Diversity of North American map and sawback turtles
(Testudines: Emydidae: *Graptemys*)

Peter Praschag\textsuperscript{1,*} | Flora Ihlow\textsuperscript{2,3,*} | Morris Flecks\textsuperscript{2} | Melita Vamberger\textsuperscript{3} | Uwe Fritz\textsuperscript{3}

\textsuperscript{1}Turtle Island, Graz, Austria
\textsuperscript{2}Herpetology Section, Zoologisches Forschungsmuseum Alexander Koenig (ZFMK), Bonn, Germany
\textsuperscript{3}Museum of Zoology, Senckenberg Dresden, Dresden, Germany

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

Map turtles of the genus *Graptemys* are native to North America, where a high degree of drainage endemism is believed to have shaped current diversity. With 14 species and one additional subspecies, *Graptemys* represents the most diverse genus in the family Emydidae. While some *Graptemys* species are characterized by pronounced morphological differences, previous phylogenetic analyses have failed yet to confirm significant levels of genetic divergence for many taxa. As a consequence, it has been debated whether *Graptemys* is taxonomically inflated or whether the low genetic divergence observed reflects recent radiations or ancient hybridization. In this study, we analysed three mtDNA blocks (3228 bp) as well as 12 nuclear loci (7844 bp) of 89 specimens covering all species and subspecies of *Graptemys*. Our analyses of the concatenated mtDNA sequences reveal that the widespread *G. geographica* constitutes the sister taxon of all other *Graptemys* species. These correspond to two clades, one comprised of all broad-headed *Graptemys* species and another clade containing the narrow-headed species. Most species of the broad-headed clade are reciprocally monophyletic, except for *G. gibbonsi* and *G. pearlensis*, which are not differentiated. By contrast, in the narrow-headed clade, many currently recognized species are not monophyletic and divergence is significantly less pronounced. Haplotype networks of phased nuclear loci show low genetic divergence among taxa and many shared haplotypes. Principal component analyses using coded phased nuclear DNA sequences revealed eight distinct clusters within *Graptemys* that partially conflict with the terminal mtDNA clades. This might be explained by male-mediated gene flow across drainage basins and female philopatry within drainage basins. Our results support that *Graptemys* is taxonomically oversplit and needs to be revised.

1 | INTRODUCTION

With 15 currently recognized species-level taxa (14 species and one additional subspecies), map and sawback turtles (*Graptemys* Agassiz, 1857) represent the most diverse genus in the family Emydidae (Ennen, Lovich, Kreiser, Selman, & Qualls, 2010; Ernst & Lovich, 2009; Lamb, Lydeard, Walker, & Gibbons, 1994; Lindeman, 2013; TTWG, 2014). *Graptemys* species are small- to medium-sized freshwater turtles reaching a maximum carapace length of 33 cm (Ernst & Lovich, 2009). The carapace of most species is characterized by knobby or elevated, staggered vertebral scutes, giving

*Both authors contributed equally.*
a characteristic serrated profile (Figure 1) and being eponymous for their common name ‘sawback turtles’. Some species exhibit an extreme sexual size and shape dimorphism, with large molluscivorous females with grotesquely enlarged heads and significantly smaller narrow-headed males, reaching only about half of the females’ size. Map and sawback turtles are predominantly riverine and found throughout eastern and central North America (Ernst & Lovich, 2009; Lindeman, 2013; TTWG, 2014; Figure 1), where they largely inhabit the drainages of the Gulf of Mexico (Ernst & Lovich, 2009; Lamb et al., 1994; Lindeman, 2013; TTWG, 2014). While Graptemys geographic a, G. ouachitensis and G. pseudogeographica are widely distributed, the remaining species are confined to single drainage systems, with closely related species having allopatric or parapatric ranges (Ernst & Lovich, 2009; Lamb et al., 1994; Lindeman, 2013; TTWG, 2014). Considering the high degree of drainage endemism combined with these turtles’ aquatic nature, historical alterations of river courses are thought to have shaped speciation in Graptemys (Lamb et al., 1994; Lindeman, 2013). Recently, all aspects of natural history, distribution and systematics of Graptemys species have been exhaustively reviewed by Lindeman (2013) and the interested reader is referred to this monograph, also for detailed distribution maps.

