

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

1 Somewhere I belong: phylogeny and morphological evolution in a species-rich lineage of
2 ectoparasitic flatworms infecting cichlid fishes.

3 Armando J. Cruz-Laufer¹, Antoine Pariselle^{2,3}, Michiel W. P. Jorissen^{1,4}, Fidel Muterezi
4 Bukinga⁵, Anwar Al Assadi⁶, Maarten Van Steenberge^{7,8}, Stephan Koblmüller⁹, Christian
5 Sturmbauer⁹, Karen Smeets¹, Tine Huyse^{4,7}, Tom Artois¹, Maarten P. M. Vanhove^{1,7,10}

6

7 *1 UHasselt – Hasselt University, Faculty of Sciences, Centre for Environmental Sciences,*
8 *Research Group Zoology: Biodiversity and Toxicology, Agoralaan Gebouw D, 3590*
9 *Diepenbeek, Belgium.*

10 *2 ISEM, Université de Montpellier, CNRS, IRD, Montpellier, France.*

11 *3 Faculty of Sciences, Laboratory “Biodiversity, Ecology and Genome”, Research Centre*
12 *“Plant and Microbial Biotechnology, Biodiversity and Environment”, Mohammed V*
13 *University, Rabat, Morocco.*

14 *4 Department of Biology, Royal Museum for Central Africa, Tervuren, Belgium.*

15 *5 Section de Parasitologie, Département de Biologie, Centre de Recherche en Hydrobiologie,*
16 *Uvira, Democratic Republic of the Congo.*

17 *6 Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Nobelstraße 12,*
18 *70569 Stuttgart, Germany.*

19 *7 Laboratory of Biodiversity and Evolutionary Genomics, KU Leuven, Charles*
20 *Deberiotstraat 32, B-3000, Leuven, Belgium.*

21 *8 Operational Directorate Taxonomy and Phylogeny, Royal Belgian Institute of Natural*
22 *Sciences, Vautierstraat 29, B-1000 Brussels, Belgium.*

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van

Steeberge M, Koblmüller S, Sturmbauer C, Huyse T, Smeets K, Artois T, Vanhove MPM

23 *9 Institute of Biology, University of Graz, Universitätsplatz 2, 8010, Graz, Austria.*

24 *10 Department of Botany and Zoology, Faculty of Science, Masaryk University, Kotlářská 2,*

25 *CZ-611 37, Brno, Czech Republic.*

26

27 Corresponding author: Armando J. Cruz-Laufer, armando.cruzlaufer@uhasselt.be

28

29 ABSTRACT

30 A substantial portion of biodiversity evolved through adaptive radiation. However, the effects

31 of explosive speciation on species interactions remain poorly understood. Metazoan parasites

32 infecting radiating host lineages could improve our knowledge because of their intimate host

33 relationships. Yet limited molecular, phenotypic, and ecological data discourage multivariate

34 analyses of evolutionary patterns and encourage the use of discrete characters. Here, we

35 assemble new molecular, morphological, and host range data widely inferred from a species-

36 rich lineage of parasites (*Cichlidogyrus*, Platyhelminthes: Monogenea) infecting cichlid

37 fishes to address data scarcity. We infer a multi-marker (28S/18S rDNA, ITS1, COI mtDNA)

38 phylogeny of 58/137 species and characterise major lineages through synapomorphies

39 inferred from mapping morphological characters. We predict the phylogenetic position of

40 species without DNA data through shared character states, a combined molecular-

41 morphological phylogenetic analysis, and a classification analysis with support vector

42 machines. Based on these predictions and a cluster analysis, we assess the systematic

43 informativeness of continuous characters, search for continuous equivalents for discrete

44 characters, and suggest new characters for morphological traits not analysed to date. We also

45 model the attachment/reproductive organ and host range evolution using the data of 136/137

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

46 described species and multivariate phylogenetic comparative methods (PCMs). We show that
47 discrete characters can mask phylogenetic signals but can be key for characterising species
48 groups. Regarding the attachment organ morphology, a divergent evolutionary regime for at
49 least one lineage was detected and a limited morphological variation indicates host and
50 environmental parameters affecting its evolution. However, moderate success in predicting
51 phylogenetic positions, and a low systematic informativeness and high multicollinearity of
52 morphological characters call for a reevaluation of characters included in species
53 characterisations.

