ORIGINAL ARTICLE

How many species of angulate tortoises occur in Southern Africa? (Testudines: Testudinidae: *Chersina*)

Cäcilia Spitzweg[1] [ID] | Melita Vamberger[1] [ID] | Flora Ihlow[1] [ID] | Uwe Fritz[1] [ID] |
Margaretha D. Hofmeyr[2][†] [ID]

[1]Museum of Zoology (Museum für Tierkunde), Senckenberg Dresden, Dresden, Germany

[2]Chelonian Biodiversity and Conservation, Department of Biodiversity and Conservation Biology, University of the Western Cape, Bellville, South Africa

Correspondence
Uwe Fritz, Museum of Zoology (Museum für Tierkunde), Senckenberg Dresden, Dresden, Germany.
Email: uwe.fritz@senckenberg.de

Abstract

Using range-wide sampling and 1,143 bp of mtDNA (cytochrome *b* gene) and 14 microsatellite loci, we examined genetic differentiation in the widely distributed Southern African angulate tortoise (*Chersina angulata*). We found evidence for two genealogical lineages that differ in both genetic marker systems and their preferred habitat conditions. According to a fossil-calibrated molecular clock for all African tortoise lineages using 1,870 bp mitochondrial and 1,416 bp nuclear DNA, the two lineages of *C. angulata* diverged in the Pliocene (approx. 3.8 million years ago). Species distribution models reveal that the ranges of the two lineages shifted little since the Last Glacial Maximum, which is in agreement with the demographic population descriptors suggestive of stationary populations that did not experience expansion. One lineage occurs in the west, and the other in the south of the extant distribution range. In the geographic contact zone, the two lineages hybridize extensively, providing evidence for their conspecificity under the biological species concept. Each lineage could be recognized as a distinct subspecies, but the ill-defined geographic origins of the type material of the available names prevent their identification with any taxon. With respect to the nuclear genomic markers, the western lineage shows further north-south substructuring. A few genetically mismatched tortoises are interpreted as human translocations. Our study underlines that confiscated or captive angulate tortoises of unknown geographic provenance should not be released without prior genetic screening to avoid genetic pollution of wild populations.

KEYWORDS

cyt *b*, distribution, Last Glacial Maximum, MAXENT, microsatellite loci, species distribution model

1 | INTRODUCTION

Africa harbours an outstanding chelonian diversity with many endemic species. Eleven of the 16 extant testudinid genera occur on this continent, and of these, six genera with 13 species live in South Africa (Buhlmann et al., 2009; Hofmeyr & Baard, 2014; Hofmeyr, Vamberger, Branch, Schleicher, & Daniels, 2017; Mittermeier, van

Dijk, Rhodin, & Nash, 2015; TTWG, 2017). In particular, the south-west of this country is renowned for its rich biodiversity, with the Cape Floristic Region and Succulent Karoo representing two global hotspots of both floral and faunal endemicity with seven tortoise species (Hofmeyr & Baard, 2014; Myers, Mittermeier, Mittermeier, da Fonseca, & Kent, 2000). One of these is the angulate tortoise _Chersina angulata_ (Schweigger, 1812), endemic to South Africa and adjacent south-western Namibia. _Chersina angulata_ is widely distributed across the coastal plains from extreme south-western Namibia to the south-western Cape region of South Africa and eastwards to East London in the Eastern Cape Province. Thus, the range of _C. angulata_ stretches across a number of biomes characterized by different climate regimes and vegetation types, with low annual precipitation (<100 mm) and Succulent Karoo vegetation in the north-west, a Mediterranean climate (annual precipitation _ca_. 250 mm) and Fynbos experiencing non-seasonal precipitation in the south-west, and highest annual rainfall (600–700 mm) and Dune Thicket in the east (Hofmeyr, 2009; Lesia, Hofmeyr, & Amato, 2003).

A study on regional differences in body size, gular shield length, and plastron concavity from three regions, namely the Eastern Cape, the Western Cape, and Namaqualand, found no significant differences in adult body size for males. However, in the Western Cape, females were significantly smaller than those from the other two regions (van den Berg & Baard, 1994). Angulate tortoises of both sexes from Dassen Island are known to be large-sized, with maximum shell lengths of up to 35 cm opposed to 17–19 cm on the continent (Hofmeyr, 2009). Lesia et al. (2003) and Daniels, Hofmeyr, Henen, and Crandall (2007) examined the genetic differentiation and population structure of _C. angulata_. Using mitochondrial cyt _b_ sequences of 25 individuals from the West Coast National Park, Dassen Island, and Kleinsee, Lesia et al. (2003) found the Kleinsee population to represent a lineage distinct from both south-western populations and argued for a detailed range-wide investigation.

Daniels et al. (2007) used a geographically wider sampling and three partial mitochondrial genes (cyt _b_: 320 bp, COI: 599 bp, ND4: 791 bp). These authors found two parapatrically distributed mitochondrial lineages, in the south and west of the species' range, and suggested that these lineages represent distinct taxa. No investigation for nuclear genomic differentiation has been conducted so far to scrutinize these findings.

Estimates for the divergence times of _C. angulata_ differ considerably. Based on mtDNA sequences and the assumption that _C. angulata_ diverged from its sister _Chersobius_ 18 million years ago (mya), Daniels et al. (2007) dated the onset of the radiation of _C. angulata_ to the Miocene (10.4–8.4 mya). In contrast, using representatives of all African testudinid genera and mitochondrial and nuclear DNA sequences for a fossil-calibrated molecular clock, Hofmeyr et al. (2017) suggested that _Chersobius_ and _Chersina_ separated 29.5 mya and that the divergence within _C. angulata_ commenced only in the Pliocene (3.3 mya).

The present study aims at reexamining the phylogeography and divergence times within _Chersina_ using comprehensive range-wide sampling. In particular, we (a) infer population structuring of _C. angulata_ using the mitochondrial cyt _b_ gene and 14 polymorphic microsatellite loci, (b) reassess divergence times of genealogical lineages of _C. angulata_ to elucidate its biogeography using the same approach as in Hofmeyr et al. (2017), and (c) correlate and discuss our findings in the light of species distribution models for _C. angulata_ under current and paleoclimatic conditions during the Last Glacial Maximum (LGM).

2 | MATERIALS AND METHODS

2.1 | Sampling and laboratory procedures

Two hundred forty-six samples of _Chersina angulata_ were studied, including 167 samples previously processed by Daniels et al. (2007) plus 79 new samples (Table S1). While extracted DNA was already available for the samples used by Daniels et al. (2007), DNA of new blood and tissue samples was isolated using the InnuPrep Blood DNA Mini Kit and the InnuPrep DNA Mini Kit (both kits: Analytik Jena AG).

The highly informative cyt _b_ gene, the most frequently used mitochondrial marker for phylogeographic and taxonomic differentiation of chelonian taxa (e.g., Spinks, Shaffer, Iverson, & McCord, 2004; Spitzweg, Hofmeyr, Fritz, & Vamberger, 2019; Vamberger et al., 2015), was chosen, and PCR was performed using the AccuStart II GelTrac PCR SuperMix (Quanta BioSciences; modified protocol: _Taq_-Mix 1:2 dilution). The total reaction volume of 25 μl contained 20–40 ng of DNA and 0.4 μM of the CytbG primer (Spinks et al., 2004) and 0.4 μM of the newly designed primer Cytb-GAC-rev (5′-GAC CAA TGC TTT GTT ATT AAG CTA C-3′). Initial denaturation lasted for 2 min at 94°C, followed by 35 cycles with denaturation at 95°C for 30 s, annealing at 50°C for 30 s, and elongation at 72°C for 1 min, and a 7-min-long final elongation step at 72°C. PCR products were purified using the ExoSAP-IT PCR Product Cleanup Reagent (Applied Biosystems, Life Technologies Corporation; 1:20 dilution; modified protocol: 30 min at 37°C, 15 min at 80°C). The BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) was used for sequencing under the following conditions: initial denaturation for 1 min at 94°C, followed by 25 cycles with denaturing at 96°C for 10 s, annealing at 50°C for 5 s, and elongation at 60°C for 75 s. For sequencing, the internal primer L-430cytb (Stuart & Fritz, 2008) and the new primer Cytb-GTT-int-rev (5′-GTT TGA TCC AGT TTC GTG GA-3′) were used. Cycle sequencing reaction products

were purified using the Performa DTR V3 96-Well Short Plate (Edge Biosystems), with each well filled with 400 µl Sephadex (GE Healthcare; 1:20 solution). Sequences were determined on an ABI 3730 DNA Analyzer (Applied Biosystems). For this purpose, the sequencing reaction was diluted 1:2 in 0.1 mM EDTA (Roth). The resulting sequences correspond to the complete cyt *b* gene and were 1,143 bp long.

Primers for 54 microsatellite loci developed for other turtle species were tested for cross-amplification in *C. angulata* (Table S2). Tests were run using cycling conditions as described in Table S3 using a reaction volume of 10 µl containing 10–20 ng of total DNA, 0.5 units of *Taq* polymerase (Bioron), the buffer recommend by the supplier (without $MgCl_2$, 10×), 1.5 mM $MgCl_2$ (Bioron), 0.2 mM of each dNTP (Roth), 2 µg of BSA (Thermo Fisher Scientific), and 0.3 µM of each primer. In case of successful amplification, the presence of the microsatellite was confirmed by sequencing with unlabelled reverse primers according to the PCR and sequencing conditions for cyt *b*.