Over the past decades, the number of Graptemys species increased steadily. However, the validity of several taxa has also been debated for a long time (e.g. Ennen, Kreiser, Qualls, & Lovich, 2010; Ennen et al., 2014; Freeman, 1970; Lindeman, 2003, 2013; Vogt, 1980, 1993) and in many cases species boundaries remain unclear. Most descriptions of new taxa were exclusively or mainly based on more or less pronounced morphological differences, typically coloration and pattern (e.g. Ennen, Lovich, et al., 2010; Folkerts & Mount, 1969; Lovich & McCoy, 1992; see also the review in Lindeman, 2013). An early study using mitochondrial markers found shallow genetic divergences, with only three clearly distinct clades corresponding to G. geographic a, a broad-headed and a narrow-headed species group (Lamb et al., 1994) and otherwise largely unresolved relationships. Since then, no further genetic investigation using broad taxon sampling has been conducted. Later studies focusing on relationships within emydid turtles in general or on individual Graptemys species applied additional genetic markers, including microsatellites and nuclear genes, but revealed also only little differentiation compared to other emydid genera (e.g. Ennen, Kreiser, et al., 2010; Ennen, Lovich, et al., 2010; Ennen et al., 2014; Fritz et al., 2012; Spinks, Thomson, McCartney-Melstad, & Shaffer, 2016; Wiens, Kuczynski, & Stephens, 2010).

This low genetic divergence suggests that the genus might be taxonomically inflated (Stephens & Wiens, 2003; Walker & Avise, 1998) and that morphological variation could be rather associated with resource use and environmental factors or that variation is population-specific and not taxonomic. To shed new light onto this question, we analyse three mitochondrial DNA (mtDNA) blocks (3228 bp) and 12 nuclear loci (7844 bp) using 89 specimens representing all currently recognized species and subspecies of Graptemys, including the recently synonymized subspecies G. nigrinoda delticola (Ennen et al., 2014). In particular, we aim at assessing genetic and taxonomic differentiation and whether the genus is taxonomically oversplit.

**FIGURE 1** Geographical distribution of all presently recognized Graptemys species across eastern North America. Hatching and cross-hatching indicate range overlap. Circles mark localities of samples used in the present study. Inset: Graptemys pearlensis, Bogue Chitto River, Louisiana. Photograph: F. Ihlow
2 MATERIAL AND METHODS

2.1 Taxon sampling

Samples were mainly taken from long-term captives housed in the live collection of Peter and Reiner Praschag in Graz, Austria. These turtles have been collected in the wild between the 1970s and the early 2000s. Deceased individuals have been or will be deposited in the herpetological collection of the Natural History Museum Vienna, Austria, upon their natural death. Voucher photographs of all specimens are stored in the Museum of Zoology, Senckenberg Dresden, Germany, and the Natural History Museum Vienna. Eighty-nine map and sawback turtles were studied, corresponding to all currently recognized species and subspecies plus the recently synonymized subspecies *Graptemys nigrofons delticola* and one hybrid (*G. flavimaculata* × *G. oculifera*). Each taxon was represented by four to ten samples (Table S1). Tissues of living individuals were obtained by clipping off a small piece of the webbing of the toes or by extracting muscle tissue (thigh muscle) from deceased turtles. Samples were preserved in pure ethanol and stored at −20°C until processing. Remaining tissue and DNA samples are permanently kept at −80°C in the tissue sample collection of the Museum of Zoology, Senckenberg Dresden, Germany. For details on the collection sites, see Table S1.

