas are conserved between samples (making genome skimming data relatively easy to combine), the difference in coverage caused by the diverse nuclear genome structure of plants may result in missing data when combining sequences from multiple samples. For instance, an equivalent amount of reads from a species with a small genome will result in a higher average coverage of organelle sequences than a set number of reads from a species with a large genome (Twyford and Ness, 2017).

The lack of a universal probe set with adequate phylogenetic breadth has hindered the viability of target enrichment-based DNA barcoding despite its potential. A possible remedy to this problem has emerged in recent years – the Angiosperms353 probe set (Brewer et al., 2019). Compared to the well-established DNA barcoding technique based on the Sanger sequencing of the plastid marker-based core plant barcode, targeted sequence capture, like many of the other NGS approaches, produces a lot more data, which increases the potential capacity for species discrimination. The method is also effective with poor-quality material, and coverage is not affected by fluctuations in genome size like genome skimming is.

Nevertheless, additional proof of the Angiosperms353 probe set's ability to discriminate between species is needed to develop it into a practical barcoding tool. Although evidence is growing that it is highly informative in lower-level phylogenetic studies (e.g., Murphy et al., 2020), few studies exist directly comparing its performance to standard plant DNA barcodes. Stewart (2021) found that the Angiosperms353 probe set provided modest discrimination gains compared to utilising standard plant DNA barcodes with additional loci (*matK, rbcLa, psbA-trnH, trnL-F, ycf1*, and ITS1) in *Gasteria* Duval. Because they were created to resolve deeper nodes in the Angiosperm tree of life, the low copy number nuclear loci targeted by universal bait kits, such as Angiosperms353, are fairly conserved genes (McDonnell et al., 2021; Woudstra et al., 2021) – limiting its application in clades that have recently and/or rapidly diversified. Additionally, the Angiosperm353 panel's gene recovery is typically subpar in Monocot clades (e.g., less than 37 % in *Cyperus* L., Larridon et al., 2020), which further restricts its utility. In order to address this and achieve very high resolution and sequencing success spanning large phylogenetic distances, a more comprehensive assessment of the best instances, methods, and scales for combining taxon-specific bait sets with a universal bait set is required (Manzanilla et al., 2022). The need for taxon-specific bait sets, however, undermines the three core tenets of DNA barcoding – standardisation, minimalism, and scalability (Hollingsworth et al., 2011).

Targeted sequence capture-based DNA barcoding implementation faces additional problems, such as the requirement to create a reference library and the related costs and technological difficulties (in both lab and bioinformatics). Additionally, traditional plant DNA barcoding markers should also be captured by a targeted sequencing approach to barcoding (Hollingsworth et al., 2016) – allowing for continued use of the reference libraries that are already available. Until these shortcomings are adequately addressed, it is unlikely that a target-capture-based DNA barcoding approach will see broad-scale implementation within South Africa in the foreseeable future.

NGS has also been used to leverage existing DNA barcode libraries. DNA meta-barcoding is a method used for detecting many species in one sample (often an environmental sample) by amplifying and sequencing multiple DNA barcodes at once. By combining traditional DNA barcoding with NGS, it allows for the identification and quantification of various taxa in a complex biological sample – including environmental samples like water, sediment, or soil samples containing eDNA (genetic material shed into the environment by all organisms, often without any obvious signs of biological source material; Hajibabaei, 2012; Thomsen and Willerslev, 2015; Aylagas et al., 2016). Unlike conventional DNA barcoding, which focuses on identifying individual specimens at the species level, meta-barcoding is designed to detect a wide range of species present in a sample, often to the family level or higher taxonomic groups. Meta-barcoding can be used in many ways including biodiversity surveys, ecological research, food authentication, forensic science, and environmental assessment. In meta-barcoding, the ability to distinguish between closely related species will also depend on the variability of the chosen barcode region (Cristescu, 2014).

The progress in the broad-scale implementation of DNA barcoding technology has transformed it into a valuable tool for biologists across various research fields and applications, including species identification, conservation initiatives, monitoring biological invasions and biodiversity conservation (Krishnamurthy and Francis, 2012). The versatility of DNA barcoding is evident in species identification, where it offers a cost-effective and standardised approach for discovering new species objectively, independent of morphology or life stage (Hebert et al., 2003). DNA barcodes contribute significantly to habitat conservation by enhancing phylogenetic diversity mapping, aiding in biodiversity assessments and conservation planning with minimal fieldwork and cost-efficiency (Farooq et al., 2020).

The success of DNA barcoding in accurately identifying plant species relies heavily on the availability and quality of reference libraries that contain DNA sequences from known species. These reference libraries are a comparison tool for analysing the DNA sequences obtained from unknown plant samples during barcoding. The Barcode of Life Data System (BOLD) is a global barcode database that allows different regions to build a reference library particular to the species in their region. South Africa, for example, has several dedicated DNA barcoding initiatives such as "The African Centre for DNA Barcoding" (ACDB) and “The South African Node of the Barcode of Life Consortium” (SABOL), which focuses on the building of barcoding reference libraries specifically for South African species (animal and plant). These initiatives aim to enhance the accuracy and efficiency of plant identification in the region by providing researchers with a valuable resource for comparison and analysis.

5. Methodology

To assess the status of plant barcoding in South Africa, we conducted an analysis in July 2023, focusing on all Tracheophyta (land plants) phyla, which include "Cycadophyta", "Gnetophyta", "Magnoliophyta", "Pinophyta", and "Pteridophyta". Bryophyta were excluded due to the limited number of records (53) available. Data were retrieved from the Barcode of Life Data Systems (BOLD) public database, chosen for its status as the world’s largest repository of curated DNA barcodes; BOLD also mines sequence data from Genbank.

Information on locality (province and GPS coordinates) and taxonomic classification was collected for each record. This dataset was then compared with the South African 2023 plant checklist using R to identify knowledge gaps. The compositional breakdown between barcode genes and other genetic markers was analysed by querying the South African plant checklist on BOLD (checklist number CL-SAPLA, managed by Johandré van Rooyen).

6. Status of plant barcoding in South Africa

In total, 12,456 published specimen records were identified from ca. 3449 species, encompassing 965 genera and 159 families. According to BOLD, South Africa has the third most Magnoliophyta records, which, at the time of query, had 11,810 publicly available plant records. Our paper aims to provide an up-to-date and comprehensive review of the current progress on plant barcoding in South Africa.

6.1. Historical overview of plant barcoding initiatives in South Africa

The first plant barcoding project in South Africa (SA) was launched in 2007 by the African Centre for DNA Barcoding (ACDB), intending to barcode the Flora of the Kruger National Park (Lahaye et al., 2008). Since then, enormous amounts of resources, including time and funding, have been spent generating and contributing records to various

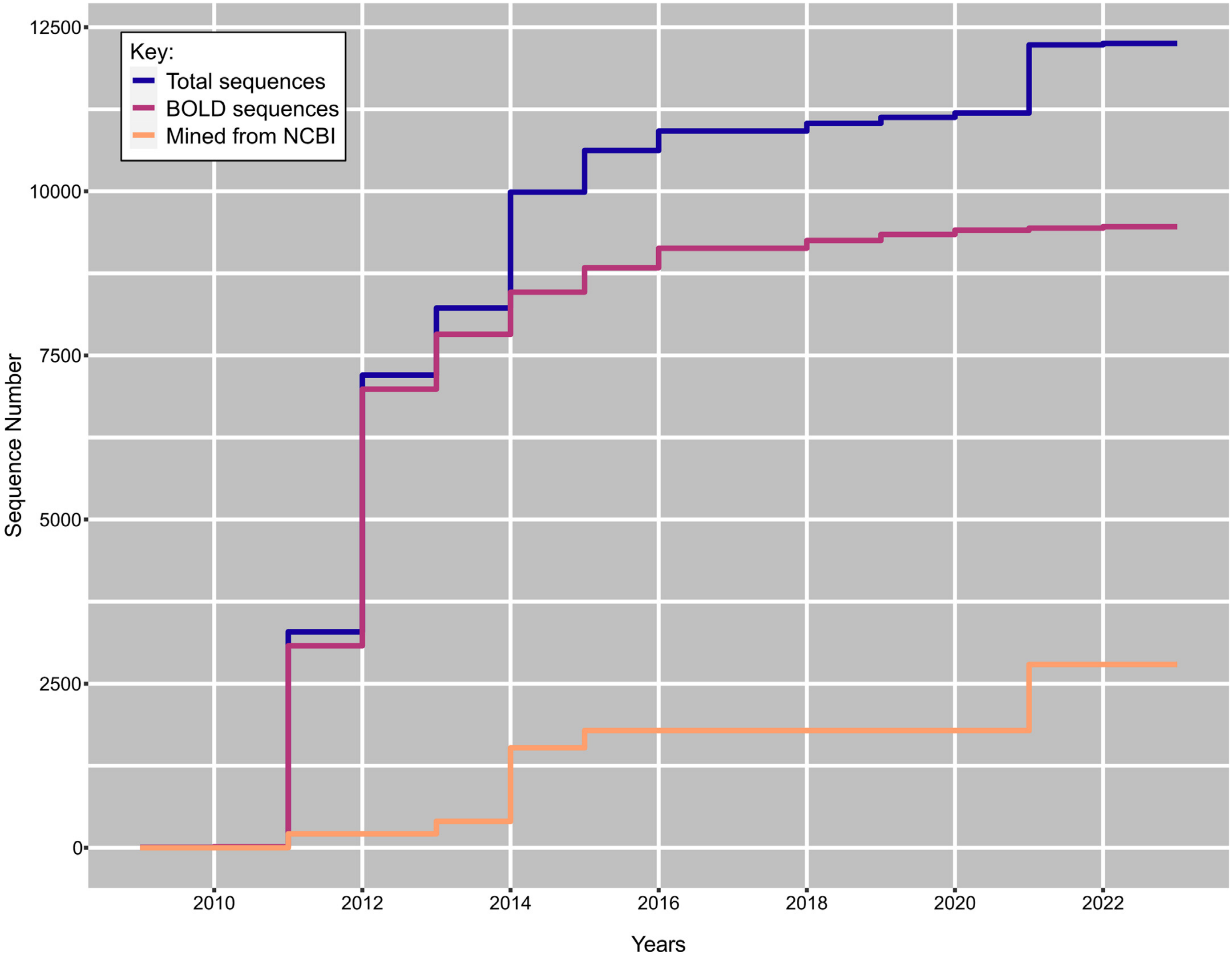

Fig. 1. Accumulation of sequence records contributed through time from individuals uploading to BOLD (purple), mining from NCBI/GenBank (orange), and the total contributions (blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