54 KEY WORDS

55 Adaptive radiation; *Cichlidogyrus*; Dactylogyridae; Monogenea; systematic informativeness.

56 ACKNOWLEDGEMENTS

57 We thank J.-F. Agnèse (Institut de Recherche pour le Développement & Université de
58 Montpellier) for taking responsibility for part of the molecular work, and F. A. M. Volckaert
59 (KU Leuven) and J. Snoeks (Royal Museum for Central Africa & KU Leuven) for their
60 guidance since the early stages of this research. J. Bamps, A. F. Grégoir, L. Makasa, J. K.
61 Zimba (Department of Fisheries), C. Katongo (University of Zambia), T. Veall, O. R.
62 Mangwangwa (Rift Valley Tropicals), V. Nshombo Muderhwa, T. Mulimbwa N'sibula, D.
63 Muzumani Risasi, J. Mbirize Ndalozi, V. Lumami Kapepula (Centre de Recherche en
64 Hydrobiologie-Uvira), the Schreyen-Brichard family (Fishes of Burundi), F. Willems
65 (Kasanka Trust), S. Dessein (Botanic Garden Meise), A. Chocha Manda, G. Kapepula
66 Kasembele, E. Abwe, B. Katemo Manda, C. Mukweze Mulelenu, M. Kasongo Ilunga Kayaba
67 and C. Kalombo Kabalika (Université de Lubumbashi), M. Collet and P. N'Lemvo (Institut
68 Congolais pour la Conservation de la Nature), D. Kufulu-ne-Kongo (École Muilu Kiawanga),
69 L. Matondo Mbela (Université Kongo), S. Wamuini Lunkayilakio, P. Nguizani Bimbundi, B.

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van Steeberge M, Koblmüller S, Sturmhuber C, Huyse T, Smeets K, Artois T, Vanhove MPM, Boki Fukiakanda, P. Ntiama Nsiku (Institut Supérieur Pédagogique de Mbanza-Ngungu), P. Nzialu Mahinga (Institut National pour l'Etude et la Recherche Agronomiques—Mvuazi/Institut Supérieur d'études agronomiques—Mvuazi) and M. Katumbi Chapwe are thanked for administrative, field and lab support, making this study possible. We thank A. Avenant-Oldewage, M. Geraerts, T. Gobbin, P. C. Igeh, N. Kmentová, M. Mendlová, and C. Rahmouni for providing raw species measurements used for their respective publications. Furthermore, we thank J. M. Mirande and T. Moons for technical advice on the use of *TNT*. Field work was supported by travel grants V.4.096.10.N.01, K.2.032.08.N.01 and K220314N from the Research Foundation—Flanders (FWO-Vlaanderen) (to MPMV), two travel grants from the King Leopold III Fund for Nature Conservation and Exploration (to MPMV and MVS), FWO-Vlaanderen Research Programme G.0553.10, the University Development Cooperation of the Flemish Interuniversity Council (VLIR-UOS) South Initiative ZRDC2014MP084, the OCA type II project S1_RDC_TILAPIA and the Mbisa Congo project (2013–2018), the latter two being framework agreement projects of the RMCA with the Belgian Development Cooperation. When data collection for this study commenced, MVS and MPMV were PhD fellows, and TH a post-doctoral fellow, of FWO-Vlaanderen. Part of the research leading to results presented in this publication was carried out with infrastructure funded by the European Marine Biological Research Centre (EMBRC) Belgium, Research Foundation – Flanders (FWO) project GOH3817N. MWPJ was supported by the Belgian Federal Science Policy Office (BRAIN-be Pioneer Project (BR/132/PI/TILAPIA), and a BOF Reserve Fellowship from Hasselt University. AJCL is funded by Hasselt University (BOF19OWB02), and MPMV receives support from the Special Research Fund of Hasselt University (BOF20TT06). We thank the reviewers

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

93 including J. M. Mirande and an anonymous reviewer for suggesting improvements to the
94 manuscript.

95 CONFLICT OF INTEREST

96 The authors declare that they have no conflict of interest.

97 DATA AVAILABILITY STATEMENT

98 The morphological data that support the findings of this study are openly available in
99 MorphoBank at www.morphobank.org, at <https://dx.doi.org/XXXXXXXX>. The DNA
100 sequence data are openly available in the GenBank Nucleotide Database at
101 <https://www.ncbi.nlm.nih.gov/genbank>, accession numbers XXXXXX–XXXXXX.
102 Phylogenetic trees and data matrices are openly available in TreeBase at <https://treebase.org>,
103 accession number XXXXXX.