For the selected loci, eight multiplex PCRs were run using a reaction volume of 10 µl as described above but with 3.0 mM $MgCl_2$ (Bioron) and a specific set of primers (Table S4). Thermocycling conditions are summarized in Table S5. Fragment lengths were determined on an ABI 3730 DNA Analyzer (Applied Biosystems). For this purpose, 1 µl of the PCR product was diluted in 49 µl double-distilled water; 1 µl of this dilution was mixed with 8.5 µl Hi-Di Formamide (Applied Biosystems), 0.25 µl double-distilled water, and 0.25 µl GeneScan-600 LIZ Size Standard (Applied Biosystems). The software PEAK SCANNER 1.0 (Life Technologies) was used for scoring fragment lengths. Null allele frequencies were examined using MICRO-CHECKER 2.2.3 (van Oosterhout, Hutchinson, Wills, & Shipley, 2004). The TWQ113 locus was monomorphic, and the TWL61 locus showed an overrepresentation of homozygotes; the TWI161, Maucas24, GmuD55, Test71, GAL45, and GAL127 loci could not be amplified for many individuals and were therefore not used, so that 14 informative microsatellite loci of 224 tortoises were available for further analyses.

2.2 | Data processing

2.2.1 | Mitochondrial DNA

Cyt *b* sequences could be generated for 180 tortoises because only small amounts of DNA remained of many samples previously used by Daniels et al. (2007). In these instances, we prioritized genotyping the samples using microsatellites. In a few other cases, especially in material originating from carcasses found in the wild, only the short microsatellite sequences could be amplified, but not the longer PCR products for cyt *b*.

Cyt *b* sequences were aligned using BIOEDIT 7.0.9.0 (Hall, 1999). Sequences from *Chersobius boulengeri* (European

Nucleotide Archive accession number: LR697071) and *Chelonoidis denticulatus* (LT599485), both generated in our laboratory (Kehlmaier et al., 2017; Kehlmaier, Graciá, et al., 2019), were used as outgroup taxa. *Chersobius* is the sister taxon of *Chersina*, while *Chelonoidis* represents a more distantly related tortoise genus (Hofmeyr et al., 2017; Kehlmaier et al., 2017).

For phylogenetic analyses, cyt *b* sequences were collapsed into haplotypes using the TCS algorithm (Clement, Posada, & Crandall, 2000) implemented in POPART (http://popart.otago. ac.nz); for accession numbers of haplotypes, see Table S1. POPART was also used to draw a TCS network for haplotypes. Evolutionary relationships of haplotypes were inferred using Maximum Likelihood (ML) and Bayesian analyses (BA). The best-fitting partitioning scheme and evolutionary models for phylogenetic analyses were determined in PARTITIONFINDER2 (Lanfear, Frandsen, Wright, Senfeld, & Calcott, 2017) using the Bayesian information criterion. Consequently, the data set was partitioned by codon position for Bayesian and ML analyses, using the K80 + I model for the first codon position, the HKY + I model for the second codon position, and the TIM + I model for the third codon position.

RAxML-NG (Kozlov, Darriba, Flouri, Morel, & Stamatakis, 2019) was used to infer phylogenetic relationships of haplotypes under ML. Five independent searches were run using different starting conditions, the described partitioning, and the fast bootstrap algorithm to explore the robustness of the results by comparing the best trees. Then, non-parametric thorough bootstrap replicates were calculated until they converged with a cut-off of 3% and plotted against the best tree. Bayesian analyses were performed using MRBAYES 3.2.1 (Ronquist et al., 2012) with two parallel runs (each with four chains) and default parameters. The chains ran for 10 million generations with every 100th generation sampled. Calculation parameters were analysed using a burn-in of 2.5 million generations to assure that both runs converged. Subsequently, only the plateau of the most likely trees was sampled, and a 50% majority rule consensus tree was generated.

Mismatch distributions were computed under the demographic expansion model (10,000 bootstrap replicates) for all samples and the two clades separately using ARLEQUIN 3.5.2.1 (Excoffier & Lischer, 2010). In addition, genetic diversity and population dynamic indices were calculated using the same software. Uncorrected *p* distances between the two clades were obtained with MEGA 6.0 (Tamura, Stecher, Peterson, Filipski, & Kumar, 2013).

2.2.2 | Microsatellite data

Population structuring of *Chersina angulata* was inferred using data for 14 microsatellite loci of 224 individuals and

the unsupervised clustering approach of STRUCTURE 2.3.4 (Hubisz, Falush, Stephens, & Pritchard, 2009; Pritchard, Stephens, & Donnelly, 2000), the admixture model, and correlated allele frequencies. All calculations were run for $K = 1–10$, and the most likely number of clusters (K) was inferred using STRUCTURE HARVESTER (Earl & vonHoldt, 2012) and the implemented ΔK method (Evanno, Regnaut, & Goudet, 2005). Calculations were repeated ten times for each K using a MCMC chain of 1,000,000 generations for each run and a burn-in of 250,000 generations. Population structuring and individual admixture were then visualized using DISTRUCT 1.1 (Rosenberg, 2004). Individuals with a membership proportion below 80% were treated as having admixed ancestries (Barilani et al., 2007; Randi, 2008).

Since STRUCTURE is known to be sensitive to uneven sample sizes (Puechmaille, 2016), additional Principal Component Analyses (PCAs) were run using the package ADEGENET 2.0.1 (Jombart, 2008) for R 3.5.3 (https://www.R-project.org/). PCAs are less biased by sample size and independent from population genetic presumptions so that their parallel application is advised (Puechmaille, 2016). Three distinct PCAs were calculated. For two PCAs, microsatellite data of all 224 tortoises were used. The individual symbols were coloured according to their assignment by STRUCTURE to two or three clusters ($K = 2$ or $K = 3$). For the third PCA, only individuals were included for which the mitochondrial lineage was known (175), and symbol colours were selected that corresponded to the respective mitochondrial lineage. Results were visualized three-dimensionally using the R package RGL 0.100.19 (Adler et al., 2019).

2.2.3 | Diversity and divergence parameters

For each STRUCTURE cluster, diversity and divergence parameters were estimated using microsatellite and mtDNA data. For doing so, a frequency table for microsatellite alleles was obtained using CONVERT 1.31 (Glaubitz, 2004), and values for locus-specific allelic richness were computed with FSTAT 2.9.3.2 (Goudet, 1995). For microsatellite data of individual clusters, locus-specific observed (H_O) and expected heterozygosities (H_E) and AMOVAs (10,000 permutations) were calculated in ARLEQUIN. The same software was also used to obtain F_{ST} values. Haplotype and nucleotide diversities were computed in DNASP 5.10.01 (Librado & Rozas, 2009). For calculations using microsatellite data, individuals with mixed ancestry were excluded.

2.3 | Molecular clock

To estimate when genetic divergence within *Chersina angulata* commenced, the data set of Hofmeyr et al. (2017) for

African Testudinidae was used, consisting of three mitochondrial genes and two nuclear loci. For this purpose, the original sequences of *C. angulata* were replaced by new sequences for 12 individuals representing all mitochondrial clades and nuclear genomic clusters identified in the present study. European Nucleotide Archive accession numbers of our new sequences are listed in Table S6. These sequences were produced according to the laboratory approaches of Hofmeyr et al. (2017). Our final data set contained concatenated sequences of the mitochondrial 12S (723 bp), 16S (532 bp), and ND4 (615 bp) genes plus sequences of the nuclear Prolactin (514 bp) and R35 (902 bp) loci, corresponding to a total length of 3,286 bp. However, the Prolactin locus did not amplify for our samples of *C. angulata*, despite repeated efforts, and was replaced in the alignment by Ns for calculations. The fossil-calibrated molecular clock was then rerun using the same calibration points and settings as in Hofmeyr et al. (2017).

2.4 | Species distribution models

Correlative species distribution models (SDMs) were computed based on occurrence records of our *Chersina angulata* samples and a set of environmental predictor variables. Separate SDMs were calculated for all *C. angulata* records and for subsets corresponding to the mitochondrial clades identified in the present study.