databases such as the Barcode of Life Data System (BOLD) and NCBI by several institutions across the country. These initiatives were funded through various funding efforts, including the Government of Canada through Genome Canada and the Ontario Genomics Institute, the International Development Research Centre (Canada), the Consortium for the Barcode of Life (CBOL) through Google's Global Impact Award Programme, the Royal Society (UK), and Toyota SA (Bezeng et al., 2017). Between 2009 and 2023, the Foundational Biodiversity Information Programme (FBIP), managed by the NRF and SANBI and funded by the Department of Science and Innovation (DSI), allocated R15 million towards funding over 90 small-scale projects. This resulted in 204 publications between 2013 and 2021 – with approximately 3 % (six) being DNA barcoding studies (source: https://fbip.co.za/publications/ – accessed on 21/02/2024). However, in 2023, a noticeable shift in funding priorities towards generating DNA barcode data was observed, with the FBIP supporting 70 projects. Despite this increase, there is still a lack of plant barcoding research being done, with only 7 % (five) of the funded studies focusing on plants (source: https://fbip.co.za/publications/ – accessed on 21/02/2024).

Over the last 14 years (2009–2023), 12,254 plant DNA sequences have been uploaded to BOLD Systems (Fig. 1), with the ACDB at the UJ contributing 46 % (5789) of the total records (Fig. 2). The next closest contribution was "records mined from the NCBI", with 2,797 records (22 %); these sequences are mined from GenBank to bolster the overall availability of sequence data on BOLD. Furthermore, SANBI has contributed 1,431 sequences over the same period (2009–2023), followed by the University of Cape Town (UCT) with 1,002 sequences.

6.2. *Overview of contributions to databases such as bold and NCBI's GenBank*

Although Lahaye et al. (2008) conducted one of the earliest plant barcoding studies in South Africa, the associated sequence data was not uploaded to BOLD until 2011 and 2012. This delay stemmed from BOLD's initial lack of support for plant barcoding data in 2008, with the platform only accepting plant barcodes in 2009 (Sujeevan Ratnasingham, Personal communication). Consequently, sequence records were uploaded to GenBank without accompanying metadata, such as locality data, rendering the country of origin indeterminable when accessed through BOLD. This discrepancy has artificially separated records on BOLD from South African DNA plant barcode records submitted over time. Subsequently, researchers often opt to submit their sequence data to GenBank due to its less stringent metadata requirements, as exemplified by Botha et al. (2023), who submitted their

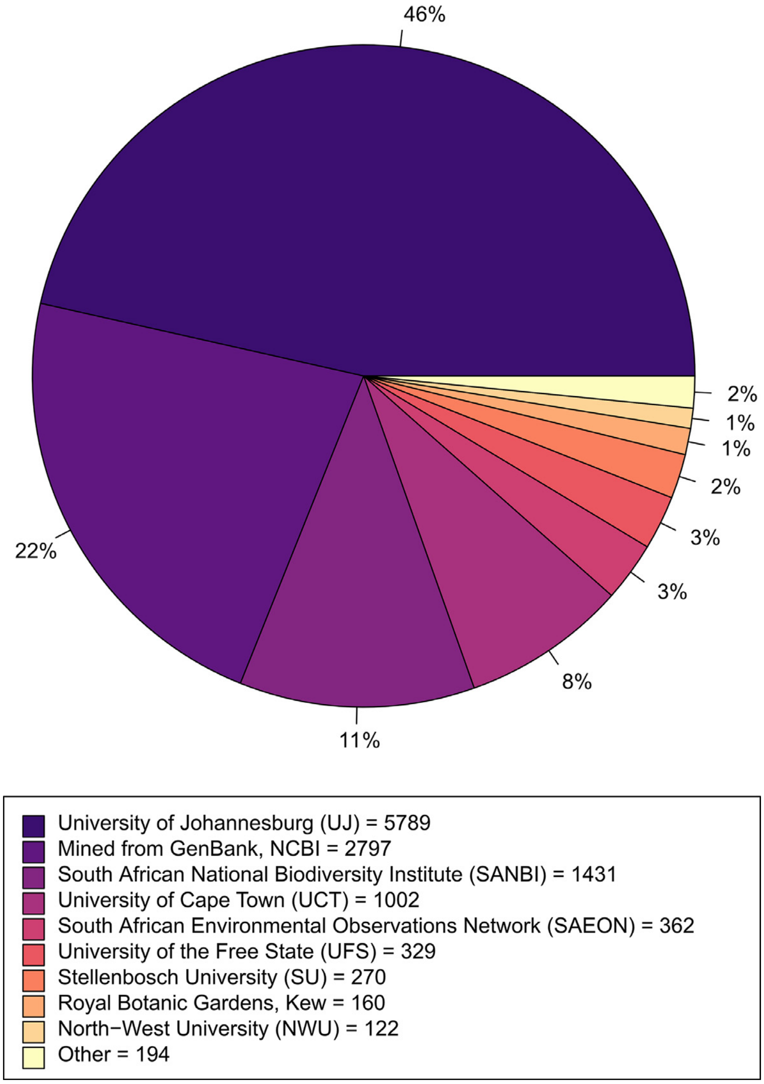

Fig. 2. The number of BOLD records submitted by the top nine contributors in South Africa.

new sequence data (2159 sequences; 1238 *rbcL* and 921 *trnL*) directly to GenBank. GenBank's appeal lies in its status as a comprehensive repository for various genetic data, extending beyond DNA barcoding – thereby potentially serving a broader range of research applications.

Even though GenBank sequence data stands as a significant resource in DNA barcoding applications, its utilisation has sparked several debates (Collins et al., 2012; Niemann et al., 2022). Despite its advantages, the more likely absence of preserved vouchers and validated identifications accompanying sequence data often undermines its direct utility for identification (Harris, 2003; Meier et al., 2006; Collins et al., 2012). Conversely, BOLD data boasts superior curation and higher quality standards but is not immune to misidentifications to some degree (Meier et al., 2006; Ratnasingham and Hebert, 2007). Recognising that no database is error-proof, users must remain cognizant of alternative explanations for observed patterns, inherent limitations in analytical techniques for taxon identification using DNA barcoding, and the necessity for ongoing dataset curation.

While BOLD serves as a sequence data repository and an informatics platform facilitating DNA barcode data storage, analysis, and publication, South African researchers have not fully capitalised on its capabilities. Instead, they tend to favour GenBank for data upload, possibly due to a lack of comprehensive understanding regarding BOLD's functionalities. Consequently, researchers may exclusively choose GenBank, which offers a simpler submission process with fewer requirements for supporting information. It is essential to note that within BOLD, an option exists for the platform to submit records to GenBank on behalf of users, allowing both GenBank and BOLD records to have the appropriate metadata available. This feature could bridge the gap between the two platforms, encourage more researchers to utilise BOLD effectively and leverage having their data available on both platforms.

