Received: 12 April 2019 | Revised: 29 July 2019 | Accepted: 6 August 2019

DOI: 10.1111/jzs.12338

ORIGINAL ARTICLE

JOURNAL OF ZOOLOGICAL SYSTEMATICS AND EVOLUTIONARY RESEARCH

WILEY

Unraveling the diversification and systematic puzzle of the highly polymorphic *Psammobates tentorius* (Bell, 1828) complex (Reptilia: Testudinidae) through phylogenetic analyses and species delimitation approaches

Zhongning Zhao[1] | Neil Heideman[1] | Paul Grobler[2] | Adriaan Jordaan[1] | Phillip Bester[3] | Margaretha D. Hofmeyr[4]

[1]Department of Zoology and Entomology, University of the Free State, Bloemfontein, South Africa

[2]Department of Genetics, University of the Free State, Bloemfontein, South Africa

[3]Department of Virology, University of the Free State and National Health Laboratory Service (NHLS), Bloemfontein, South Africa

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

Correspondence
Zhongning Zhao, Department of Zoology and Entomology, University of the Free State, Biology building B19, 205 Nelson Mandela Dr, Park West, Bloemfontein, South Africa.
Email: orochi19851020@yahoo.com

Funding information
University of the Western Cape; National Research Foundation, Grant/Award Number: IFR150216114248; Universiteit van die Vrystaat, Grant/Award Number: A1999/158110

Abstract

The high level of phenotypic diversity in southern African tent tortoises (*Psammobates tentorius* complex) has for decades prevented systematists from developing a stable taxonomy for the group. Here, we used a comprehensive DNA sequence dataset (mtDNA: *Cytb*, *ND4*, *ND4* adjacent *tRNA-His*, and *tRNA-Ser*, *12S*, *16S*; and nDNA: *PRLR* gene) of 455 specimens, and the latest phylogenetic and species delimitation analytical procedures, to unravel the long-standing *P. tentorius* complex systematic puzzle. Our results for mtDNA and nDNA were incongruent, with the poorly supported nDNA phylogeny differentiating the three recognized subspecies, and showing potential hybridization in some regions. In contrast, the concatenated mtDNA phylogeny identified seven operational taxonomic units, with strong support. Clades 1, 4, 5, and 7 corresponded to tortoises identified as *P. t. tentorius*, clade 3 to *P. t. trimeni*, and clades 2 and 6 to *P. t. verroxii*. Our analyses showed conflicting topologies for the placement of C6 (*P. t. verroxii* north of the Orange River), with stronger support for it being sister to C2 + C3 than to the other clades. Clades 1, 2, and 6 had significantly higher genetic diversity than clades 3, 4, 5, and 7, perhaps because these clades inhabit substantially larger areas. The potential for future cladogenic radiations seems high in C1 and C6, particularly in C6 for which the within-clade diversification level was highest. Further research involving microsatellite DNA, phylogeographic evaluations, and morphological variation among clades is crucial for understanding the adaptive radiation of the *P. tentorius* complex and for modifying their taxonomy.

KEYWORDS
mtDNA, phylogeny, reptile, southern Africa, tent tortoise

Contributing authors: Neil Heideman (heidemannj@ufs.ac.za), Paul Grobler (GroblerJP@ufs.ac.za), Adriaan Jordaan (JordaanA2@ufs.ac.za), Phillip Bester (BesterPA@ufs.ac.za) and Margaretha D. Hofmeyr (mdhofmeyr@gmail.com)

1 | INTRODUCTION

Southern Africa is recognized worldwide as a biodiversity hotspot for genera and species of the family Testudinidae (Branch, 2008; Cunningham, 2002; Fritz & Bininda-Emonds, 2007; Hofmeyr, Boycott, & Baard, 2014). It is home to 14 species of tortoises in six different genera, with strong indications that the current taxonomy, based on morphology, underestimates diversity (Hofmeyr, Vamberger, Branch, Schleicher, & Daniels, 2017). Using mitochondrial and nuclear DNA (mtDNA and nDNA), Kindler et al. (2012) distinguished eight sub-Saharan *Kinixys* species in contrast to six previously recognized by morphology. The importance of molecular studies to reveal chelonian diversity is exemplified by the freshwater turtle *Pelomedusa*, formerly considered monotypic (*P. subrufa*), but recently shown to consist of 10 species with at least five more candidate species (Vamberger, Hofmeyr, Ihlow, & Fritz, 2018; Vargas-Ramírez et al., 2010).

Several species delimitation approaches based on Bayesian algorithms have become available for clarifying species boundaries (Jones, Aydin, & Oxelman, 2014; Reid, 2014; Yang & Rannala, 2010; Zhang, Kapli, Pavlidis, & Stamatakis, 2013). These are particularly valuable when species boundaries are diffuse (species complexes), especially if they help to reach agreement on the standardization of criteria for delineating operational taxonomic units (OTUs, Caron et al., 2009) and unifying species concepts, which have been controversial for decades (Aldhebiani, 2018; De Queiroz, 2007; Kunz, 2013; Scudder, 1974). Since OTUs are fundamental units of biodiversity (Balakrishnan, 2005; Cracraft, 1983; De Queiroz, 2007; Hull, 1977; Mysara et al., 2017) and play ecologically important roles as indicators of dynamic changes to their environment (Caron et al., 2009; Preheim, Perrotta, Martin-Platero, Gupta, & Alm, 2013), using molecular phylogenies to delineate various aspects of OTUs have become critically important to modern conservation management.

These modern technological and methodological advances have greatly strengthened our ability to answer long-standing phylogenetic and taxonomic questions, and seem appropriate to investigate the taxonomic uncertainty of the southern African tent tortoises (*Psammobates tentorius* complex). The high level of phenotypic diversity shown by this species has prevented systematists for decades from developing a stable taxonomy for the group. Tent tortoises from South Africa and southern Namibia are morphologically highly variable with respect to color patterns, body shape, and the smoothness of their scutes. This led to the description of many species and subspecies in the past (Hewitt, 1933, 1934). In their taxonomic revision of African tortoises and turtles, Loveridge and Williams (1957) synonymized most *P. tentorius* taxa to recognize only one species with three subspecies: *P. t. tentorius*, *P. t. trimeni*, and *P. t. verroxii* (Figure S1). This taxonomy was accepted by most subsequent authors (Boycott & Bourquin, 2000; Branch, 1998, 2008; Branch, Benn, & Lombard, 1995; Greig & Burdett, 1976; Hofmeyr et al., 2014) although some of these authors indicated that there may be unresolved taxonomic issues. A recent phylogenetic evaluation of southern African tortoises identified four lineages for *P. tentorius*, of which two occur within *P. t. verroxii* (Hofmeyr et al., 2017).

This study used molecular systematic techniques to unravel the puzzling diversity of the morphologically variable tent tortoise complex. Our main objectives were to (a) verify whether the previous advocated "three subspecies" assumption was valid, (b) resolve the phylogenetic structure and understand the genetic diversity within and among different clades of the *P. tentorius* species complex, (c) address and clarify the OTUs of the complex, and (d) predict the potential for future cladogenesis and adaptive radiation to aid conservation management. To do this, we used both fast-evolving mtDNA and slow-evolving nDNA markers to infer the phylogenetic relationships within the *P. tentorius* complex.

2 | MATERIALS AND METHODS

2.1 | Sampling strategy

We obtained data of the distribution ranges of the subspecies of *P. tentorius* from the literature (Boycott & Bourquin, 2000; Branch, 1998, 2008; Greig, 1975; Greig & Burdett, 1976; Hewitt, 1933, 1934; Hofmeyr et al., 2014), the online Virtual Museum records (ADU), iSpot distribution records, Bayworld Museum (Port Elizabeth) and Ditsong Museum of Natural History (Pretoria). We mapped all the records using Google Earth Pro. Finally, we selected representative collecting sites to maximize the geographic area covered for each subspecies. A total of 455 specimens of *P. tentorius* were processed from 76 localities throughout its range in South Africa and Namibia, which comprehensively covered nearly the entire distribution range of *P. tentorius* according to the literature (Boycott & Bourquin, 2000; Branch, 1998, 2008; Greig, 1975; Greig & Burdett, 1976; Hewitt, 1933, 1934; Hofmeyr et al., 2014). Details of the areas sampled are given in Table S1 and Figure S2. Outgroup species selection (Table S2) to provide structure to the tree topology was based on previous studies (Cunningham, 2002; Hofmeyr et al., 2017; Le, Raxworthy, McCord, & Mertz, 2006).

We collected samples from both live and dead animals. For live animals, we collected fresh tissue from the tail tip (15–25 mg) or blood (0.1 ml) from the subclavian vein. Material from dead animals fell into three categories. We collected (a) connective tissue from inside shells that were in good condition; (b) bone samples from old museum or field shells; and (c) muscle tissue from preserved museum specimens from near the femur or tibia because these tissues had less exposure to preservation chemicals. Fresh tissue was preserved in 96% ethanol, and blood was dissolved in 1% of a 10M sodium–EDTA anticoagulant solution. Connective tissue and bone were kept dry in vials, whereas preserved muscle was stored in 96% ethanol. All samples were stored at −80°C in the laboratory until we did DNA extraction.

2.2 | DNA Extraction, DNA marker selection, Amplification, and Sequencing

We used QIAGEN DNeasy Blood and Tissue kits (QIAGEN, Germany) to extract DNA from all sample types. The manufacturer's protocol

was followed for fresh tissue and connective tissue from shells, except that for connective tissue the elution buffer volume was reduced from 200 to 50 µl in the final elusion stage to increase the final DNA concentration.

For preserved museum specimens, we used muscle tissue attached to the surface of the femur and tibia, since samples from these areas often provided better quality DNA. Bone and preserved muscle tissue were pre-treated before extracting DNA and were processed in a room where no *Psammobates* samples had been processed before. To prevent contamination, the work bench was exposed for 13 hr to UV light radiation and air flow between procedures. Bone samples were washed twice with double-distilled water and 96% ethanol to eliminate possible foreign DNA. Washed bones were then dried naturally and ground into powder with a sterilized mortar and pestle. A total of 25 mg bone powder of each sample was transferred into a microcentrifuge tube. The preserved muscle tissue was first rehydrated in microcentrifuge tubes filled with double-distilled water and incubated at 40°C for 2–6 days, occasionally up to 2 weeks. This procedure accelerated lysis and digesting processes and helped to dissolve unknown chemicals. Rehydrated samples were then washed two times with 96% ethanol. If the solution still showed color, we washed it three times with double-distilled water using a vortex. We then removed the ethanol and washed the tissue three times with a 10M PBS solution, using the vortex, to eliminate as much formalin as possible as well as other chemical reagents used during preservation.

When extracting DNA from bone powder, incubation time was increased from 24 to 48 hr, while 5–10 days were allowed for the extraction from preserved museum samples. To increase the total DNA yield from bone and preserved muscle, we added 30 µl Proteinase K (100 µg/ml) on day one and subsequently added 15 µl Proteinase K every day until all tissue was digested. After the digestion stage, the standard QIAGEN extraction protocol was followed, except that the volumes of the rest of the reagents were doubled before the elution stage. Furthermore, to optimize the final DNA yield and increase the final DNA concentration and thus the possibility of obtaining longer fragments of genomic DNA, we added only 15–25 µl of QIAGEN elution buffer during the elution stage. The elution buffer was preheated to 70°C to increase the final yield in terms of DNA concentration. A similar approach was used with museum samples by Daniels, Hofmeyr, Henen, and Crandall (2007) in a study of the phylogeny of the tortoise *Chersina angulata*. In order to check that no foreign DNA was introduced during DNA extraction, a blank control group was set up and treated in the same way as the bone and preserved muscle tissue, that is, using the same reagents but no tissue samples.

Tortoises are generally slow-evolving compared to most other reptiles (Avise, Bowen, Lamb, Meylan, & Bermingham, 1992); thus, fast-evolving mtDNA markers rather than slow-evolving nDNA genes should be used to accurately reconstruct their phylogenetic trees, particularly when studying small-scale species complexes (Lourenço, Claude, Galtier, & Chiari, 2012). Studies showed that slow-evolving nuclear markers are only suitable for studying the phylogeny of organisms with a wide time span range, while faster-evolving mtDNA markers become more accurate when focusing on recent events with relatively short time spans (Lourenço et al., 2012; Zheng, Peng, Kuro-o, & Zeng, 2011). This study therefore focused on six widely used fast-evolving mtDNA loci to study the phylogeny of the *P. tentorius* complex: *12S rRNA* gene (*12S*, Kocher et al., 1989), *16S rRNA* gene (*16S*, Palumbi et al., 1991), *cytochrome b* gene (*Cytb*, Kocher et al., 1989; Pääbo, 1990; Whiting, Bauer, & Sites, 2003), *NADH dehydrogenase subunit 4* gene (*ND4*, Stuart & Parham, 2004), and the two *ND4* adjacent tRNA genes, *tRNA-His* and *tRNA-Ser* (Stuart & Parham, 2004). In addition to these, we used the fastest evolving nDNA gene *Prolactin Receptor Coding* gene (*PRLR*, Townsend, Alegre, Kelley, Wiens, & Reeder, 2008).

