Syst. Biol. 55(6):912–927, 2006

ISSN: 1063-5157 print / 1076-836X online
DOI: 10.1080/10635150601058014

Extensive Morphological Convergence and Rapid Radiation in the Evolutionary History of the Family Geoemydidae (Old World Pond Turtles) Revealed by SINE Insertion Analysis

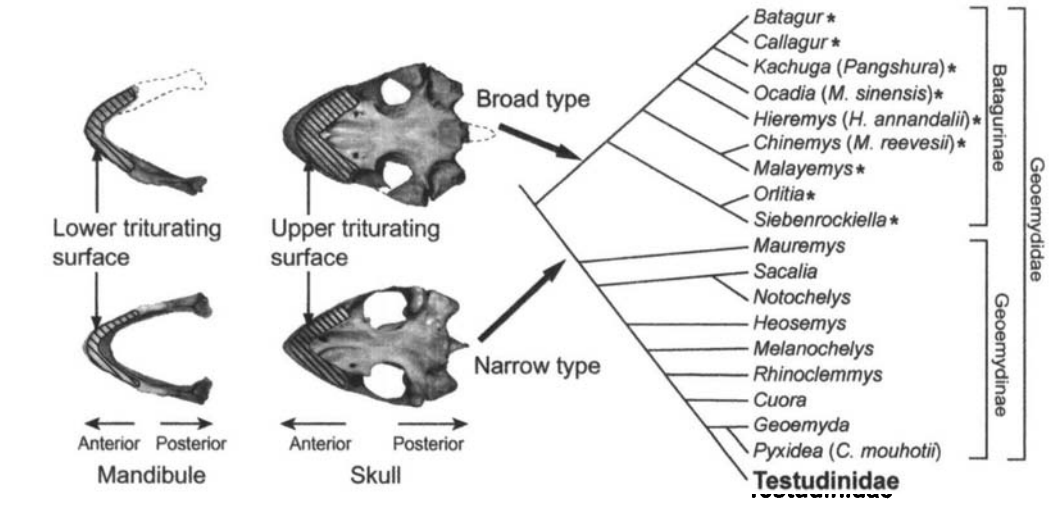
TAKESHI SASAKI,[1] YUICHIROU YASUKAWA,[2] KAZUHIKO TAKAHASHI,[1] SEIKO MIURA,[1] ANDREW M. SHEDLOCK,[3] AND NORIHIRO OKADA[1,4]

[1]*Department of Evolutionary Biology and Biodiversity, National Institute for Basic Biology, Myodaiji, Okazaki, Japan*
[2]*Tropical Biosphere Research Center, University of the Ryukyus, Nishihara, Okinawa, Japan*
[3]*Department of Organismic and Evolutionary Biology and Museum of Comparative Zoology, Harvard University, 26 Oxford Street, Cambridge, Massachusetts 02138, USA*
[4]*Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Yokohama, Japan; E-mail: nokada@bio.titech.ac.jp*

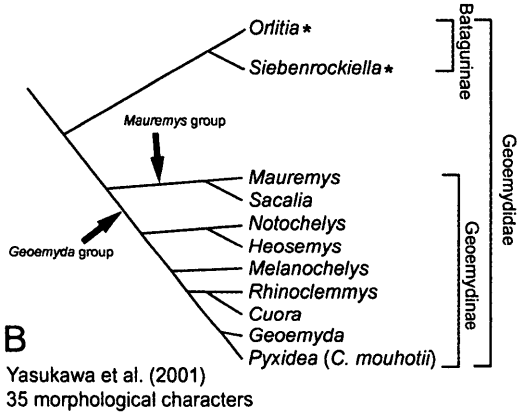
Abstract.— The family Geoemydidae is one of three in the superfamily Testudinoidea and is the most diversified family of extant turtle species. The phylogenetic relationships in this family and among related families have been vigorously investigated from both morphological and molecular viewpoints. The evolutionary history of Geoemydidae, however, remains controversial. Therefore, to elucidate the phylogenetic relationships of Geoemydidae and related species, we applied the SINE insertion method to investigate 49 informative SINE loci in 28 species. We detected four major evolutionary lineages (Testudinidae, *Batagur* group, *Siebenrockiella* group, and *Geoemyda* group) in the clade Testuguria (a clade of Geoemydidae + Testudinidae). All five specimens of Testudinidae form a monophyletic clade. The *Batagur* group comprises five batagurines. The *Siebenrockiella* group has one species, *Siebenrockiella crassicollis*. The *Geoemyda* group comprises 15 geoemydines (including three former batagurines, *Mauremys reevesii*, *Mauremys sinensis*, and *Heosemys annandalii*). Among these four groups, the SINE insertion patterns were inconsistent at four loci, suggesting that an ancestral species of Testuguria radiated and rapidly diverged into the four lineages during the initial stage of its evolution. Furthermore, within the *Geoemyda* group we identified three evolutionary lineages, namely *Mauremys*, *Cuora*, and *Heosemys*. The *Heosemys* lineage comprises *Heosemys*, *Sacalia*, *Notochelys*, and *Melanochelys* species, and its monophyly is a novel assemblage in Geoemydidae. Our SINE phylogenetic tree demonstrates extensive convergent morphological evolution between the *Batagur* group and the three species of the *Geoemyda* group, *M. reevesii*, *M. sinensis*, and *H. annandalii*. [Convergent evolution; Geoemydidae; incomplete lineage sorting; phylogeny; retroposon; SINE; turtle.]

The family Geoemydidae (the name Bataguridae is a synonym of Geoemydidae [McCord et al., 2000; Spinks et al., 2004]) comprises approximately 60 extant species of 20 genera, and it is the largest family of turtles (Ernst et al., 2000; Yasukawa et al., 2001; Spinks et al., 2004; Stuart and Parham, 2004; Diesmos et al., 2005). Geoemydidae species are broadly distributed in southern Europe, northern Africa, tropical, temperate or arid Asia, Central America, and northern South America. Hirayama (1984) extensively investigated the phylogenetic relationships within Geoemydidae at the species level, based mainly on morphological characters (Fig. 1A). He suggested that Geoemydidae species be divided into two major groups based on the type of secondary palate development, namely a broad or narrow triturating surface. The triturating surfaces are dorsal surfaces of dentaries on the lower jaw, and ventral surfaces of premaxillae and maxillae on the upper one, both of which are covered with a rhamphotheca (Fig. 1A). The narrow type is considered the primitive state in the Geoemydidae lineage, based on a comparison with the other Testudinoidea (Hirayama, 1984). Gaffney and Meylan (1988) elevated each of the broad and narrow triturating surface groups to the subfamily level (Batagurinae and Geoemydinae, respectively) based on the results of cladistic analysis by Hirayama (1984). In contrast, Claude et al. (2004) suggested that differences in diet behaviors (e.g., herbivorous, omnivorous, carnivorous, and conchifrageous) play an important role in shaping of the turtle skulls, especially that of the triturating surfaces.

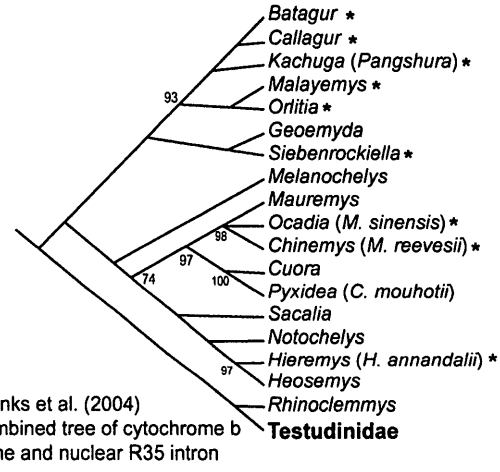
Yasukawa et al. (2001) suggested that two morphological groups, namely the *Mauremys* group and the *Geoemyda* group, are in the subfamily Geoemydinae clade (Fig. 1B). As a consensus from morphological studies, Batagurinae species are grouped together and can be characterized by relatively abundant derived characters (synapomorphies) in Geoemydidae. However, the Geoemydinae lacks apparent synapomorphies as most of character traits of this subfamily seem primitive (Yasukawa et al., 2001). On the other hand, molecular phylogenetic studies have suggested that some Batagurinae species are rather closely related to Geoemydinae species (Fig. 1C and D; Wu et al., 1998; McCord et al., 2000; Honda et al., 2002a, 2002b; Barth et al., 2004; Sasaki et al., 2004; Spinks et al., 2004; Diesmos et al., 2005). For example, the two batagurine genera *Ocadia* and *Chinemys* are included in a clade composed of *Mauremys* species of Geoemydinae, and the genus *Hieremys* (Batagurinae) forms a clade with the genus *Heosemys* of Geoemydinae (Fig. 1C). Accordingly, the genera *Ocadia* and *Chinemys*, and the genus *Hieremys*, were recently synonymized and reclassified into *Mauremys* and *Heosemys*, respectively (Spinks et al., 2004; Diesmos et al., 2005). These molecular studies suggest that convergent morphological evolution occurred in *Ocadia*, *Chinemys*, and *Hieremys* species. If the genus *Siebenrockiella*, in addition to these genera, is excluded from the assemblage of batagurines, monophyly of the remaining batagurines is supported by molecular phylogenetic studies (Fig. 1C: Spinks et al., 2004; Fig. 1D: Diesmos et al., 2005). Despite such extensive



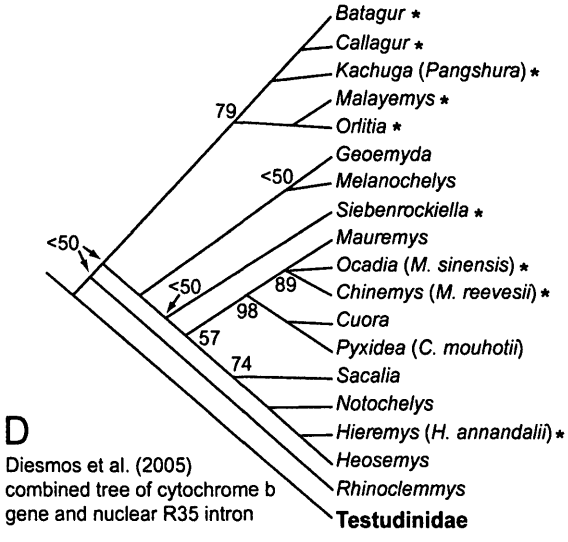
A Hirayama (1984) [Gaffney and Meylan (1988)] 82 morphological and four chromosmal characters



B
Yasukawa et al. (2001)
35 morphological characters



C
Spinks et al. (2004)
combined tree of cytochrome b gene and nuclear R35 intron



D
Diesmos et al. (2005)
combined tree of cytochrome b gene and nuclear R35 intron

FIGURE 1. Previous phylogenetic hypotheses of the family Geoemydidae and related species. A: Morphological cladogram by Hirayama (1984) (this tree was also supported by Geffney and Meylan [1988]). The photographs of mandibles and skulls are dorsal and ventral views of turtle bones, respectively. The bones are from *C. borneoensis* of Batagurinae (upper) and *C. mouhotii* of Geoemydinae (lower). Hatched lines indicate the triturating surfaces of each broad and narrow bone that support the split of the Geoemydidae into two subfamilies. B: Cladogram by Yasukawa et al. (2001). C: Molecular cladogram by Spinks et al. (2004). D: Molecular cladogram by Diesmos et al. (2005). Asterisks denote the batagurine genera. The taxonomic names used in this study are shown in parentheses. All of these previously published trees were modified to include the species used in the present study. The numbers at the nodes in each tree in C and D are bootstrap proportions that were calculated in each study.

molecular data, there is no morphological evidence supporting a polyphyletic Batagurinae. Moreover, for species in *Cuora, Mauremys, Ocadia,* and *Sacalia,* the generation of hybrid species has complicated classification in this family (Stuart and Parham, 2004; Spinks et al., 2004). Thus, Geoemydidae continues to undergo taxonomic changes at the species and/or genus level.

