

18 **Abstract**

19 Adaptive radiation of fishes was long thought to be possible only in lacustrine environments.
20 Recently, several studies have shown that also riverine and stream environments provide the
21 ecological opportunity for adaptive radiation. In this study, we report on a riverine adaptive
22 radiation of six ecomorphs of cyprinid hillstream fishes of the genus *Garra* in a river located in
23 the Ethiopian Highlands in East Africa. *Garra* are predominantly highly specialized algae-scrapers
24 with a wide distribution ranging from Southeastern Asia to Western Africa. However, adaptive
25 phenotypic diversification in mouth type, sucking disc morphology, gut length and body shape
26 have been found among these new species in a single Ethiopian river. Moreover, we found two
27 novel phenotypes of *Garra* ('thick-lipped' and 'predatory') that were not described before in this
28 species-rich genus (>160 species). Mitochondrial and genome-wide data suggest monophyletic,
29 intra-basin evolution of *Garra* phenotypic diversity with signatures of gene flow from other local
30 populations. Although sympatric ecomorphs are genetically distinct and can be considered to being
31 young species as suggested by genome-wide SNP data, mtDNA was unable to identify any genetic
32 structure suggesting a recent and rapid speciation event. Furthermore, we found evidence for a
33 hybrid origin of the novel 'thick-lipped' phenotype, as being the result of the hybridization of two
34 other sympatrically occurring species. Here we highlight how, driven by ecological opportunity,
35 an ancestral trophically highly specialized lineage is likely to have rapidly adaptively radiated in
36 a riverine environment, and that this radiation was promoted by the evolution of novel feeding
37 strategies.

38 **Introduction**

39 Unravelling the mechanisms underpinning the biological diversity remains a major challenge in
40 evolutionary biology. With more than 28,000 species, teleost fishes are the most diverse lineage
41 of vertebrates, and thus an ideal system to address questions regarding diversification. The
42 stunning phenotypic diversity of bony fishes has largely been produced through the process of
43 adaptive radiation, the rapid proliferation of multiple ecologically distinct species from a common
44 ancestor (Schluter, 2000). One of the most extraordinary examples of both adaptive radiation and
45 explosive diversification is represented by the cichlid fishes inhabiting the East African Great
46 Lakes (Kocher, 2004). According to Losos (2010) and Givnish (2015) adaptive radiation and
47 explosive diversification are distinct phenomena: the former may or may not result in, or be
48 accompanied by the latter. The evolutionary success of the cichlids, unmatched among vertebrates,
49 has been promoted by a combination of different factors, where a dominant role has been played,
50 for example, by limited dispersal (because of territoriality and mouth-brooding) and sexual
51 selection for nuptial coloration and mating behavior (Henning & Meyer, 2014; Meyer, Kocher,
52 Basasibwaki, & Wilson, 1990; Seehausen, 2000; Wagner, Harmon, & Seehausen, 2012). It has
53 been suggested, however, that trophic radiation had preceded the diversification driven by other
54 factors at least in cichlids of Lake Tanganyika (Muschick et al., 2014), a cradle of all other East
55 African haplochromine radiations (Salzburger, Mack, Verheyen, E., & Meyer, 2005). Adaptive
56 radiations and diversification bursts were found not only in cichlids, but also in other fish groups,
57 even though in smaller scale, and often in a parallel manner - coregonids, Arctic charrs, and
58 sticklebacks (e.g. Broderson, Post, & Seehausen, 2018; DeFaveri & Merila, 2013; Jacobs et al.,
59 2020; McKinnon & Rundle, 2002; Præbel et al., 2013; Peichel et al., 2001; Schluter, 2000;
60 Skúlason, 1999; Terekhanova et al., 2014) - some of the best known examples of intralacustrine
61 radiations.

62 The most supported cases of monophyletic, closely related fish species that are believed to
63 have arisen through an adaptive radiation event have been described from lakes rather than rivers

64 (Meyer et al. 1990; Seehausen, 2006; Sturmbauer, 1998; Taylor, 1999). For long time, riverine
65 environment has not been considered suitable for adaptive radiation because of its unstable
66 hydrological regimes, reduced habitat diversity and the commonly shallow and narrow
67 watercourses that might facilitate gene flow (Seehausen & Wagner, 2014). However, during the
68 last two decades, examples of fish adaptive radiations occurring in rivers have been reported
69 (Burress et al., 2018; Dimmick, Berendzen, & Golubtsov, 2001; Levin, Simonov, Dgebuadze,
70 Levina, & Golubtsov, 2020; Melnik, Markevich, Taylor, Loktyushkin, & Esin, 2020; Piálek,
71 Říčan, Casciotta, Almirón, & Zrzavý, 2012; Schwarzer, Misof, Ifuta, & Schliewen, 2011;
72 Whiteley, 2007). Although several cases of riverine diversification of cichlid fishes are considered
73 as remnants of adaptive radiations occurred in the palaeo-Lake Makgadikgadi before it dried up
74 back in the Holocene (Joyce et al., 2005), mounting evidence suggests that some fish species flocks
75 of other species than cichlids have diversified within rivers (Burress et al., 2018; Levin et al., 2019;
76 2020; Melnik et al., 2020; Piálek et al., 2012)

77 In the present study, we investigate a highly diverse fish group that presumably adaptively
78 radiated in riverine environments. The genus *Garra* is a species-rich lineage of labeonine cyprinids
79 comprising more than 160 species and is distributed from Southeast Asia to West Africa (Fricke,
80 Eschmeyer, & Van der Laan, 2021; Yang et al., 2012). *Garra* are mostly moderate-sized fish
81 (usually less than 20 cm in length) with sucking gular disc that inhabit the rhithron zone of river
82 systems (Kottelat, 2020). They are predominantly highly specialized algae scrapers that graze
83 periphyton from rocks and stones using widened jaws equipped with horny scrapers. However,
84 adaptations to still waters such as caves or lacustrine environment have been documented in the
85 *Garra*, although rarely, accompanied by a reduction of the gular disc and a change of the foraging
86 strategy from algae scraping to planktivory (Geremew, 2007; Kottelat, 2020; Segherloo et al.,
87 2018; Stiassny & Getahun, 2007; www.briancoad.com).

88 The Ethiopian Highlands are recognized as a center of *Garra* diversity within Africa
89 (Golubtsov, Dgebuadze, & Mina, 2002; Stiassny & Getahun, 2007), where 13 described species

90 out of the total 23 found in Africa are recorded (Moritz, El Dayem, Abdallah, & Neumann, 2019).
91 An assemblage of six *Garra* ecomorphs exhibiting extreme morphological diversity was recently
92 discovered in the Sore River (the White Nile Basin) in southwestern Ethiopia during a survey of
93 the Ethiopian fishes (Golubtsov, Cherenkov, & Tefera, 2012). In particular, two of the six forms
94 display features not found elsewhere within the generic range: a form with a pronounced predatory
95 morphology (large-sized, large-mouthed, with reduced sucking disk and a short gut that is equal
96 to body length) and one with ‘rubber’ lips and prolonged snout region (Fig. 1, Table 1). The other
97 four forms from the Ethiopian *Garra* assemblage drastically differ in mouth and gular disc
98 morphology as well as in body shape (Fig. 1).

99 Fig. 1.

100 Our goals were twofold: i) to investigate the morpho-ecological relationships of six *Garra*
101 sympatric ecomorphs from the Sore River, and ii) to test whether this assemblage has evolved
102 sympatrically. In detail, we aimed at elucidating the population structure and evolutionary history
103 of these ecomorphs using both mitochondrial DNA (mtDNA, cytochrome *b*) and genome-wide
104 nuclear loci obtained with a double digest restriction-site associated DNA (ddRAD) approach.

105 Table 1.

106

107 **Materials and Methods**

108 *Study area*

109 The Sore River is a headwater tributary of the Baro-Akobo-Sobat drainage in the White Nile basin,
110 (south-western Ethiopia, northern East Africa). It drains the Ethiopian Highlands close to the
111 south-western escarpment. The region is covered by moist Afromontane forest that is drastically
112 shrinking in the last decades due to agricultural development (Dibaba, Soromessa, & Workineh,
113 2019). The Sore is a rather little river with a length of *ca.* 160 km, its catchment area is *ca.* 2000
114 km² and characterized by substantial seasonal variation of rainfall (dry season from December to
115 March) (Kebede, Diekkrüger, & Moges, 2014). In comparison, the Italian Tiber River length is

116 406 km, its catchment area is 17375 km² (<https://en.wikipedia.org/wiki/Tiber>). Elevation
117 difference between the Sore source (altitude of ca. 2215 m asl, above sea level) and its confluence
118 with the Gabba (Geba) River (alt. 963 m asl) is 1.25 km. The Sore River basin shares drainage
119 boundaries with two of six major watersheds of Ethiopia: Blue Nile in the north-east and Omo-
120 Turkana in the south-east.

121 We sampled the middle reaches of the Sore River at two sites: (1) at the City of Metu
122 (8°18'42" N 35°35'54" E, alt. 1550 m asl) and (2) ca. 35 km downstream along the river course
123 (8°23'56" N 35°26'18" E, alt. 1310 m asl). The river width at the rapids sampled was 20-40 m at
124 the beginning of the rainy season, depth <1 m, bottom consisted of rocks and large boulders. Fish
125 fauna of the river segment under consideration includes (apart from *Garra* spp.) a species flock of
126 *Labeobarbus* (Levin et al., 2020), *Enteromius* cf. *pleurogramma* (Boulenger 1902), *Labeo* cf.
127 *cylindricus* Peters 1852, *Labeo forskalii* Rüppell 1835, *Chiloglanis* cf. *niloticus* Boulenger 1900
128 (at the lower site only), and introduced *Coptodon zillii* (Gervais 1848). Presence of the stony loach
129 (*Afronemacheilus*) reported by Getahun and Stiassny (1998) from the Sore River at Metu could no
130 longer be confirmed (Melaku, Abebe Getahun, & Wakjira, 2017; Prokofiev & Golubtsov, 2013;
131 present study). Attempts to re-sample a stony loach by intensive electrofishing in 2012 have
132 resulted in the discovery of the enormous morphological *Garra* diversity in the Sore River
133 (Golubtsov et al., 2012). A hundred kilometers westward, from the lowland part (alt. ca. 500 m
134 asl) of the same river drainage >100 fish species are recorded (Golubtsov & Darkov, 2008;
135 Golubtsov, Darkov, Dgebuadze, 1995;) and >115 species from the Sudd and White Nile in Sudan
136 and South Sudan (Moritz et al., 2019; Neumann, Obermaier, & Moritz, 2016;).

137

138 *Sampling*

139 *Garra* samples from the Sore River were collected using a battery driven electrofishing device
140 (LR-24 Combo Backpack, Smith-Root, USA), cast and frame nets in June 2012 and April 2014.
141 In 2011-2014 comparative *Garra* samples were collected from nine sites in six main Ethiopian

142 basins (Fig. 2, Table S1). Fish sampling was conducted under the umbrella of the Joint Ethiopian-
143 Russian Biological Expedition (JERBE) with the permissions of National Fisheries and Aquatic
144 Life Research Center (NFALRC) under Ethiopian Institute of Agricultural Research (EIAR) and
145 Ethiopian Ministry of Science and Technology (presently Ministry of Innovation and Technology).
146 Fish were killed with an overdose of an anesthetic MS-222, first preserved in 10% formalin and
147 then transferred to 70% ethanol. From each specimen fin tissue samples were fixed with 96%
148 ethanol. Some fish specimens were pictured using a Canon EOS 50D camera. All specimens
149 (Supplementary Table S1) are deposited at the A.N. Severtsov Institute of Ecology and Evolution,
150 at the Russian Academy of Sciences, Moscow, under provisional labels of JERBE.

151

152 Fig. 2.

153

154 ***Morphological analysis***

155 *Morphometry*

156 The 28 morphometric characters from 107 individuals of all ecomorphs from the Sore River were
157 examined following Hubbs and Lagler (1958) with additions from Menon (1964): standard length
158 (SL), head length (HL), snout length (R), eye diameter (O), postorbital distance (PO), interorbital
159 distance (IO), head width (HW), head height at nape (HH), head height at mid-of-eye (Hh), mouth
160 width (MW), disc length (DL), disc width (DW), maximal body height (H), minimal body height
161 at caudal peduncle (h), predorsal length (PL), postdorsal length (PDL), prepelvic length (PPL),
162 preanal length (PAL), caudal peduncle length (CPD), dorsal fin base length (DFL), dorsal fin depth
163 (DFP), anal fin base length (AFL), anal fin depth (AFD), pectoral fin length (PFL), ventral fin
164 length (VFL), pectoral-ventral fin distance (PV), ventral-anal fin distance (VA), and distance
165 between anal opening and anal fin (DAA). Measurements were done using a digital caliper (to
166 nearest 0.1 mm). All measurements were performed by one operator for the purpose of consistency
167 as recommended by Mina, Levin, and Mironovsky (2005).

168 Measured individuals had body length varied from 43.6 to 185.0 mm SL: ecomorph 1 (71.5-
169 151.0), ecomorph 2 (70.9-160.2), ecomorph 3 (49.3-100.6), ecomorph 4 (49.3-90.6), ecomorph 5
170 (43.6-81.0; one individual had outstanding length - 185.0), ecomorph 6 (118.4; 139.4) (defined as
171 in Fig. 1 and Table 1), intermediate phenotypes (59.3-105.2). The proportions of head and body
172 were used for principal component analysis (PCA) - measurements of head parts were divided for
173 head length and measurements of body parts were divided for standard length. Data was scaled.
174 The gular disc in some specimens of ecomorph 5 was greatly reduced which hampered the
175 detection of its borders. For the purpose of justification of the values of this character, the identical
176 intermediate values were arbitrarily assigned for all specimens of this ecomorph. PCA was done
177 using *prcomp* script implemented in R with a variance-covariance matrix.

178

179 *Gut length and preliminary assay of a diet*

180 Intestines were taken out from the body cavity of 62 preserved specimens of all ecomorphs except
181 for no. 6 (represented by only two specimens), and measured using a ruler to the nearest 1 mm.
182 The sample size for each ecomorph is provided in Table 2. The standard length (SL) of examined
183 individuals varied from 40 to 131 mm, one individual of ecomorph 5 had outstanding length - 185
184 mm. The ratio of gut length (GL) to SL was used for subsequent analyses. The Kruskal-Wallis
185 test for multiple independent samples with Benjamini-Hochberg method of control of false
186 discovery rate (FDR) (Benjamini & Hochberg, 1995) of p -value was applied to check a
187 significance of differences at $p < 0.05$. The dependence of GL on SL was visualized using
188 scatterplots and regressions. R-packages *ggplot2* and *PMCMR* were used to create plots and to test
189 statistical significance of differences.

190 Diet was assessed for the same individuals, whose intestine length was measured. The main
191 ecological and systematic groups were registered using stereo-microscope Micromed MC-2-
192 ZOOM and microscope Olympus CX41. A composite measure of diet, an index of relative
193 importance, IRI (Hart, Calver, & Dickman, 2002), was used to assess contribution of different

194 components to a diet. The diet components were grouped in several items i) periphyton, ii) benthos,
195 iii) macrophytes, and iv) others.

196

197 *DNA sampling, extraction, amplification, and sequencing - mtDNA data*

198 DNA samples (n=107) were collected from *Garra* inhabiting the Sore River near the City of Metu
199 in 2012 and 2014 from all six forms (see Table 2 for details). For comparison additional DNA
200 samples (n=20) were collected from 8 *Garra* species inhabiting all main drainages of Ethiopia (10
201 localities – see map of sampling in Fig. 2). Total genomic DNA was extracted from ethanol-
202 preserved fin tissues using the BioSprint 15 kit for tissue and blood (Qiagen). Sequences of the
203 mitochondrial gene, cytochrome *b* (*cytb*) of 989 bp length, were amplified (see PCR conditions in
204 Supplementary Material S2; Palumbi, 1996; Perdices & Doadrio, 2001). PCR products were
205 visualized on 1% agarose gels, purified with ExoSAP-IT™ and sequenced at the Papanin Institute
206 of Biology of Inland Waters (Russian Academy of Sciences) using an ABI 3500 sequencer. All
207 new sequences were deposited in GenBank (Accession Numbers: xxx -will be provided upon
208 acceptance, see Supplementary Table S1).

209

210 Table 2.

211

212 *Analysis of mtDNA data*

213 All sequences were aligned and edited using the MUSCLE algorithm (Edgar, 2004) as
214 implemented in MEGA 6.0 (Tamura, Stecher, Peterson, Filipski, & Kumar, 2013). A final set that
215 includes also comparative material from Genbank (African and non-African *Garra* as well as
216 outgroups) encompassed 143 *cytb* sequences (<https://www.ncbi.nlm.nih.gov>) (Table S1).

217 *Akrokolioptax bicornis* and *Crossocheilus burmanicus* were included as outgroups according to
218 previously published phylogenies (Yang et al., 2012).

219 Gene tree reconstruction was performed using both maximum-likelihood (ML) and Bayesian
220 inference (BI) approaches. Prior to these analyses all sequences were collapsed into common
221 haplotypes using ALTER software (Glez-Peña, Gómez-Blanco, Reboiro-Jato, Fdez-Riverola, &
222 Posada, 2010). We determined the best fit models of nucleotide substitution for each codon
223 position of *cytb* and optimal partitioning scheme using either ModelFinder (as implemented in IQ-
224 TREE 1.6.12; Kalyaanamoorthy, Minh, Wong, Von Haeseler, & Jermin, 2017; Nguyen, Schmidt,
225 Von Haeseler, & Minh, 2015) or PartitionFinder 2.1.1 (Lanfear, Calcott, Ho, & Guindon, 2012)
226 under Bayesian Information Criterion (BIC). The partition scheme selected by ModelFinder
227 (codon position 1 - K2P+R2; codon position 2 - HKY+F+I; codon position 3 - TN+F+G4) was
228 subsequently used in ML search with IQ-TREE, using 1 000 bootstrap replicates.