We used the polymerase chain reaction (PCR) to amplify *12S*, *16S*, *Cytb*, *ND4*, *tRNA-His* & *tRNA-Ser*, and *PRLR* genes. Primer details are given in Table S3; note that we used two primer pairs for *Cytb* to accommodate fresh and ancient DNA. PCRs were performed using KAPA2G Robust HotStart, USA. Optimal annealing temperatures for *12S*, *16S*, *Cytb*, *ND4* with *tRNA-His* & *tRNA-Ser*, and *PRLR* were determined through temperature gradient tests. PCRs for ancient DNA samples (here defined as bone and preserved muscle tissue) were performed independently from other samples, to minimize the chance of cross-contamination. To ensure that no foreign DNA was introduced during the PCR stage, a blank control group was again set up without adding any template DNA; instead, 2 µl from the extraction control group was added. PCRs were performed in a BIO-RAD T 100™ Thermal Cycler (Singapore) under the following parameters: an initial 4 min denaturation step at 94°C, followed by 37 cycles (43 for ancient DNA) of 30 s denaturation at 94°C, 30 s annealing (62°C for *12S*; 50°C for *16S*; 51°C for *Cytb*; 61°C for *ND4* with *tRNA-His* & *tRNA-Ser*; 58°C for *PRLR*), and 1 min extension at 72°C, with a final 10 min extension step at 72°C. The PCR products were electrophoresed in 1% agarose gel, visualized under UV light, and purified using a BioFlux PCR Purification Kit (Bioer Technology, China). Purified PCR products were cycle sequenced using BigDye (ABI PRISM® BigDye Terminator v3.1 Cycle Sequencing Kits, USA) and standard methods (with annealing temperatures of 60°C for *12S*; 50°C for *16S*; 50°C for *Cytb*; 57°C for *ND4* with *tRNA-His* & *tRNA-Ser*; 57°C for *PRLR*). The BigDye PCR products were purified by Zymo DNA Sequencing cleanup kit (Epigenetics Company, USA), prior to sequencing in an ABI 3,500 genetic analyzer.

2.3 | Sequence alignment, modeling, treatment of indels, and partitioning

All Sanger sequences were analyzed using ABI Prism Sequencing Analysis software v.3.7 (Applied Biosystems), aligned with MUSCLE v.3.2 (Edgar, 2004) and manually checked with MEGA v.7 (Kumar, Stecher, & Tamura, 2016). We employed PartitionFinder v.1.0.1 (Lanfear, Calcott, Ho, & Guindon, 2012) under Python v.2.7 (Python Software Foundation, 2010) to determine the best partition scheme for the concatenated dataset. We also used jModeltest v.2 (Darriba, Taboada, Doallo, & Posada, 2012) to determine the best substitution model and parameter settings for each data partition via AIC (Table S4). Nucleotide base biases across different partitions were determined through the homogeneity test implemented in PAUP

v.4.0 (Swofford, 1998). Substitution saturation tests were performed using DAMBE v.6.1.9 (Xia, 2013) at both gene and partition (on protein coding gene) levels and visually plotted with transition and transversion against GTR model modified genetic distance diagrams. This was done to investigate potential substitution saturation, particularly at the third codon position in protein coding fast-evolving mtDNA. The DAMBE v.6.1.9 software was used to read sequence frames in order to determine the codon positions. All mtDNA sequences were concatenated through SequenceMatrix (Vaidya, Lohman, & Meier, 2011) thereafter.

2.4 | Phylogenetic analyses

2.4.1 | Phylogenetic inferences

The preliminary maximum-likelihood (ML) analysis with RAxML v.8 (Stamatakis, 2014) using RDP4 (Martin, Murrell, Golden, Khoosal, & Muhire, 2015) with 1,000 bootstrap replications (Felsenstein, 1985; results not shown) showed significant tree topology conflicts between the mtDNA and nDNA loci, but not among the six mtDNA loci. We therefore ran mtDNA and nDNA separately in all subsequent analyses. We employed four different phylogenetic inference approaches independently for the concatenated mtDNA dataset and the nDNA dataset: ML, maximum parsimony (MP), and two Bayesian inference analyses (BI).

Maximum-likelihood analyses were performed using RAxML-VI-HPC (Stamatakis, 2014) implemented in CIPRES Science Gateway (Miller, Pfeiffer, & Schwartz, 2010), using non-parametric bootstraps to determine the power of support for each node. The ML analyses results inferred with the mtDNA concatenated dataset and nDNA dataset were partitioned based on results of PartitionFinder, each partition with rate heterogeneity and substitution model GTR-CAI, and 1,000 non-parametric bootstrap replications.

Maximum parsimony analyses were conducted using PAUP v.4.0 with default settings, and all gaps were treated as "new states," since gaps are informative sources that can potentially contribute phylogenetic information. Both mtDNA and nDNA datasets were treated with 1,000 bootstrap replications, using a full heuristic search consensus tree and only retaining groups with a frequency > 50%.

The traditional Bayesian inference (BI) analyses (Huelsenbeck & Ronquist, 2001; Huelsenbeck, Ronquist, Nielsen, & Bollback, 2001) were conducted with MrBayes 3.2 (Ronquist et al., 2012) on the mtDNA and nDNA datasets independently, using the evolutionary substitution model on each partition specified, based on results from PartitionFinder with the applicable prior parameter settings. Different substitution rates were allowed at different partitions. Tree topologies and tree lengths were linked across partitions, and other parameters treated were unlinked. Indel sites were treated as independent partitions. Analyses consisted of four independent runs (each with 4 chains) of 10 million generations with minimum sampling frequency every 5,000 generations, discarding the first 25% of the samples as burn-in. We used average standard deviation of split frequencies (ASDSF) to examine the mixing of chains and convergence of each run from the posterior probability distribution and ESS (effective sample size) to determine whether sampling was sufficient. We ran analyses until the ASDSF value was close to 0.02 and the ESS > 200 for most of the parameters.

Multispecies coalescent (MSC) BI analysis using the species tree inference approach (Edward, 2009; Liu, Yu, Pearl, & Edwards, 2009) was performed with the StarBeast package (Drummond, Suchard, Xie, & Rambaut, 2012; Heled & Drummond, 2009) implemented in BEAST v.2.4.8 (Bouckaert et al., 2014) for the mtDNA and nDNA datasets separately. This approach can evaluate the degree of discordance between gene and species trees by sorting incomplete evolutionary lineages. This method requires inputs of all individuals from different predefined phylogenetic groups, which are evaluated and retrieved based on results of the traditional BI phylogenetic analysis with MrBayes. The multispecies coalescent analysis permits the simultaneous estimation of species tree topology, gene tree topology, and population size, among others (Heled & Drummond, 2009). We employed the same substitution models and partition scheme previously used for the MrBayes analysis on the mtDNA dataset. Trees from different partitions were linked, and a Yule model selected. For MCMC setting, each analysis comprised four independent runs of 55 million generations, sampling every 5,000 generations and discarding the first 10% of samples as burn-in. Tracer 1.6 (Rambaut, Suchard, Xie, & Drummond, 2014) was used to check if the effective sample size (ESS) reached the threshold value of 200.

In all cases, the consensus trees yielded by the different phylogenetic inference methods were visualized and drawn with FigTree 1.4.3 (Rambaut, 2012) or TreeGraph 2 (Stöver & Müller, 2010). We considered strongly supported nodes as those with bootstrap support (BP) > 70% for ML, and MP analyses (Hillis & Bull, 1993) and with posterior probability (PP) > 0.95 for Bayesian inference (Huelsenbeck & Rannala, 2004).

2.4.2 | Tree topology test

For the conflicting tree topology from ML result: (((C1 + C4) + C7) + C5) + C6 and C2 + C3, and topology retrieved from all other results: ((C1 + C4) + C7) + C5 and (C2 + C3) + C6, we performed a Bayesian tree topology test under MrBayes, in order to test which assumption had greater likelihood. Different tree topology scenarios were set as constraints to test alternatives. Each hypothesis run had similar parameter settings with Bayesian analysis, allowing rate multipliers among gene partitions, each gene partition incorporating different substitution models and parameters. Each test run was split into two runs with MCMC of 10 million generations, sampling every 5,000 generations with the stepping-stone method used to perform the test. The rest of the steps were identical to traditional MrBayes analysis. The marginal likelihood generated from each independent run of each hypothesis was used for comparison, assuming the one with the highest marginal likelihood mean as the best assumption. In terms of threshold, a likelihood difference in the range of 5 log units was considered as strong evidence in favor of the better model (Kass & Raftery, 1995).

2.5 | Haplotype networks and molecular multivariant (Motif PCA) analysis

We employed TCS v.2.1 (Clement, Posada, & Crandall, 2000) to reconstruct haplotype groups and PopART (Leigh & Bryant, 2015) to draw parsimony networks for the two most informative mtDNA genes, *Cytb* and *ND4*, respectively. We performed a motif PCA on the mtDNA dataset with DAMBE v.6 to determine the significant motif clusters for the *P. tentorius* complex.

2.6 | Species delimitation analyses

In order to clarify the species boundaries and assign a hierarchical position to each retrieved clade, nine different species delimitation methods were used to determine whether each clade represented a valid putative species.

2.6.1 | ABGD approach

The ABGD (Automatic Barcode Gap Discovery) species delimitation approach (Puillandre, Lambert, Brouillet, & Achaz, 2012) has been shown to delineate species on the "barcoding gap" between distributions of interspecific pairwise genetic distances and intraspecific pairwise distances. This single-locus, distance-based method effectively eliminates the bias of traditional pure p-distance-based methods. We applied the ABGD approach on delimited putative species for mtDNA loci *12S*, *16S*, *Cytb*, *ND4 + tRNA-His + tRNA-Ser*, and nDNA *PRLR*, independently.

2.6.2 | The rjMCMC BPP

We used the Bayesian reversible-jump MCMC (rjMCMC) algorithm (Rannala & Yang, 2003; Yang, 2015) with an MSC model-based species delimitation approach implemented in the program BPP (Yang, 2015) to delineate seven retrieved clades to putative species. This approach can calculate posterior probabilities in the MSC model on a given species tree by considering parameters tau (divergence times) and θ (population size). Again, the phylogenetic consensus tree retrieved from BEAST (but with all outgroups removed) was used as guide tree. Two prior parameter settings were used for both tau (shallow divergence: 2, 2,000; deep divergence: 1, 10) and θ (small population size: 2, 2,000; large population size: 1, 10) parameters, resulting in a total of four possible combinations that were tested by the A00 method implemented in BPP. All four tests were conducted with algorithm 1 and fine-tune parameters alpha = 2 and m = 1. Each test was run for 300,000 generations with sampling frequency adjusted to 5% and 10% as burn-in. Each test was repeated three times with different initial seeds to ensure that all analyses generated similar results with different initial seeds. The guide tree used in all four analyses was based on a species tree generated from BI analysis (mtDNA dataset). The mtDNA dataset used in this analysis was partitioned by different gene loci beforehand, to allow rate heterogeneity across different loci.

2.6.3 | STACEY MSC

We used a BEAST MSC approach with the STACEY package (Jones, 2015) on the mtDNA dataset partitioned by PartitionFinder results. Optimization of the substitution model and parameters was based on the PartitionFinder and jModeltest results. Samples were assigned to clades based on the phylogenetic consensus tree retrieved from BEAST. Analyses involved runs with 55 million generations, sampling every 5,000 generations and discarding the first 25% as burn-in. Results were loaded into Tracer to check sample mixing with ESS and convergence diagnostics as in advanced phylogenetic analyses. Both gene and species trees generated from BI analyses were loaded into the TreeAnnotator package implemented in BEAST after the first 10% of trees was discarded as burn-in. The species trees were also loaded into the program DensiTree v.2.0 (Bouckaert, 2010) to visualize the topological overlapping of all trees and to visualize conflict branches and nodes after discarding 10% of the trees as burn-in, to determine the stability of each node and branch.

2.6.4 | Marginal likelihood-based MSC

A marginal likelihood estimate (MLE) based MSC analysis conducted with the STACEY package implemented in BEAST was used for defining the species boundaries of the clades retrieved with the phylogenetic analysis (C1 to C7, see Results section). Seven different species allocation scenarios (from one to seven species, see below) were set up by BEAUti under the STACEY template with the same substitution model and parameter settings as used above, employing the stepping-stone method for testing species boundaries based on the Bayesian factor (BF = 2lnBfs = 2 × (best scenario MLE − alternative scenario MLE)) to determine the best putative clustering scheme (Grummer, Bryson, & Reeder, 2013). This could be (a) a one species assumption: C1 + C2 + C 3 + C4 + C5 + C6 + C7; (b) a two species assumption: C1 + C4 + C5 + C7 and C2 + C3 + C6; (c) a three species assumption: C1 + C4 + C5 + C7 and C2 + C3 and C6; (d) a four species assumption: C1 + C4 + C7 and C5 and C2 + C3 and C6; (e) a five species assumption: C1 + C4 + C7 and C5 and C2 and C3 and C6; (f) a six species assumption: C1 + C4 and C7 and C5 and C2 and C3 and C6; or (g) a seven species assumption C1 and C4 and C7 and C5 and C2 and C3 and C6. Each scenario was imported into the BEAST package Path sampler to calculate the MLE with 10 million generations, sampling divided into nine different threads, with the first thread removed as burn-in. Bayes factors were generated with 2lnBf = 0–2 meaning "not worth more than a bare mention," 2lnBf = 2–6 meaning "positive" support, 2lnBf = 6–10 meaning "strong" support, and 2lnBf > 10 meaning "decisive" support evidence in favor of the better model (Grummer et al., 2013; Kass & Raftery, 1995).

2.6.5 | GMYC and bGMYC species delimitation

The GMYC and bGMYC species delimitation analysis was performed with the mtDNA dataset. Since both GMYC and bGMYC species delimitation methods require the input tree to be ultrametric, we removed all redundant sequences (i.e. identical sequences belonging to the

same haplotype), thus ensuring that there were no identical sequences in the concatenated mtDNA supermatrix. All non-congeneric outgroups were truncated from the mtDNA supermatrix before the analyses were performed. We thereafter applied the same BI inference and parameter settings as mentioned in the BI MSC above, using BEAST.