Geoemydidae belongs to the superfamily Testudinoidea, which is one of four superfamilies in the suborder Cryptodira (hidden-necked turtles) (Mlynarski, 1969; Gaffney and Meylan, 1988; Ernst et al., 2000). Formerly, Testudinoidea contained the two families Emydidae (semiaquatic or terrestrial turtles) and Testudinidae (terrestrial tortoise). The Emydidae was composed of the two subfamilies Emydinae and Batagurinae (Batagurinae = the present Geoemydidae), but evidence from subsequent morphological studies suggested that Batagurinae is more closely related to Testudinidae than to the subfamily Emydinae (McDowell, 1964; Hirayama, 1984). Subsequently, the Batagurinae were elevated to the family level, as were the Geoemydidae, Testudinidae, and Emydidae (Gaffney and Meylan, 1988). Joyce et al. (2004) assigned a new taxon, "Testuguria," to the clade of Geoemydidae and Testudinidae. Morphological studies have not resolved the exact phylogenetic relationships between these two families. Hirayama (1984) suggested that Geoemydinae is a paraphyletic group with respect to Testudinidae (Fig. 1A), whereas Yasukawa et al. (2001) argued for the monophyly of Geoemydidae based on two morphological synapomorphies: well-developed axillary and inguinal foramina of musk ducts and an anterolaterally flared iliac blade. Molecular studies have provided several different hypotheses on this issue. Spinks et al. (2004) suggested a paraphyly of Geoemydidae with respect to *Rhinoclemmys* species based on the mitochondrial cytochrome b gene and nuclear R35 intron (Fig. 1C). Honda et al. (2002a) suggested that Geoemydidae species constitute a sister taxon to Testudinidae species based on mitochondrial rRNA genes. Diesmos et al. (2005) also proposed sister relationships between Geoemydidae and Testudinidae species in Testudinoidea, although their dataset was nearly the same as that of Spinks et al. (2004) (Fig. 1D). Basal divergences in these alterntive trees are characterized by low bootstrap values.

SINEs

SINEs are retroposons that proliferate in the genome by retroposition (Singer, 1982; Weiner et al., 1986; Okada 1991a, 1991b; Deininger and Batzer, 1993). Insertion of a SINE sequence at a site in the genome occurs irreversibly, and it is highly unlikely that SINEs would have become inserted into exactly the same site in different lineages (Okada et al, 1997; Shedlock et al., 2004). Hence, a locus into which a SINE was inserted in an ancestral species would be inherited by successive generations, and its insertion could be shared in all derived descendants. Therefore, when a SINE insertion is observed at an orthologous locus in two or more lineages, it can be considered as a synapomorphy. This method of phylogenetic analysis, termed "the SINE method," has been applied to phylogenetic analysis of many vertebrate animals (e.g., Murata et al., 1993; Shimamura et al., 1997; Nikaido et al., 1999, 2001, 2006; Takahashi et al., 2001a, 2001b; Terai et al., 2003; Sasaki et al., 2004, 2006; Piskurek et al., 2006) including primates (Salem et al., 2003; Roos et al., 2004; Ray et al., 2005; Xing et al., 2005) and is a powerful tool for identifying clades (Miyamoto, 1999; Shedlock and Okada, 2000; Rokas and Holland, 2000; Okada et al., 2004; Shedlock et al., 2004).

The tortoise PolIII/SINE has been identified in the genomes of turtles belonging to Cryptodira (Endoh and Okada, 1986; Endoh et al., 1990). Sasaki et al. (2004) isolated and extensively characterized subgroups of tortoise PolIII/SINE from the genomes of 8 of 11 families of Cryptodira. They identified two major subgroups, namely Cry I and Cry II, which were further subdivided into types A to D and types A to E, respectively. This characterization of tortoise PolIII/SINE indicated that Cry I types B and C and Cry II type A have proliferated in the genomes of species belonging to Testudinoidea. Subgroups that were amplified frequently in a past genome can be informative about species divergence (Kido et al., 1991; Shedlock and Okada, 2000). By focusing on such proliferative retrotransposed SINE subgroups, Sasaki et al. (2004) applied the SINE method to Testudinoidea species, and their work represented the first such phylogenetic analysis in reptilians. The present study applies the SINE method to phylogenetic analysis of Geoemydidae and its related species of Testudinoidea using 6 species of Batagurinae, 15 species of Geoemydinae, 5 species of Testudinidae, and 2 species of Emydidae. We attempt to resolve the current ambiguous phylogeny of testudinoids based on SINE insertion patterns and to verify whether convergent evolution might have occurred repeatedly in these turtles.

MATERIALS AND METHODS

Samples and DNA Extraction

To estimate phylogenetic relationships of Geoemydidae and its relatives by SINE insertion patterns, we analyzed 28 testudinoid species (see Fig. 2). Total genomic DNA was extracted from liver, muscle or blood using the phenol/chloroform protocol (Sambrook et al., 1989) and stored at 4°C in TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA) until use.

Construction of Genomic Libraries, and Isolation and Sequencing of SINE-Inserted Clones

We identified SINE-inserted loci using seven species: *Testudo horsfieldii* and *Kinixys erosa* of Testudinidae, *Mauremys sinensis, Cuora mouhotii, Heosemys annandalii, Heosemys grandis,* and *Melanochelys trijuga* of Geoemydidae. To identify the species from which the loci were identified, we assigned a species-specific

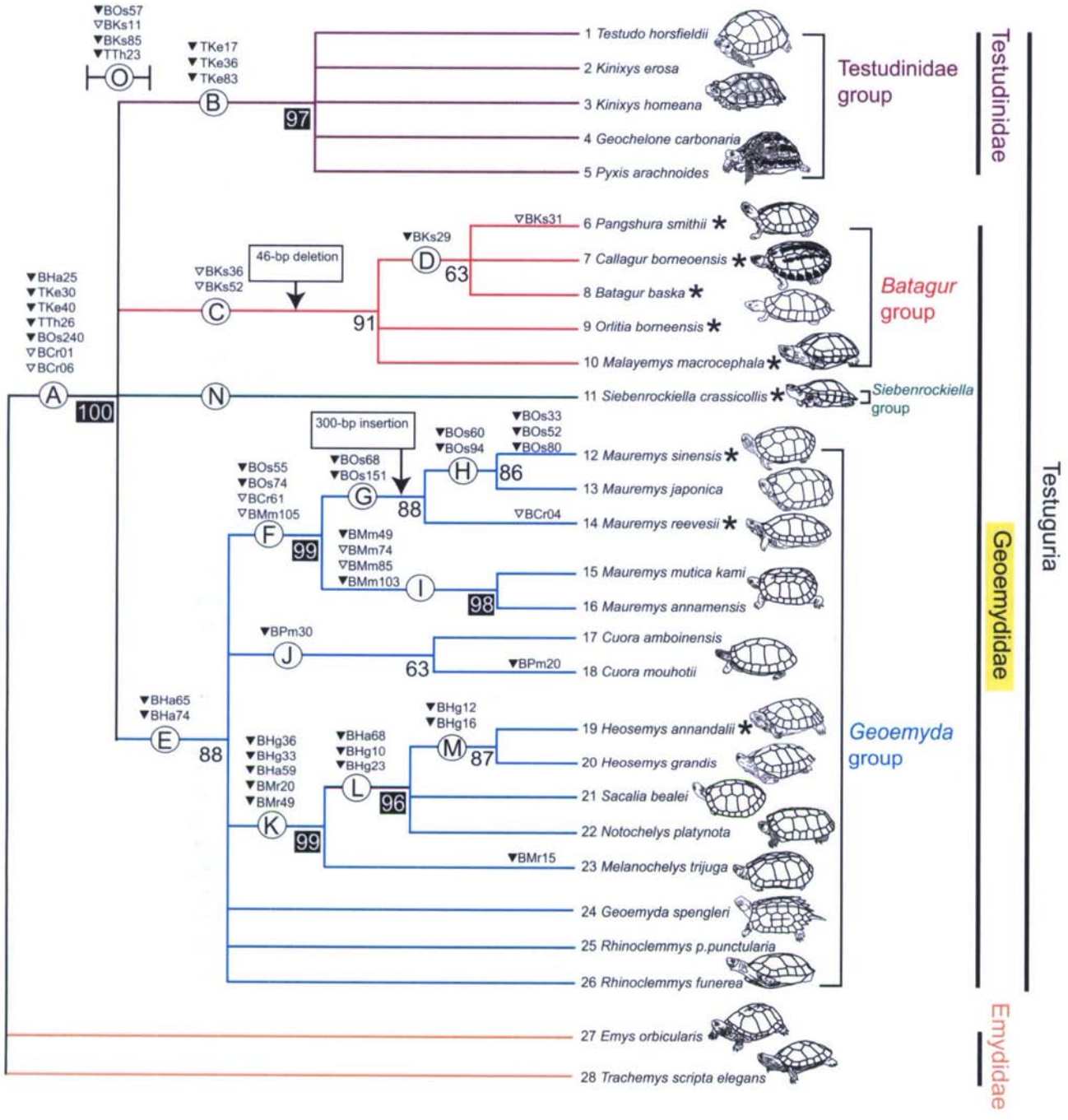


FIGURE 2. Phylogenetic tree of Geoemydidae and related species inferred from SINE insertion patterns. Arrowheads denote SINE insertions into loci that were investigated this study (closed and open arrowheads indicate newly and previously isolated loci, respectively). Numbers on each node are bootstrap values calculated by MP analysis. The bootstrap values with a black background indicate branches that are significantly supported by the likelihood test at $P < 0.05$. The clades labeled with letters denote support by SINE-inserted loci and correspond to the alphabetical panels in Figure 3. Asterisks denote the morphologically defined batagurine genera.

three-letter abbreviation to each locus as follows; 'TTh' derived from *T. horsfieldii*, 'TKe' derived from *K. erosa*, 'BKs' derived from *P. smithii*, 'BOs' derived from *M. sinensis*, 'BCr' derived from *M. reevesii*, 'BMm' derived from *M. mutica kami*, 'BPm' derived from *C. mouhotii*, 'BHa' derived from *H. annandalii*, 'BHg' derived from *H. grandis*, and 'BMr' derived from *M. trijuga*. A number following the three letters indicates a discrete locus, and this number represents the clone number assigned in each genomic library. We used two kinds of probes to screen for tortoise PolIII/SINE-inserted loci in genomic libraries. The Cry1R probe (5′-GATATACCAATCTCCTAGAA-3′) corresponded to the sequence of the Cry I SINE subfamily 3′-tail region.

Probe TEF1 (5′-GGGAGGGATAGCTCAGTGGT-3′) corresponded to common initial sequences in the Cry I and Cry II SINE subfamilies (Ohshima et al., 1996). A detailed description of the uses of these two probes is given in Sasaki et al. (2004). The construction of genomic libraries as well as the screening and sequencing of clones containing SINEs were described previously (Sasaki et al., 2004).