6.3. Impact of technological advancements on cost reduction

A noteworthy observation is the absence of significant contributions of sequence data by individuals to BOLD after 2015, likely attributed to the country's lack of funding for barcoding-related projects. Large-scale sequencing poses challenges primarily due to its high costs (capital cost = US$500 K; annual services = US$50 K), as Hebert et al. (2023) indicated. This financial barrier is beyond the means of many South African institutions. However, a potential transformation is on the horizon as Hebert et al. (2023) have demonstrated the barcoding of 100,000 specimens in a single Oxford Nanopore run, substantially reducing the costs of obtaining a DNA barcode to US$0.10. Leveraging the Oxford Nanopore technology makes sequence generation cost-effective (sequencer = US$1000; flow cell = US$90). By lowering the initial cost of large-scale sequencing, it is possible to assist

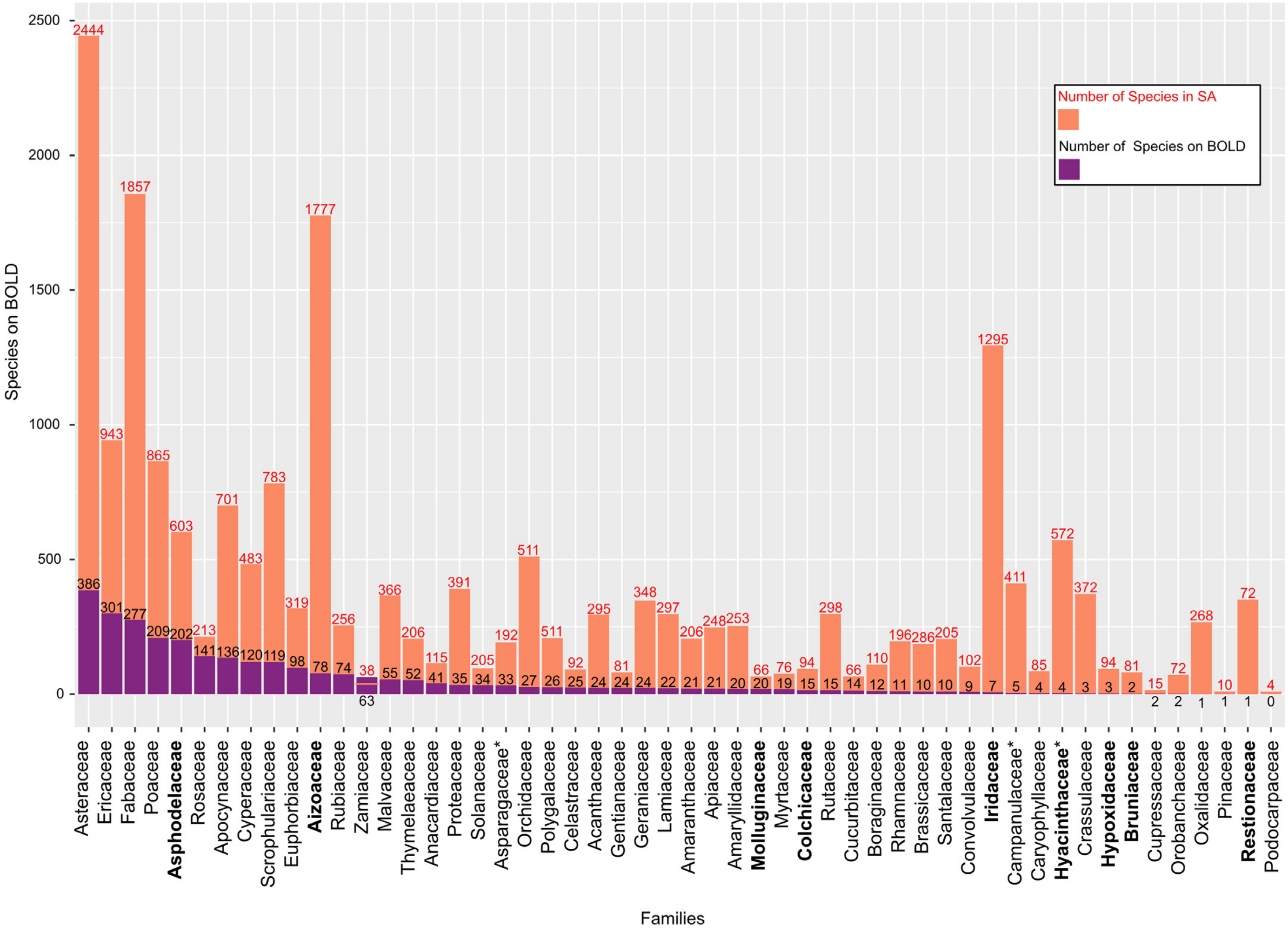

Fig. 3. Number of species records from South Africa on BOLD (purple) across families and the number of South African species found in each family. The families in bold have more than 50 % South African endemic species. The graph is sorted from the highest to the least number of species records from BOLD. The data was collected on 27/07/2023 from BOLD (http://www.boldsystems.org/). * Indicates families that have not been fully updated to APG IV.

DNA barcoding initiatives in developing countries, including South Africa. However, this is for large-scale barcoding efforts and is not necessarily suitable for small-scale projects.

6.4. Analysis of taxonomic coverage across different plant families

While global research on plant DNA barcoding is burgeoning (Cahyaningsih et al., 2022; Wang et al., 2022; Schweikle et al., 2023; Tnah et al., 2024), the trend in the number of species records on BOLD within specific South African plant families appears markedly deficient, as depicted in Fig. 3. Notably, only one family exceeds the number of barcode records compared to the number of species found in South Africa – Zamiaceae. This discrepancy can be attributed to the allocation of funding by the South African Government and Google's Global Impact Award Programme towards establishing a robust DNA barcoding reference library for *Encephalartos* to counter the illegal poaching of this endangered species. However, a significant concern arises when considering families predominantly endemic to South Africa (bold labels in Fig. 3 and Table 1), which have not managed to exceed 30 % of South African species records being uploaded to BOLD. For example, Aizoaceae, with approximately 1900 species globally (Christenhusz and Byng, 2016) and 1700 indigenous to South Africa (SA-Plant-Checklist-2023; http://opus.sanbi.org/

Table 1
South African families and the number of genera and species in the world; genera and species in SA; the number of records in BOLD; the number of SA BOLD records, genera, and species; and the percentage of the family that has been barcoded. The families in bold have more than 50 % South African endemic species. * Indicate families that have not been fully updated to APG IV (Asparagaceae – includes Ruscaceae; Campanulaceae – includes Lobeliaceae; Hyacinthaceae – should be in Asparagaceae).

Plant Family	World genera (Christenhusz and Byng, 2016)	World species (Christenhusz and Byng, 2016)	SA Genera	SA Species	World BOLD Records	SA BOLD Records	SA BOLD Genera	SA BOLD Species	Percentage Barcoded
Asteraceae	1623	24,700	259	2444	24,470	601	105	386	16 %
Ericaceae	124	4250	2	943	4899	335	1	301	32 %
Fabaceae	751	19,500	164	1857	22,935	537	68	277	15 %
Poaceae	780	12,000	185	865	19,555	391	87	209	24 %
Asphodelaceae	39	900	16	603	821	441	6	202	33 %
Rosaceae	91	2950	19	213	9354	279	5	141	66 %
Apocynaceae	366	5100	87	701	3332	175	62	136	19 %
Cyperaceae	90	5500	35	483	9564	201	18	120	25 %
Scrophulariaceae	62	1830	44	783	959	144	29	119	15 %
Euphorbiaceae	209	6252	32	319	6759	132	18	98	31 %
Aizoaceae	121	1900	120	1777	410	99	35	78	4 %
Rubiaceae	590	13,620	62	256	8244	103	40	74	29 %
Zamiaceae	9	230	2	38	720	298	2	63	166 %
Malvaceae	244	4225	35	366	4277	78	18	55	15 %
Thymelaeaceae	46	913	8	206	927	88	7	52	25 %
Anacardiaceae	83	860	15	115	1038	62	11	41	36 %
Proteaceae	83	1660	19	391	1099	41	8	35	9 %
Solanaceae	100	2600	13	96	4031	149	7	34	35 %
Asparagaceae*	114	2900	3	192	2078	35	12	33	17 %
Orchidaceae	736	28,000	53	511	20,126	30	6	27	5 %
Polygalaceae	21	900	4	208	865	34	3	26	13 %
Celastraceae	96	1350	22	92	1829	40	14	25	27 %
Geraniaceae	2	803	4	348	991	26	4	24	7 %
Acanthaceae	510	4000	38	295	1822	28	15	24	8 %
Gentianaceae	102	1735	10	81	2601	42	5	24	30 %
Lamiaceae	241	7530	41	297	7201	29	13	22	7 %
Apiaceae	442	3575	48	248	5386	32	12	21	8 %
Amaranthaceae	165	2040	39	206	2084	32	15	21	10 %
Amaryllidaceae	75	1600	22	253	1697	34	11	20	8 %
Molluginaceae	9	80	9	66	281	62	8	20	30 %
Myrtaceae	132	5950	12	76	3301	36	7	19	25 %
Rutaceae	148	2070	18	298	2301	23	10	15	5 %
Colchicaceae	17	170	9	94	342	0	4	15	16 %
Cucurbitaceae	95	965	14	66	2387	24	6	14	21 %
Boraginaceae	135	2535	22	110	3578	17	5	12	11 %
Rhamnaceae	55	950	10	196	1095	26	7	11	6 %
Santalaceae	43	1000	7	205	560	17	6	10	5 %
Brassicaceae	328	3628	28	186	6928	11	3	10	5 %
Convolvulaceae	53	1660	17	102	2786	17	4	9	9 %
Iridaceae	66	2244	38	1295	2607	10	5	7	1 %
Campanulaceae*	81	2300	18	411	2445	6	3	5	1 %
Hyacinthaceae*	46	900	18	572	101	4	2	4	1 %
Caryophyllaceae	81	2625	22	85	4405	7	3	4	5 %
Crassulaceae	35	1400	6	372	1793	3	2	3	1 %
Hypoxidaceae	4	159	4	94	111	3	2	3	3 %
Bruniaceae	6	81	6	81	127	2	2	2	2 %
Orobanchaceae	98	1960	14	72	3658	2	2	2	3 %
Cupressaceae	29	149	7	15	1325	4	1	2	13 %
Restionaceae	51	572	16	351	370	2	2	1	0 %
Oxalidaceae	5	570	1	268	622	1	1	1	0 %
Pinaceae	11	228	1	10	2823	1	1	1	10 %
Podocarpaceae	19	187	2	4	1006	0	0	0	0 %

