Canadian Science Publishing
NRC Research Press journals | Publishing Services

Genome / Génome

Ten years of barcoding at the African Centre for DNA barcoding

Journal:	*Genome*
Manuscript ID	gen-2016-0198.R2
Manuscript Type:	Review
Date Submitted by the Author:	24-Feb-2017
Complete List of Authors:	Bezeng, Simeon Bezeng; African Centre for DNA Barcoding, University of Johannesburg, PO Box 524, Auckland Park 2006, Johannesburg, South Africa Davies, T. Jonathan; McGill University, Department of Biology; African Centre for DNA Barcoding, University of Johannesburg, PO Box 524, Auckland Park 2006, Johannesburg, South Africa Daru, Barnabas H; Harvard University Department of Organismic and Evolutionary Biology Kabongo, Ronny Mukala; African Centre for DNA Barcoding, University of Johannesburg, PO Box 524, Auckland Park 2006, Johannesburg, South Africa Maurin, Olivier; African Centre for DNA Barcoding, University of Johannesburg, PO Box 524, Auckland Park 2006, Johannesburg, South Africa; Royal Botanic Gardens YESSOUFOU, KOWIYOU; Department of Geography, Environmental Management and Energy Studies, University of Johannesburg, APK Campus, 2006, Johannesburg, South Africa van der Bank, Herman; African Centre for DNA Barcoding, University of Johannesburg, PO Box 524, Auckland Park 2006, Johannesburg, South Africa Van der Bank, Michelle; African Centre for DNA Barcoding, University of Johannesburg, PO Box 524, Auckland Park 2006, Johannesburg, South Africa
Please Select from this Special Issues list if applicable:	
Keyword:	DNA barcoding, African flora and fauna, species delimitation, ecological applications, CO1

SCHOLARONE™
Manuscripts

Ten years of barcoding at the African Centre for DNA barcoding

B.S. Bezeng[1], T.J. Davies[1,2], B.H. Daru[3], R.M. Kabongo[1], O. Maurin[1,5], K. Yessoufou[4], H. van der Bank[1], M. van der Bank[1]

[1]*African Centre for DNA Barcoding, University of Johannesburg, PO Box 524, Auckland Park 2006, Johannesburg, South Africa*

[2]*Department of Biology, McGill University, 1205 Avenue Docteur Penfield, Montreal, QC H3A 0G4, Quebec, Canada*

[3]*Department of Organismic and Evolutionary Biology, Harvard University, 22 Divinity Avenue, Cambridge, MA 02138, USA*

[4]*Department of Geography, Environmental Management and Energy Studies, University of Johannesburg, APK Campus, 2006, Johannesburg, South Africa*

[5]*Royal Botanic Gardens, Kew, Richmond, Surrey, TW93AB, UK*

Draft

Abstract

The African Centre for DNA Barcoding (ACDB) was established in 2005 as part of a global initiative to accurately and rapidly survey biodiversity using short DNA sequences. The mitochondrial cytochrome c oxidase 1 gene (*CO1*) was rapidly adopted as the *de facto* barcode for animals. Following the evaluation of several candidate loci for plants, the Plant Working Group of the Consortium for the Barcoding of Life in 2009 recommended that the two plastid genes, *rbcLa* and *matK*, be adopted as core DNA barcodes for terrestrial plants. To date, numerous studies continue to test the discriminatory power of these markers across various plant lineages. Over the past decade, we at the African Centre for DNA Barcoding, have used these core DNA barcodes to generate a barcode library for southern Africa. To date, the ACDB has contributed more than 21 000 plant barcodes and over 3 000 *CO1* barcodes for animals to the Barcode of Life Database (BOLD). Building upon this effort, we at the ACDB have addressed questions related to community assembly, biogeography, phylogenetic diversification, and invasion biology. Collectively, our work demonstrates the diverse applications of DNA barcoding in ecology, systematics, evolutionary biology, and conservation.

Draft

Keywords: DNA barcoding, African flora and fauna, *matK, rbcLa, CO1,* species delimitation, ecological applications

1. Introduction

Species morphology has provided the basis for description and classification of life on Earth. However, morphological features can have complex evolutionary histories, and the identification of appropriate homologous structures on which to group taxa is a major challenge (see review by Pereira et al. 2008). Molecular sequence data provides a set of alternative characters that might more reliably capture species differences and evolutionary relationships. It is not completely certain as to when the term 'DNA barcoding' was first applied to molecular markers that allowed discrimination between species (Adamowicz 2015), but the seminal paper by Hebert and colleagues, published in 2003, brought the concept into the mainstream. Hebert et al. (2003) proposed the construction of a reference DNA barcode library of life for species identification. Hence, DNA barcoding found its primary applications mainly in the fields of systematics and taxonomy. The main focus of early work was on testing the discriminatory power of various markers as potential barcodes (Hebert et al. 2003; Hajibabaei et al. 2006; Burgess et al. 2011; Clement and Donoghue 2012; Franzini et al. 2013; Gere et al. 2013). A consensus on the reliability of *COI* as DNA barcode for animals was achieved relatively quickly. However, the quest to find a suitable barcode for plant identification took longer, and it was not until more than a decade after Hebert et al.'s seminal publication that the combination of *rbcLa* + *matK* was adopted by the Plant Working Group of the Consortium for the Barcoding of Life (CBOL) as the barcode for plants (CBOL Plant Working Group 2009). The African Centre for DNA Barcoding (ACDB) contributed significantly in affirming *matK* as a reliable barcode locus for land plants (see Lahaye et al. 2008). Although *rbcLa* and *matK* loci have been adopted as the universal plant barcode regions, many studies

continue to test the reliability of alternative plant DNA barcodes (e.g. Clement and Donoghue 2012; Xu et al. 2015; Daru and Yessoufou 2016).

In recognition of the major contributions of the ACDB to the barcoding effort, the Centre was officially recognised as a major node in Africa for DNA barcoding in 2010. Since its establishment, ACDB's primary aim has been to fill knowledge gaps, and to strengthen research frameworks for regional, international, and inter-institutional co-operation focused on African biodiversity science. We at the ACDB have worked towards validating the application of DNA barcoding to the species-rich flora of southern Africa (e.g. Gere et al. 2013; Mankga et al. 2013; Van der Bank et al. 2013; Maurin et al. 2014; Bello et al. 2015; Hoveka et al. 2016), and over the past decade we have demonstrated not only how DNA barcoding can advance species taxonomy in Africa, but also how barcoding can be integrated within a multidisciplinary framework that can aid conservation decisions and further our understanding of the fundamental processes by which species diversity is generated, distributed and maintained.

Here, we review and synthesize 10-years of DNA barcoding at the ACDB. We emphasize how these studies have advanced ecological and evolutionary knowledge of African biodiversity, and the centre's contribution to documenting global biodiversity, a central goal of the DNA barcoding effort.

2. Brief History of the ACDB

The ACDB is a division of the departments of Botany and Plant Biotechnology, and Zoology of the University of Johannesburg, South Africa. Formally established in

2010 with a vision to remain at the forefront of research on DNA barcoding and its applications for biosecurity, surveillance and biodiversity surveys in Africa, the ACDB, in collaboration with its many partners, has contributed to large scale capacity building across numerous African countries including Angola, Benin, Botswana, Cameroon, Democratic Republic of Congo, Egypt, Gabon, Ghana, Kenya, Lesotho, Madagascar, Malawi, Mozambique, Namibia, Nigeria, South Africa, Swaziland, Tanzania, Zambia, and Zimbabwe (Figure 1). The formation of a highly efficient network of partnerships with the ACDB has allowed the collection of new research data, aided biodiversity protection, and the facilitation of skill development programmes, including but not limited to training of interns, postgraduate students and border control officials.