Amplification of Orthologous Loci by PCR

To determine the presence or absence of a SINE unit at each locus in various species, we designed a pair of primers to recognize sequences surrounding the SINE unit, based on the sequences obtained from the isolated clones. The primer sequences used in this study are shown in Appendix 1. Polymerase chain reaction (PCR) amplification was conducted in 25 μl containing 100 ng genomic DNA, 0.5 pM of each primer, 0.2 mM dNTPs, 2.5 μl *Ex Taq* buffer (Takara, Japan), and 1 U of TaKaRa *Ex Taq* polymerase. The cycling program was: 30 cycles of denaturation at 94°C for 45 s, annealing at 47°C to 57°C (details were shown in Appendix 1) for 45 s, and extension at 72°C for 45 s. The PCR products were electrophoresed in 2.0% Agarose L03 (Takara, Japan) in 0.5 × TBE. DNA bands were stained with ethidium bromide and visualized under UV irradiation. The presence or absence of a SINE unit was assessed by noting differences in PCR product size; i.e., products containing a SINE unit were longer than those lacking a SINE unit. Finally, the presence or absence of a SINE at each locus was confirmed by sequencing the bands from representative species. Sequences were deposited in GenBank (AB249019 to AB249371; see the details in Fig. 5, available at the Society of Systematic Biologists website, http://systematicbiology.org). If a locus of a certain species yielded multiple PCR bands, then we subcloned and sequenced the bands to confirm the origin of each band in that species. All PCR products were ligated into pGEM-T (Promega) and subcloned. Thus, each subclone represented a single PCR product, and we further investigated the putative SINE sequences in each case.

We used SINE insertion data to reconstruct phylogenetic trees using the following strategy. If a SINE is present at a particular locus in two or more species, then the insertion can be considered a synapomorphy between or among those species; as such, the species form a clade in the resulting phylogenetic tree. On the other hand, the remaining species that lack a SINE at that locus are treated as an outgroup with respect to that locus. For example, at locus BCr06 (see Fig. 3A), species 1 to 26 have longer PCR products, implying the presence of a SINE; species 27 and 28, however, have shorter PCR products due to the absence of a SINE. Based on these results, we consider that species 1 to 26 are monophyletic and that species 27 and 28 constitute an outgroup. This result is reflected in the resulting phylogenetic tree as a SINE insertion event supporting clade A in Figure 2. Finally, the entire phylogeny can be reconstructed by collecting many informative loci that support various branches.

Evaluation of a SINE Phylogenetic Tree

We analyzed phylogenetic relationships of 28 Testudinoidea species with the maximum parsimony method (MP) using PAUP 4.0b10 (Swofford, 2001). This analysis was based on the data matrix in Appendix 2, which shows the character states for 49 loci in 28 taxa (1 = a SINE is present; 0 = no SINE is present; ? = missing data). The data matrix for phylogenetic analysis were available from TreeBASE (http://www.treebase.org) under the accession number S1558. The tree was constructed by heuristic search with tree-bisection-reconnection (TBR) branch swapping. In the setting of character types, all characters are of the type 'irrev.up' based on the character of irreversible insertion of SINEs and have equal weight. There were 42 parsimony-informative characters. Bootstrap values were calculated using 1000 replications. In addition, we also estimated the statistical significance of each branch of the resulting tree by using the likelihood model established by Waddell et al. (2001).

RESULTS

SINE Insertion Patterns Resolve Early Divergence in Testudinoidea

We isolated 38 SINE-inserted loci from seven genomic libraries, including five geoemydids and two testudinids. In addition, we reanalyzed 11 loci that were previously characterized as informative for phylogenetic analysis of Testudinoidea species (Sasaki et al., 2004). Finally, we assessed the presence or absence of a SINE unit at 49 loci and reconstructed phylogenetic relationships among the species of geoemydids and its testudinoid relatives.

By using 49 SINE loci, we identified 13 clades designated A to M (Fig. 2) in the superfamily Testudinoidea. Figure 3 shows a collection of PCR patterns of representative SINE loci, each of which represents a clade. A detailed characterization of all informative SINE loci is presented in Supplemental Figures 1 to 4 (available online at http://systematicbiology.org). BCr06 represents clade A. Longer PCR products were observed in species belonging to Testuguria. In contrast, two emydids yielded shorter products due to the absence of a SINE unit at this locus (lanes 27 and 28). We determined the sequences of the BCr06 locus from four species to confirm whether each PCR product was derived from an orthologous locus. An alignment of these sequences is shown in Supplemental Figure 5, subsection 7 (available online at http://systematicbiology.org), and the matrix in Appendix 2 indicates the presence (1) or absence (0) of a SINE. Characterization of other SINE loci representing this clade is shown in Supplemental Figure 1A (available online at http://systematicbiology.org). The Testuguria clade (clade A) was characterized by seven SINE-inserted loci and, therefore, was significantly supported by the likelihood test at the $P = 0.005$ level (Waddell et al., 2001).

Clade B is represented by loci TKe17, TKe36 and TKe83 (Fig. 3B for TKe83 and Supplemental Fig. 1B for the

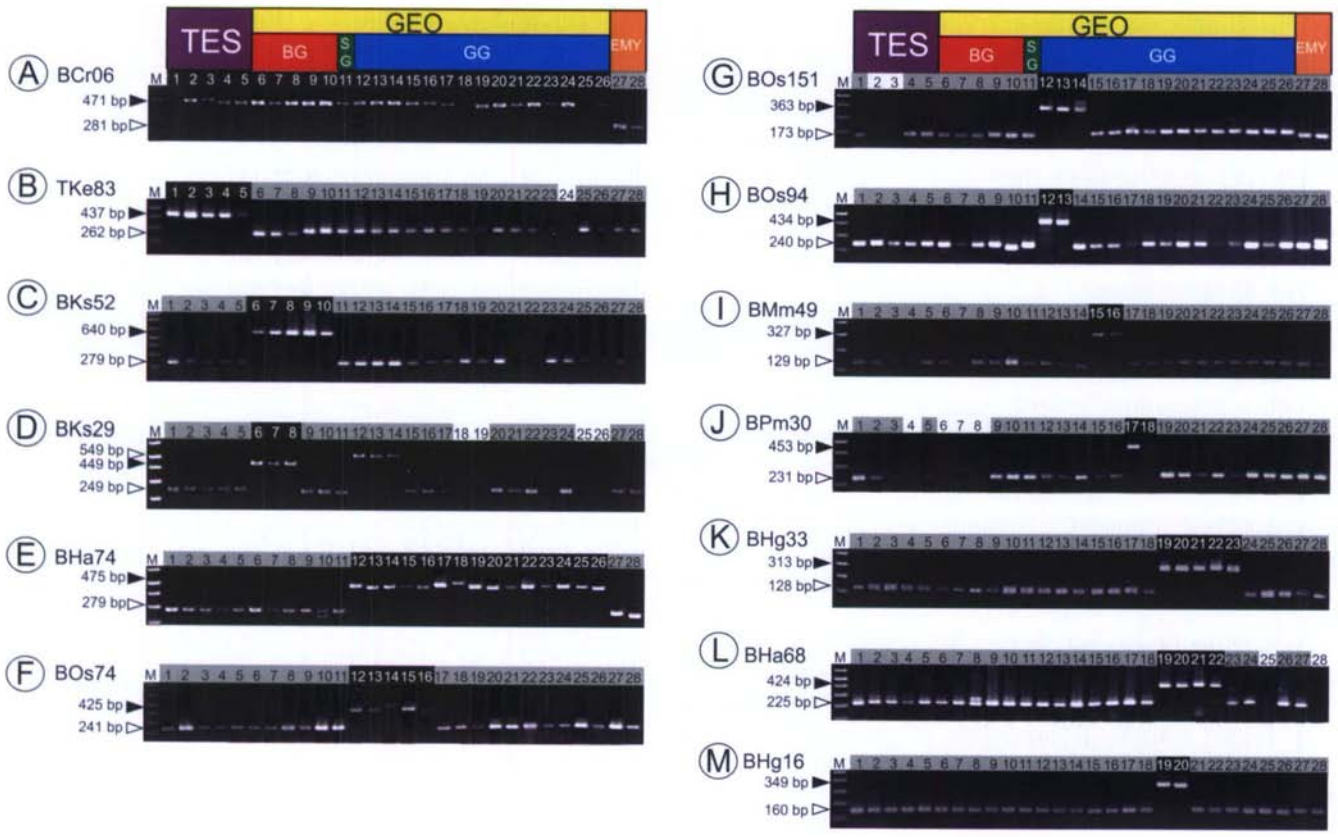


FIGURE 3. Electrophoretic profiles of PCR products of SINE-inserted loci. The lane numbers indicate various species. Lane numbers and species correspond to the numbers assigned to each species in Figure 2. Lane numbers with a black background indicate the presence of a SINE unit, those with a gray background indicate the absence of a SINE unit, and those with a white background indicate that the presence or absence of a SINE unit is unknown. Arrowheads pointing to PCR bands indicate the presence (filled) or absence (open) of a SINE unit. The colored rectangles, containing abbreviations on the upper side in each box, indicate the classification of each group in the present study. The abbreviations are as follows: TES = Testudinidae (purple); GEO = Geoemydidae (yellow); EMY = Emydidae (orange); BG = *Batagur* group (red); SG = *Siebenrockiella* group (green); GG = *Geoemyda* group (blue).

other two loci, available online at http:// systematicbiology.org). A long product containing a SINE unit was specifically observed in *Testudo horsfieldii, Kinixys erosa, Kinixys homeana, Geochelone carbonaria*, and *Pyxis arachnoides*, belonging to Testudinidae. A short product lacking a SINE unit was observed in the respective loci of the species of Geoemydidae and Emydidae except for several species that did not yield a product. For specimens that yielded no PCR product at certain loci (i.e., "unknown"; Fig. 3B for lane 24), the MP analysis supported that those species form a clade that differs from the one suggested by the loci in Figure 3B (88% BP, clade E in Fig. 2). Overall, the results of these three SINE loci (Fig. 3B and Supplemental Fig. 1B, available at online http://systematicbiology.org) indicate that the five testudinids constitute a monophyletic group in the Testuguria. Because the Testudinidae group (clade B) was characterized by three loci with no conflicting loci, the likelihood test evaluated this clade as significant at the $P = 0.05$ level.

Clade C is represented by loci BKs36 and BKs52 (Fig. 3C for BKs52 and Supplemental Fig. 1C for BKs52, available online at http://systematicbiology.org). In BKs52, a long product was obtained for *Pangshura smithii* (lane 6), *Callagur borneoensis* (lane 7), *Batagur baska* (lane 8), *Orlitia borneensis* (lane 9), and *Malayemys macrocephala* (lane 10), whereas a short product was obtained for the remaining Geoemydidae species (lanes 11 to 26) together with the species of Testudinidae and Emydidae (lanes 1 to 5, 27, and 28). The results in Figure 3C indicate that *P. smithii, C. borneoensis, B. baska, O. borneensis*, and *M. macrocephala* form a clade in the geoemydids. For descriptive purposes, we propose to call this clade the "*Batagur* group."

Clade D is represented by the BKs29 locus (Fig. 3D). A long product (449 bp) was observed for *P. smithii* (lane 6), *C. borneoensis* (lane 7), and*B. baska* (lane 8), whereas a short product (249 bp) was observed for the other species of the *Batagur* group together with the remaining testudinoids, except for four species that did not yield a PCR product. *Mauremys sinensis* (lane 12), *M. japonica* (lane 13), and *M. reevesii* (lane 14) yielded longer products (549 bp) than those containing a SINE unit (449 bp). The difference was due to an additional insertion of a non-SINE-related sequence (300 bp) upstream of the SINE 5′ flanking region in the genome of a presumptive

common ancestor of these species (see the alignment in Supplemental Figs. 5, subsection 15, available online at http://systematicbiology.org). For specimens that yielded no PCR product at locus BKs29 (lanes 18, 19, 25, and 26 in Fig. 3D), MP analysis supported placement of those species in a different clade (88% BP, clade E in Fig. 2) than that indicated by the locus in Figure 3D (63% BP, clade D in Fig. 2). The results shown in Figure 3D indicate that *P. smithii*, *C. borneoensis*, and *B. baska* are a monophyly in the *Batagur* group.