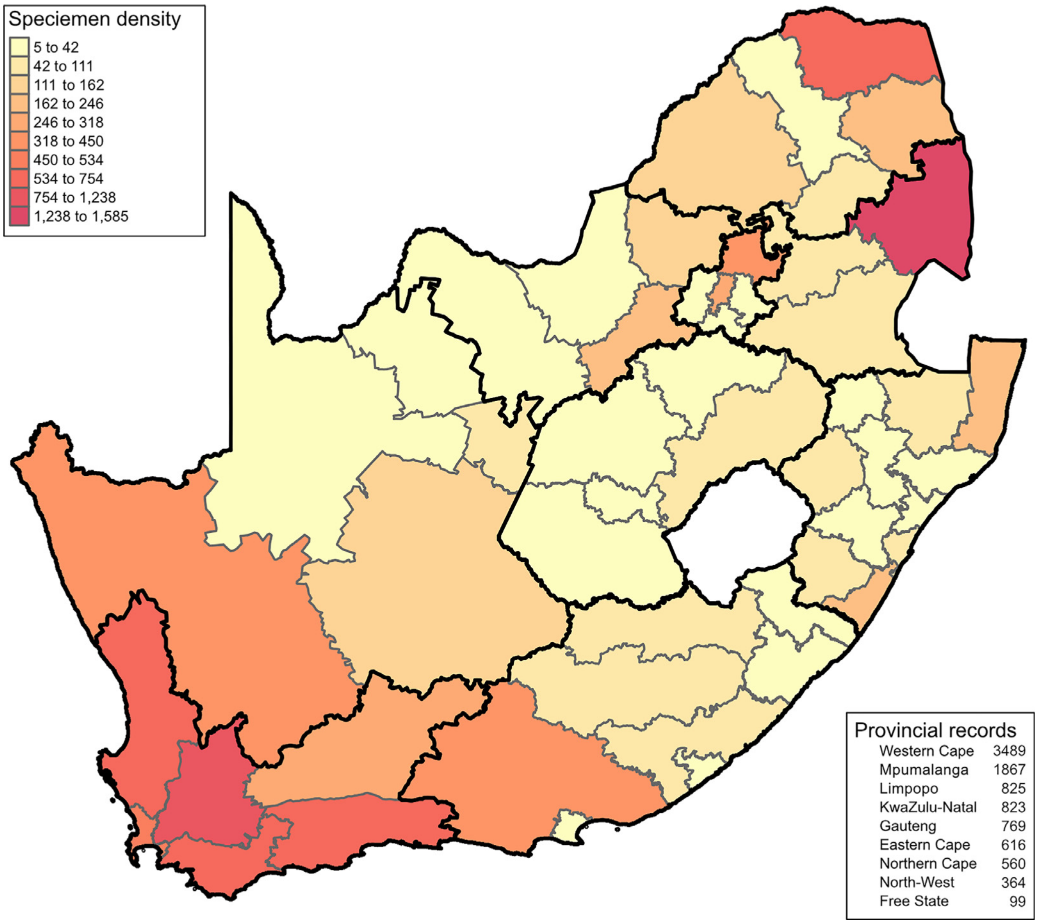

Fig. 4. The number of South African plant records from BOLD. The colour of the district municipalities is scaled by the number of BOLD records found within their boundaries. The data was collected on 27/07/2023 from BOLD (http://www.boldsystems.org/). For species distributions, see Supplementary Figure 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

handle/20.500.12143/6880), has only 4 % species coverage on BOLD. Similarly, Bruniaceae, with 81 species occurring exclusively in South Africa, has only two species – *Berzelia lanuginosa* (L.) Brongn. and *Berzelia albiflora* (E.Phillips) Class.-Bockh. & E.G.H. Oliv. – with available records on BOLD. Families with substantial endemicity such as Asphodelaceae Juss., Aizoaceae Martinov, Molluginaceae Bartl., Colchicaceae DC., Iridaceae Juss., Hyacinthaceae Batsch ex Borkh., Hypoxidaceae R.Br., and Restionaceae R.Br. should be designated as high priority for DNA barcoding initiatives by South African researchers and institutions. Notably, the Asphodelaceae is being specifically addressed through an international DNA barcoding project that involves nuclear genomics. This project includes many South African endemic species of *Aloe* L. (Woudstra et al., 2021). Such focused efforts highlight the critical need to improve species coverage and support conservation strategies, emphasising the importance of advanced genomic tools in biodiversity research and management (Woudstra et al., 2021). In their study, Hoveka et al. (2020) concentrated on angiosperm plant families alongside various South African endemic species possessing sequences available on GenBank. Their analysis unveiled that a mere 36 % of endemic species featured accessible DNA sequences on GenBank, with the less diverse families notably underrepresented in terms of sequence availability. Among the species devoid of sequence data, 68 % were classified as posing low conservation concerns, 20 % were flagged as threatened, and 12 % were categorised as data deficient. In particular, Hoveka et al. (2020) highlighted the presence of 2698 Proteaceae sequences on GenBank, while our study revealed only 318 sequences on BOLD, with contributions from *rbcLa* totalling 171 and *matK* comprising 147 sequences.

Several studies have constructed robust reference libraries with notable contributions to the plant records available on BOLD. Maurin et al. (2014) stand out in particular – contributing a substantial 1400 taxa representing 117 families and 562 genera of Angiosperms and Gymnosperms. This pivotal study not only shed light on the origins of underground forests in Africa but also provided essential data for exploring advanced ecological and evolutionary questions, such as the timing of herbivore-adapted savanna evolution (Charles-Dominique et al., 2016; Davies et al., 2020).

Notably, Mankga et al. (2013) investigated the feasibility of employing DNA barcoding to identify commonly used medicinal plants in South Africa. Their effort enriched the construction of a comprehensive reference library by contributing over 180 plant specimens to BOLD. Niemann et al. (2022) also curated an extensive macrophyte barcode reference library encompassing 539 species, with a noteworthy contribution of 120 specimens to BOLD.

Additionally, Botha et al. (2023) constructed a reference library utilising *rbcL* and *trnL* markers to support metabarcoding analysis of foraged plants. Their *rbcL* reference library consisted of 1238 sequences (representing 318 genera and 562 species), and the *trnL* reference library included 921 sequences (representing 270 genera and 562 species). Their marker choice was informed by the relative ease of amplifying shorter segments of these markers from heavily degraded material (i.e., herbivorous faecal samples) as well as the availability of these two plastid sequences on GenBank – with *rbcL* totalling 307,756 sequences and *trnL* 337,987. These collective efforts represent crucial strides in expanding the breadth and depth of the plant DNA barcoding resources of South Africa.

Preserving the rich and diverse ecosystems of South Africa entails addressing critical taxonomic, genomic, and spatial knowledge gaps. Analysing biodiversity sampling within the BOLD framework (Fig. 4) reveals that the southwestern regions, including the Cape Floristic Region and the Succulent Karoo, boast higher record numbers. Conversely, the eastern coast, particularly within the Maputaland-Pondoland-Albany hotspot, remains relatively underrepresented, with sampling concentrated primarily in areas overlapping with the Kruger National Park. Moreover, other provinces appear notably under-sampled beyond the Western Cape and Mpumalanga, with the Free State province exhibiting 99 plant records in total. Rectifying these gaps is paramount for fostering a more holistic understanding and effectively preserving South Africa's diverse plant life.

6.5. Evaluation of various barcode markers and their discriminatory power

Our analysis showed that most sequences accessible on BOLD originate from the core barcode regions (Fig. 5). However, for several South African plant families, including Ericaceae Durande, Gentianaceae Juss., Oxalidaceae R.Br., and Santalaceae R.Br., other regions such as *trnL-F* or ITS dominate the available data. This observation suggests that researchers may still evaluate which markers offer superior discriminatory power.

A notable trend emerges toward the increased utilisation of *trnL-F* (De Groot et al., 2011; Mallott et al., 2018; Omelchenko et al., 2022; Wang et al., 2022; Botha et al., 2023). This preference can partly be attributed to the gene's ability to generate small DNA segments, with the entire intron (254–767 bp) displaying significant conservation across various land plants (Taberlet et al., 1991, 2007). Furthermore, the growing emphasis on dietary research has underscored the efficiency of *trnL-F* in species-level identification – even when working with degraded samples (Ferri et al., 2015; Lee et al., 2018; Botha et al., 2023).

