Molecular Phylogenetics and Evolution 63 (2012) 656–667

Contents lists available at SciVerse ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Species boundaries and phylogenetic relationships in the critically endangered Asian box turtle genus *Cuora*

Phillip Q. Spinks [a,*], Robert C. Thomson [a,2], YaPing Zhang [b], Jing Che [b], Yonghua Wu [c], H. Bradley Shaffer [a,1]

[a] Department of Evolution and Ecology and Center for Population Biology, University of California, Davis, CA 95616, USA
[b] State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China
[c] School of Life Sciences, Northeast Normal University, Changchun City, Jilin Province 130024, China

A R T I C L E I N F O

Article history:
Received 15 November 2011
Revised 9 February 2012
Accepted 12 February 2012
Available online 3 March 2012

Keywords:
Cuora
Geoemydidae
Species tree
Introgression
Lineage sorting
Hybridization

A B S T R A C T

Turtles are currently the most endangered major clade of vertebrates on earth, and Asian box turtles (*Cuora*) are in catastrophic decline. Effective management of this diverse turtle clade has been hampered by human-mediated, and perhaps natural hybridization, resulting in discordance between mitochondrial and nuclear markers and confusion regarding species boundaries and phylogenetic relationships among hypothesized species of *Cuora*. Here, we present analyses of mitochondrial and nuclear DNA data for all 12 currently hypothesized species to resolve both species boundaries and phylogenetic relationships. Our 15-gene, 40-individual nuclear data set was frequently in conflict with our mitochondrial data set; based on its general concordance with published morphological analyses and the strength of 15 independent estimates of evolutionary history, we interpret the nuclear data as representing the most reliable estimate of species boundaries and phylogeny of *Cuora*. Our results strongly reiterate the necessity of using multiple nuclear markers for phylogeny and species delimitation in these animals, including any form of DNA "barcoding", and point to *Cuora* as an important case study where reliance on mitochondrial DNA can lead to incorrect species identification.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

Hybridization is an integral part of the evolutionary process. As multilocus nuclear and mitochondrial genetic data sets continue to accumulate, instances of hybridization and introgression are being found across a wide range of taxa in nature (Bossu and Near, 2009; Burrell et al., 2009; Good et al., 2008; Keck and Near, 2009; Linnen and Farrell, 2007; Melo-Ferreira et al., 2005; Peters et al., 2007; Sang and Zhong, 2000). However, habitat modification, animal translocations and wildlife farming have also led to increased rates of anthropogenic hybridization and introgression between native and introduced animals (Barilani et al., 2005; Kidd et al., 2009; Norén et al., 2006; Rhymer and Simberloff, 1996; Riley et al., 2003). Distinguishing between cases of natural and anthropogenic hybridization is crucial for conservation biology, particularly in taxa that are bred for commercial purposes (Allendorf et al., 2001). A possible consequence of hybridization is the introgression of alleles from one species into another, leading to incongruence among gene trees for the same individuals or taxa (Avise, 1994; Buckley et al., 2006; Moore, 1995; Funk and Omland, 2003; Sang and Zhong, 2000). However, incongruence among gene trees can also be caused by incomplete coalescence and gene duplication and loss (Maddison, 1997), both of which can be difficult to discriminate from introgression (Holder et al., 2001; Joly et al., 2009; Mallet, 2005).

Here we report on the Asian box turtle genus *Cuora*, a phylogenetically and taxonomically enigmatic group of approximately 12 semiaquatic turtle species distributed across southern China through Indonesia and the Philippines (Fig. 1). The species composition of *Cuora* is uncertain and contentious, and the validity of five of the 12 named taxa has been questioned in the last decade. In addition, the phylogenetic relationships among *Cuora* species remains elusive despite numerous molecular analyses from several independent laboratories (He et al., 2007; Honda et al., 2002; Parham et al., 2004; Spinks and Shaffer, 2007; Spinks et al., 2009). Delimiting species boundaries, and recovering phylogenetic relationships among *Cuora* species represent particularly difficult

* Corresponding author. Present addresses: Department of Ecology and Evolutionary Biology, University of California, 621 Charles E. Young Dr. South, Los Angeles, CA 90095-1606, USA. La Kretz Center for California Conservation Science, Institute of the Environment and Sustainability, University of California, La Kretz Hall, Suite 300, 619 Charles E. Young Dr. South, Los Angeles, CA 90095-14966, USA.
E-mail address: pqspinks@ucla.edu (P.Q. Spinks).
[1] Present addresses: Department of Ecology and Evolutionary Biology, University of California, 621 Charles E. Young Dr. South, Los Angeles, CA 90095-1606, USA. La Kretz Center for California Conservation Science, Institute of the Environment and Sustainability, University of California, La Kretz Hall, Suite 300, 619 Charles E. Young Dr. South, Los Angeles, CA 90095-14966, USA.
[2] Present address: Department of Biology, University of Hawaii at Manoa, Honolulu, Hawaii 96822, USA.

1055-7903/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2012.02.014

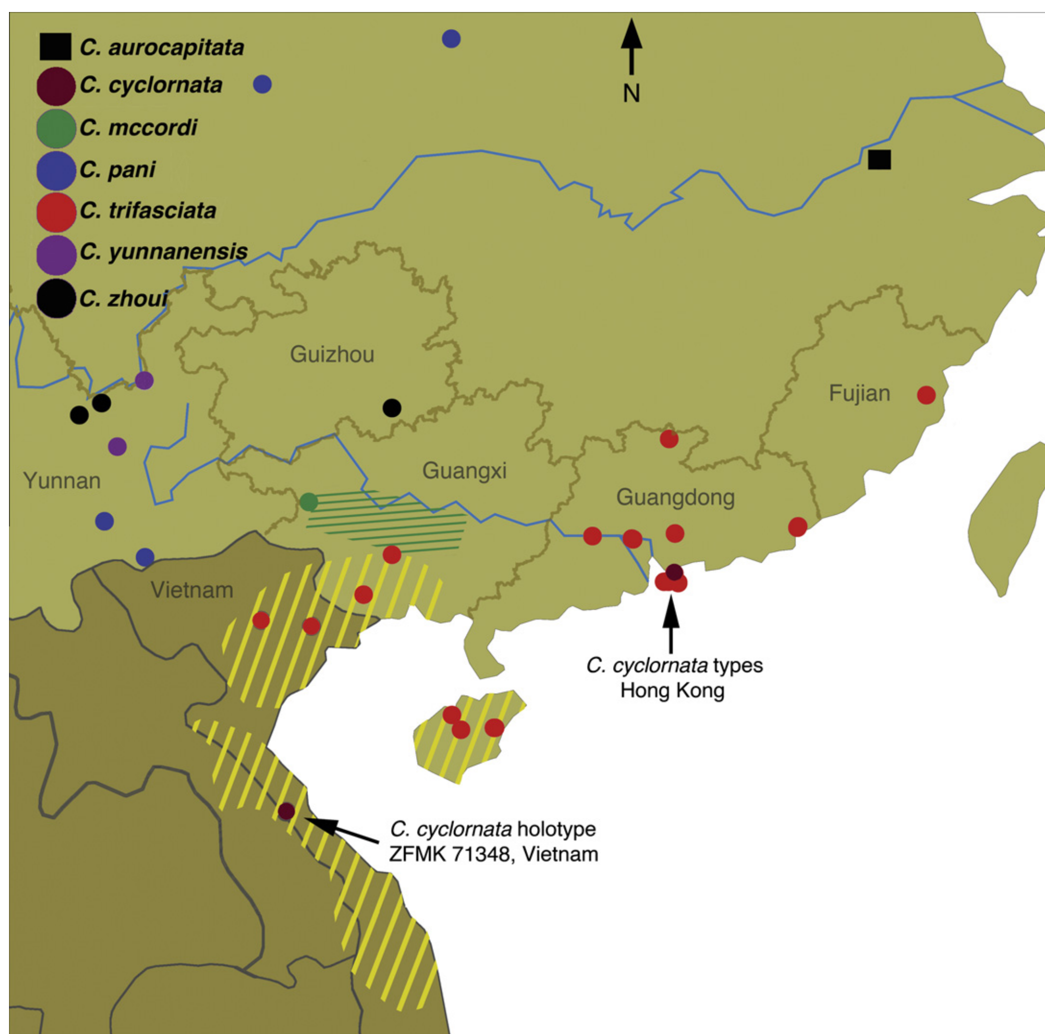

Fig. 1. Map of southern China–Southeast Asia showing collection localities for six *Cuora* species and one putative hybrid taxon (*C. cyclornata*). Green shaded area is the putative range of *C. mccordi* (Zhou et al., 2008). Collection localities for the remaining *Cuora* species are not shown. However, the yellow shaded area indicates the range for members of the *C. galbinifrons* species complex. The holotype of *C. cyclornata* (ZFMK 71348) was collected in central Vietnam (Ziegler and Herrmann, 2000), and individuals matching the phenotype of *C. cyclornata* were collected from the wild in Hong Kong, China (Spinks and Shaffer, 2007). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

challenges because hybridization (natural and anthropogenic) is rampant within *Cuora* (Parham et al., 2001; Parham and Shi, 2001; Shi et al., 2008a; Spinks et al., 2004; Spinks and Shaffer, 2007; Stuart and Parham, 2004; Wink et al., 2001), and because samples from wild-caught specimens with locality data are extremely rare. The over-collection of turtles throughout Asia for the food and traditional Chinese medicine markets and the global pet trade (Gong et al., 2009; van Dijk, 2000) has also placed these turtles in the center of a conservation crisis for Asian turtle species. As of this writing, nine *Cuora* species are critically endangered, two are endangered, and the single remaining species is vulnerable (IUCN, 2011).