229 Bayesian phylogenetic inference (BI) was carried out in MrBayes v. 3.2.6 (Ronquist et al.,
230 2012). The selected partition scheme was following: codon position 1 with K80+I+G, codon
231 position 2 with HKY+I, and codon position 3 with GTR+G. Two simultaneous analyses were run
232 for 10^7 generations, each with four MCMC chains sampled every 500 generations. Convergence
233 of runs was assessed by examination of the average standard deviation of split frequencies and the
234 potential scale reduction factor. In addition, stationarity was confirmed by examining posterior
235 probability, log likelihood, and all model parameters by the effective sample sizes (ESSs) in the
236 program Tracer v1.6 (Rambaut, Suchard, Xie, & Drummond, 2014). The gene trees resulting in
237 ML and BI analyses were visualized and edited using FigTree v.1.4.4 (Rambaut, 2014). A
238 haplotype network was constructed using the median joining algorithm (Bandelt, Forster, & Röhl,
239 1999) in PopArt 1.7 (Leigh & Bryant, 2015).

240

241 *ddRAD-seq library preparation*

242 High molecular weight DNA was isolated from fin tissue preserved in ethanol using QIAamp DNA
243 Mini Kit (Qiagen, Germany) or obtained by purification of salt method extracted DNA (Aljanabi
244 & Martinez, 1997) using CleanUp Standard kit (Evrogen, Moscow). The dsDNA quantity was

245 measured using dsDNA HS Assay Kit for fluorometer Qubit 3 (Life Technologies, USA). ddRAD-
246 library was constructed following the quaddRAD protocol (Franchini, Monné Parera, Kautt, &
247 Meyer, 2017) using restriction enzymes *Pst*I and *Msp*I. In total, 77 DNA samples of *Garra*
248 ecomorphs from the Sore River (see Table 2) and 11 DNA samples from five other species of
249 Ethiopian *Garra* from adjacent basins were sequenced by two independent runs of Illumina
250 HiSeq2500 and Illumina X Ten (2 x 150 bp paired-end reads). The raw sequencing data were
251 demultiplexed by the sequencing provider using outer Illumina TruSeq dual indexes.

252

253 *Processing of RAD-seq data*

254 The resulting reads were trimmed for remaining adapters and low quality reads Cutadapt
255 implemented in the Trim Galore 0.4.5 package (<https://github.com/FelixKrueger/TrimGalore> -
256 Martin, 2011). Read quality was assessed with FastQC 0.11.7 (Andrews & Krueger, 2010) and
257 MultiQC 1.7 (Ewels, Magnusson, Lundin, & Källér, 2016) before and after trimming. Further
258 demultiplexing of individually barcoded samples, construction and cataloging of RAD-loci, and
259 SNP calling were done with STACKS 2.41 package (Catchen, Hohenlohe, Bassham, Amores, &
260 Cresko, 2013). Identification and removal of PCR duplicates were done using the ‘*clone_filter*’
261 module of STACKS). STACKS module ‘*process_radtags*’ was used to demultiplex reads by the
262 dual index inner barcodes and obtain separate fastq files for each individual. Samples that failed
263 to produce more than 100 000 reads were excluded from further processing. To additionally
264 evaluate data quality and identify possible contaminated samples, the reads were mapped to the
265 reference genome of common carp *Cyprinus carpio* (GCF_000951615.1) using bowtie2 2.3.5
266 (Langmead & Salzberg, 2012) with ‘--local-sensitive’ presettings. Then, only Read 1 (R1) files
267 were used for downstream processing and analyses. Prior to next steps, these R1 reads were
268 trimmed at their 3` ends to a uniform length of 130 bp to reduce the influence of sequencing error
269 (due to declined base quality at 3` end).

270 The *de novo* pipeline of STACKS was used to assemble loci and perform genotype calling.

271 We selected optimal parameters using the approach suggested by Paris, Stevens, & Catchen
272 (2017). Following the aforementioned procedure, we found that minimum stack depth ($-m$) of 5,
273 distance allowed between stacks ($-M$) of 3, and the maximum distance required to merge catalog
274 loci ($-n$) of 5 provided the best balance between data quality and quantity for our dataset (Fig. S1).

275

276 *Population genomic analyses*

277 Individual genotypes of sympatric *Garra* ecomorphs from the Sore River were exported to a vcf
278 file using the ‘populations’ module of STACKS with the following settings: (i) loci genotyped in
279 at least 90% of samples ($-r$ 0.90) were kept; (ii) SNPs with a minor allele frequency ($--min-maf$)
280 less than 0.04 and a maximum observed heterozygosity ($--max_obs_het$) above 0.99 were pruned;
281 (iii) only single SNP per RAD locus was retained, to avoid inclusion of closely linked SNPs. We
282 applied VCFtools 0.1.16 (Danecek et al., 2011) for further filtering of the dataset based on mean
283 coverage and fraction of missing data for each sample. Samples with more than 20% of missing
284 data were blacklisted and excluded from further analyses. Thus, a high-quality dataset of 679 SNPs
285 and 77 individuals was obtained and used for downstream population genetics analyses.

286 First, Principal Component Analysis (PCA) was performed using the ‘*glPca*’ function of the
287 R-package *adegenet* 2.1.1 (Jombart, 2008; Jombart & Ahmed, 2011). Next, *rmaverick* 1.0.5
288 (former MaverickK; Verity & Nichols, 2016) was used to infer population structure. This program
289 estimates evidence for different numbers of populations (K), and different evolutionary models via
290 generalised thermodynamic integration (GTI). A range of K values between 1 and 10 were
291 explored, using 300 000 burn-in MCMC iterations and 10 000 sampling iterations. Convergence
292 of MCMC was automatically tested every 1 000 burn-in iterations by activating option
293 ‘auto_converge’. This allows exit burn-in iterations when convergence is reached and immediately
294 proceeds to sampling iterations. Parameter ‘rungs’ was set to 10 (number of multiple MCMC
295 chains with different ‘temperature’ to run simultaneously). Both no admixture and admixture
296 models were run, and compared by plotting values of the posterior distribution and overall model

297 evidence in log space (log-evidence) (Fig. S2-S5). According to this comparison, the admixture
298 model is decisively supported over the no admixture model, and used here to report the results.
299 The same protocol was followed for consecutive hierarchical *rmaverick* runs for the identified
300 clusters. Finally, global and pairwise Reich-Patterson F_{ST} values (Reich, Thangaraj, Patterson,
301 Price, & Singh, 2009) with respective 95% confidence intervals for ecomorphs/genetic clusters
302 were calculated using the R script from Junker et al. (2020). Basic genetic diversity statistics were
303 calculated using the ‘*populations*’ module of STACKS.

304 To test for the gene flow between ecomorphs/genetic clusters, we used the Patterson’s D
305 statistic (ABBA-BABA test), along with the f_4 -ratio statistic (Patterson et al., 2012) and its f -
306 branch metric (Malinsky et al., 2018), as implemented in Dsuite 0.4 software package (Malinsky,
307 Matschiner, & Svardal, 2021). Patterson’s D statistic is a widely used and robust tool to detect
308 introgression between populations or closely related species, and to distinguish it from incomplete
309 lineage sorting (ILS). The f_4 -ratio statistic is a similar method aiming to estimate an admixture
310 fraction. The f -branch metric is based on f_4 -ratio results and serves to assign gene flow evidence
311 to specific branches on a phylogeny. These tests were performed on a group containing
312 ecomorphs/genetic clusters 2b, 3, 4, and 6, while the rest were used as outgroup (in accordance
313 with the results of our phylogenomic analysis).

314

315 *Phylogenomic analyses*

316 IQ-TREE 2.0.5 (Minh et al., 2020) was used for ML phylogenetic analyses of RAD-seq data. First
317 dataset included one to three specimens of each *Garra* ecomorph from the Sore river and other
318 Ethiopian *Garra* species from adjacent basins. Multiple sequence alignments of all loci and
319 respective partition files were created using the ‘--phylip-var-all’ option of ‘*populations*’ module
320 of STACKS package. Heterozygous sites within each individual were encoded using IUPAC
321 notation. During the analysis each RAD-locus was treated as a separate partition with independent

322 best-fit substitution model. Node support values were obtained using ultrafast bootstrap procedure
323 (Hoang, Chernomor, von Haeseler, Minh, & Vinh, 2018) with 1 000 replicates. We also used
324 SVDQuartets algorithm (Chifman & Kubatko, 2014) as implemented in PAUP* 4.0a168
325 (Swofford, 2003) to perform species-tree inference under the multi-species coalescent model using
326 18,988 SNPs (single random SNP per locus, minor allele frequency cutoff 0.04, maximum
327 observed heterozygosity cutoff: 0.99). Node support was estimated with 1 000 bootstrap replicates.

328 The second dataset consisted of all genotyped specimens of sympatric *Garra* ecomorphs
329 from the Sore River and a single, most closely related outgroup (*G. cf. dembeensis* from the
330 Barokalu River, as revealed by the analysis of the first phylogenomic dataset that included samples
331 from all the localities in Figure 2). It was analysed with IQ-TREE as described above, except for
332 GTR+G substitution model was used for each partition. The phylogenetic trees were visualized
333 and edited using FigTree 1.4.4 (Rambaut & Drummond, 2008).

334

335

336 **Results**

337 *Trophic Morphology*

338 PCA of head and body proportions of six sympatric ecomorphs from the Sore River revealed five
339 well-defined clusters (Fig. 3A). Four clusters represent ecomorphs 3, 4, 5, and 6, while the fifth
340 includes individuals from ecomorphs 1 and 2. The ecomorph 5 is the most divergent. PC1
341 explained 72.3% of the total variance, while PC2 10.2%. The eigenvector with the highest
342 eigenvalues for PC1 were head proportions - nine of ten most loaded ones (especially gular disc
343 proportions, mouth width, interorbital distance, and snout length). The same pattern was detected
344 for PC2 - nine of ten most loaded characters belonged to head proportions (mainly disc length,
345 mouth width, height of head at nape and at eyes etc. - see Table S2 for details).

346 After excluding ecomorph 5, the ecomorphs 1 and 2 became more distinguishable with low
347 overlapping (Fig. 3B). The PC1 explained 73.8% of variance, while PC2 8.1%. The most loaded

348 eigenvectors of both PC1 and PC2 were from head proportions with few more contributions of
349 some body proportion characters (see Table S3). The difference between ecomorphs 1 and 2
350 revealed in PC2 is explained by height of head at both nape and eyes, interorbital distance, head
351 width, body height as well as other characters (Table S3).

352 Fig. 3.

353

354

355 *Gut length and preliminary data on diet*

356 Gut length broadly varied consistently between ecomorphs (Fig. 3C). Shortest guts (107-160 %
357 SL) were detected in ecomorph 5 suggested a predatory trophic type, while the longest guts were
358 recorded in ecomorphs 1 (285-799 % SL) and 2 (354-555 % SL) that possessed the well-developed
359 gular disc and therefore are specialized algal grazers, as also shown by their gut contents (see
360 below). Other ecomorphs had intermediate values gut lengths: ecomorph 3 - 124-295 % SL, and
361 ecomorph 4 - 175-513 % SL, respectively. Broad intra-group variation is explained by increase of
362 gut length with body length detected in some ecomorphs (Fig. 3D). Nevertheless, the similar-sized
363 individuals are divergent in gut length at the same manner that presented in Fig. 3C. Ecomorph 5
364 having the shortest gut displays even a slight decrease of gut length ontogenetically that was
365 previously reported for piscivorous mode of feeding among African cyprinids (Levin et al., 2019).

366 The preliminary inspection of gut content revealed differences in the diet between some
367 ecomorphs. Ecomorphs 1 and 2 had permanently filled intestines full of periphyton (diatom, green,
368 and charophyte algae; IRI = 99.98% for ecomorph 1, and IRI = 97.99% for ecomorph 2) and, rarely
369 other items (larvae of water insects - mayflies, chironomids, simuliids). The ecomorph 3 had a half-
370 filled gut with dominating periphyton (IRI = 86.3%) with a notable portion of insect larvae (7.62%
371 - predominantly chironomids, also mayflies, and simuliids) and macrophytes (5.97%). Ecomorph
372 4 had fewer filled intestines compared to ecomorph 3 however with strongly dominating
373 periphyton in diet (IRI = 99.49%). The gut of ecomorph 5 (shortest gut) frequently was empty

374 including the largest individual (SL=185 mm). When guts were filled, benthos-associated prey
375 was strongly prevalent (IRI = 99.31%; mayflies and chironomids).

376

377 **Mitochondrial data**

378 Both BI and ML analyses of *cytb* revealed monophyly of the *Garra* from the Sore River (Fig. 4A).
379 The closest relative (and ancestor lineage) is from the Barokalu River, a tributary of the Baro River
380 (White Nile drainage). Both Sore and Barokalu rivers share watershed in the Baro system and
381 sampled localities are separated just ca. 50 km by land. Divergence between *Garra* populations
382 from the Sore and Barokalu is low (p -distance = 0.0105 ∓ 0.0028) and comparable with maximum
383 intra-divergence in the Sore radiation (p -distance = 0.0111 ∓ 0.0033). Being combined together
384 White Nile lineage is a sister to the large clade of Ethiopian *Garra* from Blue Nile and Lake Tana,
385 Atbara-Nile, Ethiopian Rift Valley, and Omo-Turkana basins.

386 At the same time, our phylogenetic analyses revealed that Ethiopian *Garra* are non-
387 monophyletic (Fig. 4A). Some lineages are of more ancient origin and closer to Asian lineages (*G.*
388 *tibanica* from Indian Ocean basin) or to lineages from West Africa (e.g. *G. vinciguerra* from Blue
389 Nile basin). Matrilineal tree of Ethiopian *Garra* includes up to 12 lineages. Taking into account
390 some species cluster together in one lineage like three species from Lake Tana or that some species
391 were unavailable, we conclude cladogenesis of *Garra* in Ethiopia Highlands has been more
392 diversified than considered previously (Stiassney & Getahun, 2007).

393

394 Fig. 4.

395

396 The Sore lineage is composed of two sub-lineages or haplogroups highlighted by yellow and
397 green (Fig. 4A-B). Haplotype net constructed on 107 *cytb* sequences confirms presence of two
398 main haplogroups. The core haplotypes of these haplogroups are separated by 5 substitutions. Four
399 of six ecomorphs (2, 3, 4, and 5) share both haplogroups. The ‘green’ haplogroup is prevalent in

400 number of haplotypes (18), and number of individuals (88), and found in five ecomorphs.
401 Ecomorph 1 is presented exclusively in this haplogroup. In contrast, the ‘yellow’ haplogroup (Fig.
402 4B) is smaller, with only different 9 haplotypes found in 19 individuals (= 17.7 % of the individuals
403 analyzed). One individual of ecomorph 4 is rather distant (6 substitutions) from the core haplotype
404 of this haplogroup. ‘Yellow’ haplogroup consists of five ecomorphs as well. However, ecomorph
405 4 is much more frequently represented in this haplogroup (42 % of all individuals) compared to
406 ‘green’ one (6.97 %).

407

408 **RAD-seq data**

409 Raw reads statistics is given in Supplementary File S1.

410

411 *Nuclear phylogeny*

412 The phylogeny of Ethiopian *Garra* based on a concatenated set of RAD-loci sequences (23,365
413 partitions and 3,075,180 total sites with 0% missing data) is generally similar to that based on
414 mtDNA data (Fig. 4) but it has more strongly supported nodes, as it is based on many more variable
415 sites (Fig. 5A). Sympatric ecomorphs clustered together and form monophyletic lineages, sister to
416 the population from the same riverine basin - Baro drainage in White Nile system (Fig. 5A-B).
417 Closest relative to *Garra* from White Nile system is *Garra* lineage in the *G. dembeensis* complex
418 from neighbor drainage - Omo-Turkana system. The *G. vinciguerrae* from the Blue Nile (which
419 recorded in Ethiopia for the first time in the current study) is ancestor lineage for both White Nile
420 and Omo-Turkana lineages. The most divergent lineages, *G. makiensis* and *G. tibanica*, are from
421 Ethiopian Rift Valley and Indian Ocean basins, respectively.

422

423 Fig. 5.

424

425 Compared to mitochondrial data, the nuclear phylogenomic tree shows much better
426 segregation of *Garra* ecomorphs from the Sore River (Fig. 5A). Ecomorphs 3, 4, and 6 form
427 monophyletic clusters, while other ecomorphs are divided into two (nos. 1 and 5) or even three
428 (no. 2) clusters. We assign two distantly located branches of both ecomorph 1 (generalized) as
429 1a/1b as well as ecomorph 2 (stream-lined) as 2a/2b according to population genomics analyses
430 done below (Fig. 6-8). Ecomorphs 1 and 2 from one hand, and other ecomorphs from another hand
431 form two clusters within Sore River adaptive radiation according to SVDQ species tree (Fig. 5B).
432 Ecomorphs 3 (narrow-mouth), 4 (wide-mouth), and 6 (thick-lipped) are most recently diverged
433 branches according to SVDQ-tree but the nodes are weakly supported (Fig. 5B).