To prevent false inflation of model performance, each data set of localities was spatially rarefied to an Euclidian distance of 5 km using the SDMTOOLBOX 1.1c (Brown, 2014) in ARCGIS 10.3, retaining 149 unique records for *C. angulata* on the whole, 88 for the western 'blue clade', and 38 for the southern 'red clade'. A set of nine uncorrelated ($R^2 < .75$) bioclimatic variables with a spatial resolution of 2.5 arc-minutes (~4 km at the equator) was obtained for present climatic conditions from http://worldclim.org: annual mean temperature (bio 1), mean diurnal temperature range (bio 2), isothermality (bio 3), minimum temperature of the coldest month (bio 6), mean temperature of the wettest quarter (bio 8), annual precipitation (bio 12), precipitation of the driest month (bio 14), precipitation seasonality (bio 15), and precipitation of the coldest quarter (bio 19). To examine the impact of historical climate fluctuations and sea level changes on the distribution of *C. angulata* and its maternal lineages, we obtained corresponding predictor variables for three projections for the LGM (approx. 22,000 years ago = ya), derived from global circulation models through the Paleoclimate Modelling Intercomparison Project Phase II (Braconnot et al., 2012): namely the Community Climate System Model (CCSM3; Otto-Bliesner et al., 2006), the Max-Planck-Institute Earth System Model P (MPI-ESM-P), and the Model for Interdisciplinary Research on Climate (MIROC; Hasumi

& Emori, 2004). A circular buffer with a radius of 200 km surrounding each record was used for model training, while a rectangular area was used as projection background. Models were computed for each data set using the machine learning algorithm MAXENT 3.4.1 (Phillips, Anderson, & Schapire, 2006; Phillips & Dudík, 2008). A bootstrapping approach was applied using 100 replicates randomly subdividing the data set into 80% used for model training and 20% for model evaluation. The feature classes linear, quadratic, and hinge were selected. The resulting models were then projected onto the three LGM scenarios. Model performance was evaluated using the area under the curve (AUC; Swets, 1988). The average projection across all replicate runs was used for further processing, wherein the minimum training presence logistic threshold was applied as presence–absence threshold.

3 | RESULTS

3.1 | Mitochondrial DNA

Both tree-building methods revealed for the cyt *b* sequences weak structuring. The sequences clustered in two clades that were only under BA well supported (Figure S1). One clade (blue in Figure S1) is largely distributed in the western part of the distribution range, while the other clade (red) occurs mainly in the south. The distribution ranges of both clades abut and slightly overlap in the Western Cape Province of South Africa. Twenty-two samples from the population with large-sized tortoises on Dassen Island were included in the blue clade, without genetic signatures for differentiation. The two clades differed on average by an uncorrected *p* distance of 2.61%.
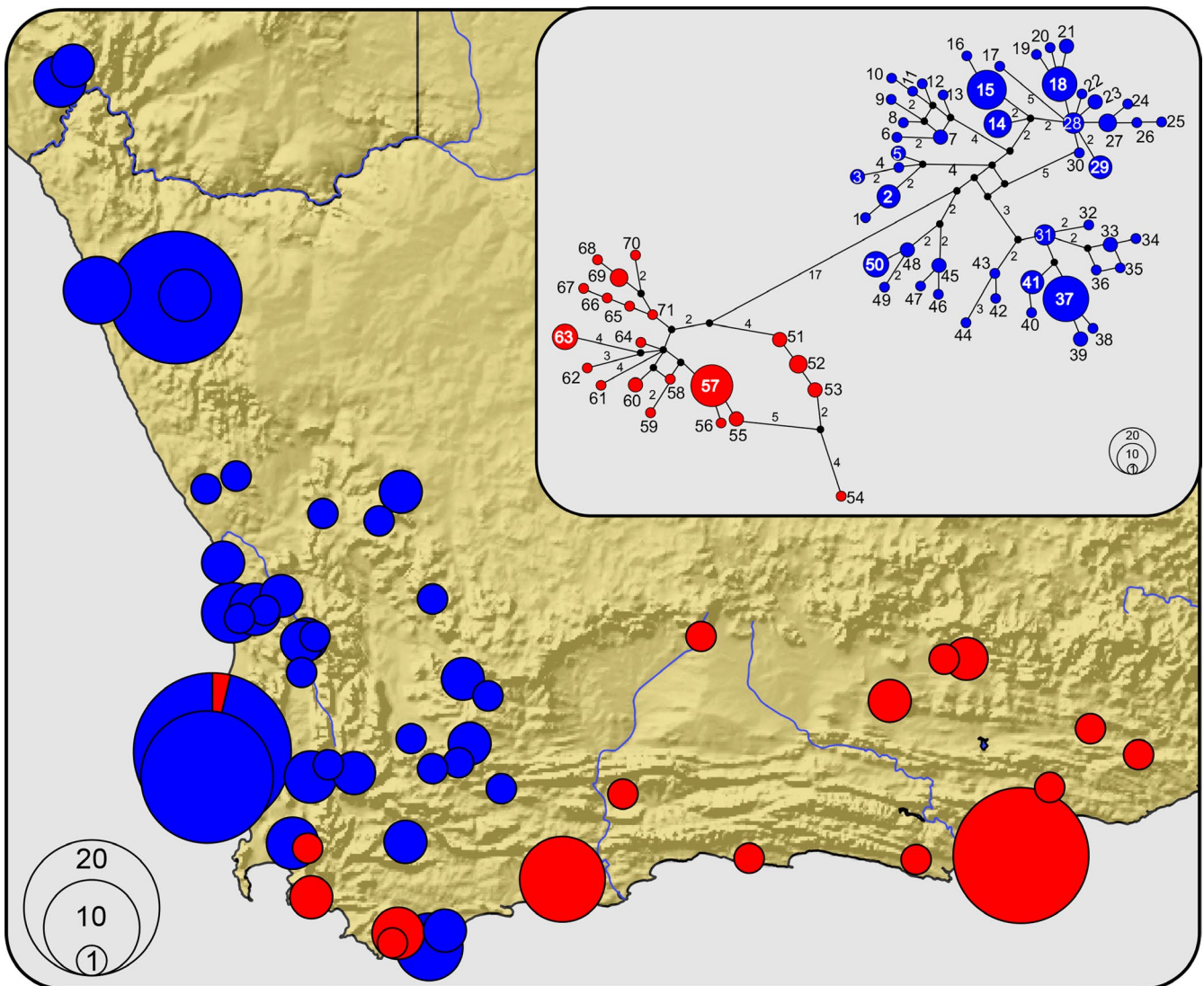

FIGURE 1 Geographic distribution of the cyt *b* clades of *Chersina angulata* (*n* = 180). Symbol sizes of sampling sites correspond to sample sizes. Slices (percentages) reflect syntopic occurrences. Inset: haplotype network (scale differs from that of the map). In the network, missing node haplotypes correspond to small black circles; lines connecting haplotypes represent one mutation step, if not otherwise indicated by numbers along lines [Colour figure can be viewed at wileyonlinelibrary.com]
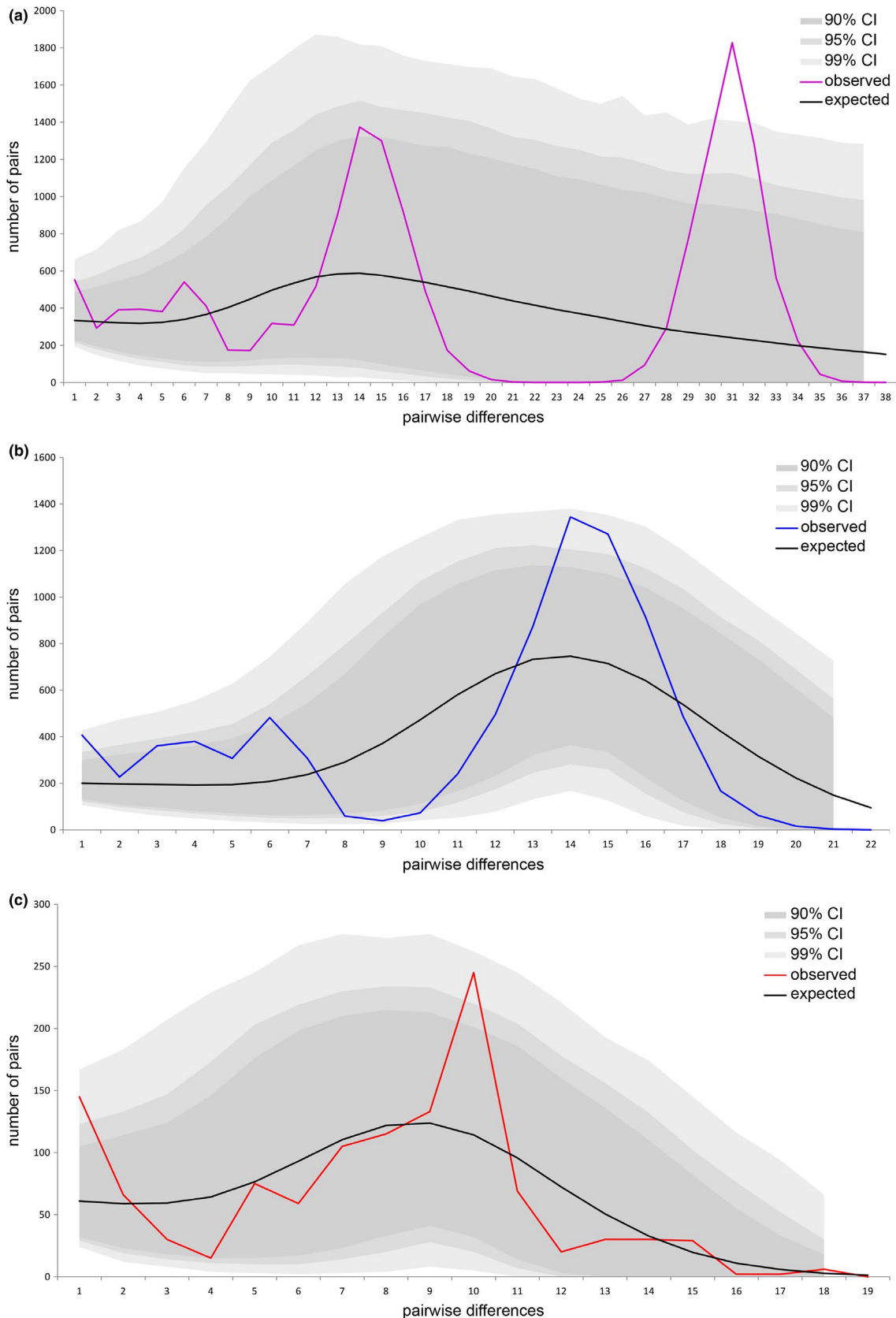
FIGURE 2 Observed and simulated mismatch distribution for mitochondrial sequences of (a) all *Chersina angulata* samples (*n* = 180), (b) samples of the blue clade (*n* = 131), and (c) samples of the red clade (*n* = 49). Coloured lines correspond to observed data. Black lines show expected mismatch distribution from simulated data under the demographic expansion model (10,000 bootstrap replicates). Grey areas are confidence intervals of simulated data [Colour figure can be viewed at wileyonlinelibrary.com]

