Biological Journal of the Linnean Society, 2022, **135**, 1–24. With 7 figures.

Taxonomic inflation due to inadequate sampling: are girdled lizards (*Cordylus minor* species complex) from the Great Karoo one and the same?

KRYSTAL A. TOLLEY[1,2,*], NICOLAS S. TELFORD[1], JODY M. TAFT[1,2], MICHAEL F. BATES[3,4], WERNER CONRADIE[5,6], BUYISILE G. MAKHUBO[3,7] and GRAHAM J. ALEXANDER[2]

[1]*South African National Biodiversity Institute, Kirstenbosch Research Centre, Private Bag X7, Claremont 7735, South Africa*
[2]*School of Animal, Plant and Environmental Sciences, University of the Witwatersrand, P.O. Wits 2050, Johannesburg, South Africa*
[3]*Division of Herpetology, Department of Animal and Plant Systematics, National Museum, P.O. Box 266, Bloemfontein 9300, South Africa*
[4]*Department of Zoology & Entomology, University of the Free State, P.O. Box 339, Bloemfontein 9300, South Africa*
[5]*Port Elizabeth Museum (Bayworld), Gqeberha 6013, South Africa*
[6]*Department of Nature Conservation Management, Natural Resource Science and Management Cluster, Faculty of Science, George Campus, Nelson Mandela University, George 6529, South Africa*
[7]*School of Life Sciences, University of KwaZulu-Natal, Private Bag X54001, Durban 4000, South Africa*

Received 9 April 2021; revised 16 July 2021; accepted for publication 22 July 2021

The Great Karoo and Namaqualand of South Africa are home to a species complex of morphologically conserved lizards that occur in allopatry (Karoo: *Cordylus aridus*, *Cordylus cloetei*, *Cordylus minor*; Namaqualand: *Cordylus imkeae*). However, there are negligible morphological differences and a lack of obvious physical or climatic barriers, particularly among the three Karoo species. We hypothesized that poor geographic coverage in previous studies and lack of an explicit species concept has caused taxonomic inflation. We therefore tested species boundaries by examining multiple criteria: multi-gene phylogenetics, niche distribution modelling and re-examination of diagnostic morphological features with a larger sample size. We found that *C. aridus*, *C. cloetei* and *C. minor* lack diagnosable differences for both genetics and morphology. Distribution modelling, ranging from present day to the last interglacial period, show connectivity has been maintained especially during cooler periods. Conversely, *C. imkeae* is morphologically diagnosable, genetically distinct and lacks connectivity with the other taxa. By evaluating multiple operational criteria, we conclude that the *C. minor* species complex comprises only two species, *C. minor* (with *C. aridus* and *C. cloetei* as junior synonyms) and *C. imkeae*, demonstrating that species defined from inadequate data and lack of an explicit species concept can lead to taxonomic inflation.

ADDITIONAL KEYWORDS: Africa – Cordylidae – General Lineage Concept – lizards – reptiles – species – taxonomic inflation.

INTRODUCTION

Modern analytical methods in systematics have revolutionized the way biological diversity is assessed and catalogued, and recently developed techniques have transformed analyses of species boundaries so that their delineation is now more objective (Carstens *et al.*, 2013; Luo *et al.*, 2018). In many cases, reassessment of taxonomy using these techniques has revealed previously hidden diversity, resulting in the recognition of lineages that represent new species and a better assessment of the evolutionary

*Corresponding author. Email: k.tolley@sanbi.org.za

history of member taxa (e.g., Adams *et al.*, 2009; Engelbrecht *et al.*, 2019; Vacher *et al.*, 2020). As with traditional methods, modern analyses are dependent on adequate and geographically dispersed data sets so that genetic differences between samples can be confidently ascribed to either geographic distance effects, or to genetic isolation resulting in species level divergence (Cicero *et al.*, 2021). Geographic gaps from clustered sampling can result in the demarcation of species boundaries where none actually exist, leading to taxonomic inflation (Isaac *et al.*, 2004; Wiemers & Fiedler, 2007). This may be especially prevalent in cases where inadequate sampling erroneously leads to the conclusion that populations are either geographically isolated, or that there is genetic isolation due to falsely perceived barriers. Thus, findings using modern techniques are only as good as the data sets they interrogate; however, recent trends suggest that they are often applied using a formulaic approach with little consideration for the quality of the data set, the biology of the taxa or any underlying species concept (Freitas *et al.*, 2020).

Geographic gaps in sampling tend to be prevalent for species that occur in rugged and remote landscapes where access is limited, and such landscapes occur over much of South Africa. For example, the Great Escarpment (uplifted 180–120 Mya) extends from the interior of the Western Cape Province, eastwards and then northwards from the interior of the Eastern Cape and KwaZulu-Natal provinces, forming the Eastern Escarpment and Drakensberg Mountains that extend into Mpumalanga Province (McCarthy & Rubidge, 2005). To the south, the ancient Cape Fold Mountains (uplifted *c*. 250 Mya) stretch largely parallel to the Great Escarpment. Both mountain ranges include dramatically rugged landscapes that provide habitat for many species of rupicolous lizards some of which are range-restricted endemics. Because parts of these mountains are inaccessible (Fig. 1A), herpetological sampling tends to be patchy, with extensive areas being unsampled (see Telford *et al.*, In press; Supporting Information, Fig. S1). The resulting spatial unevenness of distribution records and the consequent spatial bias of genetic sampling greatly diminishes the rigour of taxonomic assessments of species from the area, and this could result in either under- or over-estimation of diversity.

The Cordylidae are an exclusively African family of lizards, with highest diversity in South Africa where 43 of the 70 recognized species occur (Reissig, 2014). Of the ten genera that make up the family, the most species-rich is *Cordylus* (girdled lizards), and nearly half of the 23 species are endemic to South Africa. Species in the *Cordylus minor* complex are small-bodied, morphologically conserved girdled lizards that occur in the arid, rugged interior of the south-western parts of South Africa (Fig. 2; Supporting Information, Table S1). The most recently published distribution maps (Bates *et al.*, 2014) suggest that species in the complex occur allopatrically. Three of the species (*Cordylus aridus*, *Cordylus cloetei* and *C. minor*) occur in the Great Karoo and along the southern Great Escarpment (Fig. 1B). *Cordylus minor* and *C. cloetei* have been recorded at elevations of 1000–1700 m a.s.l., whereas *C. aridus* has been recorded south of the Great Escarpment at lower elevations of 900–1000 m a.s.l. The fourth member of the complex, *Cordylus imkeae*, occurs about 400 km to the north-west of the other species in an isolated mountainous region of Namaqualand, which is an arid coastal region that extends into Namibia. Closely related congeners, *Cordylus mclachlani* and *Cordylus macropholis* (Stanley *et al.*, 2011), occur at least 100 km and 130 km, respectively, toward the western coastal margin to the south of Namaqualand. A more distantly related congener, *Cordylus cordylus*, is partly sympatric with all these species except for *C. imkeae*.

The four species in the *C. minor* species complex are prime candidates for taxonomic re-evaluation given that they are morphologically difficult to distinguish, and poor sampling in the region may have biased perception of their presumed restricted, allopatric distributions. *Cordylus minor* was originally described as a subspecies of *C. cordylus* based mostly on the presence of a higher number of longitudinal ventral and dorsal scale rows (FitzSimons, 1943) and later elevated to a full species based on a more detailed multivariate analysis (Mouton & van Wyk, 1989). At that time, *C. minor* included an apparently isolated population to the east that was later described as *C. aridus* (Mouton & van Wyk, 1994). Two additional, presumably isolated populations of morphologically similar cordylids were also described as new, namely *C. cloetei* from the Great Escarpment and *C. imkeae* from northern Namaqualand (Mouton & van Wyk, 1994). Despite similarities in their phenotypes (Supporting Information, Table S1), *C. cloetei* was reported to have a larger head (Supporting Information, Table S2), *C. minor* to have an additional supralabial scale, and *C. aridus* to have 28–31 (average 30) transverse rows of transverse dorsal scale rows rather than 26–30 (average 28) in the other species. All other meristic characters examined overlapped between species (e.g., number of suboculars and temporal scale rows). The geographically isolated *C. imkeae* is the only species in this group that showed consistent morphological differences from the other species in the shape of the interparietal, the separation of the parietals by the interparietal and the lower number of suboculars (3 vs. 4). The negligible morphological differences observed between these supposedly allopatric populations were considered sufficient to designate them as full species (Mouton & van Wyk, 1994). Moreover, a genus

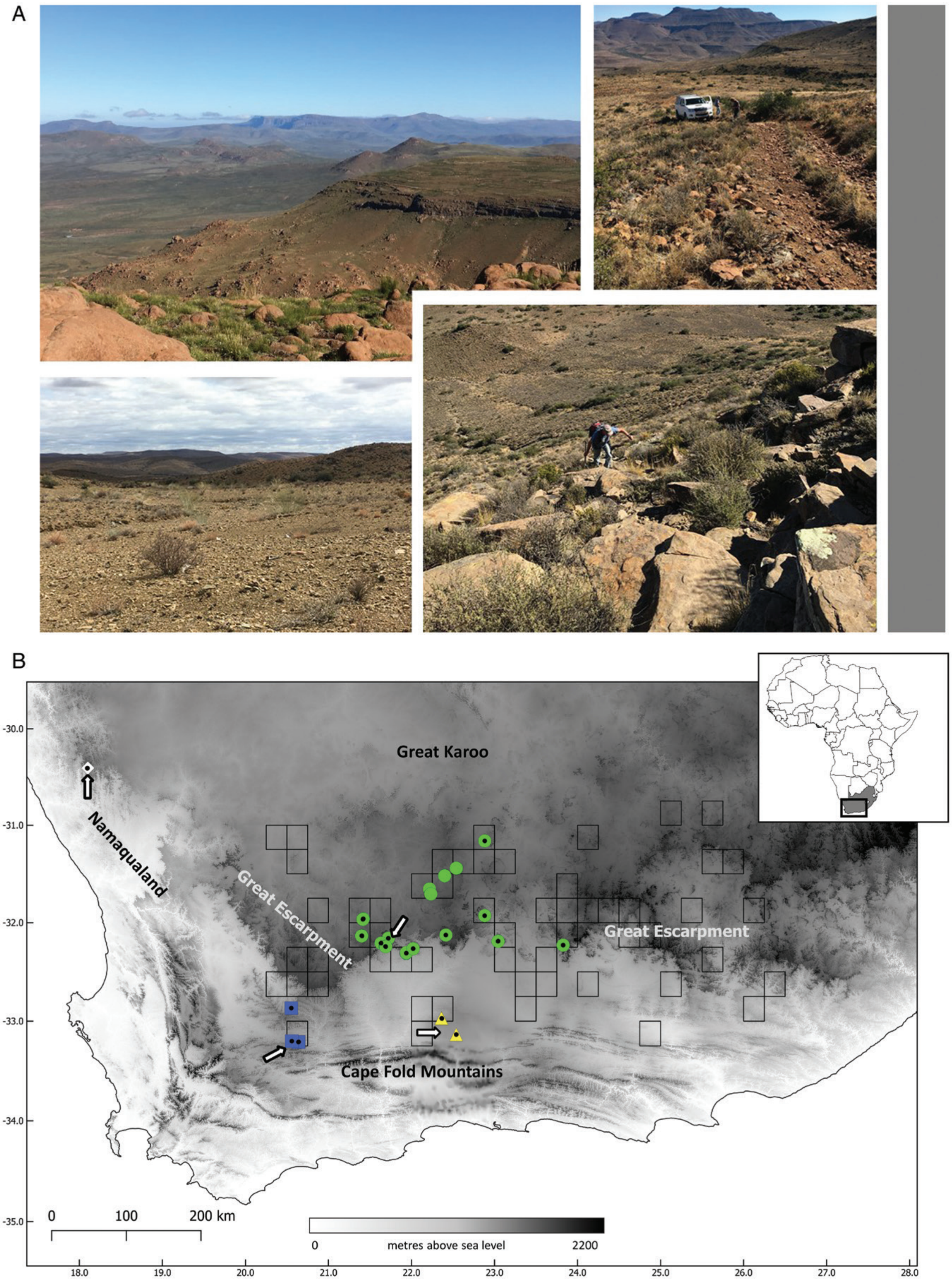

Figure 1. (A) Terrain in the Karoo and the Great Escarpment, South Africa, and (B) Map of the study area with records for taxa in the *Cordylus minor* species complex (*Cordylus aridus* – triangles, *C. cloetei* – circles, *C. imkeae* – diamond, *C. minor* – squares). Symbols with a black dot indicate localities of samples included in the genetic analyses. Recent grid cells surveyed are indicated by squares, and the type localities for each species are indicated by arrows.

Figure 2. Girdled lizards in the *Cordylus minor* species complex from South Africa according to the original taxonomy – (A) *C. aridus* (type locality), (B) *C. cloetei* (near type locality), (C) *C. minor* (type locality), (D) *C. imkeae* (type locality), (E) *C. mclachlani*, (F) *C. cordylus*.

level phylogenetic analysis showed that *C. aridus*, *C. minor* and *C. imkeae* form a monophyletic clade (Stanley *et al.*, 2011). Divergences of approximately 5–12 Myr between pairs of those taxa have been estimated (see Zheng & Wiens, 2016); however, those divergence estimates are in error (see *Material and Methods* below).

