NRC Research Press

ARTICLE

Testing the reliability of standard and complementary DNA barcodes for the monocot subfamily Alooideae from South Africa

Barnabas H. Daru, Michelle van der Bank, Abubakar Bello, and Kowiyou Yessoufou

Abstract: Although a standard DNA barcode has been identified for plants, it does not always provide species-level specimen identifications for investigating important ecological questions. In this study, we assessed the species-level discriminatory power of standard (*rbcLa* + *matK*) and complementary barcodes (ITS1 and *trnH-psbA*) within the subfamily Alooideae (Asphodelaceae), a large and recent plant radiation, whose species are important in horticulture yet are threatened. Alooideae has its centre of endemism in southern Africa, with some outlier species occurring elsewhere in Africa and Madagascar. We sampled 360 specimens representing 235 species within all 11 genera of the subfamily. With three distance-based methods, all markers performed poorly for our combined data set, with the highest proportion of correct species-level specimen identifications (30%) found for ITS1. However, when performance was assessed across genera, the discriminatory power varied from 0% for all single markers and combinations in *Gasteria* to 63% in *Haworthiopsis*, again for ITS1, suggesting that DNA barcoding success may be related to the evolutionary history of the lineage considered. Although ITS1 could be a good barcode for *Haworthiopsis*, the generally poor performance of all markers suggests that Alooideae remains a challenge. As species boundaries within Alooideae remain controversial, we call for continued search for suitable markers or the use of genomics approaches to further explore species discrimination in the group.

Key words: Asphodelaceae, barcoding gap, barcode candidates, DNA barcoding, specimen identification.

Résumé : Bien qu'un code à barres de l'ADN standard ait été identifié chez les plantes, il ne procure pas toujours une identification au niveau de l'espèce pour étudier d'importantes questions en écologie. Dans ce travail, les auteurs ont examiné la capacité de discrimination des espèces du code à barres standard (*rbcLa* + *matK*) et des codes à barres complémentaires ITS1 et *trnH-psbA* au sein de la sous-famille des Alooideae (Asphodelaceae), une grande sous-famille ayant connu une radiation récente et dont les espèces sont importantes en horticulture et pourtant menacées. Les Alooideae ont leur centre d'endémisme en Afrique australe, avec quelques cas particuliers d'espèces provenant d'ailleurs en Afrique ou de Madagascar. Les auteurs ont récolté 360 spécimens représentant 235 espèces au sein des 11 genres de cette sous-famille. En faisant appel aux trois méthodes fondées sur une mesure de distance, tous les marqueurs ont offert une piètre performance au sein de cette collection; la plus forte proportion d'espèces correctement identifiées ayant été observée avec ITS1. Cependant, en examinant la performance au sein des genres, le pouvoir discriminant variait entre 0% pour tous les marqueurs individuellement ou en combinaisons chez le genre *Gasteria* jusqu'à 63% chez le genre *Haworthiopsis* pour ITS1. Cela suggère que l'efficacité du codage à barres de l'ADN serait liée à l'histoire évolutive des branches considérées. Bien que le marqueur ITS1 puisse être utile chez le genre *Haworthiopsis*, la piètre performance globale de tous les marqueurs suggère que la sous-famille des Alooideae demeure problématique. Comme les frontières entre les espèces au sein des Alooideae demeurent sujet à débats, les auteurs suggèrent de continuer la recherche en

Received 29 November 2015. Accepted 29 October 2016.

Corresponding Editor: Sarah Adamowicz.

B.H. Daru. Department of Organismic and Evolutionary Biology and Harvard University Herbaria, Harvard University, Cambridge, MA 02138, USA; Department of Plant Science, University of Pretoria, Private Bag X20, 0028 Hatfield, Pretoria, South Africa.
M. van der Bank. African Centre for DNA Barcoding, University of Johannesburg, APK Campus, PO Box 524, Auckland Park 2006, Johannesburg, South Africa.
A. Bello. Bolus Herbarium, Biological Sciences Department, University of Cape Town, Private Bag X3, Rondebosch 7700, South Africa; Department of Biology, Umaru Musa Yarádua University, Katsina, Nigeria.
K. Yessoufou. Department of Geography, Environmental Management and Energy Studies, University of Johannesburg, APK Campus, PO Box 524, Auckland Park 2006, Johannesburg, South Africa.

Corresponding author: Barnabas H. Daru (email: barnabas_daru@fas.harvard.edu).

vue de trouver des marqueurs adéquats ou le recours à des approches génomiques pour distinguer les espèces au sein de ce groupe. [Traduit par la Rédaction]

Mots-clés : Asphodelaceae, lacune dans le codage à barres, séquences candidates de codes à barres, codage à barres de l'ADN, identification de spécimens.

Introduction

The alooids subfamily Alooideae (Asphodelaceae) is a group of rosulate succulents comprising 11 genera (Table 1). Early taxonomic studies of Alooideae were based on morphological characters (e.g., floral traits, size, shape, arrangement, and combination of leaves and markings; Smith and van Wyk 1991). Taxonomic classification and the study of species boundaries have a long and illustrious history, including taxon-based works by Linnaeus (1753), Duval (1809), Salm-Dyck (1836–1863), and Uitewaal (1947); karyotype-based studies by Taylor (1925), among others; and a number of more recent studies, such as those by Smith and van Wyk (1991) and Klopper and Smith (2011). These studies have led to taxonomic changes on several occasions; even recent studies that combined morphology and DNA-based phylogeny to reassess taxa delimitation within the subfamily (Daru 2012; Daru et al. 2013; Manning et al. 2014*a*) found pitfalls that led to taxonomic change (Manning et al. 2014*a*). However, there is an increasing interest in the use of phylogenetic data to disentangle the evolutionary relationships within the subfamily in addition to, or in support of, the observed morphology-based patterns (Treutlein et al. 2003*a*, 2003*b*; Ramdhani et al. 2011). Although these studies provide useful insights into the taxonomy of the subfamily, they are often based on sparse taxonomic sampling, and the reconstructed phylogeny is still unresolved. This lack of resolution is problematic if we are to discriminate between the over 500 species described in the subfamily, but we note that a fully resolved phylogeny is not necessarily needed for accurate specimen identification to species level.

In an attempt to provide a better understanding of evolutionary relationships within the group, a more recent study (Daru et al. 2013) combined molecular and morphological data to identify some important pitfalls in the current classification (e.g., homoplasious characters, morphological traits not consistent enough to distinguish species within the genera, etc.), prompting a new treatment of the subfamily (e.g., re-circumscribing the Alooideae genera into monophyletic entities; see Grace et al. 2013; Manning et al. 2014*a*). Given these pitfalls and this new treatment, identifying species within Alooideae becomes even more problematic.

The subfamily Alooideae is widely distributed in Africa, with its main centre of diversity found in southern Africa and outliers found in the Arabian Peninsula, Madagascar, and other islands in the western Indian Ocean (Reynolds 1966; Viljoen 1999; Glen and Hardy 2000; Klopper and Smith 2007). However, the horticultural appeal of members of the subfamily has motivated illegal collections in the wild, which has been a major threat to the plants (Smith et al. 2000; Raimondo et al. 2009). There is therefore a need for conservation actions, which require an accurate assessment of species diversity in the group, taking into account genetic-based species delineation in addition to morphological data (Eaton et al. 2010; Lowe and Cross 2011).