7. Recommendations for future research and policy interventions

The status of plant DNA barcoding in South Africa reveals persistent gaps and challenges that hinder comprehensive biodiversity conservation efforts. To address these constraints, actionable recommendations are necessary, encompassing prioritising research areas and target groups, policy advocacy, capacity building, technological advancements, and integration with conservation strategies. Prioritising research areas and taxa is crucial for advancing biodiversity research efforts, focusing on validating and optimising DNA barcoding markers for native plant species and investigating the genetic diversity of endemic taxa. Effective policy interventions should allocate funding, establish regulatory measures and frameworks, and collaborate with stakeholders to develop conservation policies. Capacity-building initiatives and collaborations require investments

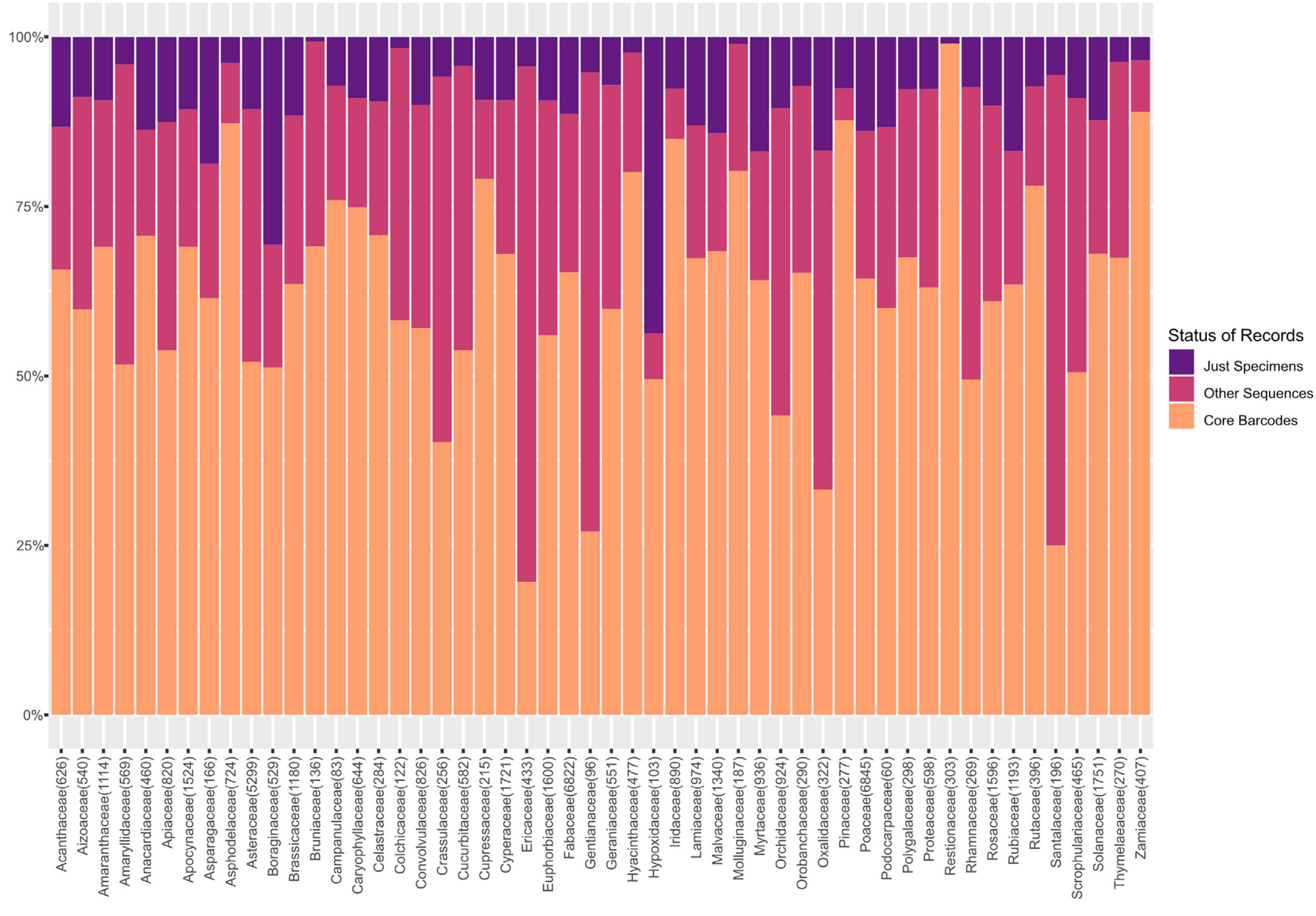

Fig. 5. The sequences composition of records on BOLD for each of the South African plant families, using orange for core barcode regions (*rbcLa* and *matK*), pink for non-core barcode sequences, and purple for entries with no associated sequence data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

in taxonomic training, laboratory infrastructure, and interdisciplinary partnerships to strengthen research capabilities. Embracing emerging technologies such as high-throughput sequencing (NGS platforms and Angisperms353) and bioinformatics is essential, necessitating resource allocation for research infrastructure and data repositories to support large-scale sequencing initiatives and facilitate data sharing. Integrating DNA barcoding data into conservation strategies and incorporating genetic information into biodiversity monitoring, conservation planning, and invasive species management is paramount. Such actions are crucial in safeguarding South Africa's plant biodiversity and preserving its natural heritage for future generations.

8. Conclusions

The comprehensive evaluation of 14 years of plant DNA barcoding in South Africa unveils both progress and persistent challenges. With a dataset comprising over 12,000 published records spanning 965 genera, 159 families and ca. 3,449 species, the nation ranks third globally in the number of Magnoliophyta records. However, critical gaps persist – particularly in endemic families and biodiversity hotspots, with only ca. 16 % of the toral flora having been barcoded. Despite a substantial contribution by several institutions nationwide (i.e., ACDB, SANBI and UCT), a noteworthy portion of data is not associated with the South African BOLD records, suggesting a potential loss of crucial information.

The underutilisation of BOLD Systems, coupled with researchers' preference for GenBank submissions, presents a significant obstacle to expanding high-quality, appropriately curated plant barcoding records within the country. Financial constraints have further impeded contributions to barcoding efforts. However, promising strides in sequencing technologies, exemplified by Oxford Nanopore sequencing technology, offer cost-effective avenues for broader participation in data generation. Notably, endemic families such as Aizoaceae and Bruniaceae remain underrepresented – highlighting the urgent need for targeted DNA barcoding initiatives. Exploring unique biodiversity hotspots across South Africa, particularly in under-sampled provinces (all provinces except the Western Cape and Mpumalanga) and prioritising those encompassing the Succulent Karoo and Maputaland-Pondoland-Albany hotspots, underscores the imperative to comprehensively address taxonomic, geographic and genomic gaps for effective preservation efforts.

This review highlights the pressing need for concerted efforts to enhance DNA barcoding utilisation, harmonise databases, increase contributions, and prioritise sampling of underrepresented taxa. These measures are crucial for comprehensively capturing South Africa's rich plant diversity and ensuring effective conservation strategies.

Declaration of competing interest

The authors declare no conflict of interest. The funders had no role in the study's design, in the collection, analyses, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

CRediT authorship contribution statement

Ryan D. Rattray: Writing – original draft, Writing – review & editing. **Ross D. Stewart:** Writing – original draft, Methodology, Formal analysis, Writing – review & editing. **Hendrik J. Niemann:** Writing – review & editing, Writing – original draft. **Oluwayemisi D. Olaniyan:** Writing – original draft, Writing – review & editing. **Michelle van der Bank:** Writing – review & editing, Resources, Investigation, Funding acquisition, Conceptualization, Writing – original draft.

Acknowledgements

The authors thank the Centre for Biodiversity Genomics, University of Guelph, Guelph Canada and the BOLD team for their continued support. Further, the authors thank the anonymous reviewers for their valuable suggestions that enhanced the manuscript.

Financial support

This work was supported by the University of Johannesburg and the African Centre for DNA Barcoding.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.sajb.2024.07.055.

References

Aylagas, E., Borja, Á., Irigoien, X., Rodríguez-ezpeleta, N., Elliott, M., 2016. Benchmarking DNA metabarcoding for biodiversity-based monitoring and assessment. Front. Mar. Sci. 3, 96. https://doi.org/10.3389/fmars.2016.00096.

Bezeng, B.S., Davies, T.J., Daru, B.H., Kabongo, R.M., Maurin, O., Yessoufou, K., Van Der Bank, H., Van Der Bank, M., 2017. Ten years of barcoding at the African Centre for DNA Barcoding. Genome 60, 629–638. https://doi.org/10.1139/gen-2016-0198.

Bohmann, K., Mirarab, S., Bafna, V., Gilbert, M.T.P., 2020. Beyond DNA barcoding: the unrealized potential of genome skim data in sample identification. Mol. Ecol. 29 (14), 2521–2534. https://doi.org/10.1111/mec.15507.

Botha, D., du Plessis, M., Siebert, F., Barnard, S., 2023. Introducing an *rbcL* and a *trnL* reference library to aid in the metabarcoding analysis of foraged plants from two semi-arid eastern South African savanna bioregions. PLoS ONE 18, 1–17. https://doi.org/10.1371/journal.pone.0286144.

Brewer, G.E., Clarkson, J.J., Maurin, O., Zuntini, A.R., Barber, V., Bellot, S., Biggs, N., Cowan, R.S., Davies, N.M.J., Dodsworth, S., Edwards, S.L., Eiserhardt, W.L., Epitawalage, N., Frisby, S., Grall, A., Kersey, P.J., Pokorny, L., Leitch, I.J., Forest, F., Baker, W.J., 2019. Factors affecting targeted sequencing of 353 nuclear genes from herbarium specimens spanning the diversity of angiosperms. Front. Plant Sci. 10, 1102. https://doi.org/10.3389/fpls.2019.01102.

Buys, M.H., Smith, G.F., 2006. Descriptive taxonomy* and DNA: two abreast, or different strokes for different blokes? S. Afr. J. Sci. 102 (5–6), 191–192.

Carlin, E., Teren, G., Ganswindt, A., 2020. Non-invasive assessment of body condition and stress-related fecal glucocorticoid metabolite concentrations in African Elephants (*Loxodonta africana*) roaming in Fynbos Vegetation. Animals 10 (5), 814. https://doi.org/10.3390/ani10050814.

Cahyaningsih, R., Compton, L.J., Rahayu, S., Brehm, J.M., Maxted, N., 2022. DNA barcoding medicinal plant species from Indonesia. Plants 11, 1–22. https://doi.org/10.3390/plants11101375.