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van

Steeberge M, Koblmüller S, Sturmhuber C, Huyse T, Smeets K, Artois T, Vanhove MPM

104 INTRODUCTION

105 *Adaptive radiations and host-parasite interactions*

106 Adaptive radiation is one of the most important processes of species formation. These
107 explosive speciation events might explain a substantial part of the biodiversity on the planet
108 (Glor, 2010). Adaptive radiations are characterised by a rapid diversification resulting from
109 adaptation to newly available ecological niches (Losos, 2010). Famous examples include
110 Darwin's finches (Grant, 1999), Caribbean lizards of the genus *Anolis* Daudin, 1802 (Mahler
111 et al., 2013), cichlid fishes (Salzburger, 2018), and Hawaiian silverworts (Landis et al.,
112 2018). Despite the attention these species have receive, few studies have investigated the
113 evolution of ecological interactions involving these groups and other organisms of other
114 groups beyond the feeding ecology of the former (e.g. Guerrero and Tye, 2009; Takahashi
115 and Koblmüller, 2011; but see Karvonen and Seehausen, 2012; Blažek et al., 2018). How do
116 organisms evolve that accompany the rapid diversification process of an adaptive radiation?

117 Metazoan parasites could provide answers to this question as they form often intimate
118 relationships with their hosts. Thus, their evolutionary regime can be strongly impacted by
119 the host's evolutionary history (Huyse et al., 2005). Multiple evolutionary regimes might
120 apply to parasites infecting radiating host lineages. First, parasites that are strongly dependent
121 on their hosts, might experience a selection pressure to remain competitive in the arms race
122 with the host's defences (Kaltz and Shykoff, 1998). Structures relevant to this arms race such
123 as attachment organs might experience a strong stabilising selection pressure as they are the
124 parasite's main physical connection to the host. In this scenario, the attachment organ
125 evolution would be heavily impacted by host and environmental parameters (Kaltz and
126 Shykoff, 1998). Second, parasites could follow the example of their hosts in terms of an

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

127 (adaptive) radiation. Such co-radiations have previously been reported for insects parasitic on
128 plants or other insects (Weiblen and Bush, 2002; Forbes et al., 2009) but quantitative support
129 for these hypotheses in the form of evolutionary models has not been provided to date. Third,
130 parasites could have structures that are not under strong selection pressure and, therefore,
131 their morphology might randomly diverge over time following a pattern associated with a
132 ‘random walk’, i.e. genetic drift or randomly fluctuating selection (Losos, 2008).
133 Reproductive organ structures, e.g. in flatworms infecting the gills of cyprinid fish (Šimková
134 et al., 2002), have been suggested to follow this pattern causing reproductive isolation
135 between species.

136 *Species-rich but missing data: a case study for host-parasite systems*

137 Extensive evolutionary analyses often require large molecular and morphological
138 datasets. However, such datasets remain scarce for parasitic organisms. Metazoan parasites
139 are often small, which makes identifying these organisms to species level notoriously
140 difficult (de Meeûs et al., 2007). Furthermore, only a fraction of the known species is
141 genetically characterised (Poulin et al., 2019) and most species are yet to be discovered and
142 described (Poulin et al., 2020). Among the best-known models in adaptive radiation research,
143 African cichlids (Cichliformes, Cichlidae) are possibly the best-studied system (Salzburger,
144 2018) with approximately 2000 (described and undescribed) species reported from Eastern
145 Africa alone (Turner et al., 2001; Salzburger et al., 2014). This knowledge is rooted in a long
146 and productive tradition of international cichlid research (see Ronco et al., 2020; Van
147 Steenberge et al., 2011). Resulting from this scientific interest, one group of gill parasites
148 (Fig. 1a) infecting these fishes, the monogenean flatworms belonging to *Cichlidogyrus* sensu
149 Paperna, 1960 [incl. the nested genus *Scutogyrus* Pariselle & Euzet, 1995 (Wu et al., 2007),
150 together referred to as *Cichlidogyrus* in the following] (Fig. 1b, c), has been studied in more
151 depth than other species-rich parasite genera, in particular from the African continent (Cruz-