There is an impressive body of literature devoted to morphology-based species delimitation within the subfamily Alooideae (Reynolds 1966; Smith and van Wyk 1991; Viljoen 1999; Glen and Hardy 2000; Klopper and Smith 2007) and comparatively poorer attention to genetic diversity. While DNA barcoding was originally developed as a system for specimen identification based solely on DNA sequences (Hebert et al. 2003), it is increasingly recognized as a key tool to complement morphology-based specimen identification (Edwards et al. 2008; Sun et al. 2012; Gere et al. 2013). The performance of DNA barcoding has, however, been mixed for various plant taxa: while some limitations have been documented in some groups, for example *Viburnum* (Adoxaceae; Clement and Donoghue 2012), *Agalinis* (Orobanchaceae; Pettengill and Neel 2010), *Tetrastigma* (Vitaceae; Fu et al. 2011), Lemnaceae (Wang et al. 2010), *Berberis* (Berberidaceae; Roy et al. 2010), and *Parnassia* (Parnassiaceae; Yang et al. 2012), strong and reliable performance of DNA barcodes has been reported in many other studies of specimen identification (Burgess et al. 2011; Gere et al. 2013; Mankga et al. 2013). These mixed findings discount the generalization power of DNA barcoding across all taxonomic groups but reinforce the need for case-by-case study (e.g., Clement and Donoghue 2012; Gere et al. 2013; Daru and Yessoufou 2016).

The use of a phylogenetic approach in ecology is now a common practice; this requires a fully resolved phylogeny (Davies et al. 2012) that barcode-based approaches do not always provide. Key ecological questions related to extinction risk, the origin of diversification of a taxonomic group, the role of historical climate in triggering and controlling the temporal dynamics of speciation, and phylogenetically informed conservation decisions can be better understood only with a species-level resolved phylogeny. Phylogenetic ethnobotany is also gaining momentum (e.g., Saslis-Lagoudakis et al. 2012; Yessoufou et al. 2015) and requires fully resolved phylogenies to test whether closely related species share similar bioactive compounds or bioactivity against a specific ailment. Because the phylogeny recovered for the subfamily Alooideae using the standard barcode does not provide well-resolved phylogenetic relationships among species

Table 1. Summary of global richness of species within Alooideae genera versus total number of species sampled in this study (indicated in parentheses).

Genus	No. of currently known species in the genus	No. of species (no. of samples)	Percent sampling completeness (%)	References
Aloe L.	~400	150 (214)	38	Reynolds (1966); Viljoen (1999); Glen and Hardy (2000); Klopper and Smith (2007)
Aloiampelos Klopper & Gideon F. Sm.	7	5 (7)	71	Grace et al. (2013)
Aloidendron (A. Berger) Klopper & Gideon F. Sm.	7	5 (12)	71	Grace et al. (2013)
Aristaloe Boatwr. & J.C. Manning	1	1 (3)	100	Manning et al. (2014*a*)
Astroloba Uitewaal	6	6 (9)	100	Roberts Reinecke (1965); Groen (1987)
Gasteria Duval	23	18 (20)	78	Duval (1809); Van Jaarsveld (2007)
Gonialoe (Baker) Boatwr. & J.C. Manning	3	1 (2)	33	Manning et al. (2014*a*)
Haworthia Duval	42	32 (52)	76	Bayer (1999)
Haworthiopsis G.D. Rowley	18	12 (30)	67	Rowley (2013)
Kumara Medik.	2	1 (3)	50	Glen and Hardy (2000)
Tulista Raf.	4	4 (8)	100	Rowley (2013)

(see Daru 2012), there is a need for a continued commitment to searching for DNA markers that can provide such resolved phylogenies to allow future detailed studies of the phylogenetic ecology of Alooideae. In addition to exploring species-level identification, our study also partially addresses this important issue of phylogenetics by examining species-level resolution, i.e., the tips of the phylogeny.

Two DNA loci, *matK* and *rbcLa*, have been proposed as the core barcodes for land plants (CBOL Plant Working Group 2009) that can be supplemented by *trnH-psbA* and internal transcribed spacer (ITS) (Hollingsworth et al. 2011; Liu et al. 2011; Gere et al. 2013). The core barcodes have been shown to yield high levels of specimen identification to species and sequence recoverability (Burgess et al. 2011; Mankga et al. 2013). However, the taxonomic sampling in some studies is sparse. If few species are included per genus, the performance of DNA barcoding in specimen identification can be inflated. We consider only ITS1 here because of its higher performance than ITS2 in disentangling phylogenetic relationships in Alooideae (Treutlein et al. 2003*a*, 2003*b*; Ramdhani et al. 2011) or in eukaryotes in general (Wang et al. 2015). Additionally, a preliminary PCR amplification of Alooideae using available ITS2 primers proved unsuccessful (Daru et al. 2013).

Most available molecular studies of Alooideae examined chloroplast markers (usually not more than six, including *rbcLa*, *matK*, *trnH-psbA*, *trnL-F*, and *rps16*) and sometimes nuclear regions (ITS). Since Chase et al. (2000) provided one of the first molecular phylogenetic evaluations of the subfamily Alooideae based on *rbcL* and demonstrated that Alooideae is monophyletic, other molecular studies have focused on different lineages within Alooideae using different markers. For instance, Treutlein et al. (2003*b*) used chloroplast sequencing and genomic fingerprinting of Alooideae to demonstrate that genera and species of Alooideae are polyphyletic. A noteworthy contribution was made by Ramdhani et al. (2011), who also confirmed the polyphyly of *Haworthia* using *trnL-trnF*, *trnH-psbA*, and ITS1. Recent phylogenetic studies of Alooideae used more comprehensive taxon sampling to reveal instead the paraphyly of *Aloe* and *Haworthia*, which led to taxonomic revisions of the subfamily (Daru et al. 2013; Grace et al. 2013; Manning et al. 2014*a*). Although these later studies form the baseline upon which our study rests, they do not explicitly assess the species-level discriminatory power of either the standard DNA barcode or the additional markers they used.

In this study, we used the most comprehensive molecular data yet available for the subfamily Alooideae, with about 50% sampling completeness of species within the subfamily (Table 1), to test the DNA barcode potential of four DNA markers (*trnH-psbA*, *matK*, *rbcLa*, and ITS1) abundantly used in phylogenetic studies of Alooideae (e.g., Daru et al. 2013).

Material and methods

Data and taxonomic sampling

We used all available DNA sequences for Alooideae for four molecular markers: *trnH-psbA*, *matK*, *rbcLa*, and ITS1, sequences that our group previously generated comprehensively for this subfamily (see Daru 2012; Daru et al. 2013; Manning et al. 2014*a*). Additional sequences for ITS1 for 85 taxa were obtained from Grace et al. (2015) (see the supplementary data, Table S1[1]). These previous studies

[1]Supplementary data are available with the article through the journal Web site at http://nrcresearchpress.com/doi/suppl/10.1139/gen-2015-0183.

(Daru 2012; Daru et al. 2013; Manning et al. 2014*a*; Grace et al. 2015) followed commonly used taxonomic concepts in Alooideae (Roberts Reinecke 1965; Reynolds 1966; Groen 1987; Bayer 1999; Glen and Hardy 2000; Van Jaarsveld 2007). All other sequences were derived from our group (Daru 2012; Daru et al. 2013; Manning et al. 2014*a*). DNA sequences were aligned using default settings in SEAVIEW v.4 (Gouy et al. 2010), setting the alignment options to "clustalo" for the combined data set and also separately for each genus and gene region. For data analysis, gaps were considered as missing data. The alignments were manually checked and adjusted in MESQUITE v.2.5 (Maddison and Maddison 2008) in cases of misalignment, and for ITS1 in particular, alignments were done for each genus separately. The final sequences used for the analysis are a combination of data derived from our group (Daru 2012; Daru et al. 2013; Manning et al. 2014*a*) and Grace et al. (2015) and include 235 species (n = 360 samples) belonging to all 11 currently known Alooideae genera, with more than 50% sampling completeness for the subfamily (Table 1). The sampling covers the geographical range of the subfamily, mainly in southern Africa but also Madagascar (e.g., *Aloe haworthioides*) and Somalia (e.g., *Aloidendron eminens*).