Previous mitochondrial DNA (mtDNA) analyses of *Cuora* recovered strong support for most interspecific relationships, but also suggested extensive introgression among members of the *C. trifasciata* species complex (sensu Parham et al., 2004), which includes *C. aurocapitata, C. pani,* and *C. trifasciata*, and *C. cyclornata* sensu Blanck et al. (2006), should that taxon be resolved as valid. In particular, mtDNA analyses of multiple individuals suggested that all four species were non-monophyletic (Parham et al., 2004; Spinks and Shaffer, 2007; Spinks et al., 2009; Stuart and Parham, 2004). On the other hand, phylogenetic analyses of three nuclear loci for the *C. trifasciata* complex were largely uninformative with respect to species monophyly or relationships (Spinks and Shaffer, 2007). The taxonomy of the *C. galbinifrons* species complex (*C. bourreti, C. galbinifrons*, and *C. picturata*, sensu Stuart and Parham, 2004) is also somewhat contentious. Most recent workers recognize all three as species because they are morphologically distinguishable, reciprocally monophyletic and highly divergent based on mtDNA (Stuart and Parham, 2004: Spinks and Shaffer, 2007), although Fritz et al. (2006) recently challenged this conclusion.

Given the propensity for hybridization and introgression among *Cuora* species and between *Cuora* and other genera, testing and confirming hypothesized species validity and detecting patterns of introgression are key goals for both conservation efforts and evolutionary analyses. In response to the clear conservation crisis for most members of the genus, captive breeding programs for endangered *Cuora* are now in place in the US, Europe and Asia (Turtle Conservation Fund, 2002). However, in many cases, the provenance and even the identity of much of the captive breeding stock is uncertain or unknown. Here we use phylogenetic analysis of multiple nuclear and mitochondrial markers, in conjunction with published accounts of phenotypic variation to evaluate species validity, assess hybridization, and construct a well-resolved

phylogeny among *Cuora* species. To explore the extent of mitochondrial introgression among the problematic *C. trifasciata* species complex, we assembled a relatively large single-gene mtDNA data set for both captive assurance colony turtles and all available known-locality *C. trifasciata* complex specimens in our tissue collection. Our taxon sampling also included breeding stock from five assurance colonies (Appendix) plus sequence data from a juvenile *Cuora* collected in Vietnam (ZFMK 71348) that was designated as the holotype for *Cuora cyclornata* (Blanck et al., 2006), a taxon that has also been interpreted as a hybrid (Spinks et al., 2009, see below). Importantly, our sampling included a single individual of *C. yunnanensis*, a recently rediscovered species thought to be extinct (Zhou and Zhao, 2004). To sample additional, non-mitochondrial loci from across the genome, we generated sequence data from up to 16 nuclear loci, and assessed potential hybridization and introgression using phylogenetic trees and networks.

2. Materials and methods

2.1. Sampling strategy

As with most Asian turtle species, our sampling was hindered by the paucity of field-collected samples with reliable locality data. However, we assembled a set of 10 *C. trifasciata* with locality data (Appendix). *Mauremys mutica* is a representative of the sister clade of *Cuora* and was used as the outgroup (Honda et al., 2002; Spinks et al., 2004).

Our previous work includes a mtDNA phylogeny for all species of *Cuora* based on nicotinamide adenine dinucleotide dehydrogenase subunit 4 (ND4) gene sequences (Spinks et al., 2009). However, the number of species contained in the *C. trifasciata* species complex remains somewhat contentious (Blanck et al., 2006; Spinks and Shaffer, 2007; Spinks et al., 2009), and three of our goals in the present paper were to (1) assess mtDNA diversity within *Cuora*, (2) assess the extent of mitochondrial introgression among our samples, and (3) determine the species composition of this complex. Nuclear mitochondrial pseudogenes (numts) of the cytochrome c oxidase subunit I (COI) and ND4 genes are widespread among the *C. trifasciata* species complex, but numts of the ND1 gene have thus far rarely been observed (Spinks and Shaffer, 2007). Therefore, we generated new mtDNA sequence data from the ND1 gene (Appendix). In total, our *ND1* data set contained 137 sequences including 38 sequences downloaded from GenBank plus an additional 99 sequences generated here. Our total sampling included at least two sequences of each *Cuora* species except *C. bourreti*, for which we had only a single individual (one other individual contained a numt), and *C. evelynae*, which is sometimes considered a subspecies of *C. flavomarginata* and sometimes a full species (the most recent literature considers it a subspecies, see Turtle Taxonomy Working Group, 2011). The 99 new sequences generated here included 80 *C. trifasciata*, 18 *C. pani*, and one *C. yunnanensis* (Appendix). Most additional *C. trifasciata* and *C. pani* sequenced here are housed in zoos or private collections, and therefore lack museum voucher specimens. However photo vouchers for most non-museum specimens (e.g. Lehn et al., 2007) are available from the Shaffer lab website (http://www.eeb.ucla.edu/Faculty/Shaffer/).

Our nuDNA data were collected primarily from a subset of the individuals for which we had mtDNA. Our sampling consisted of sequence data from 39 individual turtles including two of each species except *C. yunnanensis* (for which we only have one individual) plus two representatives of each major *C. trifasciata* mtDNA lineage recovered in the *ND1* analysis (Appendix). We assembled nuDNA sequences from 16 loci including three introns generated by us previously (Spinks and Shaffer, 2007), plus data from an additional four introns, six exons and three anonymous loci generated for this analysis (Appendix).

2.2. DNA extraction and sequencing

DNA was extracted from blood or soft tissue samples using a salt extraction protocol (Sambrook and Russell, 2001). Partial sequences of all loci were generated using 20 μl volume PCR reactions, with an initial denaturation of 60 s at 95 °C, followed by 40 cycles of denaturation (94 °C for 30 s), annealing (45 s at 60°–65 °C), and extension (72 °C for 60–90 s) with a final extension period (72 °C for 10 min) (see Table 1 for locus-specific annealing temperatures, extension times and primers). One sample (ZFMK 71348) failed to amplify for any locus and the extracted DNA appeared to be highly degraded (data not shown). We sent this sample to Lakehead University Paleo-DNA laboratory (a lab specializing in DNA extraction from degraded samples) where DNA was successfully extracted from this sample (documentation available from the corresponding author). The DNA from ZFMK 71348 was degraded so we designed sets of short (~300 bp) overlapping primers for *ND1* and *AHR* (our most robust nuclear marker) for PCR and sequencing of this sample (Table 1). For these PCR reactions, we used the same conditions as above, but decreased the extension time to 30 s. We successfully amplified the *ND1* fragments for ZFMK 71348. Although the *AHR* reactions amplified for positive controls, we were unable to amplify ZFMK 71348 for this marker and we took this as an indication that generating nuDNA from this degraded sample would not be possible for this sample. We did not attempt to generate nuclear sequence data for ZFMK 71348 from additional nuclear loci. All PCR products were sequenced by Beckman Coulter Genomics (https://psf.beckmangenomics.com/) except for the *C. yunnanensis* sequences which were generated at the Kunming Institute of Zoology, Kunming China.

2.3. Phylogeny and network analyses

The mtDNA data were partitioned by codon for analysis while the nuDNA were analyzed both as single loci and as a concatenated data set. Due to the large number of possible partitions for our concatenated nuDNA data set, we partitioned these data by locus only. Coding regions (exons, see Table 1) were translated using Geneious v5.1 (Drummond et al., 2011) to check for pseudogenes, and we tested for recombination within each locus using the pairwise homoplasy index (PHI) statistic (Bruen et al., 2006, implemented in SplitsTree v4.10). We reconstructed alleles for each nuclear locus using the Phase2.1.1 software (Stephens et al., 2001; Stephens and Donnelly, 2003) and replicated all analyses on both phased and unphased data. We selected models of molecular evolution for each marker separately using the Akaike information criterion (AIC) in jModelTest v0.1 (Posada, 2008; Guindon and Gascuel, 2003) and performed Bayesian phylogenetic analyses using MrBayes V3.1.2 (Ronquist and Huelsenbeck, 2003). These Bayesian analyses consisted of two runs with four incrementally heated chains per run for 5,000,000 generations, with samples of the posterior distribution drawn every 1000 generations. We checked that the runs had reached stationarity by ensuring that the average standard deviation of split frequencies between independent runs approached 0, and the potential scale reduction factor equaled 1. In addition, we used Tracer (Rambaut and Drummond, 2009) to examine the MCMC output to ensure that all chains were sampling from the same target distribution. The first 25% of samples was discarded as burnin provided the chains had reached stationarity prior to this point.

Because our previous results demonstrated gene-tree discordance between mtDNA and nuDNA phylogenies (Spinks and Shaffer, 2007), we approached the issue of congruence among data

Table 1
Marker name, type and associated model of sequence evolution selected via jModelTest. Also listed are primers, PCR conditions (annealing temperature/extension time) and literature sources for primers.