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van Steeberge M, Koblmüller S, Sturmbauer C, Huyse T, Smeets K, Artois T, Vanhove MPM Laufer et al., 2021). Species of *Cichlidogyrus* rival (Cruz-Laufer et al., 2021) and possibly exceed (see Poulin, 2014) their hosts in terms of species numbers: 137 parasite species have been described on 126 fish species (Cruz-Laufer et al., 2021). Monogenean flatworms have provided insight into parasite speciation (Meinilä et al., 2004; Šimková et al., 2013; Vanhove et al., 2015), population dynamics (Kmentová et al., 2021b), anthropogenic introductions (Šimková et al., 2019; Jorissen et al., 2020) as well as host biogeography (Barson et al., 2010; Pariselle et al., 2011; Vanhove et al., 2013, 2016). This variety of evolutionary research paired with the model system status of the host species of species of *Cichlidogyrus*, has led to the suggestion of the cichlid-*Cichlidogyrus* species network as model system for parasite speciation research and the evolution of host-parasite interactions (Pariselle et al., 2003; Vanhove et al., 2016). Recent advances in immunological (Zhi et al., 2018), pathological (Igeh and Avenant-Oldewage, 2020), genomic (Vanhove et al., 2018; Caña-Bozada et al., 2021), and microscopy research (Fannes et al., 2015) have brought the vision of a cichlid-*Cichlidogyrus* model system closer to reality [see Cruz-Laufer et al. (2021) for a detailed review of model system qualities].

The species-richness of *Cichlidogyrus* has previously been attributed to co-divergence with their hosts (Pariselle et al., 2003; Vanhove et al., 2015), host switching (Pariselle et al., 2003; Messu Mandeng et al., 2015), and within host speciation (Mendlová et al., 2012; Vanhove et al., 2015). Similar to other parasites, morphological characters of monogenean attachment and reproductive organ structures have formed an essential part of species descriptions. Evolutionary studies have paid considerable attention to the evolutionary patterns of these (mostly) sclerotised structures (e.g. Šimková et al., 2002; Mendlová et al., 2012). For species of *Cichlidogyrus*, the attachment organ includes a haptor with several anchors, hooks, and bars and reproductive organs with a male copulatory organ (MCO) and,

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

176 at times, a sclerotised vagina (Fig. 1b, c). Based on the similar phenotypes observed in related
177 species, some studies have suggested phylogenetic constraints, i.e. ancestry of the parasite
178 species, as the sole determinants of the morphology of the attachment organ in *Cichlidogyrus*
179 (Vignon et al., 2011). This claim has been questioned more recently. At least one species
180 appears to have changed its attachment organ morphology compared to its close relatives as a
181 response to a host switch towards non-cichlid fishes (Messu Mandeng et al., 2015). These
182 host switches are considered rare as monogenean representatives have been reported for 27
183 (Carvalho Schaeffner, 2018) of 48 (Lévêque et al., 2008) fish families in Africa and most
184 harbour some other dactylogyridean lineage of gill parasites (see Carvalho Schaeffner, 2018).
185 Despite this discovery, no study has investigated the macroevolution of the morphological
186 characters of the attachment and reproductive organs for almost a decade (see Mendlová et
187 al., 2012).

188 The present study aims to investigate the evolutionary processes that have shaped the
189 morphology of species of *Cichlidogyrus*. However, out of the 137 species that are currently
190 described, DNA sequences of only 18 species were included in the most recent phylogenetic
191 study (Messu Mandeng et al., 2015). Furthermore, morphological and ecological data have
192 not been collected in centralised databases inhibiting the progress of large-scale meta-
193 analytical studies (Cruz-Laufer et al., 2021). Previous studies have evaded this scarcity of
194 information by loosely defining discrete states for the attachment organ morphology and the
195 host repertoire base on the taxonomic literature. For instance, Vignon et al. (2011) and
196 Mendlová et al. (2012) grouped species by the relative size of the sclerotised structures of the
197 attachment organ (for an overview of the terminology, see Fig. 1c) using generic terms like
198 ‘large’ and ‘small’ instead of coding these characters based on continuous measurements.
199 Similarly, Mendlová and Šimková (2014) proposed an index of specificity (IS) to group
200 species as species-, genus-, and tribe-specific, or generalist based on the most recent host

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van

Steeberge M, Koblmüller S, Sturmbauer C, Huyse T, Smeets K, Artois T, Vanhove MPM

201 classification at the time. Yet discretisation has been known to cause information loss as the
202 variability of the otherwise continuous parameters is largely ignored (Altman and Royston,
203 2006; Goloboff et al., 2006; Parins-Fukuchi, 2018). These obstacles emphasise the need for
204 new approaches and more extensive and accessible datasets to gain insight into the evolution
205 of *Cichlidogyrus* and the phylogenetic information of morphological and ecological
206 characters.