Previous phylogenetic work did not include *C. cloetei*–with only a single *C. minor* and two each of *C. aridus* and *C. imkeae* included, with all samples of these latter species each collected from single localities. Thus, the insufficient geographic and taxon sampling in the previous studies, as well as the morphological similarities between the four species has not allowed for a full assessment of the validity of these species. Given the lack of a comprehensive data set, coupled to the lack of obvious geographic barriers, particularly between *C. cloetei* and *C. minor*, it is possible that the rugged terrain along the Great Escarpment allows for connectivity that has been undetected due to poor sampling. Although *C. aridus* is considered isolated south of the Great Escarpment (Mouton & van Wyk, 1994), the landscape is characterized by undulating hills and ridges of suitable habitat that could provide ample connectivity (Fig. 1A, bottom left). Alternatively, if the species have allopatric distributions that have been maintained over time, gene flow would have been absent between the populations and vicariance could have led to speciation, with their phenotypic similarity being the result of morphological conservatism.

Using a balance of evidence approach, we assessed whether the described species in the *C. minor* species complex represent valid species. We applied a General Lineage Concept, whereby species are considered as separately evolving metapopulation lineages (de Queiroz, 1998, 2007) diagnosable by integrating information from a combination of features such as morphology, ecology, genetics/clade monophyly, geographic isolation and reproductive isolation maintained by vicariance and/or mate-recognition (e.g., Paterson, 1985; de Quieroz, 1998; Padial *et al.*, 2010; Cicero *et al.*, 2021). For our work, we focused on assessing morphology, genetic divergence and geographic isolation. We collected new data (tissue samples, voucher specimens and distributional data) from across the region to carry out comprehensive phylogenetic and population level genetic analyses, as well as to enhance the existing data set of morphological features to better assess inter-taxon variation. Furthermore, our augmented locality data set allowed us to carry out species distribution modelling to examine the extent of overlap in climatic space of the species at present day and into the past.

The currently accepted taxonomic hypothesis is that these taxa are cryptic species that are morphologically similar due to niche conservatism but have been reproductively isolated and would have therefore diverged genetically. If this is true, we would expect to find strong genetic divergence and long-term disjunctions in their distributions. An alternative scenario that would support the cryptic species hypothesis is that the taxa have recently entered separate evolutionary trajectories and have diverged in parapatry. This would be expressed by present-day disjunct distributions that were initiated in the recent past. The new disjunctions would have disrupted gene flow causing shallow genetic differences, detectable by lack of haplotype/allele sharing but no pattern of isolation by distance. Furthermore, some morphological differentiation would be expected given selection of the potentially dissimilar niches, coupled with the effect of genetic drift on the phenotype due to local adaptation. If these conditions are not met, then it is likely that the taxa are not cryptic species, but instead, a single species.

MATERIAL AND METHODS

Phylogenetic analyses and species delimitation

We carried out field surveys across the Great Karoo from 2016–2018 to collect locality records, specimens and tissue samples of reptiles, including *Cordylus* species (Fig. 1B; Supporting Information, Fig. S2). Target sites, each covering one pentad (8 × 8 km^2), were chosen in advance (Fig. 1B). Each site was searched by three to four people over a period of 3 days, targeting all habitat types including rocky areas where *Cordylus* might occur. For field identification of specimens, we assessed the diagnostic morphological characters from the original species descriptions (Mouton & van Wyk, 1989, 1994) but found that none of the individuals could be identified to species level based on the diagnostic characters. We therefore assigned a provisional field identification based solely on the proximity to the type locality of each species. We acknowledge that this method of identification is inherently problematic for morphologically similar species because misidentifications will be common and this will lead to inaccurately mapped distributions upon which new identifications are made (e.g., Meier & Dikow, 2004; Stephens *et al.*, In review). However, we chose this approach because we could not otherwise assign a field identification to the specimens.

For new material collected, tissue samples were taken in the form of tail tips for animals that were released and liver from voucher specimens (*c*. 5–10 mg of tail tip or liver). Tissue samples were preserved in 99% ethanol or DMSO/NaCl (N = 38) and voucher specimens (N = 12) were fixed in 10% formalin and transferred to 70% ethanol. Voucher specimens were deposited in the National Museum (NMB) or Port Elizabeth Museum (PEM) (Table 1).

To place the Karoo girdled lizards in a phylogenetic context, we sequenced 38 individuals of *C. aridus*, *C. cloetei*, *C. minor* (under their provisionally assigned identifications) and the congener *C. cordylus* which is broadly sympatric with the Karoo taxa. Additional sequence data for these and other *Cordylus* species were downloaded from GenBank for a total of 76 individuals in the ingroup and five individuals in the outgroup (Table 1). Some GenBank sequences for the *C. minor* species complex were excluded as the sequences were of dubious quality given the presence of internal stop codons and numerous unlikely amino acid changes, or they were a positive match to other species in different genera of the Cordylidae, as assessed by the Blast Local Alignment Search Tool: https://blast.ncbi.nlm.nih.gov/Blast.cgi (see footnotes in Table 1). It should be noted that these sequences have been used in previous phylogenetic studies (i.e., Stanley *et al.*, 2011; Zheng & Wiens, 2016) resulting in inflated, misleading divergence estimates between taxa in Zheng & Wiens (2016).

For new samples, tissues were dried in a vacuum centrifuge prior to DNA extraction. Total genomic DNA was extracted using a salt extraction protocol (Aljanabi & Martinez, 1997). PCR amplification of two mitochondrial genes (*ND2* and 16S) and one nuclear

Table 1. *Cordylus* samples included in this study with museum voucher numbers/field numbers and GenBank accession numbers for three genes. New records of the *C. minor* species complex were assigned provisional identifications that are indicated in the species column in parentheses. Voucher specimens indicated with museum accession numbers: NMB – National Museum, PEM – Port Elizabeth Museum. The subset of samples included in the phylogeny are indicated (Y), while all data were included in the remaining analyses. Locality information is given, with coordinates provided where available. T – material from type locality, NT – material from near type locality (within 15 km). EC – Eastern Cape Province, NC – Northern Cape Province, WC – Western Cape Province, RSA – Republic of South Africa

Genus	Species	Voucher or field no.	16S	*ND2*	*PRLR*	Phylogeny subset	Latitude	Longitude	Locality
Cordylus	(*aridus*) *minor* (NT)	PEM R16377*	HQ167169	HQ166958†	HQ167498	Y	-33.1344	22.5389	Bruinrante, 20 km N of Meringspoort, WC, RSA
Cordylus	(*aridus*) *minor* (NT)	PEM R16376	HQ167170	HQ166959†	HQ167499	Y	-33.1344	22.5389	Bruinrante, 20 km N of Meringspoort, WC, RSA
Cordylus	(*aridus*) *minor* (NT)	PEM R26513 / S152	MZ619094	MZ646166	MZ646199	Y	-32.9715	22.3639	Farm Kleinwaterval, 13 km N of Botterkraal, WC, RSA
Cordylus	(*aridus*) *minor* (NT)	S159	MZ619095	MZ646167	MZ646200	Y	-32.9729	22.3665	Farm Kleinwaterval, 13 km N of Botterkraal, WC, RSA
Cordylus	*beraduccii*	JB6	KT941403	KT941393	KT941390	Y			Tanzania
Cordylus	*beraduccii*	JB7	KT941404	KT941394	KT941391	Y			Tanzania
Cordylus	(*cloetei*) *minor*	FP069	MZ619097	MZ646169	MZ646202		-31.9287	22.8815	near Taaiboschfontein Farm, WC, RSA
Cordylus	(*cloetei*) *minor* (T)	HB337	MZ619099	MZ646171	MZ646204		-32.1586	21.7214	Farm De Hoek, WC, RSA
Cordylus	(*cloetei*) *minor* (T)	HB338	MZ619100	NA	MZ646205		-32.1586	21.7214	Farm De Hoek, WC, RSA
Cordylus	(*cloetei*) *minor*	MFB2011.7.9vi	MZ619101	MZ646172	MZ646206		-32.2651	22.0199	Paalhuisberg, Beaufort West, WC, RSA
Cordylus	(*cloetei*) *minor*	MFB2011.7.9vii	MZ619102	MZ646173	MZ646207		-32.2088	21.6341	Teekloof Pass, Beaufort West, WC, RSA
Cordylus	(*cloetei*) *minor*	NMB R9368	MZ619103	MZ646174	MZ646208	Y	-32.1262	22.414	Farm Klavervlei, Beaufort West, WC, RSA
Cordylus	(*cloetei*) *minor*	NMB R10213	MZ619104	MZ646175	MZ646209	Y	-32.2651	22.0199	Paalhuisberg, Beaufort West, WC, RSA
Cordylus	(*cloetei*) *minor*	NMB R10214	MZ619105	MZ646176	MZ646210		-32.2649	22.0204	Paalhuisberg, Beaufort West, WC, RSA
Cordylus	(*cloetei*) *minor*	NMB R10215	MZ619106	MZ646177	MZ646211		-32.2649	22.0204	Paalhuisberg, Beaufort West, WC, RSA
Cordylus	(*cloetei*) *minor*	NMB R10219	MZ619107	MZ646178	MZ646214	Y	-32.2088	21.6341	Teekloof Pass, Beaufort West, WC, RSA
Cordylus	(*cloetei*) *minor*	NMB R10220	MZ619108	MZ646179	MZ646212		-32.2091	21.6338	Teekloof Pass, Beaufort West, WC, RSA
Cordylus	(*cloetei*) *minor*	NMB R10221	MZ619109	MZ646180	MZ646213		-32.2088	21.6338	Teekloof Pass, Beaufort West, WC, RSA

Table 1. Continued

Genus	Species	Voucher or field no.	16S	*ND2*	*PRLR*	Phylogeny subset	Latitude	Longitude	Locality
Cordylus	*(cloetei) minor*	NMB R11596 / FP062	MZ619096	MZ646168	MZ646201		-31.929	22.8814	near Taaiboschfontein Farm, WC, RSA
Cordylus	*(cloetei) minor*	NMB R11598 / FP085	MZ619098	MZ646170	MZ646203		-32.3108	21.9377	Tafelkop, NC, RSA
Cordylus	*(cloetei) minor*	NMB R11600 / S483	MZ619121	NA	MZ646224		-31.961	21.4192	Droogvoetsfontein, Fraserburg, NC, RSA
Cordylus	*(cloetei) minor*	P120	MZ619110	MZ646181	MZ646215		-31.1614	22.8825	Kwekwa Farm, Victoria West, NC, RSA
Cordylus	*(cloetei) minor*	P122	MZ619111	MZ646182	MZ646216		-31.1613	22.8825	Kwekwa Farm, Victoria West, NC, RSA
Cordylus	*(cloetei) minor*	P228	MZ619112	MZ646183	MZ646217		-32.1352	21.4045	Eselfontein Farm, NC, RSA
Cordylus	*(cloetei) minor*	P243	MZ619113	MZ646184	NA		-32.1337	21.40138	Eselfontein Farm, NC, RSA
Cordylus	*(cloetei) minor*	P327	MZ619114	MZ646185	MZ646218		-32.2421	21.6875	Muggefontein Farm, WC, RSA
Cordylus	*(cloetei) minor*	P348	MZ619115	MZ646186	MZ646219		-32.2442	21.6894	Muggefontein Farm, WC, RSA
Cordylus	*(cloetei) minor*	P732A	MZ619116	MZ646187	MZ646220		-32.2294	23.8267	Waalplaats Farm, Aberdeen, EC, RSA
Cordylus	*(cloetei) minor*	P733	MZ619117	NA	MZ646221		-32.2294	23.8267	Waalplaats Farm, Aberdeen, EC, RSA
Cordylus	*(cloetei) minor*	S326	MZ619118	MZ646188	MZ646222		-32.1881	23.0441	Farm Kamferskraal, Beaufort West, WC, RSA
Cordylus	*(cloetei) minor*	S329	MZ619119	NA	NA		-32.1897	23.0427	Farm Kamferskraal, Beau-fort West, WC, RSA
Cordylus	*(cloetei) minor*	S481	MZ619120	MZ646189	MZ646223		-31.961	21.4192	Droogvoetsfontein, Fraserburg, NC, RSA
Cordylus	*(cloetei) minor*	S508	MZ619122	NA	MZ646225		-31.9656	21.4149	Droogvoetsfontein, Fraserburg, NC, RSA
Cordylus	*(cloetei) minor*	S518	MZ619123	MZ646190	MZ646226		-31.9658	21.4151	Droogvoetsfontein, Fraserburg, NC, RSA
Cordylus	*cordylus*	NMB R9302	MZ619124	MZ646191	MZ646227	Y	-31.4448	26.6937	Droogefontein, Wodehouse, EC, RSA
Cordylus	*cordylus*	NMB R9313	MZ619125	MZ646192	MZ646228	Y	-32.5757	26.9398	Amatole Mountains, Cathcart, EC, RSA
Cordylus	*cordylus*	NMB R9523	MZ619126	MZ646193	MZ646229	Y	-32.2318	22.4588	Beaufort West, WC, RSA
Cordylus	*cordylus*	P653	MZ619127	MZ646194	MZ646230	Y	-31.9213	24.1012	Toon Botha's Hoek, WC, RSA
Cordylus	*cordylus*	P714	MZ619128	MZ646195	MZ646231	Y	-31.8784	24.1016	Wintershoek Farm, Middelberg Peak, WC, RSA
Cordylus	*cordylus*	PEM R15012	HQ167189	HQ166977	HQ167518	Y	-33.7992	25.7694	St. Croix Island, EC, RSA