CBOL Plant Working Group, 2009. A DNA barcode for land plants. Proc. Natl. Acad. Sci. 106 (13), 12794–12797.

Charles-Dominique, T., Davies, T.J., Hempson, G.P., Bezeng, B.S., Daru, B.H., Kabongo, R.M., Maurin, O., Muasya, A.M., Van Der Bank, M., Bond, W.J., 2016. Spiny plants, mammal browsers, and the origin of African savannas. Proc. Natl. Acad. Sci. U. S. A. 113, E5572–E5579. https://doi.org/10.1073/pnas.1607493113.

Christenhusz, M.J.M., Byng, J.W., 2016. The number of known plants species in the world and its annual increase. Phytotaxa 261, 201–217. https://doi.org/10.11646/phytotaxa.261.3.1.

Coissac, E., Hollingsworth, P.M., Lavergne, S., Taberlet, P., 2016. From barcodes to genomes: extending the concept of DNA barcoding. Mol. Ecol. 25 (7), 1423–1428. https://doi.org/10.1111/mec.13549.

Coleman, C.O., 2015. Taxonomy in times of the taxonomic impediment– examples from the community of experts on amphipod crustaceans. J. Crust. Biol. 35 (6), 729–740.

Collins, R.A., Armstrong, K.F., Meier, R., Yi, Y., Brown, S.D.J., Cruickshank, R.H., Keeling, S., Johnston, C., 2012. Barcoding and border biosecurity: identifying cyprinid fishes in the aquarium trade. PLoS ONE 7, e28381. https://doi.org/10.1371/journal.pone.0028381.

Cristescu, M.E., 2014. From barcoding single individuals to metabarcoding biological communities: towards an integrative approach to the study of global biodiversity. Trends Ecol. Evol. 29 (10), 556–571. https://doi.org/10.1016/j.tree.2014.08.001.

Davies, T.J., Daru, B.H., Bezeng, B.S., Charles-Dominique, T., Hempson, G.P., Kabongo, R.M., Maurin, O., Muasya, A.M., van der Bank, M., Bond, W.J., 2020. Savanna tree evolutionary ages inform the reconstruction of the paleoenvironment of our hominin ancestors. Sci. Rep. 10, 1–8. https://doi.org/10.1038/s41598-020-69378-0.

De Groot, G.A., During, H.J., Maas, J.W., Schneider, H., Vogel, J.C., Erkens, R.H.J., 2011. Use of *rbcL* and *trnL-F* as a two-locus DNA barcode for identification of NW-European ferns: an ecological perspective. PLoS ONE 6. https://doi.org/10.1371/journal.pone.0016371.

Department of Environmental Affairs (DEA), 2019. Environment Quarterly (No. October-December 2019). Department of Environmental Affairs, Pretoria. URL https://www.dffe.gov.za/sites/default/files/docs/publications/environmentquarterly_octoberdecember2019.pdf.

Department of the Environment, Forestry and Fisheries (DFFE), 2022. Identifying Biodiversity Knowledge Gaps For Conserving South Africa's Endemic Flora (No. February 2022). Department of the Environment, Forestry and Fisheries, Pretoria.

DeSalle, R., 2006. Species discovery versus species identification in DNA barcoding efforts: response to Rubinoff. Conserv. Biol. 20 (5), 1545–1547. https://doi.org/10.1111/j.1523-1739.2006.00543.x.

DeSalle, R., 2007. Phenetic and DNA taxonomy; a comment on Waugh. Bioessays 29 (12), 1289–1290. https://doi.org/10.1002/bies.20667.

Erckie, L., Adedoja, O., Geerts, S., Van Wyk, E., Boatwright, J.S., 2022. Impacts of an invasive alien Proteaceae on native plant species richness and vegetation structure. S. Afr. J. Bot. 144, 332–338. https://doi.org/10.1016/j.sajb.2021.09.017.

Farooq, Q., Shakir, M., Ejaz, F., Zafar, T., Durrani, K., Ullah, A., 2020. Role of DNA barcoding in plant biodiversity conservation. SIJB 03, 48–52. https://doi.org/10.36348/sijb.2020.v03i03.002.

Fazekas, A.J., Burgess, K.S., Kesanakurti, P.R., Graham, S.W., Newmaster, S.G., Husband, B.C., Percy, D.M., Hajibabaei, M., Barrett, S.C.H., 2008. Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. PLoS. One 3 (7). https://doi.org/10.1371/journal.pone.0002802.

Ferri, G., Corradini, B., Ferrari, F., Santunione, A.L., Palazzoli, F., Alu', M., 2015. Forensic botany II, DNA barcode for land plants: which markers after the international agreement? Forensic Sci. Int. Genet. 15, 131–136. https://doi.org/10.1016/j.fsigen.2014.10.005.

Forrest, L.L., Hart, M.L., Hughes, M., Wilson, H.P., Chung, K.-F., Tseng, Y.-H., Kidner, C.A., 2019. The limits of Hyb-Seq for herbarium specimens: impact of preservation techniques. Front. Ecol. Evol. 7, 439. https://doi.org/10.3389/fevo.2019.00439.

Fu, C.-N., Wu, C.-S., Ye, L.-J., Mo, Z.-Q., Liu, J., Chang, Y.-W., Li, D.-Z., Chaw, S.-M., Gao, L.-M., 2019. Prevalence of isomeric plastomes and effectiveness of plastome super-barcodes in yews (*Taxus*) worldwide. Sci. Rep. 9 (1), 2773. https://doi.org/10.1038/s41598-019-39161-x.

Fu, C.N., Mo, Z.Q., Yang, J.B., Cai, J., Ye, L.J., Zou, J.Y., Qin, H.T., Zheng, W., Hollingsworth, P.M., Li, D.Z., Gao, L.M., 2022. Testing genome skimming for species discrimination in the large and taxonomically difficult genus *Rhododendron*. Mol. Ecol. Resour. 22, 404–414. https://doi.org/10.1111/1755-0998.13479.

Gibbs, J., 2018. DNA barcoding a nightmare taxon: assessing barcode index numbers and barcode gaps for sweat bees. Genome 61 (1), 21–31. https://doi.org/10.1139/gen-2017-0096.

Goldblatt, P., 1997. Floristic diversity in the Cape Flora of South Africa. Biodivers. Conserv. 6, 359–377. https://doi.org/10.1023/A:1018360607299.

Good, R., 1947. The Geography of the Flowering Plants. Longmans. Green and Co, New York, pp. 29–31.

Guo, C., Luo, Y., Gao, L.-M., Yi, T.-S., Li, H.-T., Yang, J.-B., Li, D.-Z., 2023. Phylogenomics and the flowering plant tree of life. J. Integr. Plant Biol. 65, 299–323. https://doi.org/10.1111/jipb.13415.

Habibullah, M.S., Din, B.H., Tan, S.H., Zahid, H., 2022. Impact of climate change on biodiversity loss: global evidence. Environ. Sci. Pollut. Res. 29, 1073–1086. https://doi.org/10.1007/s11356-021-15702-8.

Hajibabaei, M., 2012. The golden age of DNA metasystematics. Trends Genet. 28 (11), 535–537. https://doi.org/10.1016/j.tig.2012.08.001.

Harris, J., 2003. Can you bank on GenBank? Trends Ecol. Evol. 18, 317–319. https://doi.org/10.1016/S0169-5347(03)00150-2.

Hebert, P.D.N., Cywinska, A., Ball, S.L., DeWaard, J.R., 2003. Biological identifications through DNA barcodes. Proceed. Biol. Sci. 270, 313–321. https://doi.org/10.1098/rspb.2002.2218.

Hebert, P.D.N., Floyd, R., Jafarpour, S., Prosser, S.W.J., 2023. Barcode 100K Specimens: in a single nanopore run. bioRxiv. 1–28. https://doi.org/10.1101/2023.11.29.569282 2023.11.29.569282.

Hollingsworth, P.M., Graham, S.W., Little, D.P., 2011. Choosing and using a plant DNA barcode. PLoS ONE 6 (5), e19254. https://doi.org/10.1371/journal.pone.0019254.

Hollingsworth, P.M., Li, D.-Z., Van Der Bank, M., Twyford, A.D., 2016. Telling plant species apart with DNA: from barcodes to genomes. Phil. Trans. R. Soc. B 371, 20150338. https://doi.org/10.1098/rstb.2015.0338.

Hoveka, L.N., van der Bank, M., Bezeng, B.S., Davies, T.J., 2020. Identifying biodiversity knowledge gaps for conserving South Africa's endemic flora. Biodivers. Conserv. 29, 2803–2819. https://doi.org/10.1007/s10531-020-01998-4.

Hoveka, L.N., Van Der Bank, M., Davies, T.J., 2022. Winners and losers in a changing climate: how will protected areas conserve red list species under climate change? Diversity Distrib. 28, 782–792. https://doi.org/10.1111/ddi.13488.

Intergovernmental Panel on Climate Change (IPCC), 2019. Special Report On Climate change, desertification, Land degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems. IPCC, Switzerland.

Kane, N., Sveinsson, S., Dempewolf, H., Yang, J.Y., Zhang, D., Engels, J.M.M., Cronk, Q., 2012. Ultra-barcoding in cacao (*Theobroma* spp.; Malvaceae) using whole chloroplast genomes and nuclear ribosomal DNA. Am. J. Bot. 99 (2), 320–329. https://doi.org/10.3732/ajb.1100570.

