Central American *Trachemys* revisited: New sampling questions current understanding of taxonomy and distribution

Uwe Fritz, Christian Kehlmaier, Rodney J. Scott, Raúl Fournier,
James R. McCranie, Natalia Gallego-García

Supplementary Information

For 40 *Trachemys* samples (Table S1) either the complete mitochondrial genome or the following mitochondrial genes were targeted: 12S (partial), ND4L (complete), ND4 (complete), and cyt *b* (complete plus part of the adjacent tRNA-Thr gene). In addition, the sequences of the partial nuclear loci Cmos, ODC, R35, Rag1, and Rag2 were obtained as far as possible.

Sanger sequencing. For 10 blood and tissue samples stored at -80°C, mtDNA and nDNA sequences were obtained as described in Fritz et al. (2012) and Praschag et al. (2017); all reaction products were purified using SephadexTM G-50 fine (GE Healthcare).

Next Generation Sequencing (NGS) and in-solution hybridization capture. For 17 DNA samples, 12 samples from historic museum specimens, and a sample from a roadkill, DNA sequences were generated by an NGS approach including two rounds of in-solution hybridization capture. The historic material was processed in a cleanroom facility, physically isolated from the main laboratory to avoid contamination by foreign DNA (Fulton and Shapiro 2019). DNA extraction was performed using Qiagen's DNeasy Blood & Tissue Kit according to the manufacturer's protocol, also in 2012 for the 17 samples for which now only a few microliters of DNA were available.

Prior to the preparation of single-indexed, double-stranded DNA libraries according to Meyer and Kircher (2010), the DNA was sheared down to an average length of 150 bp using a Covaris M220 ultrasonicator. DNA concentration and quality of the samples and DNA libraries were assessed using a Qubit 3.0 Fluorometer (Thermo Fisher Scientific) and a 4200 TapeStation system (Agilent).

In order to increase the amount of the targeted loci in all DNA libraries, two-rounds of in-solution hybridization capture (Maricic et al. 2010; Horn 2012) were performed in a dedicated capture-only workspace in the main laboratory using DNA baits generated from modern PCR products. For the mtDNA bait library, two long-range PCR reactions were performed (LR1 and LR2) using a sample of *Trachemys venusta callirostris* (T 4728), yielding amplicons with an overlap of 1136 bp and an individual length of 11,760 bp (LR1) and 6686 bp (LR2). The combined long-range PCR products covered most of the mitochondrial genome from tRNA-Phe (situated before 12S) to the 3'-end of the control region, missing out approximately 200 bp. By aiming at large stretches of mtDNA, the risk of amplifying nuclear copies of mitochondrial DNA (numts), which are an issue in *Trachemys*, is minimized (Bensasson et al. 2001; Fritz et al. 2012). For each long-range PCR, a 50 µl volume was used, containing 1 unit of TaKaRa LA Taq DNA Polymerase, Hot-Start Version (Clontech Laboratories Inc.), and the reaction mixture recommended by the manufacturer. PCR conditions comprised initial denaturation at 93°C for 3 min, followed by 35 cycles of 93°C for 20 sec, 57°C for 30 sec, 68°C for 12 min, and a final elongation step at 68°C for 20 min. For primer sequences see Table S2. PCR products were visualized and excised from a 2% agarose gel and purified using the NucleoSpin Gel and PCR Clean-up Kit (Macherey-Nagel).

After pooling both long-range products at an equimolar rate, the baits were sheared down to 150 bp and converted into the mtDNA bait library. The nDNA bait library was produced from PCR products of *T. v. callirostris* (T 4728) and *T. scripta elegans* (T 12680) obtained as detailed in Fritz et al. (2012) and Praschag et al. (2017); for primer sequences, see Table S2. The PCR products for the five nuclear loci were pooled at an equimolar rate, sheared down to 150 bp and converted into the nDNA bait library. Prior to capturing, the mtDNA and nDNA bait libraries were adjusted to the same molarity and mixed at a ratio of 1:5 to account for the lower numbers of nuclear target molecules in the DNA libraries of the individual samples.