Table 1. Continued

Genus	Species	Voucher or field no.	16S	*ND2*	*PRLR*	Phylogeny subset	Latitude	Longitude	Locality
Cordylus	*cordylus*	PEM R17394	HQ167232	HQ167015	HQ167561	Y	-33.6357	25.5562	Grassyridge, Coega, EC, RSA
Cordylus	*imkeae*	NMB R11303	HQ167197	HQ166985	HQ167526	Y	-30.4044	18.1017	Rooiberg Mountain, NC, RSA
Cordylus	*imkeae*	NMB R11304	HQ167198	HQ166986	HQ167527	Y	-30.4044	18.1017	Rooiberg Mountain, NC, RSA
Cordylus	*jonesii*	AMB8310	HQ167199	HQ166987	HQ167528	Y	-22.688	29.521	N of Soutpansberg, Limpopo Province, RSA
Cordylus	*jonesii*	AMB8396	HQ167200	HQ166988	HQ167529	Y	-24.055	28.404	W of Mokopane, Limpopo Province, RSA
Cordylus	*machadoi*	PEM R18007	KT941145	KT941259	KT941310	Y	-14.9619	13.335	Humpata, Huíla Province, Angola
Cordylus	*machadoi*	KTH09-059	KT941146	KT941260	KT941311	Y	-14.9619	13.335	Humpata, Huíla Province, Angola
Cordylus	*machadoi*	PEM R18008	KT941144	KT941258	KT941309	Y	-14.9619	13.335	Humpata, Huíla Province, Angola
Cordylus	*macropholis*	AMB8873	HQ167206	HQ166993	HQ167535	Y	-32.11	18.3036	Lamberts' Bay, WC, RSA
Cordylus	*macropholis*	AMB8874	HQ167207	HQ166994	HQ167536	Y	-32.1103	18.3039	Lambert's Bay, WC, RSA
Cordylus	*mclachlani*	AMB8855	HQ167208	HQ166995	HQ167537	Y	-33.2722	19.6283	NW of Ceres, WC, RSA
Cordylus	*mclachlani*	CmclSU1	HQ167209	HQ166996	HQ167538	Y	-32.2017	19.0978	Cederberg Mountains, WC, RSA
Cordylus	*merculae*	PEM R16165	HQ167211	NA	HQ167540	Y	-12.053	37.6469	Summit Serra Mecula, Nissa Province, Mozambique
Cordylus	*merculae*	PEM R16202	HQ167233	NA	HQ167562	Y	-12.0461	37.6225	Western slopes Serra Mecula, Nissa Province, Mozambique
Cordylus	*merculae*	PEM R16203	HQ167234	NA	HQ167563	Y	-12.0461	37.6225	Western slopes Serra Mecula, Nissa Province, Mozambique
Cordylus	*minor*	CminSU	HQ167212	HQ166997[§]	NA[‡]		-32.868	20.5528	40 km N of Matjiesfontein, WC, RSA
Cordylus	*minor* (NT)	P373	MZ619129	MZ646196	MZ646232	Y	-33.2089	20.6417	6 km W of Matjiesfontein, WC, RSA
Cordylus	*minor* (NT)	P377	MZ619130	MZ646197	MZ646233	Y	-33.2099	20.6389	5 km W of Matjiesfontein, WC, RSA
Cordylus	*minor* (NT)	PEM R19654 / SVN-661	MZ619131	MZ646198	MZ646234	Y	-33.2011	20.5586	3.5 km NE of Matjiesfontein, WC, RSA
Cordylus	*niger*	CnigSU1	HQ167217	HQ167000	HQ167546	Y	-32.9844	17.8769	Saldanha, WC, RSA
Cordylus	*niger*	Saldanha1	AY519748	AY519688	NA	Y	-33.0208	17.9492	Saldanha, WC, RSA
Cordylus	*niger*	Saldanha2	AY519747	AY519689	NA	Y	-33.0208	17.9492	Saldanha, WC, RSA
Cordylus	*oelofseni*	AMB8860	HQ167221	HQ167004	HQ167550	Y	-32.7697	18.7028	Piketberg, WC, RSA

Table 1. Continued

Genus	Species	Voucher or field no.	16S	*ND2*	*PRLR*	Phylogeny subset	Latitude	Longitude	Locality
Cordylus	*oelofseni*	CoelSU1	HQ167219	HQ167002	HQ167548	Y	-34.04	18.9983	Hottentots-Holland Mountains, WC, RSA
Cordylus	*oelofseni*	CoelSU2	HQ167220	HQ167003	HQ167549		-34.04	18.9983	Hottentots-Holland Mountains, WC, RSA
Cordylus	*oelofseni*	AMB8851	HQ167218	HQ167001	HQ167547		-32.9094	19.035	Grootwinterhoek Mountains, WC, RSA
Cordylus	*rhodesianus*	ELSPET4	HQ167230	HQ167013	HQ167559	Y			Unknown (captive)
Cordylus	*rhodesianus*	ELSPET5	HQ167231	HQ167014	HQ167560	Y			Unknown (captive)
Cordylus	*tropidosternum*	JB8	KT941405	KT941397	KT941392	Y			Tanzania
Cordylus	*tropidosternum*	WRB0038	HQ167236	NA	HQ167565	Y			Tanzania
Cordylus	*tropidosternum*	WRB0042	HQ167235	KT941398	HQ167564	Y			Tanzania
Cordylus	*ukingensis*	PET1	JQ389808	NA	NA	Y			Tanzania
Cordylus	*ukingensis*	WRB0039	HQ167237	KT941399	HQ167566	Y	-8.133	35.679	Tanzania
Cordylus	*vittifer*	AMB8274	HQ167242	HQ167020	HQ167571	Y			Limpopo Province, RSA
Cordylus	*vittifer*	PEM R17561 / AMB8603	HQ167243	HQ167021	HQ167572	Y	-25.3031	30.1475	De Berg Pass, N of Dullstroom, Mpumalanga Province, RSA
Outgroup									
Hemicordylus	*capensis*	AMB8859	HQ167255	HQ167033	HQ167584	Y			WC, RSA
Hemicordylus	*capensis*	HcapSU1 (AMB8857)	HQ167254	HQ167032	HQ167583	Y			WC, RSA
Hemicordylus	*capensis*	PEM R16378	HQ167253	HQ167031	HQ167582	Y	-33.4161	22.6922	Blesberg Peak, Eastern Swartberg Mountains, WC, RSA
Hemicordylus	*nebulosus*	HnebSU1	HQ167262	HQ167040	HQ167591	Y			WC, RSA
Hemicordylus	*nebulosus*	HnebSU2	HQ167261	HQ167039	HQ167590	Y			WC, RSA

*Museum accession number provided here is correct, whereas the incorrect number (PEM R16371) is attached to the GenBank record from Stanley *et al*. (2011).
†GenBank sequence for two samples of *C. aridus* have internal stop codons. Where included in analyses, the sequences were trimmed to 220 bp.
‡GenBank sequence for *C. minor* (HQ167541) not included because the sequence matches >99% similarity to *Ninurta coeruleopunctatus*.
§GenBank sequence for *C. minor* (HQ166997) not included because of quality issues and the presence of internal stop codons.

gene (*PRLR*) were carried out using the following primer sets–*ND2*: L4437 and H5540 (Macey *et al.*, 1997); 16S: 16Sa and 16Sb (Palumbi *et al.*, 1991); *PRLR*: F1 and R3 (Townsend *et al.*, 2008). An initial denaturation step was carried out for 4 min at 95 °C, followed by 35 cycles of denaturation (94 °C, 45 s), annealing (51–58 °C, 45 s) and extension (72 °C, 1 min). This was followed by a final extension at 72 °C for 10 min. PCR products were quantified by electrophoresis on a 1% agarose gel. Sanger sequencing was carried out at Macrogen (Amsterdam, Netherlands) using the forward primers for each marker. Sequences were aligned in Geneious v.11 (Kearse *et al.*, 2012).

A genus-level phylogeny was run for a subset of individuals of two to four individuals per species from our target taxa (Table 1). The analysis also included multiple representatives of other *Cordylus* species, overall covering 17 of the 23 species in the genus, plus two species of *Hemicordylus* that were included as outgroup taxa (Table 1). Bayesian inference and maximum likelihood (ML) analyses were run on the combined data set of 1919 characters with a total of 53 individuals. The Bayesian analysis was run using MrBayes v.3.2.6 (Huelsenbeck & Ronquist, 2001) at the Cyberinfrastructure for Phylogenetic Research (CIPRES) Science Gateway v.3.3 (Miller *et al.*, 2010). Data were partitioned by gene with 939 bp for *ND2*, 507 bp for 16S and 475 bp for *PRLR* (10 bases were excluded from the 16S alignment due to ambiguous alignment of hypervariable regions). jModelTest (Guindon & Gascuel, 2003; Darriba *et al.*, 2012) was used to assess the evolutionary model that best fitted each of the partitions using the Akaike information criterion (AIC) test, and this was incorporated into the Bayesian analysis (16S: nst = 6+G+I; *ND2*: nst = 6+G; *PRLR*: nst = 6+G). The Markov Chain Monte Carlo (MCMC) was run for 20 million generations with a burnin of 10%. Tracer v.1.7 (Rambaut *et al.*, 2018) was used to verify that the effective sample size (ESS) was above 200 for all parameters. A ML analysis was run using RAxML (Stamatakis, 2014) through the CIPRES portal. The data set was partitioned by gene, applying the GTR+I+G model for each partition with 1000 bootstrap replicates.

We used several approaches to investigate species boundaries. Firstly, a distance-based ‘barcoding’ approach was used, whereby pairwise sequence divergences for the combined mitochondrial markers (16S and *ND2*, 920 bp) were used to generate frequency distributions of intra- and interspecific sequence divergence using SpeciesIdentifier v.1.8 (Meier *et al.*, 2006). With the barcoding approach, intraspecific and interspecific divergence values should not overlap (e.g., the ‘barcode gap’), because genetic divergences should be low within species, but high between species. The threshold between intra- and interspecific comparisons is therefore a rough starting point for species delimitation (Lefébure *et al.*, 2006). For this analysis, each individual must be pre-assigned to a species. Therefore, we avoided taxonomic bias from our own species assignments of the study taxa by generating the intra- and interspecific frequency distributions from an input data set that included all *Cordylus* species except our four study taxa. Additionally, sequence divergences between *Cordylus* species were estimated using uncorrected net p-distances separately for each gene and for the combined mitochondrial genes using MEGA v.7 (Kumar *et al.*, 2016). This allowed for a comparison of the interspecific sequence divergence values between species in the genus, including the study taxa, which could then be compared to the frequency distributions generated by SpeciesIdentifier. Nineteen base pairs of the hypervariable region of 16S were excluded from the analysis, and any other missing data were excluded pairwise.

To examine haplotype/allele sharing among taxa, we generated TCS haplotype networks (Clement *et al.*, 2000) for each gene using PopArt v.1.7 (Leigh & Bryant, 2015). The networks included *C. aridus*, *C. cloetei*, *C. imkeae*, *C. minor* and for comparative purposes, the closely related species, *C. mclachlani*. Some individuals were however excluded from the networks due to short sequences and due to quality issues the two *C. aridus* sequences from GenBank were excluded (16S, $N = 38$; *ND2*, $N = 33$; *PRLR*, $N = 37$).

Species delimitation was also examined with a Bayesian general mixed Yule-coalescent model (bGMYC) in R using the package bGMYC v.3.0.1 (Reid & Carstens, 2012; R Core Team, 2013). This method accounts for error in phylogenetic estimation and model parameters by integrating the uncertainty in tree topology and branch lengths, accounting for the number of substitutions along branches between speciation events to identify the point (e.g., node) where the branching shifts from a Yule to a coalescent process. The bGMYC was run using the set of gene trees from (1) the two loci (three genes) data set and (2) a single locus data set composed of the two mitochondrial genes generated in BEAST v.2.5 (Bouckaert *et al.*, 2014). The latter analysis was carried out given that the bGMYC analysis is best suited to a single locus. To run BEAST, xml files were created in BEAUTi v.2, setting up unlinked partitions (one for each gene), a linked relaxed-clock and a Yule speciation model. Partition model priors were guided from model selection using jModelTest (Guindon & Gascuel, 2003; Darriba *et al.*, 2012) and evolutionary rates along branches followed an uncorrelated lognormal distribution. The analysis was run for 100 million generations at the CIPRES Science Gateway v.3.3 (Miller *et al.*, 2010), saving trees every 5000 generations to produce a set of 20 000 trees followed by a 50% burnin. The log files were checked in Tracer (Rambaut *et al.*, 2018) to examine tree likelihood and parameter estimates for evidence mixing and convergence, evaluated by effective sample size (ESS) greater than 200 (post-burnin). TreeAnnotator v.2.1.2 (Bouckaert *et al.*, 2014) was used to produce a maximum clade credibility tree for the set of post-burnin trees, setting the posterior probability limit to 0.5.