All GenBank and European Bioinformatics Institute accession numbers are provided in the supplementary information (Tables S1, S2)[1]. Additionally, complete data including GPS coordinates, pictures, and DNA barcodes are available on the Barcode of Life Data Systems (BOLD; http://www.boldsystems.org; Ratnasingham and Hebert 2007) within the publicly available project "Alooideae of Africa".

DNA barcoding analysis

We evaluated four single DNA markers: three chloroplast regions (*rbcLa*, *matK*, and *trnH-psbA*) and one nuclear marker (ITS1). We also tested the four genes in different combinations: (1) *rbcLa* + *matK* (i.e., the core barcodes; CBOL Plant Working Group 2009); (2) *rbcLa* + *matK* + *trnH-psbA*; (3) *rbcLa* + *matK* + ITS1; and (4) *rbcLa* + *matK* + *trnH-psbA* + ITS1. First, we subdivided the combined aligned matrix into subsets of matrices of each gene as input files for further analysis. Second, we used two criteria commonly used in DNA barcoding analyses, i.e., the barcode gap of Meyer and Paulay (2005) and discriminatory power, to assess the performance of individual and combined markers. The presence of a barcode gap for each species was defined as discontinuity between minimum interspecific pairwise Kimura's 2-parameter distance calculated by setting the analysis parameters to remove missing data as implemented in the R package ape (Paradis et al. 2004) and maximum intraspecific divergence obtained by plotting a line plot for the four gene regions and combinations. We also calculated the distribution of range, mean, and standard deviation of both intra- and interspecific distances. The nearest neighbour distance method was used for the calculation of interspecific distances.

All DNA sequences were labelled with the names of the species from which the sequences were generated. Then, each query was considered as an unknown, but all other sequences in the data set (i.e., the 360 specimens in this study) were considered as the reference DNA barcode database. If the ID of the query corresponded to the sequence label in the reference, the identification test was scored as "correct", and the overall proportion of correct identifications corresponded to the discriminatory power of the DNA marker tested. Three approaches were used for the test: "best close match" (Meier et al. 2006), "near neighbour", and BOLD criteria using, respectively, the functions *bestCloseMatch*, *threshID*, and *nearNeighbour* implemented in the program Spider 1.1-1 (Brown et al. 2012). Prior to the tests, we determined, for each data set (marker including combinations and all genera), the optimized genetic distance suitable as a threshold for specimen identification. For this purpose, we used the function *localMinima*, also implemented in Spider (Brown et al. 2012).

The function *bestCloseMatch* conducts the "best close match" analysis of Meier et al. (2006), searching for the closest individual in the reference data set. If the closest specimen is within the threshold, the identification is "correct". If it is greater than the threshold, the outcome is scored as "no ID" (no identification). However, when more than one species are tied for closest match, the identification result is scored as "ambiguous". When all matches within the threshold are species different to the query, the result is scored as "incorrect".

The function *threshID* conducts a threshold-based analysis using a threshold distance of 1%. It is more inclusive than *bestCloseMatch* in that it considers all sequences within the threshold of 1%. Four outcomes are also possible: "correct", "incorrect", "ambiguous", and "no ID".

The *nearNeighbour* function finds the closest specimens and returns the score "true", i.e., correct ID, if their names are the same; however, when the names are different, the outcome is scored as "false", i.e., incorrect ID.

Two additional analyses were conducted. We assessed the PCR success rate and sequence quality. The success rate for each marker was evaluated qualitatively based on the proportion of PCR products with strong PCR bands as scored by one of the authors (BHD), scaled arbitrarily as follows: <50% = poor PCR success; 50%–70% = moderate PCR success; and 71%–100% = high PCR success. Because PCR bands are not good indicators of successful sequencing, we then evaluated the quality of the final sequences of all extracted specimens quantitatively as the percent quality of all sequence trace files for each marker that our group generated previously (Daru 2012; Daru et al. 2013; Manning et al. 2014*a*) using Sequencher v.3.1 (Gene Codes, Ann Arbor, Mich., USA). The sequence trace file quality is a confidence score generated as an

Table 2. Summary statistics indicating the ranges and means of intra- and interspecific distances for the gene regions and combinations tested for species of Alooideae, based on Kimura's 2-parameter model of DNA evolution.

DNA barcoding region	Average sequence length (range) (bp)	Alignment length (bp)	Range inter	Mean inter (±SD)	Range intra	Mean intra (±SD)	Threshold (%)
matK	762 (251–786)	786	0–0.110	0.0170±0.0086	0–0.034	0.00290±0.0046	0.058
rbcLa	545 (272–552)	552	0–0.062	0.0055±0.0044	0–0.015	0.00091±0.0020	0.075
trnH-psbA	498 (356–504)	514	0–0.083	0.0270±0.0160	0–0.050	0.00710±0.0110	0.610
ITS1	314 (227–362)	393	0–0.310	0.0650±0.0350	0–0.055	0.00850±0.0100	9.860
rbcLa+matK	1306 (535–1338)	1341	0–1.380	0.0230±0.0480	0–0.310	0.00890±0.0370	0.360
rbcLa+matK+trnH-psbA	1817 (1037–1837)	1857	0–0.240	0.0220±0.0230	0–0.020	0.00620±0.0050	1.710
rbcLa+matK+ITS1	1623 (738–1697)	1711	0–0.930	0.0480±0.0450	0–0.200	0.01000±0.0220	2.490
rbcLa+matK+trnH-psbA+ITS1	2132 (1670–2195)	2210	0–0.150	0.0340±0.0180	0–0.035	0.00890±0.0080	1.520

Note: The genetic distances are means of pairwise divergence distances. The interspecific distances are averages of the nearest neighbour distances. SD, standard deviation; inter, interspecific; intra, intraspecific; threshold, the distance cutoff for specimen identifications by conducting a threshold-based analysis, similar to the "Identify Specimen" tool on the Barcode of Life Data Systems (www.boldsystems.org).

integral part of the chromatogram file that is obtainable directly for each specimen upon sequencing as a Phred file and can be viewed in Sequencher v.3.1 or similar programs such as Applied Biosystems's KB base caller. The program generates a quality score for each sequence, defined as the percentage of bases meeting or surpassing a Phred score of 20. We used the quality scores as our measure of sequence quality. For instance, a quality score of 60% indicates that 40% of the bases are low quality and vice versa (Gene Codes Corporation 2016). Percent sequence quality was calculated for each sequence trace file for each sample and for each marker.

Lastly, given the possibility that the performance of markers could vary among taxa (Gere et al. 2013), we further assessed the performance of the best barcode within five genera having the largest sample sizes: *Aloe*, *Astroloba*, *Gasteria*, *Haworthia*, and *Haworthiopsis*. The other Alooideae genera were not evaluated here owing to lack of sufficient DNA sequences.

Altogether, we identified the best barcode for the subfamily as the region or the combined regions that simultaneously exhibit a barcode gap and the highest score of correct identification at the species level. The results are summarized for each genus separately.

Results

Genetic variation within each DNA marker

We assessed and compared genetic variation among single loci using multiple approaches. We found that ITS1 had the highest mean interspecific variation between nearest neighbouring species (0.065 ± 0.035, n = 248), with the remaining markers possessing variability in the following order: ITS1 > *trnH-psbA* > *matK* > *rbcLa* (Table 2). The same holds for mean intraspecific distance, for which we found the same order, i.e., ITS1 > *trnH-psbA* > *matK* > *rbcLa*. For combinations of DNA markers, *rbcLa* + *matK* + ITS1 yielded the highest mean interspecific genetic distance for Alooideae identification (0.048 ± 0.045, n = 248).

All DNA regions or combinations showed a low barcoding gap, i.e., discontinuity between intra- and interspecific genetic divergences (Fig. 1), with the percentages of species with gaps ranging from 5% for *rbcLa* to 40% for ITS1 (Table 3).