Sequencing was performed in-house on an Illumina MiSeq platform, generating 75 bp-long paired-end reads.

Bioinformatics. NGS sequence data were assembled using the following pipeline: After adapter trimming with Skewer 0.2.2 (Jiang et al. 2014), read merging (minimum length 35 bp), quality filtering (minimum Q-score 20), and duplicate removal with BBmap-suite 37.24 (https://sourceforge.net/projects/bbmap/) (Bushnell et al. 2017), the remaining reads were screened for contamination using FastQScreen 0.11.4 (Wingett and Andrews 2018) and a set of predefined mitochondrial and nuclear sequences (Table S3), including the mitochondrial genome of *T. s. elegans* (GenBank accession number KM216748) and a concatenated sequence of the five nuclear loci of one individual of *T. medemi* (LT883198, LT883245, LT883260, LT883222, LT883234)—the individual loci being separated by stretches of 2000 ambiguous sites (Ns) to prevent mapping artifacts (Fig. S1). The identified non-*Trachemys* reads were excluded from the individual readpools, and the remaining reads were stored as readpool-1. Reads only mapping to the mt-genome of *T. s. elegans* were stored as readpool-2. Reads only mapping to the nuclear loci of *T. medemi* were stored as readpool-3. The mitochondrial genome was assembled with MITObim (Hahn et al. 2013), a two-step baiting and iterative mapping approach, using readpool-2, an allowed mismatch value of 2, and sequence KM216748 (*T. s. elegans*) as a starting seed. For step 1 of the mapping procedure, readpool-2 was reduced to 30,000 randomly selected reads. To double-check doubtful regions of the mitochondrial genomes of the 17 DNA samples, three short fragments of ND5-ND6, cyt *b*, and the control region were Sanger sequenced additionally (see Table S2 for details). The nuclear loci of these samples were also assembled with MITObim, using readpool-3, an allowed mismatch value of 2, and sequences of the above-mentioned *T. medemi* as a starting seed. For step 1 of the mapping procedure, readpool-3 was reduced to 5000 randomly selected reads. For the nuclear assembly of the 13 historic samples a different mapping approach was pursued, using the Burrows-Wheeler Aligner (Li and Durbin 2009) with its Maximal Exact Match algorithm (BWA-MEM), applying a relaxed mismatch threshold of 0.001 (corresponding to approximately eight mismatches in 100 bp), readpool-1 including all non-merged quality-filtered reads due to the low number of available merged reads, and the concatenated nuclear sequences of *T. medemi* as a mapping reference. Resulting scaffolds were visualized and checked for assembly artifacts in Tablet (Milne et al. 2013). Assembly artifacts were manually removed from the assembled contigs and all positions with a coverage below 3-fold masked as ambiguous (N). Sequence length distribution of mapped reads was calculated with a customized awk command and Microsoft Excel. An exemplary sample documentation is provided in Figs S1–S4 for the holotype of *Trachemys venusta uhrigi* (FLMNH 157800). The summarized mapping details are provided in Table S4.

Alignment preparations. Previously published sequence data of 80 *Trachemys* and related taxa (Table S1) were downloaded from GenBank and aligned in eight individual files (12S, ND4L/ND4, cyt *b* plus tRNA-Thr, Cmos, ODC, R35, Rag1, Rag2) using BioEdit 7.0.5.2 (Hall

1999). Sanger sequence data were inspected by eye for base-calling errors and added to the alignments. NGS sequence data were added by extracting the appropriate loci from the mitochondrial and nuclear assemblies. The individual alignments were adjusted manually and cropped to their final lengths. Each protein-coding gene was screened for the presence of internal stop codons using MEGA X (Kumar et al. 2018). For the three mitochondrial alignments, problematic sequence features (stop codons—7 bp, frameshift—1 bp, spacer DNA—5 bp) were removed manually, before the sequences were concatenated for analysis. The mtDNA alignment used for calculations was 3226 bp long (12S: 398 bp; cyt *b*: 1137 bp; tRNA-Thr: 26 bp; ND4L: 294 bp; ND4: 1371 bp).

