PCR reactions for bait library preparation of the
in-solution hybridization capture approach

Sample GK35 (WAM TR1429) was used for preparation of the individual bait libraries. Long-range PCRs for the mitochondrial baits were performed in 50 μl volumes, containing 1 μl of high-molecular DNA extract and 1 unit of TaKaRa LA Taq DNA Polymerase, Hot-Start Version (Clontech Laboratories), according to the reaction mixture recommended by the manufacturer. PCR conditions comprised an initial denaturation at 93°C for 3 min, followed by 35 cycles of 93°C for 15 sec, 60 or 63°C for 30 sec, 68°C for 10 min and a final elongation step at 68°C for 20 min. For primer sequences, annealing temperatures, and fragment lengths see Table S2. PCR products were visualized on a 1% agarose gel. The combined long-range PCR products covered most of the mitochondrial genome from tRNA-Phe (situated before 12S) to the 5´-half of the control region, missing out approximately 350 bp. PCR conditions for the baits for the ten nuclear loci (AHR, BMP2, HMGB2, HNF1A, ODC, PAX1P1, PSMC1, RAG-1, TB01, TB73) comprised an initial denaturation at 95°C for 5 min, followed by 35–40 cycles of 95°C for 30 sec, 57 to 65°C for 30 sec, 72°C for 1 min 30 sec, and a final elongation step at 72°C for 10 min. For primer sequences, annealing temperatures and fragment lengths see Table S3. PCR reactions were performed in 20 μl volumes, containing 1 μl of high molecular DNA extract and 1 unit of Bioron DFS-Taq polymerase (Bioron GmbH), 2 μl PCR buffer 10× incl. MgCl2 (25 mM), and ultrapure H2O. PCR products were visualized on a 1% agarose gel and, if necessary, excised from a 2% agarose gel and purified using the NucleoSpin Gel and PCR Clean-up Kit (Macherey-Nagel GmbH).

Selection of nuclear loci

Thomson et al. (2021) examined sequence data of 15 nuclear loci for most turtle genera. This dataset was downloaded from GenBank and screened for variable sites within the genera *Emydura* and *Chelodina*. The nine loci with the largest number of variable sites were selected for the current study (Table S4). The ODC locus, known to be informative for resolving relationships of turtles (Fritz et al. 2012, 2023, Praschag et al. 2017) was added.

Details on the mitochondrial alignment and phylogenetic analyses

The 20 new near-complete mitogenomes of *Emydura* were aligned with 15 previously published sequences in an alignment of 15,851 bp length. *Elseya flaviventralis* (KY776454) was added as the outgroup. The alignment was adjusted manually and cropped to its final length. Each protein-coding gene was screened for the presence of internal stop codons using MEGA X (Kumar et al. 2018). Problematic sequence features were deleted prior to phylogenetic analyses: (1) Stop codons of coding genes, as these do not code for any amino acid; (2) gene overlap between coding regions and between tRNAs, as these short regions cannot be attributed to a single locus and may underlie a distinct evolutionary model; (3) alignment positions that cause frame shifts in coding regions; and (4) non-coding spacer DNA. In total, 242 positions were removed.

Tables S5–S7 provide the annotation of the mitochondrial alignment and the best substitution models proposed by PartitionFinder2 (Lanfear et al. 2016) for both phylogenetic analyses.

Details on the nuclear alignment for the SplitsTree analyses

The concatenated nuclear alignment of the ten nuclear loci comprised 8,799 bp with 26 *Emydura* spp. samples (10 from GenBank; 16 newly generated). For the lengths of the individual loci, see Table S1.

Before running SplitsTree4 4.19.0 (Bryant & Huson 2023), each sequence was phased using DnaSP 6.12 (Rozas et al. 2017) and default parameters. The SplitsTree analyses were executed with default parameters except for turning on the “ignore ambiguous states” option.

Additional references

Kumar, S., G. Stecher, C. Knyaz & K. Tamura (2018): MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Molecular Biology and Evolution, **35**: 1547–1549.

Lanfear, R., P. B. Frandsen, A. M. Wright, T. Senfeld & B. Calcott (2016): PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Molecular Biology and Evolution, **34**: 772–773.

**Table S1.** Studied material and used GenBank sequences and their accession numbers. Museum acronyms are: MSNG – Museo Civico di Storia Naturale di Genova; NHM(UK) – Natural History Museum (formerly British Museum of Natural History, BMNH), London; WAM R – Western Australian Museum, Perth (Herpetological Collection); WAM TR – Western Australian Museum, Perth (Herpetological Tissue Collection).

🡪 Separate Excel spreadsheet.