Clade E is represented by loci BHa65 and BHa74 (Fig. 3E for BHa74 and Supplemental Fig. 1D for BHa65, available online at http://systematicbiology.org). A long product containing a SINE unit was observed in *M. sinensis* (lane 12), *M. japonica* (lane 13), *M. reevesii* (lane 14), *M. mutica kami* (lane 15), *M. annamensis* (lane 16), *Cuora amboinensis* (lane 17), *C. mouhotii* (lane 18), *Heosemys annandalii* (lane 19), *H. grandis* (lane 20), *Sacalia bealei* (lane 21), *Notochelys platynota* (lane 22), *Melanochelys trijuga* (lane 23), *Geoemyda spengleri* (lane 24), *Rhinoclemmys punctularia punctularia* (lane 25), and *R. funerea* (lane 26). A short product lacking a SINE unit was obtained for the remaining Geoemydidae species (lanes 6 to 11) together with the species of Testudinidae and Emydidae (lanes 1 to 5, 27, and 28). *M. macrocephala* (lane 10) at locus BHa74 yielded multiple bands. Sequencing revealed that the upper band of *M. macrocephala* represented the respective orthologous locus (see the alignment in Supplemental Figs. 5, subsection 12, available online at http://systematicbiology.org). Thus, the results in Figure 3E as well as Supplemental Figure 1D (available onnline at http://systematicbiology.org) indicate that 15 geoemydids form a clade in Testuguria. We propose that this clade be called the "*Geoemyda* group."

The Three Lineages of Species in the Geoemyda Group

Clade F is supported by four loci, BOs55, BOs74, BCr61, and BMm105. For BOs74 (Fig. 3F), a long product was observed in five *Mauremys* species (*M. sinensis*, lane 12; *M. japonica*, lane 13; *M. reevesii*, lane 14; *M. m. kami*, lane 15; and *M. annamensis*, lane 16), whereas a short product was observed in the other species of the *Geoemyda* group (lanes 17 to 26) together with the remaining species of Testudinoidea (lanes 1 to 11, 27, and 28). Several faint additional bands in gels were shown to be non-specific by sequencing. Clade F in Figure 2 shows that the five *Mauremys* species form one of the major lineages in the *Geoemyda* group. For descriptive purposes, we propose that the clade F supported by the SINE insertions in Figure 3F be called the "*Mauremys* lineage." Because the *Mauremys* lineage was characterized by four loci with no conflicting loci; the likelihood test evaluated this clade as significant at the $P = 0.05$ level.

Clade G is represented by loci BOs68 and BOs151 (Fig. 3G for BOs151 and Supplemental Fig. 3A for BOs68, available online at http://systematicbiology.org). In BOs151, a long product was observed in *M. sinensis* (lane12), *M. japonica* (lane13), and *M. reevesii* (lane14), whereas a short product was observed in *M. m. kami* (lane15), and *M. annamensis* (lane16) together with the remaining testudinoids, except for two species that did not yield a PCR product. For specimens that yielded no PCR product at certain loci (lanes 2 and 3 in Fig. 3G), MP analysis supported placement of those species in a different clade (clade B, containing the species of lanes 2 and 3 with 97% BP; Fig. 2) from the one indicated by the loci in Figure 3G (88% BP, clade G in Fig. 2). Clade G in Figure 2 indicated that *M. sinensis*, *M. japonica*, and *M. reevesii* form a monophyletic group in the *Mauremys* lineage.

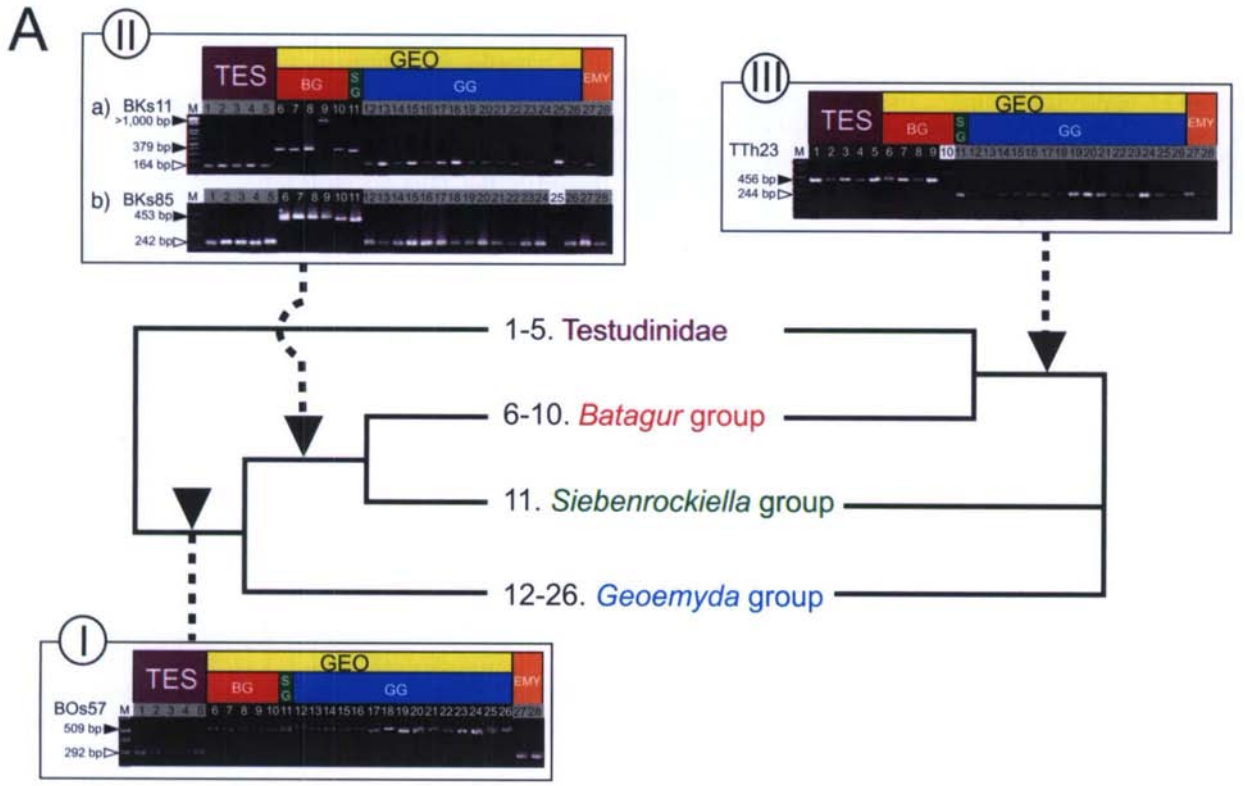
Clade H is represented by loci BOs60 and BOs94 (Fig. 3H for BOs94 and Supplemental Fig. 3B for BOs60, available online at http://systematicbiology.org). A long product was observed for *M. sinensis* (lane 12) and *M. japonica* (lane 13), whereas a short product was observed for *M. reevesii* (lane 14), *M. m. kami* (lane 15), and *M. annamensis* (lane 16) together with the remaining testudinoids. We sequenced two bands per specimen observed in *Trachemys scripta elegans* (lane 28 in Fig. 3H); the upper band was nonspecific and the lower band represented an orthologous locus. Clade H in Figure 2 indicates that *M. sinensis* and *M. japonica* form a sister relationship within a clade that includes *M. reevesii*.

Clade I is represented by loci BMm49, BMm103, BMm74, and BMm85 (Fig. 3I for BMm49 and Supplemental Fig. 3C for the other three loci, available online at http://systematicbiology.org). For BMm49, a long product was observed in *M. m. kami* (lane 15) and *M. annamensis* (lane 16), whereas a short product was observed in *M. sinensis*, *M. japonica*, and *M. reevesii* together with the remaining species. Because clade I was characterized by four loci with no conflicting loci, the likelihood test evaluated this clade as significant at the $P = 0.05$ level.

The SINE-inserted patterns in the *Mauremys* lineage (clade F) indicate two diverged lineages, namely the *M. sinensis*, *M. japonica*, and *M. reevesii* lineage and the *M. m. kami* and *M. annamensis* lineage. Additionally, the non-SINE-related insertion at locus BKs29 (Fig. 3D) also suggests monophyletic relationships between *M. sinensis*, *M. japonica*, and *M. Reevesii* if we treat it as a synapomorphy based on a rare molecular change (Rokas and Holland, 2000). Furthermore, *M. sinensis* and *M. japonica* were closely related in the former lineage (clade H in Fig. 2).

Clade J is represented by locus BPm30 (Fig. 3J). A long product was observed in *C. amboinensis* (lane 17) and *C. mouhotii* (lane 18), whereas a short product was observed in the other geoemydids together with the testudinids and emydids, except four species that did not yield a PCR product. In this regard, for locus BPm30 (lanes 4, 6, 7, and 8 in Fig. 3J), MP analysis supported placement of these species in different clades (Fig. 2; clade B containing the species of lane 4 with 97% BP, clade C containing the species of lanes 6, 7, and 8 with 91% BP) from the clade indicated by the locus in Figure 3J (63% BP, clade J in Fig. 2). Thus, the SINE insertion at locus BPm30 occurred in a common ancestor of *C. amboinensis* and *C. mouhotii*.

Clade K is represented by loci BHg33, BHg36, BHa59, BMr20, and BMr49 (Fig. 3K for BHg33 and Supplemental Fig. 2C for the other four loci, available online at http://systematicbiology.org). In BHg33, a long product was



B

BOs57 locus (A-I)

```
1.Testudo           GAGTTATTCCTATAAAGATATAGATAT----------------------------------TCCTAGAAGCTTCCATCTTAAAAAGGA
6.Pangshura         GAATTATTCCTATAAAGATATAGATAT  tortoise pol/III SINE              TATTATAAGCTTCCATCTTAAAAAGGA
11.Siebenrockiella  GAATTATTCCTATAAAGATACAGATAT  tortoise pol/III SINE              TATTATAAGCTTCCATCTTAAAAAGGA
14.Mau.reevesii     GAATTATTCCTATAAAGATATAGATAT  tortoise pol/III SINE              TATTATAAGCTTCCATCTTAAAAAGGA
                    **.*****************.******                                     *..**.**********************
```

BKs11 locus (A-II-a)

```
1.Testudo           TTTTT-ATAGTTTTGTAATTAAGAAA----------------------------------AGTATTTTCTTCAAATATAGC
6.Pangshura         TTTTTTACAGCTTTGTAATTAAGAAA  tortoise pol/III SINE              ACTATTTTCTTCAAAAATAGC
11.Siebenrockiella  TTTTT-ATAATTTTGTAATTAAGAAA  tortoise pol/III SINE              ACTATTTTCTTCAAATATAGC
14.Mau.reevesii     TTTTTTATAGTTTTGTAATTAAGAAA----------------------------------ACTATTTTCTTCAAATATAGC
                    ***** *.*..***************                                    *.*************.*****
```

BKs85 locus (A-II-b)

```
1.Testudo           GAAGACAAACCATCCTTCTGAACAACA----------------------------------CTTACAGTTTATATCTCCATTAACTTG
6.Pangshura         GAAGACAAACCATCCTTCTGAACAACA  tortoise pol/III SINE              CTTACAGTTTGTATCTCCATTAACTTG
11.Siebenrockiella  GAAGACAAACCATCCTTCTGAACAACA  tortoise pol/III SINE              CTTACAGTTTGTATCTCCATTAACTTG
14.Mau.reevesii     GAAGACAAACCATCCTTCTAAACCACA----------------------------------CTTACAGTTTGTATCTCCATTAACTTA
                    ******************.***.***                                    **********.***************.
```

TTh23 locus (A-III)

```
1.Testudo           TGG-AAATGGGGGCTGTTGGAATCTT  tortoise pol/III SINE              ATCTTTTATGTTTCTTGAGTTGGCAAT
6.Pangshura         TGG-AAATGGGGGCTGTTGGAACCTT  tortoise pol/III SINE              ATCTTTTATGTTTCTTGAGCTGGCAAT
11.Siebenrockiella  TGG-AAATGGGGGCTGTTGGAACCTT----------------------------------ATCTTTTATGTTTCTTGAGTTGGCAAT
14.Mau.reevesii     TGGGAAATGGGTGCTGTTGGAACCTT----------------------------------ATCTTTTACGTTTCTTGAGTTGGCAAT
                    *** *******.**********.***                                   ********.**********.*******
```

FIGURE 4. Inconsistent SINE insertion patterns that were observed in relationships among four groups. A: Electrophoresis profiles of PCR products of SINE-inserted loci and phylogenetic relationships among the four groups inferred from four SINE-inserted loci. Boxes I–III: The loci are arranged to correspond to each clade of the phylogenetic trees. The display method here is defined in Figure 3. B: Compilation of part of the sequences of four loci from representative species. The corresponding lane numbers are shown to the left of each genus (or species) name. Asterisks indicate an exact consensus at that position, and dots indicate a difference in one species.

obtained for *H. annandalii* (lane 19), *H. grandis* (lane 20), *S. bealei* (lane 21), *N. platynota* (lane 22), and *M. trijuga* (lane 23) in the *Geoemyda* group, whereas a short product was obtained for the other species of the *Geoemyda* group together with the remaining testudinoids. Thus, the five SINE insertion patterns in clade K shown in Figure 2 indicate that *H. annandalii, H. grandis, S. bealei, N. platynota*, and *M. trijuga* were derived from a common ancestor and that these species form one of the major evolutionary lineages in the *Geoemyda* group. For descriptive purposes, we propose that the clade K be called the "*Heosemys* lineage." The *Heosemys* lineage was characterized by five SINE-inserted loci and, therefore, was significantly supported by the likelihood test at the $P = 0.005$ level.