The alignment of the five nuclear loci had a length of 3409 bp (Cmos: 563 bp; ODC: 621 bp; R35: 974 bp; Rag1: 614 bp; Rag2: 637 bp).

Phylogenetic analyses. Phylogenetic analyses were performed for the mtDNA dataset, applying Maximum Likelihood (ML) and Bayesian Inference (BI) approaches using RAxML 8.0.0 (Stamatakis 2014) and MrBayes 3.2.6 (Ronquist et al. 2012). The best evolutionary models and partitioning schemes (Tables S5, S6) were determined with PartitionFinder2 (Lanfear et al. 2016) applying the greedy search scheme and the Bayesian Information Criterion. For ML, 10 independent searches were carried out using the GTR + G substitution model, different starting conditions, and the rapid bootstrap option. Subsequently, 1000 non-parametric thorough bootstrap replicates were calculated and the values plotted against the best tree. For BI, four parallel runs (each with eight chains) were performed with 2 million generations (burn-in 0.25; print frequency 1000; sample frequency 1000). Calculation parameters were analysed using Tracer 1.7.1 (Rambaut et al. 2018).

References

Bensasson D, Zhang D-X, Hartl DL, Hewitt GM (2001) Mitochondrial pseudogenes: Evolution's misplaced witnesses. Trends in Ecology and Evolution 16: 314–321. https://doi.org/10.1016/s0169-5347(01)02151-6

Bushnell B, Rood J, Singer, E (2017) BBMerge—accurate paired shotgun read merging via overlap. PLoS One 12: e0185056. https://doi.org/10.1371/journal.pone.0185056

Friesen VL, Congdon BC, Kidd MG, Birt TP (1999) Polymerase chain reaction (PCR) primers for the amplification of five nuclear introns in vertebrates. Molecular Ecology 8: 2147–2149. https://doi.org/10.1046/j.1365-294x.1999.00802-4.x

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

Fujita MK, Engstrom TN, Starkey DE, Shaffer HB (2004) Turtle phylogeny: Insights from a novel nuclear intron. Molecular Phylogenetics and Evolution 31: 1031–1040. https://doi.org/10.1016/j.ympev.2003.09.016

Fulton TL, Shapiro B (2019) Setting up an ancient DNA laboratory. In: Shapiro B, Barlow A, Heintzman PD, Hofreiter M, Paijmans JLA, Soares AER (Eds) Ancient DNA: Methods and Protocols. Methods in Molecular Biology, Vol. 1963. Totowa, NJ, Humana Press, 1–13. https://doi.org/10.1007/978-1-4939-9176-1_1

Hahn C, Bachmann L, Chevreux B (2013) Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—A baiting and iterative mapping approach. Nucleic Acids Research 41: 1–9.
https://doi.org/10.1093/nar/gkt371

Hall TA (1999) BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41: 95–98.

Horn S (2012) Target enrichment via DNA hybridization capture. In: Shapiro B, Hofreiter M (Eds) Ancient DNA: Methods and Protocols. Methods in Molecular Biology, Vol. 840. Springer, Berlin, 177–188.
https://doi.org/10.1007/978-1-61779-516-9_21

Jiang H, Lei R, Ding SW, Zhu S (2014) Skewer: A fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinformatics 15: 182.
https://doi.org/10.1186/1471-2105-15-182

Kumar S, Stecher G, Knyaz C, Tamura K (2018) MEGA X: Molecular Evolutionary Genetic Analysis across computing platforms. Molecular Biology and Evolution 35: 1547–1549.
https://doi.org/10.1093/molbev/msy096

Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B (2016) PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Molecular Biology and Evolution 34: 772–773.
https://doi.org/10.1093/molbev/msw260

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

Le M, McCord WP, Iverson JB (2007) On the paraphyly of the genus *Kachuga* (Testudines: Geoemydidae). Molecular Phylogenetics and Evolution 45: 398–404.
https://doi.org/10.1016/j.ympev.2007.05.002

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25: 1754–1760.
https://doi.org/10.1093/bioinformatics/btp324

Maricic T, Whitten M, Pääbo S (2010) Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. PLoS One 5: e14004.
https://doi.org/10.1371/journal.pone.0014004

Meyer M, Kircher M (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harbor Protocols 2010: pdb.prot5448.
https://doi.org/10.1101/pdb.prot5448

Milne I, Stephen G, Bayer M, Cock PJ, Pritchard L, Cardle L, Shaw PD, Marshall D (2013) Using Tablet for visual exploration of second-generation sequencing data. Briefings in Bioinformatics 14: 193-202.
https://doi.org/10.1093/bib/bbs012

Praschag P, Ihlow F, Flecks M, Vamberger M, Fritz U (2017) Diversity of North American map and sawback turtles (Testudines: Emydidae: *Graptemys*). Zoologica Scripta 46: 675–682.
https://doi.org/10.1111/zsc.12249

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA (2018) Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Systematic Biology 5: 901–904. https://doi.org/10.1093/sysbio/syy032

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology 61: 539–542. https://doi.org/10.1093/sysbio/sys029

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

Starkey DE, Shaffer HB, Burke RL, Forstner MR, Iverson JB, Janzen FJ, Rhodin AG, Ultsch GR (2003) Molecular systematics, phylogeography, and the effects of Pleistocene glaciation in the painted turtle (*Chrysemys picta*) complex. Evolution 57: 119–128. https://doi.org/10.1111/j.0014-3820.2003.tb00220.x

Wiens JJ, Kuczynski CA, Stephens PR (2010) Discordant mitochondrial and nuclear gene phylogenies in emydid turtles Implications for speciation and conservation. Biological Journal of the Linnean Society 99: 445–461. https://doi.org/10.1111/j.1095-8312.2009.01342.x

Wingett SW, Andrews S (2018) FastQ Screen: A tool for multi-genome mapping and quality control [version 2; referees: 4 approved]. F1000Research 7: 1338. https://doi.org/10.12688/f1000research.15931.2

Table S1. Samples and DNA sequences used in the present study. → See extra Excel spreadsheet.

Table S2. Number of PCR cycles, annealing temperatures, and PCR primer pairs used for amplicon sequencing and bait-library preparation.

