

## Overview of the supplementary material provided for the dissertation

“Phylogeny, divergence time estimates and systematics of African cichlids (Cichlidae: Pseudocrenilabrinae), with a focus on the rheophilic cichlids of East and Central Africa”

presented by Frederic D.B. Schedel

Supplementary material provided for **chapter 4**: “East African cichlid lineages (Teleostei: Cichlidae) might be older than their ancient host lakes: new divergence estimates for the east African cichlid radiation”

Table S1. Overview of the taxon sampling with corresponding Genbank accession numbers and, where applicable, corresponding repository numbers of specimens and their origin. (DOCX 43 kb)

Table S2. Comprehensive list of mean divergence ages and their corresponding 95% HPD age ranges of selected nodes. (Node numbers 1 to 65 correspond to numbers depicted in Fig. 2). (DOCX 32 kb)

Table S3. Overview of the taxon sampling for the nuclear markers (RAG1, ENC1, Rh1 and ttna TMO) with corresponding Genbank accession numbers. (DOCX 75 kb)

Additional File 1: Justification for exclusion of several cichlid fossils from calibration. (DOCX 24 kb)

Supplementary Material provided for **chapter 5**: “The ghost of a hybrids past uncovers ancient river tectonic captures”

### Supplementary Material

Figure S1. ML-based phylogenetic hypothesis (RAxML) based on the “AC mitochondrial dataset” including ten protein coding mitochondrial genes (ND1, ND2, COX1, COX2, ATP8, ATP6, COX3, ND3, ND4L, ND4) and 464 specimens with representatives of almost all major ovalentarian lineages with focus on the family Cichlidae (N = 330). Numbers at nodes refer to bootstrap-values. (PDF 34 kb)

Figure S2. ML-based phylogenetic hypothesis (RAxML) based on the “Austrotilapiine mtDNA dataset” including ten protein coding mitochondrial genes (ND1, ND2, COX1, COX2, ATP8, ATP6, COX3, ND3, ND4L, ND4) and 103 specimens. Numbers at nodes refer to bootstrap-values. (PDF 9 kb)

Figure S3. ML-based phylogenetic hypothesis (RAxML) based on the “Austrotilapiine ddRAD Set” including a concatenated alignment of 610 ddRAD loci (113,578 bp) for 103 individuals. Numbers at nodes refer to bootstrap-values. See Table S1 for specimen information of the corresponding tip labels (ddRAD IDs). (PDF 9 kb)

Figure S4. ML-based phylogenetic hypothesis (RAxML) based on the “Introgression-test ddRAD Set” including a concatenated alignment of 14,750 ddRAD loci (2,751,967 bp) for 60

individuals. Numbers at nodes refer to bootstrap-values. See Table S1 for specimen information of the corresponding tip labels (ddRAD IDs). (PDF 6 kb)

Figure S5. Coalescent species tree (ASTRAL) based on the “Full ddRAD Set” and 4000 individual gene (ddRAD loci) trees. Numbers at nodes are support values from local posterior probabilities for the main topology (pp1) and the two alternative topologies (pp2 & pp3). See Table S1 for specimen information of the corresponding tip labels (ddRAD IDs). (PDF 16 kb)

Figure S6. Coalescent species tree (ASTRAL) based on the “Austrotilapiine ddRAD Set” and 610 individual gene (ddRAD loci) trees. Numbers at nodes are support values from local posterior probabilities for the main topology (pp1) and the two alternative topologies (pp2 & pp3). See Table S1 for specimen information of the corresponding tip labels (ddRAD IDs). (PDF 9 kb)

Figure S7. Time tree (BEAST, relaxed normal molecular clock) based on the “AC mitochondrial dataset” (including in total 330 cichlid species, including several undescribed species, and 134 taxa representing the remaining major ovalentarian lineages). For detailed information on constrained nodes and applied calibration scheme please refer to Table 1. Node bars indicate 95% HPD intervals of divergence events. Numbers at nodes refer to Bayesian Posterior Probabilities (PDF 49 kb)

Figure S8. Time tree (BEAST, relaxed normal molecular clock) based on the “Austrotilapiine ddRAD Set” (including 97 taxa). For detailed information on constrained nodes and applied calibration scheme please refer to Table 1. Node bars indicate 95% HPD intervals of divergence events. Numbers at nodes refer to Bayesian Posterior Probabilities. See Table S1 for specimen information of the corresponding tip labels (ddRAD IDs). (PDF 11 kb)

Figure S9. Time tree (BEAST, relaxed normal molecular clock) based on the “Austrotilapiine mtDNA dataset” (including 97 taxa). For detailed information on constrained nodes and applied calibration scheme please refer to Table 1. Node bars indicate 95% HPD intervals of divergence events. Numbers at nodes refer to Bayesian Posterior Probabilities. (PDF 11 kb)

## **Supplementary Tables**

Table S1. Overview of individuals genotyped using ddRAD sequencing, information on taxonomy, repository numbers of specimens, availability of corresponding partial mitochondrial genomes and where applicable origin of samples are provided. (XLSX 50 kb)

Table S2. Overview of cichlid specimens for which partial or complete mitochondrial genomes were sequenced for this study, information on taxonomy, repository numbers of specimens, and where applicable origin of samples are provided. (XLSX 35 kb)

Table S3. Overview of mitochondrial genome data retrieved from GenBank with corresponding information on specimen metadata, taxonomy and GenBank accession number. (XLSX 42 kb)

Table S4. Overview of input Reads for PyRAD and recovered loci per sample for the various PyRAD analysis conducted. (XLSX 66 kb)

Table S5. Overview of D-statistics and corresponding Bonferroni-corrected p-values for all possible ABBA-BABA species trios tested. (XLSX 1164 kb)