Marker	Locus	Model	Primers	Temp./time	Primer source
AHR	Exon	JC	AHRex11f3, R3	64.8/60 s	Townsend et al. (2008), Barley et al. (2010)
AIING	Exon	GTR	NGF f50, f30	63/60 s	Sehgal and Lovette (2003)
BDNF	Exon	NA	BDNF F, R	65/60 s	Noonan and Chippindale (2006)
BMP2	Exon	HKY + G	BMP2 f6, r2	60/60 s	Townsend et al. (2008)
HNF-1a	Intron	HKY	HNFAL F, R	65/60 s	Primmer et al. (2002)
HMGB2	Intron	HKY	NB17483_fmod, R2	61/60 s	Backström et al. (2008), Barley et al. (2010)
NB22519	Intron	GTR	NB22519 F, R	60/60 s	Backström et al. (2008)
P26s4	Intron	HKY	NB17367 F, R	62/60 s	Backström et al. (2008)
PAX1P1	Intron	HKY + G	PAXIP1 20F, 21R	61/60 s	Kimball et al. (2009)
R35	Intron	HKY	R35 EX1, EX2	62/90 s	Fujita et al. (2004)
RAG	Exon	HKY	RAG f1/r2, or RAG f2/r2	62/60 s	Krenz et al. (2005)
RELN	Intron	HKY	RELN61 F, R	61/90 s	Spinks and Shaffer (2007)
TB01	Anonymous	JC	TB01 F, R	61/60 s	Thomson et al. (2008)
TB29	Anonymous	JC	TB29 F, R	61/60 s	Thomson et al. (2008)
TB73	Anonymous	JC	TB73 F, R	61/60 s	Thomson et al. (2008)
ZEB2	Exon	HKY + G	ZFHX1B_f1, r2	61/90 s	Townsend et al. (2008)
ND1	mtDNA	GTR + G	ND1F, R	62/60 s	Spinks and Shaffer (2007)
ND1b	mtDNA		F primer: Cuora_ND1_Fb 5′-CCATCTCATACGAAGTAACCC-3′	54/30 s	This study
	mtDNA		R primer: Cuora_ND1_Rd 5″-GAGTGTTTATTATTAGGATG -3′		
ND1c	mtDNA		F primer: Cuora_ND1_Fd 5′-CATCCTAATAATAAACACTC-3′	60/30 s	This study, Spinks and Shaffer (2007)
	mtDNA		R primer: ND1R		
ND1d	mtDNA		F primer: Cuora_ND1_Fa 5′-TTCATCAATKACATTATTTAC-3′	58/30 s	This study
	mtDNA		R primer: Cuora_ND1_Rc 5′-GGTCAGGAGGAGAATATTAAG-3′		

partitions (i.e. mtDNA vs. nuDNA) in two ways. First, we assessed congruence between our mtDNA and nuDNA phylogenies using a Bayesian test of monophyly following the strategy outlined by Linnen and Farrell (2007). We filtered the post-burnin trees from our concatenated nuDNA analysis against a constraint tree arising from the mtDNA analysis. The fraction of the posterior sample of nuDNA trees that are congruent with the mtDNA can be taken as a measure of the probability that the nuDNA and mtDNA datasets are congruent. If the two datasets are strongly incongruent, then combining them into a single concatenated analysis is inappropriate. To determine those parts of the mtDNA and concatenated nuDNA trees that might be congruent, we used PAUP* to generate agreement subtrees from the mtDNA and concatenated nuDNA Bayesian consensus trees. For this final analysis, we excluded one mitochondrial clade of *C. trifasciata* that our previous work demonstrated comprised introgressed *C. aurocapitata* mitotypes (clade A from Spinks et al., 2009).

Phylogenetic reconstructions are predicated on the assumption that nucleotide sequence data are generated through a branching process. However, recombination and introgression can produce reticulate relationships, patterns more appropriately visualized using phylogenetic networks (Huson, 1998). Therefore, we generated phylogenetic networks from individual nuclear loci, and from our concatenated nuDNA data set. We generated phylogenetic networks with Splitstree v4.10 (Huson, 1998; Huson and Bryant, 2006), utilizing the NeighborNet algorithm (Bryant and Moulton, 2004) with uncorrected "p" genetic distances, and bootstrap analyses with 1000 pseudoreplicates.

Cuora cyclornata has been interpreted as both a valid species (Blanck et al., 2006) and as a hybrid taxon (Spinks et al., 2009). To assess the impact of this potentially hybrid taxon on our results, we repeated all phylogenetic and network analyses on the concatenated nuDNA data after excluding our three *C. cyclornata* samples (samples BCC09, KFBG 030567630, KFBG 030552847).

3. Results

3.1. mtDNA phylogenies

Our *ND1* data set was composed of up to 726 base pairs (bp) for each of 137 individuals (135 *Cuora*, and 2 *M. mutica* outgroups). The matrix was almost complete with 0.4% missing data (see Appendix for GenBank accession numbers). The majority-rule consensus of the posterior distribution of trees from the Bayesian analysis was nearly equivalent to previous results (Spinks et al., 2009), with a few important differences. All *Cuora* species were recovered as monophyletic except those of the *C. trifasciata* species complex, which fell into four topologically disjunct, but individually well-supported clades, two of which are novel with respect to our previous results (Fig. 2). As in previous analyses, the *C. cyclornata* holotype sequence (ZFMK 71348) fell into a topologically disjunct clade along with four additional *C. cyclornata* sequences and one *C. trifasciata* sequence; together, they formed the sister clade (clade C, Fig. 2) to the *C. galbinifrons* species complex. However, 15 *C. trifasciata* sequences formed a novel sister clade to the *C. zhoui* sequences with strong support (clade E, Fig. 2). In addition, 19/20 *C. pani* formed a novel sister clade (clade D) to a composite clade containing 41 *C. trifasciata*, three *C. aurocapitata*, and the remaining *C. pani* (clade A, Fig. 2). The remaining 34 *C. trifasciata* (clade B) formed the sister clade to clades A + D (Fig. 2). Finally, all of the remaining species for which we had more than one sample were monophyletic with strong support (Fig. 2).

3.2. Features of the nuDNA data set

We generated sequence data from up to 16 nuclear loci for 40 individual turtles (GenBank accession numbers provided in the Appendix). Sequencing chromatograms for several individuals were partially mixed at three loci (*HNFL*, RELN and TB73), and displayed patterns indicative of heterozygous length polymorphisms (see Bhangale et al., 2005). We used the Indelligent v.1.2 software (Dmitriev and Rakitov, 2008) to reconstruct nucleotide sequences from chromatograms disrupted by heterozygous length polymorphisms (available at http://ctap.inhs.uiuc.edu/dmitriev/indel.asp). In addition, we detected a significant level of recombination within the RELN locus ($P = 0.016$) for the two *C. mccordi* samples, but no significant recombination within the remaining 15 loci ($P \geqslant 0.12$). To explore the effect of this recombination, we performed a phylogenetic network analysis on the RELN locus and all species were recovered as monophyletic, except for the two samples of *C. mccordi*, which were widely disjunct from one another in the network (not shown). We then excluded the *C. mccordi*

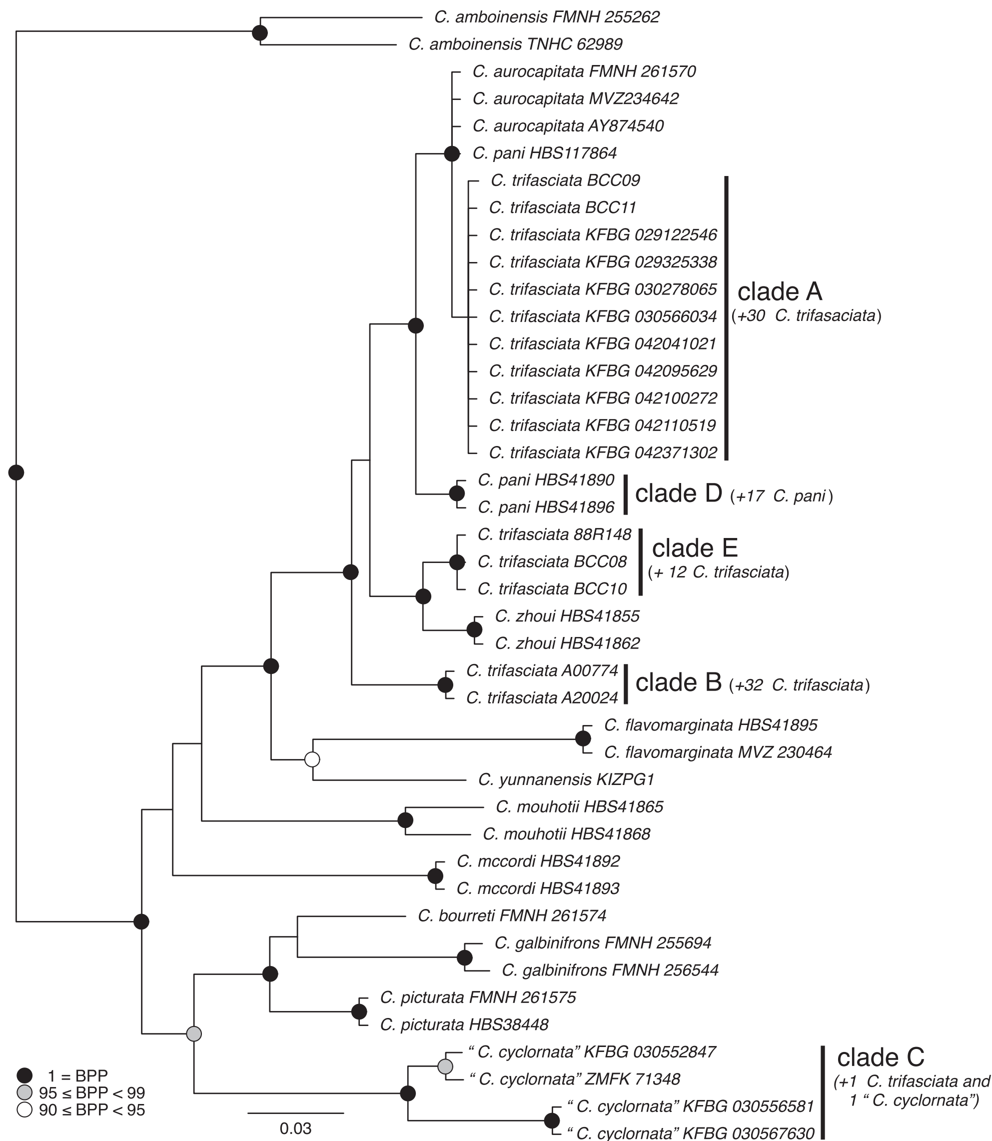

Fig. 2. Majority-rule consensus of the posterior distribution of trees from the Bayesian analysis of the concatenated *ND1* data set (42 taxon, 726 bp), estimated under the GTR + G model of sequence evolution. Many *C. trifasciata* and *C. pani* (76 and 17, respectively) were removed to simplify the presentation, but clade membership of all individuals is shown in the Appendix. Bayesian posterior probabilities (BPP) as indicated. The *M. mutica* outgroup was removed for clarity of presentation.

sequences, repeated the recombination test, and found no significant signal of recombination ($P = 0.33$). Therefore, we included the *RELN* data in subsequent analyses, but with the *C. mccordi* sequences excluded for this marker only. In addition, our analyses revealed extremely low levels of variation for *BDNF*. Because *BDNF* contained very little information we excluded it from further analyses. We analyzed 15 loci individually and based on examination of trace files, five loci and the concatenated data set had failed to reach stationarity due presumably to over-parameterized models (Rannala, 2002). Therefore, we reanalyzed these five data sets with simplified models (Table 1); these analyses achieved convergenceed in all cases.