The generalized mixed Yule-coalescent (GMYC) model (Fontaneto et al., 2007; Fujisawa & Barraclough, 2013; Pons et al., 2006) implemented in R (R Core Team, 2017) with single threshold GMYC model (Reid & Carstens, 2012) was used to perform GMYC species delimitation analysis by R package Split (Ezard, Fujisawa, & Barraclough, 2009) on the BEAST mtDNA tree file.

We also performed bGMYC Bayesian coalescent model-based analysis to delineate putative species from the BI MSC trees (generated from BEAST) with multiple thresholds. This approach allowed determining the putative species by producing a pairwise probability heat map among tree leaves to visualize the probability that two nested tree leaves are members of the same putative species. This approach was applied by using the R package GMYC (Reid & Carstens, 2012). We first ran the algorithms on a single random tree from randomly selected 100 BEAST MSC trees, ran MCMC with 100,000 generations, cutoff 50,000 as burn-in, thinning = 10, leaves prior interval was set between 2 and 200 with start vector value = 25 to confirm sampling is sufficient and well mixed. Then, we ran MCMC with 100,000 generations, cutoff 50,000 as burn-in, using a thinning interval of 100 from the 100 random trees. Besides plotting a probabilities heatmap for nested tips, a log posterior sampling plot, a coalescent ratio plot and a sampling frequency plot were plotted simultaneously to confirm sampling was sufficient and even.

2.6.6 | PTP and mPTP

A PTP model-based (Zhang et al., 2013) Bayesian species delimitation was performed on bPTP web server (Zhang et al., 2013) with the mtDNA RaxML ML tree, with parameters settings: MCMC generations 500,000, 100 of thinning and cutoff 10% as burn-in. Finally, we also performed an mPTP model (Kapli et al., 2017), which is based on PTP methods but with multiple rates and faster ML algorithms.

2.6.7 | TCS network punctuation

We employed TCS v.2.1 (Clement et al., 2000) to determine the number of significantly punctuated clades, with each isolated clade regarded as a valid putative species, since it represented the clade as being substantially isolated from the others.

2.6.8 | Traditional *p*-distance method

We used MEGA v.7 to calculate the uncorrected *p*-distances for the *Cytb* gene among clades, within each clade, as well as the *p*-distances between outgroups. We used *Cytb* as the reference gene to calculate *p*-distances, as it has been widely used in testudine studies (Cunningham, 2002; Fritz et al., 2012; Kindler et al., 2012; Vamberger et al., 2018; Vargas-Ramírez et al., 2010).

2.7 | Genetic diversity and population differentiation

In order to investigate the genetic structure and genetic diversity of each clade, the concatenated mtDNA supermatrix was transformed into a haplotype sequences matrix file using DnaSP v.5 (Librado & Rozas, 2009) before being imported into Arlequin v. 3.5 (Excoffier & Lischer, 2010). Only 423 samples with full coverage for all mtDNA loci were included in this analysis. Analysis of molecular variance (AMOVA) (Excoffier, Smouse, & Quattro, 1992) was used to compute genetic structure, genetic distance (mean pairwise nucleotide differences) and F_{ST} variation with corresponding significance levels among, and within each of the groups and populations using 10,000 permutations. Fu's F_s test (Fu, 1997) and Tajima's *D* test (Tajima, 1989) were used to test whether selection on the different allele sites was even and neutral using a coalescent simulation in Arlequin. A population differentiation analysis method was used to determine whether individuals of each assigned population were randomly distributed. Molecular diversity indices θs were calculated by Arlequin, and these θ values were used to evaluate molecular diversity among different clades. Finally, we performed a population differentiation test in Arlequin to determine whether the seven clades differed significantly from each other.

3 | RESULTS

Despite extensive protocol modifications for DNA extraction from historical museum specimens, our success rate was generally poor. For 150 preserved and dry samples from the Port Elizabeth Museum, only 23 samples (15.3%) were successfully amplified for the short sequences of *12S* and *Cytb*.

A 407-bp fragment of *12S* (*12S* alignment provided as File S1), 539 bp of *16S* (*16S* alignment provided as File S2), 465 bp (short sequence for recovering ancient DNA from museum preserved specimens) and 693 bp (long sequence) of *Cytb* (*Cytb* alignment provided as File S3), 678 bp of *ND4* (*ND4* alignment provided as File S4), 120 bp of the *tRNA-His* and *tRNA-Ser* (*tRNA-His* and *tRNA-Ser* alignment provided as File S5), and 525 bp of *PRLR* (*PRLR* alignment provided as File S6). All the nucleotide sequences generated for this study have been deposited in NCBI GeneBank (*12S*: MH175524-MH175534, MK693228-MK693442: *16S*: MH175535-MH175540, MK693454-MK693668; *Cytb*: MH175541-MH175633, MK693742-MK693956, *ND4*: MH175634-MH175674, MK693957-MK694171; *tRNA-His* & *tRNA-Ser*: MK694288-MK694716 and *PRLR*: MH175675-MH175700, MK694172-MK694287); for details, see Table S1.

No significant substitution saturation was detected, neither from DAMBE (Iss.c > Iss with $p < .0001$ in all cases, for both symmetrical and asymmetrical topology assumptions). The GTR model-based substitution saturation plots (for 1st and 3rd codon positions for *Cytb* and *ND4* genes, non-protein-coding genes *12S*, *16S* and *tRNA-His* & *tRNA-Ser*) did not reveal any signs of saturation; results not shown).

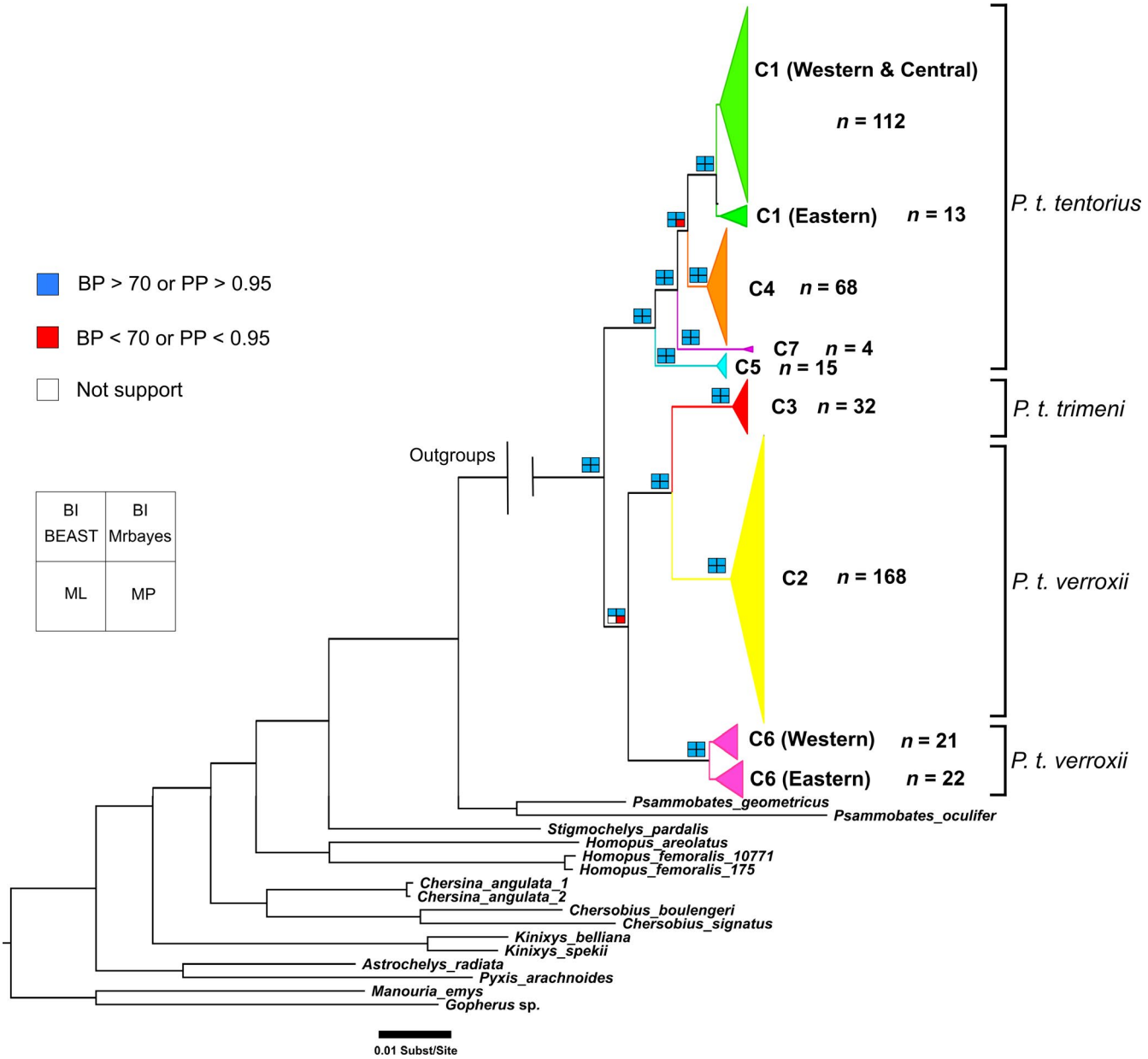

FIGURE 1 Phylogenetic relationships among seven clades of the *P. tentorius* complex inferred from BI (Mrbayes) analysis of the combined mtDNA dataset (*12S*, *16S*, Cytb, *ND4*, *tRNA-His*, and *tRNA-Ser*), detail given in text. The strength of support values from four different approaches was visualized as three different colors as indicated in the figure. Top two are posterior probabilities (PP) for Bayesian inference (BI) from BEAST and MrBayes, while, the bootstrap values (BP) of maximum likelihood (ML) and maximum parsimony (MP) are shown at the bottom. Color scheme: C1 colored in green, C2 in yellow, C3 in red, C4 in orange, C5 in aqua, C6 in pink, and C7 in purple. The same color scheme will be used throughout this paper

3.1 | Partitions and substitution models

PartitionFinder suggested the best partition scheme to be: Subset 1: *12S*, Subset 2: *16S*, Subset 3: *Cytb* 1st codon position, Subset 4: *ND4* 2nd codon position, *Cytb* 2nd codon position, Subset 5: *Cytb* 3rd codon position, Subset 6: *ND4* 1st codon position, Subset 7: *ND4* 3rd codon position and Subset 8: *tRNA-His* & *tRNA-Ser*. Homogeneity tests showed that the third codon positions of *Cytb* and *ND4* mtDNA genes were significantly different from other partitions in their nucleotide composition ($p < .0001$). None of the other partitions were significantly different ($p > .05$). The optimal substitution models with parameters for different partitions are specified in Table S4.

3.2 | Phylogeny of the *P. tentorius* complex

For the mtDNA dataset (an alignment of 2,437 bp in total, provided as File S7), the MP analysis results recovered a phylogenetic tree

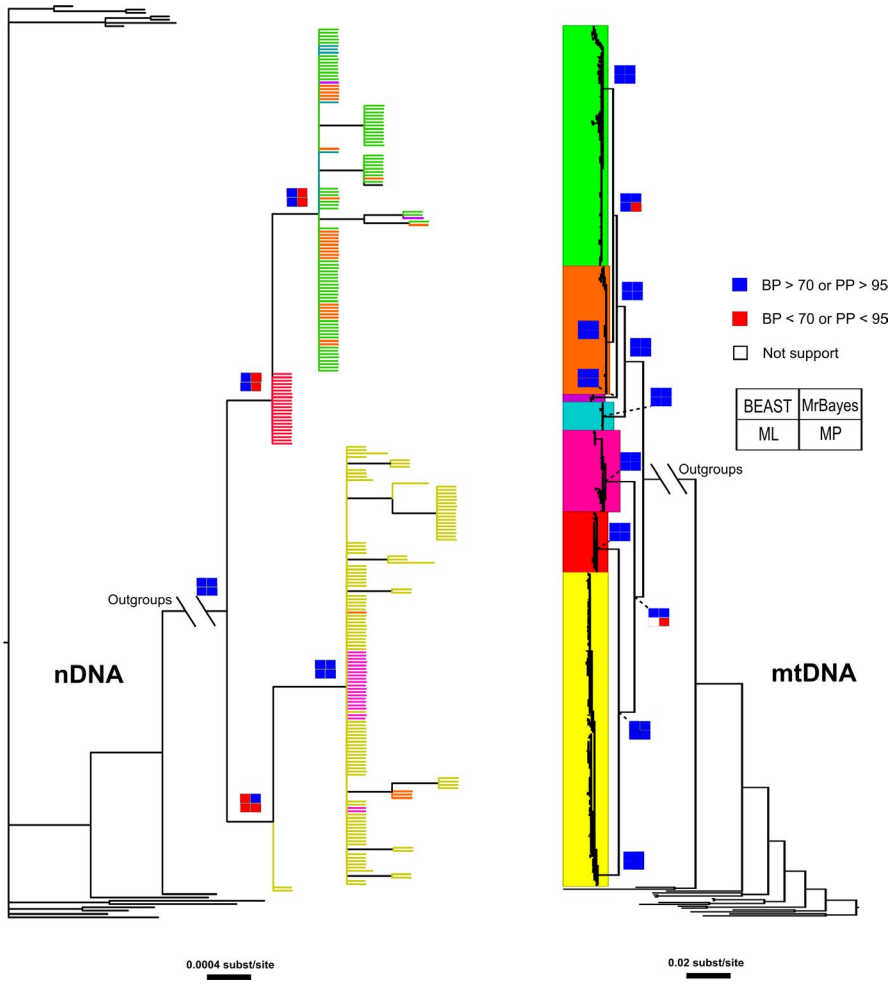

FIGURE 2 Phylogenetic relationships among seven clades of the *P. tentorius* complex inferred by BI (Mrbayes) analyses from the nDNA dataset (*PRLR*) (left) and mtDNA dataset (*12S*, *16S*, Cytb, *ND4*, *tRNA-His*, and *tRNA-Ser*) (right). Detail given in text. The strength of support values from four different approaches was visualized as three different colors as indicated in the figure. Top two are posterior probabilities (PP) for Bayesian inference (BI) from BEAST and MrBayes, while the bootstrap values (BP) of maximum likelihood (ML) and maximum parsimony (MP) are shown at the bottom

with total tree length of 3,075 steps, CI (Consistency index) = 0.37, RI (Retention index) = 0.92, and HI (Homoplasy index) = 0.63. Maximum-likelihood analysis generated a best tree with ML optimization likelihood = −13,866.01. The phylogenetic analyses retrieved seven distinct clades, each with strong support. Apart from ML, the phylogenetic reconstructing approaches generated similar tree topologies ((C1 + C4) + C7) + C5 and (C2 + C3) + C6 across all major nodes and branches (Figure 1). Each node generally showed high support values (BP > 70 and PP > 0.95) except for ML (Figure 1), where the results revealed a weakly supported (BP < 70) topology (((C1 + C4) + C7) + C5) + C6 and C2 + C3. Such topological conflicts were also visible from BEAST MSC analysis (both gene trees and species trees) by using DensiTree (Figure S3). Topology tests, however, suggested the most likely topology to be (C2 + C3) + C6, since its marginal likelihood value was almost 100 log units higher than that of the alternative topology.