For the bGMYC species delimitation, 1000 randomly sampled trees from the post-burnin posterior distribution of the sets of ultrametric trees were used. Each data set was run for 1 million generations, sampling every 1000 generations, with a 10% burnin. A heatmap of groupings of the terminal tips in the phylogeny was produced, with probabilities ≥ 0.90 considered supported as conspecific (Reid & Carstens, 2012). The efficiency of the analysis was checked by the distribution of ratios (and the log ratios) of coalescence to speciation events to ensure that these ratios were above 0, as this would indicate that the frequency of speciation events is higher than the divergences at a population level (Reid & Carstens, 2012).

To examine whether genetic distance could be explained by geographic distance between sample localities, rather than by taxa that have isolated allopatric distributions, an analysis of isolation by distance (IBD) for *C. aridus*, *C. cloetei* and *C. minor* was run for the combined mitochondrial genes (N = 29). The IBD analysis was run in Alleles in Space (Miller, 2005) using genetic and geographic distance between all pairs of individuals with input data consisting of DNA sequences and the coordinates of the collection localities. This analysis does not allow for missing data, so the *ND2* sequences were trimmed to 220 bp to match a shorter portion of the GenBank sequences of *C. aridus* that we considered reliable after scrutinizing those sequences for quality (see footnotes in Table 1). This allowed us to retain all individuals of *C. aridus* in the analysis, albeit with a shorter gene fragment. The resulting scatterplot of genetic and geographic distance was interpreted in light of a larger data set (N = 33) which included individuals of the closely related *C. imkeae* and *C. macropholis* to compare the influence of interspecific divergence on IBD patterns.

Species distribution modelling

Occurrence records used in distribution models were gathered from the Karoo surveys in addition to existing records (Supporting Information, Fig. S2). Because there were only two unique data points for *C. imkeae*, it was excluded from the modelling. The analysis was run under two different scenarios: (1) a three taxa analysis with data points assigned to one of three species (*C. aridus*, *C. cloetei* or *C. minor*) based on the original museum records and from our provisional species assignments and (2) a single taxon analysis with data points assigned based on synonymy of the Karoo taxa (*C. aridus*, *C. cloetei* and *C. minor* as a single species). To reduce spatial autocorrelation, records were spatially rarefied to a distance of 5 km using the package spThin (Aiello-Lammens *et al.*, 2015) run in R. This resulted in a total of 68 unique data points.

Nineteen bioclimatic variables were downloaded from www.worldclim.org at a 30 sec and 2.5 arc min resolution. A terrain ruggedness index map was created using the package raster in R with the WorldClim v.2.1 30 sec elevation layer (Riley *et al.*, 1999; Fick & Hijmans, 2017). To reduce the effects of collinearity, a Pearson's correlation coefficient test was performed on all environmental variables using the package ENMTools in R (Warren *et al.*, 2010). Variables that had an r ≥ 0.75 were inspected and variables considered important for the distributions of the reptiles were retained. The remaining variables were BIO1 – annual mean temperature; BIO2 – mean diurnal temperature range; BIO3 – isothermality; BIO6 – minimum temperature of coldest month; BIO12 – annual precipitation; BIO19 – precipitation of coldest quarter; and terrain ruggedness index.

Species distribution modelling was carried out using the maximum entropy approach in Maxent v.3.3.3 (Phillips *et al.*, 2006), as it performs better than other approaches when using a low number of occurrence localities (Elith *et al.*, 2006, 2011). The parameter settings used when constructing distribution models have significant effects on model outcomes, therefore species-specific tuning is recommended to improve model performance (Elith *et al.*, 2011). ENMeval was used to construct models with different parameter settings and perform model evaluation to identify the most optimum model (Muscarella *et al.*, 2014). Models were built with different combinations of the linear (L), quadratic (Q), hinge (H), product (P) and threshold (T) feature classes (LQHPT; LQHP; LQH; L; LQ; H) and varying the regularization multiplier (0 to 4.5 with 0.5 increments). Data were partitioned into testing and training bins using the 'jack-knife' method since this is the recommended method with sample sizes smaller than 25 (Muscarella *et al.*, 2014). To account for spatial sampling bias, 10 000 background points were randomly selected across the study area (Phillips *et al.*, 2006).

Optimal model parameters were selected using a variety of criteria. The Akaike Information Criterion (AIC) corrected for small sample sizes was first considered. The model with the lowest AIC value indicates a balance between the best goodness of fit and complexity (Warren & Seifert, 2011). The threshold-independent metric (AUC), difference between test and training AUC (AUCdiff), minimum training presence omission rate (ORmtp) and the training omission rate (OR10) were also inspected to ensure that the models were not overfitting (Anderson & Gonzalez, 2011). Variable contributions in the optimum model were inspected, and the most important variables were noted as per the permutation importance (Phillips *et al.*, 2006).

Fluctuations in climate are known to affect species distributions (Rosenzweig *et al.*, 2008; Ikeda *et al.*, 2016), with the most recent large-scale climatic shifts being after the Last Interglacial (120 Kya), the Last Glacial Maximum (21 Kya) and subsequent changes during the mid-Holocene (6 Kya). Therefore, projected distributions at these time-slices were modelled using palaeoclimate environmental variables downloaded from WorldClim [Palaeoclimate Modelling Intercomparison Project Phase II (PMIP2): Braconnot *et al.* (2007)], derived from the general circulation models (GCMs; CCSM-4 and MPI-ESM-P: Hijmans *et al.*, 2005) based on CMIP5 (Taylor *et al.*, 2012) data. These data are widely used when constructing palaeoclimate models incorporating climate cycles (e.g., Brown & Knowles, 2012). Suitable climate during the palaeoclimate time-slices were predicted by projecting the reduced set of bioclimatic variables from the optimized current climate model. For all models, a 10% training presence logistic threshold was used when identifying suitable and non-suitable habitat.

Morphology

Newly-collected material and additional voucher specimens are in the collections of the National Museum, Bloemfontein (NMB), Port Elizabeth Museum, Gqeberha (PEM), South African Museum, Cape Town (SAM) and Ditsong National Museum of Natural History, Pretoria (TM) (Supporting Information, Appendix S1). Morphological features that have been used as diagnostic characters to discriminate species of the *C. minor* species complex (Mouton & van Wyk, 1989, 1994; see also Supporting Information, Materials and Methods) were examined and assessed. All type specimens of the four species in the complex were examined, excluding only the holotype of *C. minor* (all paratypes examined), as was new material collected during our surveys, and additional museum specimens (N = 62, Supporting Information, Appendix S1). For comparison, we included morphological data for *C. cordylus* (N = 20), a congener that is partly sympatric with the *C. minor* complex.

Specimens were examined under stereo-microscopes for scalation and morphometrics following Mouton & van Wyk (1994). Scale characters examined were the numbers of supralabials, suboculars, transverse rows of temporal scales, chin-shields in contact with first pair of sublabials, dorsal scale rows longitudinally and transversely, ventral scale rows longitudinally and transversely, subdigital lamellae of 4th toe, femoral pores, differentiated/glandular femoral scales (additional details in the Supporting Information, Materials and Methods). Measurements were taken using digital vernier callipers for snout-to-vent length, head length, head width and head depth (additional details in the Supporting Information, Materials and Methods).

RESULTS

Phylogenetic analyses and species delimitation

Each of the phylogenetic analyses resulted in the same topology with *C. aridus*, *C. cloetei* and *C. minor* forming a well-supported clade that is sister to *C. imkeae* (Fig. 3). The majority of described *Cordylus* species included in the analysis were supported, in agreement with the existing comprehensive Cordylidae phylogeny (Stanley *et al.*, 2011).

The frequency distribution of sequence divergences showed no overlap between intra- and interspecific comparisons for the combined mitochondrial genes, with the transition from intra- to interspecific values (barcoding gap) around 5%. Comparisons between *C. aridus*, *C. cloetei* and *C. minor* based on their preliminary identifications were within the intraspecific range (Fig. 4; Supporting Information, Table S3a). The values for *C. imkeae* were several times greater than the comparisons between the three Karoo species, falling in the interspecific range.

The networks show a clustering of haplotypes for *C. aridus*, *C. cloetei* and *C. minor* for the two mitochondrial genes as compared to *C. imkeae* and *C. mclachlani* which are both separated by many additional mutational steps (Fig. 5). There is some haplotype sharing between *C. aridus* and *C. cloetei*, with *C. minor* generally being separated by additional mutational steps. This could possibly suggest greater historical connectivity between *C. aridus* and *C. cloetei* than with *C. minor*. The network for the nuclear gene showed allele sharing between all three Karoo taxa, but with distinct alleles for *C. imkeae* and *C. mclachlani* (Fig. 5).

The bGMYC analysis using the three-gene data set supported most described species. *C. cloetei*, *C. aridus* and *C. minor* were supported as a single taxon at ≥ 0.9 probability (Fig. 3; Supporting Information, Fig. S3), which is considered strong support (Reid & Carstens, 2012). Most other described species were supported by this analysis although *C. cordylus* and *C. oelofseni* were both sub-divided. In contrast, the mitochondrial-only data set grouped several described species as a single taxon, including *C. cloetei*, *C. aridus*, *C. minor* and *C. imkeae* and *C. cordylus* with *C. oelofseni* (Fig. 3; Supporting Information, Fig. S4). Given that most species outside of the *C. minor* clade were represented by only two individuals, model performance could have been an issue (Reid & Carstons, 2012). Therefore, these results, particularly the three-gene analysis, were used to guide our interpretations, rather than being an unequivocal finding.

There was significant isolation by distance (r = 0.60, $P < 0.001$) within *C. aridus*, *C. cloetei* and *C. minor* (Fig. 6A). This indicates that genetic distance between these taxa can be explained by increasing geographic distance rather than barriers to gene flow as would be expected between species. Comparatively, intraspecific pairwise comparisons of genetic and geographic distance that included *C. imkeae* and *C. macropholis* are substantially higher and do not fit the pattern of isolation by distance (Fig. 6B). This suggests that there is some barrier to gene flow between *C. imkeae*, *C. macropholis* and the three Karoo taxa, but not among the three Karoo taxa (i.e., *C. aridus, C. cloetei* and *C. minor*).

Species distribution modelling

Contributions of the original variables to the models differed slightly according to scenario and spatial resolution (Supporting Information, Table S4), although overall the most important variables were terrain ruggedness, annual mean temperature (Bio 1), mean diurnal temperature range (Bio 2) and precipitation of coldest quarter (Bio 19; see also Supporting Information, Fig. S5A–E). Model performance was good, with evaluation metrics of the optimum models for each scenario meeting the expected thresholds (i.e., ΔAIC values were zero and most OR_{mtp} values were < 0.1, Supporting Information, Table S5). Slightly elevated values for *C. minor* suggest the model could be marginally overfitting possibly due to the few occurrence points for this taxon.

Species distribution modelling for the single taxon and the three taxa scenarios suggest there is some degree of connectivity at the present day between the ranges of *C. aridus*, *C. cloetei* and *C. minor* (Fig. 7; Supporting Information, Fig. S6). The three taxa occupy different areas, but the inferred ranges based on the models are not isolated or disjunct (Fig. 7). Similarly, the single taxon scenario does not demonstrate any potential disjunctions in the range (Supporting Information, Fig. S6). Although *C. imkeae* was not included in the modelling, the single taxon model shows the area where *C. imkeae* occurs as suitable for the Karoo taxa during most periods (Supporting Information, Fig. S6), and likely points to similarity of the environmental niche between the Karoo taxa and *C. imkeae*.

The models suggest that connectivity was much greater during the Mid-Holocene and the Last Glacial Maximum (LGM). In contrast, the last interglacial period shows a similar pattern to present day, with connectivity maintained but patchier as compared to the mid-Holocene and LGM. In contrast, the models do not demonstrate connectivity between *C. imkeae* and other members of the *C. minor* species complex at any time period, and this could suggest a persistent lack of connectivity of *C. imkeae* with the other species throughout Plio-Pleistocene glacial-interglacial cycling over the last *c.* 2.6 Myr.

Morphology

The widespread congener *C. cordylus* can be separated from the *C. minor* group by its larger size (max. recorded SVL 81.7 mm vs. 70.5 mm), lower number of longitudinal dorsal (16–20 vs. 20–26) and longitudinal ventral scales (12–14 vs. 13–16) and flattened infranasals, but all other values for characters examined are overlapping (Table 2; Supporting Information, Results; Table S6).