We calculated the optimized genetic distance (threshold distance) with which to evaluate the discriminatory power of different gene regions and combinations. Apart from ITS1, for which we found a threshold of 9.86%, all other single regions had an optimized threshold of <1% (Table 2). The thresholds increased slightly above 1% for all combinations except *rbcLa* + *matK* + ITS1, which had a threshold of 2.49%. Using these cutoffs, we evaluated the discriminatory power of the different gene regions. For single regions based on the best close match method, again ITS1 provided the highest rate (20%, n = 248) of discrimination followed by *matK* and *trnH-psbA* (both 11%, n = 360 and 202, respectively), with *rbcLa* assigning only 5% (n = 360) of the individuals to the correct species (Table 3). The same order of performance was observed for the near neighbour method but with greater identification success for ITS1 (30%, n = 248), *matK* (28%, n = 360), *rbcLa* (20%, n = 360), and *trnH-psbA* (19%, n = 202).

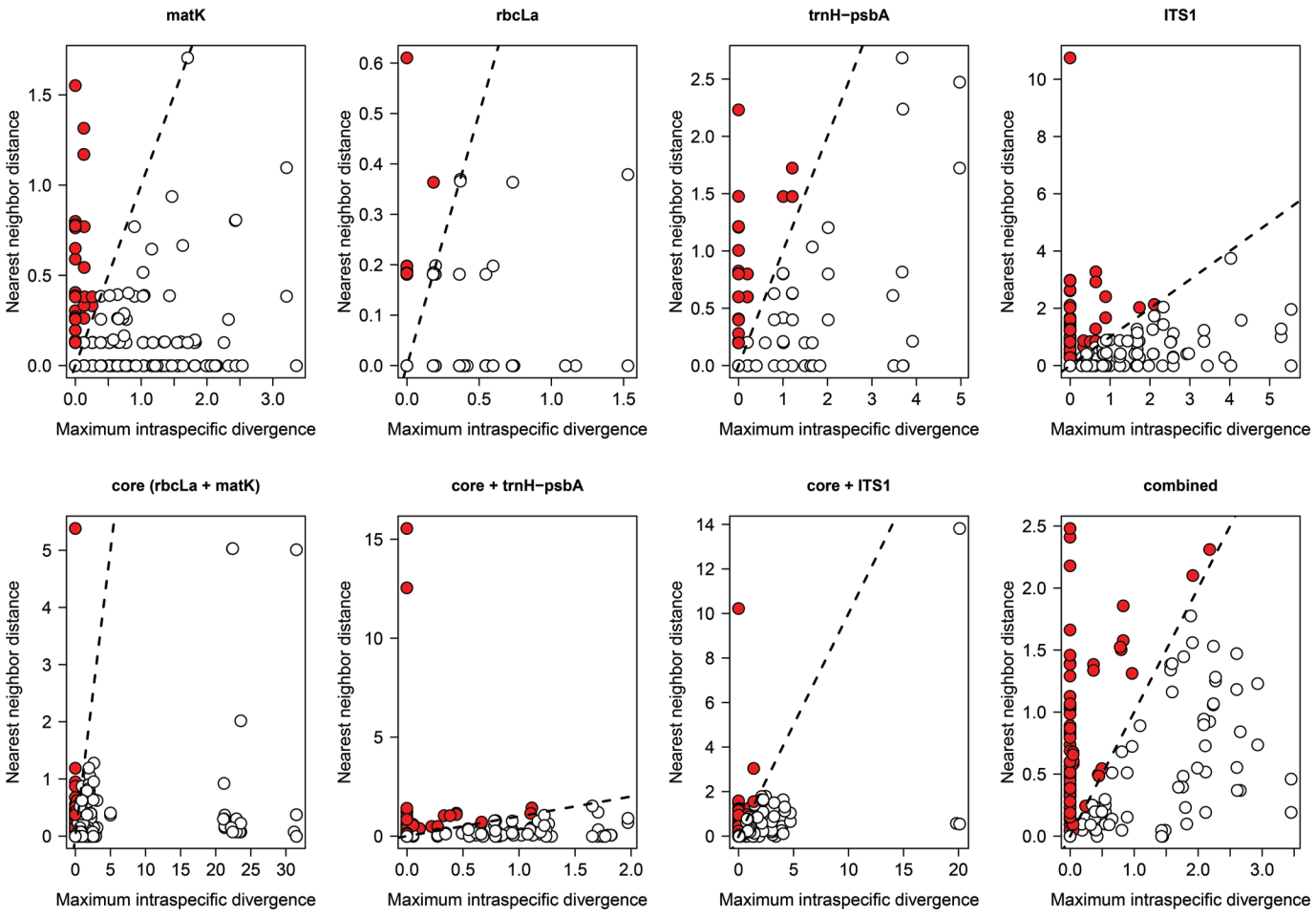
For the combined regions under both best close match and near neighbour methods, inclusion of either ITS1 or *trnH-psbA* with the core barcodes (*rbcLa* + *matK*) did not improve identification success rate (best close match: ITS1 + *matK* + *rbcLa* = 20% and *trnH-psbA* + *matK* + *rbcLa* = 17%; near neighbour: ITS1 + *matK* + *rbcLa* = 25% and *trnH-psbA* + *matK* + *rbcLa* = 22%; Table 3).

Within single genera, we found that the combination of ITS1 with the core barcodes (*matK* + *rbcLa*), i.e., ITS1 + *matK* + *rbcLa*, improved specimen identification in *Aloe* from 7% (ITS1 alone) to 14% (for ITS1 + *matK* + *rbcLa*) and in *Haworthia* from 20% to 24% (ITS1 alone vs. *matK* + *rbcLa* + ITS1, respectively; Table 4). For *Astroloba*, there was no improvement in species discrimination (25% for both ITS1 alone and *matK* + *rbcLa* + ITS1), whereas we found a reduction in species discrimination in *Haworthiopsis* from 63% to 50% when ITS1 was combined with the core barcodes (ITS1 alone vs. *matK* + *rbcLa* + ITS1, respectively).

Amplification success and quality of sequence trace files

The amplification success varied from poor to high (Fig. 2). A poor PCR success rate was found with ITS1

Fig. 1. Line plots of DNA barcode gap for four gene regions and combinations for Alooideae specimen identification. For each marker and combination, closed circles above the 1:1 line indicate the presence of a barcode gap, whereas open circles below the line indicate no barcode gap. Species included in the barcode gap analysis are represented by at least two sequences. [Colour online.]



(30.6%), whereas the success rate was moderate with *trnH-psbA* (52.8%) and *matK* (64.8%). The highest success rate was observed with *rbcLa* (83.1%). The primers recommended by the CBOL Plant Working Group (2009) for *rbcLa* and *matK* (rbcL-barcode-F: rbcL-barcode-R and 3F_KIM: 1R_KIM, respectively) and the other two tested in this study (*trnH-psbA* and ITS1) were successful, such that no internal priming was required for any of the DNA regions. The quality of sequence trace files followed a similar trend (*rbcLa* > *matK* > *trnH-psbA* > ITS1); for all species, 90.25% ± 13.21% of specimens yielded a readable trace file for *rbcLa* versus 84.57% ± 12.94% for *matK* and 75.65% ± 18.79% for *trnH-psbA*, with often little or no editing. ITS1 yielded the lowest percent sequence quality (59.71% ± 20.10%) and often with considerable editing of the chromatograms.