Kraaij, T., Baard, J.A., Arndt, J., Vhengani, L., Van Wilgen, B.W., 2018. An assessment of climate, weather, and fuel factors influencing a large, destructive wildfire in the Knysna region, South Africa. Fire Ecol. 14, 4. https://doi.org/10.1186/s42408-018-0001-0.

Krishnamurthy, P.K., Francis, R.A., 2012. A critical review on the utility of DNA barcoding in biodiversity conservation. Biodivers. Conserv. 21, 1901–1919.

Lahaye, R., Van Der Bank, M., Bogarin, D., Warner, J., Pupulin, F., Gigot, G., Maurin, O., Duthoit, S., Barraclough, T.G., Savolainen, V., 2008. DNA barcoding the floras of biodiversity hotspots. Proc. Natl. Acad. Sci. U. S. A. 105, 2923–2928. https://doi.org/10.1073/pnas.0709936105.

Larridon, I., Villaverde, T., Zuntini, A.R., Pokorny, L., Brewer, G.E., Epitawalage, N., Fairlie, I., Hahn, M., Kim, J., Maguilla, E., Maurin, O., Xanthos, M., Hipp, A.L., Forest, F., Baker, W.J., 2020. Tackling rapid radiations with targeted sequencing. Front. Plant Sci. 10, 1655. https://doi.org/10.3389/fpls.2019.01655.

Lee, T.R.C., Alemseged, Y., Mitchell, A., 2018. Dropping Hints: estimating the diets of livestock in rangelands using DNA metabarcoding of faeces. Metabarcoding. Metagenom. 2, 1–17. https://doi.org/10.3897/mbmg.2.22467.

Letsiou, S., Madesis, P., Vasdekis, E., Montemurro, C., Grigoriou, M.E., Skavdis, G., Moussis, V., Koutelidakis, A.E., Tzakos, A.G., 2024. DNA Barcoding as a Plant Identification Method. Appl. Sci. 14 (4), 1415. https://doi.org/10.3390/app140414151−12.

Li, X., Yang, Y., Henry, R.J., Rossetto, M., Wang, Y., Chen, S., 2015. Plant DNA barcoding: from gene to genome. Biol. Rev. Camb. Philos. Soc. 90 (1), 157–166. https://doi.org/10.1111/brv.12104.

Malakasi, P., Bellot, S., Dee, R., Grace, O.M., 2019. Museomics clarifies the classification of Aloidendron (Asphodelaceae), the iconic African tree aloes. Front. Plant Sci. 10, 1227. https://doi.org/10.3389/fpls.2019.01227.

Mallott, E.K., Garber, P.A., Malhi, R.S., 2018. *TrnL* outperforms *rbcL* as a DNA metabarcoding marker when compared with the observed plant component of the diet of wild white-faced capuchins (*Cebus capucinus*, primates). PLoS ONE 13, 1–16. https://doi.org/10.1371/journal.pone.0199556.

Mamathaba, M.P., Yessoufou, K., Moteetee, A., 2022. What does it take to further our knowledge of plant diversity in the megadiverse South Africa? Diversity 14, 748. https://doi.org/10.3390/d14090748.

Mani, S., Osborne, C.P., Cleaver, F., 2021. Land degradation in South Africa: Justice and climate change in tension. People Nat 3, 978–989. https://doi.org/10.1002/pan3.10260.

Mankga, L.T., Yessoufou, K., Moteetee, A.M., Daru, B.H., van der Bank, M., 2013. Efficacy of the core DNA barcodes in identifying processed and poorly conserved plant materials commonly used in South African traditional medicine. Zookeys 365, 215–233. https://doi.org/10.3897/zookeys.365.5730.

Manzanilla, V., Teixidor-Toneu, I., Martin, G.J., Hollingsworth, P.M., De Boer, H.J., Kool, A., 2022. Using target capture to address conservation challenges: population-level tracking of a globally-traded herbal medicine. Mol. Ecol. Resour. 22, 212–224. https://doi.org/10.1111/1755-0998.13472.

Margulies, J.D., Moorman, F.R., Goettsch, B., Axmacher, J.C., Hinsley, A., 2022. Prevalence and perspectives of illegal trade in cacti and succulent plants in the collector community. Conserv. Biol. 37, 1–15. https://doi.org/10.1111/cobi.14030.

Maurin, O., Davies, T.J., Burrows, J.E., Daru, B.H., Yessoufou, K., Muasya, A.M., van der Bank, M., Bond, W.J., 2014. Savanna fire and the origins of the "underground forests" of Africa. New. Phytol. 204, 201–214. https://doi.org/10.1111/nph.12936.

McDonnell, A.J., Baker, W.J., Dodsworth, S., Forest, F., Graham, S.W., Johnson, M.G., Pokorny, L., Tate, J., Wicke, S., Wickett, N.J., 2021. Exploring Angiosperms353: developing and applying a universal toolkit for flowering plant phylogenomics. Appl. Plant Sci. 9 (7), e11443. https://doi.org/10.1002/aps3.11443.

Meier, R., Shiyang, K., Vaidya, G., Ng, P.K.L., 2006. DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. Syst. Biol. 55, 715–728. https://doi.org/10.1080/10635150600969864.

Meyer, C.P., Paulay, G., 2005. DNA barcoding: error rates based on comprehensive sampling. PLoS Biol. 3, e422. https://doi.org/10.1371/journal.pbio.0030422.

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). Mol. Phylogenet. Evol. 144, 106668. https://doi.org/10.1016/j.ympev.2019.106668.

Nichols, R., 2001. Gene trees and species trees are not the same. Trends Ecol. Evol. 16 (7), 358–364. https://doi.org/10.1016/s0169-5347(01)02203-0.

Niemann, H.J., Bezeng, B.S., Orton, R.D., Kabongo, R.M., Pilusa, M., van der Bank, M., 2022. Using a DNA barcoding approach to facilitate biosecurity: identifying invasive alien macrophytes traded within the South African aquarium and pond plant industry. South African J. Bot. 144, 364–376. https://doi.org/10.1016/j.sajb.2021.08.041.

Okuyama, Y., Kato, M., 2009. Unveiling cryptic species diversity of flowering plants: successful biological species identification of Asian *Mitella* using nuclear ribosomal DNA sequences. BMC Evol. Biol. 9 (1), 1–16. https://doi.org/10.1186/1471-2148-9-105.

Omelchenko, D.O., Krinitsina, A.A., Kasianov, A.S., Speranskaya, A.S., Chesnokova, O.V., Polevova, S.V., Severova, E.E., 2022. Metabarcoding of Poaceae Pollen. Diversity 14, 1–14.

Pezzini, F.F., Ferrari, G., Forrest, L.L., Hart, M.L., Nishii, K., Kidner, C.A., 2023. Target capture and genome skimming for plant diversity studies. Appl. Plant Sci. 11 (4), e11537. https://doi.org/10.1002/aps3.11537.

Phillips, J.D., Griswold, C.K., Young, R.G., Hubert, N., Hanner, R.H., 2024. A measure of the DNA barcode gap for applied and basic research. In: DeSalle, R. (Ed.), DNA Barcoding, Methods in Molecular Biology. Humana, New York, NY. https://doi.org/10.1007/978-1-0716-3581-0_24. vol 2744.

Pirie, M.D., Oliver, E.G.H., Mugrabi De Kuppler, A., Gehrke, B., Le Maitre, N.C., Kandziora, M., Bellstedt, D.U., 2016. The biodiversity hotspot as evolutionary hotbed: spectacular radiation of *Erica* in the Cape Floristic Region. BMC Evol. Biol. 16. https://doi.org/10.1186/s12862-016-0764-3.

Rach, J., DeSalle, R., Sarkar, I.N., Schierwater, B., Hadrys, H., 2008. Character-based DNA barcoding allows discrimination of genera, species and populations in Odonata. Proceed. R. Soc. B 275 (1632), 237–247. https://doi.org/10.1098/rspb.2007.1290.

Raphela, T.D., Duffy, K., 2022. The impact of *Lantana camara* on invertebrates and plant species of the Groenkloof Nature Reserve, South Africa. Zool. Stud. 61, 1–10. https://doi.org/10.6620/ZS.2022.61-33.

Ratnasingham, S., Hebert, P.D.N., 2007. BOLD: the barcode of life data system (www.barcodinglife.org). Mol. Ecol. Notes 7 (3), 355–364. https://doi.org/10.1111/j.1471-8286.2007.01678.x.

Reynaert, S., De Boeck, H.J., Verbruggen, E., Verlinden, M., Flowers, N., Nijs, I., 2021. Risk of short-term biodiversity loss under more persistent precipitation regimes. Glob. Chang. Biol. 27, 1614–1626. https://doi.org/10.1111/gcb.15501.

Richardson, D.M., Foxcroft, L.C., Latombe, G., Le Maitre, D.C., Rouget, M., Wilson, J.R., 2020. Chapter 3: The biogeography of south african terrestrial plant invasions, in: Van Wilgen, B.W., Measey, J., Richardson, D.M., Wilson, J.R., Zengeya, T.A. (Eds.), Biological Invasions in South Africa, Invading Nature - Springer Series in Invasion Ecology. Springer Nature, Switzerland. https://doi.org/10.1007/978-3-030-32394-3

Rieseberg, L.H., Soltis, D.E., 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. Evol. Trends Plants 5 (1), 65–84.