Clade L is represented by loci BHg10, BHg23 and BHa68 (Fig. 3L for BHa68 and Supplemental Fig. 4A for the other two loci, available online at http:// systematicbiology. org). *H. annandalii* (lane 19), *H. grandis* (lane 20), *S. bealei* (lane 21), and*N. platynota* (lane 22) yielded long products, whereas *M. trijuga* (lane 23) yielded a short product together with the remaining species except for two species that did not yield a PCR product. For specimens that yielded no PCR product at certain loci (lanes 25 and 28 in Fig. 3L), MP analysis supported placement of species in lanes 25 and 28 outside of the *Heosemys* clade (99% BP, clade K in Fig. 2). Clade L in Figure 2 indicates that *M. trijuga* diverged first in the *Heosemys* lineage and that the remaining four species were derived from a common ancestor. Because clade L was characterized by three loci with no conflicting loci, the likelihood test evaluated this clade as significant at the $P = 0.05$ level.

Clade M is represented by loci BHg12 and BHg16 (Fig. 3M for BHg16 and Supplemental Fig. 4B for BHg12, available online at http://systematicbiology.org). In BHg16, a long product was observed in *H. annandalii* (lane 19) and *H. grandis* (lane 20), whereas a short product was observed in the remaining species. The clade M in Figure 2 indicates that *H. annandalii* and *H. grandis* form a monophyletic group within the *Heosemys* lineage. Based on morphology, *H. annandalii* was formerly classified as the monotypic genus *Hieremys* of subfamily Batagurinae (Hirayama, 1984; Gaffney and Meylan, 1988). The SINE analysis, however, clearly demonstrates that *H. annandalii* is positioned deeply in the *Geoemyda* group, and this conclusion is supported by the molecular phylogenies shown in Figure 1C and D (Spinks et al., 2004; Diesmos et al., 2005).

Thus, the SINE insertion patterns indicate three distinct lineages in the *Geoemyda* group: (1) *Mauremys* lineage (clade F); (2) *Cuora* lineage (clade J); and (3) *Heosemys* lineage (clade K).

We also identified seven loci at which a SINE unit was specifically inserted into each of the species (see Supplemental Fig. 6, available online at http:// systematicbiology.org). Of course, it is possible that SINE insertions at these loci also occurred in species that were not investigated in this study.

Inconsistent Phylogeny in the Early Divergence of Testuguria

As shown Figure 2, the SINE insertion patterns reveal an evolutionary history of the divergence of species in Testuguria. The assembly of Testuguria species used in the present analysis comprises at least four groups: Testudinidae, *Batagur*, *Siebenrockiella*, and *Geoemyda*. The *Siebenrockiella* group contains a single species in the present analysis, although it is possible that other species that were not analyzed here belong in this group. Interestingly, concerning the phylogenetic relationships among these four groups, we characterized four loci among which there were inconsistent SINE insertion patterns. In the analysis of locus BOs57 (Fig. 4A-I), the species of lanes 6 to 26 yielded a long product, whereas the remaining species yielded a short product. For *M. macrocephala* (lane 10 in Fig. 4A-I) locus BOs57, we sequenced the two PCR products and confirmed that the upper band represented an orthologous locus and that the lower band was nonspecific. The SINE insertion pattern for this locus indicates that the species belonging to the *Geoemyda* group and the *Batagur* group were derived from a common ancestor (left-side tree in Fig. 4A).

For loci BKs11 and BKs85 (Fig. 4A-II), a long product was observed in species of the *Batagur* group and *S. crassicollis*, whereas a short product was observed in the remaining species except for *R. p. punctularia* at locus BKs85, for which we sequenced the two PCR products (lane 25 in Fig. 4A-II-b). Both products were nonspecific. *O. borneensis* yielded long products that were longer than that expected for insertion of a single SINE unit at locus BKs11 (lane 9 in Fig. 4A-II-a). Sequencing revealed that a fragment of >1 kbp was inserted at the SINE flanking region at the 5′ end (see alignment in Supplemental Figs. 5, subsection 40 available online at http://systematicbiology.org).

For specimens that yielded no PCR product at locus BKs85 in the genome of *R. p. punctularia* (lane 25 in Fig. 4A-II-b), MP analysis supported placement of this species in the *Geoemyda* group (clade E, containing the species of lane 25 with 88% BP; Fig. 2). The results in Figs. 4A-II indicate that the *Batagur* group and *S. crassicollis* form a clade (left-side tree in Fig. 4A).

For locus TTh23 (Fig. 4A-III), a long product was observed for species of Testudinidae (lanes 1 to 5) and the *Batagur* group except for *M. macrocephala* (lane 10), whereas a short product was observed in the remaining species (lanes 11 to 28). The product observed in *M. macrocephala* (lane 10) was nonspecific. Therefore, the presence or absence of a SINE unit at locus TTh23 in this species remains unclear. The results in Figure 4A-III indicate that species of Testudinidae and the *Batagur* group, except for *M. macrocephala*, form a clade that crosses the boundaries of the family classification (right-side tree in Fig. 4A).

Figure 4A summarizes two possible phylogenies among four groups, in which loci BOs57, BKs11, and BKs85 and locus TTh23 apparently are contradictory. To confirm the SINE insertion data, we determined the

flanking sequences of each locus, and the alignments are shown in Figure 4B (more detailed alignments were shown in Supplemental Figs. 5, subsection 42, available online at http://systematicbiology.org). The data demonstrate that the SINE insertion at each locus was derived from the orthologous locus. Furthermore, based on the phylogenetic relationships inferred from these loci, we attempted to validate the following three hypotheses based on a statistical test for SINE insertion patterns (Waddell et al., 2001); (1) monophyly of Geoemydidae (= BOs57); (2) monophyly of the *Batagur* group and *Siebenrockiella* group (= BKs11 and BKs85); and (3) monophyly of the Testudinidae group and *Batagur* group (= TTh23). However, none of these hypotheses had significant statistical support ($P = 0.3333$ for hypotheses 1 and 3, and $P = 0.1481$ for hypothesis 2). We therefore concluded that the divergence order of the four groups should be treated as a polytomy. Such incongruencies with respect to topologies constructed by SINE insertion patterns have been reported in phylogenetic studies of African cichlids (Takahashi et al., 2001a, b; Terai et al., 2003), baleen whales (Nikaido et al., 2006; Sasaki et al., 2006), and primates (Salem et al., 2003; Ray, et al., 2005; Xing et al., 2005). There are two possible explanations for such SINE insertion homoplasies. One possibility is that SINE elements may be precisely removed from the genome, as has happened to some *Alu* elements in primate genomes (van de Lagemaat et al., 2005). If such precise deletion occurred in certain species (or lineages) at a particular locus, the SINE insertion pattern would be represented as an insertion homoplasy as shown in Figure 4A. The other possibility is that incomplete lineage sorting occurred in SINE-inserted alleles. This putative incomplete lineage sorting could have been caused by rapid speciation in the ancestral population before fixation of alleles that either contained or lacked a SINE unit (Shedlock, et al., 2000, 2004). If the SINE insertions in Figure 4A occurred in ancestral species of Testuguria, and the ancestral population that retained the polymorphisms of SINE-inserted alleles diverged into the four lineages, then the fixation of alleles containing or lacking a SINE insertion could have undergone genetic drift in each lineage to produce the incongruent insertion patterns observed.

DISCUSSION

Close Relationships among Testuguria Species

The SINE insertion patterns at loci BHa25, TKe30, TKe40, TTh26, BOs240, BCr01, and BCr06 strongly suggest that the Geoemydidae are closely related to the Testudinidae rather than Emydidae (stem A in Fig. 2). The families Testudinidae and Geoemydidae are classified as Testuguria (Joyce et al., 2004), and its monophyly was suggested by earlier morphological phylogenetic studies (Hirayama, 1984; Gaffney and Meylan, 1988) as well as by recent molecular studies (Shaffer et al., 1997; Spinks et al., 2004; Krenz et al., 2005). Preliminary SINE insertion analysis (Sasaki et al., 2004) suggested close relationships among the species of Testuguria based on two SINE-inserted loci (BCr01 and BCr06). In the present analysis, we included 12 additional species of Testudinoidea. By using 28 species total, we showed significant close relationships among the Testuguria species by characterizing an additional five SINE-inserted loci (seven loci total).

Rapid Radiation during the Initial Divergence of the Four Major Lineages of Testuguria

In the present study, we identified four major evolutionary groups in the monophyletic Testuguria, namely Testudinidae (clade B in Fig. 2), *Batagur* group (clade C in Fig. 2), *Siebenrockiella* group (clade N in Fig. 2), and *Geoemyda* group (clade E in Fig. 2). We obtained an inconsistent phylogenetic topology using the SINE insertion patterns for loci BOs57, BKs11, BKs85, and TTh23 (period O in Fig. 2). We considered the effects of precise removal of SINE elements or, alternatively, incomplete lineage sorting of SINE alleles as underlying evolutionary processes that might produce the incongruent topologies observed.

We expect precise removal or parallel insertion of nonorthologous repeats to be rare events based on the lack of any known mechanism that precisely removes SINEs from the genome and the overall paucity of cases described to date for this phenomenon (reviewed by Ray et al., 2006). For instance, an examination of more than 11,000 *Alu* loci in primates at the nucleotide sequence level found evidence for insertion homplasy at the rate of only 0.05% (Ray et al., 2006). Similarly, an informatics-based comprehensive review of *Alu* loci in human and chimp genome assemblies confirmed only 9 precise SINE excisions in the human assembly and just 27 in the chimpanzee, due presumably to illegitimate homologous recombination events between target site duplications (van de Laemaat et al., 2005). Copy number of tortoise polIII/SINEs in the sea turtle genome (*Chelonia mydas*) has been estimated to be $\sim 2.7 \times 10^4$ based on our counts of positive plaques to total plaques in a screen for SINE-inserted loci, and copy number of *Alus* is even greater, on the order of $\sim 10^6$, in primates (Deininger et al., 1981). Assuming an *Alu*-like model of insertion homoplasy is realistic in turtles, we might expect as few as only 10 to 20 PolIII/SINEs in the turtle genomes we examined to have been precisely removed in manner that cannot be detected experimentally. The possibility that all four independent loci diagrammed in Figure 4 were affected by this process seems exceedingly remote based on statistical considerations and available comparative genomic data.