Clade 1, C7, and C5 included specimens previously recognized as *P. t. tentorius* (Boycott & Bourquin, 2000), whereas C4 comprised individuals previously considered *P. t. trimeni* but recently identified as *P. t. tentorius* from the Kamiesberg, Hamtam Karoo and Roggeveldberge (KHR) region of the western Great Escarpment (GE; Rhodin et al., 2017). Clade 3 included samples of *P. t. trimeni* from the West Coast Succulent Karoo (WCSK) region (Boycott & Bourquin, 2000; Rhodin et al., 2017), and appeared as the sister group of C2, which comprised *P. t. verroxii* (Boycott & Bourquin, 2000; Rhodin et al., 2017) south of the Orange River. Clade 6 also comprised of *P. t. verroxii* populations, occurring north of the Orange River (NOR), being the sister group of C2 + C3. These results indicate that *P. t. verroxii* is not monophyletic and thus not a systematically valid taxon.

The nDNA-based phylogenetic analyses revealed substantial topological conflicts compared with the mtDNA results. Our *PRLR*-based phylogeny retrieved two clades, each with two subclades, with moderate-to-low support (Figure 2). Clade 1 consisted of one subclade composed solely of individuals assigned to mtDNA C3 (*P. t. trimeni*), whereas the second subclade consisted of mtDNA C1, C4, C5, and C7, all forming part of *P. t. tentorius*. The second clade included individuals assigned mainly to mtDNA C2 and C6, corresponding to *P. t. verroxii* from south and north of the Orange River. The major subclade of clade 2 also included four samples corresponding to mtDNA C4 (*P. t. tentorius*): three of these were from Sutherland and one from Kliprand, both regions where the distributions of C4 and C2 abut and/or overlap. Two individuals (mtDNA C1) from Victoria West (a region where the distributions of C1 and C2 overlap) formed a distinct subclade of clade 2, possibly indicating intergradation between the two mtDNA clades. The nDNA phylogeny, in general, delineated the three recognized subspecies of *P. tentorius*.

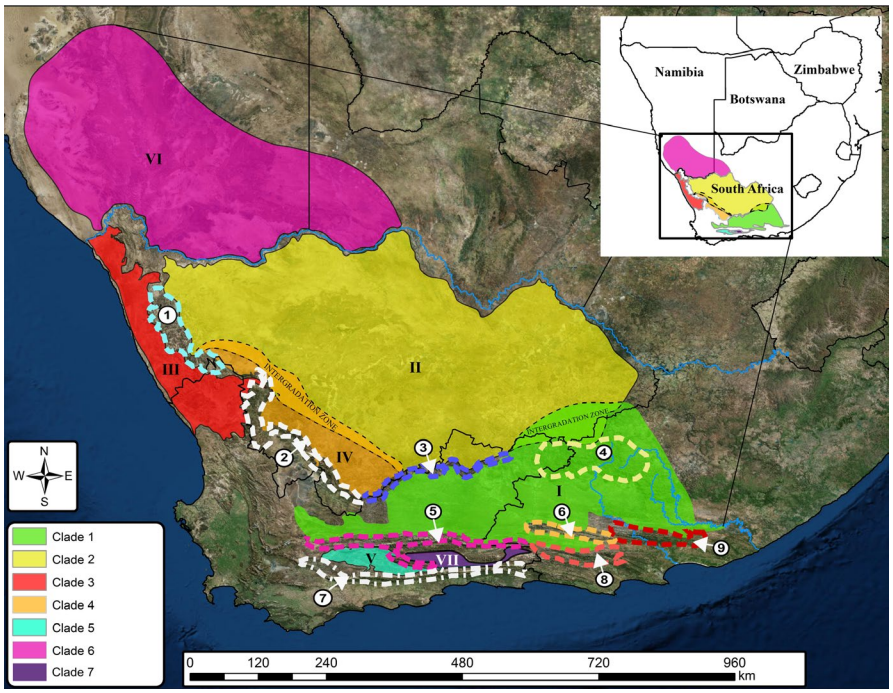

FIGURE 3 The map shows geographic distribution of the seven mtDNA clades throughout their range. The thin light blue belt represents a major river system between South Africa and Namibia. The major mountain barriers are indicated (1) Kamiesberg, (2) Hantamberge and Roggeveldberge, (3) Nuweveldberge, (4) Sneeuberge and Kompasberg, (5) Swartberge and Rooiberg, (6) Grootrivierberge, (7) Langeberge, (8) Baviaanskloofberge, (9) Suurberge

Based on our mtDNA results, C1 occurs throughout the lower and central Karoo (from the Tankwa Karoo on the western side to the Fish River Valley to the east) and is mainly present below the GE, except where it overlaps with C2 at the corridor between the Nuweveldberge and Sneeuberge. Clade 2 occurs throughout the upper Karoo and Bushmanland in the Northern Cape Province and overlaps the range of C4 in the KHR region. Clade 3 is found in the WCSK region, from the Richtersveld in the north to the Knersvlakte in the south. Clade 4 occurs in the KHR region and clade 5 in the southwestern Little Karoo, whereas clade 6 is only present NOR. Finally, C7 is found in the Oudtshoorn Basin, a region between Calitzdorp and Oudtshoorn, possibly extending into the Uniondale area (Figure 3).

3.3 | Networks and motif PCA

The TCS network retrieved six significantly punctuated haplotype groups for *Cytb* (Figure S4), and between C1 and C4, there were only ten mutation steps and a zero median vector, while, for *ND4* (Figure S4), only five significantly punctuated haplotype groups were generated, and there were only four mutation steps between C1 and C4. The motif PCA generated a similar clustering pattern and TCS network. The first two principal components explained 74.35% of the total variation and were used to visualize the scatterplot for all seven clades (Figure S5). Clade 1, C4, C7, and C5 were very close to each other, although none of the seven clusters overlapped.

3.4 | Species delimitation

Our ABGD approach in general supported a "seven putative taxa" assumption for *Cytb* and *16S*, but not for the ND4, *16S* and *PRLR* genes. The ABGD advocated "six putative species" for *ND4* and *12S*, but the ABGD of *PRLR* only retrieved "three putative species," which was congruent with the nDNA phylogenetic results (Table 1).

The BPP method generally strongly supported seven clades as valid OTUs for all four scenarios (Table 1, Figure S3). All four possible scenarios: (a) small population size (θ: 2, 2,000) with shallow divergence (tau: 2, 2,000); (b) small population size (θ: 2, 2,000) with deep divergence (tau: 1, 10); (c) large population size (θ: 1, 10) with shallow divergence (tau: 2, 2,000); and (d) large population size (θ: 1, 10) with deep divergence (tau: 1, 10) all showed strong support (PP > 0.95) at all nodes across topologies. Therefore, it may imply that the depth of divergence and population size do not influence BPP delimitation in the *P. tentorius* complex.

The Bayesian STACEY MSC species tree showed that all nodes within the *P. tentorius* complex were strongly supported (PP > 0.95). Notwithstanding, the DensiTree revealed the uncertainty about the relationship between C6 and C2 + C3, and the existence of conflicting placements of C6. Nevertheless, the BF-based stepping-stone method strongly supported all seven clades as seven valid OTUs (since the "seven" species scenario generated significantly better MLE and BF scores (see Table 2).

The GMYC method strongly suggested seven OTUs (ML of null model = 172.46, ML of GMYC = 219.39, likelihood ratio = 93.86, and LR test: $p < .0001$). The bGMYC retrieved seven major nested tip groups as putative OTUs (Figure S6), since the pairwise posterior probability differences between all clades were significant ($p < .05$ in all cases). Besides, the MCMC outputs from both single-tree runs and multiple-tree runs were well mixed (Figures S7 and S8), implying that the MCMC sampling was sufficient in both cases. The "checkrates" R function revealed that the distribution of ratios of the Coalescence to Yule rates sampled in the analysis was all above zero, indicating that the model may be a good approximation of the data (Figure S9). The

TABLE 1 The summary of different species delimitation approaches with different data partitions. The number of putative species is indicated in each case with its putative species cluster components. For the BPP methods: 6a) small population size (θ: 2, 2000) with shallow divergence (tau: 2, 2,000), 6b) small population size (θ: 2, 2,000) with deep divergence (tau: 1, 10), 6c) large population size (θ: 1, 10) with shallow divergence (tau: 2, 2,000), 6d) large population size (θ: 1, 10) with deep divergence (tau: 1, 10)

Putative species no.	1) ABGD (*ND4 + tRNA*)	2) ABGD (*PRLR*)	3) ABGD (*12S*)	4) ABGD (*16S*)	5) ABGD (*Cytb*)
	6	3	6	7	7
C1	C1 = C4	C1 = C4 = C5 = C7	C1 = C4	C1	C1
C2	C2	C2 = C6	C2	C2	C2
C3	C3	C3	C3	C3	C3
C4	C1 = C4	C1 = C4 = C5 = C7	C1 = C4	C4	C4
C5	C5	C1 = C4 = C5 = C7	C5	C5	C5
C6	C6	C2 = C6	C6	C6	C6
C7	C7	C1 = C4 = C5 = C7	C7	C7	C7

	6a) BPP mtDNA	6b) BPP mtDNA	6c) BPP mtDNA	6d) BPP mtDNA	7) STACEY- MSC (mtDNA)	8) STACEY- BF (mtDNA)
	7	7	7	7	7	7
C1	C1	C1	C1	C1	C1	C1
C2	C2	C2	C2	C2	C2	C2
C3	C3	C3	C3	C3	C3	C3
C4	C4	C4	C4	C4	C4	C4
C5	C5	C5	C5	C5	C5	C5
C6	C6	C6	C6	C6	C6	C6
C7	C7	C7	C7	C7	C7	C7

	9) GMYC (mtDNA)	10) bGMYC (mtDNA)	11a) bPTP (mtDNA)	11b) mPTP (mtDNA)	12) TCS (*Cytb*)	13) TCS (*ND4*)
	7	7	7	7	6	5
C1	C1	C1	C1	C1	C1 = C4	C1 = C4 = C7
C2	C2	C2	C2	C2	C2	C2
C3	C3	C3	C3	C3	C3	C3
C4	C4	C4	C4	C4	C1 = C4	C1 = C4 = C7
C5	C5	C5	C5	C5	C5	C5
C6	C6	C6	C6	C6	C6	C6
C7	C7	C7	C7	C7	C7	C1 = C4 = C7

mtDNA dataset used in GMYC methods was provided as File S8. Both bPTP and mPTP methods retrieved seven significant clusters as OTUs.

The *Cytb* gene-based pairwise distance analysis results (see Table S5) revealed that the greatest among-clades *p*-distance occurred between C6 and C7 (*p*-distance = 0.083), while the smallest among-clades *p*-distance was found between C1 and C4 (*p*-distance = 0.019). At within-clade level, C2 and C6 were the clades with the highest within-clade *p*-distance (*p*-distance = 0.007), while C7 had with the lowest within-clade *p*-distance (*p*-distance = 0.000).

3.5 | Genetic diversity and population differentiation

The AMOVA results showed that the demarcations among the seven mtDNA clades were all statistically significant (*df* = 422, sum of squares = 13,373.04, total variance = 41.37 $p < .0001$), 94.71% of the total variation being from among the clades (*df* = 6, sum of squares = 12,462.64, total variance = 39.18, $p < .0001$), and only 5.29% from within the clades (*df* = 416, sum of squares = 910.4, total variance = 2.19, $p < .0001$). The Tajima's *D* test suggested neutral selection on nucleotide substitution in all seven clades and the *P. tentorius* complex overall ($p > .05$ in all cases), while Fu's F_s test gave incongruent results for C1-C5, for which the neutral selection assumption was rejected ($p < .05$ in all cases). The results indicated that C1, C2, and C6 had higher haplotype diversity, molecular diversity (at both $\Theta\pi$ and Θ k), and genetic diversity than the other clades (Table 3, Figure 4, and Table S6). The population differentiation test revealed that the seven clades were significantly different from each other (pairwise $p < .0001$ in all cases). The global differentiation test also showed significant differences among the seven clades (Exact

TABLE 2 The Bayesian factor (BF) and marginal likelihood estimation (MLE) based BEAST MSC species delimitation results of the seven putative species scenarios. The best scenario with highest MLE and BF is in bold

Scenario	Putative species	Topology hypothesis	MLE	BF	Rank
S1	1	(C1 + C2+C3 + C4 + C5 + C6 + C7)	−8515.35	174.82	5
S2	2	(C1 + C4 + C5 + C7) + (C2 + C3 + C6)	−8561.29	266.7	6
S3	3	(C1 + C4 + C5 + C7) + ((C2 + C3) + C6))	−8455.72	55.56	2
S4	4	((C1 + C4 + C7) + C5) + ((C2 + C3) + C6)	−8499.69	143.5	4
S5	5	((C1 + C4 + C7) + C5)) + (((C2) + C3) + C6)	−8596.15	336.42	7
S6	6	(((C1 + C4) + C7) + C5) + (((C2) + C3) + C6)	−8492.96	130.04	3
S7	**7**	**((((C1) + C4) + C7) + C5) + ((((C2) + C3) + C6))**	**−8427.94**	*n*/a	**1**

$p < .0001$). The pairwise F_{ST} and average pairwise differences (Pxy) among the seven clades denoted significant diversification among all pairwise clades (all pairwise p-values of $F_{ST} < 0.0001$, all pairwise p-values of Pxy < 0.0001).