Draft

Due to the location of the ACDB in South Africa, our work over the past decade has focussed mainly on southern Africa. In South Africa, the ACDB is affiliated with the South African Department of Environmental Affairs (DEA), providing training for interns in laboratory techniques and biodiversity science. The centre has also worked with various governmental and nongovernmental organisations, including the Working for Water Programme, the South African National Biodiversity Institute, and the Buffelskloof Nature Reserve. Other African institutions with close ties include the University of Lagos in Nigeria and the Namibian Tourism and Marine Resource Centre. In addition, the ACDB has established formal collaborations with other institutions in North America, Europe, and Asia.

2.1 Species sampling and DNA barcoding **—** Over the past ten years, the ACDB has conducted extensive field data collection across the continent (Figure 2). These

collections have been concentrated in southern Africa, a region rich in biodiversity, with more than 24 000 plant species, encompassing five biodiversity hotspots (the Cape Floristic Region, Maputaland-Pondoland-Albany, Succulent Karoo, parts of the Coastal Forests of East Africa, and the eastern Afromontane) distributed across diverse vegetation types (Figure 3). Specimen collection has focused mainly on the woody flora, but includes some herbs and fauna (i.e. fish, mammals, insects, molluscs etc; Figure 4). While we have adopted the standard DNA barcodes for plants (*rbcLa* + *matK*) and animals (*COI*) (Hajibabaei et al. 2005; CBOL working group 2009), we have also conducted sequencing projects using additional molecular markers, including *psbA-trnH, trnL-F,* and nrITS for plants, and 16S and 18S for animals.

All specimen metadata, including photographs, taxonomic information, collection details and barcode sequences are submitted to BOLD and GenBank/EBI. To date, the ACDB has contributed some 24, 000 records to the global barcode initiative, encompassing 67 orders of plants and animals (see Table 1).

3. Taxonomic, ecological and evolutionary applications of ACDB barcode data

Since its inception, over 50 scientific publications have used data generated by the ACDB (Figure 5). Research focus has shifted since the establishment of the ACDB, with a move from molecular systematics and phylogenetics, towards a more ecological focus. The majority of studies (89%) were on plants but a number (11%) of studies on animals have also been published. Taxonomic studies make up approximately 56% of all publications; however, an important additional component (44%) of the output from the ACDB has addressed ecological and evolutionary questions, demonstrating the utility of barcode data beyond the fields of taxonomy

and systematics (see also Joly et al. 2013). In the following sections we highlight a few of the key studies that have resulted from this effort.

3.1 Plant taxonomic studies at the ACDB

Plant molecular systematics has been a major focus of our work at the ACDB. We have published a number of taxonomic studies on the African flora, which have been cited over 1200 times in total, including important revisions to African *Acacias*, a significant component of the tree flora of southern Africa, and aloes, a species rich lineage of succulent plants. Although there is a clear distinction between DNA barcoding and molecular systematics—the former has been developed mostly for species identification/discovery, whilst the latter attempts to resolve phylogenetic relationships among species within particular taxonomic groupings—we at the ACDB have integrated DNA barcodes in systematics studies to (i) contribute to the global barcoding campaign and (ii) supplement existing data, or to resolve relationships on 'backbone' phylogenies. It is important to note, however, that in most systematics studies conducted at the ACDB we do not rely solely on barcodes, but rather we use barcodes to supplement DNA studies using additional molecular markers including *psbA-trnH, trnL, trnL-trnF, rpl32-trnL, psaA-ycf3* and ITS1 (see Daru et al. 2013; Kyalangalilwa et al. 2013; Manning et al. 2014; Boatwright et al. 2015; Maurin et al. 2016 and other examples below).

Re-circumscribing African Acacias **—** The ACDB joined the global effort to transfer African *Acacia* s.l. to the genera *Senegalia* Raf. and *Vachellia* Wight & Arn. following the acceptance of the proposed re-typification of the genus with an Australian species. At the ACDB, we used the *matK* barcoding locus to help

Draft

disentangle relationships among *Acacia* species. We demonstrated a polyphyletic relationship within the subgenus *Acacia,* supporting the need for taxonomic revision. Kyalangalilwa et al. (2013) formally transferred the African *Acacia* species to the two genera (i.e. *Senegalia* and *Vachellia*). Despite this advance, the position of the Madagascan species remained unclear from a phylogenetic point of view. However, Boatwright et al. (2015) generated and included sequences for seven Madagascan species into the existing matrix of Acacia s.l. (see Kyalangalilwa et al. 2013) and analyzed the placement of these species in the broader phylogeny. Results indicated that species from this region are placed in either *Senegalia* or *Vachellia*, which is in agreement with the morphological characters used to distinguish between them (Boatwright et al. 2015). Initially, there was hesitation among African botanists in adopting this new classification, but such changes are gaining acceptance in the wider botanical literature.

Draft

Improving Alooideae systematics — We combined DNA barcode data and morphological characters to provide a better understanding of the evolutionary relationships of aloes and related genera within subfamily Alooideae (Daru et al. 2013). Notably, this study identified important shortcomings in the previously accepted classification, a result of homoplasy and the difficulty in identifying consistent characters to distinguish species within genera, suggesting the need for a new treatment of the subfamily, re-circumscribing Alooideae genera into monophyletic groupings (Daru et al. 2013; Manning et al. 2014).

Taxonomic revision and species delimitation within Combretaceae — In 2006 a phylogenetic project was initiated at the ACDB with the aim of investigating species

delimitation and the phylogenetic relationships within the Combretaceae. The dataset generated so far comprises all the genera and key taxa at the infra-generic level, with over 200 specimens sequenced for the core barcode markers plus some additional regions including the plastid *psaA-Ycf3* spacer. These studies supported the classification of the family into two subfamilies, Strephonematoideae and Combretoideae, but also showed a lack of resolution within subtribe Terminaliinae, with *Terminalia* not supported as a monophyletic group. These results led to the reinstatement of two genera, *Grandiflora* and *Trichopetala*, which had previously been sunk into subgenus *Cacoucia* (Maurin et al. 2010; Jordaan et al. 2011a, b), and highlight the need for major taxonomic revisions within this family. This work additionally demonstrated high efficacy of DNA barcoding for species delimitation within African Combretaceae (Gere et al. 2013).

3.2. Faunal diversity studies at the ACDB

While the main focus of the ACDB has been on the flora of southern Africa, the Centre has also addressed a number of problematic animal clades, including species delineation of the African silver catfish, and the cryptic diversity of scale insects, spiders, and gastropods.

Taxonomy and phylogeography of the African silver catfish — Members of the catfish genus, *Schilbe*, notably *Schilbe intermedius,* have great economic and nutritional value because they are widely distributed and abundant across African river systems, and they are a good source of protein and omega three oils. This catfish was traditionally divided into different genera based on the presence or absence of adipose fins. However, researchers subsequently observed a rudimentary adipose fin on the

holotype for *Eutropius* that may have been overlooked by early taxonomists, and synonymized *Eutropius depressirostris* and *Schilbe mystus* with *Schilbe intermedius* (De Vos and Skelton 1990). Our results, using the *CO1* barcode, not only revealed that they are different species, but also showed a strong phylogeographical pattern in their distribution (Van der Bank et al. 2012).

Cryptic diversity: scale insects and spiders — Scale insects (Coccoidea) are potentially important crop pests; however, identifying specimens is complicated because they are often immature or damaged, making morphological characters difficult or in some cases impossible to analyse. We combined the *CO1* DNA barcode regions together with the nuclear regions *18S* and *28S* across 10 scale insect families, and showed that it was possible to identify species with up to ~92% accuracy (Sethusa et al. 2014). Correct and rapid identification of scale insects is critical because delays at ports of entry could cause damage to imported food. In a separate study on tent-web spiders, we combined the *COI* barcode and sequence information for the nuclear H3 gene along with morphological characters, and revealed multiple species of tent-web spiders in South Africa (Franzini et al. 2013). This result was significant since the diversity of the southern African species has been little explored, and only the cosmopolitan *Cyrtophora citricola* was previously reported to occur in South Africa.