To better understand the source and nature of rare SINE and LINE homoplasy and to rigorously establish phylogenetic hypotheses, we strongly advocate the examination of multiple independent loci when inferring common ancestry among species (Shedlock and Okada, 2000; Shedlock et al., 2000, 2004; Okada et al., 2004). In addition to analyzing multilocus insertion matrices with maximum parsimony and bootstrap replication (e.g., methods in PAUP*; Swofford, 2001), likelihood-ratio

tests provide useful statistical probabilities for evaluating clades supported with data from different numbers of independent loci (Waddel et al., 2000). Moreover, informatics-based strategies can be useful for screening non-orthologous loci from large data sets of repeats available from genome sequencing projects. For example, Bashir et al. (2006) recently surveyed more than 1000 repeat loci and used skewed divergence levels in SINE and LINE flanking versus repeat sequences to minimize the likelihood of including nonorthologous insertions in their in silico phylogenetic analysis of 28 mammal species. By evaluating large numbers of loci in combination with sequence divergence profiles, a single case of possible insertion homplasy was detected in a SINE repeat shared in the domestic cat and rat, to the exclusion of other species examined. A similar putative case of insertion homplasy was previously suggested for a single SINE locus on the Y chromosome in cats, *SMCY* SN327 (Pecon-Slattery et al., 2004). However, until more than one retroelement locus is evaluated in the bobcat and related lineages, the conclusions of Pecon-Slattery et al. (2004) about the relative importance of insertion homoplasy to accurately inferring phylogeny with SINE insertions remain premature.

An alternative explanation for the incongruent insertion patterns we observed is that they represent incomplete lineage sorting of SINE alleles among the four major testugurian groups. Relationships between the Geomydidae and Testudinidae inferred from morphology and nucleotide sequences vary depending on the genes analyzed and the methods used to reconstruct trees (Fig. 1). A similar phenomenon was reported in ambiguous phylogenetic relationships in baleen whales with regard to the order of early divergence of species among balaenopterids and gray whales (Rychel et al., 2004; Arnason, 2005; Rychel et al., 2005). Based on a comprehensive phylogenetic study of mitochondrial genomic data, Sasaki et al. (2005) suggested that rapid radiation occurred during the period of early divergence of the four major lineages of the clade composed of balaenopterids and gray whales. Moreover, a SINE method–based phylogenetic study of baleen whales (Nikaido et al., 2006) indicated that SINE insertion patterns among the four lineages indicated an inconsistent topology, and they suggested that such incomplete lineage sorting of SINE alleles was caused by rapid radiation during the initial divergence of the four lineages. Inconsistent SINE insertion patterns due to an ancestral polymorphism also have been observed in primate phylogenies (Salem et al., 2003; Xing et al., 2005). Thus, there are several examples for which SINE insertion patterns have revealed past rapid radiation events in lineages that were unclear based on sequence comparisons of genomic or mitochondrial DNAs.

We suggest that the inconsistent SINE insertion patterns observed in this study are a consequence of rapid radiation of the Testuguria in the Eocene, based on increased fossil diversity during this period (Hutchison, 1998; Claude et al., 2004). Hutchinson (1998) suggested that this rapid divergence was promoted by warmer polar climates that enabled Asian emydids and testudinids to expand their range via radiation in Eurasia and migration to North America.

Evolutionary History of the Four Groups in Testuguria

We examined just 5 of about 55 species (Ernst et al., 2000) of Testudinidae (clade B in Fig. 2). Due to this small sample size, the data do not necessarily imply that SINE insertions at loci TKe17, TKe36, and TKe83 occurred in the common ancestor of the family Testudinidae. The phylogenetic interrelationships of the *Batagur* group (clade C in Fig. 2) are consistent with those from previous molecular studies (Fig. 1C: Spinks et al., 2004; Fig. 1D: Diesmos et al., 2005), but inconsistent with those of previous morphological studies (Fig. 1A: Hirayama, 1984; Gaffney and Meylan, 1988). Based on the shared possession of broad triturating surfaces of jaws (Fig. 1) and additional characters (Yasukawa et al., 2001), *S. crassicollis, M. sinensis, M. reevesii,* and *H. annandalii* should be included in the the *Batagur*. Nevertheless, the present study together with previous molecular systematic studies demonstrate the polyphyly of the Batagurinae group with respect to the phylogenetic position of *M. sinensis, M. reevesii, H. annandalii,* and *S. crassicollis* (Fig. 1C, D) (Honda et al., 2002b; Barth et al., 2004; Spinks et al., 2004; Diesmos et al., 2005). These studies strongly argued that the previous classification of Batagurinae sensu Gaffney and Meylan (1988) did not reflect evolutionary history. On the other hand, the monophyly of *Pangshura, Callagur* and*Batagur* within the *Batagur* group is consistent with both morphological and molecular studies (Fig. 1A, C, and D).

Siebenrockiella crassicollis is part of the polytomy that we suggest reflects a period of rapid radiation among the four groups (clade N in Fig. 2). Morphological studies suggested that this species forms a clade with *O. borneensis*, and this clade has been regarded as the most primitive lineage in the subfamily Batagurinae (Fig. 1A) (Hirayama, 1984; Gaffney and Meylan, 1988; Yasukawa et al., 2001). Spinks et al. (2004) suggested that a clade comprising *S. crassicollis* and *G. spengleri* was sister to the *Batagur* group (Fig. 1C), whereas Diesmos et al. (2005) suggested that the genus *Siebenrockiella* forms a clade with the species of the *Geoemyda* group (sensu this study, but excluding *Rhinoclemmys*) (Fig. 1D). However, neither grouping received high bootstrap support (Fig. 1C, D). Thus, the phylogenetic position of *S. crassicollis* remains unclear. Our results suggest that it diverged early, and existed as a single independent lineage longer than previously thought. Diesmos et al. (2005) assigned *Siebenrockiella leytensis*, formerly included in *Geoemyda* or *Heosemys*, to this genus. We have no samples of this species, and hence could not test whether it is a member of the *Siebenrockiella* group.

The *Geoemyda* group, assembled for the first time in the present study, comprises 15 geoemydines, including three former batagurines, *M. reevesii*, *M. sinensis*, and *H. annandalii* (clade E in Fig. 2). It is noteworthy that *Rhinoclemmys* and *Geoemyda* species are included in this clade, contrary to previous molecular studies (Fig 1C, D). The composition of the *Geoemyda* group recognized here is nearly equivalent to the subfamily Geoemydinae based on morphological data (Fig. 1A, B), but with the addition of the three former batagurines, *M. sinensis*, *M. reevesii*, and *H. annandalii* recently moved to this subfamily (Spinks et al., 2004; Diesmos et al., 2005). Given that most morphological characters uniting members of the *Geoemyda* group are primitive (with the possible exception of the large and elongated foramen palatinum posterius), there are no obvious synapomorphies the group (Yasukawa et al., 2001). In contrast, our SINE data provides clear support for this group.

We recognize three principal lineages in the *Geoemyda* group (clades F, J, and K in Fig. 2). Phylogenetic relationships among the *Mauremys* lineage (clade F in Fig. 2) are consistent with the phylogeny suggested by molecular studies (Wu et al., 1998; Honda et al., 2002a, b; Barth et al., 2004; Feldman and Parham, 2004; Spinks et al., 2004; Diesmos et al., 2005). These molecular studies, however, failed to obtain significant statistical support for the grouping of *M. sinensis* and *M. japonica* (Barth et al., 2004; Spinks et al., 2004). In contrast, our SINE insertion results clearly indicate this monophyletic relationship. Still, our study does not include the three western Palearctic *Mauremys* species (*M. caspica*, *M. rivulata*, and *M. leprosa*), and thus further analysis including these species will reveal detailed relationships among *Mauremys* species.

The *Heosemys* lineage comprises four different genera (*Heosemys*, *Sacalia*, *Notochelys*, and *Melanochelys*) (clade K in Fig. 2). In this lineage, *M. trijuga* forms a sister taxon to a clade of *H. annandalii*, *H. grandis*, *S. bealei*, and *N. platynota*, thereby defining a novel evolutionary history.

Morphological Convergent Evolution of M. sinensis, M. reevesii, *and* H. annandalii

Our SINE insertion analysis clearly shows that the former batagurines, *M. sinensis*, *M. reevesii*, and *H. annandalii*, form a clade with the species of Geoemydinae (clades G and M in Fig. 2). Polyphyly of Batagurinae had been suggested by molecular phylogenetic studies (Honda et al., 2002a, 2002b; Barth et al., 2004; Feldman and Parham, 2004; Spinks et al., 2004; Diesmos et al., 2005), which is confirmed by our study. Monophyly of this subfamily has been based on morphological (principally skull) characters. However, our results suggest that *M. sinensis*, *M. reevesii*, and *H. annandalii* are members of the *Geoemyda* group (sensu this study) that independently evolved several morphological traits shared with batagurine genera. The three species share two character states, expanded triturating surfaces comprised of large premaxillae and maxillae and an expanded anteromedial portion of the lower triturating surface of the dentaries, which are regarded as unique synapomorphies of Batagurinae (Fig. 1A; Hirayama, 1984; Gaffney and Meylan, 1988). The sharing of both characters is probably the result of convergent evolution. In the *Mauremys* lineage, *M. sinensis*, and *M. reevesii* are paraphyletic with respect to *M. japonica*. If the putative convergent evolution occurred in a common ancestor of these species, regressive evolution of the morphs may have once again occurred in the ancestral species of *M. japonica*. In addition, the two *Mauremys* species and *H. annandalii* in the *Geoemyda* group have convergently acquired similar characters. Claude et al. (2004) used geometric morphometric analysis to investigate the correlation between morphological variation of the turtle skull and adaptive radiation, and they suggested that morphological evolution of the skulls of Testudinoidea species was easily affected by environmental factors (e.g., habitat and diet) without phylogenetically constrained morphological evolution. We hypothesize that feeding habits of the three species guided the evolution of the shape of triturating surfaces in a manner similar to that of batagurines. However, feeding habits vary widely even among Batagurinae species, as evidenced by the inclusion of omnivorous, herbivorous, carnivorous, and conchifrageous species in this subfamily (Ernst et al., 2000). Moreover, among the three species mentioned above, the feeding habits are different from one another and affected by body size and sexes. Juveniles and adult males of *M. sinensis* are omnivorous, whereas the adult females are mostly herbivorous (Chen and Lue, 1999). *Mauremys reevesii* is usually omnivorous but large individuals are strongly conchifrageous (Aoki, 1990). On the other hand, their closest relative (Honda et al., 2002a; Barth et al., 2004; Feldman and Parham, 2004; Spinks et al., 2004), *M. japonica*, is consistently omnivorous (Yasukawa et al., unpublished data). Each species of *Mauremys* is largely semiaquatic (Ernst et al., 2000). In the genus *Heosemys*, *H. annandalii*, as well as *H. grandis* (somewhat omnivorous in juveniles) and *H. spinosa* are almost herbivorous, but *H. depressa* is omnivorous (Iverson and McCord, 1997; Lim and Das, 1999; Ernst et al., 2000; Yasukawa, unpublished data). In addition, *H. annandalii* is highly aquatic, but the other congeneric species are more terrestrial (Iverson and McCord, 1997; Lim and Das, 1999; Ernst et al., 2000). Thus, at present it is difficult to correlate apparent feeding habits or habitats with morphological convergent evolution in these three species. It seems certain, however, that selection dominated by the adaptation of some genetic factor(s) to environmental circumstances facilitated convergent evolution among these three species, resulting in broad triturating surfaces. Further detailed research on morphogenetic mechanisms that could determine the morphology of triturating surfaces may provide an explanation for how convergent evolution occurred in these species.