**Table S2.** Fragment lengths, primer sequences, number of PCR cycles, and annealing temperatures of long-range PCR reactions.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Long-range PCR** | **Length of PCR product excl. primers** | **Overlap** | **5´- For -3´** | **5´- Rev -3´** | **PCR cycles** | **Annealing temperature** |  |
| part 1 | 8,641 bp | 203 bp | ATGGCACTGAAGCTGCCAAGATG | ATGGGCTTTGGTTAACTATGTGG | 35 | 63°C |  |
| part 2 | 7,643 bp | ATTACAGCAAACYTAACAGCAGG | TGAACGTARGTCCAGTCTAATAG | 35 | 60°C |  |

**Table S3.** Fragment lengths, primer sequences, number of PCR cycles, and annealing temperatures of PCR reactions for nuclear loci.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Locus** | **Length of PCR product excl. primers** | **Overlap** | **5´- For -3´** | **5´- Rev -3´** |  **PCR cycles** | **Annealing temperature** |  |
| AHR | 539 bp |  | CARGATGAGTCTRTKTATCTCT | GYRAACATSCCATTRACTTGCAT | 40 | 60°C |  |
| BMP2 | 530 bp |  | CTAAGGACCCTGCCACAAGAC | TCAAGGTAGAGCATGGAAATAGC | 35 | 65°C |  |
| HMGB2 | 510 bp |  | GAAATGTGGTCTGARCAGTC | GGCRCGRTATGCWGCAATATC | 40 | 60°C |  |
| HNF1A | 683 bp |  | GTTAGACAAGTCCATGACACC | ACTCAGGATCTCCTACTATGG | 35 | 62°C |  |
| ODC | 750 bp |  | GACTCCAAAGCAGTTTGTCGTCTCAGTGT | TCTTCAGAGCCAGGGAAGCCACCACCAAT | 40 | 60°C |  |
| PAX1P1 | 847 bp |  | CCCTCAGACACTGGATTAYGAATCAT | CCAAGGATTCCGAAGCAGTAAG | 40 | 57°C |  |
| PSMC1 | 770 bp |  | GGAGTCCTGATGGATGACAC | GGTGGCTTTATACCCATCTC | 40 | 60°C |  |
| Rag1 part 1 | 1,663 bp | 42 bp | TGAAAAGGCACCTTCTGGAGA | TCACAAGACTCCTTCACCACCA | 40 | 60°C |  |
| Rag1 part 2 | 647 bp | GAAGACATCTTGGAAGGCATGA | TGCATTGCCAATGTCACAGTG | 40 | 58°C |  |
| TB01 | 679 bp |  | CCGGGCTGACATATACGAAA | GACTGCTGTTGGGTTGTTGA | 40 | 57°C |  |
| TB73 | 579 bp |  | GTTAAAGMAGGGAATCCACTGC | TAAARCAAACCTAACCCTGGC | 35 | 65°C |  |

**Table S4.** Number of informative sites within the genera *Emydura* and *Chelodina* based on the dataset of Thomson et al. (2021). The loci highlighted in green were selected for the current study. Obtained R35 sequences were of substandard quality, which is why these sequences were discarded.

|  |  |  |
| --- | --- | --- |
| **Locus** | ***Chelodina*** | ***Emydura*** |
| AHR | 14 | 10 |
| BDNF | 3 | 1 |
| BMP2 | 7 | 6 |
| HMGB2 | 25 | 10 |
| HNF1A | 29 | 13 |
| NGF | 10 | 2 |
| PAX1P1 | 27 | 21 |
| PSMC1 | 25 | 16 |
| R35 | 42 | 8 |
| Rag1 | 22 | 8 |
| TB01 | 18 | 1 |
| TB29 | 18 | No data |
| TB73 | 35 | 18 |
| TMPPE | 5 | 6 |
| ZFHX1B | 11 | 2 |

**Table S5.** Data block used for PartitionFinder2 (BIC & greedy search schemes), including the annotation of the alignment.

Gene1 tRNA = 1–22;

Gene2 12S = 23–1002;

Gene3 tRNA = 1003–1073;

Gene4 16S = 1074–2698;

Gene5 tRNA = 2699–2774;

Gene6 ND1 pos1 = 2775–3734\3;

Gene6 ND1 pos2 = 2776–3734\3;

Gene6 ND1 pos3 = 2777–3734\3;

Gene7 tRNA = 3735–3943;

Gene8 ND2 pos1 = 3944–4984\3;

Gene8 ND2 pos2 = 3945–4984\3;

Gene8 ND2 pos3 = 3946–4984\3;

Gene9 tRNA = 4985–5339;

Gene10 coxI pos1 = 5340–6875\3;

Gene10 coxI pos2 = 5341–6875\3;

Gene10 coxI pos3 = 5342–6875\3;

Gene11 tRNA = 6876–7015;

Gene12 coxII pos1 = 7016–7693\3;

Gene12 coxII pos2 = 7017–7693\3;

Gene12 coxII pos3 = 7018–7693\3;

Gene13 tRNA = 7694–7766;

Gene14 atp8 pos1 = 7767–7916\3;

Gene14 atp8 pos2 = 7768–7916\3;

Gene14 atp8 pos3 = 7769–7916\3;

Gene15 atp6 pos1 = 7917–8585\3;

Gene15 atp6 pos2 = 7918–8585\3;

Gene15 atp6 pos3 = 7919–8585\3;

Gene16 coxIII pos1 = 8586–9362\3;

Gene16 coxIII pos2 = 8587–9362\3;