Discussion

Several criteria have been defined for the identification of the best DNA barcode candidate (Hebert et al. 2004; Kress and Erickson 2007; Lahaye et al. 2008; CBOL Plant Working Group 2009). First, the candidate must provide maximal discrimination among species. We measured the discriminatory power of four candidates using "barcode gap" (Meyer and Paulay 2005) and distance-based methods (Kress et al. 2005; Lahaye et al. 2008; Hollingsworth et al. 2009). Although the core barcodes (*matK* and *rbcLa*) may not exhibit a barcode gap for several genera (e.g., *Parnassia*; Yang et al. 2012), we also found that all our markers (including the core barcodes) exhibit low prevalence of a gap for Alooideae. Misidentifications and phylogenetic reticulation are commonplace in rapidly evolving lineages such as the Alooideae (Viljoen 1999), for example, *Haworthia* (Bayer 1999), *Astroloba* (Treutlein et al. 2003*a*), and Aloineae (Riley and Majumdar 1979), which may have led to the low discrimination rates of the DNA barcodes in this study. Such cases of reticulation have led to the adoption of other PCR-based methods, such as inter simple sequence repeat fingerprinting, for detecting hybrids (Wink et al. 2001; Treutlein et al. 2003*b*). Nonetheless, ITS1 shows relatively high interspecific variation, irrespective of the metric used. These findings indicate that ITS1, regardless

Table 3. Efficacy of candidate DNA barcodes in identification of Alooideae based on discriminatory potential using distance methods: near neighbour, BOLD ID, and best close match.

			Near neighbour		BOLD ID				Best close match			
DNA barcoding region	No. of species (no. of samples)	Proportion of species with barcode gap (%)	TRUE (%)	FALSE (%)	Ambiguous (%)	Correct (%)	Incorrect (%)	No ID (%)	Ambiguous (%)	Correct (%)	Incorrect (%)	No ID (%)
matK	189 (360)	9	28	72	72	1	27	0	46	11	42	1
rbcLa	189 (360)	5	20	80	79	0	21	0	66	5	29	0
trnH-psbA	130 (202)	25	19	81	42	3	50	5	26	11	58	5
ITS1	158 (248)	40	**30**	70	38	**9**	43	10	15	20	55	10
rbcLa + *matK*	189 (360)	16	29	71	71	0	27	2	31	18	49	2
rbcLa+matK+trnH-psbA	130 (202)	50	22	78	46	3	47	4	11	17	67	5
rbcLa+matK+ITS1	158 (248)	51	25	75	36	4	51	9	5	20	66	9
rbcLa+matK+trnH-psbA+ITS1	122 (183)	62	26	74	29	7	50	14	0	**23**	63	14

Note: "TRUE" indicates instances when the near neighbour method finds the closest individual in the data set and the name is the same; "FALSE", the name is different. "Correct", "Incorrect", "Ambiguous", and "No ID" are used in the best close match method when the name of the closest match is the same, the name is different, more than one species is the closest match, or no species is within the threshold distance, respectively. Values highlighted in bold represent the DNA region with the highest TRUE/Correct identification among the DNA barcoding regions tested.

of the generally low specimen identification rate of the markers tested in this study, could be a more favourable barcode for the subfamily.

Second, a good DNA barcode should be easily amplified with universal primers (CBOL Plant Working Group 2009). In our study, the plastid genes *matK*, *rbcLa*, and *trnH-psbA* were easily amplified with universal primers. Although ITS1 was more difficult to amplify, leading to the poorest PCR success and sequence quality we found, it was consistently retrieved as the best-performing locus in the genetic variation analysis. The low sequence quality recorded for ITS1 could be an artefact of errors in homopolymeric regions, where sequences of identical bases occur in tandem (Bizzaro and Marx 2003). This could be overcome through the use of anchored primers (Thomas et al. 1993) or primers that anneal at a different position. It could also be due to multiple variants within single individuals, as is the case in Alooideae, with high rates of hybridization (Ramdhani et al. 2011). Previous molecular taxonomic studies in different Alooideae lineages (e.g., Treutlein et al. 2003*a*, 2003*b*; Ramdhani et al. 2011; Daru et al. 2013; Manning et al. 2014*a*) have consistently favourably appraised the utility of ITS1 in species discrimination and disentangling phylogenetic relationships, as in many angiosperm families (Baldwin et al. 1995). This relatively high resolution of ITS1, compared with other markers, is an indication of better species discrimination, confirming ITS1 as a better barcode for the subfamily. Given the high interspecific variation of ITS1, we argue that if universal primers that could boost its amplification success could be designed, this marker could be a suitable barcode for Alooideae.

In general, the ITS region as a potential barcode has been controversial (see Kress et al. 2005), and recent studies have raised some potential concerns about its suitability (e.g., incomplete lineage sorting, inhomogeneous concerted evolution, divergent paralogous copies within individuals, and pseudogenes; Alvarez and Wendel 2003; Chase et al. 2007; Starr et al. 2009; Hollingsworth et al. 2011). However, a more recent test of ITS on a large data set revealed that these drawbacks are not sufficiently severe to preclude consideration of this marker (China Plant BOL Group 2011), giving further support to our advocacy of ITS1 for the monocot subfamily considered in this study (see also Liu et al. 2011 for the genus *Taxus* and Yang et al. 2012 for the genus *Parnassia*).

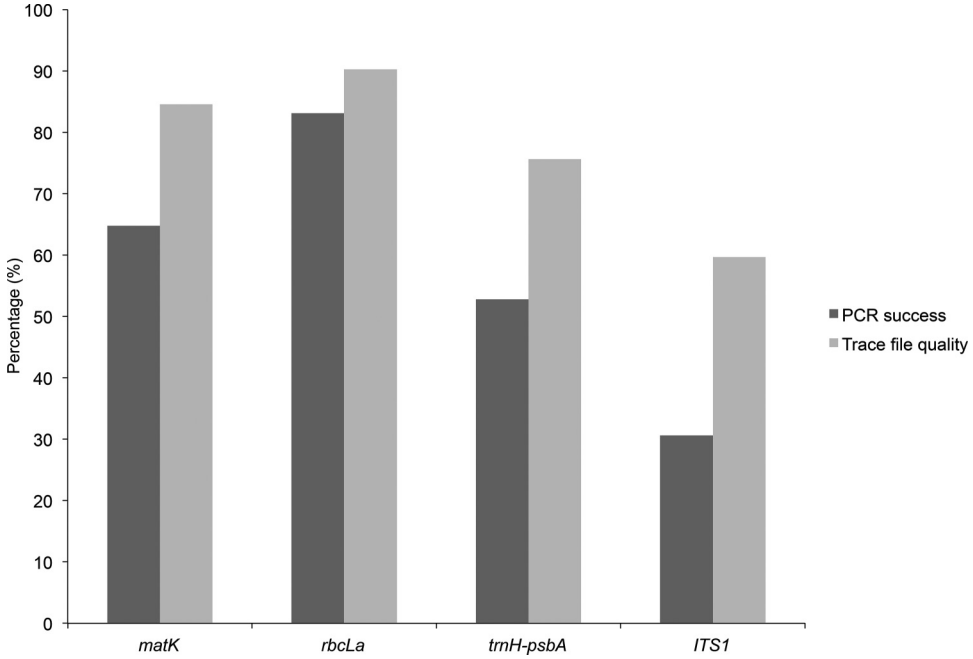
Looking at other potential barcodes, we found that *rbcLa* has shown the lowest intra- and interspecific distances (see Lahaye et al. 2008; Clerc-Blain et al. 2010; Zuo et al. 2011). This marker not only has high universality and sequence quality (see CBOL Plant Working Group 2009) but is also well known for its high discrimination power at higher taxonomic levels, i.e., generic and familial levels (Kress and Erickson 2007). However, in this study, as in previous studies (e.g., Lahaye et al. 2008), it showed relatively low discriminating power between

Table 4. Comparisons of efficacy of core barcodes and best barcode within Alooideae genera using the best close match method.