*Diversity and taxonomy of sea snail**s*** — Species within *Oxystele,* a genus of sea snail endemic to southern Africa are abundant on our shores and an important component of littoral biodiversity in rocky intertidal habitats. We used DNA barcoding to delimit five molecular operational taxonomic units (MOTUs) within the genus and

Draft

distinguish *O. impervia* from *O. variegata* (Van der Bank et al. 2013), revealing discrepancies between MOTUs and morphology-based species identification. In a related study, we investigated the diversity of sea snails on Robben Island (Van der Bank and Greenfield 2015). We expected that the species composition of the island would be similar to that of the Cape Peninsula (South Africa) since they were connected approximately 10 000 years ago. However, we showed differences in species composition, for example, the South American bisexual mussel, which is extremely abundant and invasive on the South African mainland, was not detected on Robben Island.

3.3 Ecological and evolutionary applications of DNA barcoding

While DNA barcoding was developed to aid species identification, and the majority of publications applying barcode tools have addressed species taxonomy and delineation (see Figure 5 for studies published by the ACDB), barcode data can also inform ecological and evolutionary studies (Joly et al. 2013). Perhaps most obviously, barcodes can help resolve evolutionary relationships among taxa. Although the phylogenetic information content of barcodes is relatively low, by combining barcodes with additional DNA markers it is possible to generate well resolved phylogenetic trees for species rich groups (Joly et al. 2013). In plants, the deeper evolutionary relationships among families and orders have been well studied, using multiple markers (see APG IV 2016). It is therefore possible to constrain phylogenetic analyses by including a well-supported backbone topology, and then use barcode markers to resolve fine-scale relationships within families and genera, and to calibrate phylogenetic branch lengths to provide estimates of divergence times. At the ACDB, we have used a barcode phylogeny of the trees of southern Africa to explore a number

of ecological and evolutionary hypotheses that have provided new insights into the history of the flora in the region.

Plant community dynamics in response to large mammal disturbances — There is an on-going debate on the role of large herbivores in the structuring of plant communities, and on the impacts of generalist versus specialist herbivores (Cavender-Bares et al. 2009; Kartzinel et al. 2015). Here, using DNA barcode data, we assembled the first complete phylogeny for all trees and shrubs of the Kruger National Park (KNP) in South Africa (Lahaye et al. 2008), a protected area of ~20 000 km^2 that encompasses the Maputo-Pondoland Albany biodiversity hotspot. Using this phylogeny and data from long term experimental plots, we explored the impact of large herbivores on plant community structure. We showed that the exclusion of large herbivores results in impoverished plant species diversity (Yessoufou et al. 2013 but see also Kalwij et al. 2010). However, we revealed that plant community responses to herbivore exclusion depend on the initial structure of the community: communities that are initially phylogenetically over-dispersed (comprised of distant relatives) tend to become under-dispersed (comprised of more closely related species) after herbivore exclusion, and communities that are initially under-dispersed tend to become over-dispersed after herbivore exclusion.

Phylogenetically informed biogeographical regionalization — The extensive DNA barcoding campaign of southern African trees conducted at the ACDB has enabled the construction of a time-calibrated phylogeny for the region that sheds light on evolutionary affinity of the biomes within it (Daru et al. 2016). This study showed that existing biogeographic divisions in the region represent discrete phylogenetic

units, with the Fynbos biome being the most evolutionarily distinct. We were also able to identify two additional regions, the 'Gariep Karoo' and the 'Zambezian transition zone', that were phylogenetically distinct but which had been overlooked by more traditional species-level approaches.

DNA Barcoding and conservation of regional plant diversity —Using the same dated barcode phylogeny as above, we also explored the spatial distribution of different dimensions of tree diversity across southern Africa, and evaluated the efficiency of hotspots in capturing complementary areas of species richness and phylogenetic diversity (Daru et al. 2015). We compared patterns of species richness, phylogenetic diversity, phylogenetic endemism, species endemism, and 'evolutionary distinctiveness and global endangerment', and revealed spatial incongruence between these different facets of diversity, with no hotspot area shared among all diversity metrics, and almost 70% of hotspot areas unique to a single diversity metric (Daru et al. 2015). The use of DNA barcode data thus aids integrative conservation planning, and allowed us to evaluate the efficacy of protected areas in conserving evolutionary history.

The evolutionary history of African cycads: the role of underground stems —Cycads are a unique plant group that likely originated ~300 Ma, and extant cycads have been regarded as living fossils. However, recent findings suggest that extant cycads are not as old as previously thought (Nagalingum et al. 2011; Condamine et al. 2015). Using a dated DNA barcode phylogeny, we revisited the evolutionary diversification of Africa's endemic cycad *Encephalartos* (Yessoufou et al. 2014). Patterns of Cycad diversification through time suggested a mass extinction event followed by a rate shift

that most likely resulted from a rapid recolonization of emptied niches, and correlates with the diversification of the southern Africa's cycads with underground stems (Yessoufou et al. 2014). This underground stem may have been a key innovation associated with the arid environments that prevailed during the Pliocene–Pleistocene transition in southern Africa.

Africa's trees go underground in response to fire — Extending our analyses of southern African trees, we used barcode data to explore the phylogenetic placement of 'geoxyles suffrutices', so called 'underground trees', to examine the evolutionary drivers of this unusual life-form (Maurin et al. 2014). Phylogenetic analysis revealed multiple independent origins for the geoxyle life-form, and suggested that most geoxylic species arose in the Pliocene, coincident with the diversification of cycads with underground stems, and with the origin of cerrado woodland savanna in Brazil (see Pennington et al. 2006; Simon et al. 2009). Our analysis indicated that the geoxylic life-form most likely evolved in response to the effects of frequent fires where high precipitation occurs (Maurin et al. 2014; Davies et al. 2016).

Mammals, spiny plants and the savanna story — The origin and timing of the spread of savannas in Africa has been much debated. One explanation for the spread of savanna is an increase in fire frequency, promoted by the expansion of C_4 grasses (Ehleringer et al. 1997). An alternative explanation is that large mammal browsers pushed back the forest and allowed the expansion of the savanna biome in habitats otherwise suitable for trees. Using a dataset consisting of over 1 800 tree species across Africa, we explored the importance of mammal browsers in opening up forest in Africa, using spinescence as a marker of mammal herbivory. We found that

spinescence evolved at least 55 times but that the diversification of spiny plants occurred long after the evolution of Afrotherian proboscideans and hyracoids (Charles-Dominique et al. 2016). However, the diversification of spiny lineages coincides with the appearance of antelopes (Charles-Dominique et al. 2016), suggesting that herbivore-adapted savannas evolved in response to a new type of mammalian 'predator' and several million years before fire-maintained savannas.

Darwin's naturalization conundrum and invasive species in South Africa — Of the many threats to southern Africa's flora, alien species may be the single most detrimental (Mack et al. 2000; Richardson and Van Wilgen 2004; Pimental et al. 2005). Through its national invasive and Toyota enviro outreach (http://www.klipbokkop.co.za/logisticalsupport/toyotaoutreach.html) programmes, the ACDB has generated and assembled an extensive DNA barcode library for non-native terrestrial and aquatic plants in the region. By placing non-native trees within our DNA barcode phylogeny, we were able to test Darwin's Naturalization Hypothesis, which suggests that "introduced species are more likely to become naturalized and invasive in new environments where (phylogenetically) close relatives are absent" (Daehler 2001). We showed that non-native plants distantly related to native plants were more likely to become invasive, whereas non-native plants closely related to native species were more likely to establish in the region, but were less likely to become invasive (Bezeng et al. 2015). Our use of a DNA barcode phylogeny thus helps us move towards resolving Darwin's naturalization conundrum (Diez et al. 2008).