The same genetic structure as in the trees was also reflected by the haplotype network (Figure 1). Two haplotype clusters corresponded to the two clades. Even though there was some variation in frequencies of individual haplotypes, no cluster showed a star-like structure with a common ancestral haplotype and rare tip haplotypes. Instead, intricate haplotype connections with several loops were found, suggestive of stationary populations that have not experienced recent expansion. This was supported by mismatch analyses (Figure 2) and population demography descriptors (Table 1). Mismatch analyses found multimodal frequency distributions for all samples together and for each clade. The sum of the squared deviations (SSD) from the demographic expansion model was non-significant only for the red clade, and the raggedness index (RI) was non-significant only for all samples lumped together, supporting stationary populations, in agreement with the network topology. Tajima's D values for all samples and the two clades were negative but not significantly different from zero, suggestive of drift-mutation equilibrium resulting from non-expanding population demographies. All values for Fu's F_S, an estimator for deviations from mutation-drift-equilibrium based on singleton mutations, were negative but not significant in the red clade. However, F_S values were significantly different from zero for pooled samples ($p < .05$) and the blue clade ($p < .01$), reflecting an excess of singletons in the blue clade. This excess of singletons of the blue clade is also obvious in the haplotype network (Figure 1) and from the generally higher genetic diversity estimates (Table 1).

3.2 | Microsatellite data

The ΔK approach revealed $K = 3$ as the most likely number of clusters (Figure S2). We present, however, also the results for $K = 2$ to compare how two clusters match with the mitochondrial structure (Figure 3). With respect to geographic distribution, there is an excellent agreement between mitochondrial clades and microsatellite clusters under $K = 2$, with some individuals in the geographic contact zone of the mitochondrial clades showing evidence for nuclear genomic admixture or cyto-nuclear discordance. However, within the distribution range of the 'blue microsatellite cluster', some tortoises were assigned to the 'red cluster' by STRUCTURE. Unfortunately, no mtDNA sequences could be generated for these samples. The tortoises from Dassen Island were assigned to the blue cluster.

Under $K = 3$, the red cluster remains unchanged, while the blue cluster is divided into two clusters, with higher percentages of admixture among all three clusters. The former blue cluster corresponds now to one cluster in the very north (dark blue), and another cluster (light blue) distributed in the remaining range of the blue cluster. Dassen Island tortoises occurred, together with tortoises from normal-sized populations from the adjacent mainland, in the light blue cluster.

Under $K = 2$, the two clusters have similar average numbers of alleles per locus and similar values for the allelic richness. The remaining diversity indices for microsatellites and mtDNA of the two clusters vary according to sample sizes (Table 2: top). The AMOVAs revealed that 6.34% of the molecular variance of the microsatellite data occurs between clusters and 93.66% within clusters. For the mtDNA, 51.32% occurs between and 48.68% within clusters (Table 3: left).

Under $K = 3$, diversity values for microsatellites and mtDNA varied according to sample size (Table 2; bottom), except for the observed heterozygosity, which was highest in the dark blue cluster having the lowest sample size. The AMOVAs revealed for $K = 3$ a similar picture as for $K = 2$, even though for mtDNA a higher percentage (60.21%) of the molecular variance occurred among clusters (Table 3: right).

The Principal Component Analyses (PCAs) for microsatellite data largely confirmed the STRUCTURE results, even though

	All samples (n = 180)	Red clade (n = 49)	Blue clade (n = 131)
Number of haplotypes n_{ht}	71	21	50
Polymorphic sites S	122	41	79
Mean pairwise differences π	17.8628	6.5476	10.3787
Gene diversity	0.9657 ± 0.0055	0.8767 ± 0.0388	0.9522 ± 0.0088
Nucleotide diversity	0.0156 ± 0.0077	0.0057 ± 0.0031	0.0091 ± 0.0046
Tajima's D	−0.4894 *n.s.*	−0.9844 *n.s.*	−0.9027 *n.s.*
Fu's F_S	−14.8381*	−3.9053 *n.s.*	−14.4141**
Sum of squared deviations SSD	0.032**	0.0242 *n.s.*	0.0226**
Raggedness Index RI	0.011 *n.s.*	0.0441*	0.0147*

TABLE 1 Genetic diversity and population dynamic indices for mtDNA (cyt *b* gene)

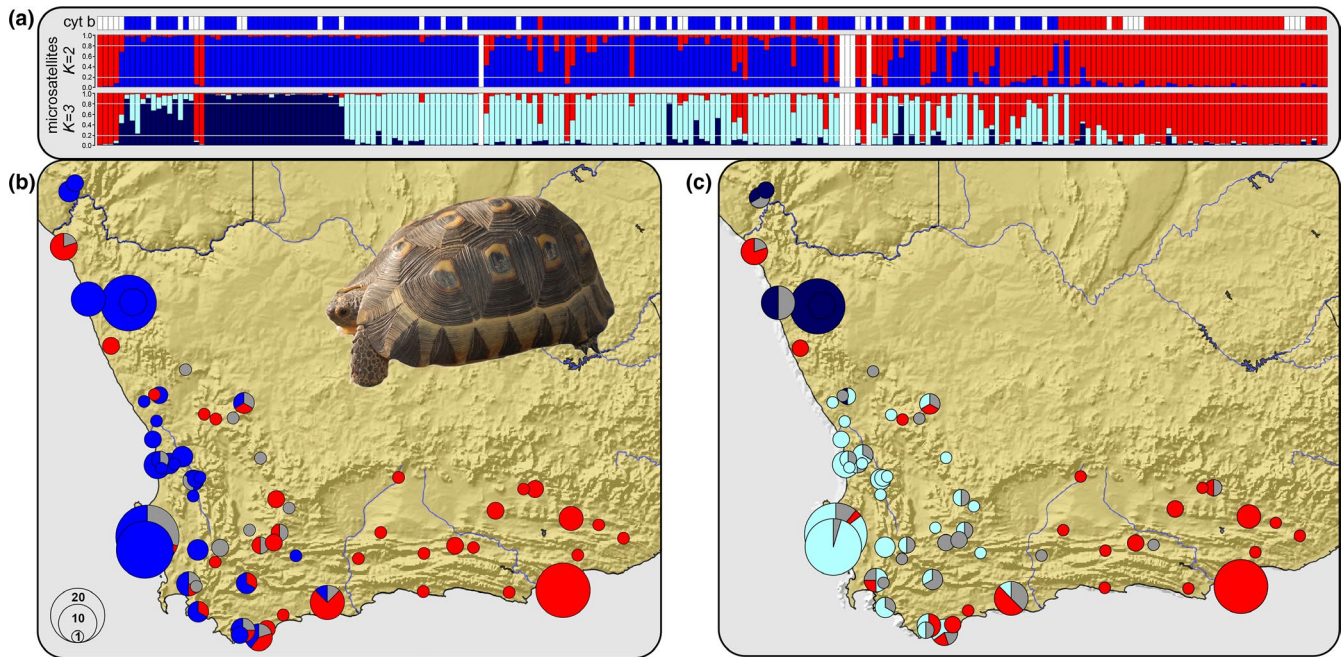Abbreviation: *n.s.*, not significant.

*$p < .05$; **$p < .01$.