Morphological features examined were overlapping for *C. aridus*, *C. cloetei* and *C. minor* (Table 2; Supporting Information, Results, Table S6). Although Mouton & van Wyk (1994) separated *C. minor* from the other three species on the basis of usually having six (vs. five) supralabials on either side of the head, posterior parietals smaller than anterior ones and the frequent occurrence (vs. absence) of a post-interparietal scale, our expanded data set did not support their observations. Notably, the number of supralabials varies (usually 5–6) although the posterior parietals are often the smallest in *C. aridus* and usually equal in size to the anterior ones in *C. cloetei*. A post-interparietal scale is occasionally present in *C. aridus* and *C. cloetei*. Mouton & van Wyk (1994) reported that the infranasals were slightly protruding in *C. cloetei* vs. flattened in *C. aridus*. However, we found that the character is variable and the flattened and protruding state is present in similar proportions of individuals of *C. aridus* and *C. cloetei*. *Cordylus imkeae* is differentiated from the other species by almost always having two chin shields in contact with the anterior pair of sublabials (vs. 1–2), distinct postnasals that are larger than the nostril, as many as 17 differentiated/glandular femoral scales (vs. 6–8) and usually only three suboculars

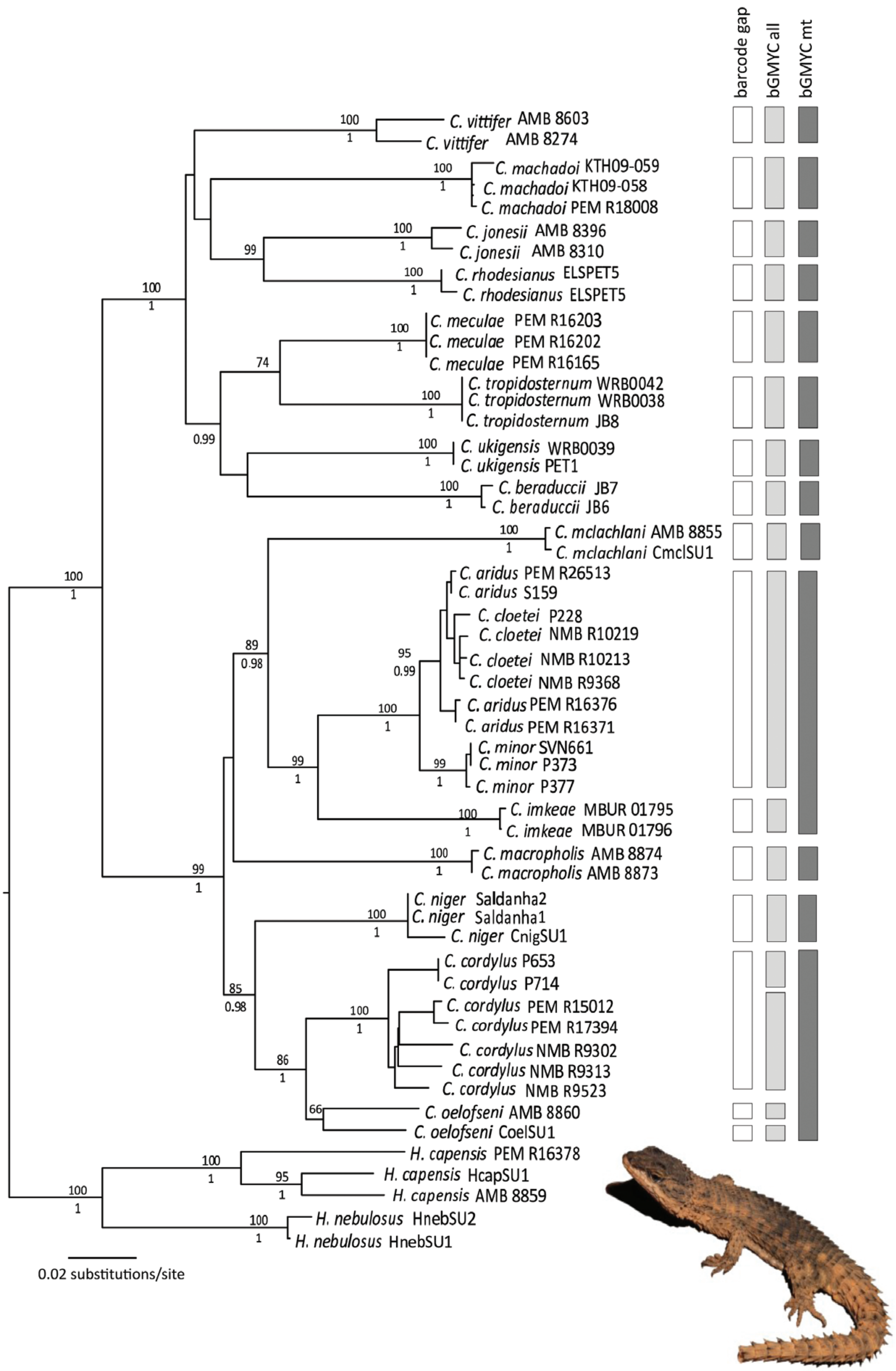

Figure 3. Maximum likelihood consensus tree for *Cordylus* with bootstrap values (top) and Bayesian posterior probabilities (bottom). Support values not shown for intraspecific nodes or for nodes with ≤ 0.90 pp/70% bootstrap. Species delimitation groupings are indicated by the bars for the barcoding analysis, bGMYC for the three-gene (all), and the mitochondrial only (mt) analyses.

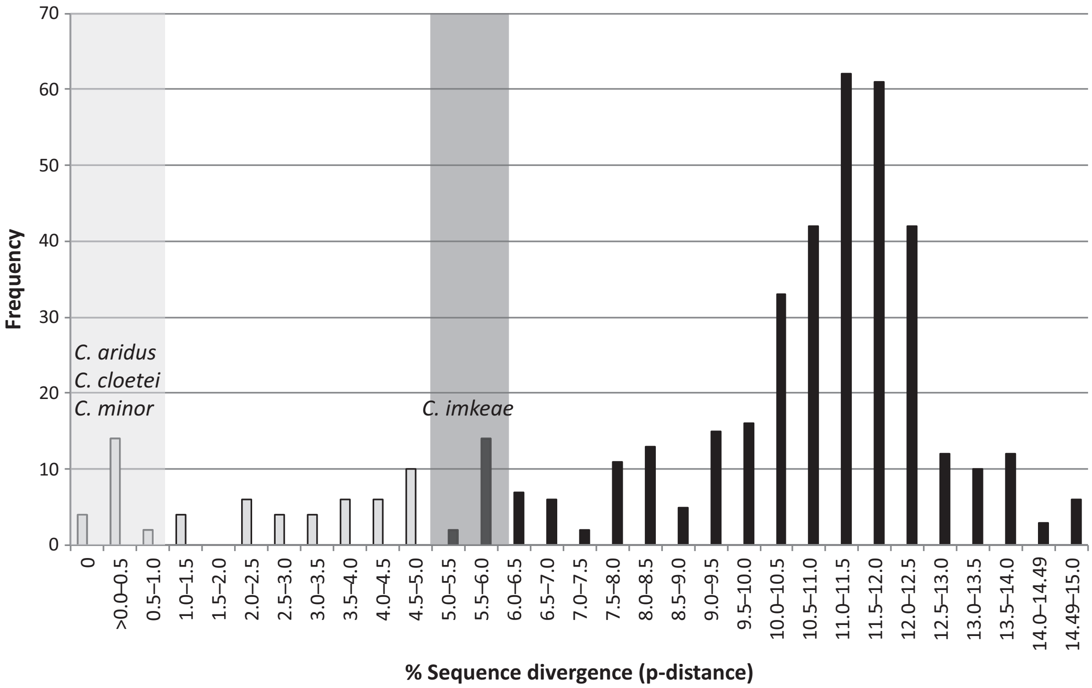

Figure 4. Frequencies of pairwise sequence divergence values for *Cordylus*. Interspecific distances are shown by the black bars, intraspecific distances are shown by grey bars. The range of pairwise sequence divergence values that were estimated for the study taxa (Supporting Information, Table S3) are shown by the grey shading. Light grey shading shows the range for *C. aridus*, *C. cloetei* and *C. minor* and dark grey shading shows the range for *C. imkeae*.

(vs. 4), although the ranges narrowly overlap with the other species (Table 2).

DISCUSSION

Through the application of several complementary data sets and approaches, there is broad agreement that the three Karoo taxa (*C. aridus*, *C. cloetei* and *C. minor*) represent a single taxon, separate to *C. imkeae* from Namaqualand. For the Karoo taxa, our findings refute the hypothesis of three cryptic species as these taxa do not meet the necessary conditions. The Karoo taxa do not have any diagnosable morphological differences between them. They also do not appear to be reproductively isolated as they share alleles and haplotypes, suggesting gene flow has not been disrupted. While this could be the result of shared ancestral polymorphism and a lack of sufficient time for these alleles/haplotypes to have shifted frequency, the species distribution modelling shows connectivity, even over differing (palaeo) environmental conditions. Thus, the lack of physical or environmental barriers over time has allowed gene flow to be maintained. In contrast, *C. imkeae* meets the requirements to be considered a divergent, but phenotypically cryptic species. Although it is morphologically very similar, it does show some diagnosable differences. It also appears to be reproductively isolated, probably due to a long-term environmental barrier that has caused vicariance, disrupting gene flow causing genetic differentiation through genetic drift and/or selection.

The inference that the Karoo taxa (*C. aridus*, *C. cloetei*, *C. minor*) are a single species can be justified through the integration of several lines of evidence. Firstly, there are no diagnostic morphological differences between the taxa. Secondly, sequence divergence is shallow between these taxa and is lower than expected between species. The networks show haplotype and allele sharing, or separation by only a few mutational steps, in contrast to clear separation by multiple mutational steps for *C. imkeae*. Overall, the genetic diversity within the Karoo taxa is best explained by isolation by distance with greater genetic distance between individuals as geographic distance increases. In contrast, genetic distances between other conspecifics are several times greater than between the Karoo taxa, and the maximum genetic distance within the three Karoo taxa does not exceed the interspecific threshold.

The Bayesian species delimitation analysis also supports the synonymy of the three Karoo taxa with strong support. Although Bayesian multispecies coalescent methods such as bGMYC are prone to

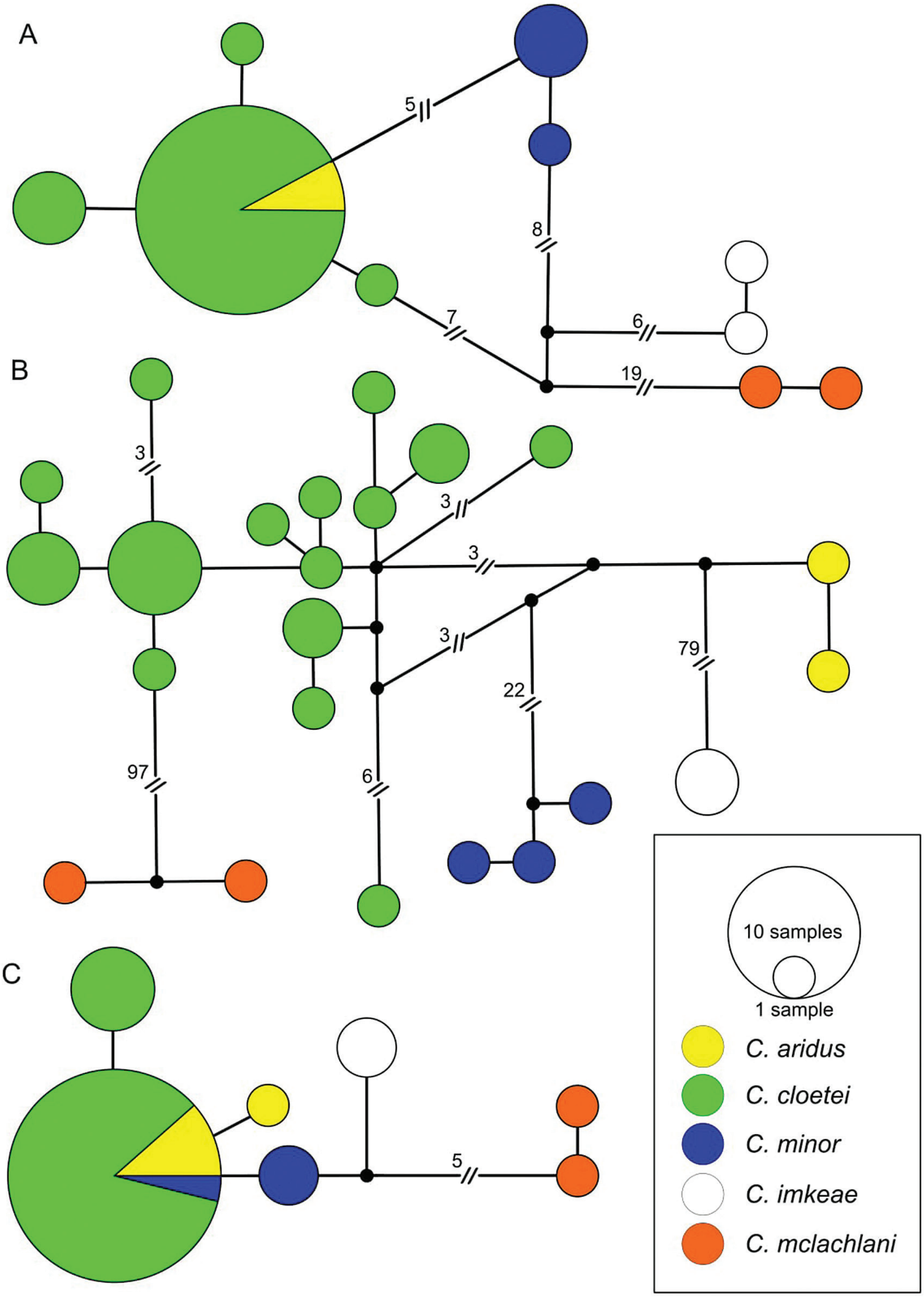

Figure 5. Network of (A) 16S, (B) *ND2* and (C) *PRLR* for the *Cordylus minor* species complex. The size of the circles is proportional to the frequency of individuals with that haplotype/allele, and the branch lengths are proportional to the number of mutations. The branches interrupted by hatch marks are shortened, with the number of mutations along that branch indicated. The colours represent the proportion of taxa with that haplotype/allele.