3.3. nuDNA phylogeny and phylogenetic networks: individual gene trees

The Bayesian test of monophyly revealed strong incongruence between our mtDNA and nuDNA ($P < 0.001$). Therefore, we did not concatenate the mtDNA and nuDNA for any analyses. The major discrepancies between the mtDNA and the nuDNA trees tended to be the relative positions of *C. mccordi* and *C. cyclornata* (see online Supplement for single gene trees and phylogenetic networks). Based on single-gene phylogenetic analyses, *C. trifasciata* was paraphyletic (with strong support) with respect to *C. mccordi* at *AHR*, *BMP2*, *P26s4*, and *TB73* (Figs. S1, S3, S7, and S14, respectively), and *C. mccordi* was contained within the *C. trifasciata* species complex with strong support at *HMGB2*, *HNFL*, *NB22519*, and *PAX* (Figs. S4–S6 and S8, respectively). *Cuora cyclornata* was never recovered as a clade (Supplementary files S1–S15), even with weak support. Furthermore, for three loci at least one allele from a *C. cyclornata* individual clustered more closely with *C. aurocapitata* and *C. pani* than it did with the remaining *C. trifasciata*. For example, at the AHR locus, all *C. trifasciata, cyclornata*, and *C. mccordi* alleles formed a monophyletic group except for a single allele from sample *C. cyclornata* KFBG 030552847, which clustered with *C. aurocapitata* + *C. pani* (Fig. S1). Likewise, at the PAX locus, both alleles from *C. cyclornata* KFBG 030552847 clustered with *C. pani* (Fig. S8). Finally, both alleles from sample *C. cyclornata* BCC09 clustered with *C. aurocapitata* + *C. pani* at the TB73 locus (Fig. S14).

Network analyses of individual loci were qualitatively similar to the gene trees, but they revealed more complex relationships among several key groups that are not apparent in the correspond-

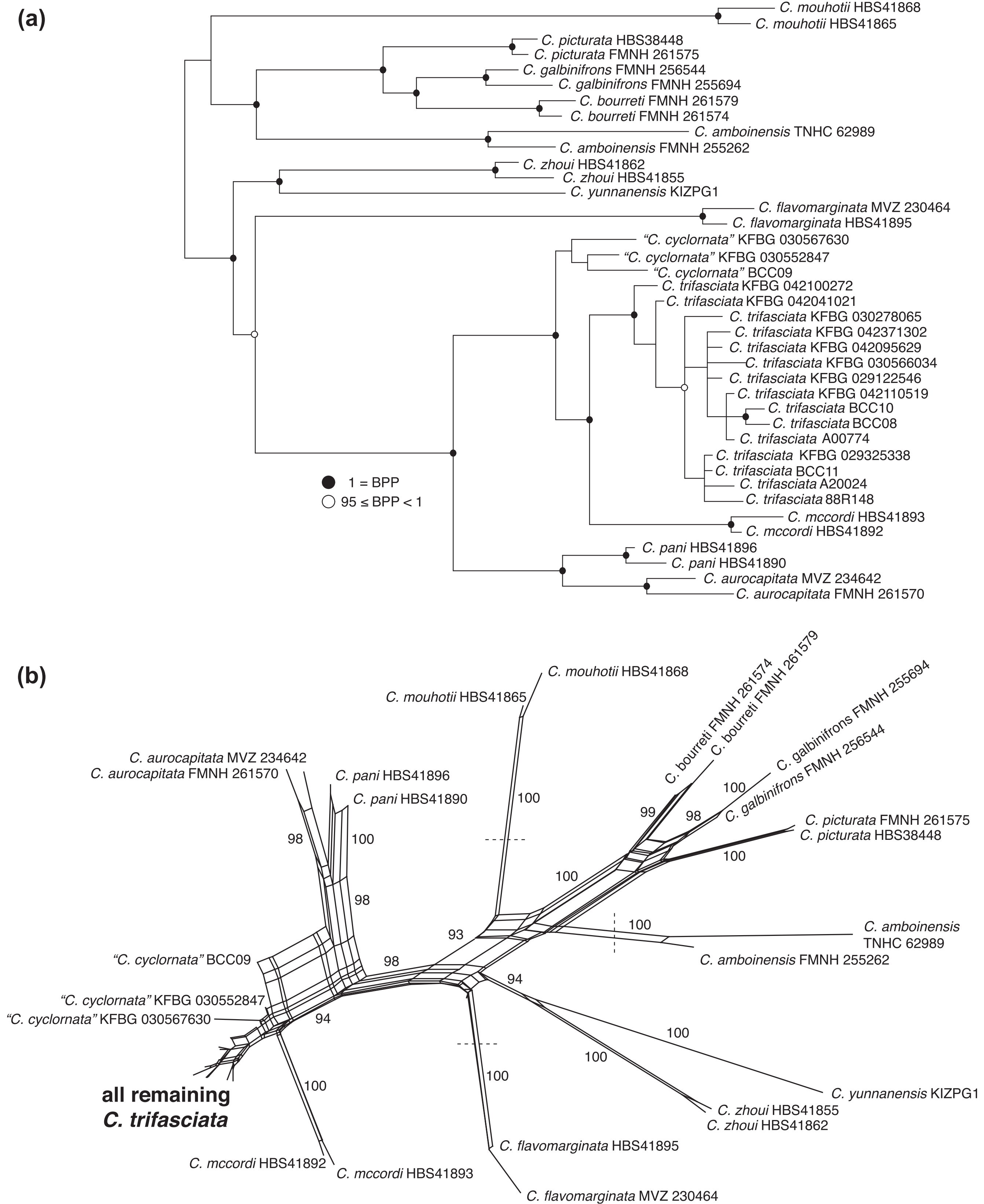

Fig. 3. Analyses of the concatenated nuDNA data set (40 taxon, 10357 bp). (a) Majority-rule consensus of the posterior distribution of trees from the partitioned-model Bayesian analysis. Models of molecular evolution for parameter estimation are shown in Table 1. Bayesian posterior probabilities (BPP) as indicated. The outgroup has been removed for clarity of presentation. (b) Phylogenetic network generated using the neighbor-net algorithm with uncorrected-P distances. Branches have been artificially shortened at dashed lines. Numbers along edges are bootstrap support values (1000 pseudoreplicates). A single branch or edge indicates a lack of conflicting phylogenetic signal while numerous edges are indicative of incompatible splits or conflicting phylogenetic signal. For example, the edges connecting *C. cyclornata*, *C. trifasciata*, and *C. mcordi* to the remaining *Cuora* are well supported (bootstrap support value = 94), but also indicate that there is some conflicting phylogenetic signal for *C cyclornata* BCC09 (which generates many novel connections to the *C. aurocapitata/pani* lineage), and to a lesser degree *C. cyclornata* KFBG 030552847.

ing gene trees. For example, at the AIING locus, the *C. mccordi* alleles show a close, but ambiguously resolved relationship to both *C. flavomarginata* and *C. amboinensis* (Fig. S2). Further, at the HNFL locus (Fig. S5), one allele each from *C. cyclornata* KFBG 030552847 and *C. cyclornata* KFBG 0305567630 plus all alleles from *C. mccordi* fell along the branch connecting a differentiated *C. aurocapitata* + *C. pani* lineage to the remaining *C. trifasciata* alleles. Similarly, at the NB22519 locus, there is clear evidence that *C. pani, C. aurocapitata, C. trifasciata*, and *C. mccordi* form a cluster, but how that cluster splits into species is difficult to resolve. In particular, both alleles from *C. cyclornata* KFBG 030552847, and one allele from *C. cyclornata* BCC09 formed an unresolvable link between the main cluster