ACKNOWLEDGMENTS

We thank Mr. K. Honda, Mr. K. Kawaguchi, Mr. H. Koieyama, Mr. M. Udagawa, Mr. E. Chida, and Dr. T. Sato for providing precious tissue samples. This work supported by a Grant-in-Aid to N.O. from the Ministry of Education, Science, Sports and Culture of Japan and from JSPS.

REFERENCES

Aoki, R. 1990. Freshwater turtles of Japan. Biol. Japan 4:60–65. [In Japanese.]

Arnason, U. 2005. Comments on the paper "Phylogeny of mysticete whales based on mitochondrial and nuclear data" by Amanda L. Rychel, Tod W. Reeder, and Annalisa Berta [Mol. Phylogenet. Evol. 32 (2004) 892–901]. Mol. Phylogenet. Evol. 35:309–310.

Barth, D., D. Bernhard, G. Fritzsch, and U. Fritz. 2004. The freshwarter turtle genus *Mauremys* (Testudines, Geoemydidae)—a textbook example of an east-west disjunction or a taxonomic misconcept? Zool. Scr. 33:213–221.

Bashir, A., C. Ye, A. L. Price, and V. Bafna. 2006. Orthologous repeats and mammalian phylogentic inference. Genome Res. 15:998–1006.

Chen, T.-H., and K.-Y. Lue. 1999. Food habits of the Chinese stripe-necked turtle, *Ocadia sinensis*, in the Keelung River, northern Taiwan. J. Herpetol. 33:463–471.

Claude, J., P. Pritchard, H. Tong, E. Paradis, and J.-C. Auffray. 2004. Ecological correlates and evolutionary divergence in the skull of turtles: a geometric morphometric assessment. Syst. Biol. 53:933–948.

Deininger, P. L., and M. A. Batzer. 1993. Evolution of retroposons. Pages 157–196 *in* Evolutionary Biology, volume 27 (M. K. Hecht, eds.). Plenum Press, New York.

Deininger, P. L., D. J. Jolly, C. M. Rubin, T. Friedmann, and C. W. Schmid. 1981. Base sequences studies of 300 nucleotide renatured repeated human DNA clones. J. Mol. Biol. 151:17–33.

Diesmos, A. C., J. F. Parham, B. L. Stuart, and R. M. Brown. 2005. The phylogenetic position of the recently rediscovered Philippine forest turtle (Batagurinae: *Heosemys leytensis*). Proc. Calif. Acad. Sci. 56:31–41.

Endoh, H., S. Nagahashi, and N. Okada. 1990. A highly repetitive and transcribable sequence in the tortoise genome is probably a retroposon. Eur. J. Biochem. 189:25–31.

Endoh, H., and N. Okada. 1986. Total DNA transcription *in vitro*: A procedure to detect highly repetitive and transcribable sequences with tRNA-like structures. Proc. Natl. Acad. Sci. USA 83:251–255.

Ernst C. H., R. G. M. Altenburg, and R. W. Barbour. 2000. Turtles of the World (CD-ROM). Macintosh version 1.2. Editions ETI, Springer, Berlin.

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

Gaffney, E. S., and P. A. Meylan. 1988. A phylogeny of turtles. Pages 157–219 *in* The phylogeny and classification of the tetrapods, volume 1 (M. J. Benton, ed.). Oxford University Press, New York.

Hirayama, R. 1984. Cladistic analysis of batagurine turtles (Batagurinae: Emydidae: Testudinoidea); a preliminary results. Pages 141–157 *in* Stvdia geologica salamanticensia, volume 1 especial 1 (F. de Broin and E. Jiménez-Fuentes, eds.). Ediciones Universidadde Salamanca.

Honda, M., Y. Yasukawa, R. Hirayama, and H. Ota. 2002a. Phylogenetic relationships of the Asian box turtles of the genus *Cuora* sensu lato (Reptilia: Bataguridae) inferred from mitochondrial DNA sequences. Zool. Sci. 19:1305–1312.

Honda, M., Y. Yasukawa, and H. Ota. 2002b. Phylogeny of the Eurasian freshwater turtles of the genus *Mauremys* Gray, 1869 (Testudines), with special reference to a close affinity of *M. japonica* with *Chinemys reevesii*. J. Zool. Syst. Evol. Res. 40:195–200.

Hutchison, J. H. 1998. Turtles across the Paleocene/Eocene epoch boundary in west-central North America. Pages 401–408 *in* Late Paleocene-early Eocene climatic and biotic events in the marine and terrestrial records (M. P. Aubry, S. G. Lucas, and W. A. Berggren, eds.). Columbia University Press, New York.

Iverson, J. B., and W. P. McCord. 1997. Redescription of the Arakan forest turtle *Geoemyda depressa* Anderson 1875 (Testudines: Bataguridae). Chelonian Conserv. Biol. 2:384–389.

Joyce, W. G., J. F. Parham, and J. A. Gauthier. 2004. Developing a protocol for the conversion of rank-based taxon names to phylogenetically defined clade names, as exemplified by turtles. J. Paleontol. 78:989–1013.

Kido, Y., M. Aono, K. Matsumoto, T. Yamaki, M. Saneyoshi, and N. Okada. 1991. Shaping and reshaping of salmonid genomes by amplification of tRNA-derived retroposons during evolution. Proc. Natl. Acad. Sci. USA 88:2326–2330.

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

Lim, L. B., and I. Das. 1999. Turtles of Borneo and Peninsular Malaysia. Natural History Publications, Borneo.

McCord, W. P., J. B. Iverson, P. Q. Spinks, and H. B. Shaffer. 2000. A new genus of geoemydid turtle from Asia. Hamadryad 25:86–90.

McDowell, S. B. 1964. Partition of the genus *Clemmys* and related problems in the taxonomy of the aquatic Testudinidae. Proc. Zool. Soc. Lon. 143:239–279.

Miyamoto, M. M. 1999. Perfect SINEs of evolutionary history? Curr. Biol. 9:R816–R819.

Mlynarski, M. 1969. Fossile Schidkröten. A. Ziemsen Verlag, Wittenberg Lutherstadt.

Murata, S., N. Takasaki, M. Saitoh, and N. Okada. 1993. Determination of the phylogenetic relationships among Pacific salmonids by using short interspersed elements (SINEs) as temporal landmarks of evolution. Proc. Natl. Acad. Sci. USA 90:6995–6999.

Nikaido, M., H. Hamilton, H. Makino, T. Sasaki, K. Takahashi, M. Goto, N. Kanda, L. A. Pastene, and N. Okada. 2006. Baleen whale phylogeny and a past extensive radiation event revealed by SINE insertion analysis. Mol. Biol. Evol. 23:866–873.

Nikaido, M., F. Matsuno, H. Hamilton, R. L. Brownell, Jr., Y. Cao, W. Ding, Z. Zuoyan, A. M. Shedlock, R. E. Fordyce, M. Hasegawa, and N. Okada. 2001. Retroposon analysis of major cetacean lineages: The monophyly of toothed whales and the paraphyly of river dolphins. Proc. Natl. Acad. Sci. USA 98:7384–7389.

Nikaido, M., A. P. Rooney, and N. Okada. 1999. Phylogenetic relationships among cetartiodactyls based on insertions of short and long interspersed elements: Hippopotamuses are the closest extant relatives of whales. Proc. Natl. Acad. Sci. USA 96:10261–10266.

Ohshima, K., M. Hamada, Y. Terai, and N. Okada. 1996. The 3′ ends of tRNA-derived short interspersed repetitive elements are derived from the 3′ ends of long interspersed repetitive elements. Mol. Cell. Biol. 16:3756–3764.

Okada, N. 1991a. SINEs. Curr. Opin. Genet. Dev. 1:498–504.

Okada, N. 1991b. SINEs: Short interspersed repeated elements of the eucaryotic genome. Trends Ecol. Evol. 6:358–361.

Okada, N., M. Hamada, I. Ogiwara, and K. Ohshima. 1997. SINEs and LINEs share common 3′ sequences: A review. Gene 205:229–243.

Okada, N., A. M. Shedlock, and M. Nikaido. 2004. Retroposon mapping in molecular systematics. Pages 189–226 *in* Mobile Genetic Elements (W. J. Miller, and P. Capy, eds.). Humana Press, New Jersey.

Pecon-Slattery, J., A. J. P. Wilkerson, W. J. Murphy, and S. J. O'Brien. 2004. Phylogenetic assessment of introns and SINEs within the Y chromosome using the cat family felidae as a species tree. Mol. Biol. Evol. 21:2299–2309.

Piskurek, O., C. C. Austin, and N. Okada. 2006. Sauria SINEs: Novel short interspersed retroposable elements that are widespread in reptile genomes. J. Mol. Evol. 62:630–644.

Ray, D. A., J. Xing, D. J. Hedges, M. A. Hall, M. E. Laborde, B. A. Anders, B. R. White, N. Stoilova, J. D. Fowlkes, K. E. Landry, L. G. Chemnick, O. A. Ryder, and M. A. Batzer. 2005. *Alu* insertion loci and platyrrhine primate phylogeny. Mol. Phylogenet. Evol. 35:117–126.

Ray, D. A., J. Xing, A. H. Salem, and M. A. Batzer. 2006. SINEs of a nearly perfect character: Applications of the SINE method to phylogeny and population biology. Syst. Biol. 55(6).

Rokas, A., and P. W. Holland. 2000. Rare genomic changes as a tool for phylogenetics. Trends Ecol. Evol. 15:454–459.

Roos, C., J. Schmitz, and H. Zischler. 2004. Primate jumping genes elucidate strepsirrhine phylogeny. Proc. Natl. Acad. Sci. USA 101:10650–10654.

Rychel, A. L., T. W. Reeder, and A. Berta. 2004. Phylogeny of mysticete whales based on mitochondrial and nuclear data. Mol. Phylogenet. Evol. 32:892–901.

Rychel, A.L., T. W. Reeder, and A. Berta. 2005. Response to Ulfur Arnason: Where's the beef? Mol. Phylogenet. Evol. 35:311–312.

Salem, A. H., D. A. Ray, J. Xing, P. A. Callinan, J. S. Myers, D. J. Hedges, R. K. Garber, D. J. Witherspoon, L. B. Jorde, and M. A. Batzer. 2003. *Alu* elements and hominid phylogenetics. Proc. Natl. Acad. Sci. USA 100:12787–12791.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: Laboratory manual, 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sasaki, T., M. Nikaido, H. Hamilton, M. Goto, H. Kato, N. Kanda, L. A. Pastene, Y. Cao, R. E. Fordyce, M. Hasegawa, and N. Okada. 2005. Mitochondrial phylogenetics and evolution of Mysticete whales. Syst. Biol. 54:77–90.

Sasaki, T., M. Nikaido, S. Wada, T. K. Yamada, Y. Cao, M. Hasegawa, and N. Okada. 2006. *Balaenoptera omurai* is a newly discovered baleen whale that represents an ancient evolutionary lineage. Mol. Phylogenet. Evol. 41:40–52.

Sasaki, T., K. Takahashi, M. Nikaido, S. Miura, Y. Yasukawa, and N. Okada. 2004. First application of the SINE (short interspersed repetitive element) method to infer phylogenetic relationships in reptiles: An example from the turtle superfamily Testudinoidea. Mol. Biol. Evol. 21:705–715.

Shaffer, H. B., P. Meylan, and M. L. McKnight. 1997. Tests of turtle phylogeny: Molecular, morphological, and paleontological approaches. Syst. Biol. 46:235–268.

Shedlock, A. M., M. C. Milinkovitch, and N. Okada. 2000. SINE evolution, missing data, and the origin of whales. Syst. Biol. 49:808–817.