434 Relationships among the Sore River sympatric ecomorphs based on analysis of all samples
435 and full RAD-loci sequences (> 7000 loci and > 0.96 Mbp length sequences) are presented in Fig.
436 6. The ML analysis highly support the monophyly of each ecomorph except for ecomorph 2. The
437 most basal lineage is ecomorph 2, which in turn, is paraphyletic, suggesting, possibly, that there is
438 another 7th cryptic species that we could not distinguish phenotypically. Four individuals along
439 with one individual of intermediate phenotype represent another lineage that we call 2b (Fig. 6).
440 Lineage 2a is sister all other ecomorphs that are divided for two subclades - one includes only
441 ecomorph 1 individuals (which, in turn is subdivided into what we call - 1a-1b), while another
442 includes all other ecomorphs - 3, 4, 5, 6, and above mentioned 2b. That latter lineage is composed
443 of lineages, each containing samples of particular ecomorphs except for several samples which
444 were intermediate in their phenotypes (Fig. 6). Ecomorph 6 (thick-lipped mouth) is resolved as
445 sister to the 2b lineage albeit with an apparent rather deep last common ancestor. Generally, the
446 placement of clade 2a as sister to all other *Garra* from the Sore River, that is characterized by a
447 well-developed gular disc (type C), might suggest that this an ancestral condition of this radiation.

448

449 Fig. 6.

450

451 *Population genomics*

452 Principle component analyses of the 679 nuclear SNPs of sympatric ecomorphs revealed
453 several well-defined clusters that correspond to the phenotypic differentiation (Fig. 7). Ecomorph
454 1 (composed of two genetic sub-clusters 1a-1b), genetic cluster 2a, ecomorphs 3 and 4 are not
455 overlapping, while clusters of 2b and ecomorph 5 broadly overlap. Thick-lipped ecomorph (6)
456 interestingly (although it is difficult to place since we only found two individuals that we could
457 include in this study) could not be identified by PCA as a distinct cluster.

458

459 Fig. 7.

460

461 The analysis of population structure with admixture revealed an optimum of three genomic
462 clusters that correspond to the i) ecomorph 1 + 2a lineage, ii) ecomorphs 3 + 4, and iii) ecomorph
463 5 + 2b lineage (Fig. 8, Upper row, K3). Ecomorph 6 is characterized by admixture of two clusters
464 from ecomorphs 3 and 4.

465

466 Fig. 8.

467

468 Subsequent analysis of each cluster (=lineage) revealed hierarchical subdivision. Thus
469 ecomorph 1 and genetic lineage 2a each are also identified as cluster in the admixture analysis
470 (Fig. 8 middle row, K=2). Although ecomorphs 3, 4, 5, and lineage 2b are supported as
471 independent evolutionary units based on several types of genetic analyses, few individuals in all
472 of these show signs of historical gene flow based on the admixture analysis (Fig. 8). While the two
473 individuals from ecomorph 6 in our study seem most clearly be composed of genetic contributions
474 by ecomorphs 3 (36.8-47.5%) and genetic lineage 2b (51.3-62.3%), possibly supporting a hybrid
475 origin hypothesis. Interestingly, one more individual with combination of the same genomic
476 clusters but with the opposite ratio (54.0% from ecomorph 3 and 43.9 % from lineage 2b) had no

477 thick-lipped features (the main phenotypic diagnostic feature for ecomorph 6) and was
478 phenotypically assigned to ecomorph 3 (Fig. 8). One more level of population subdivision was
479 detected in ecomorph 1 (Fig. 8) with two genomic clusters (lineages 1a and 1b) of high degree of
480 admixture. It suggests heterogeneous genomic structure of the generalized ecomorph as a result of
481 secondary contact.

482 All Reich F_{ST} pairwise comparisons were statistically significant with values ranging from
483 0.10 (lineages 1a vs. 1b) to 0.46 (ecomorphs 2b vs. 6) (Fig. 9). The ecomorph 6 F_{ST} values were
484 the highest (0.39-0.46).

485

486 Fig. 9.

487

488 As the *rmaverick* analysis suggested a notable level of admixture between lineage 2b and
489 ecomorphs 3, 4, and 6 (Fig X), which form a single monophyletic cluster in our phylogenomic
490 analysis (Fig 8), we performed a number of tests to distinguish between gene flow (introgression)
491 and incomplete lineage sorting (ILS). The obtained D statistic was positive and significant for a
492 number of comparisons (Table 3.). Visualization of f -branch metric (which is based on f_4 -ratio
493 results) highlighting introgression between ecomorphs/genetic lineages 2b and 3, 6 and 3, 5 and 3
494 (Fig 9).

495

496 Table 3.

497

498 The eighth genetic clusters possess from three (ecomorph 6) to 38 private alleles (ecomorph 4)
499 (Table 4). The ecomorph 6 has also the lowest heterozygosity ($H_o = 0.00058$) as well as nucleotide
500 diversity ($P_i = 0.00054$) compared to all other ecomorphs ($H_o = 0.00104$ - 0.00128 ; $P_i = 0.00121$ -
501 0.00091) (Table 4).

502

503 Table 4.

504

505 **Discussion**

506 Our study provides genetic support for the hypothesis of the evolution of an adaptive radiation in
507 a riverine environment. By analyzing trophic features and sucking disc variation, as well as trophic
508 ecology, we show morpho-ecological diversification of the cyprinid fish *Garra dembeensis* into
509 six distinct ecomorphs. First, diversification of two novel phenotypes (thick-lipped and predatory)
510 in the Sore River has evolved rapidly, an event that can be classified as burst of speciation sensu
511 Givnish (2015). Second, adaptive radiation resulted in the origin of several highly specialized
512 lineages of algae scrapers, i.e. specialized ancestor adaptively radiates giving rise to eco-
513 morphological diverse lineages, that seem to be not only ecologically, but also reproductively
514 isolated from each other and can be considered the new species.

515

516 ***Eco-morphological diversification and adaptive radiation of Garra***

517 The genus *Garra* is currently comprised of more than 160 species (Fricke et al., 2021; Yang et al.,
518 2012). Only 23 of which occur in Africa (Moritz et al., 2019). So far, 13 described species were
519 reported from Ethiopia (Golubtsov et al., 2002; Stiassny & Getahun, 2007). In this study, we
520 discovered six additional distinct ecomorphs that originated through adaptive radiation in the Sore
521 River, and thus might warrant the description of five-six new African *Garra* species.

522 The ecomorphs of the Sore's *Garra* are exceptionally diverse in trophic and sucking disc
523 morphology. Two novel phenotypes for the whole genus *Garra* – 'thick-lipped' and 'predatory' -
524 have superficial similarities to Lake Tana large barbs species/morphotypes, e.g., thick-lipped barb
525 *L. negdia* (Rüppell, 1836) and predatory *L. gorguari* (Rüppell, 1836) (Nagelkerke & Sibbing,
526 1997). This high degree of variation in the sucking disc in Sore's *Garra* can be observed - from
527 well-developed disc with free posterior margin to complete absence. Such a degree of

528 morphological diversity concentrated in one riverine spot of Ethiopian Highlands would seem to
529 satisfy the requirements of a diversification burst (sensu Givnish, 2015).

530 Divergent feeding-related morphology and gut content analysis suggest trophic
531 specialization of *Garra* sympatric forms. This is consistent with other cases of adaptive radiation
532 among Ethiopian cyprinids, where trophic resource partitioning promoted diversification -
533 *Labeobarbus* spp. in Lake Tana (Sibbing, Nagelkerke, Stet, & Osse, 1998) as well as in the Genale
534 River (Levin et al., 2019). The most common foraging strategy among *Garra* is scraping of
535 periphyton from stones and rocks (Hamidan, Jackson, & Britton, 2016; Matthes, 1963). This is
536 predominant in Sore's *Garra* ecomorphs 1 and 2 that have long gut (4-5 times longer than body
537 length) filled with periphyton and detritus. The ecomorphs 1 and 2 are divergent mainly in body
538 shape. The latter has streamlined appearance and probably is adapted for life in more rapid flowing
539 water. Ecomorph 3 has shorter gut length (ca. 2-times longer than body length) and a mixed diet
540 with significant additions of benthic invertebrates. Ecomorph 5 has an extremely short gut, whose
541 length is as long as the fish body. Short gut is a strong marker for predatory/piscivory feeding
542 strategy in fishes, including cyprinids (Nagelkerke, 1997; Sibbing et al., 1998; Wagner, McIntyre,
543 Buels, Gilbert, & Michel, 2009, Zandoná, Auer, Kilham, & Reznick, 2015). Predatory *Garra* from
544 the Sore River have 4-5-times shorter gut length than congeneric periphyton feeders and twice
545 shorter gut than that of piscivory large-mouthed ecomorph of *Labeobarbus* from the Genale River,
546 Ethiopia (Levin et al., 2019). We found an empty gut in many individuals of ecomorph 5, while
547 small-sized fishes had gut filled with insects. Ecomorph 4 has a rather long intestine and
548 predominantly periphyton in diet, but it is characterized by distinctly divergent mouth phenotype
549 compared to ecomorphs 1 and 2 (Fig. 3). The gut of thick-lipped phenotype (ecomorph 6) was not
550 analyzed because of the extreme rarity of samples. Hypertrophied lips (or 'rubber lips') of fishes
551 is an adaptation to foraging on benthos hidden between rock crevices on pebble and rock fragments
552 via increased sucking power by sealing cracks and grooves (Baumgarten, Machado-Schiaffino,
553 Henning, & Meyer, 2015; Machado-Schiaffino, Henning, & Meyer, 2014; Matthes, 1963; Ribbink,

554 Marsh, Marsh, & Sharp, 1983). This phenotype is widely distributed among other cyprinid fish,
555 the *Labeobarbus* spp., inhabiting lakes and rivers of Ethiopian Highlands (Mina, Mironovsky, &
556 Dgebuadze, 1996; Mironovsky, Mina, & Dgebuadze, 2019; Nagelkerke, Sibbing, van den
557 Boogaart, Lammens, & Osse, 1994) including the Sore River (Levin et al., 2020), but it was never
558 detected among *Garra* species. Our study shows that the thick-lipped mouth phenotype represents
559 an evolutionary novelty within the *Garra* lineage that most probably resulted from hybridization
560 events between ecomorphs 2 (lineage 2b) and 3 because its genome had an admixture from these
561 genetic lineages. Hybridogenic origin of the *Garra*'s thick-lipped phenotype may corroborate
562 results of recent experimental study demonstrating the importance of hybridization in generating
563 of functional novelty of ecological relevance in relation to trophic resources unavailable for
564 parental species in cichlids (Selz & Seehausen, 2019). The origin of novel thick-lipped phenotype
565 in the genus *Garra* is of particular interest in light of knowledge of non-hybrid origin of
566 hypertrophied lips from ancestors with normally developed lips in cichlid fishes (Baumgarten et
567 al., 2015; Machado-Schiaffino et al., 2017). Interestingly, there might only be a single locus
568 involved in producing the hypertrophied cichlid phenotype (Kautt et al., 2020), the genomic basis
569 of the lip phenotypes in *Garra* remains unknown.

570 Another novel phenotype for *Garra* detected in the Sore River is the “predatory” niche. A
571 conspicuously piscivory trophic strategy is rare among Cypriniformes, presumably because they
572 have a toothless jaw. Nevertheless, this feeding strategy is quite common among cyprinid fishes
573 inhabiting water bodies of Ethiopian Highlands. For example, seven of the total 15 endemic
574 *Labeobarbus* spp. found in Lake Tana are predatory on fish (Nagelkerke et al., 1994; Sibbing et
575 al., 1998); that evolved multiple times among riverine populations of the genus *Labeobarbus*
576 (Levin et al., 2020).

577 To our knowledge, only one sympatric diversification has previously suggested for *Garra* –
578 the intralacustrine complex including three species inhabited Lake Tana in Ethiopia (Geremew,
579 2007; Stiassny & Getahun, 2007). This diversification resulted in divergent phenotypes (gular

580 discs varies from well-developed to reduced size) and ecology (one form is pelagic - *G. tana*) and
581 can be considered as a recent speciation as suggested by the absence of mtDNA divergence among
582 these species (Tang, Getahun, & Liu, 2009). Unfortunately, little is known about morpho-
583 ecological and genetic diversity of this Lake Tana radiation. Sympatric divergence was also
584 recently proposed as the most likely mechanisms for the origin of two blind *Garra* species, *G.*
585 *typhlops* and *G. lorestanensis*, inhabited the same cave in Zagros Mountains, Iran (Segherloo et
586 al., 2018).

587

588 ***Possible scenarios of evolution of Garra's adaptive radiation in the Sore River***

589 Both mtDNA and genome-wide SNPs data support monophyly of the Sore's *Garra* as well as their
590 recent speciation based on low genetic divergence between the nearest ancestor and Sore River's
591 ecomorphs. The closest relative and ancestor of the Sore River diversification inhabits the same
592 subbasin of the White Nile in Ethiopia, therefore suggesting an intra-basin diversification of *Garra*
593 there. On the one hand, mtDNA data might have failed to distinguish sympatric ecomorphs
594 because of high level of shared genetic diversity caused by ILS and introgression, this latter
595 highlighted by D-statistic calculated with the genome-wide nuclear data. On the other hand, the
596 SNP data support a reproductive isolation among closely-related ecomorphs despite few
597 individuals having intermediate phenotypes and genetic admixture. Hybrid origin of intermediate
598 phenotypes might suggest that reproductive isolation barriers are not complete yet.

599 Patterns of haplotype net (numerous haplotypes occurring in the same phenotypes) as well
600 as SNP data (presence of more genetic clusters than phenotypes) could also suggest secondary
601 contact of local sub-isolated populations. The riverine net of Ethiopian Highlands was significantly
602 influenced by several episodes of dramatic volcanism and tectonism until the Quaternary
603 (Ferguson et al., 2010; Hutchison et al., 2016; Prave et al., 2016). Thus, riverine net fragmentation,
604 isolation or sub-isolation of some riverine parts, and captures of headwaters is a likely scenario
605 given the geological history of Ethiopian Highlands (Mège, Purcell, Pochat, & Guidat, 2015), also

606 supported by genetic studies on other Ethiopian fishes (Levin et al., 2019; 2020). Concerning the
607 Sore River, while waterfalls and rapids are rather frequent, no geological data that support its
608 connection to other basins are known. In our view, the most reliable evolutionary scenario for the
609 origin of the riverine adaptive radiation in the *Garra* species group draws upon a combination of
610 allopatric and sympatric stages of speciation with hybridization and admixture. A comparable
611 evolutionary history was detected in the *Labeobarbus* adaptive radiation in the Genale River
612 (Ethiopia), which is part of the extended ancient riverine net in Juba-Wabe-Shebelle drainage
613 (Levin et al., 2019).

614 Speciation with gene flow was detected in several studies (e.g. Feder, Egan, & Nosil, 2012;
615 Fruciano, Franchini, Raffini, Fan, & Meyer, 2016; Kautt, Machado-Schiaffino, & Meyer, 2016;
616 Kautt et al., 2018; Kautt et al., 2020; Machado-Schiaffino et al., 2017; Malinsky et al., 2018; Puebla,
617 2009; Rougeux, Bernatchez, & Gagnaire, 2017; Schwarzer et al., 2011; Smadja & Butlin, 2011;
618 Zheng & Ge, 2010). Notably, it has been shown as genetic admixture between divergent
619 populations/lineages may be a key factor in promoting rapid ecological speciation (Jacobs et al.,
620 2020; Kautt et al., 2016; Kautt et al., 2020; Martin et al., 2015; Marques, Meier, & Seehausen,
621 2019). Moreover, ancient hybridization is widely considered one of the most important factors
622 driving the spectacular cichlid adaptive radiations in the Great African Lakes (Irissari et al., 2018;
623 Meier et al., 2017; Verheyen, Salzburger, Snoeks, & Meyer, 2003). Seemingly, ancient
624 introgressive hybridization could be a trigger for small-scaled repeated adaptive radiations among
625 the Arctic charrs *Salvelinus* (Lecaudey et al., 2018). Furthermore, hybridization is the main
626 mechanism generating polyploid lineages in fishes (tetraploid, hexaploid etc. - Braasch &
627 Postlethwait, 2012), whose complex genomes constitute the raw material for the rapid origin of
628 sympatric forms (e.g. *Schizothorax* in Central Asia - Berg, 1914; Burnashev, 1952; Terashima,
629 1984; *Labeobarbus* in Africa - Levin et al., 2020; Mina et al., 1996; Nagelkerke et al., 1994;
630 Vreven, Musschoot, Snoeks, & Schliewen, 2016). Nevertheless, all described *Garra*, including the
631 Ethiopian species, have diploid genomes (Krysanov & Golubtsov, 1993).

632

633 *Adaptive radiation in riverine environment*

634 Most adaptive radiations of fishes were reported from the lacustrine environment (e.g., Fryer &
635 Iles 1972; Seehausen & Wagner, 2014). However, increasing evidence suggest that adaptive
636 radiation can take place in other aquatic environments (e.g., marine, riverine) (Burress et al., 2018;
637 Dimmick et al., 2001; Feulner, Kirschbaum, & Tiedemann, 2008; Levin et al., 2019; 2020; Melnik
638 et al., 2020; Matchiner, Hanel, & Salzburger, 2011; Piálek et al., 2012; Puebla, 2009; Whiteley,
639 2007). Several other cases of potential riverine adaptive radiations that includes ≥ 3 sympatric
640 ecomorphs exist, although they were not been tested with genetic methods yet - for instance, snow
641 trout from Central Asia (Berg, 1914; Burnashev, 1952), barbs *Poropuntius* and *Neolissochilus*
642 from Southeastern Asia (Roberts, 1998; Roberts & Khaironizam, 2008). Among cichlids, one of
643 the first riverine adaptive radiations examined genetically were from Southern Africa (Joyce et al.,
644 2005). However, the authors of this study suggested that the adaptive radiation occurred in the
645 lacustrine environment in the palaeo lake Makgadikgadi that dried up in the Holocene (Joyce et
646 al., 2005). Other cichlid adaptive radiations from the rivers of Western Africa (Schwarzer et al.,
647 2011), Southern America (Burress et al., 2018; Piálek et al., 2012;) as well as four independently
648 evolved riverine radiations of labeobarbs from East Africa (Levin et al., 2020), have instead took
649 place in riverine drainages without known lacustrine conditions in the past.