Genus (no. of species)	DNA region (no. of samples)	Mean inter (±SD)	Threshold (%)	Best close match			
				Ambiguous (%)	Correct (%)	Incorrect (%)	No ID (%)
Aloe (72)				15	7	77	1
	rbcLa+matK (98)	0.027±0.065	1.95	10	11	79	0
	rbcLa+matK+ITS1 (98)	0.032±0.058	3.95	3	14	82	1
Astroloba (6)	ITS1 (8)	0.038±0.021	2.92	0	25	63	12
	rbcLa+matK (8)	0.044±0.079	10.31	0	25	63	12
	rbcLa+matK+ITS1 (8)	0.038±0.06	8.16	0	25	63	12
Gasteria (19)	ITS1 (22)	0.0035±0.0037	0.17	27	0	50	23
	rbcLa+matK (22)	0.0035±0.0025	0.50	0	0	95	5
	rbcLa+matK+ITS1 (22)	0.0029±0.0017	0.85	0	0	100	0
Haworthia (37)	ITS1 (70)	0.046±0.032	1.89	24	20	51	5
	rbcLa+matK (70)	0.0068±0.0049	1.76	31	9	60	0
	rbcLa+matK+ITS1 (70)	0.009±0.0055	0.47	16	24	56	4
Haworthiopsis (12)	ITS1 (32)	0.041±0.02	6.78	0	63	37	0
	rbcLa+matK (32)	0.006±0.0025	0.34	22	34	38	6
	rbcLa+matK+ITS1 (32)	0.014±0.0057	1.15	0	50	44	6

Note: "Correct", "Incorrect", "Ambiguous", and "No ID" mean that the name of the closest match is the same, the name is different, more than one species is the closest match, or no species is within the threshold distance, respectively. The mean interspecific distance refers to Kimura's 2-parameter divergence between congenerics.

Fig. 2. Percentages for PCR efficiency (based on the quality of PCR bands) and trace file sequence quality for the candidate DNA barcodes (*rbcLa*, *matK*, *trnH-psbA*, and ITS1) in identifying species of Alooideae.



species and therefore could not be recommended as a potential barcode for the subfamily at the species level.

The discriminatory power of the DNA regions for species-level resolution differed among genera, with fair performance in *Haworthiopsis* and poor performance in *Aloe* and *Haworthia*. The poor performance is not surprising owing to the generally low genetic variation often found in Alooideae lineages (Ramdhani et al. 2011; Daru et al. 2013; Grace et al. 2015). In addition, our study indicates that species discrimination within a large taxonomic group with closely related taxa should be tested within genera, with dense sampling of species (see also Gere et al. 2013). With the growing availability of next-generation sequencing, a more extensive approach, for example multi-marker analysis methods, chloroplast sequencing, or using more parts of the nuclear genome, could be required to yield additional discriminating regions.

Going forward, we suggest a three-pronged approach to reduce the high rate of incorrect specimen identifica-

tions in Alooideae. First, including more replicates per species would allow comparison of intra- and interspecific genetic divergence. However, this option would not likely change our findings significantly because our sampling included some replication within species (see Table 1), yet we found poor discriminatory power, as in previous studies (e.g., Clement and Donoghue 2012; Yang et al. 2012). Second, more multigene methods in search of variable markers should be developed. However, this option may be counterintuitive given that the purpose of DNA barcoding is to ease specimen identification and achieve universality in specimen discrimination. Third, DNA barcoding could also be tested using a tree-based method in a phylogenetic context (see Mankga et al. 2013). This is being tested in some plant groups with good results, for example Combretaceae (Gere et al. 2013) and medicinal plants (Mankga et al. 2013).

Overall, we suggest that the use of ITS1 alone or in combination with the core barcodes (*rbcLa* + *matK*) has fair barcode potential for the subfamily Alooideae. However, the barcode potential of these regions might vary across the different Alooideae genera. The taxonomy of the alooids is still rife with uncertainty and controversy, such that new classification systems are rapidly emerging (Grace et al. 2013; Rowley 2013; Manning et al. 2014*a*, 2014*b*). We hope that our study will quickly be followed by others where new and more universal ITS1 primers could be investigated to boost amplification success.

Implications for conservation

Various species within Alooideae have restricted populations and are also of high horticultural appeal and are therefore threatened by illegal and excessive collection in the wild. For instance, *Kumara disticha* is listed in CITES Appendix II, implying that the species is of conservation concern and international trade should be limited. Since DNA barcoding has been used to track down illegal trade in endangered species, for example fin whale trade (Baker et al. 2010) and illegal logging of protected tree species (Degen and Fladung 2007), it follows that DNA barcoding could also assist conservationists in managing and tracking down species of Alooideae that are highly threatened, for example *Aloidendron pillansii* (critically endangered), *Astroloba rubriflora* (vulnerable), *Haworthia pubescens* (critically endangered), *Haworthiopsis longiana* (endangered), and *Tulista kingiana* (endangered) (www.redlist.sanbi.org). Thus, an identification tool such as DNA barcoding that can reliably identify species of Alooideae will go a long way to help preserve these species along with their horticultural appeal.

Acknowledgements

We would like to thank the following organizations for funding and logistic support: the University of Johannesburg, the Royal Society of London, and the National Research Foundation of South Africa. Part of this project was also funded by the Government of Canada through Genome Canada and the Ontario Genomics Institute (2008-OGI-ICI-03). We thank the Associate Editor Sarah Adamowicz and two anonymous referees for comments on an earlier draft of the manuscript.

References

Alvarez, I., and Wendel, J.F. 2003. Ribosomal ITS sequences and plant phylogenetic inference. Mol. Phylogenet. Evol. **29**(3): 417–434. PMID:14615184.

Baker, C.S., Steel, D., Choi, Y., Lee, H., Kim, K.S., Choi, S.K., et al. 2010. Genetic evidence of illegal trade in protected whales links Japan with the US and South Korea. Biol. Lett. **6**: 647–650. PMID:20392716.

Baldwin, B.G., Sanderson, M.J., Porter, J.M., Wojciechowski, M.F., Campbell, S.G., and Donoghue, M.J. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Ann. Mo. Bot. Gard. **82**: 247–277. doi:10.2307/2399880.

Bayer, M.B. 1999. *Haworthia* revisited. A revision of the genus. Umdaus Press, Hatfield, South Africa.

Bizzaro, J.W., and Marx, K.A. 2003. Poly: a quantitative analysis tool for simple sequence repeat (SSR) tracts in DNA. BMC Bioinform. **4**: 22–27. PMID:12791171.

Brown, S.D., Collins, R.A., Boyer, S., Lefort, M.C., Malumbres-Olarte, J., Vink, C.J., and Cruickshank, R.H. 2012. Spider: an R package for the analysis of species identity and evolution, with particular reference to DNA barcoding. Mol. Ecol. Resour. **12**(3): 562–565. PMID:22243808.

Burgess, K.S., Fazekas, A.J., Kesanakurti, P.R., Graham, S.W., Husband, B.C., Newmaster, S.G., et al. 2011. Discriminating plant species in a local temperate flora using the *rbcL* plus *matK* DNA barcode. Methods Ecol. Evol. **2**(4): 333–340. doi:10.1111/j.2041-210X.2011.00092.x.

CBOL Plant Working Group 2009. A DNA barcode for land plants. Proc. Natl. Acad. Sci. U.S.A. **106**(31): 12794–12797. PMID:19666622.

Chase, M.W., De Bruijn, A.Y., Cox, A.V., Reeves, G., Rudall, P.J., Johnson, M.A.T., and Eguiarte, L.E. 2000. Phylogenetics of Asphodelaceae (Asparagales): an analysis of plastid *rbcL* and *trnL-F* DNA sequences. Ann. Bot. **86**(5): 935–951.