FIGURE 3 Nuclear genomic clusters as inferred by STRUCTURE for *Chersina angulata* (*n* = 224) using 14 microsatellite loci under *K* = 2 and *K* = 3 (a) and their respective geographic distributions (b,c). Distinct clusters are colour-coded to match with mitochondrial clades; grey symbols and slices represent admixed tortoises. In A, each vertical segment represents an individual tortoise showing its assignment to a cluster; different colours in one segment indicate mixed ancestry (white = missing data). The mitochondrial clade for each individual is shown above STRUCTURE diagrams. White lines correspond to the 20% threshold for admixture. Inset: *Chersina angulata* (Photograph: M. Vamberger) [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 2 Genetic diversity of STRUCTURE clusters with admixed individuals excluded

Cluster	Microsatellites							mtDNA				
	n	n_A	$n_{\bar{A}}$	n_p	*AR*	H_O	H_E	*n*	n_{HT}	n_{pHT}	*h*	*π*
K = 2												
Red	76	228	16.29	62	15.96	0.542	0.731	52	27	20	0.894	0.0127
Blue	122	253	18.07	88	16.05	0.632	0.769	103	44	37	0.944	0.0098
K = 3												
Red	57	190	13.57	36	11.84	0.526	0.714	40	19	16	0.826	0.0085
Dark blue	36	155	11.07	13	11.00	0.661	0.725	29	7	6	0.608	0.0011
Light blue	92	238	17.00	72	13.41	0.612	0.758	81	41	38	0.958	0.0096

Abbreviations: *AR*, allelic richness; *h*, haplotype diversity; H_E, average expected heterozygosity; H_O, average observed heterozygosity; *n*, number of individuals; $n_{\bar{A}}$, average number of alleles per locus; n_A, number of alleles; n_{HT}, number of haplotypes per cluster; n_P, number of private alleles; n_{pHT}, number of private haplotypes per cluster; *π*, nucleotide diversity.

TABLE 3 AMOVA results for STRUCTURE clusters

	K = 2		*K* = 3	
	Microsatellites	mtDNA	Microsatellites	mtDNA
Between populations	6.34	51.32	7.91	60.21
Within populations	93.66	48.68	92.09	39.79

the 95% confidence intervals overlapped widely (Figure 4). In particular, the PCA for those individuals for which cyt *b* sequences and microsatellites were available revealed massive overlap, suggestive of weak differentiation. This is also supported by low F_{ST} values for microsatellite data (Table 4).

3.3 | Molecular clock

Our fossil-calibrated molecular clock (Figure S3) produced younger divergence estimates than in Hofmeyr et al. (2017),
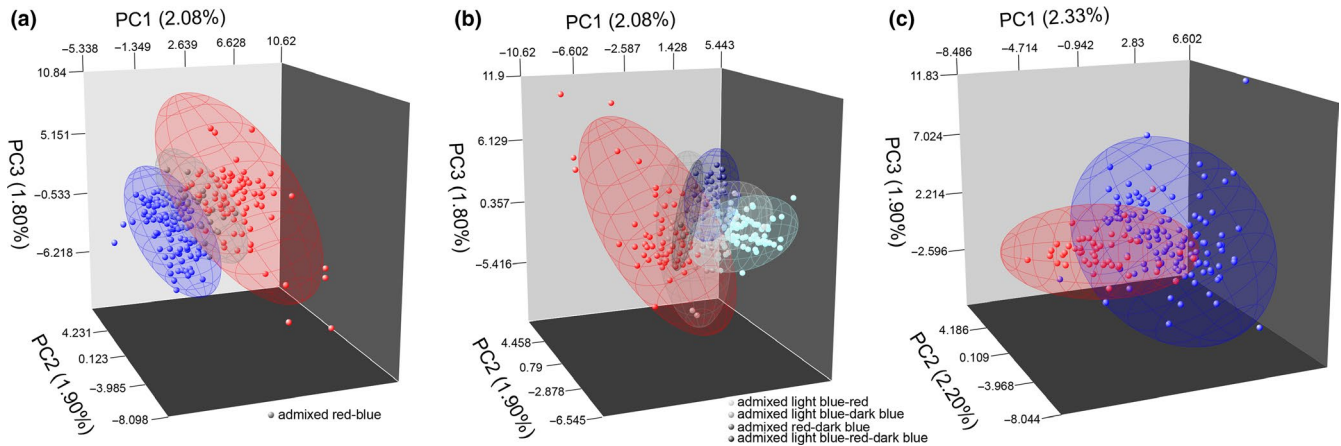
FIGURE 4 Principal Component Analyses (PCAs) for microsatellite data of *Chersina angulata*. (a) PCA using 224 angulate tortoises coloured according to the STRUCTURE clusters for $K = 2$; (b) PCA using 224 angulate tortoises coloured according to the STRUCTURE clusters for $K = 3$; (c) PCA using 175 angulate tortoises for which cyt *b* data are available, coloured according to mtDNA clades. Percentages of variance explained per components are shown along the axes. Spheres indicate 95% confidence intervals [Colour figure can be viewed at wileyonlinelibrary.com]

$K = 2$	Red	Blue	$K = 3$	Red	Dark blue	Light blue
Red	—	0.513*	Red	—	0.752*	0.609*
Blue	0.063*	—	Dark blue	0.103*	—	0.429*
			Light blue	0.075*	0.067*	—

TABLE 4 F_{ST} values for the STRUCTURE clusters. Above the diagonal are F_{ST} values for mtDNA; below the diagonal, for microsatellite data

*$p < .05$.

especially for the deeper nodes, even though the same settings and calibration points were used. The only difference compared to Hofmeyr et al. (2017) was that the three *Chersina angulata* from the original alignment were replaced by sequences for 12 other individuals. Nevertheless, the results of both studies were consistent in that the 95% HPD intervals overlapped widely, and the mean estimates fell within both intervals.

Our calculations suggested that *Chersina* branched off from its sister group *Chersobius* 26.35 (18.85–34.36) million years ago (mya). The two *Chersina* clades were inferred to have split 3.79 (2.18–5.82) mya. Divergence within the red clade was estimated to have commenced only 0.23 (0.01–0.70) mya, while the onset of the divergence within the blue clade was dated to 2.97 (1.73–4.53) mya.

3.4 | Species distribution models

For *Chersina angulata* as a whole, model performance under current climatic conditions was high ($AUC_{training} = 0.87$; $AUC_{test} = 0.81$), indicating that the model discriminates well between suitable and unsuitable space. Mean temperature of the wettest quarter (bio 8) had the highest contribution (43.9%) to the model, followed by precipitation of the coldest quarter (bio 19:23.4%), minimum temperature of the coldest month (bio 6:7.6%), mean diurnal temperature range (bio 2:7.2%), isothermality (bio 3:7.1%), and precipitation

seasonality (bio 15:6.8%). The remaining predictors contributed less than 5% to the model (Table S7).

The location and extent of habitat predicted as suitable under current climatic conditions matches well with the known distribution range. Suitable habitat is restricted to lower elevations south and west of the Great Escarpment, with highest suitability across the coastal plains. The model identifies a large continuous stretch of highly suitable habitat that extends from Kleinsee in the north-west to Vleesbaai and is predominantly restricted to the Succulent Karoo and Fynbos biomes. Two additional but much smaller areas of high suitability appear to be located in St Francis Bay and in the vicinity of Alexandria. While St Francis Bay falls into the Fynbos biome, Alexandria lies in a region covered by thicket vegetation. However, predicted suitability appears to decrease towards the southern and eastern distribution limits of the species (Figure 5).

Model performance for the respective LGM projections was equally high (CCSM3: $AUC_{training} = 0.87$; $AUC_{test} = 0.82$; MPI-ESM-P: $AUC_{training} = 0.86$; $AUC_{test} = 0.82$; MIROC: $AUC_{training} = 0.87$; $AUC_{test} = 0.82$). Variable contribution across all three LGM projections was similar to current conditions (Table S7).