In a separate study, we were able to demonstrate that the *psbA-trnH* gene-region was a reliable DNA barcode for freshwater non-native plant species in South Africa

(Hoveka et al. 2016). We then used our *psbA-trnH* barcode library to show that some prohibited aquatic invaders are currently being sold in the local aquarium trade (Hoveka et al. 2016), highlighting the importance of DNA barcodes in monitoring invasive species. These results together with current research on the influence of aquarium trade in the spread of aquatic non-native species have important implications for the countrywide campaign to promptly detect and rapidly eradicate non-native species before they become problematic.

Traditional medicine and socio-economic applications — Poverty in Africa and South Africa in particular, has favoured the use of traditional medicine in many rural communities. With rapid population growth, the traditional medicine industry has expanded greatly in recent years, and plant specimens (e.g. barks, bulbs, roots, and leaves or in some cases whole plant) are often harvested in an unsustainable manner. This unsustainable harvesting of plant materials from the wild has been identified as a leading biodiversity challenge in the 21st century (Williams et al. 2013). However, policing the collection of plant material is difficult. Once harvested, specimen identification becomes extremely challenging since most morphological features that aid identification are often damaged or missing (see also Hollingsworth et al. 2016). At the ACDB, we have used DNA barcoding to test the species identity of plant materials commonly traded in South Africa's traditional medicine markets. In one study, Williamson et al. (2016) sampled 45 cycad specimens from the Faraday and Warwick markets located in Johannesburg and Durban respectively, and found five *Encephalartos* spp, which are listed in CITES *Appendix I* as threatened or protected species in terms of the National Environmental Management: Biodiversity Act (NEM:BA) of 2004. Trade in such species is prohibited except for non-commercial

Draft

purposes (e.g. scientific research). In a similar study Lekganyane et al. (2015), used DNA barcoding to identify commonly traded bulb species for sale at the Faraday traditional medicine market, and demonstrated species identification increased from 36%, using morphological features, to 90%.

4. Challenges Ahead

Since 2005 the ACDB has been at the forefront of DNA barcoding science in Africa. We at the ACDB have contributed over 24, 000 records to BOLD and published 50 papers in international peer reviewed journals using DNA barcoding data. However, considerable challenges remain. First, Africa is a megadiverse continent, home to nine out of the world's 34 biodiversity hotspots, five of which are found within southern Africa (UNEP 2008). Documenting this diversity will require dedicated efforts of trained taxonomists, yet the number of taxonomists in this scarce skills field is steadily declining in Africa and worldwide (Hopkins and Freckleton 2002). Training of the next generation of taxonomists is a major long-term goal of the ACDB. Second, efforts to barcode the flora and fauna of Africa will require new collaborations with additional institutions and research agencies across the African continent. To date, our collection efforts have been biased towards southern Africa. Thus, a key challenge over the next few years will be to extend the existing network of formal collaborations and the training of scientists in DNA barcoding and its applications. This initiative will be showcased at the up-coming International Barcode of Life Conference (http://dnabarcodes2017.org), which will be hosted in South Africa in 2017.

"Metabarcoding" at the ACDB — The ACDB has extended the application of traditional DNA barcoding techniques over the last decade (see Figure 5); however,

Draft

there has been an increasing awareness of the limitations of traditional DNA barcode tools, for example, in resolving relationships among closely related species, and in their high sensitivity to contaminants. Analyzing environmental samples, which often contain a mixture of non-target species, presents additional challenges (Kress et al. 2005; Handley 2015). These limitations may be particularly important when applying DNA barcoding techniques to mixed samples, especially for the tracking of wildlife crime, early detection of alien invasive species at ports of entry, and adulteration of herbal medicines. Metabarcoding—a technique that allows bulk DNA extraction, amplification and high throughput sequencing of multiple taxa—offers one exciting pathway moving forward (Hajibabaei et al. 2011). The ACDB is looking towards leveraging this potential. In collaboration with various research institutions, including McGill University, Canada, and Imperial College, London, UK, current research efforts are using metabarcoding methods to explore whether invasive species can be identified from environmental DNA present in ballast water, and to characterize the species composition of ecological communities through the analysis of environmental DNA in watering holes.

Draft

5. Summary

The global DNA barcoding campaign has grown tremendously over the last decade, resulting in a rapid accumulation of DNA sequence data for numerous lineages. The ACDB has been part of this movement, contributing over 24,000 records, across plants and animals. Work by the ACDB has contributed to a better understanding of the systematics of several problematic and economically important groups in southern Africa. In addition, we at the ACDB have pioneered the use of DNA barcode data in ecology and conservation across the African continent. Key studies have examined

the regional assembly of the southern African flora, the drivers of plant community structure in Savanna, and patterns of diversification and alien species invasion. Nonetheless, major challenges need to be addressed in the future, these include training the next generation of taxonomists (a scarce skills field), extending the network of formal and informal collaborations to research institutes and NGOs across Africa, and extending current efforts to understudied taxonomic groups such as beetles, flies, and disease causing organisms or pest species. New advances will provide the potential to further extend the applications of DNA barcode approaches. For example, novel metagenomic methods allow for better analyses of degraded environmental samples, improving the detection of adulteration in herbal medicines and dietary supplements, and track wildlife crime. At the ACDB, we are looking to pioneer the application of novel barcoding approaches in plants, including the sequencing of whole plastid genome. These new data will allow us to address questions beyond those related to taxonomy and systematics, which have traditionally been core to DNA barcoding, and help establish DNA barcoding as a key tool in environmental research and forensic science, and the ACDB as a global leader in the field.

Draft

6. Acknowledgements

We thank the Government of Canada through Genome Canada and the Ontario Genomics Institute (2008-OGI-ICI-03), the International Development Research Centre (Canada), the Consortium for the Barcode of Life (CBOL) through the Google's Global Impact Award Programme, the University of Johannesburg Analytical Facility (South Africa), South African National Research Foundation, the Royal Society (UK), the Invasive Species Programme of the South African National

Biodiversity Institute and the Working for Water Programme, Natural Resource Management Branch, Department of Environmental Affairs (South Africa), Toyota for financial support, and various local and international authorities who granted plant collection permits. Lastly, we would like to thank three anonymous reviewers for comments on an earlier draft of this manuscript.

Figure Legend

Figure 1: African countries with formal collaborations with the ACBD and field collection sites.

Figure 2: Fauna and flora field data collection efforts at the ACDB.

Figure 3: Major vegetation types where collection efforts at the ACDB have been concentrated. (A) Olifants River, Kruger National Park, Mpumalanga/Limpopo, South Africa. (B) Royal Natal National Park, Drakensberg, KwaZulu-Natal, South Africa. (C) Table Mountain, Table Mountain National Park, Western Cape, South Africa. (D) Lake Sibaya, ISimangaliso Wetland Park, KwaZulu-Natal, South Africa. (E) West Coast National Park, Western Cape, South Africa. (F) Kunene River, Epupa Falls, Kunene, Namibia. (G) Gorongosa limestone gorges, Gorongosa National Park, Sofala Province, Mozambique. (H) Zambesian flooded grasslands, Western Province, Zambia. Photographs: A-H by Olivier Maurin.

Figure 4: Representation of taxonomic coverage of barcoding efforts at the ACDB following Raupach and Radulovici 2015.

Figure 5: Taxonomic and ecological studies published by the ACDB since its inception.

Draft

Table Legend

Table 1: Number of taxonomic records representing orders, families, genera and species submitted by ACDB to the BOLD database.

	Orders	Families	Genera	Linnaean species (specimens barcoded)
Plants	60	239	1 924	7 452 (21 102)
Animals	7	27	444	521 (3 751)

5. References

Adamowicz, S.J. 2015. International Barcode of Life: Evolution of a global research community. Genome. **58**(5): 151-162, doi:10.1139/gen-2015-0094.

APG IV: The Angiosperm Phylogeny Group. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. Bot. J. Linn. Soc. **181**(1): 1–20.

Bello, A., Daru, B.H., Stirton, C.H., Chimphango, S.B.M., Van der Bank, M., Maurin, O., and Muasya, A.M. 2015. DNA barcodes reveal microevolutionary signals in fire response trait in two legume genera. AoB PLANTS. **7**(27): plv124. doi:10.1093/aobpla/plv124.