4 | DISCUSSION

4.1 | Phylogeny

Our phylogenetic analyses of the *P. tentorius* complex showed discordance between the mtDNA and nDNA results. The mtDNA data retrieved seven clades, with strong support, with each clade occupying a distinct geographic region. The nDNA data, in contrast, were broadly consistent with the three recognized subspecies and also indicated the possibility of hybridization in regions where mtDNA clades abut or overlap. Based on morphology, Loveridge and Williams (1957) and Boycott and Bourquin (1988) regarded the southwestern Karoo as a three-way contact zone where all three subspecies are present and presumably hybridize. Both our mtDNA and nDNA results, however, showed that only one mtDNA clade, C1, of *P. t. tentorius* occurred there. Nevertheless, additional sampling in the contact zones of mtDNA clades is needed to ascertain which clades hybridize and how extensive hybridization is.

Our mtDNA results indicated that *P. t. tentorius* consisted of four clades, each occurring in a different region, generally isolated from each other. Based on morphology, tortoises above the GE in the KHR regions were previously considered to be *P. t. trimeni* (Boycott & Bourquin, 1988; Branch, 2008) but our results confirmed the findings reported in Hofmeyr et al. (2014) and Rhodin et al. (2017) that these tortoises belong to *P. t. tentorius* (clade 4). Our sampling showed that *P. t. trimeni* is only present west of the GE in the WCSK region of South Africa. Despite extensive sampling in southwestern Namibia, all individuals from there belonged to C6 (*P. t. verroxii*); consequently, there is reason to doubt that the range of *P. t. trimeni* extends into Namibia as has been reported (Boycott & Bourquin, 2000; Branch, 2008; Greig & Burdett, 1976). The paraphyly of *P. t. verroxii* highlights the need for a taxonomic revision of C2 and C6, a finding also addressed by Hofmeyr et al. (2017).

The fact that morphological characteristics are unreliable to differentiate taxa in a highly polymorphic species such as *P. tentorius* is illustrated by the contention of Hewitt (1933, 1934), Loveridge and Williams (1957) and Branch (2008) that the uniformly brown colored "*P. bergeri*" morph may deserve recognition as a new taxon. In our study, we found this uniformly brown morph in the Tankwa Karoo (mtDNA assigned to C1), in the Upington and Helmeringhausen regions (mtDNA assigned to C6) and in northern Bushmanland, near Pofadder and Kenhardt (mtDNA assigned to C2). Therefore, "*P. bergeri*" is merely a color morph present in three separate clades and should not be considered a separate taxon.

Inconsistent phylogenetic patterns of mtDNA and nDNA have been found across a range of organisms (Gonçalves, Martínez-Solano, Ferrand, & García-París, 2007; Rato, Carranza, Perera, Carretero, & Harris, 2010; Shaw, 2002), including turtles (Vargas-Ramirez, Carr, & Fritz, 2013). Such incongruences are usually caused by variation in

TABLE 3 The AMOVA results showing polymorphic sites, number of haplotypes, Tajima's *D* and Fu's F_s test results (selective neutral test), and quantification of the genetic diversity at Θ π and Θ k of the seven clades, as well as for the entire *P. tentorius* complex

Clade	*N*	Polymorphic sites	No. Haplotype	Tajima's *D* test	Fu's F_s test	Θ (π)	Θ (k)	Gene diversity
C1	123	39	52	$p > .05$	$p < .001$	3.95	33.45	0.95
C2	151	52	64	$p > .05$	$p < .001$	5.91	41.40	0.97
C3	32	13	11	$p > .05$	$p < .05$	1.90	5.50	0.68
C4	63	9	12	$p > .05$	$p < .05$	1.40	4.12	0.59
C5	12	10	7	$p > .05$	$p < .05$	2.05	6.15	0.83
C6	38	38	22	$p > .05$	$p > .05$	7.48	20.95	0.95
C7	4	4	3	$p > .05$	$p > .05$	2.00	3.77	0.83
Entire complex	423	288	171	$p > .05$	$p > .05$	63.38	106.25	0.98

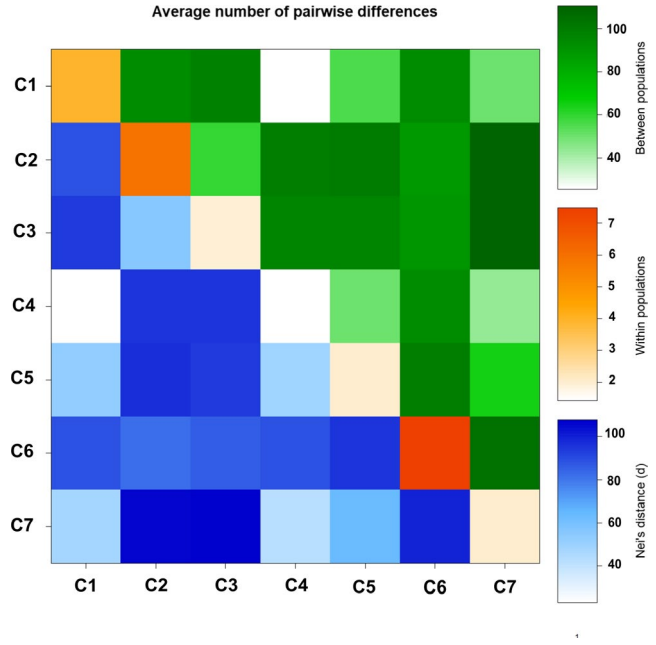

FIGURE 4 The heatmap visualizes the average number of pairwise differences between clades and within clades, and also Nei's distance

the depth of coalescent processes at different gene loci, incomplete lineage sorting resulting in mismatching between gene trees, rate heterogeneity and species trees (Baum & Smith, 2013; Leliaert et al., 2014; Rubinoff & Holland, 2005), as well as stochastic bias and errors from algorithms, various models, and programs (Rubinoff & Holland, 2005). Mitochondrial and nuclear DNA carry different evolutionary information and their selection and inheritance patterns may be incongruent with each other, but both nevertheless represent fundamental elements of genomic evolution (Rubinoff & Holland, 2005).

When dealing with a slow-evolving group, such as tortoises (Avise et al., 1992; Tollis et al., 2017), the rapid evolving mtDNA markers are possibly more suitable and reliable for elucidating a phylogeny (Rubinoff & Holland, 2005) than the slower evolving nDNA markers. Phylogenetic studies of chelonians (tortoises and turtles) often rely on mtDNA loci because nDNA markers frequently provide incongruent and poorly resolved trees with poor support (e.g., Kindler et al., 2012; Petzold et al., 2014). Caccone et al. (2004) found that nDNA divergence was nearly 30 times slower than that of mtDNA for Galápagos tortoises. Nonetheless, there are pitfalls in using mtDNA to elucidate phylogenetic history. For example, when gene trees bifurcate before speciation, the number of taxa can be overestimated, and in the case of incomplete lineage sorting, gene trees may incorrectly reflect species trees (Baum & Smith, 2013; Leliaert et al., 2014; Rubinoff & Holland, 2005; Shaw, 2002). More research at genomic level is therefore necessary to clarify the phylogeny of the *P. tentorius* complex.

The diversification patterns not only demarcated seven cluster groups within the *P. tentorius* complex, it also showed subclades within C1 and C6 (Figure 1, Figure 2, and Figure 3). Clade 1 tends to show a clear subdivision between western and central populations, and population from the Fish River valley in the east. It is also noteworthy to mention that the eastern population usually has highly domed carapacial scutes compared to those from the central and western region. In terms of C6, populations from the west and east formed separate branches, but morphological differences between the two are not clear (Zhao et al., unpublished data). These subdivisions may imply that the relatively high level of within-group diversifications of C1 and C6 may lead to further cladogenic radiation of the *P. tentorius complex*.

4.2 | Species delimitation and determination of OTUs

Our delimitation analyses suggested five to seven putative species (Figure 5 and Table 1) for the seven mtDNA clades of the *P. tentorius* complex. The conservative nDNA (*PRLR*) results distinguished the three subspecies of *P. tentorius*, corresponding to *P. t. tentorius* (C1 + C4+C5 + C7), *P. t. trimeni* (C3), and *P. t. verroxii* (C2 + C6). It is known that population size and divergence level can impact species delimitation negatively and that the impact from incomplete lineage sorting (usually ascribed to bias from markers) can also be substantial (Luo, Ling, Ho, & Zhu, 2018; Sukumaran & Knowles, 2017). We used six mtDNA markers and one nDNA marker, and the power of the markers was possibly not strong enough to perform MSC-based coalescent delimitation. Our sample sizes for C5 and C7 were also comparatively low. Notwithstanding, a recent simulation-based study found that increasing the number of loci and sample size caused only limited improvement of multiple species delimitation approaches (Luo et al., 2018) and that even single-locus-based delimitations can be useful (Sukumaran & Knowles, 2017).

The MSC model-based species delimitation approach seems useful for delineating putative species, as it outperforms the simple distance threshold-based DNA barcoding methods (Yang & Rannala, 2017). However, a simulation-based study found that the MSC species delimitation approach identifies genetic structure rather than species, resulting in an inflated estimation of the number of species (Sukumaran & Knowles, 2017). A follow-up study based on both simulations and real datasets supported the finding that BPP may detect population splits rather than species divergences, but still advocates the method's usefulness under certain conditions (Leaché, Zhu, Rannala, & Yang, 2018). In general, the genomic data-based species delimitation approaches should be considered as hypotheses, which require additional information (e.g., ecological and phenotypic) before delineating species (Leaché et al., 2018; Sukumaran & Knowles, 2017). Moreover, the other Bayesian-based methods, and the GMYC approach may also over-split the number of species (Miralles & Vences, 2013).

To set up species delimitation boundaries so as to assign taxa to proper hierarchical positions remain challenging (de Queiroz, 2007; Fujita, Leaché, Burbrink, McGuire, & Moritz, 2012), as the delimitation boundary between species is still controversial in view of different species concepts having completely different criteria (de Queiroz, 1998, 1999, 2005a, 2005b), and none of them can comprehensively explain the delimitation thresholds. Furthermore, none of the existing species concepts can define species boundaries to fully cover species from every aspect of hierarchical systematics