Locus	Primer name	Primer sequence (5'–3')	Amplicon excl. oligos	Number of PCR cycles	Annealing temperature	Excised from gel	Source
Long-range PCR of mtDNA							
LR1	Tscripta_tPhe.for	AGCACGGCACTGAAGTTGCC	11760 bp	35	57°C	Yes	This study
	Tscripta_tLeu_a.rev	GGGTGGATGGTTTCTAAAACCAATGG					This study
LR2	Tscripta_tArg_b.for	CTTAGTTAATCGTGATTAAACTCCACGGC	6686 bp	35	57°C	Yes	Fritz et al. (2012)
	DES2	GGATTTAGGGGTTTGACGAGAAT					Starkey et al. (2003)
PCR of nDNA							
Cmos	Cmos1	GCCTGGTGCTCCATCGACTGGGATCA	590 bp	40	58°C	no	Le et a. (2006)
	Cmos3	GTAGATGTCTGCTTTGGGGGTGA					Le et a. (2006)
ODC	OD-F	GACTCCAAAGCAGTTTGTCGTCTCAGTGT	628 bp	35	60°C	no	Friesen et al. (1999)
	ODC-rvs2	ATTGGTYRTAAGATTTAGTAAGTCT					Wiens et al. (2010)
R35	R35Ex1_F (=R35Ex1)	ACGATTCTCGCTGATTCTTGC	> 1000 bp	37	53°C	no	Fujita et al. (2004)
	R35Ex2_R (=R35Ex2)	GCAGAAAACTGAATGTCTCAAAGG					Fujita et al. (2004)
Rag1	Rag1878	GAAGACATCTTGGAAGGCATGA	647 bp	38	53°C	no	Le et al. (2007)
	Rag2547	TGCATTGCCAATGTCACAGTG					Le et al. (2007)
Rag2	F2-1	TTCCAGAGCTTCAGGATGG	658 bp	40	52°C	yes	Le et a. (2006)
	R2-1	CAGTTGAATAGAAAGGAACCCAAGT					Le et a. (2006)
PCR of mtDNA to improve low-quality regions							
ND5/ND6	Tra_ND5-ND6_check_For	TGACCCAACCAGCATAGCCG	572 bp	35	60°C	no	This study
	Tra_ND5-ND6_check_Rev	TACGGTGTGGAAGGGTTTGG					This study
tRNA-Glu/ cyt *b*	Trach_1.for	GATTTAAGCCGAGACCTGTG	810 bp	33	64°C	no	This study
	Trach_Seq	TGGATAAGGGGTTGGCTGGTGTG					Praschag et al. (2017)
CR	Tra_CR_check_For	TATTCTCCATTTCTCAACCC	588 bp	35	60°C	no	This study
	Tra_CR_check_For	GCTTATTTCTCGTGATTAGG					This study

Table S3. Results of contamination screening using FastQScreen (Wingett et al. 2018) for the obtained quality-filtered reads of sample FLMNH 157800 (holotype of *Trachemys venusta uhrigi*) to assess endogenous DNA content in relation to potential contamination sources. Reported are number of reads that map to a given set of references.

Taxa	Reads processed	Unmapped	One hit / one genome	Multiple hits / one genome	One hit / multiple genomes	Multiple hits / multiple genomes
mt-genome						
Homo	731,587	727,666	37	0	3,884	0
Penicillium	731,587	731,587	0	0	0	0
Bacillus	731,587	731,586	0	1	0	0
Ecoli	731,587	731,580	1	6	0	0
Felis	731,587	727,250	3	0	4,334	0
Ursus	731,587	726,179	21	0	5,387	0
Canis	731,587	727,111	1	0	4,475	0
Sus	731,587	727,171	1	0	4,415	0
Bos	731,587	725,666	11	0	5,910	0
Gallus	731,587	723,845	75	0	7,667	0
Sula	731,587	725,328	130	0	6,129	0
Mus	731,587	726,860	170	0	4,557	0
Cyprinus	731,587	728,381	42	0	3,164	0
Trachemys scripta elegans	731,587	470,389	245,381	276	15,538	3
nuclear loci (Cmos, ODC, R35, Rag1, Rag2)						
Trachemys medemi	731,587	729,601	1,985	1	0	0

Table S4. Mapping details for the samples processed via Next Generation Sequencing. See above paragraphs on Next Generation Sequencing and Bioinformatics for the interpretation of the individual columns.