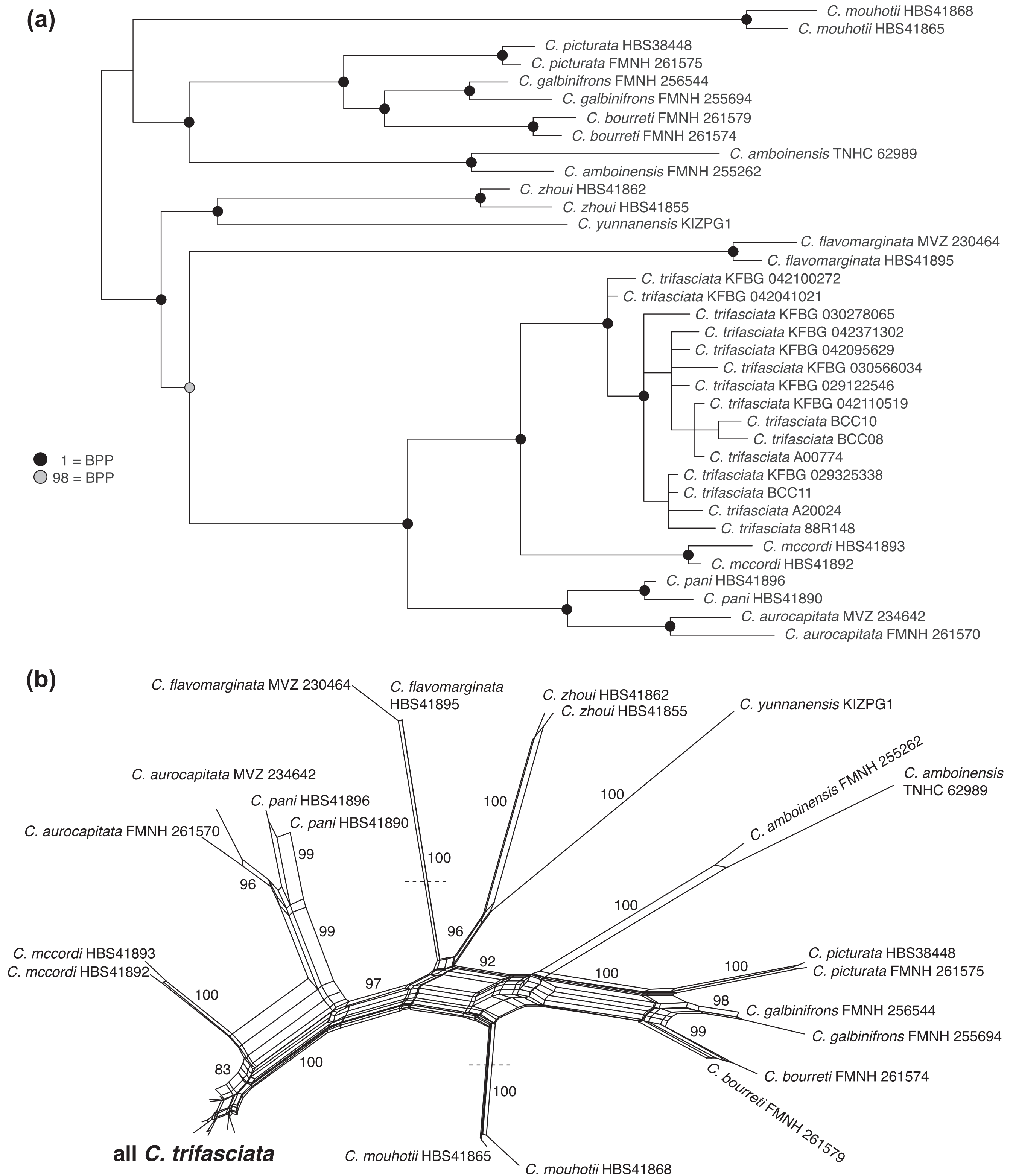

Fig. 4. Same analyses as in Fig. 3, but with *C. cyclornata* (BCC09, KFBG 030567630, and KFBG 030552847) excluded.

of *C. trifasciata*, *C. aurocapitata* and *C. pani*, and/or *C. mccordi*. In addition, one allele from BCC10 clustered with *C. mccordi* to the exclusion of other *C. trifasciata* (Fig. S6).

The *C. galbinifrons* species complex (that is, *C. galbinifrons, C. picturata, C. bourreti*) was recovered as monophyletic in five of 14 phylogenetic analyses of individual loci, *C. bourreti* was monophyletic at 4/13 loci, and *Cuora galbinifrons* and *C. picturata* were monophyletic at 5/15 and 7/14 loci, respectively. In addition, *C. bourreti* was the sister taxon to *C. galbinifrons* at two loci (RELN, TB01), while *C. galbinifrons* was the sister taxon to *C. picturata* at the HNFL locus. The network analyses revealed evidence of gene flow and/or incomplete lineage sorting among members of the *C. galbinifrons* species complex, especially between *C. bourreti* and *C galbinifrons* (Figs. S1–S15).

3.4. nuDNA phylogeny and phylogenetic networks: concatenated data

The phylogeny reconstructed from the partitioned, concatenated data set was well supported, recovering most nodes with strong support and all species as monophyletic except *C. yunnanensis* (for which we could include only one sample). The three *C. cyclornata* individuals were recovered as the monophyletic (with no statistical support) sister clade to *C. trifasciata* + *C. mccordi* (Fig. 3a). The *C. galbinifrons* complex and all three contained species were highly divergent from one another and monophyletic with strong support. Unlike the mtDNA analyses where *C. zhoui* was recovered as the sister taxon to a clade of *C. trifasciata* (clade E, Fig. 2), *C. zhoui* was recovered as the sister taxon to *C. yunnanensis* with strong support based on analyses of the concatenated nuDNA (Fig. 3a).

Results from the phylogenetic network were similar to those from Bayesian analyses of the concatenated nuDNA in that all species were recovered as monophyletic with strong support except *C. cyclornata* (Fig. 3b). The network also revealed evidence for incomplete lineage sorting or reticulation between *C. cyclornata* and *C. aurocapitata* + *C. pani* (Fig. 3b). In particular, the position of *C. cyclornata* BCC09 is consistent with this individual being the result of a relatively recent cross between *C. trifasciata* and either *C. aurocapitata* or *C. pani*. *Cuora cyclornata* KFBG 030552847 also showed some evidence of reticulation with *C. aurocapitata* or *C. pani* (Fig. 3b), although less strongly than did BCC09. Analyses of the concatenated nuDNA with *C. cyclornata* removed rendered *C. trifasciata* monophyletic with strong support in both the phylogeny (Bayesian posterior probability [BPP] = 1) and phylogenetic network (bootstrap support value [BP] = 83). The network analysis without *C. cyclornata* also suggested possible reticulation between *C. mccordi* and *C. aurocapitata/C. pani*, based on the many edges connecting these lineages (Fig. 4b). Relationships among the remaining *Cuora* species, however, were unaffected by the removal of the problematic *C. cyclornata* samples.

Finally, the agreement subtree analysis confirmed the close relationships among members of the *C. galbinifrons* species complex, as well as the close relationships among members of the *C. trifasciata* species complex (excluding *C. cyclornata*), as these groupings were common to both the mtDNA and nuDNA trees. There was equal support for either *C. zhoui* or *C. flavomarginata* as the sister clade to the *C. trifasciata* species complex (Fig. 5).

4. Discussion

Among closely related species, discordance between current taxonomy and phylogenetic patterns is fairly common, and may represent inaccurate taxonomy, incomplete lineage sorting, hybridization, gene duplication and loss, or any combination of these factors (Avise, 1989; Maddison, 1997; Funk and Omland, 2003; Zink and Barrowclough, 2008; Brito and Edwards, 2009; Petit and Excoffier, 2009). Given their established propensity for hybridization, the discordance among taxonomy, mtDNA, and nuDNA phylogenetic trees revealed here for *Cuora* may well be due to a long history of hybridization and introgression among wild turtles and, potentially, more recent hybridization between wild and farmed turtles.

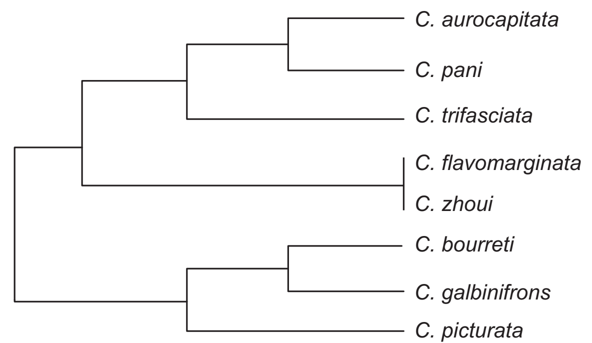

Fig. 5. Consensus of two agreement subtrees of the combined mtDNA tree (Fig. 2) plus concatenated nuDNA tree (Fig. 3).

Correctly identifying species and their phylogenetic relationships for *Cuora* has been a challenging endeavor, but interspecific relationships and species boundaries for several major lineages of *Cuora* are coming into focus. For example, *C. amboinensis*, *C. flavomarginata*, *C. mouhotii*, and *C. yunnanensis* display no evidence of introgression (at least with our sampling), and mitochondrial and nuclear phylogenies are in close agreement on relationships for several species (i.e. Fig. 5). The key point is that phylogenetic relationships and species boundaries among most *Cuora* species now appear to be well-resolved, providing some optimism that the systematics and conservation communities can move forward with a relatively sound, stable understanding of the group. In addition to a mtDNA phylogeny, we now have a complementary, well-supported nuDNA phylogeny with which we can assess evolution in this group and gain further insight into relationships and species composition for *Cuora*.

4.1. Cuora trifasciata species complex

Species boundaries in the *C. trifasciata* species complex have been contentious and difficult to delineate. Previous work has largely been based on limited individual sampling of mtDNA, which has yielded gene trees that are very different from the well supported, multi-gene, nuclear DNA tree presented here. Given the very strong support provided by the 15-gene nuclear tree and its overall concordance with current taxonomy based on morphological data, we view the mtDNA gene tree as an incorrect representation of the species tree in areas where the two disagree. Shared mitotypes among morphologically similar species are most likely due to hybridization. Highly divergent mitotypes recovered from morphologically similar taxa could be indicative of cryptic taxa or could be due to mitochondrial introgression from other, potentially unsampled clades. Under the hypothesis of deeply divergent cryptic taxa, our expectation is that along with divergent mitotypes, the putative cryptic taxa should also exhibit a parallel, elevated level of nuclear divergence. On the other hand, mitochondrial introgression would lead to taxa that contain divergent mitotypes, but exhibit little or no nuclear divergence (unless nuclear alleles are also introgressed from other taxa).

All members of the *C. trifasciata* species complex (*C. aurocapitata*, *C. pani* and *C. trifasciata* including *C. cyclornata*) are morphologically similar and somewhat variable in color pattern and shell

features. This morphological similarity led Blanck and Tang (2005) to regard *C. aurocapitata* and *C. pani* as conspecific. In addition, Blanck et al. (2006) interpreted the variation that they observed to indicate that *C. trifasciata* is composed of multiple species/subspecies, one of which they recognized as *C. cyclornata*. If correct, these hypotheses suggest that we should expect relatively little genetic divergence and no evidence of single or multi-gene monophyly between *C. aurocapitata* and *C. pani*, and significant divergence among consistent subgroups of *C. trifasciata* including individuals that are morphologically assignable to *C. cyclornata*. This pattern has been recovered in some previous analyses of mtDNA (Blanck et al., 2006; Honda et al., 2002; Parham et al., 2001; Spinks and Shaffer, 2007). However, that picture is strikingly different from the nuDNA analyses. Here, *C. aurocapitata* and *C. pani* are reciprocally monophyletic with strong support, and mitochondrial introgression readily explains the combined mtDNA and nuDNA patterns including the presence of *C. aurocapitata* mitotypes in *C. pani* and *C. trifasciata* (see below).