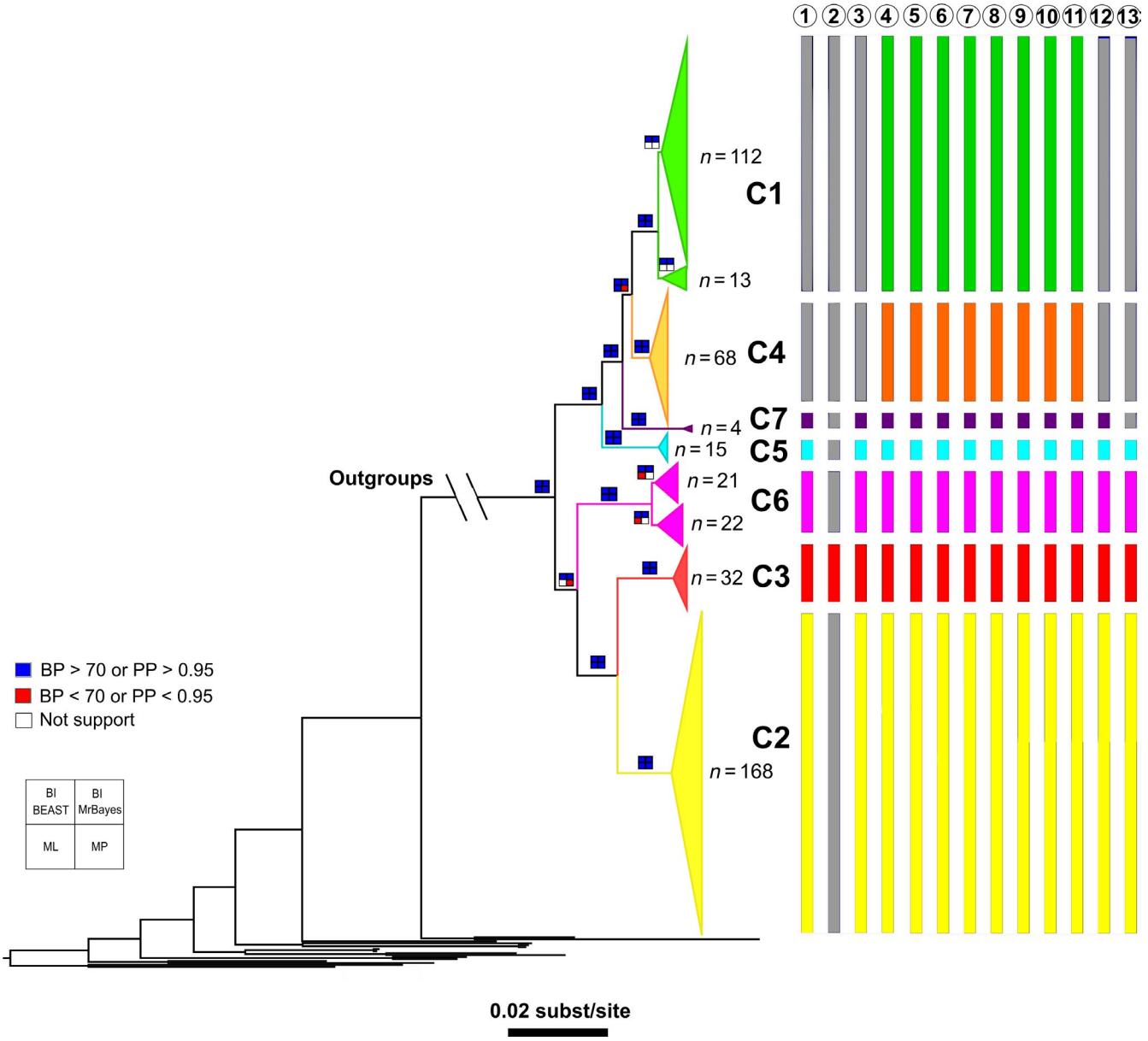

FIGURE 5 The summary graph shows results from multiple species delimitation approaches and the phylogenetic relationships among seven clades of the *P. tentorius* complex inferred from BI (Mrbayes) analysis of the combined mtDNA dataset (*12S*, *16S*, Cytb, *ND4*, *tRNA-His* and *tRNA-Ser*), detail given in text. The strength of support values from four different approaches was visualized as three different colors as indicated in the figure. Top two are posterior probabilities (PP) for Bayesian inference (BI) from BEAST and MrBayes, while, the bootstrap values (BP) of maximum likelihood (ML) and maximum parsimony (MP) are shown at the bottom. The bar plot 1–13 at right side represents multiple species delimitation results, the detail given in Table 1. The gray color represents clades which were not supported by the corresponding species delimitation approaches

(de Queiroz, 1998, 1999, 2005b, 2007). This is especially true when there is no unifying ultimate definition for a species (Mallet, 2013). Even modern likelihood and Bayesian algorithms, based on sophisticated delimitation methods, can only generate results suggesting a particular taxonomy, but are unable to determine absolute species boundaries (Fujita et al., 2012; Jones, 2015, 2017; Kass & Raftery, 1995; Leaché & Fujita, 2010).

One of the traditional approaches to assigning OTUs to hierarchical taxonomic positions and evaluate species boundaries is to consider uncorrected pairwise genetic distances (*p*-distance), particularly of *Cytb*. The literature reflects wide variation in *Cytb p*-distances among congeneric species, ranging from 1.5% to 4.92% for *Trachemys* taxa (Fritz et al., 2012), 8.6 to 18.3% for *Pelomedusa* species (Vargas-Ramírez et al., 2010), and 3.76 to 11.59% for *Kinixys* species (Kindler et al., 2012). The uncorrected *p*-distances of *Cytb* among five *Testudo* species were 6.9% to 12.7% (Fritz & Bininda-Emonds, 2007), and among the three *Indotestudo* species, 3.7% to 5.9% (Iverson, Spinks, Shaffer, McCord, & Das, 2001). With regard to *p*-distances within species level, the *p*-distances within *Chersina angulata* and *Chersobius signatus* were reported as 2.8% (Daniels et al.,

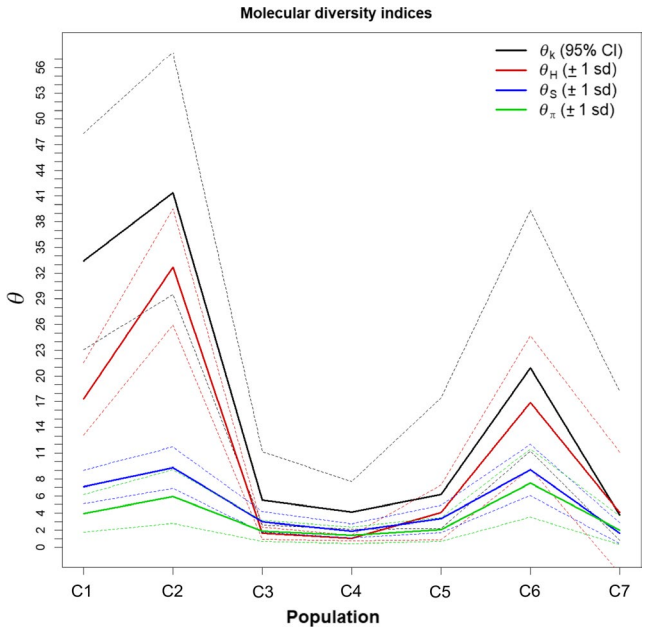

FIGURE 6 The diagram shows molecular indices θ (k), θ (π), theta (H), and theta (S) across all seven mtDNA clades

2007) and 2.5% (Daniels, Hofmeyr, Henen, & Baard, 2010), respectively; within *Aldabrachelys gigantea*, 0.9% (Austin, Arnold, & Bour, 2003); within *Testudo graeca*, 3.4% (Fritz et al., 2007); within *Testudo horsfieldii*, 1.1% (Fritz et al., 2009); and within *Chelonoidis chilensis*, 1.21% (Fritz et al., 2012). For the *P. tentorius* complex, the *Cytb* average *p*-distances among clades ranged from 1.9%–8.3% with only one comparison (C1 vs. C4 = 1.9%) being below 3.8%. When comparing the *p*-distance values within the *P. tentorius* complex to other tortoise species, it is clear that only the *p*-distances between C1 and C4 fall into the range as observed within, but all other pairwise comparison between clades fall into the range as observed between. Based on this index, it is possible that at least six mtDNA clades of *P. tentorius* may qualify for species specific status, but more research is required to investigate the level of hybridization between clades. The traditional absolute *p*-distance-based methods are arbitrary and biased, since *p*-distance may vary greatly across different organism groups for the same gene, also, the substitutions may not be evenly distributed in each gene (Nagy, Sonet, Glaw, & Vences, 2012). Nonetheless, Malone and Fontenot (2008) found a positive correlation between uncorrected *p*-distance and the degree of post-zygotic isolation in toads. This finding may imply that degree of genetic divergence (in terms of *p*-distance) is still a useful indicator reflecting the patterns of reproductive isolation and may still be a useful criterion in species delimitation.

In general, 70.59% of the species delimitation approaches used here advocated a "seven putative species" assumption, 17.65% suggested "six putative species," 5.89% supported "five putative species" and 5.89% supported the "three subspecies" assumption. According to the Relative taxonomic resolving power index (Rtax) advocated by Miralles and Vences (2013) (for details, see Table S7), we found that the 6–7 putative species assumption seemed well supported, since their Rtax indices were much higher than that of other assumptions. Therefore, it seems likely that some of the mtDNA clades may deserve full species status. Such a decision, however, should obviously be approached with caution. Results of a fine-scale genetic structure study using multiple microsatellite DNA markers (Z. Zhao, in prep) and a morphological marker-based study (Z. Zhao, in prep) are crucial before considering taxonomic revision in this species complex.

4.3 | Genetic diversity and conservation genetics

The AMOVA results and differentiation analysis also advocated seven distinct clusters in line with the seven clades retrieved from the phylogenetic analyses and species delimitation analyses, since all clades were genetically highly structured. Clade 1, C2, and C6 showed higher levels of haplotype diversity and genetic diversity than C3, C4, C5, and C7. Although such horizontal comparisons might be biased and arbitrary, since the sample sizes were not even across all clades, C4 showed the lowest genetic diversity level of all the clades (Table 3, Figures 4 and 6), even though it had the third largest sample size of all. This might imply that the diversification rate of C4 is slow, and its genetic structure generally stable. In contrast, the genetic diversity of C6 is remarkably high (Table 3, Figures 4 and 6), even though the sample size of C6 only ranked 4th among the clades. This finding corresponds with our phylogenetic results, where a clear subdivision was discovered within C6 (Figures 1, 2 and Figure S6). Furthermore, it is noteworthy that C6 had the highest within-group average pairwise differences of all clades (Figure 4). This may lead us to reasonably deduce that the diversification rate in C6 is higher than that of the other clades, and further cladogenic radiation might be expected in C6. The pairwise F_{ST} and Pxy showed that C2, C3, and C6 are closer to each other than to the other clades, that is, to C1, C4, C5, and C7 (Figure 4). The Fu's F_s test results imply that selection was not neutral in C1–C5.

Conservation efforts should focus mainly on conserving areas with cladogenic radiation potential, rather than only on preserving endemic taxa (Allendorf & Luikart, 2009). In this regard, the KHR region should be given adequate conservation attention since the area includes the intergradation zone between C2 and C4. So far, only the Akkerendam Nature Reserve has been established in this region. Similar attention is warranted for the biological corridor between the Nuweveldberge and Sneeuberge, where C1 and C2 coexist. Unfortunately, this region currently has no protected areas. In the NOR region, C6 deserves conservation attention, given that it has the highest genetic potential for further radiation. Fortunately, the region already has protected areas like Ai-Ais Trans-frontier Park and Tirasberg Conservancy.

4.4 | Future research

In this study, we used six mtDNA markers and one nDNA marker, and multiple modern species delimitation procedures to clarify the phylogenetic relationships within the *P. tentorius* complex, resulting in the identification of seven mtDNA clades and seven putative species. In order to map the full evolutionary picture of the complex, a further analysis based on multiple microsatellite DNA markers is needed to

determine the fine-scale genetic structure among populations and to quantify gene flow among them. This analysis (Z. Zhao, in prep) is also crucial for investigating whether the putative species patterns derived from mtDNA, nDNA, and microsatellite DNA markers are congruent. Furthermore, the findings of a phylogeographic analysis (Z. Zhao, in prep) will also be crucial for understanding the diversification and radiation of the *P. tentorius* complex under temporal and spatial dimensions. In this regard, the approach will allow tracing back the cladogenic history and aligning historical geographic events with the diversification of the *P. tentorius* complex from an ecological perspective.

In summary, we identified seven operational taxonomic units within the complex, rather than the currently advocated "three-subspecies assumption," though the conservative nDNA marker generated incongruent results. We found that *P. t. verroxii* was not a monophyletic group and should therefore not be considered as taxonomically valid. The genetic diversity analyses revealed that C1, C2, and C6 had significantly higher genetic diversity than C3, C4, C5, and C7. Further cladogenic radiation is expected in C1 and C6, but particularly in C6, since the within-clade genetic diversity level found in the latter clade was the highest. The establishment of protected areas at intergradation zones between C1 and C2, C2 and C4, and in the distribution range of C6, may facilitate conserving the adaptive radiation potential of the *P. tentorius* complex. Future studies at morphological level, at population level using microsatellite DNA markers and at phylogeographic level, will be crucial for understanding the evolutionary radiation of the *P. tentorius* complex, as well as for making appropriate decisions in its taxonomic review.

ACKNOWLEDGEMENTS

We thank the National Research Foundation Grant-IFR150216114248) and the Universities of the Free State (UFS Research Grant: A1999/158110) and Western Cape (SNS Grants) for providing the funding to carry out the study. We also thank the following provincial departments of Nature Conservation for providing collecting permits: Northern Cape Province (Permits No: FAUNA 1061/2/2015, FAUNA 1266/2016, FAUNA 1458/2015, FAUNA 1267/2016, FAUNA 0729/2018, FAUNA 0730/2018), Western Cape Province (Permit No: AAA007-00179-0056), and Eastern Cape Province (Permits No: CRO 171/16CR, CRO 172/16CR), as well as the Ministry of Environment and Tourism, Namibia (Permit No: 1430/2009). The Animal Research Ethics Committees of the University of the Free State (AREC References No: 180111-005, UFS-AED2015/0013) and University of the Western Cape (ScRiRC2008/39) are thanked for providing ethical clearance for the research. We thank Dr Krystal Tolley from the South African National Biodiversity Institute (SANBI) and Werner Conradie from Port Elizabeth Museum for providing additional samples for the DNA analyses. Theo Busschau, David Maguire, Jonathan Palmer, Gary Nicolau, Luke Kemp, Tyrone Ping, Marius Burger, Andries Cilliers, and Zandri Van Der Westhuizen are thanked for assisting with biological sample collections. We extend our sincere gratitude to all the farmers who allowed us to enter their properties for sampling and for being a great source of information regarding the areas on their farm to search. We are also thankful to Dr. Joaquín Verdú-Ricoy for assisting with the text editing. The High-Performance Computing Centre of the University of the Free State is thanked for assisting with the phylogenetic analyses.

ORCID

Zhongning Zhao https://orcid.org/0000-0002-6400-4743
Adriaan Jordaan https://orcid.org/0000-0002-0772-0305

REFERENCES

Aldhebiani, A. Y. (2018). Species concept and speciation. *Saudi Journal of Biological Sciences*, *25*(3), 437–440. https://doi.org/10.1016/j.sjbs.2017.04.013

Allendorf, F. W., & Luikart, G. (2009). *Conservation and the genetics of populations*. New York, NY: John Wiley & Sons.