					mtDNA						nDNA				
Taxon	Sample	DNA (ng) into dsLib prep	Raw reads	Quality filtered reads (merged)	MITObim readpool-2	MITObim assembled reads	Average read length	Average coverage	Length of final contig submitted to ENA	Ambiguous sites in final contig	BWA readpool	MITObim readpool-3	MITObim readpool-1	MITObim assembled reads	Average read length
Next-Generation-sequencing of modern samples															
Trachemys grayi emolli	Tv57	546	940,591	716,680	150,693	142,783	101	859	16,556	140		15,098	703,607	16,358	95
Trachemys grayi emolli	Tv64	143	1,227,620	915,508	174,295	160,700	101	959	16,551	140		18,018	898,507	19,424	95
Trachemys grayi emolli	Tv65	283	1,147,477	878,723	188,490	176,673	100	1,054	16,555	137		17,324	861,802	18,677	93
Trachemys grayi emolli	Tv66	207	1,163,509	881,418	188,990	179,300	100	1,071	16,555	138		17,613	864,138	19,081	94
Trachemys grayi emolli	Tv69	600	1,086,309	849,953	190,869	183,030	100	1,083	16,555	140		19,275	832,513	20,675	93
Trachemys grayi emolli	Te14	247	1,076,789	811,342	171,248	157,501	99	928	16,555	138		14,061	795,346	14,953	91
Trachemys grayi emolli	Te23	573	1,313,816	990,860	204,236	196,912	102	1,192	16,555	140		22,975	972,092	24,899	95
Trachemys grayi emolli	Te35	480	969,578	738,752	143,688	134,196	102	812	16,548	150		15,860	725,057	16,975	95
Trachemys grayi emolli	Te39	780	1,090,619	823,903	167,634	156,927	100	938	16,555	140		21,519	808,546	23,032	92
Trachemys grayi emolli	Te42	215	1,118,947	834,772	142,170	125,496	98	734	16,555	139		17,898	819,242	18,630	91
Trachemys grayi emolli	Tv47	148	788,750	603,156	120,646	114,780	100	684	16,544	140		12,863	591,225	13,836	93
Trachemys grayi emolli	Tv48	255	1,178,758	901,215	176,420	168,131	99	995	16,556	141		20,836	884,383	21,734	91
Trachemys grayi emolli	Tv49	273	1,225,218	841,835	159,247	152,560	102	924	16,547	137		13,848	827,166	15,038	95
Trachemys grayi panamensis	Tv50	231	1,161,892	784,362	168,502	159,933	102	971	16,558	138		10,839	770,345	11,698	96
Trachemys grayi panamensis	Tv51	309	1,110,346	823,774	157,798	147,757	100	884	16,558	137		15,498	808,080	16,868	93
Trachemys grayi panamensis	Tv52	205	1,180,369	890,837	210,327	198,365	100	1,180	16,558	137		15,749	872,565	17,085	93
Trachemys grayi panamensis	Tv53	289	855,863	666,209	148,827	141,128	98	827	16,547	138		10,245	653,073	11,145	91
Next-Generation-sequencing of historic samples															
Trachemys decorata	D 32042	30	833,454	498,788	274,259	269,379	94	1,510	16,558	0	969,424			302	
Trachemys decussata angusta	20622	97	1,690,779	603,906	192,165	188,484	93	1,039	16,568	0	1,880,246			1,629	
Trachemys decussata decussata	D 32041	49	1,180,327	687,567	429,467	421,462	94	2,354	16,560	0	1,333,004			748	
Trachemys stejnegeri vicina	D 31506	98	861,699	436,050	215,849	211,667	96	1,216	16,557	0	1,116,291			2,024	
Trachemys terrapen	D 34289	111	1,099,432	572,941	392,590	382,841	93	2,111	16,553	0	1,206,443			228	
Trachemys terrapen	D 34291	114	1,270,303	593,775	401,452	392,000	96	2,230	16,553	0	1,466,681			262	
Trachemys terrapen	D 34655	104	1,402,502	699,020	415,174	407,685	99	2,392	16,553	5	895,603			790	
Trachemys venusta venusta s. l.	D 41608	>0,3	1,173,403	694,544	319,672	314,701	99	1,841	16,687	51	1,243,646			280	
Trachemys venusta venusta s. l.	FLMNH 157800	>0,3	1,452,062	731,587	223,910	240,018	102	1,450	16,563	55		1,985	712,225	1,951	89
Trachemys venusta venusta s. l.	D 39071	94	1,148,780	504,007	314,605	309,652	106	1,949	16,562	100	1,415,906			474	
Trachemys venusta venusta s. l.	D 39077	116	1,212,780	763,127	454,213	447,062	90	2,400	16,561	54	1,222,411			1,077	
Trachemys venusta venusta s. l.	D 39069	44	737,327	508,915	291,325	287,008	96	1,633	16,672	55	776,799			512	
Trachemys venusta venusta s. l.	D 39078	18	1,341,760	714,591	389,660	383,301	105	2,381	16,675	140	1,309,414			1,227	