Results are less clear for *C. cyclornata*, because either introgression or incomplete coalescence could account for the observed patterns. For example, the *C. cyclornata* are recovered as the monophyletic sister group to the remaining *C. trifasciata* + *C. mccordi* in the concatenated nuDNA phylogenetic analysis (Fig. 3a), albeit with no statistical support. However, the single-gene analyses reveal that in no case is *C. cyclornata* recovered as monophyletic, with or without strong statistical support. Rather, some alleles from *C. cyclornata* are similar (sometimes identical) to *C. aurocapitata*, *C. pani* or *C. mccordi* alleles (Figs. S1, S8 and S14). This suggests that the "divergence" between *C. cyclornata* and the remaining *C. trifasciata* seen in the concatenated nuDNA analysis (Fig. 3a) could be the result of introgressed *C. aurocapitata*/*C. pani* alleles in *C. cyclornata*. Thus, the divergent phylogenetic position of the *C. cyclornata* in the concatenated nuDNA phylogeny may be an artifact of introgression. The network analysis of the concatenated nuDNA, particularly for *C. cylornata* specimen BCC09, is consistent with this interpretation (Fig. 3b). In addition, the hypothesis that *C. cylornata* represents an independently evolving, cryptic taxon is inconsistent with the combined evidence from the mtDNA and nuDNA trees. Mitochondrially, *C. cyclornata* is the sister group to the *C. galbinifrons* species complex (Fig. 2), a position that has not been suggested based on morphological similarity. However, at the nuclear level *C. cyclornata* is highly divergent from the *C. galbinifrons* species complex but falls within the *C. trifasciata* species complex (Fig. 3a). These patterns are inconsistent with expected ancestor–descendent relationships of cryptic taxa, or lineage sorting, but are consistent with introgression.

The geographic distribution of our known-locality samples of *C. cyclornata* and *C. trifasciata* is also confusing. Blanck et al. (2006) describe the range of *C. cyclornata* as central Vietnam to western Guangxi Province, China, and that of *C. trifasciata* as eastern Guangxi to Fujian Province including Hainan Island (see Fig. 1). However, two of the *C. cyclornata* analyzed here were collected from the wild in Vietnam and Hong Kong (ZFMK 71348 and KFBG 030552847, respectively) (Fig. 1). Hong Kong is directly offshore of the central part of the range of *C. trifasciata* in Guangdong Province, and geography would suggest that Hong Kong specimens should most likely be *C. trifasciata*. The presence of sympatric *C. cyclornata* with *C. trifasciata* in Hong Kong (Spinks and Shaffer, 2007) is difficult to explain if the *C. cyclornata* phenotype is indeed restricted to Vietnam; one possibility is that *C. cyclornata*, if a good species, is present in Hong Kong as a result of releases of captive specimens. Alternatively, the presence of the subtly different "*C. cyclornata*" color pattern initially identified in Vietnam specimens in Hong Kong populations could represent phenotypic plasticity, dietary-induced differentiation, or a genetically-controlled pattern that has evolved more than once.

In summary, both mitochondrial and nuclear data now available indicate that *C. pani* and *C. aurocapitata* are each well defined species separate from *C. trifasciata*. This is particularly clear from the nuDNA data, and we recommend that all three continue to be recognized as good species under a lineage-based species conceptualization (de Queiroz, 1998). There may well be some natural or human-mediated hybridization leading to mitochondrial introgression, but this has not affected the pattern of nuclear gene monophyly recovered in our analyses. The situation is less clear for *C. cyclornata*, although we find the current evidence from both nuclear and mitochondrial genes, combined with the inconsistent distribution of specimens from Hong Kong, most consistent with *C. cylornata* representing phenotypic variation within *C. trifasciata*, or perhaps a novel color pattern that may result from hybridization among other members of the *C. trifasciata* species complex. Pending further data, particularly from field-collected specimens, we recommend that *C. cyclornata*-pattern specimens be considered *C. trifasciata* rather than a distinct species.

4.2. *Cuora zhoui*, *C. mccordi*, and *C. yunnanensis*

Turtles are farmed in large numbers in southern China (Gong et al., 2009; Shi and Parham, 2001; Shi et al., 2008b) and turtle farmers often keep different genera and species in common ponds, enabling interspecific and intergeneric hybridization (Parham and Shi, 2001). These hybrid turtles are sold in markets, and some have been purchased and subsequently described as new species (reviewed in Parham et al. (2001), Shi et al. (2005) and Spinks and Shaffer (2007)). For example, during a 15-year span (1984–1998), a series of 12 new species of Asian turtles were described, only three of which (*C. aurocapitata*, *C. bourreti*, *C. pani*) utilized field-collected material. A fourth species, *C. picturata* was also described from market turtles, but populations were later found in the wild (Ly et al., 2011). The remaining eight species were described from turtles culled from the food markets of southern China; they include three *Cuora* (*C. mccordi*, *C.* "*serrata*", *C. zhoui*), four *Mauremys* (*M.* "*iversoni*", *M.* (*Ocadia*) "*glyphistoma*", *M.* (*Ocadia*) "*philippeni*", *M.* "*pritchardi*)", and *Sacalia* "*pseudocellata*". Six of these 12 new species have been confirmed as recent interspecific hybrids (those in quotation marks above) produced in turtle farms in southern China (Feldman and Parham, 2004; Parham et al., 2001; Shi et al., 2005; Spinks et al., 2004; Spinks and Shaffer, 2007; Stuart and Parham, 2004, 2007; Wink et al., 2001), and are no longer recognized as valid species (Turtle Taxonomy Working Group, 2011). Thus, of the eight new species that were described from market turtles and have not been found in the wild, only *C. zhoui* and *C. mccordi* are still regarded as candidate valid taxa. Our results provide new and important insights into the authenticity of these two taxa.

Our multilocus nuDNA data suggest that *C. zhoui* is a deeply divergent lineage that is the sister taxon to *C. yunnanensis* (Figs. 3 and 4), a species that was described over 100 years ago (Boulenger, 1906) and until recently had been considered extinct. However, analyses of mtDNA data (Fig. 2) agree with previous results that consistently recover *C. zhoui* as closely related to members of the *C. trifasciata* species complex (Honda et al., 2002; Parham et al., 2004; Spinks and Shaffer, 2007). The nuDNA analyses of Spinks and Shaffer (2007) also recovered *C. zhoui* as monophyletic with strong support, but lacked specimens of the newly rediscovered *C. yunnanensis*. Inclusion of *C. yunnanensis* appears to be critical because *C. yunnanensis* is most likely the sister species of *C. zhoui* (Fig. 4). Thus, *C. zhoui* appears to be a relatively deeply divergent lineage and a valid species, but the mitotypes recovered from samples of *C. zhoui* could be introgressed *C. trifasciata* mitotypes (see below).

The status of *Cuora mccordi* as a valid species is less clear. On the one hand, both nuclear and mitochondrial analyses consistently

recover *C. mccordi* as monophyletic, suggesting that it is a good species under a lineage-based species concept. However, the phylogenetic position of *C. mccordi* based on analyses of mtDNA are highly incongruent with those recovered from nuDNA, suggesting that introgression may have played a role in the evolutionary history of this taxon. Also important in our view is the lack of verified field-collected specimens; although Zhou et al. (2008) described the range of *C. mccordi* as Quangxi Province in southern China, to our knowledge no actual *C. mccordi* have ever been collected from the wild.

In conclusion, the combined inferences from nuclear and mitochondrial DNA suggest that both *C. zhoui* and *C. yunnanensis* are valid evolutionary lineages that should continue to be recognized as good species. We are less certain about *C. mccordi*, given the very disparate mitochondrial and nuclear phylogenetic placements of the taxon and the lack of verified material from the wild. However, based on the relatively deep divergence and inferred monophyly of *C. mccordi* for both mt and nuDNA data sets, we recommend its continued recognition as a valid species for the time being.

4.3. Clade A–E mitotypes

Given the prevalence of hybridization and subsequent mitochondrial introgression within *Cuora*, we view the multi-locus nuclear estimates of phylogeny and species boundaries as our most accurate estimates of species trees. However, discordance with the mitochondrial tree can provide important insights into past or ongoing hybridization, and thus it is important to reconcile the various mitochondrial haplotypes with the nuclear tree. *Cuora aurocapitata* and *C. pani* are reciprocally monophyletic based on nuDNA, and clade D mitotypes were recovered from 19 *C. pani*, while clade A mitotypes were found in all available *C. aurocapitata*, plus numerous *C. trifasciata*, and one *C. pani* (Fig. 2). These data strongly suggest that clade D mitotypes represent ancestral *C. pani* mitotypes. It is less certain, but likely that clade A represents *C. aurocapitata* mitotypes that are now also present in some *C. trifasciata* and *C. pani* via introgressive hybridization.

In a previous analysis, Spinks et al. (2009) hypothesized that clade C mitotypes represent genuine *C. trifasciata*. The present analysis rejects this hypothesis because clade C mitotypes are highly divergent from those of *C. aurocapitata* and *C. pani*, but nuDNA data strongly support *C. trifasciata* as the sister group to (*C. aurocapitati* + *C. pani*). Thus, authentic *C. trifasciata* mitotypes could include clade B, clade E or mitotypes recovered from *C. zhoui*, or perhaps all of these (Fig. 2). We suggest that clade C and the *C. mccordi* mitotypes might be the result of mitochondrial capture following hybridization between *C. trifasciata* and the *C. galbinifrons* species complex.

Ultimately, the complex and confusing arrangement of mitochondrial clades most strongly supports the interpretation that hybridization and introgressive movement of mitochondrial DNA has been common across *Cuora*, and particularly so for members of the *C. trifasciata* complex. Additional, genome-level nuclear data sets may help to resolve this history of introgressive hybridization, or that history may be too complex for the details to be recoverable. In any case, our strong conclusion is that species and lineage identification and DNA barcoding-like efforts should rely on multigene nuclear data, rather than mitochondrial DNA data for these turtles.