207 To address data scarcity, we provide new morphological, molecular, and host range data,
208 which are made available in public databases. We perform multiple analyses to investigate
209 the evolution of phenotypic characters to answer the following questions:

- 210 (i) Which main lineages can we infer from the phylogeny? To which lineages do species
211 without molecular data belong? Which predictions can be made through character
212 mapping, a combined (molecular and morphological) parsimony analysis, and a machine
213 learning algorithm?
- 214 (ii) How systematically informative are the morphometric measurement most widely applied
215 to this genus? Does the use of discrete characters lead to information loss in the
216 morphometric and host range data?
- 217 (iii) As species of *Cichlidogyrus* infect a host lineage known for its rapid speciation, which
218 evolutionary processes have shaped the attachment and reproductive organ morphology?
219 Are these structures under stabilising selection pressure (H_2), do the parasites mirror the
220 radiations of their hosts either as an early (H_3) or a late burst (H_4) in the character
221 evolution, or do these structures follow a pattern associated with a random walk (H_1)?
222 Furthermore, do species infecting cichlids from the Eastern African radiations follow a
223 different evolutionary regime than the species infecting other hosts (H_5)?

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

224 MATERIALS AND METHODS

225 *Sampling*

226 Fish specimens belonging to Cichlidae Bonaparte, 1835 (Cichliformes), the
227 nothobranchiid genus *Aphyosemion* Myers, 1924 (Cyprinodontiformes: Nothobranchiidae),
228 and *Polycentropsis abbreviata* Boulenger, 1901 (Ovalentaria, *incertae sedis*: Polycentridae)
229 were collected during various expeditions across Africa between 2008 and 2019 and from
230 few specimens from aquaculture facilities in France (Appendix 1). As samples collected
231 during these expeditions have been included in previous studies, sampling details can be
232 found in the respective publications (Vanhove et al., 2011, 2013; Muterezi Bukinga et al.,
233 2012; Pariselle et al., 2015a, 2015b; Jorissen et al., 2018a, 2018b, 2020). We dissected the
234 gills of the fish and stored the samples in 96% ethanol to preserve the DNA of the parasites
235 attached to the gills. Then, we screened the gills for parasitic infections of flatworm species
236 belonging to *Cichlidogyrus*, removed the parasites from the gills using dissection needles
237 under a stereomicroscope, mounted them temporarily in water for species-level identification
238 at a magnification of 1000x (100x magnification with oil immersion and 10x ocular) and
239 stored the parasite specimens individually in 96% ethanol for DNA extraction. As mostly
240 entire specimens were used for DNA extraction, we refer to the above-mentioned articles for
241 type and voucher specimens from the same host individuals deposited in curated collections
242 (see Appendix 1).

243 *DNA extraction, amplification, and sequencing*

244 For the DNA extraction of recent samples, we used the Nucleospin Kit (Macherey-
245 Nagel, USA) following manufacturer guidelines but with 60 µl instead of 100 µl of elution
246 buffer added in the final step. For some samples processed at an earlier stage, we suspended
247 DNA samples through homogenisation with no added extraction steps following the

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van

Steeberge M, Koblmüller S, Sturmhuber C, Huyse T, Smeets K, Artois T, Vanhove MPM

248 procedures described by Marchiori et al. (2015). We amplified and sequenced three partial
249 nuclear ribosomal genes including a fragment of the large subunit ribosomal DNA (28S
250 rDNA) with the primers C1 (5'-ACCCGCTGAATTTAAGCAT-3') and D2 (5'-
251 TGGTCCGTGTTTCAAGAC-3') (Hassouna et al., 1984), a fragment of the small subunit
252 ribosomal DNA (18S rDNA) and the internal transcribed spacer 1 (ITS1) with the primers S1
253 (5'-ATTCCGATAACGAACGAGACT-3') (Matejusová et al., 2001) and IR8 (5'-
254 GCAGCTGCGTTCTTCATCGA-3') (Šimková et al., 2003), and a fragment of the
255 mitochondrial gene coding for the cytochrome oxidase *c* subunit 1 protein (COI mtDNA)
256 with the primers ASmit1 (5'-TTTTTTGGGCATCCTGAGGTTTAT-3') (Littlewood et al.,
257 1997), Cox1_Schisto_3 (5'-TCTTTRGATCATAAGCG-3') (Lockyer et al., 2003), and
258 ASmit2 (5'-TAAAGAAAGAACATA ATGAAAATG-3') (Littlewood et al., 1997). Reaction
259 protocols followed the procedures of Messu Mandeng et al. (2015) for 28S rDNA, Mendlová
260 et al. (2012) for 18S rDNA, and Vanhove et al. (2015) for COI mtDNA. We purified PCR
261 products with a NucleoFast 96 PCR kit (Macherey-Nagel, USA) or with a GFX PCR DNA
262 kit and Gel Band Purification kit (GE Healthcare, USA) following manufacturer guidelines.
263 Bidirectional Sanger sequencing was conducted according to the Big Dye Terminator v3.1
264 sequencing protocol (Applied Biosystems, USA) at a 1:8 dilution with an ABI PRISM 3130
265 Avant Genetic Analyser automated sequencer (Applied Biosystems, USA) and the primers
266 included in the PCR protocols.