Table S5. The best evolutionary models and partitioning schemes for the mt-genome dataset as determined by PartitionFinder2 applying the greedy search scheme and the Bayesian Information Criterion.

RAxML		
Sub-set	**Best mod-el**	**Partition names**
1	GTR+I+G	Block 3 tRNA, Block 5 ND4 pos1, Block 4 ND4L pos1, Block 2 cytb pos1, Block 1 12S
2	GTR+I+G	Block 2 cytb pos2, Block 5 ND4 pos2, Block 4 ND4L pos2
3	GTR+G	Block 2 cytb pos3, Block 5 ND4 pos3, Block 4 ND4L pos3
MrBayes		
Sub-set	**Best mod-el**	**Partition names**
1	GTR+I+G	Block 3 tRNA, Block 5 ND4 pos1, Block 4 ND4L pos1, Block 1 12S, Block 2 cytb pos1
2	HKY+I	Block 2 cytb pos2, Block 4 ND4L pos2, Block 5 ND4 pos2
3	GTR+G	Block 4 ND4L pos3, Block 2 cytb pos3, Block 5 ND4 pos3

Table S6. Data blocks of the mtDNA alignment used for phylogenetic analyses.

Block 1 12S	= 1–398;
Block 2 cytb pos1	= 399–1535\3;
Block 2 cytb pos2	= 400–1535\3;
Block 2 cytb pos3	= 401–1535\3;
Block 3 tRNA	= 1536–1561;
Block 4 ND4L pos1	= 1562–1855\3;
Block 4 ND4L pos2	= 1563–1855\3;
Block 4 ND4L pos3	= 1564–1855\3;
Block 5 ND4 pos1	= 1856–3226\3;
Block 5 ND4 pos2	= 1857–3226\3;
Block 5 ND4 pos3	= 1858–3226\3;

Figure S1. Figured is the scaled assembly for the nDNA loci of sample FLMNH 157800 (holotype of *Trachemys venusta uhrigi*) as seen in Tablet. Assembled reads: 1951. Maximum coverage depth: 266-fold. Average coverage depth: n/a. Average consensus quality: n/a. Mismatch value: 1.0%.