Chase, M.W., Cowan, R.S., Hollingsworth, P.M., van den Berg, C., Madrinan, S., Petersen, G., et al. 2007. A proposal for a standardised protocol to barcode all land plants. Taxon, **56**(2): 295–299.

China Plant BOL Group. 2011. Comparative analysis of a large dataset indicates that ITS should be incorporated into the core barcode for seed plants. Proc. Natl. Acad. Sci. U.S.A. **108**(49): 19641–19646. PMID:22100737.

Clement, W.L., and Donoghue, M.J. 2012. Barcoding success as a function of phylogenetic relatedness in *Viburnum*, a clade of woody angiosperms. BMC Evol. Biol. **12**: 73. PMID:22646220.

Clerc-Blain, J.L.E., Starr, J.R., Bull, R.D., and Saarela, J.M. 2010. A regional approach to plant DNA barcoding provides high species resolution of sedges (*Carex* and *Kobresia*, Cyperaceae) in the Canadian Arctic Archipelago. Mol. Ecol. Resour. **10**(1): 69–91. PMID:21564992.

Daru, B.H. 2012. Molecular phylogenetics of Alooideae (Asphodelaceae). M.Sc. thesis, University of Johannesburg, South Africa.

Daru, B.H., and Yessoufou, K. 2016. A search for a single DNA barcode for seagrasses of the world. *In* DNA barcoding in marine perspectives. *Edited by* S. Trivedi, A.A. Ansari, and S.K. Ghosh. Springer International Publishing, Switzerland. pp. 313–330. doi:10.1007/978-3-319-41840-7_19.

Daru, B.H., Manning, J.C., Boatwright, J.S., Maurin, O., Maclean, N., Schaefer, H., et al. 2013. Molecular and morpho-

logical analysis of subfamily Alooideae (Asphodelaceae) and the inclusion of *Chortolirion* in *Aloe*. Taxon, **62**(1): 62–76.

Davies, T.J., Kraft, N.J.B., Salamin, N., and Wolkovich, E.M. 2012. Incompletely resolved phylogenetic trees inflate estimates of phylogenetic conservatism. Ecology, **93**(2): 242–247. doi:10.1890/11-1360.1. PMID:22624305.

Degen, B., and Fladung, M. 2007. Use of DNA-markers for tracing illegal logging. *In* Proceedings of the International Workshop "Fingerprinting Methods for the Identification of Timber Origins," Bonn Press. pp. 6–14.

Duval, H.A. 1809. Plantae succulentae, in Horto Alenconio. Parisiis apud Gabon et Socios, Paris.

Eaton, M.J., Meyers, G.L., Kolokotronis, S.O., Leslie, M.S., Martin, A.P., and Amato, G. 2010. Barcoding bushmeat: molecular identification of Central and South American harvested vertebrates. Conserv. Genet. **11**(4): 1389–1404. doi:10.1007/s10592-009-9967-0.

Edwards, D., Horn, A., Taylor, D., Savolainen, V., and Hawkins, J.A. 2008. DNA barcoding of a large genus, *Aspalathus* L. (Fabaceae). Taxon, **57**(4): 1317–1327.

Fu, Y.M., Jiang, W.M., and Fu, C.X. 2011. Identification of species within *Tetrastigma* (Miq.) Planch. (Vitaceae) based on DNA barcoding techniques. J. Syst. Evol. **49**(3): 237–245. doi:10.1111/j.1759-6831.2011.00126.x.

Gene Codes Corporation. 2016. Quality scores. Gene Codes Corporation, Ann Arbor, MI 48108, USA.

Gere, J., Yessoufou, K., Daru, B.H., Maurin, O., Mankga, L.T., and Van der Bank, M. 2013. Incorporating *trnH-psbA* to core DNA barcodes improves discrimination of species within southern African Combretaceae. ZooKeys, **365**: 127–147. doi:10.3897/zookeys.365.5728.

Glen, H.F., and Hardy, D.S. 2000. Aloaceae. *Aloe*. *In* Flora of Southern Africa. Fascicle 1: *Aloaceae* (First part): Aloe. Vol. 5. *Edited by* G. Germishuizen. National Botanical Institute, Pretoria, South Africa. pp. 1–167.

Gouy, M., Guindon, S., and Gascuel, O. 2010. SeaView Version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol. **27**(2): 221–224. PMID:19854763.

Grace, O.M., Klopper, R.R., Smith, G.F., Crouch, N.R., Figueiredo, E., Rønsted, N., and Van Wyk, A.E. 2013. A revised generic classification for *Aloe* (Xanthorrhoeaceae subfam. Asphodelaceae). Phytotaxa, **76**(1): 7–14. doi:10.11646/phytotaxa.76.1.2.

Grace, O.M., Buerki, S., Symonds, M.R., Forest, F., Van Wyk, A.E., Smith, G.F., et al. 2015. Evolutionary history and leaf succulence as explanations for medicinal use in aloes and the global popularity of *Aloe vera*. BMC Evol. Biol. **15**: 29. PMID:25879886.

Groen, L.E. 1987. *Astroloba* Uitew. (III). Succulenta, **66**: 82–87.

Hebert, P.D.N., Cywinska, A., Ball, S.L., and de Waard, J.R. 2003. Biological identifications through DNA barcodes. Proc. R. Soc. B Biol. Sci. **270**(1512): 313–321. doi:10.1098/rspb.2002.2218.

Hebert, P.D.N., Penton, E.H., Burns, J.M., Janzen, D.H., and Hallwachs, W. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proc. Natl. Acad. Sci. U.S.A. **101**(41): 14812–14817. PMID:15465915.

Hollingsworth, M.L., Clark, A.A., Forrest, L.L., Richardson, J., Pennington, R.T., Long, D.G., et al. 2009. Selecting barcoding loci for plants: evaluation of seven candidate loci with species level sampling in three divergent groups of land plants. Mol. Ecol. Resour. **9**(2): 439–457. PMID:21564673.

Hollingsworth, P.M., Graham, S.W., and Little, D.P. 2011. Choosing and using a plant DNA barcode. PLoS ONE, **6**(5): e19254. doi:10.1371/journal.pone.0019254. PMID:21637336.

Klopper, R.R., and Smith, G.F. 2007. The genus *Aloe* L. (Apshodelaceae: Alooideae) in Namqualand, South Africa. Haseltonia, **13**: 1–13.

Klopper, R.R., and Smith, G.F. 2011. The genus *Aloe* L. (Apshodelaceae: Alooideae) in the Eastern Cape Province of South Africa. Haseltonia, **16**: 16–53.

Kress, W.J., and Erickson, D.L. 2007. A two-locus global DNA barcode for land plants: The coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. PLoS ONE, **2**(6): e508. doi:10.1371/journal.pone.0000508. PMID:17551588.

Kress, W.J., Wurdack, K.J., Zimmer, E.A., Weigt, L.E., and Janzen, D.H. 2005. Use of DNA barcodes to identify flowering plants. Proc. Natl. Acad. Sci. U.S.A. **102**(23): 8369–8374. PMID:15928076.

Lahaye, R., van der Bank, M., Bogarin, D., Warner, J., Pupulin, F., Gigot, G., et al. 2008. DNA barcoding the floras of biodiversity hotspots. Proc. Natl. Acad. Sci. U.S.A. **105**(8): 2923–2928. PMID:18258745.

Liu, J., Moller, M., Gao, L.M., Zhang, D.Q., and Li, D.Z. 2011. DNA barcoding for the discrimination of Eurasian yews (*Taxus* L., Taxaceae) and the discovery of cryptic species. Mol. Ecol. Resour. **11**(1): 89–100. PMID:21429104.

Lowe, A.J., and Cross, H.B. 2011. The application of DNA to timber tracking and origin verification. IAWA J. **32**(2): 251–262. doi:10.1163/22941932-90000055.