650 The *Garra* lineage is adapted to fast and torrent waters. It possesses a morphological novelty
651 - gular sucking disc - used to cling on the bottom of swift waters. This novelty allowed *Garra* to
652 be distributed widely in highlands and montane zones from Southeastern China to Western Africa.
653 Only a few species were found in the lacustrine environment (Lake Tana – Stiassny & Getahun,
654 2007) or in caves (e.g. Banister, 1984; Coad, 1996; Kruckenhauser, Haring, Seemann, & Sattmann,
655 2011; Mousavi-Sabet & Eagderi, 2016), indicating their potential to adapt to steady waters.

656 Despite the riverine network is generally considered more open to gene flow compared to
657 landlocked water bodies, mountain and highland are an exception to this rule. The Ethiopian

658 Highlands are a volcanic massif of flood and shield volcano basalts 0.5–3.0 km thick that form
659 spectacular trap topography (1500–4500 m) flanking the Main Ethiopian Rift (Prave et al., 2016).
660 The geological history of the Ethiopian Highlands was tectonically very dynamic and rich in
661 volcanic episodes from Oligocene to Pleistocene time with very recent episodes (Prave et al.,
662 2016). The volcanic activity has been severe enough to deleteriously affect the biota and cause
663 major disruptions in ecosystems. This hypothesis found support in the inferred evolutionary
664 history of the *Labeobarbus* in East Africa. The earliest fossil records of *Labeobarbus* were found
665 in the Ethiopian Rift Valley and dated back to the late-Miocene (Stewart & Murray, 2017), but
666 most of the Ethiopian lineages are younger (Pleistocene origin) (Beshera, Harris, & Mayden, 2016;
667 de Graaf, Megens, Samallo, & Sibbing, 2010; Levin et al., 2020). The tectonic activity of the
668 region could have favored local isolation via the formation of waterfalls (e.g., 33 kya the Blue Nile
669 basaltic blockade formed Tis-Isat waterfall - Prave et al., 2016) or river net fragmentation (Juba-
670 Wabe-Shebelle drainage Mège et al., 2015) along with climatic oscillations resulted to
671 disconnection of water bodies during aridization (Benvenuti et al., 2002). Periodically, it resulted
672 in vacant habitats and ecological opportunity (reviewed by Stroud & Losos 2018) for new species
673 to exploit similar to islands or crater lakes (Burress et al., 2018).

674 The *Garra*'s diversification burst in the Sore River was detected in the riverine segment at
675 an altitude range of 1310-1550 m asl, that is within the range of four riverine diversifications of
676 the *Labeobarbus* detected throughout Ethiopian Highlands: 1050-1550 m (Levin et al., 2020).
677 Despite the generally broader elevation gradient (175-2000 m asl - Levin et al., 2020) of the
678 *Labeobarbus* species complex, the diversification bursts were only detected in mid-upper reaches.
679 We believe that a combination of two factors might explain this observation: i) fauna in mid-upper
680 reaches is poorer compared to lower reaches, where a more diversified fauna might have already
681 filled the available ecological niches necessary for an adaptive radiation to unfold; ii) the biotopes
682 are more diverse compared to the most upper reach, that means vacant niches are available.

683 Five endemic, and one introduced non-*Garra* species were recorded in the Sore River in the
684 study area (data of this study). This is an extremely low number compared to more than 110 fish
685 species (Golubtsov & Darkov, 2008, and our data) recorded in the Baro River at Gambella at 440m
686 altitude (our data) to which the drainage of the Sore River belongs with a distance of ~150km
687 between compared localities. The segment of the Sore River where *Garra*'s diversification was
688 detected is rather rich in biotope complexity - pools are alternating pools slow currents, rift areas
689 and rapids (Fig. S6). The depauperated fauna was suggested to provide the ecological opportunities
690 for riverine adaptive radiations similar to the in Southeastern cyprinids of the genus *Poropuntius*
691 (Roberts, 1998) and South America cichlids of the *Crenicichla* due to relaxed competition and
692 vacant niches might have provided ecological opportunities for sympatric speciation by trophic
693 specializations (Burruss et al., 2018).

694 We discovered six new species within the genus *Garra* in the Sore River. Given that the
695 same riverine segment is home for another riverine diversification of fishes represented by four
696 phenotypically diverged ecomorphs of the genus *Labeobarbus* (Levin et al., 2020), we consider
697 the Sore River to being a hot-spot of riverine diversification in the Ethiopian Highlands that
698 requires conservation management. The Ethiopian Highlands are home for several young fish
699 radiations - a large lacustrine diversification among cyprinids (15 species/morphotypes - Mina et
700 al., 1996; Nagelkerke et al., 1994; Nagelkerke et al., 2015) as well as small-sized diversifications
701 of *Garra* (three species – Stiassny & Getahun, 2007) and *Enteromius* (two species - de Graaf,
702 Megens, Samallo, & Sibbing, 2007; Dejen et al., 2002) - all in Lake Tana, and five riverine
703 adaptive radiations of cyprinids each including from four to seven species (Golubtsov, 2010;
704 Golubtsov, Korostelev, & Levin, 2021; Levin et al., 2019; 2020; Mina, Mironovsky, Golubtsov,
705 & Dgebuadze, 1998; current study), highlighting this region's importance as a hotspot for fish
706 speciation that is in need of additional research on ecological speciation processes.

707

708 **Acknowledgements**

709 The study was supported by Russian Science Foundation (grant no. 19-14-00218). We are grateful
710 to all members of Joint Ethiopian-Russian Biological Expedition (JERBE), who participated in
711 our field operations (S.E. Cherenkov, Genanaw Tesfaye, Fekadu Tefera, and I.S. Razgon), and
712 especially to JERBE coordinator Dr. A.A. Darkov for his permanent and invaluable aid. We are
713 thankful to O.N. Artaev for creating a map, S.E. Cherenkov for photographing the fish, A.S.
714 Komarova for data on the gut content as well as to M.V. Mina for critical notes on the first variant
715 of manuscript.

716

717 **Author contributions**

718 BL, ES, PF, NM, AG, and AM designed and contributed to the original concept of the studies.
719 BL and AG collected most of the specimens and related data, BL and NM obtained mtDNA data
720 and prepared DNA libraries for ddRAD, BL conducted morphologic analyses, ES conducted the
721 most of bioinformatics, and BL, ES, PF, and AM finalized the manuscript. All authors partici-
722 pated in project design, and read and approved the final manuscript.

723 **Data availability statement**

724 Morphologic data (body proportions and gut lengths), mtDNA subsets (cytochrome *b*), and gen-
725 otyping files (various sets of SNPs) have been uploaded to
726 Dryad: <https://doi.org/10.5061/dryad.j6q573ndp>
727

728

729 **References**

730

- 731 Aljanabi, S. M., & Martinez, I. (1997). Universal and rapid salt-extraction of high quality genomic
732 DNA for PCR-based techniques. *Nucleic acids research*, 25(22), 4692–4693.
733 doi.org/10.1093/nar/25.22.4692
734 Andrews, S., & Krueger, F. (2010). FastQC. *A quality control tool for high throughput sequence*
735 *data*, 370.
736 Bandelt, H. J., Forster, P., & Röhl, A. (1999). Median-joining networks for inferring intraspecific
737 phylogenies. *Molecular Biology and Evolution*, 16(1), 37–48.

- 738 Banister, K. E. (1984). A subterranean population of *Garra barreimiae* (Teleostei: Cyprinidae)
739 from Oman, with comments on the concept of regressive evolution. *Journal of Natural History*,
740 18(6), 927–938.
- 741 Baumgarten, L., Machado-Schiaffino, G., Henning, F., & Meyer, A. (2015). What big lips are
742 good for: on the adaptive function of repeatedly evolved hypertrophied lips of cichlid fishes.
743 *Biological Journal of the Linnean Society*, 115(2), 448–455. doi.org/10.1111/bij.12502
- 744 Berg, L. S. (1914). *Fishes*. Vol. 3, Ostariophysi, Part. 2. St. Petersburg: Izd. Imper. Akad. Nauk
745 (in Russian).
- 746 Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and
747 powerful approach to multiple testing. *Journal of the Royal statistical society: series B*
748 (*Methodological*), 57(1), 289–300.
- 749 Benvenuti, M., Carnicelli, S., Belluomini, G., Dainelli, N., Di Grazia, S., Ferrari, G. A., ... Kebede,
750 S. (2002). The Ziway–Shala lake basin (main Ethiopian rift, Ethiopia): a revision of basin
751 evolution with special reference to the Late Quaternary. *Journal of African Earth Sciences* 35,
752 247–269.
- 753 Beshera, K. A., Harris, P. M., & Mayden, R. L. (2016). Novel evolutionary lineages in
754 *Labeobarbus* (Cypriniformes; Cyprinidae) based on phylogenetic analyses of mtDNA sequences.
755 *Zootaxa*, 4093(3), 363–381. doi.org/10.11646/zootaxa.4093.3.4
- 756 Braasch, I., & Postlethwait, J. H. (2012). Polyploidy in fish and the teleost genome duplication. In
757 D. E. Soltis (Ed.), *Polyploidy and genome evolution* (pp. 341–383). Berlin, Heidelberg: Springer.
- 758 Brodersen, J., Post, D. M., & Seehausen, O. (2018). Upward adaptive radiation cascades: predator
759 diversification induced by prey diversification. *Trends in Ecology & Evolution*, 33(1), 59–70.
760 doi.org/10.1016/j.tree.2017.09.016
- 761 Burnashev, M. S. (1952). Snow trouts of the Zeravshan River. *Proceedings of the Kishinev State*
762 *University (Biology)*, 4, 111–125 (in Russian).
- 763 Burress, E. D., Piálek, L., Casciotta, J. R., Almirón, A., Tan, M., Armbruster, J. W., & Řičan, O.
764 (2018). Island-and lake-like parallel adaptive radiations replicated in rivers. *Proceedings of the*
765 *Royal Society B: Biological Sciences*, 285(1870), 20171762. doi.org/10.1098/rspb.2017.1762
- 766 Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: an
767 analysis tool set for population genomics. *Molecular Ecology*, 22(11), 3124–3140.
768 doi.org/10.1111/mec.12354
- 769 Chifman, J., & Kubatko, L. (2014). Quartet Inference from SNP Data Under the Coalescent Model.
770 *Bioinformatics*, 30(23), 3317–3324, <https://doi.org/10.1093/bioinformatics/btu530>
- 771 Coad, B. W. (1996). Threatened fishes of the world: *Iranocypris typhlops* Bruun & Kaiser, 1944
772 (Cyprinidae). *Environmental Biology of Fishes*, 46(4), 374. <https://doi.org/10.1007/BF00005015>
- 773 Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... Durbin, R.
774 (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158. doi:
775 10.1093/bioinformatics/btr330 doi.org/10.1093/bioinformatics/btr330
- 776 Data availability: <https://doi.org/10.5061/dryad.j6q573ndp>
- 777 de Graaf, M., Megens, H. J., Samallo, J., & Sibbing, F. (2007). Evolutionary origin of Lake
778 Tana's (Ethiopia) small *Barbus* species: indications of rapid ecological divergence and
779 speciation. *Animal Biology*, 57(1), 39–48. doi.org/10.1163/157075607780002069
- 780 de Graaf, M., Megens, H. J., Samallo, J., & Sibbing, F. (2010). Preliminary insight into the age
781 and origin of the *Labeobarbus* fish species flock from Lake Tana (Ethiopia) using the mtDNA
782 cytochrome *b* gene. *Molecular Phylogenetics and Evolution*, 54(2), 336–343.
783 doi.org/10.1016/j.ympev.2009.10.029
- 784 DeFaveri, J., & Merilä, J. (2013). Evidence for adaptive phenotypic differentiation in Baltic Sea
785 sticklebacks. *Journal of Evolutionary Biology*, 26(8), 1700–1715.
786 <https://doi.org/10.1111/jeb.12168>

- 787 Dejen, E., Rutjes, H. A., De Graaf, M., Nagelkerke, L. A., Osse, J. W., & Sibbing, F. A. (2002).
788 The ‘small barbs’ *Barbus humilis* and *B. trispilopleura* of Lake Tana (Ethiopia): are they ecotypes
789 of the same species?. *Environmental Biology of Fishes*, 65(4), 373–386.
790 doi.org/10.1023/A:1021110721565
- 791 Dibaba, A., Soromessa, T., & Workineh, B., 2019. Carbon stock of the various carbon pools in
792 Gerba-Dima moist Afromontane forest, South-western Ethiopia. *Carbon Balance and*
793 *Management*, 14, 1. <https://doi.org/10.1186/s13021-019-0116-x>
- 794 Dimmick, W. W., Berendzen, P. B., & Golubtsov, A. S. (2001). Genetic comparison of three
795 *Barbus* (Cyprinidae) morphotypes from the Genale River, Ethiopia. *Copeia*, 2001(4), 1123–1129.
796 [doi.org/10.1643/0045-8511\(2001\)001](https://doi.org/10.1643/0045-8511(2001)001)
- 797 Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: summarize analysis results
798 for multiple tools and samples in a single report. *Bioinformatics*, 32(19), 3047–3048.
799 doi.org/10.1093/bioinformatics/btw354
- 800 Feder, J. L., Egan, S. P., & Nosil, P. (2012). The genomics of speciation-with-gene-flow. *Trends*
801 *in Genetics*, 28(7), 342–350. doi.org/10.1016/j.tig.2012.03.009
- 802 Ferguson, D. J., Barnie, T. D., Pyle, D. M., Oppenheimer, C., Yirgu, G., Lewi, E., ... & Hamling,
803 I. (2010). Recent rift-related volcanism in Afar, Ethiopia. *Earth and Planetary Science Letters*,
804 292(3-4), 409–418. doi.org/10.1016/j.epsl.2010.02.010
- 805 Feulner, P. G., Kirschbaum, F., & Tiedemann, R. (2008). Adaptive radiation in the Congo River:
806 an ecological speciation scenario for African weakly electric fish (Teleostei; Mormyridae;
807 *Campylomormyrus*). *Journal of Physiology-Paris*, 102(4-6), 340–346.
808 doi.org/10.1016/j.jphysparis.2008.10.002
- 809 Franchini, P., Monné Parera, D., Kautt, A. F., & Meyer, A. (2017). quaddRAD: a new high-
810 multiplexing and PCR duplicate removal ddRAD protocol produces novel evolutionary insights
811 in a nonradiating cichlid lineage. *Molecular Ecology*, 26(10), 2783–2795.
812 <https://doi.org/10.1111/mec.14077>
- 813 Fricke, R., Eschmeyer, W. N. & Van der Laan, R. (Eds.) 2021. ESCHMEYER'S CATALOG OF
814 FISHES: GENERA, SPECIES, REFERENCES.
815 (<http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp>). Electronic
816 version accessed 22 Feb 2021.
- 817 Fruciano, C., Franchini, P., Raffini, F., Fan, S., & Meyer, A. (2016). Are sympatrically speciating
818 Midas cichlid fish special? Patterns of morphological and genetic variation in the closely related
819 species *Archocentrus centrarchus*. *Ecology and Evolution*, 6(12), 4102–4114.
820 <https://doi.org/10.1002/ece3.2184>
- 821 Fryer, G., & Iles, T. D. (1972). *The Cichlid Fishes of the Great Lakes of Africa*. Neptune City,
822 NY: T.H.F. Publications Inc.
- 823 Geremew, A. (2007). *Taxonomic Revision, Relative Abundance, and Aspects of the Biology of*
824 *some Species of the Genus Garra, Hamilton 1922 (Pisces: Cyprinidae) in Lake Tana, Ethiopia*
825 (Unpublished doctoral dissertation). Addis Ababa University.
- 826 Getahun, A., & Stiassny, M. L. J., 1998. The freshwater biodiversity crisis: the case of the
827 Ethiopian fish fauna. *SINET: Ethiopian Journal of Science*, 21, 207–230.
- 828 Givnish, T. J. (2015). Adaptive radiation versus ‘radiation’ and ‘explosive diversification’: why
829 conceptual distinctions are fundamental to understanding evolution. *New Phytologist*, 207(2),
830 297–303. <https://doi.org/10.1111/nph.13482>
- 831 Glez-Peña, D., Gómez-Blanco, D., Reboiro-Jato, M., Fdez-Riverola, F., & Posada, D. (2010).
832 ALTER: Program-oriented format conversion of DNA and protein alignments. *Nucleic Acids*
833 *Research*, 38(Suppl 2), W14–W18. doi.org/10.1093/nar/gkq321.
- 834 Golubtsov, A. S. (2010). Fish ‘Species Flocks’ in Rivers and Lakes: Sympatric Divergence in Poor
835 Fauna Fish Communities as Particular Modus of Evolution. In D. S. Pavlov, Y. Y. Dgebuadze &
836 M. I. Shatunovsky (Eds.), *Relevant Problems of Contemporary Ichthyology (To 100 Jubilee of G.*
837 *V. Nikolsky)* (pp. 96–123). Moscow: KMK Scientific Press.