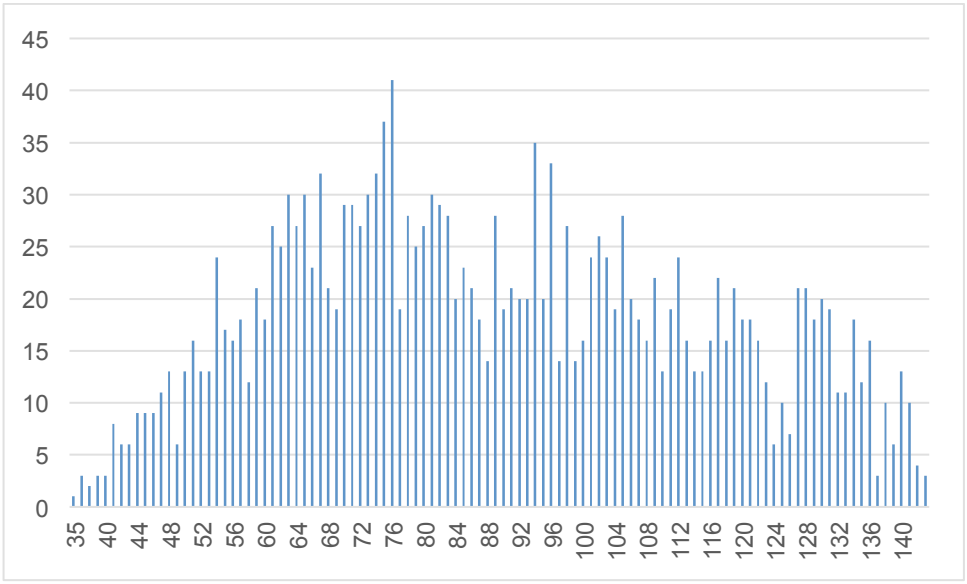

Figure S2. Figured are read lengths of 1951 mapped nuclear reads of sample FLMNH 157800 (holotype of *Trachemys venusta uhrigi*) ranging from 35 bp to 143 bp, with an average read length of 89 bp.

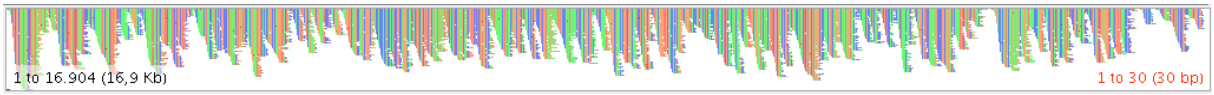

Figure S3. Figured is the scaled assembly for the mt-genome of sample FLMNH 157800 (holotype of *Trachemys venusta uhrigi*) as seen in Tablet. Assembled reads: 240018. Maximum coverage depth: 3133-fold. Average coverage depth: 1450. Average consensus quality: 89 (max. value = 90). Mismatch value: 0.1%.

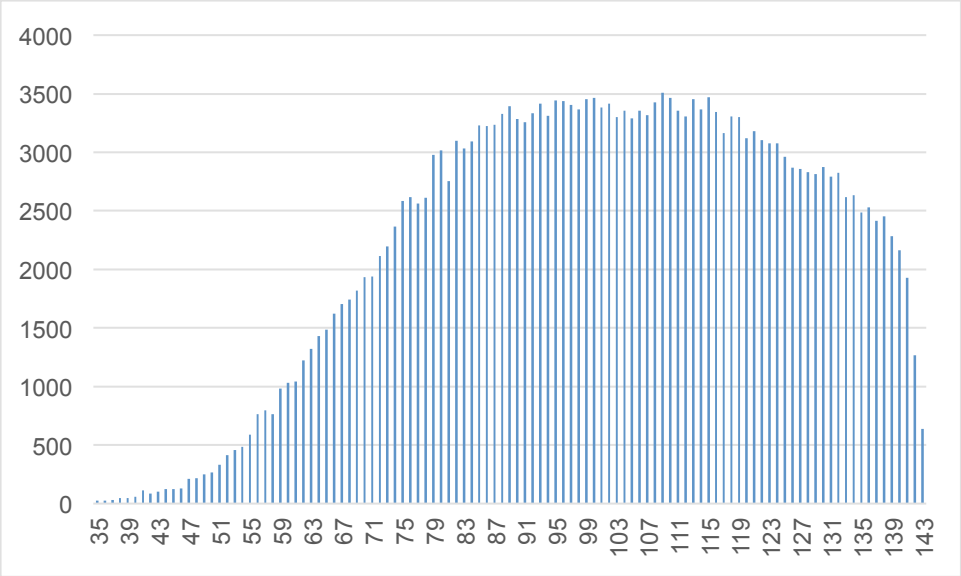

Figure S4. Figured are read lengths of 240018 mapped mitochondrial reads of sample FLMNH 157800 (holotype of *Trachemys venusta uhrigi*) ranging from 35 bp to 143 bp, with an average read length of 102 bp.