During the LGM, the extent of suitable habitat and the extent of highly suitable space were slightly larger. A comparison of the SDM for current conditions with the three LGM projections suggests that areas with predicted high suitability
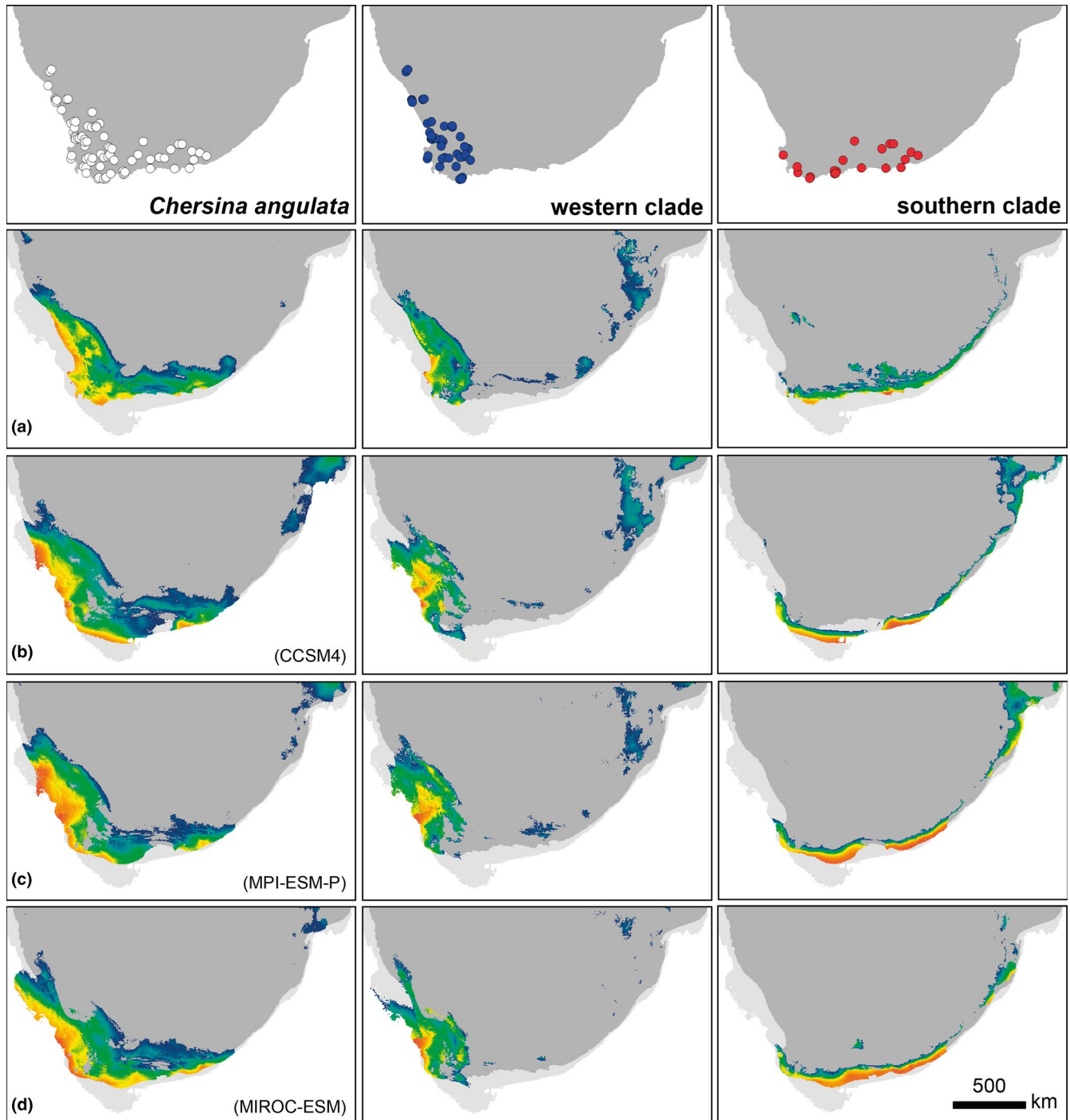
FIGURE 5   Current and past distribution of *Chersina angulata* and its two clades inferred by MAXENT. (a) Potential distribution under current environmental conditions; (b-d) projection onto paleoclimatic conditions of the Last Glacial Maximum derived from three global circulation models (CCSM4, MPI-ESM-P, and MIROC-ESM). Suitability ranges from moderate (dark blue) to high (red). Exposed land surface under current conditions in dark grey; land surface under LGM conditions, light grey [Colour figure can be viewed at wileyonlinelibrary.com]

have been continuous and extended along nowadays submerged parts of the coastal plain. Thus, suitable space shifted northwards with rising sea level.

Both for the western blue clade and the southern red clade, model performance was excellent under current conditions (blue clade: $AUC_{training} = 0.90$; $AUC_{test} = 0.84$; red clade: $AUC_{training} = 0.92$; $AUC_{test} = 0.84$). Variable contribution (Table S7) reveals the western blue clade, which inhabits the more arid Fynbos and Succulent Karoo biomes, to be strongly affected by precipitation. The southern red clade appears to be more influenced by temperature-related predictors. In contrast to the blue clade, the distribution of the

red clade extends from Fynbos into a mosaic of Thicket and Nama Karoo habitat, with the majority of occurrences falling into an area with non-seasonal precipitation and higher annual rainfall.

Suitable space for the blue clade appears to be restricted to the south-west of the range, with highest suitability across the coastal plains between Cape Town and Rietpoort. In contrast, for the red clade, suitable space was predicted for a narrow strip along the southern coastline, with patches of highest suitability around Cape Agulhas and the De Hoop Nature Reserve.

Model performances for the LGM projections for the two clades were excellent (blue clade: CCSM3: $AUC_{training} = 0.90$; $AUC_{test} = 0.83$; MPI-ESM-P: $AUC_{training} = 0.90$; $AUC_{test} = 0.84$; MIROC: $AUC_{training} = 0.90$; $AUC_{test} = 0.84$; red clade: CCSM3: $AUC_{training} = 0.92$; $AUC_{test} = 0.84$; MPI-ESM-P: $AUC_{training} = 0.92$; $AUC_{test} = 0.83$; MIROC: $AUC_{training} = 0.92$; $AUC_{test} = 0.84$). For both mitochondrial clades, variable contributions for the LGM projections resembled those under current conditions (Table S7).

When the SDMs for current and LGM conditions are compared, it is obvious that the overall extent of suitable habitat for both clades did not alter much since the LGM. However, areas with high suitability seem to have contracted and shifted northwards, which was probably correlated with the rising sea level.

4 | DISCUSSION

Chersina angulata is a well-known and charismatic tortoise species of Southern Africa. Yet, its phylogeography and taxonomy are poorly understood. The species was traditionally regarded as monotypic and the only species of the *Chersina* genus (Ernst & Barbour, 1989; Hofmeyr, 2009; Wermuth & Mertens, 1961, 1977). However, Daniels et al. (2007) suggested that two distinct species could exist, based on the putative Miocene divergence (10.4–8.4 mya) of the two mitochondrial clades identified in this study. Our understanding of distinct species is close to the Biological Species Concept of Mayr (1942) in that no large-scale gene flow should occur between species, or in other terms, there should be evidence for the presence of effective reproductive barriers impeding extensive gene flow. Thus, if the two mitochondrial clades of *C. angulata* represent distinct species, we expect nuclear genomic differentiation that varies concordantly with mitochondrial divergence, without signals for large-scale gene flow. Nuclear genomic evidence to examine differentiation and gene flow was until now lacking.

Our present study confirmed the existence of two mitochondrial clades within *C. angulata*. With respect to the cyt *b* gene, the average uncorrected *p* distance between the two clades (2.6%) falls into the range observed within other tortoise species (0.9%–3.4%), while between distinct congeneric tortoise species average values of 3.7%–12.7% are observed (see the reviews in Fritz et al., 2012 and Kindler et al., 2012). The two clades of *C. angulata* occupy different, largely mutually exclusive climatic niches (Figure 5). We inferred for these clades a much younger divergence time than Daniels et al. (2007). The estimate of Daniels et al. (2007) was based on mtDNA sequences and the a priori assumption that *Chersina* diverged from its sister group *Chersobius* 18 mya, which is a weakly justified prior derived from a previous molecular clock using constant sequence divergence rates (Cunningham, 2002). In addition, Daniels et al. (2007) included only a few outgroup taxa. Using representatives of all African tortoise genera and a combination of mitochondrial and nuclear loci, both Hofmeyr et al. (2017) and we estimated that *Chersina* and *Chersobius* diverged much earlier (29.51 mya or 26.35 mya, respectively). Despite this older divergence, a more recent age for the radiation within *Chersina angulata* was inferred (Hofmeyr et al., 2017: 3.32 mya; this study: 3.79 mya).

Both STRUCTURE analyses and PCAs (Figures 3 and 4) revealed nuclear genomic clusters whose geographic distributions largely agree with those of the mitochondrial clades, supporting that the two mitochondrial clades represent distinct genealogical lineages. One of the two lineages ('blue') includes our tortoises from the large-sized population on Dassen Island (Table S1). The 22 Dassen Island tortoises were genetically not differentiated from populations with normal-sized individuals from the adjacent mainland. This finding is in agreement with other studies that reported considerable phenotypic plasticity in tortoises with respect to maximum adult size (Fritz et al., 2007, 2010, 2012; Fritz, Široký, Kami, & Wink, 2005; Spitzweg et al., 2019).

Our microsatellite analyses indicate substructuring for the western 'blue lineage', with a northern and southern cluster within its range. This is in agreement with genetic diversity descriptors for the two lineages, revealing for mtDNA data higher variation in the blue cluster (Table 2). In the contact zone of the 'blue' and the 'red lineages', we found clear signatures of gene flow, with admixed tortoises and individuals showing cyto-nuclear discordance. In addition, we found several mismatched tortoises representing the southern red lineage deep in the distribution range of the blue lineage (Figure 3), without evidence for admixture. Unfortunately, for most of these mismatched individuals, we could generate only the microsatellite data. Repeated efforts to sequence the cyt *b* gene of these samples failed. The occurrence of these tortoises is difficult to explain. It is well known that tortoises, in particular angulate tortoises, are translocated across wide distances in South Africa (Fritz et al., 2010; Spitzweg et al., 2019), and we recorded in a previous study a *C. angulata* in KwaZulu-Natal (Spitzweg et al., 2019), far outside its native distribution range. Consequently, we cannot exclude that

the genetically mismatched tortoises revealed in the present study are another example of translocation. Nevertheless, the extensive nuclear genomic admixture and cyto-nuclear discordance in the contact zone of the blue and red lineages clearly reject the hypothesis of two distinct species, in line with the low uncorrected *p* distances of the cyt *b* gene. Thus, *C. angulata* embraces two distinct, but conspecific, genealogical lineages that are still capable of large-scale gene flow and that hybridize extensively along their geographic contact zone. Each lineage qualifies for an Evolutionarily Significant Unit (Moritz, 1994) that is confirmed by two independent genetic lines of evidence (mitochondrial and nuclear DNA), while the clusters within the blue lineage, different in the studied microsatellite markers but not in mtDNA, can be recognized as two Management Units in the sense of Moritz (1994).