Austin, J. J., Arnold, E. N., & Bour, R. (2003). Was there a second adaptive radiation of giant tortoises in the Indian Ocean? Using mitochondrial DNA to investigate speciation and biogeography of *Aldabrachelys* (Reptilia, Testudinidae). *Molecular Ecology*, *12*(6), 1415–1424. https://doi.org/10.1046/j.1365-294X.2003.01842.x

Avise, J. C., Bowen, B. W., Lamb, T., Meylan, A. B., & Bermingham, E. (1992). Mitochondrial DNA evolution at a turtle's pace: Evidence for low genetic variability and reduced microevolutionary rate in the Testudines. *Molecular Biology and Evolution*, *9*(3), 457–473. https://doi.org/10.1093/oxfordjournals.molbev.a040735

Balakrishnan, R. (2005). Species concepts, species boundaries and species identification: A view from the tropics. *Systematic Biology*, *54*(4), 689–693. https://doi.org/10.1080/10635150590950308

Baum, D. A., & Smith, S. D. (2013). *Tree thinking: an introduction to phylogenetic biology*. Greenwood Village, CO: Roberts.

Bouckaert, R. R. (2010). DensiTree: Making sense of sets of phylogenetic trees. *Bioinformatics*, *26*(10), 1372–1373. https://doi.org/10.1093/bioinformatics/btq110

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., … Drummond, A. J. (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, *10*(4), e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Boycott, R. C., & Bourquin, O. (1988). *The South African Tortoise Book. A Guide to South African Tortoises, Terrapins and Turtles*. Johannesburg, South Africa: Southern Book Publisher.

Boycott, R. C., & Bourquin, O. (2000). The southern African tortoise book: A guide to southern African tortoises, terrapins and turtles. O. Bourquin.

Branch, B. (1998). *Field guide to snakes and other reptiles of southern Africa*. Sanibel Island, FL: Ralph Curtis Books. Cape Town.

Branch, B. (2008). *Tortoise, Terrapins and Turtles of Africa*. Cape Town, South Africa: Struik Publisher.

Branch, W. R., Benn, G. A., & Lombard, A. T. (1995). The tortoises (Testudinidae) and terrapins (Pelomedusidae) of southern Africa: Their diversity, distribution and conservation. *South African Journal of Zoology*, *30*(3), 91–102. https://doi.org/10.1080/02541858.1995.11448377

Caccone, A., Gentile, G., Burns, C. E., Sezzi, E., Bergman, W., Ruelle, M., … Powell, J. R. (2004). Extreme difference in rate of mitochondrial and nuclear DNA evolution in a large ectotherm. *Galápagos Tortoises. Molecular Phylogenetics and Evolution*, *31*(2), 794–798. https://doi.org/10.1016/j.ympev.2004.02.004

Caron, D. A., Countway, P. D., Savai, P., Gast, R. J., Schnetzer, A., Moorthi, S. D., … Jones, A. C. (2009). Defining DNA-based operational taxonomic units for microbial-eukaryote ecology. *Journal of Applied & Environmental Microbiology*, *75*(18), 5797–5808. https://doi.org/10.1128/AEM.00298-09

Clement, M., Posada, D. C. K. A., & Crandall, K. A. (2000). TCS: A computer program to estimate gene genealogies. *Molecular Ecology*, *9*(10), 1657–1659. https://doi.org/10.1046/j.1365-294x.2000.01020.x

Core Team, R. C. T. R. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for statistical computing. http://www.R-project.org/

Cracraft, J. (1983). Species concepts and speciation analysis. In R. F. Johnston (Ed.), *Current ornithology* (pp. 159–187). Boston, MA: Springer.

Cunningham, J. (2002). *A molecular perspective on the family Testudinidae Batsch, 1788, Ph.D. Dissertation, University of Cape Town*.

Daniels, S. R., Hofmeyr, M. D., Henen, B. T., & Baard, E. H. W. (2010). Systematics and phylogeography of a threatened tortoise, the speckled padloper. *Animal Conservation*, *13*(3), 237–246. https://doi.org/10.1111/j.1469-1795.2009.00323.x

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*(3), 915–926. https://doi.org/10.1016/j.ympev.2007.08.010

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: More models, new heuristics and parallel computing. *Nature Methods*, *9*(8), 772. https://doi.org/10.1038/nmeth.2109

De Queiroz, K. (1998). The general lineage concept of species, species criteria, and the process of speciation: A conceptual unification and terminological recommendations. In D. J. Howard, & S. H. Berlocher (Eds.), *Endless forms: Species and speciation* (pp. 57–75). New York, NY: Oxford University Press.

De Queiroz, K. (1999). The general lineage concept of species and the defining properties of the species category. In R. A. Wilson (Ed.), *Species: New interdisciplinary essays* (pp. 49–89). Cambridge, MA: MIT Press.

De Queiroz, K. (2005a). Different species problems and their resolution. *BioEssays*, *27*(12), 1263–1269. https://doi.org/10.1002/bies.20325

De Queiroz, K. (2005b). Ernst Mayr and the modern concept of species. *Proceedings of the National Academy of Sciences*, *102*(suppl 1), 6600–6607. https://doi.org/10.1073/pnas.0502030102

De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology*, *56*(6), 879–886. https://doi.org/10.1080/10635150701701083

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, *29*(8), 1969–1973. https://doi.org/10.1093/molbev/mss075

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. https:// doi.org/10.1093/nar/gkh340

Edwards, S. V. (2009). Is a new and general theory of molecular systematics emerging? *Evolution: International Journal of Organic Evolution*, *63*(1), 1–19. https://doi.org/10.1111/j.1558-5646.2008.00549.x

Excoffier, L., & Lischer, H. E. (2010). Arlequin Suite Ver 3.5: A new series of programs to perform population genetics analyses under Linux and windows. *Molecular Ecology Resources*, *10*(3), 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, *131*(2), 479–491.

Ezard, T., Fujisawa, T., & Barraclough, T. G. (2009). *Splits: Species' limits by threshold statistics*. R Package Version, *1*(11), r29.

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, *39*(4), 783–791. https://doi.org/10.1111/j.1558-5646.1985.tb00420.x

Fontaneto, D., Herniou, E. A., Boschetti, C., Caprioli, M., Melone, G., Ricci, C., & Barraclough, T. G. (2007). Independently evolving species in asexual bdelloid rotifers. *PLoS Biology*, *5*(4), e87. https://doi.org/10.1371/journal.pbio.0050087

Fritz, U., Auer, M., Chirikova, M. A., Duysebayeva, T. N., Eremchenko, V. K., Kami, H. G., … Hundsdörfer, A. K. (2009). Mitochondrial diversity of the widespread Central Asian steppe tortoise (*Testudo horsfieldii* Gray, 1844): Implications for taxonomy and relocation of confiscated tortoises. *Amphibia-Reptilia*, *30*(2), 245–257. https://doi.org/10.1163/156853809788201135

Fritz, U., & Bininda-Emonds, O. R. (2007). When genes meet nomenclature: Tortoise phylogeny and the shifting generic concepts of *Testudo* and *Geochelone*. *Zoology*, *110*(4), 298–307. https://doi.org/10.1016/j.zool.2007.02.003

Fritz, U., Hundsdörfer, A. K., Široký, P., Auer, M., Kami, H. G., 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*(1), 97–121. https://doi.org/10.1163/156853807779799135

Fritz, U., Stuckas, H., Vargas-Ramírez, M., Hundsdörfer, A. K., Maran, J., & Päckert, M. (2012). Molecular phylogeny of Central and South American slider turtles: Implications for biogeography and systematics (Testudines: Emydidae: Trachemys). *Journal of Zoological Systematics and Evolutionary Research*, *50*(2), 125–136. https://doi.org/10.1111/j.1439-0469.2011.00647.x

Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, *147*(2), 915–925.

Fujisawa, T., & Barraclough, T. G. (2013). Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent approach: A revised method and evaluation on simulated data sets. *Systematic Biology*, *62*(5), 707–724. https://doi.org/10.1093/sysbio/syt033

Fujita, M. K., Leaché, A. D., Burbrink, F. T., McGuire, J. A., & Moritz, C. (2012). Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology & Evolution*, *27*(9), 480–488. https://doi.org/10.1016/j.tree.2012.04.012

Gonçalves, H., Martínez-Solano, I., Ferrand, N., & García-París, M. (2007). Conflicting phylogenetic signal of nuclear vs mitochondrial DNA markers in midwife toads (Anura, Discoglossidae, Alytes): Deep coalescence or ancestral hybridization? *Molecular Phylogenetics and Evolution*, *44*(1), 494–500. https://doi.org/10.1016/j.ympev.2007.03.001

Greig, J. C. (1975). Patterns In the distribution of Southern African terrestrial tortoises (Chelonia: Cryptodira: Testudinidae). *The Journal of the Herpetological Association of Africa*, *14*(1), 9–9.

Greig, J. C., & Burdett, P. D. (1976). Patterns in the distribution of southern African terrestrial tortoises (Cryptodira: Testudinidae). *African Zoology*, *11*(2), 251–273. https://doi.org/10.1080/00445096.1976.11447532

Grummer, J. A., Bryson, R. W. Jr, & Reeder, T. W. (2013). Species delimitation using Bayes factors: Simulations and application to the Sceloporus scalaris species group (Squamata: Phrynosomatidae). *Systematic Biology*, *63*(2), 119–133. https://doi.org/10.1093/sysbio/syt069

Heled, J., & Drummond, A. J. (2009). Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, *27*(3), 570–580. https://doi.org/10.1093/molbev/msp274

Hewitt, J. (1933). On the Cape species and subspecies of the genus *Chersinella* Gray. Part I. *Annals of the Natal Museum*, *7*, 255–297.

Hewitt, J. (1934). On the Cape species and subspecies of the genus *Chersinella*. Part II. *Annals of the Natal Museum*, *7*, 303–352.

Hillis, D. M., & Bull, J. J. (1993). An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology*, *42*(2), 182–192. https://doi.org/10.1093/sysbio/42.2.182

Hofmeyr, M. D., Boycott, R. C., & Baard, E. H. W. (2014). *Psammobates tentorius* (Bell, 1828). 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* (pp. 70–85). Pretoria, South Africa: South African National Biodiversity Institute.

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

Huelsenbeck, J. P., & Rannala, B. (2004). Frequentist properties of Bayesian posterior probabilities of phylogenetic trees under simple and complex substitution models. *Systematic Biology*, *53*(6), 904–913. https://doi.org/10.1080/10635150490522629

Huelsenbeck, J. P., & Ronquist, F. (2001). MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics*, *17*(8), 754–755. https://doi.org/10.1093/bioinformatics/17.8.754

Huelsenbeck, J. P., Ronquist, F., Nielsen, R., & Bollback, J. P. (2001). Bayesian inference of phylogeny and its impact on evolutionary biology. *Science*, *294*(5550), 2310–2314. https://doi.org/10.1126/science.1065889

Hull, D. L. (1977). The ontological status of species as evolutionary units. In R. E. Butts & J. Hintikka (Eds.), *Foundational problems in the special sciences* (pp. 91–102). Dordrecht, the Netherlands: Springer.

Iverson, J. B., Spinks, P. Q., Shaffer, H. B., McCord, W. P., & Das, I. (2001). Phylogenetic relationships among the Asian tortoises of the genus *Indotestudo* (Reptilia: Testudines: Testudinidae). *HAMADRYAD-MADRAS-*, *26*, 272–275.

Jones, G. (2015). *Species delimitation and phylogeny estimation under the multispecies coalescent*. BioRxiv, 010199, https://doi.org/10.1101/010199

Jones, G. (2017). Algorithmic improvements to species delimitation and phylogeny estimation under the multispecies coalescent. *Journal of Mathematical Biology*, *74*(1–2), 447–467. https://doi.org/10.1007/s00285-016-1034-0

Jones, G., Aydin, Z., & Oxelman, B. (2014). DISSECT: An assignment-free Bayesian discovery method for species delimitation under the multispecies coalescent. *Bioinformatics*, *31*(7), 991–998. https://doi.org/10.1093/bioinformatics/btu770

Kapli, P., Lutteropp, S., Zhang, J., Kobert, K., Pavlidis, P., Stamatakis, A., & Flouri, T. (2017). Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. *Bioinformatics*, *33*(11), 1630–1638. https://doi.org/10.1093/bioinformatics/btx025

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

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

Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Pääbo, S., Villablanca, F. X., & Wilson, A. C. (1989). Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences*, *86*(16), 6196–6200. https://doi.org/10.1073/pnas.86.16.6196

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, *33*(7), 1870–1874. https://doi.org/10.1093/molbev/msw054

Kunz, W. (2013). *Do species exist?: Principles of taxonomic classification*. Weinheim, Germany: John Wiley & Sons.

Lanfear, R., Calcott, B., Ho, S. Y., & Guindon, S. (2012). PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, *29*(6), 1695–1701. https://doi.org/10.1093/molbev/mss020

Le, M., Raxworthy, C. J., McCord, W. P., & Mertz, L. (2006). A molecular phylogeny of tortoises (Testudines: Testudinidae) based on mitochondrial and nuclear genes. *Molecular Phylogenetics and Evolution*, *40*(2), 517–531. https://doi.org/10.1016/j.ympev.2006.03.003

Leaché, A. D., & Fujita, M. K. (2010). Bayesian species delimitation in West African forest geckos (Hemidactylus fasciatus). *Proceedings of the Royal Society B: Biological Sciences*, *277*(1697), 3071–3077. https://doi.org/10.1098/rspb.2010.0662

Leaché, A. D., Zhu, T., Rannala, B., & Yang, Z. (2018). The spectre of too many species. *Systematic Biology*, *68*(1), 168–181. https://doi.org/10.1093/sysbio/syy051

Leigh, J. W., & Bryant, D. (2015). PopART: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, *6*(9), 1110–1116. https://doi.org/10.1111/2041-210X.12410

Leliaert, F., Verbruggen, H., Vanormelingen, P., Steen, F., López-Bautista, J. M., Zuccarello, G. C., & De Clerck, O. (2014). DNA-based species delimitation in algae. *European Journal of Phycology*, *49*(2), 179–196. https://doi.org/10.1080/09670262.2014.904524

Librado, P., & Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, *25*(11), 1451–1452. https://doi.org/10.1093/bioinformatics/btp187

Liu, L., Yu, L., Pearl, D. K., & Edwards, S. V. (2009). Estimating species phylogenies using coalescence times among sequences. *Systematic Biology*, *58*(5), 468–477. https://doi.org/10.1093/sysbio/syp031

Lourenço, J. M., Claude, J., Galtier, N., & Chiari, Y. (2012). Dating cryptodiran nodes: Origin and diversification of the turtle superfamily Testudinoidea. *Molecular Phylogenetics and Evolution*, *62*(1), 496–507. https://doi.org/10.1016/j.ympev.2011.10.022

Loveridge, A., & Williams, E. E. (1957). Revision of the African tortoises and turtles of the suborder Cryptodira. *Bulletin of the Museum of Comparative Zoology*, *115*(6), 161–557.