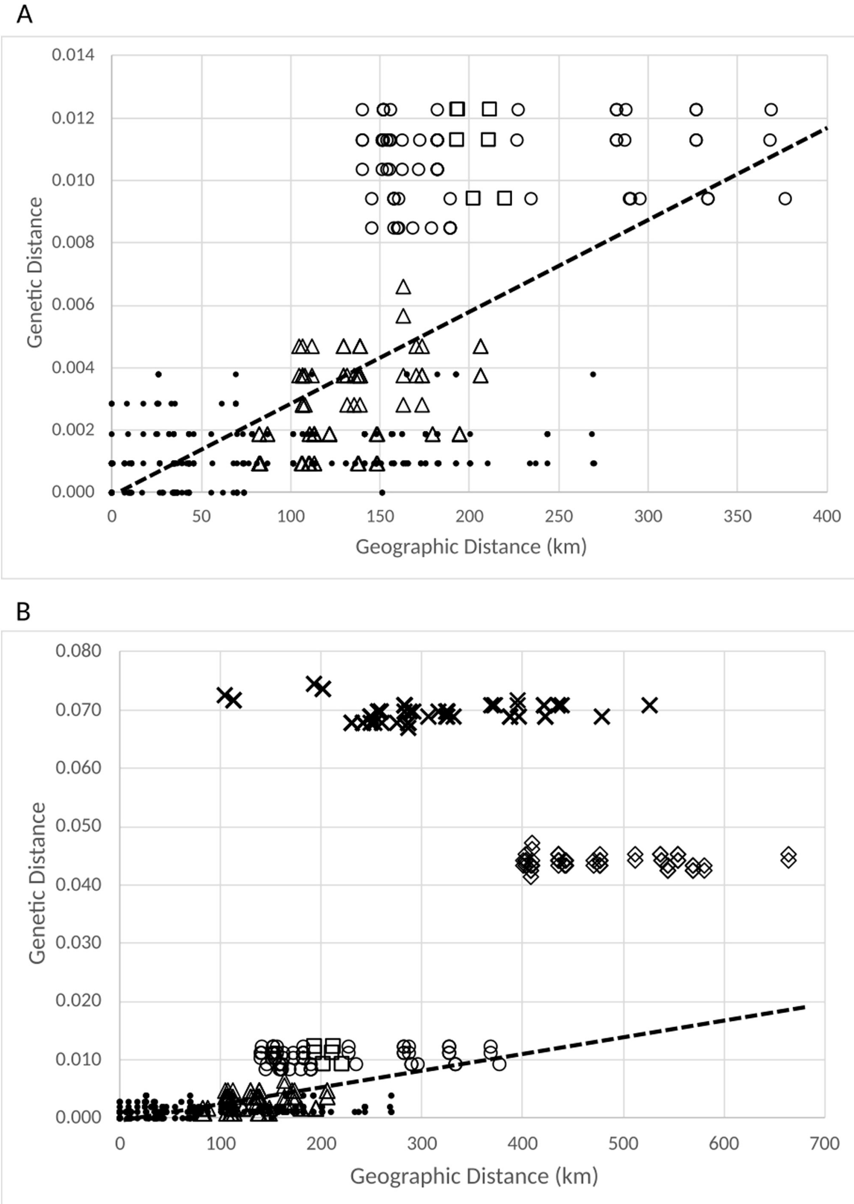

Figure 6. Isolation by distance scatterplots for pairwise intra- and interspecific comparisons for Karoo cordylid lizards. (A) The Karoo taxa only (*C. aridus*, *C. cloetei*, *C. minor*) and (B) Karoo species compared to sister taxa (*C. imkeae*, *C. macropholis*). Intraspecific comparisons denoted by black dots. Interspecific pairwise comparisons: *C. aridus*/*C. cloetei* – triangles, *C. aridus*/*C. minor* – squares, *C. cloetei*/*C. minor* – circles, three Karoo species/*C. imkeae* – diamonds, three Karoo species/*C. mclachlani* – crosses. Isolation by distance trend is shown by the dotted line. Axes in A and B are not of the same range.

over-splitting (Satler *et al.*, 2013; Luo *et al.*, 2018; Chambers & Hillis, 2020), it should be noted that our data set included only a few individuals from the other *Cordylus* species, which were often from the same locality. This lack of coverage over the full spectrum of species in the genus could have impacted model performance, yet despite this, our results do not support the hypothesis of three Karoo species. Therefore, it is most parsimonious to accept the three described Karoo species as a single species rather than falsely delimit species that do not represent independently evolving lineages (Carstens *et al.*, 2013).

Our improved sampling shows the range is much wider than had been thought and extends the combined range of the Karoo taxa more than 180 km. Our surveys have gone some way to filling parts of the sampling gaps and our findings suggest that the three Karoo taxa have a relatively continuous distribution that may be patchy in

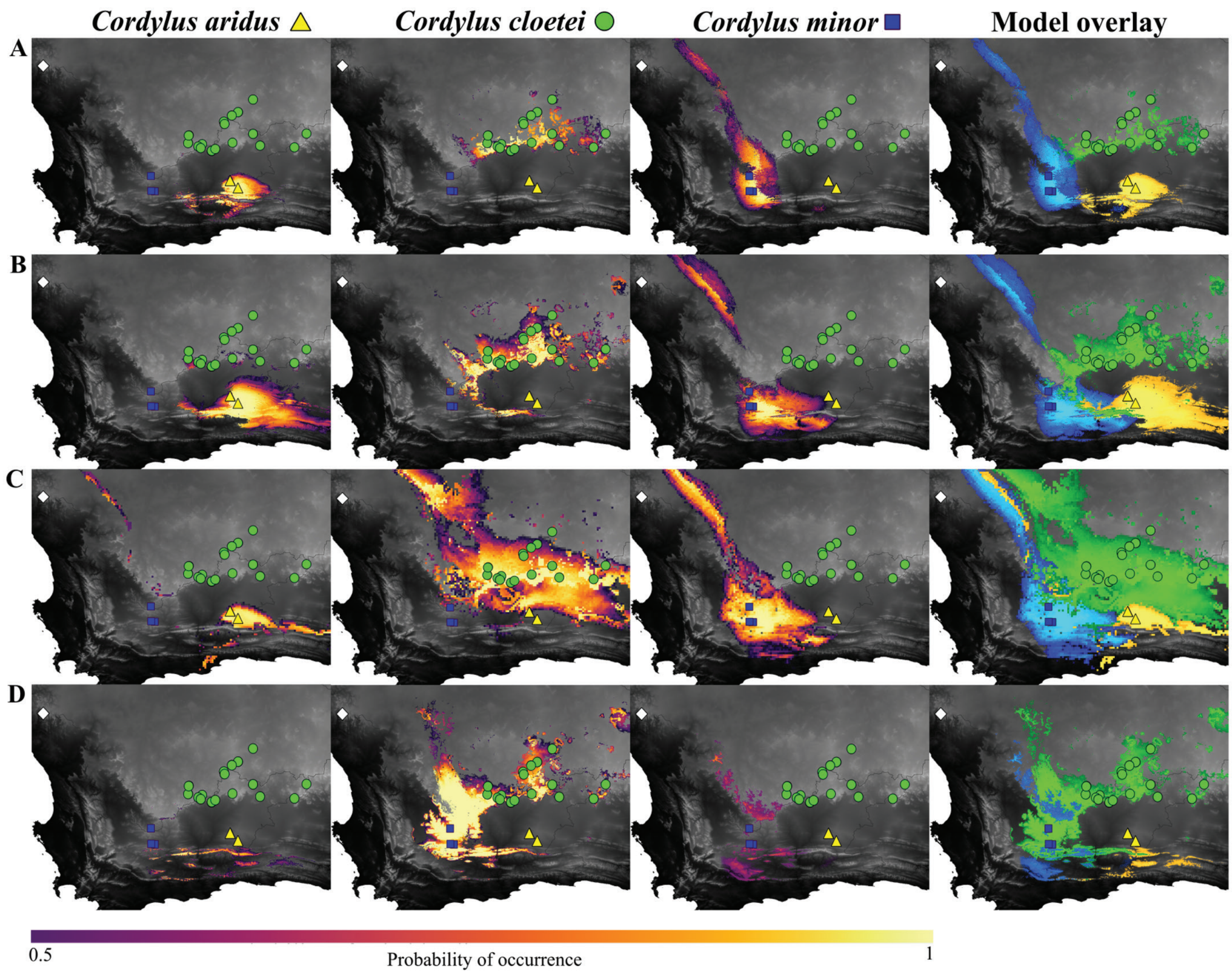

Figure 7. Species distribution models for the *Cordylus minor* species complex at four time-slices under a scenario that assumes four separate taxa (non-synonymy of the Karoo taxa) for the (A) Present, (B) Mid-Holocene (6 Kya), (C) Last Glacial Maximum (21 Kya) and (D) Last Interglacial (120 Kya) with shading showing the areas of suitability at the 10% training presence logistic threshold (Supporting Information, Table S5). The most suitable areas for each species are overlaid to show the areas of connectivity between taxa. The location of *C. imkeae* is indicated by the white diamond.

some places, but they are unlikely to be truly allopatric or isolated. The rugged, high elevation mountainous terrain along the Great Escarpment is essentially continuous providing ample connectivity between populations. Therefore, we suggest that there are no barriers as originally proposed (Mouton & van Wyk, 1994) and that the magnitude of gene flow is not significantly impeded between the three Karoo taxa.

On the balance of evidence, the fourth taxon in the group, *C. imkeae* from Namaqualand, appears to be a valid species. This taxon is diagnosable by morphology, albeit weakly, with slightly different ranges of values for four primary characters. The species delimitation analyses support it as separate, sister to the three Karoo taxa with a divergence that is within the range of interspecific divergence values for the genus. There is no evidence of haplotype/allele sharing between *C. imkeae* and the Karoo taxa, and its divergence cannot be explained by isolation by distance, despite what might be viewed as a sampling gap between *C. imkeae* and the other taxa. Notably, this 'sampling gap' is in a moderately sampled region of the western Great Escarpment (Supporting Information, Fig. S1) and the lack of records from the intervening areas suggests that the gap is real. However, if such populations were discovered, they would need to be evaluated in the current phylogenetic framework to assess their taxonomic status particularly with reference to the validity of *C. imkeae*.

Species distribution modelling

Regardless of whether the three taxa or the single taxon scenario is applied, the species distribution

Table 2. Variation in morphological characters for species in the *Cordylus minor* species complex and *Cordylus cordylus* as comparison. For *C. imkeae*, values in bold indicate character values useful for distinguishing this species from the other taxa. The maximum snout-vent length for the specimens examined is given with the corresponding museum number of that specimen. NMB – National Museum; SAM – Iziko South African Museums; PEM – Port Elizabeth Museum; TM – Ditsong National Museum of Natural History. Specimen information is given in Supporting Information Appendix S1, and additional details of morphology are in Supporting Information Table S6.

	C. minor	*C. aridus*	*C. cloetei*	*C. imkeae*	*C. cordylus*
Sample size	9	20	26	7	20
Maximum snout-to-vent length (mm)	64.3 (TM 19564)	69.7 (PEM R16376)	70.5 (NMB R11599)	67.6 (SAM 50897)	81.7 (NMB R10252)
Supralabials	5–7	5–6	4–6	4–5	5–6
Suboculars	3–4	4–5	3–5	**3–4**	3–5
Temporals transverse rows	4	4–5	4–5	4–5	3–5
Chin shields contacting pair of anterior sublabials	1–2	1–2	0–1	**2**	1–3
Dorsals transversely	26–28	27–31	27–30	27–29	24–30
Dorsals longitudinally	22–25	21–26	20–26	21–25	16–20
Ventrals transversely	22–24	22–25	21–25	22–25	20–27
Ventrals longitudinally	14–16	14–16	13–16	16	12–14
Subdigital lamellae 4th toe	10–14	11–14	11–14	11–13	13–17
Femoral pores (per thigh)	4–6	4–7	3–8	6–8	0–9
Glandular femoral scales (maximum per thigh)	8	6	8	**17**	18

models show areas of connectivity between the three Karoo taxa. Connectivity appears patchier at present day and the last interglacial (120 000 years before present [Ybp]), suggesting that the taxa contract into refugia during warmer periods, with the zone of continuous distribution interspersed with lacunae. Conversely, there are large areas of high suitability during the Mid-Holocene (6000 Ybp) and the LGM (23 000 Ybp), suggesting that expansions took place during the cooler phases. Similarly, the Great Karoo is thought to have been climatically unstable throughout the Plio-Pleistocene, showing high climatic velocity that brought about repeated shifts in habitat extent (Tolley *et al.*, 2014). Although these varied climatic conditions throughout the period of glacial cycling would have influenced the distribution of the Karoo taxa, there is no evidence of complete vicariance and of the formation of allopatric populations for the duration necessary for species-level divergence.