Bezeng, S.B., Davies, J., Yessoufou, K., Maurin, O., and Van der Bank, M. 2015. Revisiting Darwin's naturalization conundrum: explaining invasion success of non-native trees and shrubs in southern Africa. J. Ecol. **103**(4): 871–879. doi:10.1111/1365-2745.12410.

Boatwright, J.S., Maurin, O., and Van der Bank, M. 2015. Phylogenetic position of Madagascan *Acacia* s.l. species and new combinations in *Senegalia* and *Vachellia*

(Leguminosae, Mimosoideae, Acacieae). Bot. J. Linn. Soc. **179**(2): 288–294. doi: 10.1111/boj.12320.

Burgess, K.S., Fazekas, A.J., Kesanakurti, P.R., Graham, S.W., Husband, B.C., Newmaster, S.G., Percy, D.M., Hajibabaei, M., and Barrett, S.C.H. 2011. Discriminating plant species in a local temperate flora using the *rbc*L plus *mat*K DNA barcode. Methods Ecol. Evol. **2**(4): 333–340. doi:10.1111/j.2041-210X.2011.00092.x.

Cavender-Bares, J., Kozak, K.H., Fine, P.V.A. and Kembel, S.W. 2009. The merging patterns of community ecology and phylogenetic biology. Ecol Lett. **12**: 693–715.

CBOL Plant Working Group. 2009. Botanists Recommend a DNA Barcode for Land Plants. Proc. Natl. Acad. Sci. USA. **31**(106): 12794–12797. doi:10.1073/pnas.0905845106.

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. USA. **113**: E5572–E5579. doi: 10.1073/pnas.1607493113.

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. doi:10.1186/1471-2148-12-73.

Condamine, F.L., Nagalingum, N.S., Marshall, C.R., and Morlon, H. 2015. Origin and diversification of living cycads: a cautionary tale on the impact of the branching process prior in Bayesian molecular dating. BMC Evol. Biol. **15**: 65. doi:10.1186/s12862-015-0347-8.

Draft

Daehler, C.C. 2001. Darwin's naturalization hypothesis revisited. Am. Nat. **158**: 324–330.

Daru, B.H., Manning, J.C., Boatwright, J.S., Maurin, O., Maclean, N., Schaefer, H., Kuzmina, M., and Van der Bank, M. 2013. Molecular and morphological analysis of subfamily Alooideae (Asphodelaceae) and the inclusion of *Chortolirion* in Aloe. Taxon. **62**(1): 62–76.

Daru, B.H., Van der Bank, M., and Davies, T.J. 2015. Spatial incongruence among hotspots and complementary areas of tree diversity in southern Africa. Divers. Distrib. **21**(7): 769–780. doi:10.1111/ddi.12290.

Daru, B.H., Van der Bank, M., Maurin, O., Yessoufou, K., Schaefer, H., Slingsby, J.A., and Davies, T.J. 2016. A novel phylogenetic regionalization of phytogeographical zones of southern Africa reveals their hidden evolutionary affinities. J. Biogeogr. **43**(1): 155–166. doi:10.1111/jbi.12619.

Daru, B.H. and Yessoufou, K. 2016. A search for a single DNA barcode for seagrasses of the world. *DNA Barcoding in Marine Perspectives: Assessment and Conservation of Biodiversity* (ed. by S. Trivedi, A.A. Ansari, K.S. Ghosh, H. Rehman), pp. 313-330. Springer International Publishing Switzerland.

Davies, T.J., Daru, B.H., Van der Bank, M., Maurin, O., and Bond, W.J. 2016. Multiple routes underground? Frost alone cannot explain the evolution of underground trees. New Phytol. **209**(3): 910–912. doi:10.1111/nph.13795.

De Vos, L., and Skelton, P. 1990. Name changes for two common African catfishes. Rehabilitation of *Schilbe intermedius* Rüppel, 1832 (Siluriformes, Schilbeidae). Cybium. **14** (4): 323–326.

Diez, J.M., Sullivan, J.J., Hulme, P.E., Edwards, G., and Duncan, R.P. 2008. Darwin's naturalization conundrum: dissecting taxonomic patterns of species invasions. Ecol Lett. **11**: 674–681.

Ehleringer, J.R., Cerling, T.E., and Helliker, B.R. 1997. C_4 photosynthesis, atmospheric CO_2, and climate. Oecol. **112**(3): 285–299.

Franzini, P.Z.N., Dippenaar-Schoeman, A.S., Yessoufou, K., and Van der Bank, F.H. 2013. Combined analyses of genetic and morphological data indicate more than one species of *Cyrtophora* (Araneae: Araneidae) in South Africa. Int. J. Mod. Biol. Res. **1**(3): 21–34.

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

Hajibabaei, M., Shokralla, S., Zhou, X., Singer, G.A.C., and Baird, D.J. 2011. Environmental Barcoding: A Next-Generation Sequencing Approach for Biomonitoring Applications Using River Benthos. PLoS ONE. **6**(4): e17497. doi:10.1371/journal.pone.0017497

Hajibabaei, M., DeWaard, J.R., Ivanova, N.V., Ratnasingham, S., Dooh, R.T., Kirk, S.L., Mackie, P.M., and Hebert, P.D.N. 2005. Critical factors for assembling a high volume of DNA barcodes. Philos. Trans. R. Soc. Lond. B, Biol. Sci. **360**: 1959-1967.

Hajibabaei, M., Janzen, D.H., Burns, J.M., Hallwachs, W., and Hebert, P.D.N. 2006. DNA barcodes distinguish species of tropical Lepidoptera. Proc. Natl. Acad. Sci. USA. **103**(4): 968–971. doi:10.1073/pnas.0510466103.

Draft

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

Handley, L.L. 2015. How will the 'molecular revolution' contribute to biological recording? Bot. J. Linn. Soc. **115**: 750– 766.

Hollingsworth, P.M., Li, D.Z., van der Bank, M., and Twyford, A.D. 2016. Telling plant species apart with DNA: from barcodes to genomes. Phil. Trans. R. Soc. B. **371**(1702), 20150338.

Hopkins, G.W., and Freckleton, R.P. 2002. Declines in the numbers of amateur and professional taxonomists: implications for conservation. Anim. Conserv. **5**(3): 245–249. doi:10.1017/S1367943002002299.

Hoveka, L.N., Van der Bank, M., Boatwright, J.S., Bezeng, B.S., and Yessoufou, K. 2016. The noncoding *psbA-trnH* spacer, as an effective DNA barcode for aquatic freshwater plants, reveals prohibited invasive species in aquarium trade in South Africa. S. Afr. J. Bot. **102**: 208–216.

Joly, S, Davies, T. J., Archambault, A. Bruneau, A, Derry, A., Kembel, S. W. Peres-Neto, P., Vamosi, J. and Wheeler, T. A. 2013. Ecology in the age of DNA barcoding: the resource, the promise, and the challenges ahead. Mol Ecol Resour. doi: 1111/1755-0998.12173.

Jordann, M., Van Wyk, A.E., and Maurin, O. 2011a. A conspectus of Combretum (Combretaceae) in southern Africa, with taxonomic and nomenclatural notes on species and sections. *Bothalia* **41(**1): 135–160.

Jordann, M., Van Wyk, A.E., and Maurin, O. 2011b. Generic status of Quisqualis (Combretaceae), with notes on the taxonomy and distribution of Q. parviflora. *Bothalia* **41**(1): 161–169.

Draft

Kalwij, J.M., de Boer, W.F., Mucina, L., Prins, H.H.T., Skarpe, C., and Winterbach, C. 2010. Tree cover and biomass increase in a southern African savanna despite growing elephant population. Ecol Appl. **20**: 222–233.