- 838 Golubtsov, A. S., Cherenkov, S. E., & Tefera, F. (2012). High morphological diversity of the genus
839 *Garra* in the Sore River (the White Nile Basin, Ethiopia): one more cyprinid species
840 flock? *Journal of Ichthyology*, 52(11), 817–820. doi.org/10.1134/S0032945212110057
- 841 Golubtsov, A. S., & Darkov, A. A. 2008. A review of fish diversity in the main drainage systems
842 of Ethiopia based on the data obtained by 2008. In D. S. Pavlov, Y. Y. Dgebuadze, A. A. Darkov,
843 A. S. Golubtsov & M. V. Mina (Eds.), *Ecological and faunistic studies in Ethiopia, Proceedings*
844 *of jubilee meeting “Joint Ethio-Russian Biological Expedition* (pp. 69–102). Moscow: KMK
845 Scientific Press.
- 846 Golubtsov, A. S., Darkov, A. A., Dgebuadze, Y. Y., & Mina, M. V. (1995). An artificial key to
847 fish species of the Gambela region (the White Nile basin in the limits of Ethiopia). *Joint Ethio-*
848 *Russian Biological Expedition*. Addis Abeba.
- 849 Golubtsov, A. S., Dgebuadze, Y. Y., & Mina, M. V. (2002). Fishes of the Ethiopian Rift Valley.
850 In C. Tudorancea & W. D. Taylor (Eds.), *Ethiopian Rift Valley Lakes. Biology of Inland Waters*
851 *Series* (pp. 167–258). Leiden, The Netherlands: Backhuys Publishers.
- 852 Golubtsov, A. S., Korostelev, N. B., & Levin, B. A. (2021). Monsters with a shortened vertebral
853 column: A population phenomenon in radiating fish *Labeobarbus* (Cyprinidae). *Plos One*, 16(1),
854 e0239639. doi.org/10.1371/journal.pone.0239639
- 855 Hamidan, N., Jackson, M. C., & Britton, J. R. (2016). Diet and trophic niche of the endangered
856 fish *Garra ghorensis* in three Jordanian populations. *Ecology of Freshwater Fish*, 25(3), 455–464.
857 doi.org/10.1111/eff.12226
- 858 Hutchison, W., Fusillo, R., Pyle, D. M., Mather, T. A., Blundy, J. D., Biggs, J., ... & Calvert, A.
859 T. (2016). A pulse of mid-Pleistocene rift volcanism in Ethiopia at the dawn of modern humans.
860 *Nature Communications*, 7(1), 1–12. doi.org/10.1038/ncomms13192
- 861 Hart, R. K., Calver, M. C., & Dickman, C. R. (2002). The index of relative importance: an
862 alternative approach to reducing bias in descriptive studies of animal diets. *Wildlife Research*,
863 29(5), 415–421. doi.org/10.1071/WR02009
- 864 Henning, F., & Meyer, A. (2014). The evolutionary genomics of cichlid fishes: explosive
865 speciation and adaptation in the postgenomic era. *Annual Review of Genomics and Human*
866 *Genetics*, 15, 417–441. doi.org/10.1146/annurev-genom-090413-025412
- 867 Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2:
868 Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution*, 35(2), 518–
869 522. doi.org/10.1093/molbev/msx281
- 870 Hubbs, C. L., & Lagler, K. F. (1958). *Fishes of the Great Lakes region*. Ann Arbor: Univ. Mich.
871 Press.
- 872 Jacobs, A., Carruthers, M., Yurchenko, A., Gordeeva, N. V., Alekseyev, S. S., Hooker, O., ... &
873 Elmer, K. R. (2020). Parallelism in eco-morphology and gene expression despite variable
874 evolutionary and genomic backgrounds in a Holarctic fish. *PLoS Genetics*, 16(4), e1008658.
875 doi.org/10.1371/journal.pgen.1008658
- 876 Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers.
877 *Bioinformatics*, 24(11), 1403–1405. doi.org/10.1093/bioinformatics/btn129
- 878 Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP
879 data. *Bioinformatics*, 27(21), 3070–3071. doi.org/10.1093/bioinformatics/btr521
- 880 Joyce, D. A., Lunt, D. H., Bills, R., Turner, G. F., Katongo, C., Duftner, N., ... & Seehausen, O.
881 (2005). An extant cichlid fish radiation emerged in an extinct Pleistocene lake. *Nature*, 435(7038),
882 90–95. doi.org/10.1038/nature03489
- 883 Junker, J., Rick, J. A., McIntyre, P. B., Kimirei, I., Sweke, E. A., Mosille, J. B., ... & Wagner, C.
884 E. (2020) Structural genomic variation leads to genetic differentiation in Lake Tanganyika's
885 sardines. *Molecular Ecology*, 29: 3277–3298. <https://doi.org/10.1111/mec.15559>
- 886 Irisarri, I., Singh, P., Koblmüller, S., Torres-Dowdall, J., Henning, F., Franchini, P., ... & Meyer,
887 A. (2018). Phylogenomics uncovers early hybridization and adaptive loci shaping the radiation of

- 888 Lake Tanganyika cichlid fishes. *Nature communications*, 9(1), 1–12. doi.org/10.1038/s41467-
889 018-05479-9
- 890 Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., Von Haeseler, A., & Jermini, L. S. (2017).
891 ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature methods*, 14(6),
892 587–589. doi.org/10.1038/nmeth.4285
- 893 Kautt, A. F., Machado-Schiaffino, G., & Meyer, A. (2016). Multispecies outcomes of sympatric
894 speciation after admixture with the source population in two radiations of Nicaraguan crater lake
895 cichlids. *PLoS Genetics*, 12(6), e1006157. doi.org/10.1371/journal.pgen.1006157
- 896 Kautt, A. F., Machado-Schiaffino, G., & Meyer, A. (2018). Lessons from a natural experiment:
897 Allopatric morphological divergence and sympatric diversification in the Midas cichlid species
898 complex are largely influenced by ecology in a deterministic way. *Evolution Letters*, 2(4), 323–
899 340. doi.org/10.1002/evl3.64
- 900 Kautt, A. F., Kratochwil, C. F., Nater, A., Machado-Schiaffino, G., Olave, M., Henning, F., ... &
901 Meyer, A. (2020). Contrasting signatures of genomic divergence during sympatric speciation.
902 *Nature*, 588(7836), 106–111. doi.org/10.1038/s41586-020-2845-0
- 903 Kebede, A., Diekkrüger, B., & Moges, S.A., 2014. Comparative study of a physically based
904 distributed hydrological model versus a conceptual hydrological model for assessment of climate
905 change response in the Upper Nile, Baro-Akobo basin: a case study of the Sore watershed,
906 Ethiopia. *International Journal of River Basin Management*, 12(4), 299–318.
907 <http://dx.doi.org/10.1080/15715124.2014.917315>
- 908 Kocher, T. D. (2004). Adaptive evolution and explosive speciation: the cichlid fish model. *Nature*
909 *Reviews Genetics*, 5(4), 288–298. doi.org/10.1038/nrg1316
- 910 Kottelat, M. (2020). *Ceratogarra*, a genus name for *Garra cambodgiensis* and *G. fasciacauda* and
911 comments on the oral and gular soft anatomy in labeonine fishes (Teleostei: Cyprinidae). *The*
912 *Raffles Bulletin of Zoology Supplement*, 35, 156–178. DOI: 10.26107/RBZ-2020-0049
- 913 Kruckenhauser, L., Haring, E., Seemann, R., & Sattmann, H. (2011). Genetic differentiation
914 between cave and surface-dwelling populations of *Garra barreimiae* (Cyprinidae) in Oman. *BMC*
915 *Evolutionary Biology*, 11(1), 1–15. doi.org/10.1186/1471-2148-11-172
- 916 Krysanov, E. Y., & Golubtsov, A. S. (1993). Karyotypes of three *Garra* species from Ethiopia.
917 *Journal of fish biology*, 42(3), 465–467.
- 918 Lanfear, R., Calcott, B., Ho, S. Y., & Guindon, S. (2012). PartitionFinder: combined selection of
919 partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and*
920 *Evolution*, 29(6), 1695–1701. doi.org/10.1093/molbev/mss020
- 921 Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature*
922 *Methods*, 9(4), 357–359. doi: 10.1038/nmeth.1923 doi.org/10.1038/nmeth.1923
- 923 Lecaudey, L. A., Schliwen, U. K., Osinov, A. G., Taylor, E. B., Bernatchez, L., & Weiss, S. J.
924 (2018). Inferring phylogenetic structure, hybridization and divergence times within Salmoninae
925 (Teleostei: Salmonidae) using RAD-sequencing. *Molecular Phylogenetics and Evolution*, 124,
926 82–99. doi.org/10.1016/j.ympev.2018.02.022
- 927 Leigh, J. W., & Bryant, D. (2015). popart: full-feature software for haplotype network
928 construction. *Methods in Ecology and Evolution*, 6(9), 1110–1116. doi.org/10.1111/2041-
929 210X.12410
- 930 Levin, B.A., Casal-López, M., Simonov, E., Dgebuadze, Y.Y., Mugue, N.S., Tiunov, A.V., ...
931 Golubtsov, A.S. (2019). Adaptive radiation of barbs of the genus *Labeobarbus* (Cyprinidae) in an
932 East African river. *Freshwater Biology*, 64, 1721–1736. <https://doi.org/10.1111/fwb.13364>
- 933 Levin, B.A., Simonov, E., Dgebuadze, Y.Y., Levina, M., & Golubtsov, A.S. (2020). In the rivers:
934 Multiple adaptive radiations of cyprinid fishes (*Labeobarbus*) in Ethiopian Highlands. *Scientific*
935 *reports*, 10(1), 7192. <https://doi.org/10.1038/s41598-020-64350-4>
- 936 Losos, J. B. (2010). Adaptive radiation, ecological opportunity, and evolutionary determinism:
937 American Society of Naturalists EO Wilson Award address. *The American Naturalist*, 175(6),
938 623–639. DOI: 10.1086/652433

- 939 Machado-Schiaffino, G., Henning, F., & Meyer, A. (2014). Species-specific differences in
940 adaptive phenotypic plasticity in an ecologically relevant trophic trait: hypertrophic lips in Midas
941 cichlid fishes. *Evolution*, 68(7), 2086–2091. doi.org/10.1111/evo.12367
- 942 Machado-Schiaffino, G., Kautt, A. F., Torres-Dowdall, J., Baumgarten, L., Henning, F., & Meyer,
943 A. (2017). Incipient speciation driven by hypertrophied lips in Midas cichlid fishes?. *Molecular*
944 *ecology*, 26(8), 2348–2362. doi.org/10.1111/mec.14029
- 945 Malinsky, M., Svoldal, H., Tyers, A. M., Miska, E. A., Genner, M. J., Turner, G. F., & Durbin, R.
946 (2018) Whole-genome sequences of Malawi cichlids reveal multiple radiations interconnected by
947 gene flow. *Nature Ecology & Evolution*, 457, 830. doi.org/10.1038/s41559-018-0717-x
- 948 Malinsky, M., Matschiner, M., & Svoldal, H. (2021). Dsuite-fast D-statistics and related admixture
949 evidence from VCF files. *Molecular Ecology Resources*, 21(2), 584–595.
950 <https://doi.org/10.1111/1755-0998.13265>
- 951 Marques, D. A., Meier, J. I., & Seehausen, O. (2019). A combinatorial view on speciation and
952 adaptive radiation. *Trends in Ecology & Evolution*, 34(6), 531–544.
953 doi.org/10.1016/j.tree.2019.02.008
- 954 Martin, C. H., Cutler, J. S., Friel, J. P., Denning Touokong, C., Coop, G., & Wainwright, P. C.
955 (2015). Complex histories of repeated gene flow in Cameroon crater lake cichlids cast doubt on
956 one of the clearest examples of sympatric speciation. *Evolution*, 69(6), 1406–1422.
957 doi.org/10.1111/evo.12674
- 958 Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads.
959 *EMBnet journal*, 17(1), 10. doi: 10.14806/ej.17.1.200 doi.org/10.14806/ej.17.1.200
- 960 Matschiner, M., Hanel, R., & Salzburger, W. (2011). On the origin and trigger of the notothenioid
961 adaptive radiation. *PLoS One*, 6(4), e18911. doi.org/10.1371/journal.pone.0018911
- 962 Matthes, H. (1963). A Comparative Study of the Feeding Mechanisms of Some African Cyprinidae
963 (Pisces, Cypriniformes). *Bijdragen tot de Dierkunde*, 33(1), 3–24.
- 964 McKinnon, J. S., & Rundle, H. D. (2002). Speciation in nature: the threespine stickleback model
965 systems. *Trends in Ecology & Evolution*, 17(10), 480–488. [doi.org/10.1016/S0169-
966 5347\(02\)02579-X](https://doi.org/10.1016/S0169-5347(02)02579-X)
- 967 Mège, D., Purcell, P., Pochat, S., & Guidat, T. (2015). The landscape and landforms of the Ogaden,
968 Southeast Ethiopia. In P. Billi (Ed.), *Landscapes and landforms of*
969 *Ethiopia* (pp. 323–348). Dordrecht, The Netherlands: Springer.
- 970 Meier, J. I., Marques, D. A., Mwaiko, S., Wagner, C. E., Excoffier, L., & Seehausen, O. (2017).
971 Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nature communications*, 8(1),
972 1–11. doi.org/10.1038/ncomms14363
- 973 Menon, A. G. K. (1964). *Monograph of the cyprinid fishes of the genus Garra Hamilton* (Vol. 14).
974 Government of India.
- 975 Melaku, S., Abebe Getahun, A., & Wakjira, M. (2017). Population aspects of fishes in Geba and
976 Sor rivers, White Nile System in Ethiopia, East Africa. *International Journal of Biodiversity*,
977 2017, 1252604. <https://doi.org/10.1155/2017/1252604>
- 978 Melnik, N. O., Markevich, G. N., Taylor, E. B., Loktyushkin, A. V., & Esin, E. V. (2020).
979 Evidence for divergence between sympatric stone charr and Dolly Varden along unique
980 environmental gradients in Kamchatka. *Journal of Zoological Systematics and Evolutionary*
981 *Research*, 58(4), 1135–1150. doi.org/10.1111/jzs.12367
- 982 Meyer, A., Kocher, T. D., Basasibwaki, P., & Wilson, A. C. (1990). Monophyletic origin of Lake
983 Victoria cichlid fishes suggested by mitochondrial DNA sequences. *Nature*, 347(6293), 550–553.
- 984 Mina, M. V., Levin, B. A., & Mironovsky, A. N. (2005). On the possibility of using character
985 estimates obtained by different operators in morphometric studies of fish. *Journal of Ichthyology*,
986 45(4), 284–294.
- 987 Mina, M. V., Mironovsky, A. N., & Dgebuadze, Y. (1996). Lake Tana large barbs: phenetics,
988 growth and diversification. *Journal of Fish Biology*, 48(3), 383–404.

- 989 Mina, M. V., Mironovsky, A. N., Golubtsov, A. S., & Dgebuadze, Y. Y. (1998). II. Morphological
990 diversity of “large barbs”; from Lake Tana and neighbouring areas: Homoplasies or
991 synapomorphies?. *Italian Journal of Zoology*, 65(S1), 9–14.
- 992 Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A.,
993 & Lanfear, R. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference
994 in the genomic era. *Molecular Biology and Evolution*, 37(5), 1530–1534.
995 <https://doi.org/10.1093/molbev/msaa015>
- 996 Mironovsky, A. N., Mina, M. V., & Dgebuadze, Y. Y. (2019). Large African Barbs with
997 Hypertrophied Lips and their Relationship with Generalized Forms of Species of the Genus *Barbus*
998 (*Labeobarbus auctorum*). *Journal of Ichthyology*, 59(3), 327–335.
999 doi.org/10.1134/S0032945219030111
- 1000 Moritz, T., El Dayem, Z.N., Abdallah, M.A., & Neumann, D. (2019). New and rare records of
1001 fishes from the White Nile in the Republic of the Sudan. *Cybium*, 43, 137–151.
1002 <https://doi.org/10.26028/cybium/2019-423-011>
- 1003 Mousavi-Sabet, H., & Eagderi, S. (2016). *Garra lorestanensis*, a new cave fish from the Tigris
1004 River drainage with remarks on the subterranean fishes in Iran (Teleostei: Cyprinidae). *FishTaxa*,
1005 1(1), 45–54. <http://dx.doi.org/10.7508/jft.2016.01.006>
- 1006 Muschick, M., Nosil, P., Roesti, M., Dittmann, M. T., Harmon, L., & Salzburger, W. (2014).
1007 Testing the stages model in the adaptive radiation of cichlid fishes in East African Lake
1008 Tanganyika. *Proceedings of the Royal Society B: Biological Sciences*, 281(1795), 20140605.
1009 doi.org/10.1098/rspb.2014.0605
- 1010 Nagelkerke, L. (1997). *The barbs of Lake Tana, Ethiopia: morphological diversity and its*
1011 *implications for taxonomy, trophic resource partitioning, and fisheries* (Unpublished doctoral
1012 dissertation). Agricultural University of Wageningen.
- 1013 Nagelkerke, L. A., Sibbing, F. A., van den Boogaart, J. G., Lammens, E. H., & Osse, J. W. (1994).
1014 The barbs (*Barbus* spp.) of Lake Tana: a forgotten species flock?. *Environmental Biology of*
1015 *Fishes*, 39(1), 1–22.
- 1016 Nagelkerke, L. A. J., & Sibbing, F. A. (1997). A revision of the large barbs (*Barbus* spp.,
1017 Cyprinidae, Teleostei) of Lake Tana, Ethiopia, with a description of seven new species. In: *The*
1018 *barbs of Lake Tana, Ethiopia: morphological diversity and its implications for taxonomy, trophic*
1019 *resource partitioning, and fisheries* (pp. 105–170). (Unpublished doctoral dissertation).
1020 Agricultural University of Wageningen.
- 1021 Nagelkerke, L. A. J., Leon-Kloosterziel, K. M., Megens, H. J., De Graaf, M., Diekmann, O. E., &
1022 Sibbing, F. A. (2015). Shallow genetic divergence and species delineations in the endemic
1023 *Labeobarbus* species flock of Lake Tana, Ethiopia. *Journal of Fish Biology*, 87(5), 1191–1208.
1024 doi.org/10.1111/jfb.12779
- 1025 Neumann, D., Obermaier, H., & Moritz, T. 2016. Annotated checklist for fishes of the Main Nile
1026 Basin in the Sudan and Egypt based on recent specimen records (2006-2015). *Cybium*, 40: 287–
1027 317. doi.org/10.26028/cybium/2016-404-004
- 1028 Nguyen, L. T., Schmidt, H. A., Von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: a fast and
1029 effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology*
1030 *and Evolution*, 32(1), 268–274. doi.org/10.1093/molbev/msu300
- 1031 Østbye, K., Amundsen, P. A., Bernatchez, L., Klemetsen, A., Knudsen, R., Kristoffersen, R., ... &
1032 Hindar, K. (2006). Parallel evolution of ecomorphological traits in the European whitefish
1033 *Coregonus lavaretus* (L.) species complex during postglacial times. *Molecular Ecology*, 15(13),
1034 3983–4001. doi.org/10.1111/j.1365-294X.2006.03062.x
- 1035 Palumbi, S. R. (1996). Nucleic acids II: The polymerase chain reaction. In D. M. Hillis, C. Moritz
1036 & B. K. Mable (Eds.), *Molecular systematics* (pp. 205–247). Sunderland, MA: Sinauer
1037 Associates.
- 1038 Paris, J. R., Stevens, J. R., & Catchen, J. M. (2017). Lost in parameter space: a road map for stacks.
1039 *Methods in Ecology and Evolution*, 8(10), 1360–1373. doi.org/10.1111/2041-210X.12775