Shedlock, A. M., and N. Okada. 2000. SINE insertions: Powerful tools for molecular systematics. BioEssays 22:148–160.

Shedlock, A. M., K. Takahashi, and N. Okada. 2004. SINEs of speciation: Tracking lineages with retroposons. Trends Ecol. Evol. 19:545–553.

Shimamura, M., H. Yasue, K. Ohshima, H. Abe, H. Kato, T. Kishiro, M. Goto, I. Munechika, and N. Okada. 1997. Molecular evidence from retroposons that whales form a clade within even-toed ungulates. Nature 388:666–670.

Singer, M. F. 1982. SINEs and LINEs: Highly repeated short and long interspersed sequences in mammalian genomes. Cell 28:433–434.

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

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

Swofford, D.L. 2001. PAUP*: Phylogenetic analysis using parsimony, version 4. Sinauer Associates, Sunderland, Massachusetts.

Takahashi, K., M. Nishida, M. Yuma, and N. Okada. 2001a. Retroposition of the AFC family of SINEs (short interspersed repetitive elements) before and during adaptive radiation of cichlid fishes in lake Malawi and related inferences about phylogeny. J. Mol. Evol. 53:496–507.

Takahashi, K., Y. Terai, M. Nishida, and N. Okada. 2001b. Phylogenetic relationships and ancient incomplete lineage sorting among cichlid fishes in lake Tanganyika as revealed by analysis of the insertion of retroposons. Mol. Biol. Evol. 18:2057–2066.

Terai, Y., K. Takahashi, M. Nishida, T. Sato, and N. Okada. 2003. Using SINEs to probe ancient explosive speciation: "hidden" radiation of African cichlids? Mol. Biol. Evol. 20:924–930.

van de Lagemaat, L. N., L. Gagnier, P. Medstrand, and D. Mager. 2005. Genomic deletions and precise removal of transposable elements mediated by short identical DNA segments in primates. Genome Res. 15:1243–1249.

Waddell, P. J., H. Kishino, and R. Ota. 2001. A phylogenetic foundation for comparative mammalian genomics. Genome Inform. Ser. Workshop Genome Inform. 12:141–154.

Weiner, A. M., P. L. Deininger, and A. Efstratiadis. 1986. Nonviral retroposons: Genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. Ann. Rev. Biochem. 55:631–661.

Wu, P., K.-Y. Zhou, and Q. Yang. 1998. Evolution of hinged turtles in Bataguridae inferred from sequences of 12S rRNA gene. Chin. J. Appl. Environ. Biol. 4:374–378.

Xing, J., H. Wang, K. Han, D. A. Ray, C. H. Huang, L. G. Chemnick, C.B. Stewart, T. R. Disotell, O. A. Ryder, and M. A. Batzer. 2005. A mobile element based phylogeny of Old World monkeys. Mol. Phylogenet. Evol. 37:872–880.

Yasukawa, Y., R. Hirayama, and T. Hikida. 2001.Phylogenetic relationships of geoemydine turtles (Reptilia: Bataguridae). Curr. Herpetol. 20:105–133.

Received First submitted 7 January 2006; reviews returned 1 March 2006; final acceptance 13 June 2006
Associate Editor: Rod Page

APPENDIX 1. Primer sequences for amplification of each SINE-inserted locus.

Locus name	Forward primer[a]	Reverse primer	Annealing temperature
BHa25	GTTATTTGCAAACTTTTGGA	GCTTCCTTGACTTAGACAAA	47
TKe30	TGCCTCCTTTAAGTGCAGTT	AATTCACAAAAATGCCAGAG	54
TKe40	ATCACAGGTTTCCCCTCTTA	TGGCTAAAGACTTGGATTGA	55
TTh26	TTCATGATGTTTCACTCACC	TAATGCCATGAAGTTCTTCC	52
BOs240	AGGTGACCCAGTGAGTCAGT	GTGGCTCCTTGAAGACAAAC	57
BCr01	GTTCAATCCAGCCAAGTTTT	AAAAGCACCCTCATGAACTT	53
BCr06	GATTCTTCCCGGAGTTTCTA	GTGTTTTCACTGAGGGGTTA	54
TKe17	TGTTTCATCCCAGTTCTCCT	ATTATCCCAGTGCAATCCTG	55
TKe36	AATGTGTTGGATGCCGTAAT	TAAACTCCAGGAATTCCATG	54
TKe83	AGGGGTTGCTACCTGTTTAA	TGTGGGAATACCAGCAATTA	55
BHa65	CATGGAGAGAGTTTGTTGTCAAGTA	AACAGAAACCCTAAAGCAAT	51
BHa74	AGGTTGGAAAGTTACATTTGTACTG	TTCAAATGTTCTTCCTGTAAAAAAC	56
BKs36	TGCATGAATTTCTAAGCAAA	CACCAAAAGTCAATTGTGAA	53
BKs52	GAAGAGGTAGTGGGGAATTA	GCCACCAGAAATAACAAGAT	54
BKs29	CAGTTAAAGCATTTGAAAGGT	CCCCAACATTACCTGAGATA	49
BOs55	ATGGCTGCACATGTGGAGTA	TATTTAACCCCATGCACACC	55
BOs74	CTTGAAGGTCTGTTCGCATT	GCAGTGAAGGGCTCCTAAAT	55
BCr61	GCTCCTAATATGTCAGGGAT	GCTGCGTACAGTAATTTTCA	54
BMm105	TAGTGCCAGTGCAGGTAAAT	AGAACTTTGCACCATTCAGA	52
BPm30	GTTTATAACCATTTGAGGCA	AACCAGAGGCTTCTAGAGTGACC	51
BHg33	GTTTGGATGCCATCTATCTC	GTGGATTAAGCACAATAGCTCTGAA	52
BHg36	ATACAATGCCGTCAGACTGT	AGTATCTGGGTTTGGATTGT	53
BHa59	GGAAAAGAAACCTAGAAGCT	GGGGACCAGATTTTTTTAAT	51
BMr20	GACTGGGGAAATCCTTAGAG	TCTACACTAATTTTGTGCCC	53
BMr49	AACATTAGGCATCCCCAATAAGC	CTAATTGTCTCTTCCGAGCAACC	52
BOs68	CCTGTGCCAAATTCATCTAT	TACTTCAACATTGTCCCTCC	54
BOs151	TGGCGTTTCTAGAAGTTCTAAG	ACCCAAAATTGTGACAAATA	50
BOs60	CTCCATTATTACGGTACCCA	CCAGTGACACTCAGCTCTTG	55
BOs94	TCTCTTTCATTTTTCCATCCA	CACAAATCCAGCAGCCTAGT	50
BMm49	TGATGACTGCATCCATAACA	TTCAAAGAAGATAGCCCCAT	52
BMm103	TAGTGAGTGCCAACAAGTCC	CTTGCTGAGTTGTTTCCATT	52
BMm74	CCATTCAGGTCAGCTAATAT	AGCTTCAAAAGAAAAACATC	53
BMm85	GTTAGTAGATGGCCAGTGGT	ACTGCTCTTTGTACCCTCAG	53
BHg10	TCTTCCTGTGTGTTTGTACAAC	CAGCACATATGCACTCCAGAATTA	53
BHg23	GAAGGGCACCATATTTTAGG	CTAAGCCTAGACTTTGGGTC	55
BHa68	TCTTCCTTGGCATAGACTCATAGA	ACTCTGTGTTGTGGGCTATT	53
BHg12	CTTGTGGTTGGAGTGTCATC	ACAGAGATGGCAATGTCAGA	55
BHg16	TTGCTTCCTGCAGATATATAA	ATTACCTGTAAATTCCATGTGC	53
BOs57	CCAATGACAATCTGCAAAAA	CAAATACGAAAGGGTTGACA	52
BKs11	TGGAAAATGTGGTGCAGTAA	CTAATGTTTTTTCATCCACC	52
BKs85	TCCCCAACTGCAGATTTAAT	CAGAAAAAGGGCTGAAGATC	53
TTh23	AAACAGCTCTTTCAAATTTTC	TTAATTAGGACCATTTGAATG	50
BKs31	GTGCATTTTACACATGGGTA	TTTGTTCCTTTAGTGGCACA	51
BOs33	TTTTCAGATTTATTTGGGGG	TCATCAGGTGTGAGTGAAGC	54
BOs52	AGGTGAGAGGATGGAGAAAA	CATGGATCAGTGTTCACGTT	55
BOs80	TGAAAAGATAAAACCTGCAAGA	GTTGCCAGGAACTAATGAAA	50
BCr04	AGCTTTTCCCCAGATCAC	ACAGTTAATGGTCCACACCT	54
BPm20	TTTTTCAATTAAAGTGCCTG	GGCTTGTAAGCTATAGGTTGCAGTT	51
BMr15	GCATGATTTGCTTAAGAAAA	CCATTACACTCTGTTTGCAG	50

[a] The forward and reverse primers were designed to anneal to flanking sequences upstream and downstream, respectively, of SINEs.

APPENDIX 2. Presence (1) or absence (0) of a SINE unit at each locus.

		1[a]	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
1	BHa25	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	?	1	1	1	1	1	1	1	1	0	0
2	TKe30	1	1	1	1	1	1	1	?	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
3	TKe40	1	1	1	1	1	1	1	1	1	1	1	1	1	1	?	1	1	1	1	1	1	1	1	1	1	1	0	0
4	TTh26	1	?	1	1	1	1	1	1	1	?	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
5	BOs240	1	1	1	1	1	1	1	1	1	1	?	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
6	BCr01	?	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	?	1	1	1	1	?	1	1	1	0	0
7	BCr06	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
8	TKe17	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	?	0	0	0	0	0	0	0	0	0	0	0	0
9	TKe36	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	?	?	0	0
10	TKe83	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	?	0	0	0	0
11	BHa65	0	?	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
12	BHa74	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
13	BKs36	0	0	0	0	?	1	1	1	1	1	0	?	0	0	0	0	0	0	0	0	0	0	0	0	0	0	?	?
14	BKs52	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	BKs29	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	BOs55	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
17	BOs74	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
18	BCr61	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	?	0	0	0	0	?	0	0	0	0	0
19	BMm105	0	0	0	0	?	?	0	0	0	0	0	1	1	1	1	1	0	0	0	0	?	0	0	0	0	0	0	0
20	BPm30	0	0	0	?	0	?	?	?	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
21	BHg33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0
22	BHg36	?	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0
23	BHa59	0	0	0	0	0	0	0	0	0	?	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0
24	BMr20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0
25	BMr49	0	0	0	0	0	0	0	0	0	0	?	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0
26	BOs68	0	0	0	0	0	0	0	?	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	?
27	BOs151	0	?	?	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	BOs60	0	0	?	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	?	0	0	0	0	0	?	?
29	BOs94	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	BMm49	0	0	0	0	0	0	0	0	0	0	?	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
31	BMm103	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	?	0	0	0	0	0	0	0	0	0	0
32	BMm74	0	0	0	0	?	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	?	0	0	0
33	BMm85	0	0	0	0	0	0	0	0	0	0	0	?	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
34	BHg10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	?	1	1	1	1	0	0	0	0	0	0
35	BHg23	0	0	0	0	0	0	?	?	?	0	0	0	?	0	0	0	0	0	1	1	1	1	0	0	?	?	0	0
36	BHa68	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	?	0	0	?
37	BHg12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	?
38	BHg16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
39	BOs57	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
40	BKs11	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
41	BKs85	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	?	0	0	0
42	TTh23	1	1	1	1	1	1	1	1	1	?	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
43	BKs31	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
44	BOs33	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
45	BOs52	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
46	BOs80	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
47	BCr04	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
48	BPm20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
49	BMr15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0

[a] Numbers assigned to each column correspond to species shown in Figure 2. ?: no data, because of a lack of amplification by PCR.