Kartzinel, T.R., Chen, P.A., Coverdale, T.C., Erickson, D.L., Kress, W.J., Kuzmina, M.L., Rubenstein, D.I., Wang, W., and Pringle, R.M. 2015. DNA metabarcoding illuminates dietary niche partitioning by African large herbivores. Proc. Natl. Acad. Sci. USA. **112**: 8019-8024.

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

Kyalangalilwa, B., Boatwright, J.S., Daru, B.H., Maurin, O., and Van der Bank M. 2013. Phylogenetic position and revised classification of the African *Acacia* species (Fabaceae: Mimosoideae). Bot. J. Linn. Soc. **172**(4): 500–523. doi:10.1111/boj.12047.

Lahaye, R., Van der Bank, M., Bogarin, D., Warner, J., Pupulin, F., Gigot, G., Maurin, O., Duthoit, S., Barraclough, T.G., and Savolainen, V. 2008. DNA barcoding the floras of biodiversity hotspots. Proc. Natl. Acad. Sci. USA. **105**(8): 2923–2928. doi:10.1073/pnas.0709936105.

Lekganyane, D.M., Van der Bank, H., Van der Bank, M. 2015. Survey of traditional medicinal bulbs traded at a street market. Barcoding Bulletin. **6**(3): 15-16.

Mack, R.N., Simberloff, D., Lonsdale, W.M., Evans, H., Clout, M., and Bazzaz, F.A. 2000. Biotic invasion: causes, epidemiology, global consequences, and control. Ecol. Appl. **10**(3): 689–710.

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

Draft

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 phylogenetic classification of Asphodelaceae subfamily Alooideae: towards a resolution of the prickly issue of polyphyly in *Haworthia.* Syst. Bot. **39**(1): 55–74. doi:http://dx.doi.org/10.1600/036364414X678044.

Maurin, O., Davies, T.J., Burrows, J.E., Daru, B.H., Yessoufou, K., Muasya, A.M., Van der Bank, M., and Bond, W.J. 2014. Savanna fire and the origins of the "underground forests" of Africa. New Phytol. **204**(1): 201–214. doi:10.1111/nph.12936.

Maurin, O., Chase, M.W., Jordaan, M., and Van der Bank M. 2010. Phylogenetic relationships of Combretaceae inferred from nuclear and plastid DNA sequence data: implications for generic classification. Bot. J. Linn. Soc. **162**(3)**:** 453–476. doi: 10.1111/j.1095-8339.2010.01027.x.

Maurin, O., Muasya, A.M., Catalan, P., Shongwe, E.EZ., Viruel, J., Wilkin P., and van der Bank, M. 2016. Diversification into novel habitats in the Africa clade of Dioscorea (Dioscoreaceae): erect habit and elephant's foot tubers. BMC Evol Biol. **16**: 238. doi: 10.1186/s12862-016-0812-z.

Nagalingum, N.S., Marshall, C.R., Quental, T.B., Rai, H.S., Little, D.P., and Mathews, S. 2011. Recent synchronous radiation of a living fossil. Science. **334**(6057): 796–9. doi:10.1126/science.1209926.

National Environmental Management: Biodiversity Act (NEM:BA) (10/2004): Alien and Invasive Species List, 2014.

Pennington, R.T., Richardson, J.E., and Lavin, M. 2006. Insights into the historical construction of species-rich biomes from dated plant phylogenies, neutral ecological

Draft

theory and phylogenetic community structure. New Phytol. **172**(4): 605–616. doi:10.1111/j.1469-8137.2006.01902.x.

Pereira, F., Carneiro, J., and Amorim, A. 2008. Identification of Species with DNA Based Technology: Current Progress and Challenges. Recent Patents on DNA & Gene Sequences **2**(3): 187-200.

Pimentel, D., Zuniga, R., and Morrison, D. 2005. Update on the environmental and economic costs associated with alien invasive species in the United States. Ecolog. Econ. **52**: 273–288. doi:10.1016/j.ecolecon.2004.10.002.

Raupach, M.J., and Radulovici, A.E. 2015. Looking back on a decade of barcoding crustaceans. ZooKeys. **539**: 53–81.

Richardson, D.M., and Van Wilgen, B.W. 2004. Invasive alien plants in South Africa how well do we understand the ecological impact? S. Afr. J. Sci. **100**(1): 45–52.

Sethusa, M.T., Millar, I.M., Yessoufou, K., Jacobs, A., Van Der Bank, M., and Van der Bank, F.H. 2014. DNA barcode efficacy for identification of economic important scale insects (Hemiptera: Coccoidea) in South Africa. Afri. Entomol. **22**(2): 257–266. doi:http://dx.doi.org/10.4001/003.022.0218.

Simon, M.F., Grether, R., de Queiroz, L.P., Skema, C., Pennington, R.T., and Hughes, C.E. 2009. Recent assembly of the Cerrado, a Neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. Proc. Natl. Acad. Sci. USA. **106**(48): 20359–20364. doi:10.1073/pnas.0903410106.

UNEP. 2008. Biodiversity, on the Move to 2010. http://www.unep.org/Themes/Biodiversity

Van der Bank, F.H., and Greenfield, R. 2015. A pioneer survey and DNA barcoding of some commonly found gastropod molluscs on Robben Island. ZooKeys. **481**: 15–23. doi:10.3897/zookeys.481.8188.

Draft

Van der Bank, F.H., Greenfield, R., Daru, B.H., and Yessoufou, K. 2012. DNA barcoding reveals micro-evolutionary changes and river system-level phylogeographic resolution of African silver catfish, *Schilbe intermedius* (Actinopterygii: Siluriformes: Schilbeidae) from seven populations across different African river systems. Acta Ichthyolo. Piscat. **42**(4): 307–320. doi:10.3750/AIP2012.42.4.04.

Van der Bank, F.H., Herbert, D., Greenfield, R., and Yessoufou, K. 2013. Revisiting species delimitation within the genus *Oxystele* using DNA barcoding approach. ZooKeys. **365**: 337–354. doi:10.3897/zookeys.365.5356.

Williams, V.L., Victor, J.E, and Crouch, N.R. 2013. Red Listed medicinal plants of South Africa: Status, trends, and assessment challenges. S. Afr. J. Bot. **86**: 23–35.