- 1040 Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., ... & Reich, D. (2012).
1041 Ancient admixture in human history. *Genetics*, 192(3), 1065–1093.
1042 doi.org/10.1534/genetics.112.145037
- 1043 Peichel, C. L., Nereng, K. S., Ohgi, K. A., Cole, B. L., Colosimo, P. F., Buerkle, C. A., ... &
1044 Kingsley, D. M. (2001). The genetic architecture of divergence between threespine stickleback
1045 species. *Nature*, 414(6866), 901–905. doi.org/10.1038/414901a
- 1046 Perdices, A., & Doadrio, I. (2001). The molecular systematics and biogeography of the European
1047 cobitids based on mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, 19(3),
1048 468–478. doi.org/10.1006/mpev.2000.0900
- 1049 Piálek, L., Říčan, O., Casciotta, J., Almirón, A., & Zrzavý, J. (2012). Multilocus phylogeny of
1050 *Crenicichla* (Teleostei: Cichlidae), with biogeography of the *C. lacustris* group: species flocks as
1051 a model for sympatric speciation in rivers. *Molecular Phylogenetics and Evolution*, 62(1), 46–61.
1052 doi.org/10.1016/j.ympev.2011.09.006
- 1053 Præbel, K., Knudsen, R., Siwertsson, A., Karhunen, M., Kahilainen, K. K., Ovaskainen, O., ... &
1054 Amundsen, P. A. (2013). Ecological speciation in postglacial European whitefish: rapid adaptive
1055 radiations into the littoral, pelagic, and profundal lake habitats. *Ecology and Evolution*, 3(15),
1056 4970–4986. doi.org/10.1002/ece3.867
- 1057 Prave, A. R., Bates, C. R., Donaldson, C. H., Toland, H., Condon, D. J., Mark, D., & Raub, T. D.
1058 (2016). Geology and geochronology of the Tana Basin, Ethiopia: LIP volcanism, super eruptions
1059 and Eocene–Oligocene environmental change. *Earth and Planetary Science Letters*, 443, 1–8.
1060 doi.org/10.1016/j.epsl.2016.03.009
- 1061 Prokofiev, A. M. & Golubtsov A. S. (2013). Revision of the loach genus *Afronemacheilus*
1062 (Teleostei: Balitoridae: Nemacheilinae) with description of a new species from the Omo-Turkana
1063 basin, Ethiopia. *Ichthyological Exploration of Freshwaters*, 24, 1–14.
- 1064 Puebla, O. (2009). Ecological speciation in marine v. freshwater fishes. *Journal of Fish Biology*,
1065 75(5), 960–996. doi.org/10.1111/j.1095-8649.2009.02358.x
- 1066 Rambaut, A. (2014). FigTree 1.4.2 software. Institute of Evolutionary Biology, Univ. Edinburgh.
- 1067 Rambaut, A., Ho, S. Y., Drummond, A. J., & Shapiro, B. (2009). Accommodating the effect of
1068 ancient DNA damage on inferences of demographic histories. *Molecular Biology and Evolution*,
1069 26(2), 245–248. doi.org/10.1093/molbev/msn256
- 1070 Rambaut, A., Suchard, M. A., Xie, D. & Drummond, A. J. (2014). Tracer v1.6. Retrieved from
1071 <http://beast.bio.ed.ac.uk/Tracer>
- 1072 Reich, D., Thangaraj, K., Patterson, N., Price, A. L., & Singh, L. (2009). Reconstructing Indian
1073 population history. *Nature*, 461(7263), 489–494. doi.org/10.1038/nature08365
- 1074 Ribbink, A. J., Marsh, A. C., Marsh, B. A., & Sharp, B. J. (1983). The zoogeography, ecology and
1075 taxonomy of the genus *Labeotropheus* Ahl, 1927, of Lake Malawi (Pisces: Cichlidae). *Zoological*
1076 *Journal of the Linnean Society*, 79(3), 223–243.
- 1077 Richards, E. J., Servedio, M. R., & Martin, C. H. (2019). Searching for sympatric speciation in the
1078 genomic era. *BioEssays*, 41(7), 1900047. doi.org/10.1002/bies.201900047
- 1079 Roberts, T. R. (1998). Review of the tropical Asian cyprinid fish genus *Poropuntius*, with
1080 descriptions of new species and trophic morphs. *Natural History Bulletin of the Siam Society*,
1081 46(1), 105–135.
- 1082 Roberts, T. R., & Khaironizam, M. Z. (2008). Trophic polymorphism in the Malaysian fish
1083 *Neolissochilus soroides* and other old world barbs (Teleostei, Cyprinidae). *Natural History*
1084 *Bulletin of the Siam Society*, 56, 25–53.
- 1085 Rougeux, C., Bernatchez, L., & Gagnaire, P. A. (2017). Modeling the multiple facets of speciation-
1086 with-gene-flow toward inferring the divergence history of lake whitefish species pairs (*Coregonus*
1087 *clupeaformis*). *Genome Biology and Evolution*, 9(8), 2057–2074. doi.org/10.1093/gbe/evx150
- 1088 Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S., ... &
1089 Huelsenbeck, J. P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model

- 1090 choice across a large model space. *Systematic Biology*, 61(3), 539–542.
1091 doi.org/10.1093/sysbio/sys029
- 1092 Salzburger, W., Mack, T., Verheyen, E., & Meyer, A. (2005). Out of Tanganyika: genesis,
1093 explosive speciation, key-innovations and phylogeography of the haplochromine cichlid fishes.
1094 *BMC Evolutionary Biology*, 5(1), 1–15. doi.org/10.1186/1471-2148-5-17
- 1095 Schluter, D. (2000). *The ecology of adaptive radiation*. New York: Oxford University Press.
- 1096 Schwarzer, J., Misof, B., Ifuta, S. N., & Schlieven, U. K. (2011). Time and origin of cichlid
1097 colonization of the lower Congo rapids. *Plos One*, 6(7), e22380.
1098 doi.org/10.1371/journal.pone.0022380
- 1099 Seehausen, O. (2000). Explosive speciation rates and unusual species richness in haplochromine
1100 cichlid fishes: effects of sexual selection. *Advances in Ecological Research*, 31, 237–274.
1101 [doi.org/10.1016/S0065-2504\(00\)31015-7](https://doi.org/10.1016/S0065-2504(00)31015-7)
- 1102 Seehausen, O. (2006). African cichlid fish: a model system in adaptive radiation
1103 research. *Proceedings of the Royal Society B: Biological Sciences*, 273(1597), 1987–1998.
1104 doi.org/10.1098/rspb.2006.3539
- 1105 Seehausen, O., & Wagner, C. E. (2014). Speciation in freshwater fishes. *Annual Review of*
1106 *Ecology, Evolution, and Systematics*, 45, 621–651. doi.org/10.1146/annurev-ecolsys-120213-091818
- 1107
- 1108 Segherloo, I. H., Normandeau, E., Benestan, L., Rougeux, C., Coté, G., Moore, J. S., ... &
1109 Bernatchez, L. (2018). Genetic and morphological support for possible sympatric origin of fish
1110 from subterranean habitats. *Scientific Reports*, 8(1), 1–13. doi.org/10.1038/s41598-018-20666-w
- 1111 Selz, O. M., & Seehausen, O. (2019). Interspecific hybridization can generate functional novelty
1112 in cichlid fish. *Proceedings of the Royal Society B*, 286(1913), 20191621.
1113 doi.org/10.1098/rspb.2019.1621
- 1114 Sibbing, F. A., Nagelkerke, L. A., Stet, R. J., & Osse, J. W. (1998). Speciation of endemic Lake
1115 Tana barbs (Cyprinidae, Ethiopia) driven by trophic resource partitioning; a molecular and
1116 ecomorphological approach. *Aquatic Ecology*, 32(3), 217–227.
- 1117 Skúlason, S. (1999). Sympatric morphs, populations and speciation in freshwater fish with
1118 emphasis on arctic charr. In A. Magurran & R. M. May (Eds.), *Evolution of biological diversity*
1119 (pp. 71–92). New York: Oxford University Press.
- 1120 Smadja, C. M., & Butlin, R. K. (2011). A framework for comparing processes of speciation in the
1121 presence of gene flow. *Molecular Ecology*, 20(24), 5123–5140. doi.org/10.1111/j.1365-294X.2011.05350.x
- 1122
- 1123 Stiassny, M.L.J. & Abebe Getahun. 2007. An overview of labeonin relationships and the
1124 phylogenetic placement of the Afro-Asian genus *Garra* Hamilton, 1822 (Teleostei: Cyprinidae),
1125 with the description of five new species of *Garra* from Ethiopia, and a key to all African species.
1126 *Zoological Journal of the Linnean Society*, 150, 41–83. doi.org/10.1111/j.1096-3642.2007.00281.x
- 1127
- 1128 Stewart, K. M., & Murray, A. M. (2017). Biogeographic implications of fossil fishes from the
1129 Awash River, Ethiopia. *Journal of Vertebrate Paleontology*, 37(1), e1269115.
1130 doi.org/10.1080/02724634.2017.1269115
- 1131 Sturmbauer, C. (1998). Explosive speciation in cichlid fishes of the African Great Lakes: a
1132 dynamic model of adaptive radiation. *Journal of Fish Biology*, 53, 18–36. doi.org/10.1111/j.1095-8649.1998.tb01015.x
- 1133
- 1134 Swofford, D. L. 2003. PAUP*. Phylogenetic Analysis Using Parsimony (and Other Methods).
1135 Version 4. Sinauer Associates, Sunderland, Massachusetts.
- 1136 Taylor, E. B. (1999). Species pairs of north temperate freshwater fishes: evolution, taxonomy, and
1137 conservation. *Reviews in Fish Biology and Fisheries*, 9(4), 299–324.
1138 doi.org/10.1023/A:1008955229420

- 1139 Tamura, K., Stecher, G., Peterson, D., Filipiński, A., & Kumar, S. (2013). MEGA6: molecular
1140 evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, 30(12), 2725–2729.
1141 doi.org/10.1093/molbev/mst197
- 1142 Tang, Q., Getahun, A., & Liu, H. (2009). Multiple in-to-Africa dispersals of labeonin fishes
1143 (Teleostei: Cyprinidae) revealed by molecular phylogenetic analysis. *Hydrobiologia*, 632(1), 261–
1144 271. doi.org/10.1007/s10750-009-9848-z
- 1145 Terashima, A. (1984). Three new species of the cyprinid genus *Schizothorax* from Lake Rara,
1146 northwestern Nepal. *Japanese Journal of Ichthyology*, 31(2), 122–135.
- 1147 Terekhanova, N. V., Logacheva, M. D., Penin, A. A., Neretina, T. V., Barmintseva, A. E., Bazykin,
1148 G. A., ... & Mogue, N. S. (2014). Fast evolution from precast bricks: genomics of young freshwater
1149 populations of threespine stickleback *Gasterosteus aculeatus*. *PLoS Genetics*, 10(10), e1004696.
1150 doi.org/10.1371/journal.pgen.1004696
- 1151 Verheyen, E., Salzburger, W., Snoeks, J., & Meyer, A. (2003). Origin of the superflock of cichlid
1152 fishes from Lake Victoria, East Africa. *Science*, 300(5617), 325–329.
1153 DOI:10.1126/science.1080699
- 1154 Verity, R., & Nichols, R. A. (2016). Estimating the Number of Subpopulations (K) in Structured
1155 Populations. *Genetics*, 203(4), 1827–1839. doi:10.1534/genetics.115.180992
- 1156 Vreven, E. J., Musschoot, T., Snoeks, J., & Schlieuwen, U. K. (2016). The African hexaploid Torini
1157 (Cypriniformes: Cyprinidae): review of a tumultuous history. *Zoological Journal of the Linnean
1158 Society*, 177(2), 231–305. doi.org/10.1111/zoj.12366
- 1159 Wagner, C. E., Harmon, L. J., & Seehausen, O. (2012). Ecological opportunity and sexual selection
1160 together predict adaptive radiation. *Nature*, 487(7407), 366–369. doi.org/10.1038/nature11144
- 1161 Wagner, C. E., McIntyre, P. B., Buels, K. S., Gilbert, D. M., & Michel, E. (2009). Diet predicts
1162 intestine length in Lake Tanganyika's cichlid fishes. *Functional Ecology*, 23(6), 1122–1131.
1163 doi.org/10.1111/j.1365-2435.2009.01589.x
- 1164 Whiteley, A. R. (2007). Trophic polymorphism in a riverine fish: morphological, dietary, and
1165 genetic analysis of mountain whitefish. *Biological Journal of the Linnean Society*, 92(2), 253–267.
1166 doi.org/10.1111/j.1095-8312.2007.00845.x
- 1167 Yang, L., Arunachalam, M., Sado, T., Levin, B. A., Golubtsov, A. S., Freyhof, J., ... & He, S.
1168 (2012). Molecular phylogeny of the cyprinid tribe Labeonini (Teleostei: Cypriniformes).
1169 *Molecular Phylogenetics and Evolution*, 65(2), 362–379. doi.org/10.1016/j.ympev.2012.06.007
- 1170 Zandonà, E., Auer, S. K., Kilham, S. S., & Reznick, D. N. (2015). Contrasting population and diet
1171 influences on gut length of an omnivorous tropical fish, the Trinidadian guppy (*Poecilia
1172 reticulata*). *PLoS One*, 10(9), e0136079. doi.org/10.1371/journal.pone.0136079
- 1173 Zheng, X. M., & Ge, S. (2010). Ecological divergence in the presence of gene flow in two closely
1174 related *Oryza* species (*Oryza rufipogon* and *O. nivara*). *Molecular Ecology*, 19(12), 2439–2454.
1175 doi.org/10.1111/j.1365-294X.2010.04674.x
- 1176 <https://briancoad.com>
- 1177 <https://en.wikipedia.org/wiki/Tiber>
- 1178
- 1179
- 1180
- 1181
- 1182
- 1183
- 1184
- 1185
- 1186
- 1187
- 1188
- 1189

1190
1191
1192
1193
1194
1195
1196
1197
1198

Table 1. Common names of the six ecomorphs of African *Garra* from the Sore River, and the preliminary qualitative descriptions used in the field to identify each form.

Name used in the text	Basal description
No. 1, ‘generalized’	Well-developed round-shaped gular disc of type C with free posterior margin (disc classification follows Stiassny & Getahun, 2007). Body shape is generalized for <i>Garra</i> .
No. 2, ‘stream-lined’	Slender stream-line body with slim caudal peduncle and increased pectoral fins. Disc of type C.
No. 3, ‘narrow-mouth’	Disc is reduced in size, elongated, oval-shaped (closer to type A). Narrow mouth often with groove on lower jaw.
No. 4, ‘wide-mouth’	Disc is reduced in size, triangle-shaped. Wide mouth with significantly enlarged labellum (sensu Kottelat, 2020). Disc of type B in degree of development.
No. 5, ‘predator’	Completely or almost completely reduced gular disc (type A when presented). Wide head and mouth. This ecomorph achieves larger size compared to others. Largest individuals have nuchal hunch and almost terminal mouth with a bony projection on the lower jaw and matching incision on the upper jaw.
No. 6, ‘thick-lipped’	Greatly developed lips, referred to as ‘rubber lips’ (Matthes, 1963). Intermediate lobe of the lower lip is ball-shaped and unattached. Gular disc is greatly reduced, oval-shaped (type A). Only two individuals recorded.

1199
1200
1201
1202

Table 2. DNA and morphology sample numbers of *Garra* ecomorphs from the Sore River.

Ecomorphs	Morphology		mtDNA	RAD-seq
	Measurements	Gut length and diet		
1	27	18	27	22
2	17	7	19	13
3	19	13	18	11
4	20	10	17	13
5	15	14	24	11
6	2	-	2	2
Intermediate phenotype	6	-	5	5
Total	106	62	112	77

1203

1204

1205

1206

1207 Table 3. Results of Patterson's D statistic (ABBA-BABA test) and f_4 -ratio test on selected genetic

1208 clusters of *Garra* from the Sore River.