Overall, our niche models suggest that the distribution of the Karoo taxa is heavily influenced by terrain ruggedness, with occurrence most probable in the heterogeneous terrain of rocky outcrops and ridges as well as the more continuous mountainous escarpment. Therefore, the shifting and fragmentation of distribution over time as predicted by the models is presumably shaped by attributes of a changing climate superimposed on the suitable terrain. The models revealed that temperature and precipitation are important climatic components. Therefore, it appears that the range of the Karoo taxa extends northwards during cooler periods but contracts and fragments during warmer periods, such as present day, leaving small isolates along the northern range edge that form a zone of disjunct distribution from the main population (Fig. 7; Gorodkov, 1986). These isolates are associated with large inselbergs that rise about 200 m from the pediplain and are scattered over approximately one degree of latitude north of the escarpment. The climatic elevational-latitudinal relationship (Gaston, 2003) would suggest that populations can persist on these inselbergs due to their cooler microclimates, with the intervening area being unsuitable due to higher average temperatures (Supporting Information, Fig. S5). Thus, dynamics of these range edges over time are spatially and temporally complex, and this dynamic will have a direct impact on the extent of connectivity which in turn controls the magnitude of gene flow. Despite range edges becoming fragmented for the Karoo taxa, connectivity across a core region is maintained, and this should provide opportunity for gene flow to hinder divergence.

Taxonomic considerations

Based on a much-improved data set with several lines of evidence analysed by modern methods, all of which agree, we propose that the *C. minor* species complex is comprised of only two species: *C. minor* FitzSimons, 1943 and *C. imkeae* Mouton & van Wyk, 1994, and that *C. aridus* Mouton & van Wyk, 1994 and *C. cloetei* Mouton & van Wyk, 1994 should be relegated to the status of junior synonyms of *C. minor*. We therefore formally synonymize *C. aridus* and *C. cloetei* with *C. minor*. The type locality of *C. minor* 'just north of Matjiesfontein' (FitzSimons, 1943) is imprecise but given that our genetic sampling comes from within 6 km of Matjiesfontein town centre, we consider our new material as topotypic.

The original species descriptions were based on overlapping ranges of morphological traits gathered from a few specimens that were spatially clustered. We have applied an integrative taxonomic approach (Padial *et al.*, 2010), and used an improved spatially distributed data set (Cicero *et al.*, 2021), a more powerful and comprehensive analytical methodology and a robust philosophical framework entrenched in the General Lineage Concept (de Queiroz, 1998, 2007). In this concept, species are defined as independently evolving metapopulations that can be characterized by the coalescence of not only their genes but of their ecology and morphology, and these traits are unified by a reproductive isolation through vicariance and/or a specific mate-recognition system (de Queiroz, 1998). Species, therefore, can be recognized by examining various operational criteria relating to these traits, which are applied in demonstrating whether all individuals of a species have a mutually exclusive common ancestry. Agreement of our results from multiple lines of such evidence provide greater certainty for our interpretation.

CONCLUSION

The quest to describe and catalogue life on Earth (see Mora *et al.*, 2011) is vital to gain perspective on whether our planet's ecosystem can be sustained given the massive human impact over the last centuries. However, in the rush to discover and name species, superficial and formulaic approaches to systematics and taxonomy have focussed on specific operational criteria for defining species, rather than evaluating criteria that underpin a particular species concept. Furthermore, this often includes descriptions of species that are based on limited data sets so that variation within a species may not be well represented, leading to weakly defined diagnostic features. This is often coupled to subjectively defined clades in phylogenies and cut-off sequence divergence values that may vary widely between studies (Goldstein & De Salle, 2011). While these criteria can provide a rough guide for detecting cryptic species (e.g., Meier *et al.*, 2006), incremental reductions of sequence divergence cut-off values for defining species (e.g., De la Riva *et al.*, 2018) result in over-splitting (see Wiemers & Fiedler, 2007). The resulting downward trend in barcoding gaps ultimately 'lowers the bar' for clades to qualify as species.

Our study highlights an example of taxonomic inflation where species delineation, based on a scant data set and limited analytical assessment has resulted in overestimating the number of species. This may be a widespread phenomenon in taxonomy whereby populations or subspecies are described or elevated to species status erroneously due to insufficient data sets with patchy sampling and an ill-defined or non-existent species concept. Although the underestimation of species richness due to the presence of cryptic species is commonly acknowledged (e.g., Vacher *et al.*, 2020), the overestimation of species richness, as demonstrated here, is likely more common than generally acknowledged (Pérez-Ponce de León & Poulin, 2016). The over-splitting of clades to species devalues the concept of a species and diverts scarce resources to the conservation of populations that are not evolutionarily unique.

ACKNOWLEDGEMENTS

This work was carried out with funding from the National Research Foundation of South Africa (NRF Rated Research Incentive Funding: UID 85413 and Foundational Biodiversity Information Program: UID 98864). Research was conducted under provincial collection and export permits: Northern Cape (FAUNA 0052/2016; FAUNA 0053/20136; FAUNA 0278/2016 FAUNA 0279/2016), Western Cape (0056-AAA041-00115 and 0035-AAA004-00655). We are extremely grateful to all the Karoo landowners for access to their properties and their remarkable hospitality and kind assistance, to Chris Broeckhoven for donating samples of *C. cloetei* from the type locality, to Marius Burger and le Fras Mouton for use of their photographs, and to the two reviewers for their constructive comments.

DATA AVAILABILITY

The data underlying this article are available from the GenBank Nucleotide Database (https://www.ncbi.nlm.nih.gov/genbank/) and can be accessed through the accession numbers provided in Table 1.

REFERENCES

Adams DC, **Berns CM**, **Kozak KH**, **Wiens JJ. 2009.** Are rates of species diversification correlated with rates of morphological evolution? *Proceedings of the Royal Society B: Biological Sciences* **276:** 2729–2738.

Aiello-Lammens ME, **Boria RA**, **Radosavljevic A**, **Vilela B**, **Anderson RP. 2015.** spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography* **38:** 541–545.

Aljanabi SM, **Martinez I. 1997.** Universal and rapid salt-extraction of high quality genomic DNA for PCR based techniques. *Nucleic Acids Research* **25:** 4692–4693.

Anderson RP, Gonzalez JI. 2011. Species-specific tuning increases robustness to sampling bias in models of species distributions: an implementation with Maxent. *Ecological Modelling* **222**: 2796–2811.

Bates MF, **Branch WR**, **Bauer AM**, **Burger M**, **Marais J**, **Alexander GJ**, **de Villiers MS. 2014.** *Atlas and Red List of the reptiles of South Africa, Lesotho and Swaziland. Suricata 1.* Pretoria: South African National Biodiversity Institute.

Bouckaert R, **Heled J**, **Kühnert D**, **Vaughan T**, **Wu CH**, **Xie D**, **Suchard MA**, **Rambaut A**, **Drummond AJ. 2014.** BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* **10:** e1003537.

Braconnot P, **Otto-Bliesner B**, **Harrison S**, **Joussaume S**, **Peterchmitt JY**, **Abe-Ouchi A**, **Crucifix M**, **Fichefet T**, **Hewitt CD**, **Kageyama M**, **Kitoh A**, **Loutre MF**, **Marti O**, **Merkel U**, **Ramstein G**, **Valdes P**, **Weber L**, **Yu Y**, **Zhao Y. 2007.** Coupled simulations of the mid-Holocene and Last Glacial Maximum: new results from PMIP2. *Climate of the Past Discussions* **2:** 1293–1346.

Brown JL, **Knowles LL. 2012.** Spatially explicit models of dynamic histories: examination of the genetic consequences of Pleistocene glaciation and recent climate change on the American pika. *Molecular Ecology* **21:** 3757–3775.

Carstens BC, **Pelletier TA**, **Reid NM**, **Satler JD. 2013.** How to fail at species delimitation. *Molecular Ecology* **22:** 4369–4383.

Chambers EA, **Hillis DM. 2020.** The multispecies coalescent over-splits species in the case of geographically widespread taxa. *Systematic Biology* **69:** 184–193.

Cicero C, **Mason NA**, **Jiménez RA**, **Wait DR**, **Wang-Claypool CY**, **Bowie RC. 2021.** Integrative taxonomy and geographic sampling underlie successful species delimitation. *The Auk* **138:** ukab009.

Clement M, **Posada D**, **Crandall KA. 2000.** TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9:** 1657–1659.

Darriba D, **Taboada GL**, **Doallo R**, **Posada D. 2012.** jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9:** 772.

De la Riva I, **Chaparro JC**, **Castroviejo-Fisher S**, **Padial JM. 2018.** Underestimated anuran radiations in the high Andes: five new species and a new genus of Holoadeninae, and their phylogenetic relationships (Anura: Craugastoridae). *Zoological Journal of the Linnean Society* **182:** 129–172.

Drummond AJ, **Rambaut A. 2007.** BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7:** 1–8.

Elith J, **Graham CH**, **Anderson RP**, **Dudík M**, **Ferrier S**, **Guisan A**, **Hijmans RJ**, **Huettmann F**, **Leathwick JR**, **Lehmann A**, **Li J**, **Lohmann LG**, **Loiselle BA**, **Manion G**, **Moritz C**, **Nakamura M**, **Nakazawa Y**, **Overton JM**, **Peterson AT**, **Phillips SJ**, **Richardson KS**, **Scachetti-Pereira R**, **Schapire RE**, **Soberón J**, **Williams S**, **Wisz MS**, **Zimmermann NE. 2006.** Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29:** 129–151.

Elith J, **Phillips SJ**, **Hastie T**, **Dudík M**, **Chee YE**, **Yates CJ. 2011.** A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* **17:** 43–57.

Engelbrecht HM, **Branch WR**, **Greenbaum E**, **Alexander GJ**, **Jackson K**, **Burger M**, **Conradie W**, **Kusamba C**, **Zassi-Boulou AG**, **Tolley KA. 2019.** Diversifying into the branches: species boundaries in African green and bush snakes, *Philothamnus* (Serpentes: Colubridae). *Molecular Phylogenetics and Evolution* **130:** 357–365.

Fick SE, **Hijmans RJ. 2017.** WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* **37:** 4302–4315.

FitzSimons VF. 1943. The lizards of South Africa. *Transvaal Museum Memoirs* **1:** 1–528.

Freitas I, **Ursenbacher S**, **Mebert K**, **Zinenko O**, **Schweiger S**, **Wüster W**, **Brito JC**, **Crnobrnja-Isailović J**, **Halpern B**, **Fahd S**, **Santos X**, **Pleguezuelos JM**, **Joger U**, **Orlov N**, **Mizsei E**, **Lourdais P**, **Zuffi MAL**, **Strugariu A**, **Zamfirescu SR**, **Martínez-Solano I**, **Velo-Antón G**, **Kaliontzopoulou A**, **Martínez-Freiría F. 2020.** Evaluating taxonomic inflation: towards evidence-based species delimitation in Eurasian vipers (Serpentes: Viperinae). *Amphibia-Reptilia* **41:** 285–311.

Gaston KJ. 2003. *The Structure and dynamics of geographic ranges.* Oxford: Oxford University Press.

Goldstein PZ, **De Salle R. 2011.** Integrating DNA barcode data and taxonomic practice: determination, discovery, and description. *Bioessays* **33:** 135–147.

Gorodkov KB. 1986. Three-dimensional climatic model of the potential range and some of its characteristics. *Entomological Review* **65:** 1–18.

Guindon S, **Gascuel O. 2003.** A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. *Systematic Biology* **52:** 696–704.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high-resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**: 1965–1978.

Huelsenbeck JP, **Ronquist F. 2001.** MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17:** 754–755.

Ikeda DH, **Max TL**, **Allan GJ**, **Lau MK**, **Shuster SM**, **Whitham TG. 2016.** Genetically informed ecological niche models improve climate change predictions. *Global Change Biology* **23:** 164–176.

Isaac NJB, **Mallet J**, **Mace GM. 2004.** Taxonomic inflation: its influence on the macroecology and conservation. *Trends in Ecology and Evolution* **19:** 464–469.

Kearse M, **Moir R**, **Wilson A**, **Stones-Havas S**, **Cheung M**, **Sturrock S**, **Buxton S**, **Cooper A**, **Markowitz S**, **Duran C**, **Thierer T**, **Ashton B**, **Meintjes P**, **Drummond A. 2012.** Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28:** 1647–1649.

Kumar S, **Stecher G**, **Tamura K. 2016.** MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33:** 1870–1874.

Lefébure T, **Douady C**, **Gouy M**, **Gibert J. 2006.** Relationship between morphological taxonomy and molecular divergence within Crustacea: proposal of a molecular threshold to help species delimitation. *Molecular Phylogenetics and Evolution* **40:** 435–477.

Leigh JW, **Bryant D. 2015.** PopART: full-feature software for haplotype network construction. *Methods in Ecology and Evolution* **6:** 1110–1116.

Luo A, **Ling C**, **Ho SY**, **Zhu CD. 2018.** Comparison of methods for molecular species delimitation across a range of speciation scenarios. *Systematic Biology* **67:** 830–846.