4.4. Cuora galbinifrons species complex

Cuora bourreti, C. galbinifrons, and *C. picturata* are now considered distinct species by many authors (but see Fritz et al. (2006) for an alternative view), and our analysis recovers strong support for this interpretation. Thus, our position is that the weight of molecular evidence is consistent with the view that *C. bourreti, C. galbinifrons*, and *C. picturata* form a clade of three genetically and morphologically diagnosable taxa that should continue to be recognized as valid species. Within the *C. galbinifrons* complex, our concatentated nuclear data indicate that *C. bourreti* and *C. galbinifrons* are sister taxa with strong statistical support.

4.5. Concluding thoughts

Overall, the phylogeny and species composition of *Cuora* are coming into focus. We now have a well-supported phylogenetic hypothesis for *Cuora* generated from multiple nuclear loci (Figs. 3 and 4), and equally compelling evidence that introgression makes the mitochondrial gene tree an incorrect estimate of species relationships within the group. This nuclear phylogeny contains ample support for recognition of most currently recognized members of the *C. galbinifrons* and *C. trifasciata* species complexes as defensible species. In addition, extensive mito-nuclear discordance indicates that introgression and potentially mitochondrial capture are historical processes in this clade. Future analyses of this group should explicitly recognize these phenomena, make use of available nuclear markers and analytical approaches, and include more individuals with known locality information to reveal the complex phylogenetic and biogeographic history of *Cuora*.

Acknowledgments

For tissue samples, we thank Gary Ades and Paul Crow (Kadoorie Farms and Botanical Gardens), Sue Angus and Chris Banks (Melbourne Zoo), H. Artner, James Barzyk, Wolgang Böhme (Zoologisches Forschungsmuseum Alexander Koenig), Rafe Brown (University of Kansas Natural History Museum), Carla Cicero (Museum of Vertebrate Zoology), Dave Collins and Bill Hughes (Tennessee Aquarium), Penny Danielewicz (Buffalo Zoo), Paul Gibbons (Behler Conservation Center), Rick Haeffner (Denver Zoo), Rick Hudson (Fort Worth Zoo), Elmar Meier, Alan Resetar (Field Museum of Natural History), Dennis Uhrig, and Thomas Ziegler (AG Zoologischer Garten Köln). We thank Damian Goodall (Melbourne Zoo) for the *C. trifasciata* photograph. Allan Larson and two anonymous reviewers provided valuable comments. This material is based upon work supported by the National Science Foundation under Grants No. DEB-0213155, DEB-0516475, DEB-0507916 and DEB 0817042, and the UC Davis Agricultural Experiment Station.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2012.02.014.

References

Allendorf, F.W., Leary, R.F., Spruell, P., Wenburg, J.K., 2001. The problems with hybrids: setting conservation guidelines. Trends Ecol. Evol. 16, 613–622.

Avise, J.C., 1989. Gene trees and organismal histories: a phylogenetic approach to population biology. Evolution 43, 1192–1208.

Avise, J.C., 1994. Molecular Markers, Natural History and Evolution. Chapman and Hall, New York.

Backström, N., Fagerberg, S., Ellegren, H., 2008. Genomics of natural bird populations: a gene-based set of reference markers evenly spread across the avian genome. Mol. Ecol. 17, 964–980.

Barilani, M., Deregnaucourt, S., Gallego, S., Galli, L., Mucci, N., Piombo, R., Puigcerver, M., Rimondi, S., Rodríguez-Teijeiro, J.D., Spanò, S., Randi, E., 2005. Detecting hybridization in wild (*Coturnix c. coturnix*) and domesticated (*Coturnix c. japonica*) quail populations. Biol. Conserv. 126, 445–455.

Barley, A., Spinks, P.Q., Thomson, R.C., Shaffer, H.B., 2010. Fourteen nuclear genes provide phylogenetic resolution for difficult nodes in the turtle tree of life. Mol. Phylogenet. Evol. 55, 1189–1194.

Bhangale, T.R., Rieder, M.J., Livingston, R.J., Nickerson, D.A., 2005. Comprehensive identification and characterization of diallelic insertion–deletion polymorphisms in 330 human candidate genes. Hum. Mol. Genet. 14, 59–69.

Blanck, T., Tang, M., 2005. Ein neuer Fundort von Cuora pani Song, 1984 mit Diskussion **uber den taxonomischen Status von *Cuora pani* und *Cuora aurocapitata*. Sacalia 20, 16–37.

Blanck, T., McCord, W.P., Le, M., 2006. On the variability of *Cuora trifasciata* (Bell, 1825). Edition Chimaira, Frankfurt am Main.

Bossu, C.M., Near, T.J., 2009. Gene trees reveal repeated instances of mitochondrial DNA introgression in orangethroat darters (Percidae: *Etheostoma*). Syst. Biol. 58, 114–129.

Boulenger, G.A., 1906. Descriptions of new reptiles from Yunnan. Ann. Mag. Natur. Hist. Lond. 17, 567–568.

Brito, P.H., Edwards, S.V., 2009. Multilocus phylogeography and phylogenetics using sequence-based markers. Genetica 135, 439–455.

Bruen, T.C., Philippe, H., Bryant, R., 2006. A simple and robust statistical test for detecting the presence of recombination. Genetics 172, 2665–2681.

Bryant, D., Moulton, V., 2004. NeighborNet: an agglomerative algorithm for the construction of planar phylogenetic networks. Mol. Biol. Evol. 21, 255–265.

Buckley, T.R., Cordeiro, M., Marshall, D.C., Simon, C., 2006. Differentiating between hypotheses of lineage sorting and introgression in New Zealand alpine cicadas (Maoricicada Dugdale). Syst. Biol. 55, 411–425.

Burrell, A.S., Jolly, C.J., Tosi, A.J., Disotell, T.R., 2009. Mitochondrial evidence for the hybrid origin of the kipunji, *Rungwecebus kipunji* (Primates: Papionini). Mol. Phylogenet. Evol. 51, 340–348.

de Queiroz, K., 1998. The general lineage concept of species, species criteria, and the process of speciation: a conceptual uni cation and terminological recommendations. In: Howard, D.J., Berlocher, S.H. (Eds.), Endless Forms. Species and Speciation, Oxford University Press, Oxford, pp. 57–75.

Dmitriev, D.A., Rakitov, R.A., 2008. Decoding of superimposed traces produced by direct sequencing of heterozygous indels. PLoS Comput. Biol. 4, e1000113. doi:10.1371/journal.pcbi.1000113.

Drummond, A.J., Ashton, B., Buxton, S., Cheung, M., Cooper, A., Duran, C., Field, M., Heled, J., Kearse, M., Markowitz, S., Moir, R., Stones-Havas, S., Sturrock, S., Thierer, T., Wilson, A., 2011. Geneious v5.4. <http://www.geneious.com/>.

Feldman, C.R., Parham, J.F., 2004. Molecular systematics of Old World stripe-necked turtles (Testudines: *Mauremys*). Asiat. Herpetol. Res. 10, 28–37.

Fritz, U., Petzold, A., Auer, M., 2006. Osteology in the *Cuora galbinifrons* complex suggests conspecifity of *C. bourreti* and *C. galbinifrons*, with notes on shell osteology and phalangeal formulae within the Geoemydidae. Amphibia-Reptilia 27, 195–205.

Fujita, M.K., Engstrom, T.N., Starkey, D.E., Shaffer, H.B., 2004. Turtle phylogeny: insights from a novel nuclear intron. Mol. Phylogenet. Evol. 31, 1031–1040.

Funk, D.J., Omland, K.E., 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Ann. Rev. Ecol. Evol. Syst. 34, 397–423.

Gong, S., Chow, A.T., Fong, J.J., Shi, H., 2009. The chelonian trade in the largest pet market in China: scale, scope and impact on turtle conservation. Oryx 43, 213–216.

Good, J.M., Hird, S., Reid, N., Demboski, J., Steppan, S.J., Martin-Nims, T.R., Sullivan, J., 2008. Ancient hybridization and mitochondrial capture between two species of chipmunks. Mol. Ecol. 17, 1313–1327.

Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704.

He, J., Zhou, T., Rao, D., Zhang YaPing, 2007. Molecular identification and phylogenetic position of *Cuora yunnanensis*. Chinese Sci. Bull. 52, 3305–3309.

Holder, M.T., Anderson, J.A., Holloway, A.K., 2001. Difficulties in detecting hybridization. Syst. Biol. 50, 978–982.

Honda, M., Yasukawa, Y., Hirayama, R., Ota, H., 2002. Phylogenetic relationships of the Asian box turtles of the genus *Cuora* sensu lato (Reptilia: Bataguridae) inferred from mitochondrial DNA sequences. Zool. Sci. 17, 1305–1312.

Huson, D.H., 1998. SplitsTree: a program for analyzing and visualizing evolutionary data. Bioinformatics 14, 68–73.

Huson, D.H., Bryant, D., 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23, 254–267. Software available from <http://www.splitstree.org>.

IUCN, 2011. IUCN Red List of Threatened Species. Version 2011.1. <http://www.iucnredlist.org> (downloaded 07.11.11).

Joly, S., McLenachan, P.A., Lockhart, P.J., 2009. A statistical approach for distinguishing hybridization and incomplete lineage sorting. Am. Nat. 174, E54–E70.

Keck, B.P., Near, T.J., 2009. Patterns of natural hybridization in darters (Percidae: Etheostomatinae). Copeia 4, 758–773.

Kidd, A.G.J., Bowman, J., Lesbarrères, D., Hostedde, S.A.I., 2009. Hybridization between escaped domestic and wild American mink (*Neovison vison*). Mol. Ecol. 18, 1175–1186.

Kimball, R.T., Braun, E.L., Barker, F.K., Bowie, R.C.K., Braun, M.J., Chojnowski, J.L., Hackett, S.J., Han, K.L., Harshman, J., Heimer-Torres, V., Holznagel, W., Huddleston, C.J., Marks, B.D., Miglia, K.J., Moore, W.S., Reddy, S., Sheldon, F.H., Smith, J.V., Witt, C.C., Yuri, T., 2009. A well-tested set of primers to amplify regions spread across the avian genome. Mol. Phylogenet. Evol. 50, 654–660.