Ruhsam, M., Rai, H.S., Mathews, S., Ross, T.G., Graham, S.W., Raubeson, L.A., Mei, W., Thomas, P.I., Gardner, M.F., Ennos, R.A., Hollingsworth, P.M., 2015. Does complete plastid genome sequencing improve species discrimination and phylogenetic resolution in *Araucaria*? Mol. Ecol. Resour. 15 (5), 1067–1078. https://doi.org/10.1111/1755-0998.12375.

Samways, M.J., Caldwell, P.M., Osborn, R., 1996. Ground-living invertebrate assemblages in native, planted and invasive vegetation in South Africa. Agric. Ecosyst. Environ. 59, 19–32. https://doi.org/10.1016/0167-8809(96)01047-X.

Schweikle, S., Häser, A., Wetters, S., Raisin, M., Greiner, M., Rigbers, K., Fischer, U., Pietsch, K., Suntz, M., Nick, P., 2023. DNA barcoding as new diagnostic tool to lethal plant poisoning in herbivorous mammals. PLoS ONE 18, e0292275. https://doi.org/10.1371/journal.pone.0292275.

Smith, V.R., Mucina, L., 2006. The vegetation of South Africa, Lesotho and Swaziland, Strelitzia, in: Smith, V.R., Mucina, L. (Eds.), Strelitzia. South African National Biodiversity Institute, Pretoria, pp. 698–723.

Slenzka, A., Mucina, L., Kadereit, G., 2013. *Salicornia* L. (Amaranthaceae) in South Africa and Namibia: rapid spread and ecological diversification of cryptic species: *Salicornia* L. in South Africa. Bot. J. Linn. Soc. 172, 175–186. https://doi.org/10.1111/boj.12041.

Smith, G.F., Figueiredo, E., Victor, J., Klopper, R.R., 2023. Plant poaching in southern Africa is aided by taxonomy: is a return to *Caput bonae spei* inevitable? Taxon. 72, 717–723. https://doi.org/10.1002/tax.12882.

South African National Biodiversity Institute (SANBI), 2023. New Plants of southern Africa [WWW Document]. Bot. Database South. Africa URL https://posa.sanbi.org/.

South African National Biodiversity Institute (SANBI), 2024. Red List of South African Plants [WWW Document]. URL http://redlist.sanbi.org/

Statistics South Africa, 2021. Natural Capital 2: Accounts for Protected Areas, 1900 to 2020. . (Discussion Document No. D0401.2). Statistics South Africa, PretoriaURL http://www.statssa.gov.za/publications/D04012/D040122020.pdf.

Stewart, R.D., 2021. Phylogeny of the Southern African genus *Gasteria* Duval (Asphodeloideae) Based on Sanger and Next Generation Sequencing Data (Masters Dissertation). University of Johannesburg, Johannesburg, South Africa.

Stewart, R.D., Clugston, J.A.R., Williamson, J., Niemann, H.J., Little, D.P., Bank, M.Van Der, 2023. Species relationships and phylogenetic diversity of the African genus *Encephalartos* Lehm. (Zamiaceae). S. Afr. J. Bot. 152, 165–173. https://doi.org/10.1016/j.sajb.2022.12.001.

Stoeckle, M.Y., Thaler, D.S., 2014. DNA barcoding works in practice but not in (neutral) theory. PLoS ONE 9, e100755. https://doi.org/10.1371/journal.pone.0100755.

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermat, T., Corthier, G., Brochmann, C., Willerslev, E., 2007. Power and limitations of the chloroplast *trnL* (UAA) intron for plant DNA barcoding. Nucl. Acids Res. 35. https://doi.org/10.1093/nar/gkl938.

Taberlet, P., Gielly, L., Pautou, G., Bouvet, J., Biologie, L.De, Fourier, U.J., 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17, 1105–1109. https://doi.org/10.1007/Bf00037152.

Thomsen, P.F., Willerslev, E., 2015. Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. Biol. Conserv. 183, 4–18. https://doi.org/10.1016/j.biocon.2014.11.019.

Tnah, L.H., Lee, S.L., Lee, C.T., Ng, K.K.S., Ng, C.H., Zawiah, N., 2024. DNA barcode identification of cultivated and wild tropical fruit species. 3. Biotech. 14, 1–10. https://doi.org/10.1007/s13205-023-03848-w.

Twyford, A.D., Ness, R.W., 2017. Strategies for complete plastid genome sequencing. Mol. Ecol. Resour. 17 (5), 858–868. https://doi.org/10.1111/1755-0998.12626.

Van Der Colff, D., Kumschick, S., Foden, W., Raimondo, D., Botella, C., Von Staden, L., Wilson, J.R.U., 2023. Drivers, predictors, and probabilities of plant extinctions in South Africa. Biodivers. Conserv. 32, 4313–4336. https://doi.org/10.1007/s10531-023-02696-7.

Van Wilgen, B.W., Measey, J., Richardson, D.M., Wilson, J.R., Zengeya, T.A., 2020. Chapter 1: Biological invasions in South Africa: an Overview. Biological Invasions in South Africa. Springer Nature.

Victor, J.E., Smith, G.F., 2011. The conservation imperative and setting plant taxonomic research priorities in South Africa. Biodivers. Conserv. 20 (7), 1501–1505. https://doi.org/10.1007/s10531-011-0041-0.

Victor, J.E., Smith, G.F., Ribeiro, S., Van Wyk, A.E., 2015. Plant taxonomic capacity in South Africa. Phytotaxa 238 (2), 149–162. https://doi.org/10.11646/phytotaxa.238.2.3.

Villano, C., Procino, S., Blaiotta, G., Carputo, D., D'agostino, N., Di Serio, E., Fanelli, V., La Notte, P., Miazzi, M.M., Montemurro, C., Taranto, F., Aversano, R., 2023. Genetic diversity and signature of divergence in the genome of grapevine clones of Southern Italy varieties. Front. Plant Sci. 14, 1201287. https://doi.org/10.3389/fpls.2023.1201287.

Wang, J., Yan, Z., Zhong, P., Shen, Z., Yang, G., Ma, L., 2022. Screening of universal DNA barcodes for identifying grass species of Gramineae. Front. Plant Sci 13, 1–8. https://doi.org/10.3389/fpls.2022.998863.

Waugh, J., Huynen, L., Millar, C., Lambert, D., 2008. DNA barcoding of animal species—response to DeSalle. Bioessays 30 (1), 92–93. https://doi.org/10.1002/bies.20698.

Wiemers, M., Fiedler, K., 2007. Does the DNA barcoding gap exist? – A case study in blue butterflies (Lepidoptera: lycaenidae). Front. Zool. 4, 8. https://doi.org/10.1186/1742-9994-4-8.

Williamson, J., Maurin, O., Shiba, S.N.S., Van Der Bank, H., Pfab, M., Pilusa, M., Kabongo, R.M., Van Der Bank, M., 2016. Exposing the illegal trade in cycad species (Cycadophyta: *Encephalartos*) at two traditional medicine markets in South Africa using DNA barcoding. Genome 59, 771–781. https://doi.org/10.1139/gen-2016-0032.

Woudstra, Y., Viruel, J., Fritzsche, M., Bleazard, T., Mate, R., Howard, C., Rønsted, N., Grace, O.M., 2021. A customised target capture sequencing tool for molecular identification of *Aloe vera* and relatives. Sci. Rep. 11, 24347. https://doi.org/10.1038/s41598-021-03300-0.

Wu, L., Wu, M., Cui, N., Xiang, L., Li, Y., Li, X., Chen, S., 2021. Plant super-barcode: a case study on genome-based identification for closely related species of *Fritillaria*. Chin. Med. 16 (1), 1–11. https://doi.org/10.1186/s13020-021-00460-z.

Yang, J.-B., Yang, S.-X., Li, H.-T., Yang, J., Li, D.-Z., 2013. Comparative chloroplast genomes of *Camellia* species. PLoS ONE 8 (8), e73053. https://doi.org/10.1371/journal.pone.0073053.

Yu, X.-Q., Jiang, Y.-Z., Folk, R. A., Zhao, J.-L., Fu, C.-N., Fang, L., Peng, H., Yang, J.-B., and Yang, S.-X., 2022. Species discrimination in *Schima* (Theaceae): Next-generation super-barcodes meet evolutionary complexity. Molecular Ecology Resources 22, 3161–3175. https://doi.org/10.1111/1755-0998.13683.

Zhang, Y., Du, L., Liu, A., Chen, J., Wu, L., Hu, W., Zhang, W., Kim, K., Lee, S.C., Yang, T.J., Wang, Y., 2016. The complete chloroplast genome sequences of five *Epimedium* species: lights into phylogenetic and taxonomic analyses. Front. Plant Sci. 7, 306. https://doi.org/10.3389/fpls.2016.00306.

Zhang, N., Erickson, D.L., Ramachandran, P., Ottesen, A.R., Timme, R.E., Funk, V.A., Luo, Y., Handy, S.M., 2017. An analysis of Echinacea chloroplast genomes: implications for future botanical identification. Sci. Rep. 7 (1), 216. https://doi.org/10.1038/s41598-017-00321-6.

Zengeya, T.A., Wilson, J.R (Eds.), 2023. The Status of Biological Invasions and Their Management in South Africa 2022. South African National Biodiversity Institute, Kirstenbosch and DSI-NRF Centre of Excellence for Invasion Biology, Stellenbosch. https://doi.org/10.5281/zenodo.8217182.