Recently, Kindler and Fritz (2018) argued that Evolutionarily Significant Units should be nomenclaturally recognized as distinct subspecies because this facilitates communication within and beyond science. This is of particular value for conservation. Even though we fully endorse this point of view, the subspecies of *C. angulata* cannot be named at present. Three scientific names can be unambiguously identified with this species: *Testudo angulata* Schweigger, 1812, *Testudo bellii* Gray, 1828, and *Chersina angulata pallida* Gray, 1831 (TTWG, 2017). A type locality (Cape of Good Hope) is only known for *T. bellii*, while the two other taxa were described without any knowledge of the provenance of the type material. Even worse, 'Cape of Good Hope' was in the early 19th century a general term for the 'Cape Colony', later named the Cape Province, a region that includes both of the two genealogical lineages of *C. angulata*. Thus, it remains completely unclear to which taxon any name refers. This could be only clarified when the type material, putatively in the Muséum national d'Histoire naturelle, Paris, and the Natural History Museum, London, is located and genetically characterized using cutting-edge approaches for historical DNA that have been successfully applied for other type material (e.g., Kehlmaier, Zhang, et al., 2019).

The two unnamed subspecies of *C. angulata* occupy climatic niches that are clearly different but overlap in the Fynbos biome of the south-western Cape region (Figure 5), corresponding to their hybrid zone. The southern subspecies appears to tolerate a broader climatic range spanning several biomes, while the western subspecies seems to be restricted to the Fynbos and Succulent Karoo biomes. Our LGM projections indicated that the climatic niches of *C. angulata* extended then onto the exposed shelf and have not changed much since the LGM, in agreement with demographically static populations as suggested by several analyses. However, in response to the raising sea level, the distribution range of the southern subspecies shifted north during the Holocene, which may have fostered admixture with the western taxon.

The less dynamic distribution range of the western subspecies, on the other hand, could explain the excess of singleton mutations in mtDNA of the blue clade (Table 1) because such mutations had more time to accumulate locally in a stable distribution range.

In any case, the geographic ranges of the two subspecies resemble the distribution of genetic lineages within two other turtle species, *Pelomedusa galeata* (Vamberger, Hofmeyr, Ihlow, & Fritz, 2018) and *Stigmochelys pardalis* (Spitzweg et al., 2019), suggestive of a shared phylogeographic history in these taxa.

5 | CONCLUSIONS

Our study shows that the two previously identified mitochondrial clades of the angulate tortoise (*Chersina angulata*) represent genealogical lineages that also differ with respect to nuclear genomic markers, their environmental preferences, and associated habitat requirements. According to species distribution models, the ranges of the two lineages shifted little since the LGM, which is in agreement with the demographic population descriptors suggestive of stationary populations that did not experience expansion. One of the lineages is distributed in the west, and the other in the south of the distribution range. The western lineage is substructured in two population clusters that differ with respect to microsatellite loci, but not in mtDNA. In the contact zone of the two lineages, extensive gene flow provides evidence for their conspecificity under the Biological Species Concept. Both lineages could be recognized as distinct subspecies of *C. angulata*, but the ill-defined geographic origins of the type material for the three available names prevent their identification with any of the two lineages. Further research should examine whether there are consistent morphological differences matching the genetic differentiation. In addition, the type material should be located and studied genetically to reveal to which subspecies the available names refer. Confiscated or captive angulate tortoises of unknown geographic provenance should not be released without prior genetic screening to avoid genetic pollution of wild populations.

ORCID

Cäcilia Spitzweg (iD) https://orcid.org/0000-0001-9297-7633
Melita Vamberger (iD) https://orcid.org/0000-0002-1404-2469
Flora Ihlow (iD) https://orcid.org/0000-0002-0460-4210
Uwe Fritz (iD) https://orcid.org/0000-0002-6740-7214
Margaretha D. Hofmeyr (iD) https://orcid.org/0000-0002-4963-6750

REFERENCES

Adler, D., Murdoch, D., Nenadic, O., Urbanek, S., Chen, M., Gebhardt, A., …Senger, A. (2019). *RGL: 3D visualization using OpenGL. R Package Version, 0.100.19. Retrieved from* https://CRAN.R-project.org/package=rgl

Barilani, M., Sfougaris, A., Giannakopoulos, A., Mucci, N., Tabarroni, C., & Randi, E. (2007). Detecting introgressive hybridisation in rock partridge populations (*Alectoris graeca*) in Greece through Bayesian admixture analyses of multilocus genotypes. *Conservation Genetics*, *8*, 343–354. https://doi.org/10.1007/s10592-006-9174-1

Braconnot, P., Harrison, S. P., Kageyama, M., Bartlein, P. J., Masson-Delmotte, V., Abe-Ouchi, A., … Zhao, Y. (2012). Evaluation of climate models using palaeoclimatic data. *Nature Climate Change*, *2*, 417–424. https://doi.org/10.1038/nclimate1456

Brown, J. L. (2014). SDMTOOLBOX: A python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *Methods in Ecology and Evolution*, *5*, 694–700. https://doi.org/10.1111/2041-210X.12200

Buhlmann, K. A., Akre, T. S. B., Iverson, J. B., Karapatakis, D., Mittermeier, R. A., Georges, A., … Gibbons, J. W. (2009). A global analysis of tortoise and freshwater turtle distributions with identification of priority conservation areas. *Chelonian Conservation and Biology*, *8*, 116–149. https://doi.org/10.2744/CCB-0774.1

Clement, M., Posada, D., & Crandall, K. A. (2000). TCS: A computer program to estimate gene genealogies. *Molecular Ecology*, *9*, 1657–1659. https://doi.org/10.1046/j.1365-294x.2000.01020.x

Cunningham, J. (2002). *A molecular perspective on the family Testudinidae Batsch, 1788*. Doctoral Thesis, Cape Town, South Africa: University of Cape Town.

Daniels, S. R., Hofmeyr, M. D., Henen, B. T., & Crandall, K. A. (2007). Living with the genetic signature of Miocene induced change: Evidence from the phylogeographic structure of the endemic angulate tortoise *Chersina angulata*. *Molecular Phylogenetics and Evolution*, *45*, 915–926. https://doi.org/10.1016/j.ympev.2007.08.010

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*, 359–361. https://doi.org/10.1007/s12686-011-9548-7

Ernst, C. H., & Barbour, R. W. (1989). *Turtles of the world*. Washington, DC: Smithsonian Institution Press.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, *14*, 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier, L., & Lischer, H. E. L. (2010). ARLEQUIN suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*, 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Fritz, U., Alcalde, L., Vargas-Ramírez, M., Goode, E. V., Fabius-Turoblin, D. U., & Praschag, P. (2012). Northern genetic richness and southern purity, but just one species in the *Chelonoidis chilensis* complex. *Zoologica Scripta*, *41*, 220–232. https://doi.org/10.1111/j.1463-6409.2012.00533.x

Fritz, U., Daniels, S. R., Hofmeyr, M. D., González, J., Barrio-Amorós, C. L., Široký, P., … Stuckas, H. (2010). Mitochondrial phylogeography and subspecies of the wide-ranging sub-Saharan leopard tortoise *Stigmochelys pardalis* (Testudines: Testudinidae) – a case study for the pitfalls of pseudogenes and GenBank sequences. *Journal of Zoological Systematics and Evolutionary Research*, *48*, 348–359. https://doi.org/10.1111/j.1439-0469.2010.00565.x

Fritz, U., Hundsdörfer, A. K., Široký, P., Auer, M., Kami, H., Lehmann, J., … Wink, M. (2007). Phenotypic plasticity leads to incongruence between morphology-based taxonomy and genetic differentiation in western Palaearctic tortoises (*Testudo graeca* complex; Testudines, Testudinidae). *Amphibia-Reptilia*, *28*, 97–121. https://doi.org/10.1163/156853807779799135

Fritz, U., Široký, P., Kami, H., & Wink, M. (2005). Environmentally caused dwarfism or a valid species – Is *Testudo weissingeri* Bour, 1996 a distinct evolutionary lineage? New evidence from mitochondrial and nuclear genomic markers. *Molecular Phylogenetics and Evolution*, *37*, 389–401. https://doi.org/10.1016/j.ympev.2005.03.007

Glaubitz, J. C. (2004). CONVERT: A user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. *Molecular Ecology Notes*, *4*, 309–310. https://doi.org/10.1111/j.1471-8286.2004.00597.x

Goudet, J. (1995). FSTAT (version 1.2): A computer program to calculate *F*-statistics. *Journal of Heredity*, *86*, 485–486. https://doi.org/10.1093/oxfordjournals.jhered.a111627

Hall, T. A. (1999). BIOEDIT: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, *41*, 95–98.

Hasumi, H., & Emori, S., Eds. (2004). *K-1 coupled GCM (MIROC) description*. Tokyo, Japan: Center for Climate System Research, University of Tokyo (K-1 Technical Report No. 1).