Macey JR, **Larson A**, **Ananjeva NB**, **Fang Z**, **Papenfuss TJ. 1997.** Two novel gene orders and the role of light-strand replication in rearrangement of the vertebrate mitochondrial genome. *Molecular Biology and Evolution* **14:** 91–104.

McCarthy T, **Rubridge B. 2005.** *The story of earth and life: a Southern African perspective on a 4.6-billion-year journey.* Cape Town: Struik Publishers.

Meier R, **Dikow T. 2004.** Significance of specimen databases from taxonomic revisions for estimating and mapping the global species diversity of invertebrates and repatriating reliable specimen data. *Conservation Biology* **18:** 478–488.

Meier R, **Shiyang K**, **Vaidya G**, **Ng P. 2006.** DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. *Systematic Biology* **55:** 715–728.

Miller MA, **Pfeiffer W**, **Schwartz T. 2010.** Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: *Proceedings of the Gateway Computing Environments workshop (GCE)*. San Diego Supercomputer Center, New Orleans, LA, 14 November 2010, 1–8.

Miller MP. 2005. Alleles in space (AIS): computer software for the joint analysis of interindividual spatial and genetic information. *Journal of Heredity* **96:** 722–724.

Mora C, **Tittensor DP**, **Adl S**, **Simpson AG**, **Worm B. 2011.** How many species are there on Earth and in the ocean? *PLoS Biology* **9:** e1001127.

Mouton PLF, **van Wyk JH. 1989.** *Cordylus minor*: a valid species of South African lizard (Reptilia: Cordylidae). *African Zoology* **24:** 322–328.

Mouton PLF, **van Wyk JH. 1994.** Taxonomic status of geographical isolates in the *Cordylus minor* complex (Reptilia: Cordylidae): a description of three new species. *Journal of the Herpetological Association of Africa* **43:** 6–18.

Muscarella R, **Galante PJ**, **Soley-Guardia M**, **Boria RA**, **Kass JM**, **Uriarte M**, **Anderson RP. 2014.** ENMeval: an R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. *Methods in Ecology and Evolution* **5:** 1198–1205.

Padial JM, **Miralles A**, **De la Riva I**, **Vences M. 2010.** The integrative future of taxonomy. *Frontiers in Zoology* **7:** 1–14.

Palumbi S, **Martin A**, **Romano S**, **McMillan WO**, **Stice L**, **Grabowski G. 1991.** *Simple fool's guide to PCR.* Honolulu: Department of Zoology and Kewalo Marine Laboratory.

Paterson HEH. 1985. The recognition concept of species. In: Vrba ES, ed. *Species and speciation.* Pretoria: Transvaal Museum, 21–29.

Pérez-Ponce de León G, **Poulin R. 2016.** Taxonomic distribution of cryptic diversity among metazoans: not so homogeneous after all. *Biology Letters* **12:** 20160371.

Phillips SJ, **Anderson RP**, **Schapire RE. 2006.** Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **190:** 231–259.

QGIS.org. 2021. *QGIS Geographic Information System.* QGIS Association. Available at: http://www.qgis.org

de Queiroz K. 1998. The general lineage concept of species, species criteria, and the process of speciation: a conceptual unification and terminological recommendations. In: Howard DJ, Berlocher SH, eds. *Endless forms: species and speciation.* Oxford: Oxford University Press, 57–75.

de Queiroz K. 2007. Species concepts and species delimitation. *Systematic Biology* **56:** 879–886.

R Core Team. 2013. *R: a language and environment for statistical computing.* Vienna: R Foundation for Statistical Computing. Available at: http://www.R-project.org/

Rambaut A, **Drummond AJ**, **Xie D**, **Baele G**, **Suchard MA. 2018.** Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* **67:** 901.

Reid NM, **Carstens BC. 2012.** Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. *BMC Evolutionary Biology* **12:** 196.

Reissig J. 2014. *Girdled lizards and their relatives: natural history, captive care and breeding.* Frankfurt am Main: Edition Chimaira.

Riley SJ, **DeGloria SD**, **Elliot R. 1999.** A terrain ruggedness index that quantifies topographic heterogeneity. *Intermountain Journal of Sciences* **5:** 23–27.

Rosenzweig C, **Karoly D**, **Vicarelli M**, **Neofotis P**, **Wu Q**, **Casassa G**, **Menzel A**, **Root TL**, **Estrella N**, **Seguin B**, **Tryjanowski P**, **Liu C**, **Rawlins S**, **Imeson A. 2008.** Attributing physical and biological impacts to anthropogenic climate change. *Nature* **453:** 353–357.

Satler JD, Carstens BC, Hedin M. 2013. Multilocus species delimitation in a complex of morphologically conserved trapdoor spiders (Mygalomorphae, Antrodiaetidae, Aliatypus). *Systematic Biology* **62**: 805–823.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30:** 1312–1313.

Stanley EL, **Bauer AM**, **Jackman TR**, **Branch WR**, **Mouton PLFN. 2011.** Between a rock and a hard polytomy: rapid radiation in the rupicolous girdled lizards (Squamata: Cordylidae). *Molecular Phylogenetics and Evolution* **58:** 53–70.

Stephens K, **Alexander GJ**, **Makhubo BG**, **Telford NS**, **Tolley KA. In review.** Mistaken identity: reliance on locality information influences identification in the morphologically conservative genus *Trachylepis*. *African Journal of Herpetology*

Tabachnick BG, **Fidell LS. 1996.** *Using multivariate statistics, 3rd edn*. New York: HarperCollins College Publishers.

Taylor KE, **Stouffer RJ**, **Meehl GA. 2012.** An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society* **93:** 485–498.

Telford NS, **Alexander GJ**, **Becker FS**, **Conradie W**, **Jordaan A**, **Kemp L**, **le Grange A**, **Rebelo AD**, **Strauss P**, **Taft JM**, **Weeber J**, **Tolley KA. In press.** Extensions to the known Geographic distributions of reptiles in the Karoo, South Africa. *Amphibian and Reptile Conservation*.

Tolley KA, **Bowie RCK**, **Price BW**, **Measey GJ**, **Forest F. 2014.** The shifting landscape of genes since the Pliocene: terrestrial phylogeography in the Greater Cape Floristic Region. In: Allsopp N, Colville JF, Verboom T, eds. *Ecology and evolution of fynbos understanding megadiversity.* Oxford: Oxford University Press, 142–163.

Townsend TM, **Alegre RE**, **Kelley ST**, **Wiens JJ**, **Reeder TW. 2008.** Rapid development of multiple nuclear loci for phylogenetic analysis using genomic resources: an example from squamate reptiles. *Molecular Phylogenetics and Evolution* **47:** 129–142.

Vacher JP, **Chave J**, **Ficetola FG**, **Sommeria-Klein G**, **Tao S**, **Thébaud C**, **Blanc M**, **Camacho A**, **Cassimiro J**, **Colston TJ**, **Dewynter M. 2020.** Large-scale DNA-based survey of frogs in Amazonia suggests a vast underestimation of species richness and endemism. *Journal of Biogeography* **47:** 1781–1791.

Warren DL, **Glor RE**, **Turelli M. 2010.** ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography* **33:** 607–611.

Warren DL, **Seifert SN. 2011.** Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications* **21:** 335–342.

Wiemers M, **Fiedler K. 2007.** Does the DNA barcoding gap exist? – a case study in blue butterflies (Lepidoptera: Lycaenidae). *Frontiers in Zoology* **4:** 8.

Zheng Y, **Wiens JJ. 2016.** Combining phylogenomic and supermatrix approaches, and a time-calibrated phylogeny for squamate reptiles (lizards and snakes) based on 52 genes and 4162 species. *Molecular Phylogenetics and Evolution* **94:** 537–547.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Figure S1. Reptile record density from museum collections (see Bates *et al.*, 2014), public databases (iNaturalist: https://www.inaturalist.org; ReptileAtlas: http://vmus.adu.org.za/) and the present surveys. Darker (red) shaded cells show a higher density of reptile collections, and the blank grid cells have zero records. The elevation map underlies the density records (darkest shading shows highest elevation). Localities recorded for the *C. minor* species complex are shown (*C. aridus* – yellow triangles, *C. cloetei* – green circles, *C. minor* – blue squares, *C. imkeae* – white diamond).

Figure S2. Locality records for the *C. minor* species complex that were used in the species distribution modelling shown with elevation shading (darkest shading shows highest elevation). Some symbols are inclusive of multiple individual records, and for these, the red dots represent the number of individual occurrences represented. Localities recorded for the *C. minor* species complex are shown (*C. aridus* – yellow triangles, *C. cloetei* – green circles, *C. minor* – blue squares, *C. imkeae* – white diamond).

Figure S3. bGMYC heatmap for *Cordylus* from three-gene analysis, shaded by the range of marginal posterior probabilities for species identities.

Figure S4. bGMYC heatmap for *Cordylus* from the two-gene mitochondrial analysis, shaded by the range of marginal posterior probabilities for species identities.

Figure S5. Localities recorded for the *C. minor* species complex with (A) annual mean temperature (°C), (B) daily temperature range (°C) in summer (February), (C) daily temperature range (°C) in winter (August), (D) mean annual precipitation (mm), (E) median winter (August) precipitation (mm). The environmental variables mapped correspond to the most influential variables for the niche modelling i.e., Bio 1, Bio 2, Bio 12 and Bio 19, respectively but with Bio 2 represented here by both summer and winter diurnal temperature range. Occurrence records for the species are indicated – *C. aridus*, yellow triangles; *C. cloetei*, green circles; *C. minor*, blue squares; *C. imkeae* – white diamond. Map layers from Schultze (1997).

Figure S6. Species distribution models for the *C. minor* species complex at four time-slices under a scenario that assumes the synonymy of the three Karoo taxa for (A) the Present, (B) the Mid-Holocene (6 Kya), (C) the Last Glacial Maximum (21 Kya) and (D) the Last Interglacial (120 Kya) with shading showing the areas of suitability. The occurrence records for the individual species are shown – *C. aridus*, yellow triangles; *C. cloetei*, green circles; *C. minor*, blue squares. For reference, the general locality for *C. imkeae* is shown by the white diamond, but this species was not included in the model due to too few available data points.

Table S1. Traits regarded as diagnostic in the original descriptions of species in the *C. minor* species complex [data compiled from FitzSimons (1943); Mouton & van Wyk (1989, 1994)].

Table S2. Variation in head proportions (measurements from adults ≥ 50 mm SVL) between species in the *C. minor* species complex. Sample sizes are indicated for each species. Values for each proportion are the mean and standard deviation, and the range of values. Specimens from the National Museum, Bloemfontein, and Port Elizabeth Museum, Gqeberha (see Supporting Information, Appendix S1). SVL – snout-to-vent length.

Table S3. Uncorrected net p-distances for *Cordylus* species for (a) combined mitochondrial genes, (b) 16S only, (c) *ND2* only and (d) *PRLR*. Pairwise comparisons among species are in the bottom matrices, whereas intraspecific values are on the diagonal. na – not available: instances where only one individual was available (intraspecific), or no sequences were available for that species (interspecific). Comparisons between and within *C. aridus*, *C. cloetei*, *C. minor* and *C. imkeae* are shown in bold and are along the top rows of the matrices.

Table S4. The contributions (permutation importance percentage) of each variable in species distribution models for each *Cordylus* taxon and for the single taxon scenario at two resolutions (30 arc seconds – approximately 1 km^2 and 2.5 arc minutes – approximately 5 km^2). Bio 1 – annual mean temperature, Bio 2 – diurnal temperature range, Bio 3 – isothermality, Bio 6 – min temperature of coldest month, Bio 12 – annual precipitation, Bio 19 – precipitation of coldest quarter, Terrain – Terrain ruggedness index. Contributions of < 1% are indicated by a dash.

Table S5. Evaluation metrics of the optimum Maxent models for each taxon of the *C. minor* species complex. Metrics shown are feature class (L: linear, Q: quadratic, H: hinge, P: product, T: threshold), regulization parameter, Akaike Information Criterion (AIC), threshold-independent metric (AUC_{Test}), difference between test and training AUC (AUC_{Diff}), minimum training presence omission rate (OR_{mtp}), training omission rate (OR_{10}) and the 10% training presence logistic threshold (10_{TPLT}).

Table S6. Variation in morphological characters for type specimens and new material referable to the *C. minor* species complex, with *C. cordylus* as a comparative species. Scale counts on the head are given for one side, and half values (e.g., 6.5) result from differing values on either side of the head. For *C. imkeae*, values in bold indicate character values useful for distinguishing this species from other taxa in the *C. minor* species complex. Maximum snout-to-vent length for specimens examined is given, with the corresponding museum number. NMB – National Museum, Bloemfontein; SAM – Iziko South African Museums, Cape Town; PEM – Bayworld (Port Elizabeth Museum), Gqeberha; TM – Ditsong National Museum of Natural History, Pretoria. Specimen details are in Supporting Information, Appendix S1.

Appendix S1. List of specimens from the *Cordylus minor* species complex and *C. cordylus* examined for this study. *Cordylus aridus* and *C. cloetei* are now considered junior synonyms of *C. minor*. NMB–National Museum, Bloemfontein; SAM–Iziko South African Museums; PEM–Port Elizabeth Museum (Bayworld), Gqeberha; TM–Ditsong National Museum of Natural History, Pretoria.