2.2 Gene selection and laboratory procedures

Three mitochondrial DNA blocks coding for 12S ribosomal RNA (367 bp), the NADH dehydrogenase subunits 4L and 4 (ND4L and ND4, 1690 bp) and cytochrome b plus adjacent tRNA (CYTB, 1171 bp) were analysed. All of these genes have previously been successfully used for inferring taxonomy, phylogenetic and phylogeographic relationships of emydid turtles (e.g., Ennen, Lovich, et al., 2010; Fritz et al., 2012; Parham et al., 2013; Spinks et al., 2016; Wiens et al., 2010). In addition, twelve nuclear markers that have been demonstrated to be informative in the closely related genera *Pseudemys* (Spinks et al., 2013) and *Trachemys* (Fritz et al., 2012) were selected. These include coding and non-coding DNA of the following loci: oocyte maturation factor *Mos* (CMOS, 563 bp), high mobility group protein B2 (HMGB2, 550 bp), hepatocyte nuclear factor 1-alpha (HNF-1α, 844 bp), KIAA0398 protein (NB22519, 669 bp), ornithine decarboxylase (ODC, 446 bp), 26S protease regulatory subunit 4 (P26S4, 744 bp), RNA fingerprint fragment 35 (R35, 987 bp), recombination activation genes 1 and 2 (RAG1 and RAG2, 606 and 593 bp, respectively) and three anonymous loci TB01 (572 bp), TB73 (639 bp) and TB86 (631 bp).

Total genomic DNA was extracted using the innuPREP DNA Mini Kit (Analytik Jena AG, Jena, Germany). As sequencing of mtDNA blocks proved to be challenging in *Trachemys* species due to the presence of nuclear mitochondrial insertions (numts; Fritz et al., 2012), mtDNA was amplified for representatives of each taxon using long-range PCR. Following Fritz et al. (2012), the resulting 5.6-kb-long mtDNA sequence was used for sequencing the targeted individual mtDNA blocks as described below. Additional samples from the same taxa were then processed with the PCR primers for shorter mtDNA fragments (Table S2) corresponding to the genes coding for 12S rRNA, CYTB and ND4/ND4L. Using the specifically designed *Trachemys* primers for CYTB and ND4/ND4L (Fritz et al., 2012; this study), there was no evidence for numts because sequences from long-range PCRs and sequences from specific PCR primers matched perfectly.

For PCR amplification, the primers of Tables S2 and S3 were used. The long-range PCR protocol was the one from Fritz et al. (2012); for obtaining the other amplicons, a final volume of 20 μl was used, containing 1 unit DFS-Taq polymerase (Bioron, Ludwigshafen, Germany) with the buffer recommended by the supplier and a final concentration of 0.25 mM of each dNTP (Fermentas, St. Leon-Rot, Germany), 0.5 μM of the respective primer pair and 10-40 ng of total DNA. PCR conditions are explained in Table S4. PCR products were purified using the ExoSAP-IT enzymatic cleanup (Affymetrix USB, Cleveland, OH, USA; 1:20 dilution; modified protocol: 30 min at 37°C, 15 min at 80°C) and sequenced on an ABI 3730 Genetic Analyser (Life Technologies, Carlsbad, CA, USA) using the primers of Tables S2 and S3, the BigDye Terminator v3.1 Cycle Sequencing Kit (Life Technologies) and purified by gel filtration with the Performa DTR V3 96-Well Short Plate Kit (Edge BioSystems, Gaithersburg, MD, USA) and 400 μl of a 5% Sephadex solution (GE Healthcare Europe GmbH, Freiburg, Germany).