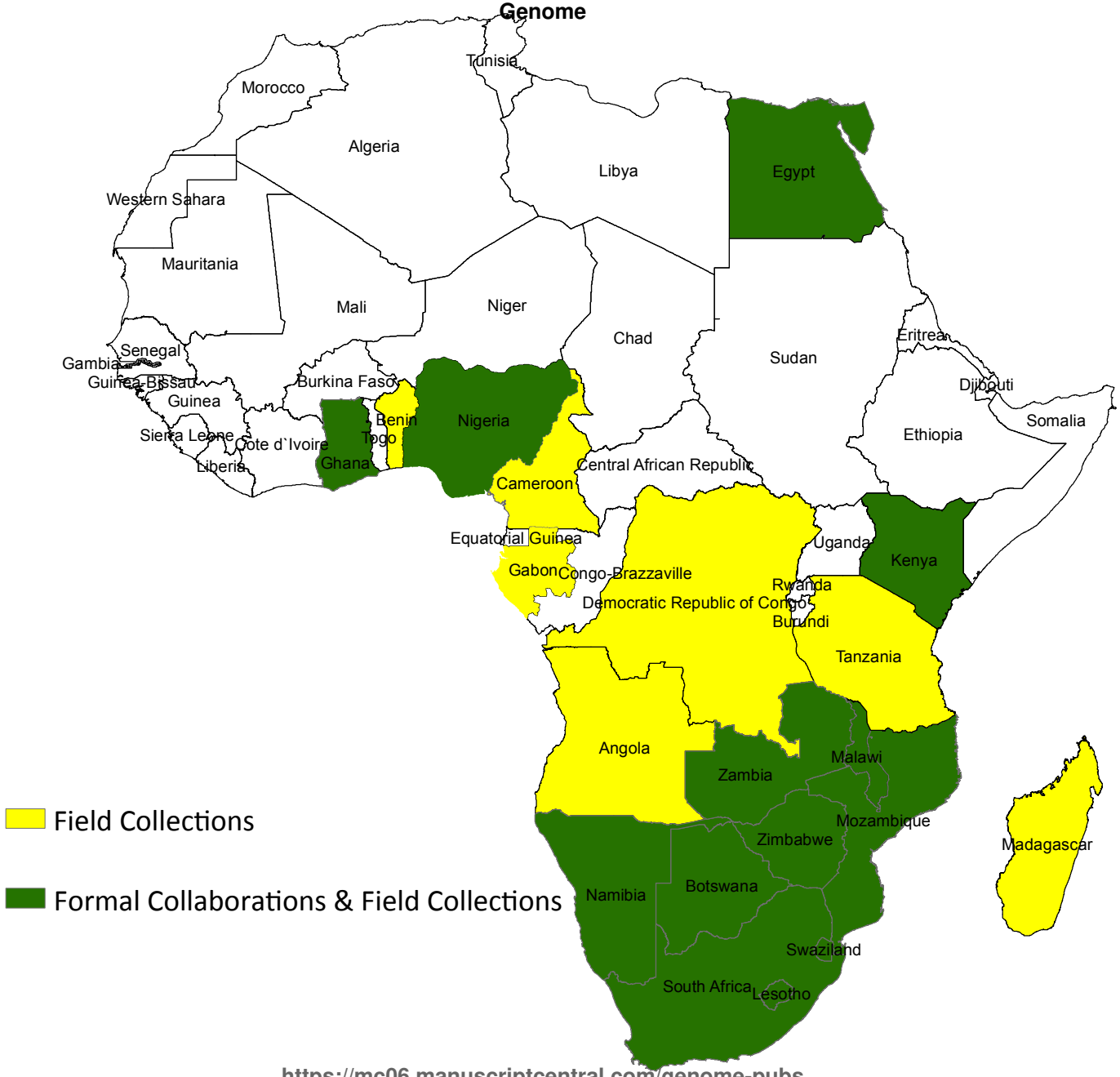
Williamson, J., Maurin, O., Shiba, S., van der Bank, H., Pfab, M., Pilusa, M., Kabongo, R.M. and van der Bank, M. 2016. Exposing the illegal trade in cycad species (Cycadophyta: *Encephalartos*) at two traditional medicine markets in South Africa using DNA barcoding. Genome. **59**: 771-781

Xu, S., Li, D., Li, J., Xiang, X., Jin, W., Huang, W., Jin, H.W., and Huang, L. 2015. Evaluation of the DNA Barcodes in *Dendrobium* (Orchidaceae) from Mainland Asia. PLoS ONE. **10**(1): e0115168.

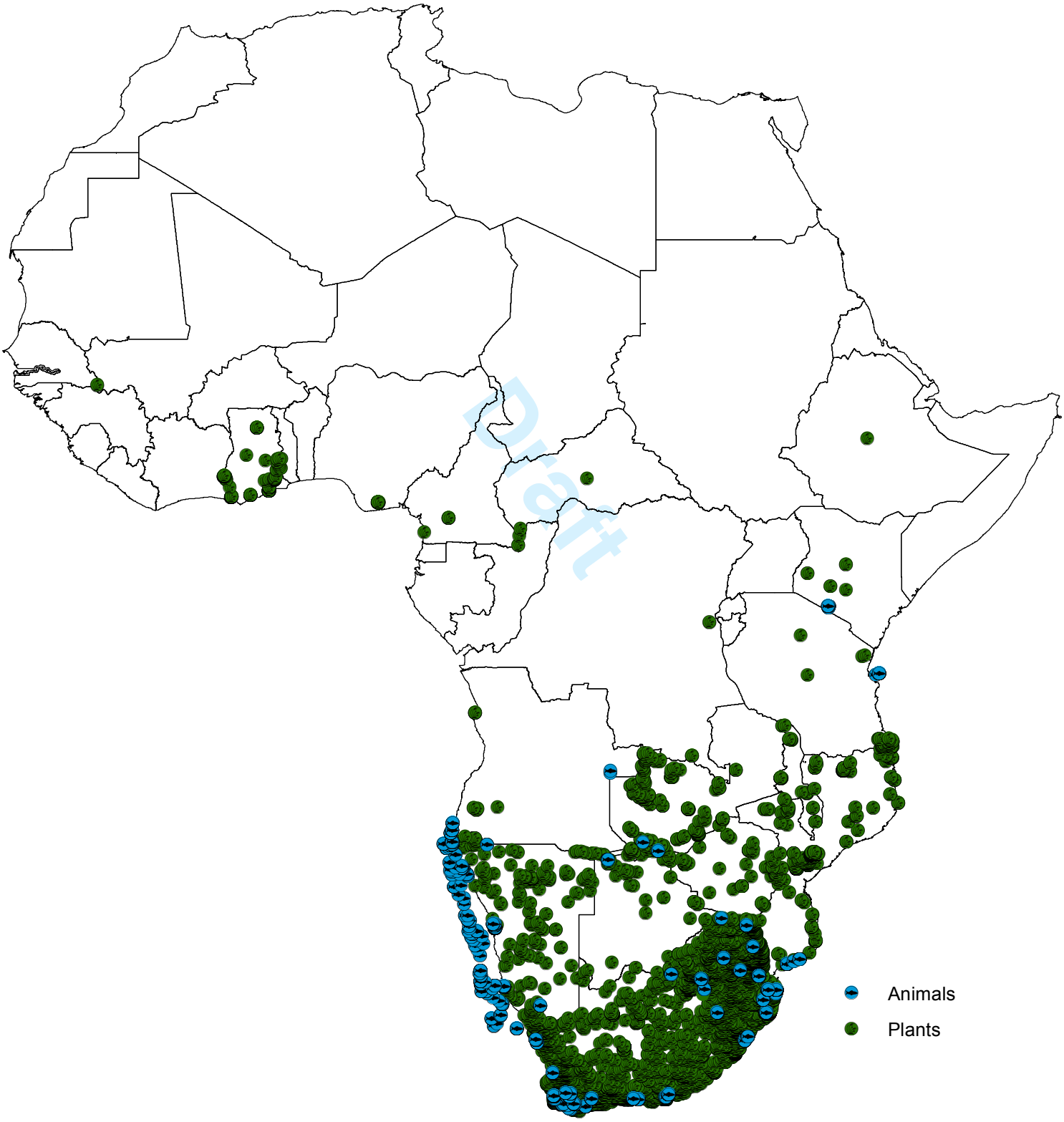
Yessoufou, K., Bamigboye, S.O., Daru, B.H., and Van der Bank, M. 2014. Evidence of constant diversification punctuated by a mass extinction in the African cycad. Ecol. Evol. **4**: 50–58.

Yessoufou, K., Davies, T.J., Maurin, O., Kuzmina, M., Schafer, H., Van der Bank, M., and Savolainen, V. 2013. Large herbivores favour species diversity but have mixed impacts on phylogenetic community structure in African savannah ecosystem. J. Ecol. **101**(3): 614–625. doi:10.1111/1365-2745.12059.