Krenz, J.G., Naylor, G.J.P., Shaffer, H.B., Janzen, F.J., 2005. Molecular phylogenetics and evolution of turtles. Mol. Phylogenet. Evol. 37, 178–191.

Lehn, C., Das, I., Forstner, M.R.J., and Brown, R.M. 2007. Responsible vouchering in turtle research: an introduction and recommendations. In: Shaffer, H.B., FitzSimmons, N.N., Georges, A., and Rhodin, A.G.J. (Eds.), Defining Turtle Diversity: Proceedings of a Workshop on Genetics, Ethics and Taxonomy of Fresh Water Turtles and Tortoises. Chelonian Research Monographs No. 4, pp. 147–156.

Linnen, C., Farrell, B., 2007. Mitonuclear discordance is caused by rampant mitochondrial introgression in *Neodiprion* (Hymenoptera: Diprionidae) sawflies. Evolution 61, 1417–1438.

Ly, T., Hoang, H.D., Stuart, B.L., 2011. Market turtle mystery solved in Vietnam. Biol. Conserv. 144, 1767–1771.

Maddison, W., 1997. Gene trees in species trees. Syst. Biol. 46, 523–536.

Mallet, J., 2005. Hybridization as an invasion of the genome. Trends Ecol. Evol. 20, 229–237.

Melo-Ferreira, J., Boursot, P., Suchentrunk, F., Ferrand, N., Alves, P.C., 2005. Invasion from the cold past: extensive introgression of mountain hare (*Lepus timidus*) mitochondrial DNA into three other hare species in Northern Iberia. Mol. Ecol. 14, 2459–2464.

Moore, W.S., 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. Evolution 49, 718–726.

Noonan, B.P., Chippindale, P.T., 2006. Vicariant origin of Malagasy reptiles supports late cretaceous Antarctic land bridge. Am. Nat. 168, 730–741.

Norén, K., Dalén, L., Kirsti, K., Anders, A., 2006. Detection of farm fox and hybrid genotypes among wild arctic foxes in Scandinavia. Conserv. Genet. 6, 885–894.

Parham, J.F., Shi, H., 2001. The Discovery of *Mauremys iversoni*-like Turtles at a Turtle Farm in Hainan Province, China: the Counterfeit Golden Coin. Asiatic Herpetol. Res. 9, 71–76.

Parham, J.F., Simison, W.B., Kozak, K.H., Feldman, C.R., Shi, H., 2001. New Chinese turtles: endangered or invalid? A reassessment of two species using mitochondrial DNA, allozyme electrophoresis and known-locality specimens. Anim. Conserv. 4, 357–367.

Parham, J.F., Stuart, B.L., Bour, R., Fritz, U., 2004. Evolutionary distinctiveness of the extinct Yunnan box turtle (*Cuora yunnanensis*) revealed by DNA from an old museum specimen. Proc. Roy. Soc. Lond. B 271, S391–S394.

Peters, J.L., Zhuravlev, Y., Fefelov, I., Logie, A., Omland, K.E., 2007. Nuclear loci and coalescent methods support ancient hybridization as cause of mitochondrial paraphyly between gadwall and falcated duck (*Anas* spp.). Evolution 61, 1992–2006.

Petit, R.J., Excoffier, L., 2009. Gene flow and species delimitation. Trends Ecol. Evol. 24, 386–393.

Posada, D., 2008. JModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25, 1253–1256.

Primmer, C.R., Borge, T., Lindell, J., Saetre, G.P., 2002. Single-nucleotide polymorphism characterization in species with limited available sequence information: high nucleotide diversity revealed in the avian genome. Mol. Ecol. 11, 603–611.

Rambaut, A., Drummond, A.J., 2009. Tracer v1.5. <http://beast.bio.ed.ac.uk/Tracer>.

Rannala, B., 2002. Identifiability of parameters in MCMC Bayesian inference of phylogeny. Syst. Biol. 51, 754–760.

Rhymer, J.M., Simberloff, D., 1996. Extinction by hybridization and introgression. Annu. Rev. Ecol. Syst. 27, 83–109.

Riley, S.P.D., Shaffer, H.B., Voss, S.R., Fitzpatrick, B.M., 2003. Hybridization between a rare, native tiger salamander (*Ambystoma californiense*) and its introduced congener. Ecol. Appl. 13, 1263–1275.

Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Sambrook, J., Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual, third ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Sang, T., Zhong, Y., 2000. Testing hybridization hypotheses based on incongruent gene trees. Syst. Biol. 49, 422–424.

Sehgal, R.N.M., Lovette, I.J., 2003. Molecular evolution of three avian neurotrophin genes: implications for proregion functional constraints. J. Mol. Evol. 57, 335–342.

Shi, H., Parham, J.F., 2001. Preliminary observations of a large turtle farm in Hainan Province, People's Republic of China. Turtle Tortoise Newsl. 3, 4–6.

Shi, H., Parham, J.F., Simison, W.B., Wang, J., Gong, S., Fu, B., 2005. A report on the hybridization between two species of threatened Asian box turtles (Testudines: *Cuora*) in the wild on Hainan Island (China) with comments on the origin of '*serrata*'-like turtles. Amphibia-Reptilia 26, 377–381.

Shi, H., Fong, J.J., Parham, J.F., Pang, J., Wang, J., Hong, M., Zang, Y.P., 2008a. Mitochondrial variation of the "eyed" turtles (*Sacalia*) based on known-locality and trade specimens. Mol. Phylogenet. Evol. 49, 1025–1029.

Shi, H., Parham, J.F., Zhiyong, F., Meiling, H., Feng, Y., 2008b. Evidence for the massive scale of turtle farming in China. Oryx 42 (1), 147–150.

Spinks, P.Q., Shaffer, H.B., 2007. Conservation phylogenetics of the Asian box turtles (Geoemydidae, *Cuora*): mitochondrial introgression, numts, and inferences from multiple nuclear loci. Conserv. Genet. 8, 641–657.

Spinks, P.Q., Shaffer, H.B., Iverson, J.B., McCord, W.P., 2004. Phylogenetic hypotheses for the turtle family Geoemydidae. Mol. Phylogenet. Evol. 32, 164–182.

Spinks, P.Q., Thomson, R.C., Shaffer, H.B., 2009. A reassessment of *Cuora cyclornata* Blanck, McCord and Le, 2006 (Testudines, Geoemydidae) and a plea for taxonomic stability. Zootaxa 2018, 58–68.

Stephens, M., Donnelly, P., 2003. A comparison of Bayesian methods for haplotype reconstruction. Am. J. Hum. Genet. 73, 1162–1169.

Stephens, M., Smith, N.J., Donnelly, P., 2001. A new statistical method for haplotype reconstruction from population data. Am. J. Hum. Genet. 68, 978–989.

Stuart, B.L., Parham, J.F., 2004. Molecular phylogeny of the critically endangered Indochinese box turtle (*Cuora galbinifrons*). Mol. Phylogenet. Evol. 31, 164–177.

Stuart, B.L., Parham, J.F., 2007. Recent hybrid origin of three rare Chinese turtles. Conserv. Genet. 8, 169–175.

Thomson, R.C., Shedlock, A.M., Edwards, S.V., Shaffer, H.B., 2008. Developing markers for multilocus phylogenetics in non-model organisms: a test case with turtles. Mol. Phylogenet. Evol. 49, 514–525.

Townsend, T.M., Alegre, R.E., Kelley, S.T., Weins, J.J., Reeder, T.W., 2008. Rapid development of multiple nuclear loci for phylogenetic analysis using genomic resources: an example from squamate reptiles. Mol. Phylogenet. Evol. 47, 129–142.

Turtle Conservation Fund, 2002. A Global Action Plan for Conservation of Tortoises and Freshwater Turtles. Strategy and Funding Prospectus 2002–2007. Conservation International and Chelonian Research Foundation, Washington, DC.

Turtle Taxonomy Working Group (van Dijk, P.P., Iverson, J.B., Shaffer, H.B., Bour, R., Rhodin, A.G.J.), 2011. Turtles of the world, 2011 update: annotated checklist of taxonomy, synonymy, distribution, and conservation status. In: Rhodin, A.G.J., Pritchard, P.C.H., van Dijk, P.P., Saumure, R.A., Buhlmann, K.A., Iverson, J.B., Mittermeier, R.A. (Eds.), Conservation Biology of Freshwater Turtles and Tortoises: A Compilation Project of the IUCN/SSC Tortoise and Freshwater Turtle Specialist Group. Chelonian Research Monographs No. 5, pp. 000.165–000.242.

van Dijk, P.P., 2000. The status of turtles in Asia. In: van Dijk, P.P., Stuart, S., Rhodin, A. (Eds.), Asian turtle Trade Proceedings of a Workshop on Conservation and Trade of Freshwater Turtles and Tortoises in Asia. Phnom Penh, Cambodia, December 1999. Chelonian Research Monographs 2. Chelonian Research Foundation. Lunenberg, pp. 39–44.

Wink, M., Guicking, D., Fritz, U., 2001. Molecular evidence for a hybrid origin of *Mauremys iversoni* Pritchard & McCord, 1991, and *Mauremys pritchardi* McCord, 1997 (Reptilia: Testudines: Bataguridae). Zoologische Abhandlungen Staatliches Museum für Tierkunde Dresden 51, 41–49.

Zhou, T., Zhao, E., 2004. On the occurrence of living *Cuora yunnanensis* since fifty-eight years, and its description. Sichuan J. Zool. 23, 325–327.

Zhou, T., Blanck, T., McCord, W.P., Li, P., 2008. Tracking *Cuora mccordi* ernst, 1988: The first record of its natural habitat; a re-description; with data on captive populations and its vulnerability. Hamadryad 32, 57–69.

Ziegler, T., Herrmann, H.W., 2000. Preliminary list of the herpetofauna of the Phong Nha – Ke Bang area in Quang Binh province, Vietnam. Biogeographica 76, 49–62.

Zink, R.M., Barrowclough, G.F., 2008. Mitochondrial DNA under siege in avian phylogeography. Mol. Ecol. 17, 2107–2121.