Hofmeyr, M. D. (2009). *Chersina angulata* (Schweigger 1812) – angulate tortoise, South African bowsprit tortoise. In A. G. J. Rhodin, P. C. H. Pritchard, P. P. van Dijk, R. A. Saumure, K. A. Buhlmann, J. B. Iverson, & R. A. Mittermeier (Eds.), *Conservation biology of freshwater turtles and tortoises: A compilation project of the IUCN/SSC Tortoise and Freshwater Turtle Specialist Group* (pp. 030.1–030.6). Lunenburg, MA: Chelonian Research Foundation (Chelonian Research Monographs No. 5).

Hofmeyr, M. D., & Baard, E. H. W. (2014). *Chersina angulata* (Schweigger, 1812). In M. F. Bates, W. R. Branch, A. M. Bauer, M. Burger, J. Marais, G. J. Alexander, & M. S. de Villiers (Eds.), *Atlas and red list of the reptiles of South Africa, Lesotho and Swaziland* (p. 71). Pretoria, South Africa: South African National Biodiversity Institute (SANBI).

Hofmeyr, M. D., Vamberger, M., Branch, W., Schleicher, A., & Daniels, S. R. (2017). Tortoise (Reptilia, Testudinidae) radiations in Southern Africa from the Eocene to the present. *Zoologica Scripta*, *46*, 389–400. https://doi.org/10.1111/zsc.12223

Hubisz, M. J., Falush, D., Stephens, M., & Pritchard, J. K. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, 9, 1322–1332. https://doi.org/10.1111/j.1755-0998.2009.02591.x

Jombart, T. (2008). ADEGENET: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24, 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Kehlmaier, C., Barlow, A., Hastings, A. K., Vamberger, M., Paijmans, J. L. A., Steadman, D. W., … Fritz, U. (2017). Tropical ancient DNA reveals relationships of the extinct Bahamian giant tortoise *Chelonoidis alburyorum*. *Proceedings of the Royal Society B*, 284, 20162235. https://doi.org/10.1098/rspb.2016.2235

Kehlmaier, C., Graciá, E., Campbell, P. D., Hofmeyr, M. D., Schweiger, S., Martínez-Silvestre, A., … Fritz, U. (2019). Ancient mitogenomics clarifies radiation of extinct Mascarene giant tortoises (*Cylindraspis* spp.). *Scientific Reports*, 9, 17487. https://doi.org/10.1038/s41598-019-54019-y

Kehlmaier, C., Zhang, X., Georges, A., Campbell, P. D., Thomson, S., & Fritz, U. (2019). Mitogenomics of historical type specimens of Australasian turtles: Clarification of taxonomic confusion and old mitochondrial introgression. *Scientific Reports*, 9, 5841. https://doi.org/10.1038/s41598-019-42310-x

Kindler, C., Branch, W. R., Hofmeyr, M. D., Maran, J., Široký, P., Vences, M., … Fritz, U. (2012). Molecular phylogeny of African hinge-back tortoises (*Kinixys*): Implications for phylogeography and taxonomy (Testudines: Testudinidae). *Journal of Zoological Systematics and Evolutionary Research*, 50, 192–201. https://doi.org/10.1111/j.1439-0469.2012.00660.x

Kindler, C., & Fritz, U. (2018). Phylogeography and taxonomy of the barred grass snake (*Natrix helvetica*), with a discussion of the subspecies category in zoology. *Vertebrate Zoology*, 68, 269–281.

Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., & Stamatakis, A. (2019). RAxML-NG: A fast, scalable, and user-friendly tool for Maximum Likelihood phylogenetic inference. *Bioinformatics*, 35, 4453–4455. https://doi.org/10.1093/bioinformatics/btz305

Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., & Calcott, B. (2017). PARTITIONFINDER 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, 34, 772–773. https://doi.org/10.1093/molbev/msw260

Lesia, M. G., Hofmeyr, M. D., & D'Amato, M. E. (2003). Genetic variation in three C*hersina angulata* (angulate tortoise) populations along the west coast of South Africa. *African Zoology*, 38, 109–117. https://doi.org/10.1080/15627020.2003.11657198

Librado, P., & Rozas, J. (2009). DNASP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25, 1451–1452. https://doi.org/10.1093/bioinformatics/btp187

Mayr, E. (1942). *Systematics and the origin of species from the viewpoint of a zoologist*. New York, NY: Columbia University Press.

Mittermeier, R. A., van Dijk, P. P., Rhodin, A. G. J., & Nash, S. D. (2015). Turtle hotspots: An analysis of the occurrence of tortoises and freshwater turtles in biodiversity hotspots, high-biodiversity wilderness areas, and turtle priority areas. *Chelonian Conservation and Biology*, 14, 2–10. https://doi.org/10.2744/ccab-14-01-2-10.1

Moritz, C. (1994). Defining 'Evolutionarily Significant Units' for conservation. *Trends in Ecology & Evolution*, 9, 373–375. https://doi.org/10.1016/0169-5347(94)90057-4

Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, 403, 853–858. https://doi.org/10.1038/35002501

Otto-Bliesner, B. L., Brady, E. C., Clauzet, G., Tomas, R., Levis, S., & Kothavala, Z. (2006). Last Glacial Maximum and Holocene climate in CCSM3. *Journal of Climate*, 19, 2526–2544. https://doi.org/10.1175/JCLI3748.1

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190, 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Phillips, S.J., & Dudík, M. (2008). Modeling of species distributions with MAXENT: New extensions and a comprehensive evaluation. *Ecography*, 31, 161–175. https://doi.org/10.1111/j.0906-7590.2008.5203.x

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945–959.

Puechmaille, S. J. (2016). The program STRUCTURE does not reliably recover the correct population structure when sampling is uneven: Subsampling and new estimators alleviate the problem. *Molecular Ecology Resources*, 16, 608–627. https://doi.org/10.1111/1755-0998.12512

Randi, E. (2008). Detecting hybridization between wild species and their domesticated relatives. *Molecular Ecology*, 17, 285–293. https://doi.org/10.1111/j.1365-294X.2007.03417.x

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., … Huelsenbeck, J. P. (2012). MRBAYES 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61, 539–542. https://doi.org/10.1093/sysbio/sys029

Rosenberg, N. A. (2004). DISTRUCT: A program for the graphical display of population structure. *Molecular Ecology Notes*, 4, 137–138. https://doi.org/10.1046/j.1471-8286.2003.00566.x

Spinks, P. Q., Shaffer, H. B., Iverson, J. B., & McCord, W. P. (2004). Phylogenetic hypotheses for the turtle family Geoemydidae. *Molecular Phylogenetics and Evolution*, 32, 164–182. https://doi.org/10.1016/j.ympev.2003.12.015

Spitzweg, C., Hofmeyr, M. D., Fritz, U., & Vamberger, M. (2019). Leopard tortoises in southern Africa have greater genetic diversity in the north than in the south (Testudinidae). *Zoologica Scripta*, 48, 57–68. https://doi.org/10.1111/zsc.12328

Stuart, B. L., & Fritz, U. (2008). Historical DNA from museum type specimens clarifies diversity of Asian leaf turtles (*Cyclemys*). *Biological Journal of the Linnean Society*, 94, 131–141. https://doi.org/10.1111/j.1095-8312.2008.00966.x

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science*, 240, 1285–1293.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution*, 30, 2725–2729. https://doi.org/10.1093/molbev/mst197

TTWG [Turtle Taxonomy Working Group] (2017). *Turtles of the world. Annotated checklist and atlas of taxonomy, synonymy, distribution, and conservation status* (8thed.). Lunenburg, MA: Chelonian Research Foundation and Turtle Conservancy (Chelonian Research Monographs No. 7).

Vamberger, M., Hofmeyr, M. D., Ihlow, F., & Fritz, U. (2018). In quest of contact: Phylogeography of helmeted terrapins (*Pelomedusa galeata*, *P. subrufa* sensu stricto). *PeerJ*, 6, e4901. https://doi.org/10.7717/peerj.4901

Vamberger, M., Stuckas, H., Sacco, F., D'Angelo, S., Arculeo, M., Cheylan, M., … Fritz, U. (2015). Differences in gene flow in a twofold secondary contact zone of pond turtles in southern Italy

(Testudines: Emydidae: *Emys orbicularis galloitalica*, *E. o. hellenica*, *E. trinacris*). *Zoologica Scripta*, *44*, 233–249. https://doi.org/10.1111/zsc.12102

van den Berg, P., & Baard, E. H. W. (1994). Regional variation in morphometric characters in the angulate tortoise, *Chersina angulata*, from South Africa. *Journal of the Herpetological Association of Africa*, *43*, 28–32. https://doi.org/10.1080/04416651.1994.9650378

van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., & Shipley, P. (2004). MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, *4*, 535–538. https://doi.org/10.1111/j.1471-8286.2004.00684.x

Wermuth, H., & Mertens, R. (1961). *Schildkröten, Krokodile, Brückenechsen*. Jena, Germany: VEB Gustav Fischer.

Wermuth, H., & Mertens, R. (1977). *Testudines, Crocodylia, Rhynchocephalia*. Das Tierreich, 100. Berlin, Germany: Walter de Gruyter.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.