Draft


Tunisia
Morocco
Algeria
Libya
Egypt
Western Sahara
Mauritania
Mali
Niger
Chad
Sudan
Eritrea
Djibouti
Senegal
Gambia
Guinea-Bissau
Guinea
Burkina Faso
Benin
Nigeria
Togo
Ghana
Sierra Leone
Cote d`Ivoire
Liberia
Ethiopia
Somalia
Central African Republic
Cameroon
Equatorial Guinea
Gabon
Congo-Brazzaville
Democratic Republic of Congo
Uganda
Kenya
Rwanda
Burundi
Tanzania
Angola
Zambia
Malawi
Mozambique
Zimbabwe
Madagascar
Namibia
Botswana
Swaziland
South Africa
Lesotho
Field Collections
Formal Collaborations & Field Collections

Collection Effort Across Africa



Draft

A

B

C

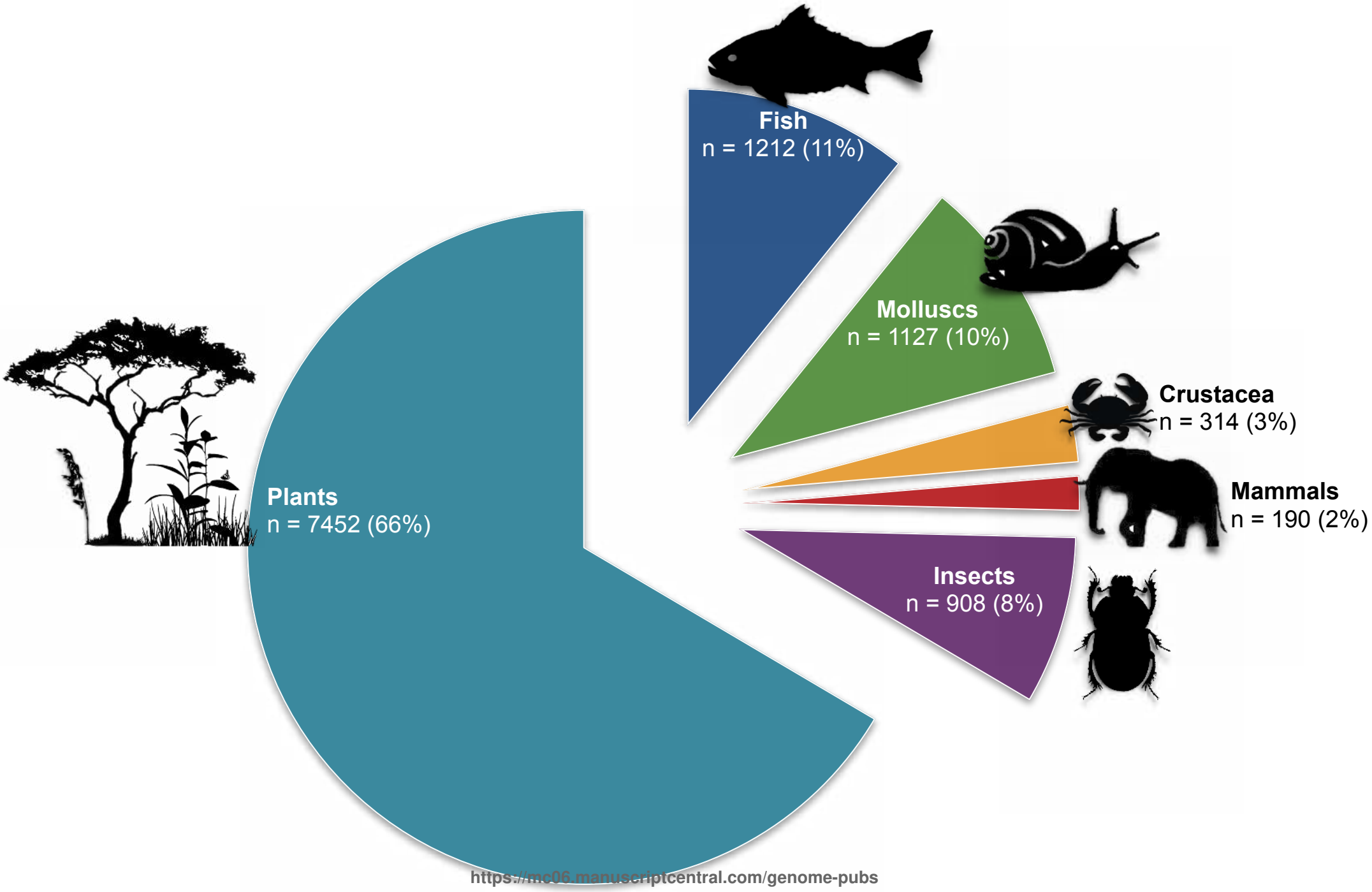
D

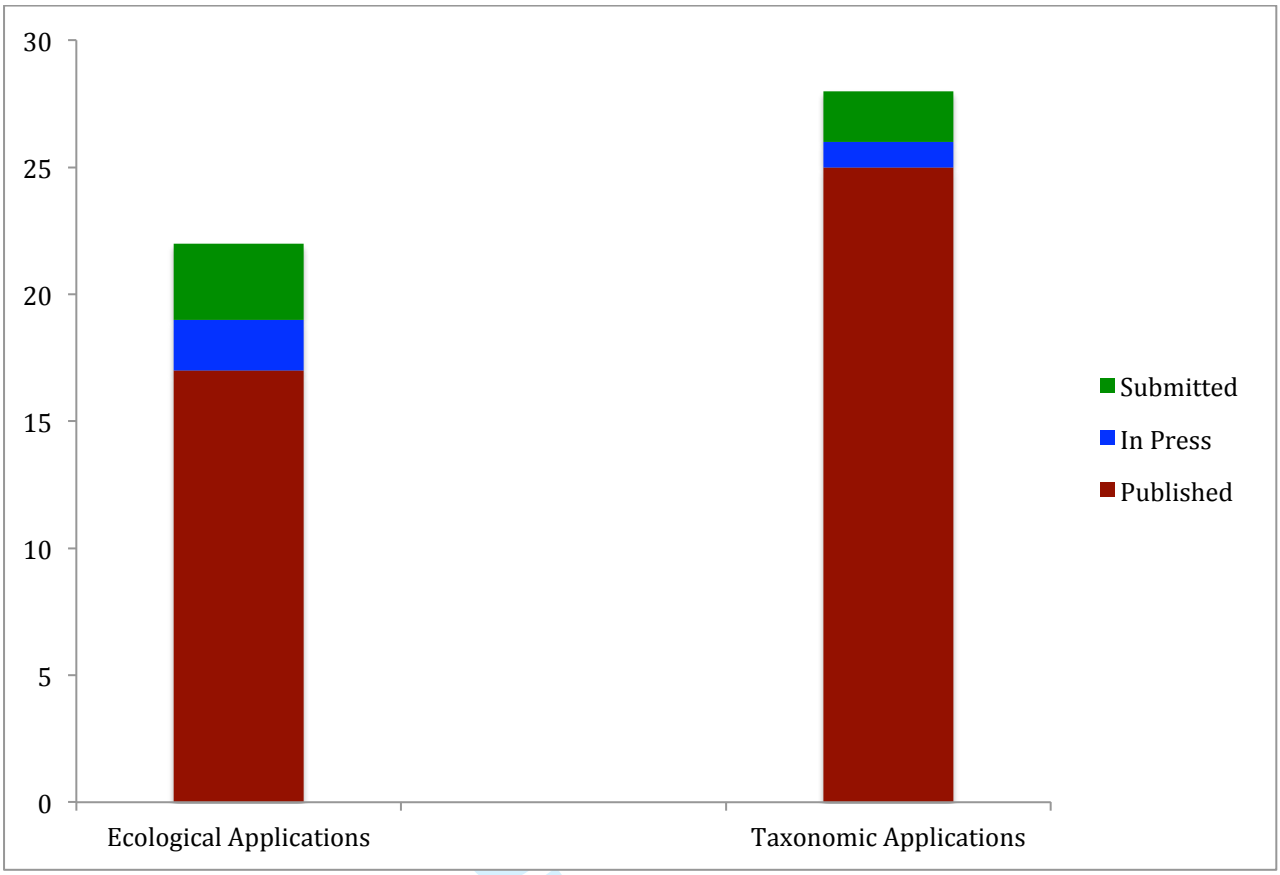
E

F

G

H


Fish
n = 1212 (11%)
Molluscs
n = 1127 (10%)
Crustacea
n = 314 (3%)
Mammals
n = 190 (2%)
Insects
n = 908 (8%)
Plants
n = 7452 (66%)


30
25
20
15
10
5
0
Ecological Applications
Taxonomic Applications
Submitted
In Press
Published

Draft