2.3 Data analysis

2.3.1 Analyses of mitochondrial DNA blocks

All sequences were checked using the original electropherograms and subsequently manually aligned in *BioEdit* 7.0.5.2 (Hall, 1999). Homologous sequences of *Chrysemys picta belli* were downloaded from GenBank (accession number KF874616, complete mitochondrial genome) and added as outgroup. The optimal partitioning scheme and the best-fitting nucleotide substitution model (Bayesian information criterion) were evaluated using *PartitionFinder* v1.1.1 (Lanfear, Calcott, Ho, & Guindon, 2012). The final concatenated data set consisted of 3228 bp and was partitioned by gene and codon position, with the following models for partition 1 (12S, tRNA, first codon positions of CYTB, ND4L, ND4): HKY+I; partition 2 (second codon positions of CYTB, ND4L, ND4): HKY+I; and partition 3 (third codon positions of CYTB, ND4L, ND4): HKY+I.
FIGURE 2  Bayesian majority rule consensus tree based on complete taxon sampling for Graptemys (89 specimens) using 3228 bp of mitochondrial DNA. Nodes with posterior probabilities greater than 0.95 are marked by circles; bootstrap support under Maximum Likelihood is coded by filling the circles as explained in the inset. On the right are PCA clusters and mitochondrial clades shown; letters correspond to Figure 3. Outgroup (Chrysemys picta bellii) removed for clarity.
of CYTB, ND4L, ND4): HKY. Phylogenetic trees were inferred using MrBayes 3.2.6 (Ronquist et al., 2012), with model parameters estimated separately for each of the three partitions by unlinking them. Four independent runs were performed with 100 million generations each and sampling every 10,000 trees. However, the analysis was stopped when the average standard deviation of split frequencies fell below 0.01. Results of the MCMC were summarized and the initial 25% of each run discarded as burn-in after checking for convergence and sufficient effective sample sizes in Tracer v1.6 (Rambaut, Suchard, Xie, & Drummond, 2014). In addition, Maximum Likelihood trees were calculated using RAxML 7.2.8 (Stamatakis, 2006) and the GTR+G substitution model across all partitions. Clade support was assessed by bootstrap and thorough bootstrap algorithms.

2.3.2 | Analyses of nuclear markers

Nuclear sequences were phased using PHASE 2.1 (Stephens, Smith, & Donnelly, 2001) and SEQPHASE (Flot, 2010). Relationships of the phased nuclear DNA sequences were assessed by building haplotype networks for each of the 12 loci using statistical parsimony. We used simple gap coding to account for length polymorphisms. Networks were built with the package 'PEGAS' for CRAN R (Paradis, 2010; Templeton, Crandall, & Sing, 1992). In addition, based on coded sequence data, principal component analyses (PCAs) were calculated using the R package ADEGENET (Jombart, 2008) for R v3.2.3. We did not analyse our nuclear data using software enforcing dichotomous splits because we argue that in closely related taxa reticulate patterns are expected that would be completely misleading for any bifurcating representation.

3 | RESULTS

The topology of the Bayesian and Maximum Likelihood trees was identical (Figure 2). In concordance with Lamb et al. (1994), the mtDNA phylogeny corresponds to three clades matching (i) the widespread Graptemys geographica, (ii) the broad-headed species group (G. pulchra group sensu Lamb et al., 1994), comprised of G. barbouri, G. ernsti, G. gibbonsi, G. pearlensis and G. pulchra and (iii) the narrow-headed species group (G. pseudogeographica group sensu Lamb et al., 1994) with G. caglei, G. flavimaculata, G. nigrinoda, G. oculifera, G. ouachitensis, G. pseudogeographica, G. sabinensis and G. versa. The broad-headed and narrow-headed groups are sister clades, and G. geographica is the deeply divergent successive sister taxon of these two clades. Compared to the shallow genetic divergence in the narrow-headed group, divergences within the broad-headed group are slightly more pronounced. However, in both clades, the species are often not reciprocally monophyletic. In the broad-headed clade, there are four terminals. Graptemys gibbonsi and G. pearlensis were not differentiated and were sister to another clade comprised of G. barbouri, G. ernsti and G. pulchra. These three species are reciprocally monophyletic and placed in an unresolved polytomy. Within the clade containing the narrow-headed taxa, sister group relationships were not well resolved, with a basal trichotomy and several subordinated terminal clades. Reciprocally monophyletic clades matched with G. caglei and G. versa, even though the monophyly of G. versa was weakly supported. The remaining terminal clades were comprised of representatives of several taxa, with sequences of G. flavimaculata, G. oculifera, G. ouachitensis and G. sabinensis in one terminal clade and sequences of G. n. nigrinoda, G. n. delticola, G. p. pseudogeographica and G. p. kohnii in the other. The subspecies of G. nigrinoda and G. pseudogeographica were not differentiated, and only very weak differentiation was found for G. ouachitensis and G. sabinensis, which clustered in an unresolved polytomy together with G. flavimaculata and G. oculifera, taxa occurring in neighbouring drainages.