Luo, A., Ling, C., Ho, S. Y., & Zhu, C. D. (2018). Comparison of methods for molecular species delimitation across a range of speciation scenarios. *Systematic Biology*, *67*(5), 830–846. https://doi.org/10.1093/sysbio/syy011

Mallet, J. (2013). Concepts of species. In S. Levin (Ed.), *Encyclopedia of biodiversity*, Vol. *6* (pp. 679–691). Boston, MA: Academic Press.

Malone, J. H., & Fontenot, B. E. (2008). Patterns of reproductive isolation in toads. *PLoS ONE*, *3*(12), e3900. https://doi.org/10.1371/journal.pone.0003900

Martin, D. P., Murrell, B., Golden, M., Khoosal, A., & Muhire, B. (2015). RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evolution*, *1*(1), vev003. https://doi.org/10.1093/ve/vev003

Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). *Creating the CIPRES Science Gateway for inference of large phylogenetic trees*. Gateway Computing Environments Workshop (GCE), *2010*, 1–8.

Miralles, A., & Vences, M. (2013). New metrics for comparison of taxonomies reveal striking discrepancies among species delimitation methods in Madascincus lizards. *PLoS ONE*, *8*(7), e68242. https://doi.org/10.1371/journal.pone.0068242

Mysara, M., Vandamme, P., Props, R., Kerckhof, F. M., Leys, N., Boon, N., … Monsieurs, P. (2017). Reconciliation between operational taxonomic units and species boundaries. *FEMS Microbiology Ecology*, *93*(4), fix029. https://doi.org/10.1093/femsec/fix029

Nagy, Z. T., Sonet, G., Glaw, F., & Vences, M. (2012). First large-scale DNA barcoding assessment of reptiles in the biodiversity hotspot of Madagascar, based on newly designed COI primers. *PLoS ONE*, *7*(3), e34506. https://doi.org/10.1371/journal.pone.0034506

Pääbo, S. (1990). Amplifying ancient DNA. In M. A. Innis, D. H. Gelfand, J. J. Sninsky, & T. J. White (Eds.), *PCR-protocols and applications – A laboratory manual* (pp. 159–166). San Diego, CA: Academic Press.

Palumbi, S. R., Martin, A., Romano, S., Mcmillan, W. O., Stice, L., & Grabowski, G. (1991). *The simple fool's guide to PCR. A collection of PCR protocols, version 2*. Honolulu, HI: University of Hawaii.

Petzold, A., Vargas-Ramirez, M., Kehlmaier, C., Vamberger, M., Branch, W. R., Du Preez, L., … Fritz, U. (2014). A revision of African helmeted terrapins (Testudines: Pelomedusidae: Pelomedusa), with descriptions of six new species. *Zootaxa*, *3795*(5), 523–548. https://doi.org/10.11646/zootaxa.3795.5.2

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., … Vogler, A. P. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, *55*(4), 595–609. https://doi.org/10.1080/10635150600852011

Preheim, S. P., Perrotta, A. R., Martin-Platero, A. M., Gupta, A., & Alm, E. J. (2013). Distribution-based clustering: Using ecology to refine the operational taxonomic unit. *Journal of Applied & Environmental Microbiology*, *79*(21), 6593–6603. https://doi.org/10.1128/AEM.00342-13

Puillandre, N., Lambert, A., Brouillet, S., & Achaz, G. (2012). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology*, *21*(8), 1864–1877. https://doi.org/10.1111/j.1365-294X.2011.05239.x

Python Software Foundation (2010). *Python Language Reference, version 2.7*. http://www.python.org

Rambaut, A. (2012). http://tree.bio.ed.ac.uk/software/figtree/

Rambaut, A., Suchard, M. A., Xie, D., & Drummond, A. J. (2014). *Tracer 1.6*. http://beast.bio.ed.ac.uk/tracer

Rannala, B., & Yang, Z. (2003). Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics*, *164*(4), 1645–1656.

Rato, C., Carranza, S., Perera, A., Carretero, M. A., & Harris, D. J. (2010). Conflicting patterns of nucleotide diversity between mtDNA and nDNA in the Moorish gecko, Tarentola Mauritanica. *Molecular Phylogenetics and Evolution*, *56*(3), 962–971. https://doi.org/10.1016/j.ympev.2010.04.033

Reid, N. M. (2014). *bGMYC: a Bayesian MCMC implementation of the general mixed Yule-coalescent model for species delimitation*. R package version: 1(2).

Reid, N. M., & Carstens, B. C. (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*(1), 196. https://doi.org/10.1186/1471-2148-12-196

Rhodin, A. G. J., Iverson, J. B., Bour, R., Fritz, U., Georges, A., Shaffer, H. B., & van Dijk, P. P. (2017). Turtles of the world: annotated checklist and atlas of taxonomy, synonymy, distribution, and conservation status (8th ed.). In:A. G. J. Rhodin, J. B. Iverson, P. P. van Dijk, R. A. Saumure, K. A. Buhlmann, P. C. H. Pritchard, & 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. 1–292). Lunenburg, Canada: Chelonian Research Monographs.

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*(3), 539–542. https://doi.org/10.1093/sysbio/sys029

Rubinoff, D., & Holland, B. S. (2005). Between two extremes: Mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. *Systematic Biology*, *54*(6), 952–961. https://doi.org/10.1080/10635150500234674

Scudder, G. G. E. (1974). Species concepts and speciation. *Canadian Journal of Zoology*, *52*(9), 1121–1134. https://doi.org/10.1139/z74-150

Shaw, K. L. (2002). Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: What mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. *Proceedings of the National Academy of Sciences*, *99*(25), 16122–16127. https://doi.org/10.1073/pnas.242585899

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stöver, B. C., & Müller, K. F. (2010). TreeGraph 2: Combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinformatics*, *11*(1), 7. https://doi.org/10.1186/1471-2105-11-7

Stuart, B. L., & Parham, J. F. (2004). Molecular phylogeny of the critically endangered Indochinese box turtle (Cuora galbinifrons). *Molecular Phylogenetics and Evolution*, *31*(1), 164–177. https://doi.org/10.1016/S1055-7903(03)00258-6

Sukumaran, J., & Knowles, L. L. (2017). Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences*, *114*(7), 1607–1612. https://doi.org/10.1073/pnas.1607921114

Swofford, D. L. (1998). *Paup 4.0 beta version for windows: Phylogenetic analysis using parsimony*. Sunderland, MA: Sinauer Associates.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, *123*(3), 585–595.

Tollis, M., DeNardo, D. F., Cornelius, J. A., Dolby, G. A., Edwards, T., Henen, B. T., … Kusumi, K. (2017). The Agassiz's desert tortoise genome provides a resource for the conservation of a threatened species. *PLoS ONE*, *12*(5), e0177708. https://doi.org/10.1371/journal.pone.0177708

Townsend, T. M., Alegre, R. E., Kelley, S. T., Wiens, J. J., & Reeder, T. W. (2008). Rapid development of multiple nuclear loci for phylogenetic analysis using genomic resources: An example from squamate reptiles. *Molecular Phylogenetics and Evolution*, *47*(1), 129–142. https://doi.org/10.1016/j.ympev.2008.01.008

Vaidya, G., Lohman, D. J., & Meier, R. (2011). SequenceMatrix: Concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics*, *27*(2), 171–180. https://doi.org/10.1111/j.1096-0031.2010.00329.x

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

Vargas-Ramirez, M., Carr, J. L., & Fritz, U. (2013). Complex phylogeography in *Rhinoclemmys melanosterna*: Conflicting mitochondrial and nuclear evidence suggests past hybridization (Testudines: Geoemydidae). *Zootaxa*, *3670*(2), 238–254. https://doi.org/10.11646/zootaxa.3670.2.8

Vargas-Ramírez, M., Vences, M., Branch, W. R., Daniels, S. R., Glaw, F., Hofmeyr, M. D., … Fritz, U. (2010). Deep genealogical lineages in the widely distributed African helmeted terrapin: Evidence from mitochondrial and nuclear DNA (Testudines: Pelomedusidae: *Pelomedusa subrufa*). *Molecular Phylogenetics and Evolution*, *56*, 428–440. https://doi.org/10.1016/j.ympev.2010.03.019

Whiting, A. S., Bauer, A. M., & Sites, J. W. Jr (2003). Phylogenetic relationships and limb loss in sub-Saharan African scincine lizards (Squamata: Scincidae). *Molecular Phylogenetics and Evolution*, *29*(3), 582–598. https://doi.org/10.1016/S1055-7903(03)00142-8

Xia, X. (2013). DAMBE5: A comprehensive software package for data analysis in molecular biology and evolution. *Molecular Biology and Evolution*, *30*(7), 1720–1728. https://doi.org/10.1093/molbev/mst064

Yang, Z. (2015). The BPP program for species tree estimation and species delimitation. *Current Zoology*, *61*(5), 854–865. https://doi.org/10.1093/czoolo/61.5.854

Yang, Z., & Rannala, B. (2010). Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences*, *107*(20), 9264–9269. https://doi.org/10.1073/pnas.0913022107

Yang, Z., & Rannala, B. (2017). Bayesian species identification under the multispecies coalescent provides significant improvements to DNA barcoding analyses. *Molecular Ecology*, *26*(11), 3028–3036. https://doi.org/10.1111/mec.14093

Zhang, J., Kapli, P., Pavlidis, P., & Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*, *29*, 2869–2876. https://doi.org/10.1093/bioinformatics/btt499

Zheng, Y., Peng, R., Kuro-o, M., & Zeng, X. (2011). Exploring patterns and extent of bias in estimating divergence time from mitochondrial DNA sequence data in a particular lineage: A case study of salamanders (Order Caudata). *Molecular Biology and Evolution*, *28*(9), 2521–2535. https://doi.org/10.1093/molbev/msr072

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1. Map showing the distribution ranges of the three currently recognised subspecies of *Psammobates tentorius* based on previous records: *P. t. tentorius* (green), *P. t. verroxii* (blue) and *P. t. trimeni* (red).

Figure S2. The map shows the distribution range of the seven clades retrieved from the phylogenetic analyses, in relation to geographic topology.

Figure S3. Left part: BEAST (STACEY) Bayesian multispecies coalescent species tree (PP > 0.95 at all nodes, not shown), with the four values on the left side of each node indicating the posterior probabilities generated from the BPP species delimitation analysis.

Figure S4. TCS networks for the *Cytb* and *ND4* genes.

Figure S5. Multivariate results of motif PCA analysis of the mtDNA dataset (*12S*, *16S*, *Cytb*, *ND4*, *tRNA-His* and *tRNA-Ser*) derived from DAMBE.

Figure S6. The bGMYC pairwise posterior probability matrix visualized as a heat map with the BEAST MSC ultra-metric tree (retrieved from mtDNA).

Figure S7. The plots are for visualizing three different parameters and log posterior values derived from an MCMC run using the single tree generated from BEAST (STACEY) BI analysis).

Figure S8. The plots are for visualizing three different parameters and log posterior values derived from an MCMC run using the 100 random trees generated from BEAST (STACEY) BI analysis).

Figure S9. The plots for visualizing the rate of branching for the coalescent process across all generations of MCMC runs (top).

Table S1. List of all samples, their corresponding localities and NCBI GeneBank accession numbers across different genes.

Table S2. GeneBank accession numbers for all outgroups used in the study.

Table S3. Primers used in the study with corresponding oligo sequences, optimized annealing temperatures and sources.

Table S4. Optimal partition scheme, substitution model, likelihood score (-InL), Gamma shape, proportion of estimated invariant and Homogeneity Test results.

Table S5. The uncorrected *p*-distance matrix for the *Cytb* gene.

Table S6. The average number of pairwise differences among the seven clades retrieved from Arlequin analyses results.

Table S7. Summary of the multiple species delimitation approaches used in the study (The detail given in the text and Table 1).

File S1. The DNA sequence alignment of *12S* gene.

File S2. The DNA sequence alignment of *16S* gene.

File S3. The DNA sequence alignment of *Cytb* gene.

File S4. The DNA sequence alignment of *ND4* gene.

File S5. The DNA sequence alignment of *tRNA–His* & *tRNA-Ser*.

File S6. The DNA sequence alignment of *PRLR* gene.

File S7. The concatenated mtDNA dataset with outgroup sequences.

File S8. The mtDNA dataset used in GMYC methods.

How to cite this article: Zhao Z, Heideman N, Grobler P, Jordaan A, Bester P, Hofmeyr MD. Unraveling the diversification and systematic puzzle of the highly polymorphic *Psammobates tentorius* (Bell, 1828) complex (Reptilia: Testudinidae) through phylogenetic analyses and species delimitation approaches. *J Zool Syst Evol Res*. 2020;58:308–326. https://doi.org/10.1111/jzs.12338