P1	P2	P3	D statistic	Z-score	p-value	f4-ratio	BBAA	ABBA	BABA
4	3	6	0.1176	5.3829	<0.0001	0.1128	227.5	235.0	185.5
2b	3	5	0.0650	3.1078	0.0009	0.4226	253.5	246.5	216.4
2b	6	3	0.0646	2.3475	0.0095	0.2854	215.6	217.3	190.9
4	3	2b	0.0624	3.8143	<0.0001	0.1237	264.6	241.4	213.0
4	3	5	0.0492	3.6742	0.0001	0.3277	276.2	247.4	224.2
2b	6	5	0.0327	1.4755	0.0700	0.2051	248.6	203.4	190.5
4	6	5	0.0304	1.5315	0.0628	0.2330	224.5	226.5	213.2
6	3	5	0.0199	0.9380	0.1741	0.1641	244.2	204.7	196.8
2b	4	5	0.0178	1.0774	0.1406	0.1134	245.9	246.3	237.7
2b	6	4	0.0040	0.1592	0.4368	0.0151	244.6	197.8	196.3

1209

1210

1211

1212

1213

1214

1215

1216

1217

1218

1219

1220

1221

1222

1223 Table 4. Summary of the ecomorphs' genetic diversity indices averaged over 89 070 loci (both
1224 variant and fixed).

Ecomorphs *	No. of private alleles, Np	No. of polymorphic loci, %	Heterozygosity		Coefficient of inbreeding (Fis) ± SE	Nucleotide diversity (Pi) ± SE
			Observed (Ho) ± SE	Expected (He) ± SE		
1a	19	0.42	0.00128± 0.00008	0.00116± 0.00007	-0.00014±0.0015	0.00121± 0.00007
1b	18	0.40	0.00128± 0.00008	0.00113± 0.00007	-0.00019±0.0011	0.00119± 0.00007
2a	27	0.41	0.00124± 0.00008	0.00114± 0.00007	-0.00007±0.0012	0.00120± 0.00007
2b	9	0.24	0.00104± 0.00008	0.00079± 0.00006	-0.00023±0.0012	0.00091± 0.00007
3	20	0.43	0.00127± 0.00008	0.00107± 0.00006	-0.00037±0.0013	0.00111± 0.00007
4	38	0.43	0.00109± 0.00007	0.001± 0.00006	-0.00008±0.0015	0.00104± 0.00006
5	33	0.44	0.00126± 0.00008	0.00115± 0.00007	-0.00011±0.0019	0.00120± 0.00007
6	3	0.10	0.00058± 0.00007	0.0004± 0.0000	-0.00006±0.0004	0.00054± 0.00006

1225 * - letters 'a' and 'b' assign genetic lineages within ecomorphs 1 and 2.

1226

1227

1228

1229

1230

1231

1232

1233

1234

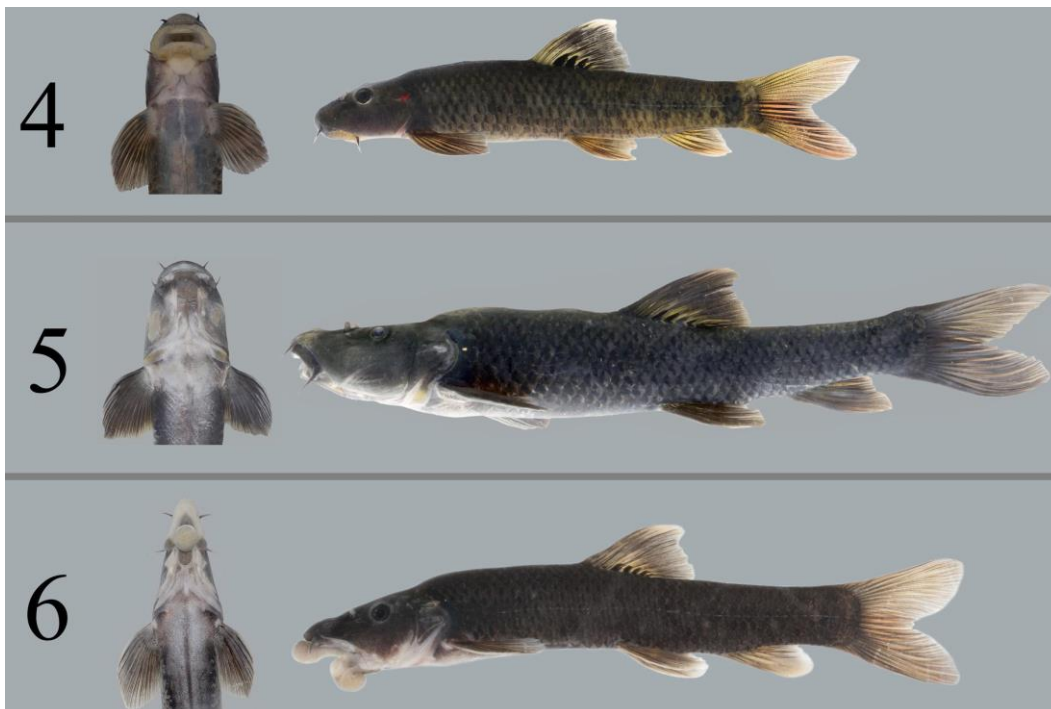


1235

1236 Fig. 1A. *Garra* ecomorphs 1-3 from the Sore River: 1 - 'generalized': 136 mm SL; 2 -
1237 'stream-lined': 99 mm SL; 3 - 'narrow-mouth': 100 mm SL.

1238

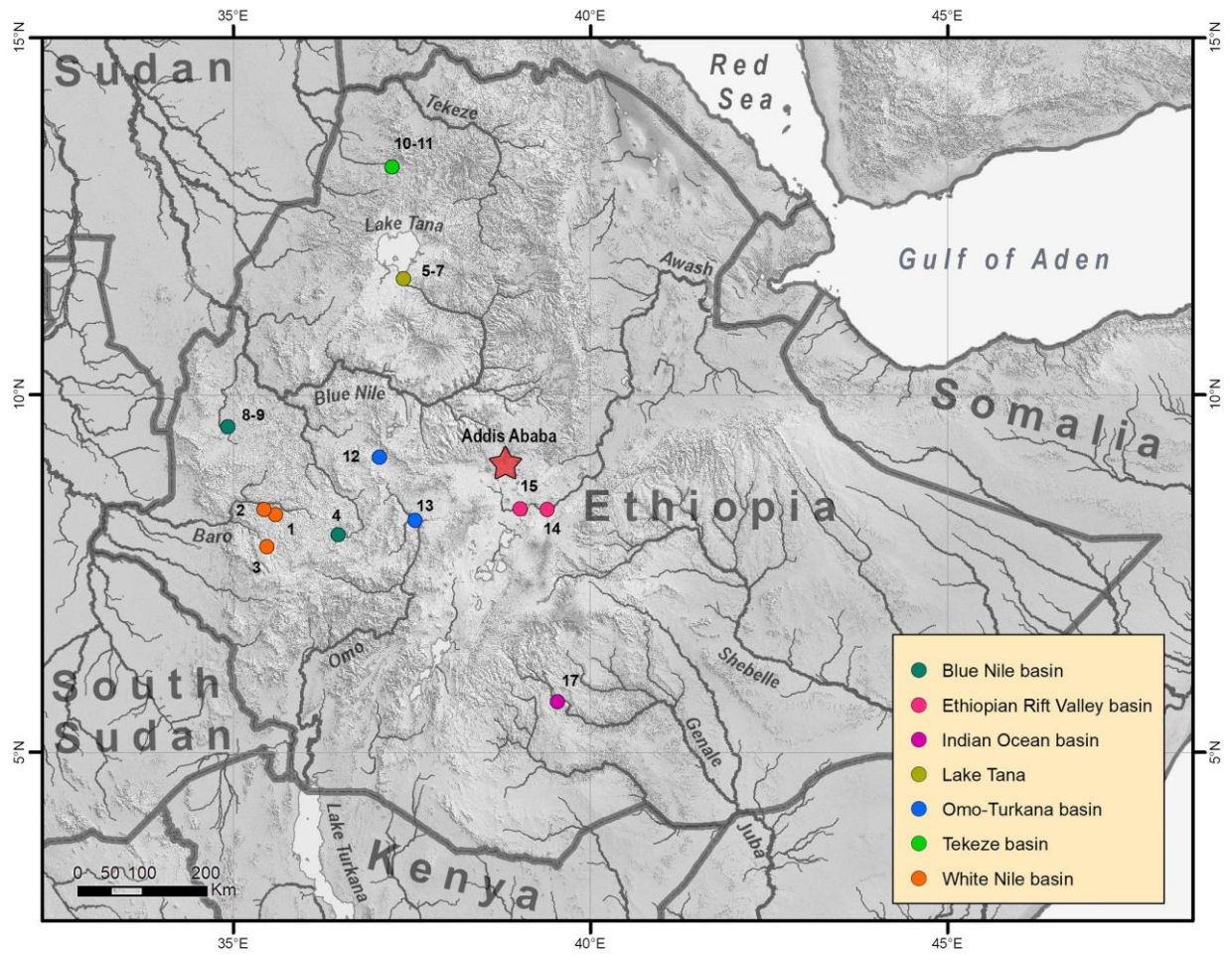
1239



1240

1241 Fig. 1B. *Garra* ecomorphs 4-6 from the Sore River: 4 - 'wide-mouth': 100 mm SL; 5 - 'predator':
1242 193 mm SL; 6 - 'thick-lipped': 128 mm SL.

1243
1244
1245
1246

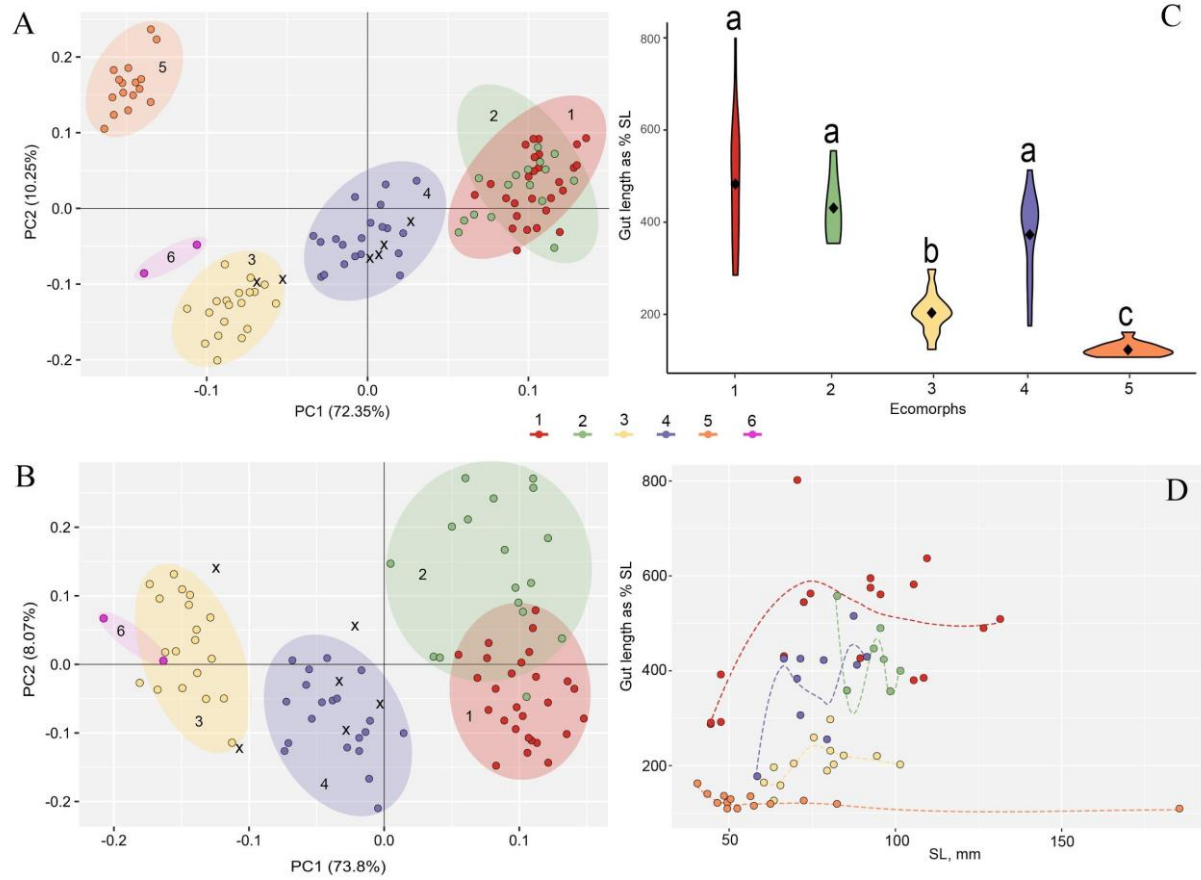


1247

1248 Fig. 2. Sampling sites of *Garra* in Ethiopian Highlands and Ethiopian Rift Valley; loc. 1-2 are in
1249 the Sore River.

1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266

1267
1268
1269
1270

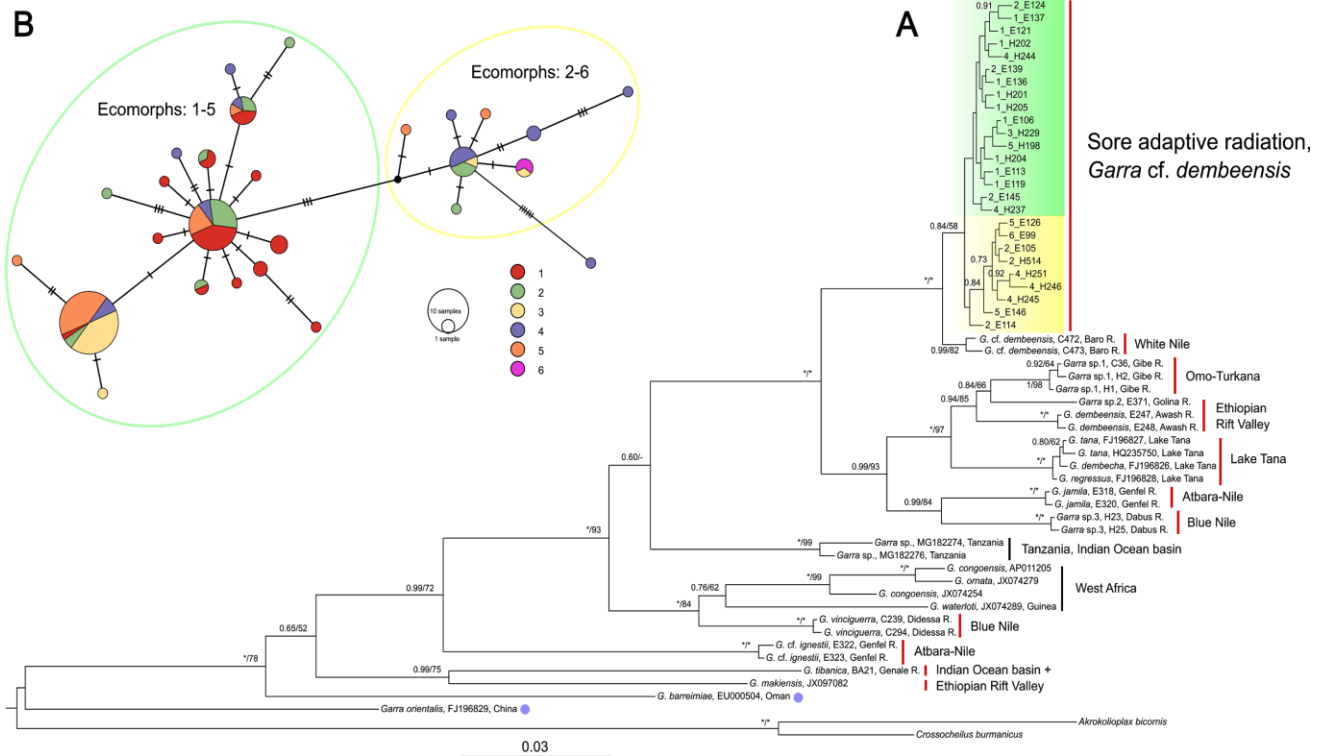


1271

1272 Fig. 3. (A) PCA of body and head proportions of six sympatric ecomorphs from the Sore River
1273 (n=107); (B) PCA of body and head proportions of five sympatric ecomorphs from the Sore River
1274 (n=90) excluding the most divergent sample, ecomorph 5. X designates intermediate phenotypes;
1275 (C) Gut length of five sympatric *Garra* ecomorphs from the Sore River represented as violin
1276 boxplots. Middle points are the means, and the box show the range respectively, samples are
1277 combined and each contains between 7 (ecomorph 2) and 18 (ecomorph 1) individuals, for a total
1278 of 62 individuals. Different lowercase letters above the boxplots indicate significant differences
1279 between ecomorphs ($p < 0.05$, Kruskal-Wallis test with BH adjustment of p -value); (D)
1280 Dependence of gut length on body length in five *Garra* ecomorphs from the Sore River with
1281 smooth local regression lines (Loess regression).

1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292

1293
1294
1295



1296

1297 Fig. 4. (A) Consensus tree of relationships among the Ethiopian *Garra* from all main drainages
1298 based on *cytb* sequences. Bayesian posterior probabilities (before slash) from BI analysis and
1299 bootstrap values from ML analysis (after slash) above 0.5/50 are shown; asterisks represent
1300 posterior probabilities/bootstrap values of 1/100. Scale bar and branch lengths provide the
1301 expected substitutions per site. The green and yellow colors highlight two branches of *Garra* in
1302 the Sore River. (B) Median-joining haplotype network of the *Garra* from the Sore River, based on
1303 107 *cytb* sequences (989 bp length). ‘Green’ haplogroup includes ecomorphs 1-5, while ‘yellow’
1304 haplogroup includes ecomorphs 2-6. Black dots represent hypothetical intermediate haplotypes.