The PCAs using nuclear data revealed eight clusters (A-H in Figures 2 and 3) of which only four are in concordance with current taxonomy and mtDNA clades, namely G. barbouri (cluster A), G. caglei (cluster B), G. geographica (cluster E) and G. versa (cluster G). Graptemys sabinensis, virtually undifferentiated in mtDNA from G. flavimaculata, G. oculifera and G. ouachitensis, constitutes another distinct cluster (H). The remaining three PCA clusters are composed of several species. Cluster C consists of G. ernsti, G. gibbonsi, G. pearlensis and G. pulchra. Cluster D contains G. flavimaculata, G. n. nigrinoda, G. n. delticola and G. oculifera. Cluster F comprises G. ouachitensis, G. p. pseudogeographica and G. p. kohnii (Figure 3). Compared to the mitochondrial phylogeny, there are some conflicts (Figure 2). In particular, within the broad-headed species group, G. ernsti, G. pulchra and the terminal clade consisting of G. gibbonsi and G. pearlensis correspond to one and the same cluster in PCAs (cluster C). Within the narrow-headed clade, a similar pattern is found with respect to G. ouachitensis and the two subspecies of G. pseudogeographica. Using mtDNA, the latter taxa are placed in distinct terminal clades, but using nuclear loci, they appear in the same PCA cluster F. The same is true for G. flavimaculata, G. oculifera and G. nigrinoda. Using mtDNA, G. nigrinoda is assigned to another terminal clade than the other two species, but is placed together with them in PCA cluster D.

When the individual nuclear loci are examined using haplotype networks, only limited variation is revealed (Figure S1), with many shared haplotypes among different taxa. Private haplotypes unique to certain species occur (Table S5), but are often only weakly differentiated. Some nuclear loci reflect
the differentiation pattern of the PCAs, while others show too little variation for species- or group-specific patterns.

4 | DISCUSSION

Using for the first time multiple nuclear loci and a complete taxon sampling, our results confirm that most of the morphologically defined species-level taxa within *Graptemys* are genetically poorly differentiated. Based on mtDNA data (3228 bp), there are three clearly distinct clades, corresponding to nine reciprocally monophyletic terminal clades. Among these nine terminal clades, six match with morphologically defined species-level taxa, while the remaining three clades are comprised of representatives of two to four distinct taxa (Figure 2). Unlike Wiens et al. (2010), who suggested that the short mitochondrial branch lengths of *Graptemys* could result from sequencing numts, we found no evidence for such biases using our long-PCR approach. With respect to the 12 nuclear loci (together 7844 bp), PCAs revealed eight genetic clusters that partly conflict with the terminal mitochondrial clades. Congruency between mitochondrial clades and nuclear genomic PCA clusters was found with respect to *G. geographica*, *G. barbouri*, *G. caglei* and *G. versa*. Within the broad-headed clade, another PCA cluster comprises the remaining species *G. gibbonsi* and *G. pearlensis* (which are also mitochondrialy not differentiated) plus *G. ernsti* and *G. pulchra* (which are mitochondrialy differentiated). Within the narrow-headed clade, there were, besides *G. caglei* and *G. versa*, three further PCA clusters revealed that conflict with mitochondrial clades. Two of these PCA clusters contain taxa of two different terminal mtDNA clades. One cluster contains *G. flavimaculata*, the two subspecies of *G. nigrinoda*, and *G. oculifera* – taxa that have been treated in the past as conspecific by some authors (Wermuth & Mertens, 1961). The second PCA cluster includes *G. ouachitensis* and the two subspecies of *G. pseudogeographica*, again taxa that have been regarded for a long time as conspecific by many authors (see the detailed review in Lindeman, 2013). The third conflicting cluster refers to *G. sabinensis*, which is only negligibly differentiated in mtDNA, but distinct in PCAs.