Gene16 coxIII pos3 = 8588–9362\3;

Gene17 tRNA = 9363–9431;

Gene18 ND3 pos1 = 9432–9780\3;

Gene18 ND3 pos2 = 9433–9780\3;

Gene18 ND3 pos3 = 9434–9780\3;

Gene19 tRNA = 9781–9849;

Gene20 ND4L pos1 = 9850–10137\3;

Gene20 ND4L pos2 = 9851–10137\3;

Gene20 ND4L pos3 = 9852–10137\3;

Gene21 ND4 pos1 = 10138–11508\3;

Gene21 ND4 pos2 = 10139–11508\3;

Gene21 ND4 pos3 = 10140–11508\3;

Gene22 tRNA = 11509–11712;

Gene23 ND5 pos1 = 11713–13482\3;

Gene23 ND5 pos2 = 11714–13482\3;

Gene23 ND5 pos3 = 11715–13482\3;

Gene24 ND6 pos3 = 13483–14007\3;

Gene24 ND6 pos2 = 13484–14007\3;

Gene24 ND6 pos1 = 13485–14007\3;

Gene25 tRNA = 14008–14076;

Gene26 cytb pos1 = 14077–15204\3;

Gene26 cytb pos2 = 14078–15204\3;

Gene26 cytb pos3 = 14079–15204\3;

Gene27 tRNA = 15205–15343;

Gene28 dloop = 15344–15851;

**Table S6.** Best substitution models for RAxML analysis as proposed by PartitionFinder2.

**Subset Best Model Partition names**

1 GTR+I+G Gene24 ND6 pos2, Gene4 16S, Gene2 12S, Gene18 ND3 pos3, Gene14 atp8 pos1, Gene1 tRNA, Gene18 ND3 pos1, Gene27 tRNA, Gene23 ND5 pos1, Gene21 ND4 pos1, Gene8 ND2 pos1, Gene24 ND6 pos1, Gene15 atp6 pos1

2 GTR+G Gene11 tRNA, Gene10 coxI pos1, Gene20 ND4L pos1, Gene6 ND1 pos1, Gene3 tRNA, Gene16 coxIII pos1, Gene12 coxII pos1, Gene13 tRNA, Gene7 tRNA, Gene26 cytb pos1, Gene18 ND3 pos2, Gene14 atp8 pos2, Gene22 tRNA, Gene19 tRNA, Gene5 tRNA, Gene9 tRNA, Gene25 tRNA, Gene17 tRNA

3 GTR+G Gene15 atp6 pos2, Gene26 cytb pos2, Gene23 ND5 pos2, Gene8 ND2 pos2, Gene20 ND4L pos2, Gene6 ND1 pos2, Gene21 ND4 pos2, Gene12 coxII pos2, Gene10 coxI pos2, Gene16 coxIII pos2

4 GTR+G Gene15 atp6 pos3, Gene20 ND4L pos3, Gene26 cytb pos3, Gene24 ND6 pos3, Gene16 coxIII pos3, Gene21 ND4 pos3, Gene23 ND5 pos3, Gene14 atp8 pos3, Gene12 coxII pos3, Gene10 coxI pos3, Gene6 ND1 pos3, Gene8 ND2 pos3

5 GTR+I+G Gene28 dloop

**Table S7.** Best substitution models for MrBayes analysis as proposed by PartitionFinder2.

**Subset Best Model Partition names**

1 HKY+G Gene24 ND6 pos1, Gene15 atp6 pos1, Gene14 atp8 pos3, Gene18 ND3 pos3, Gene18 ND3 pos1, Gene14 atp8 pos1, Gene1 tRNA

2 GTR+I+G Gene14 atp8 pos2, Gene24 ND6 pos2, Gene13 tRNA, Gene5 tRNA, Gene26 cytb pos1, Gene25 tRNA, Gene20 ND4L pos1, Gene6 ND1 pos1, Gene3 tRNA, Gene22 tRNA, Gene9 tRNA, Gene8 ND2 pos1, Gene21 ND4 pos1, Gene27 tRNA, Gene23 ND5 pos1, Gene2 12S, Gene4 16S, Gene17 tRNA, Gene19 tRNA

3 HKY+I Gene15 atp6 pos2, Gene26 cytb pos2, Gene20 ND4L pos2, Gene23 ND5 pos2, Gene8 ND2 pos2, Gene6 ND1 pos2, Gene21 ND4 pos2, Gene10 coxI pos2, Gene16 coxIII pos2, Gene11 tRNA, Gene12 coxII pos2

4 GTR+G Gene8 ND2 pos3, Gene6 ND1 pos3, Gene26 cytb pos3, Gene24 ND6 pos3, Gene20 ND4L pos3, Gene16 coxIII pos3, Gene15 atp6 pos3, Gene21 ND4 pos3, Gene23 ND5 pos3, Gene10 coxI pos3, Gene12 coxII pos3

5 K80+I Gene18 ND3 pos2, Gene10 coxI pos1, Gene7 tRNA, Gene16 coxIII pos1, Gene12 coxII pos1

6 HKY+I+G Gene28 dloop