Maddison, W.P., and Maddison, D.R. 2008. Mesquite: a modular system for evolutionary analysis. Version 2.5. Available from http://mesquiteproject.org.

Mankga, L.T., Yessoufou, K., Moteetee, A.M., Daru, B.H., and 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. doi:10.3897/zookeys.365.5730.

Manning, J.C., Boatwright, J.S., Daru, B.H., Maurin, O., and Van der Bank, M. 2014*a*. A molecular phylogeny and generic classification of Asphodelaceae subfamily Alooideae: a final resolution of the prickly issue of polyphyly in the Alooids? Syst. Bot. **39**(1): 55–74. doi:10.1600/036364414X678044.

Manning, J.C., Boatwright, J.S., and Daru, B.H. 2014*b*. *Aloe* and goodbye: a new evolutionary classification of the Alooids. Alsterworthia International, **14**: 7–15.

Meier, R.S., Kwong, S., Vaidya, G., and Ng, P.K.L. 2006. DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. Syst. Biol. **55**(5): 715–728. PMID:17060194.

Meyer, C.P., and Paulay, G. 2005. DNA barcoding: error rates based on comprehensive sampling. PLoS Biol. **3**(12): 2229–2238. doi:10.1371/journal.pbio.0030422.

Paradis, E., Claude, J., and Strimmer, K. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics, **20**(2): 289–290. doi:10.1093/bioinformatics/btg412. PMID:14734327.

Pettengill, J.B., and Neel, M.C. 2010. An evaluation of candidate plant DNA barcodes and assignment methods in diagnosing 29 species in the genus *Agalinis* (Orobanchaceae). Am. J. Bot. **97**(8): 1391–1406. doi:10.3732/ajb.0900176. PMID:21616891.

Raimondo, D., Von Staden, L., Foden, W., Victor, J.E., Helme, N.A., Turner, R.C., et al. 2009. Red List of South African Plants. Strelitzia 25. South African National Biodiversity Institute, Pretoria.

Ramdhani, S., Barker, N.P., and Cowling, R.M. 2011. Revisiting monophyly in *Haworthia* Duval (Asphodelaceae): incongruence, hybridization and contemporary speciation. Taxon, **60**(4): 1001–1014.

Ratnasingham, S., and Hebert, P.D.N. 2007. bold: The Barcode of Life Data System (http://www.barcodinglife.org). Mol. Ecol. Notes, **7**(3): 355–364. doi:10.1111/j.1471-8286.2007.01678.x. PMID:18784790.

Reynolds, G.W. 1966. The aloes of tropical Africa and Madagascar. The Trustees: The Aloes Book Fund, Mbabane, Swaziland.

Riley, H.P., and Majumdar, S.K. 1979. The Aloineae: a biosystematic survey. Kentucky: The University Press of Kentucky.

Roberts Reinecke, P. 1965. A revision of the genus *Astroloba*. MSc thesis, University of Cape Town, South Africa.

Rowley, G.D. 2013. Generic concepts in the Alooideae 3: The phylogenetic story. Alsterworthia International, **10**(Special issue): 1–7.

Roy, S., Tyagi, A., Shukla, V., Kumar, A., Singh, U.M., Chaudhary, L.B., Datt, B., et al. 2010. Universal plant DNA barcode loci may not work in complex groups: a case study with Indian *Berberis* species. PLoS ONE, **5**(10): e13674. doi:10.1371/journal.pone.0013674. PMID:21060687.

Saslis-Lagoudakis, C.H., Savolainen, V., Williamson, E.M., Forest, F., Wagstaff, S.J., Baral, S.R., et al. 2012. Phylogenies reveal predictive power of traditional medicine in bioprospecting. Proc. Natl. Acad. Sci. U.S.A. **109**(39): 15835–15840. PMID:22984175.

Smith, G.F., and van Wyk, B.E. 1991. Generic relationships in the Alooideae (Asphodelaceae). Taxon, **40**(4): 557–581. doi:10.2307/1222765.

Smith, G.F., Steyn, E.M.A., Victor, J.E., Crouch, N.R., Golding, J., and Hilton-Taylor, C. 2000. The conservation status of *Aloe* in South Africa: an updated synopsis. Bothalia, **30**: 206–211.

Starr, J.R., Naczi, R.F.C., and Chouinard, B.N. 2009. Plant DNA barcodes and species resolution in sedges (*Carex*, Cyperaceae). Mol. Ecol. Resour. **9**(s1): 151–163. PMID:21564974.

Sun, X.Q., Zhu, Y.J., Guo, J.L., Peng, B., Bai, M.M., and Hang, Y.Y. 2012. DNA barcoding the *Dioscorea* in China, a vital group in the evolution of monocotyledon: Use of *matK* gene for species discrimination. PLoS ONE, **7**(2): e32057. doi:10.1371/journal.pone.0032057. PMID:22363795.

Thomas, M.G., Hesse, S.A., McKie, A.T., and Farzaneh, F. 1993. Sequencing of cDNA using anchored oligo dT primers. Nucleic Acids Res. **21**(16): 3915–3916. PMID:8367318.

Treutlein, J., Smith, G.F., van Wyk, B.E., and Wink, M. 2003*a*. Evidence for the polyphyly of *Haworthia* (Asphodelaceae subfamily Alooideae; Asparagales) inferred from nucleotide sequences of *rbcL*, *matK*, ITS1 and genome fingerprinting with ISSR-PCR. Plant Biol. **5**(5): 513–521. doi/10.1055/s-2003-44793.

Treutlein, J., Smith, G.F., van Wyk, B.E., and Wink, M. 2003*b*. Phylogenetic relationships in Asphodelaceae (subfamily Alooideae) inferred from chloroplast DNA sequences (*rbcL*, *matK*) and from genomic fingerprinting (ISSR). Taxon, **52**(2): 193–207. doi:10.2307/3647389.

Van Jaarsveld, E.J. 2007. The genus *Gasteria*: A synoptic review (new taxa and combinations). Aloe, **44**: 84–104.

Viljoen, A.M. 1999. A chemotaxonomic study of phenolic leaf compounds in the genus Aloe. Ph.D. thesis, Rand Afrikaans University, Johannesburg, South Africa.

Wang, W., Wu, Y., Yan, Y., Ermakova, M., Kerstetter, R., and Messing, J. 2010. DNA barcoding of the Lemnaceae, a family of aquatic monocots. BMC Plant Biol. **10**: 205. PMID:20846439.

Wang, X.-C., Liu, C., Huang, L., Bengtsson-Palme, J., Chen, H., Zhang, J.-H., et al. 2015. ITS1: A DNA barcode better than ITS2 in eukaryotes? Mol. Ecol. Resour. **15**(3): 573–586. PMID:25187125.

Wink, M., Guicking, D., and Fritz, U. 2001. Molecular evidence for hybrid origin of *Mauremys iversoni* Pritchard et McCord, 1991, and *Mauremys pritchardi* McCord, 1997 (Reptilia: Testudines: Bataguridae). Zool. Abh. Staatl. Mus. Tierkunde Dresden, **51**: 41–50.

Yang, J.B., Wang, Y.P., Moller, M., Gao, L.M., and Wu, D. 2012. Applying plant DNA barcodes to identify species of *Parnassia* (Parnassiaceae). Mol. Ecol. Resour. **12**(2): 267–275. PMID:22136257.

Yessoufou, K., Daru, B.H., and Muasya, A.M. 2015. Phylogenetic exploration of commonly used medicinal plants in South Africa. Mol. Ecol. Resour. **15**(2): 405–413. PMID:25066923.

Zuo, Y.J., Chen, Z.J., Kondo, K., Funamoto, T., Wen, J., and Zhou, S.L. 2011. DNA barcoding of *Panax* species. Planta Med. **77**(2): 182–187. PMID:20803416.