We argue that the taxonomic status of *G. geographica*, *G. barbouri*, *G. caglei* and *G. versa* as distinct species is well supported, as concordantly indicated by mitochondrial and nuclear DNA markers and morphology. However, with respect to all other taxa, things are less clear, in agreement with the view that map and sawback turtles are taxonomically oversplit (Walker & Avise, 1998) and need to be revised. Our nuclear PCA data suggest that it should be considered to reinstate for *G. ouachitensis* the status of a subspecies of *G. pseudogeographica* and for *G. flavimaculata*, *G. nigrinoda* and *G. oculifera* the status of subspecies of *G. oculifera*, because these taxa correspond to only one cluster each. Moreover, within the broad-headed group, *G. ernsti*, *G. gibbonsi*, *G. pearlensis* and *G. pulchra* match with one and the same PCA cluster and are, thus, likely conspecific. In contrast to Ennen, Lovich, et al. (2010), we found *G. gibbonsi* and *G. pearlensis* genetically undifferentiated, both in nuclear and in mitochondrial DNA. Ennen, Lovich, et al. (2010),
who described *G. pearlensis* as a new species, used sequence variation of the mitochondrial control region (and the ND4 gene) and found two samples of *G. gibbonsi* and three samples of *G. pearlensis* constituting reciprocally monophyletic sister clades. Considering the confinement of map turtles to drainage systems, differentiation in rapidly evolving genetic markers, like the control region, is not unexpected and was perhaps overinterpreted. This is also supported by the recently reported differences between populations of *G. ernsti* using morphology, microsatellite loci and rapidly evolving mtDNA sequences (control region plus ND4 sequences: Ennen et al., 2016). Coloration and pattern, important characters for diagnosing *Graptemys* taxa morphologically (cf. Ernst & Lovich, 2009; Lindeman, 2013), probably also reflect such drainage-dependent population differentiation. Moreover, these characters are, at least partially, influenced by incubation conditions of eggs (Ewert, 1979; Vogt, 1980, 1993). Thus, our original hypothesis that some *Graptemys* taxa represent rather population-level than taxonomic variation has been confirmed. In the same vein, using morphology and mtDNA variation, Ennen et al. (2014) concluded that the two morphologically defined subspecies of *G. nigrinoda*, which were thought to occur within the same drainage system (Folkerts & Mount, 1969), are taxonomically not distinct. This finding is in line with our observation of the complete lack of genetic differentiation between these two taxa.

The conflicts between mitochondrial clades and PCA clusters in our data set (Figures 2 and 3) are likely related to the confinement of map and sawback turtles to different drainage systems and to female philopatry and largely male-mediated gene flow. It is obvious that such conflicts occur in taxa with weak genetic divergence having either wide and overlapping or abutting distributions (*G. ouachitensis*, *G. pseudogeographica*, *G. sabinensis*) or inhabiting neighbouring drainage basins (*G. flavimaculata*, *G. nigrinoda* and *G. oculifera*; *G. ernsti*, *G. gibbonsi*, *G. pearlensis* and *G. pulchra*). The mitochondrial similarity of *G. sabinensis* to *G. ouachitensis*, despite being distinct in PCAs of nuclear data, suggests that mitochondrial introgression could play a role here, as known for many animal species (cf., for instance, the review in Currat, Ruedi, Petit, & Excoffier, 2008), among them also freshwater turtles (Fritz et al., 2008; Irlow et al., 2016; Vamberger et al., 2017). *Graptemys sabinensis* and *G. ouachitensis* occur in neighbouring drainage basins (Figure 1), and such introgression events could easily occur during flood events or stream capture.

**ACKNOWLEDGEMENTS**

We are grateful to Reiner Praschag for providing samples for this study. Thanks for laboratory work go to Anke Müller. Oliver Hawlitschek and an anonymous reviewer made helpful comments on a previous version of this study.

**REFERENCES**


**SUPPORTING INFORMATION**

Additional Supporting Information may be found online in the supporting information tab for this article.