1305

1306

1307

1308

1309

1310

1311

1312

1313

1314

1315

1316

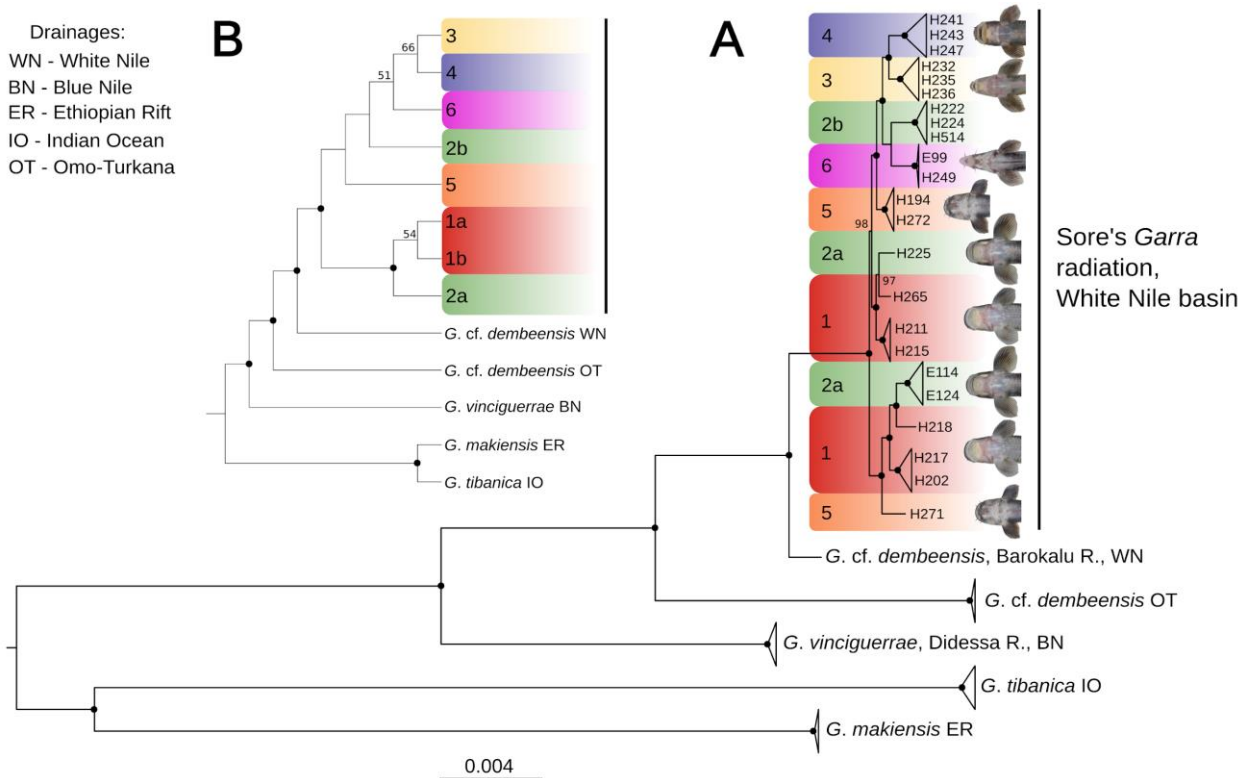
1317

1318

1319

1320

1321
1322
1323
1324
1325

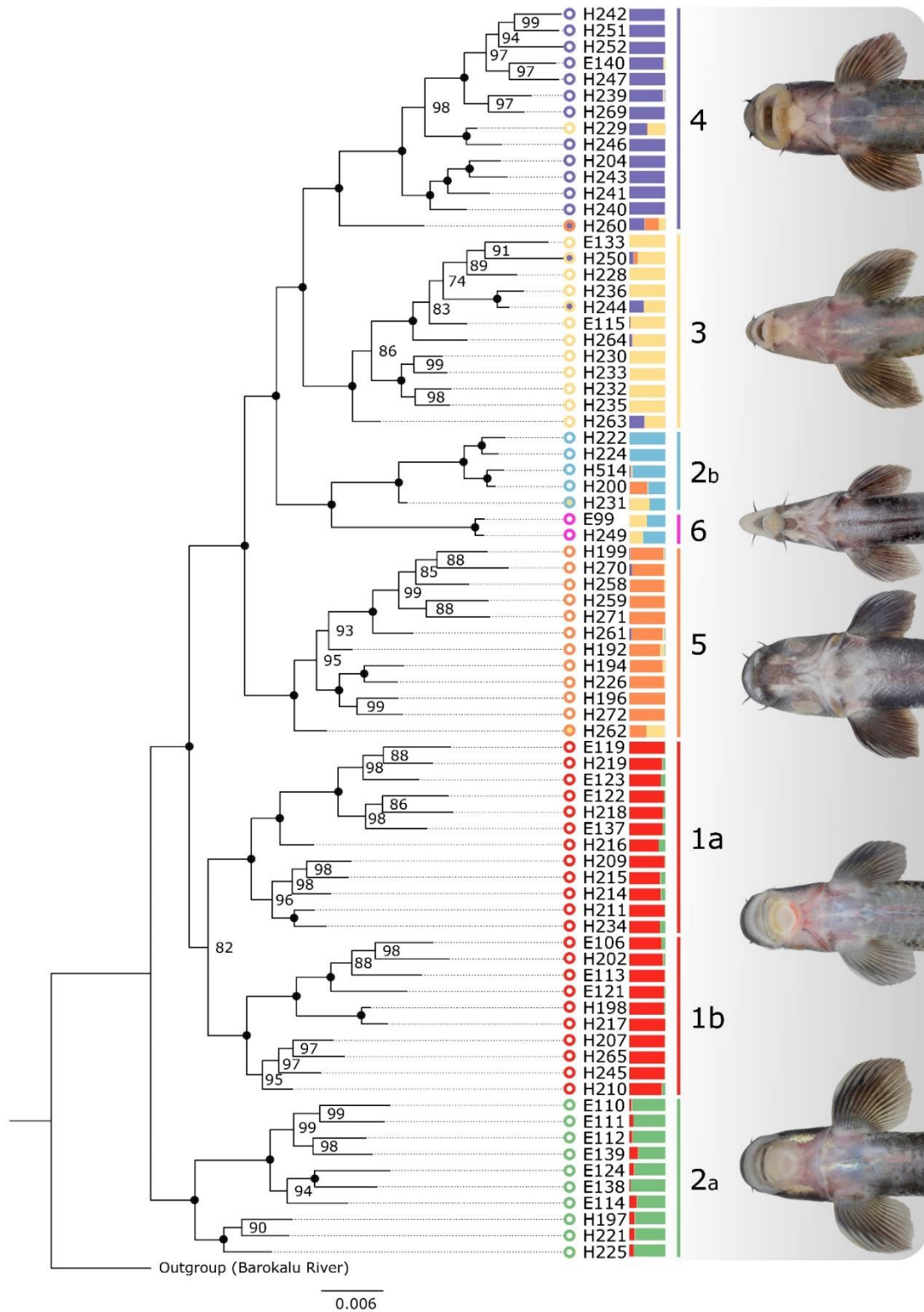


1326

1327 Fig. 5. (A) ML phylogenetic tree of Ethiopian *Garra* based on RAD-loci sequences - 23,365 loci;
1328 3,075,180 bp and (B) SVDQ species tree. Each locus was treated as a separate partition with
1329 GTR+G substitution model and heterozygous sites within each individual encoded using IUPAC
1330 notation. Black dots designate 100% bootstrap support, and only values above 50% are given.

1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349

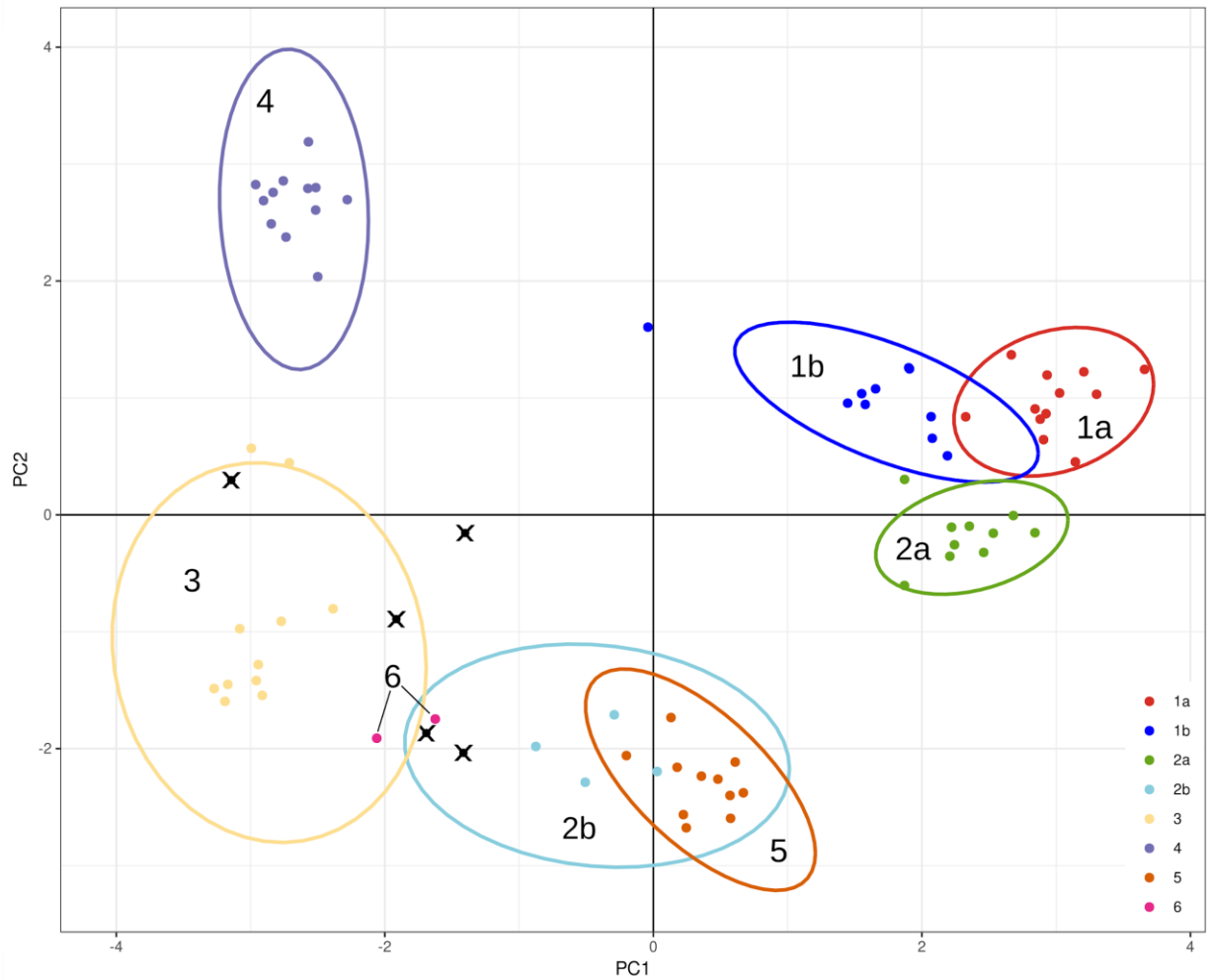
1350
1351
1352
1353
1354



1355

1356 Fig. 6. ML phylogeny of sympatric *Garra* ecomorphs from the Sore River based on concatenated
1357 RAD-loci sequences (7,370 loci; 969,450 bp). Each locus was treated as a separate partition with

1358 GTR+G substitution model. Heterozygous sites within each individual encoded using IUPAC
1359 notation. The individual samples are colored based on the color scheme of Fig. 4 and intermediate
1360 (putative hybrids) phenotypes are depicted in another color. The genetic clusters proportions
1361 inferred by *rmaverick* analysis are shown to the right of sample numbers. Black points designate
1362 100% bootstrap support.



1363

1364 Fig. 7. Principal Component Analysis (PCA) based on 679 nuclear SNPs of sympatric *Garra*
1365 ecomorphs from the Sore River. Points (individuals) and 95% confidence ellipses are colored by
1366 phenotype/genetic cluster. Crosses assign intermediate phenotypes.

1367

1368

1369

1370

1371

1372

1373

1374

1375

1376

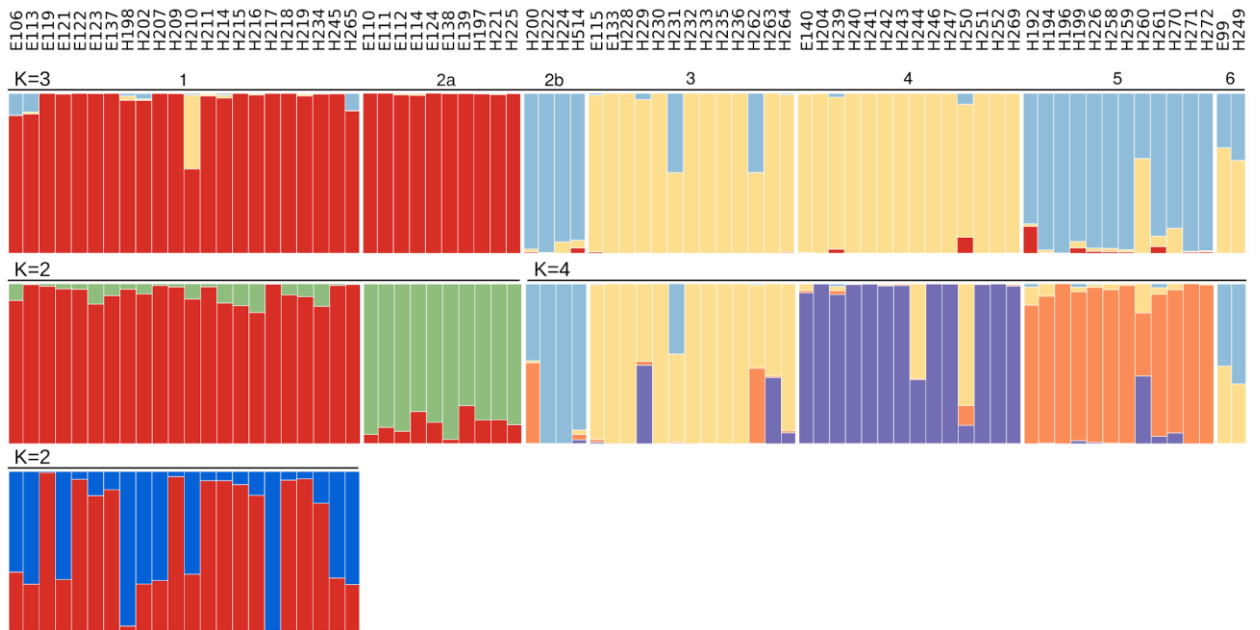
1377

1378

1379

1380

1381
1382
1383
1384
1385

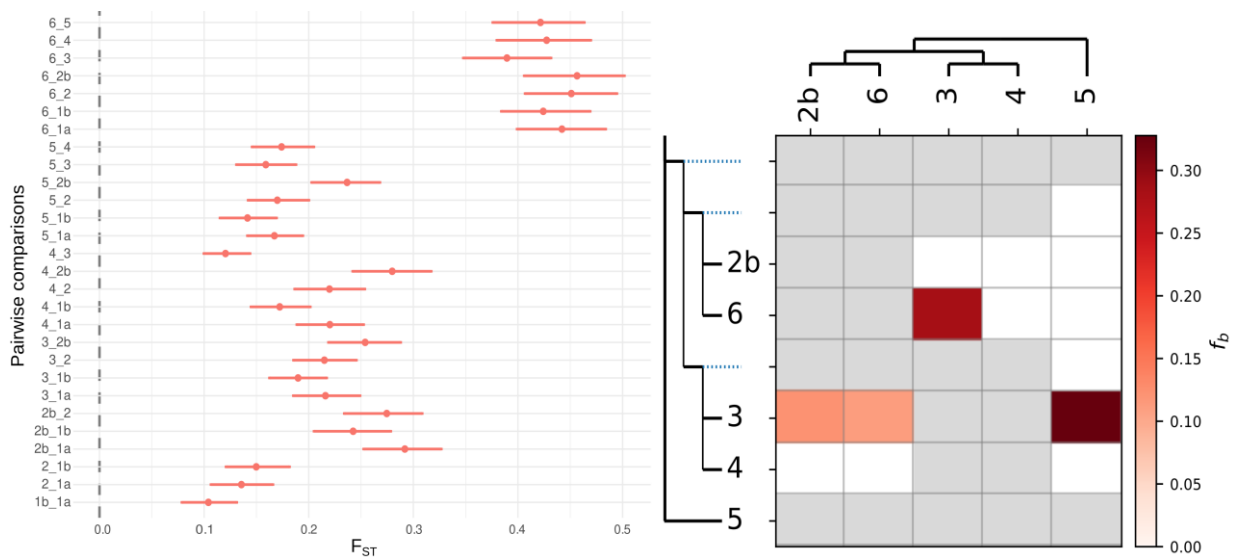


1386

1387 Fig. 8. Hierarchical *maverick* results for sympatric ecomorphs of *Garra* from the Sore River,
1388 based on 679 nuclear SNPs. Each column of the barplot shows individual assignments to one of
1389 the inferred genetic clusters. Independent runs of *maverick* are indicated by a solid black line
1390 above a plot, along with an inferred value of *K*.

1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413

1414
1415
1416
1417
1418



1419

1420 Fig. 9. Left - pairwise Reich F_{ST} values (points) with their respective 95% confidence
1421 intervals (horizontal lines) for *Garra* genetic lineages from the Sore River based on 679 SNPs.
1422 Right - heat map of f -branch metric for selected ecomorphs/lineages of the *Garra* Sore radiation.
1423 The used guide tree is shown along the x and y axes (in 'laddered' form along the y axis). The
1424 matrix shows the inferred f -branch metric, reflecting excess allele sharing between the branch of
1425 the 'laddered' tree on the y axis (relative to its sister branch) and the branches defined on the x
1426 axis.