267 *Morphological and host range data collection*

268 We assembled morphometric (Fig. 1c) and host range data for 136 species belonging
269 to *Cichlidogyrus*. Raw measurements taken through light microscopy were assembled from
270 AP's personal morphometric database and from the raw data of previous publications kindly
271 provided by the authors of the cited studies (Pariselle and Euzet, 2003; Vanhove et al., 2011;

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

272 Gillardin et al., 2012; Muterezi Bukinga et al., 2012; Pariselle et al., 2013; Řehulková et al.,
273 2013; Van Steenberge et al., 2015; Messu Mandeng et al., 2015; Kmentová et al., 2016a,
274 2016b, 2016c; Rahmouni et al., 2017, 2018; Igeh et al., 2017; Geraerts et al., 2020; Gobbin et
275 al., 2021). These raw data were deposited at MorphoBank (www.morphobank.org, project
276 XXXXX). We calculated the standard errors and standard deviations for each measurement
277 from the raw data. However, for species where no or only partial raw data were available, we
278 used the mean values provided in the literature (Douëllou, 1993; Pariselle and Euzet, 2003;
279 Řehulková et al., 2013; Rahmouni et al., 2017; Jorissen et al., 2018b, 2018a) without errors.
280 If no mean was reported, we inferred the measures from drawings in the literature (Dossou,
281 1982; Dossou and Birgi, 1984; Birgi and Lambert, 1986; Muterezi Bukinga et al., 2012) and
282 calibrating measurements through the included scale bars. Measurements of *C. dionchus*
283 Paperna, 1968 were not included as no mean values were provided in respective publications
284 (Paperna, 1960, 1968; Paperna and Thurston, 1969) and drawings were incomplete with some
285 structures, e.g. the second marginal hook, entirely missing. For the auxiliary plate (AuP) of
286 the male copulatory organ, we used the surface area and its relative error resulting of the sum
287 of relative errors of the original measurements inferred from the plate length and width
288 assuming an ellipsoid shape.

289 Beyond these continuous characters, we assigned all species to previously suggested
290 discrete characters for the haptor morphology and host range, which include the configuration
291 of the hooks (Vignon et al., 2011), the similarity of the anchors, the shape of the ventral bar
292 (Mendlová et al., 2012), and the index of specificity (IS) (Mendlová and Šimková, 2014).
293 Hook configurations were coded according to the secondary growth in the first and third hook
294 pair ('small-small', 'small-large', 'large-small', 'large-large') relative to the second pair,
295 which retains its embryonal size in adult specimens (Llewellyn, 1963). The anchor pairs were
296 categorised as 'similar' or 'dissimilar' in shape and size based on the original species

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van Steeberge M, Koblmüller S, Sturmbauer C, Huyse T, Smeets K, Artois T, Vanhove MPM

297 descriptions. The ventral bar shape was defined according to the presence of membranous
298 extensions (see Mendlová et al., 2012) and size ('with membranous extensions', 'without
299 membranous extensions', 'massive with membranous extension', 'bar supports large
300 extended plate (*Scutogyrus*'). We inferred IS data from the host-parasite interactions listed
301 by Cruz-Laufer et al. (2021) with species exclusively reported from hosts belonging to a
302 single species, genus, or tribe classified as species-, genus-, and tribe-specific, or generalist if
303 hosts belonging to two different tribes (or even families) were infected by the same parasite
304 species. For the classification of cichlid species into tribes, we followed Dunz and Schliewen
305 (2013). All discrete and continuous morphometric and host specificity data are made
306 available at TreeBase alongside the DNA sequence alignments and tree topologies
307 (www.treebase.org, accession number: XXXXX). Raw morphometric data are made
308 available at MorphoBank (www.morphobank.org, project XXXXX).

309 We also surveyed the taxonomic literature to detect additional features used to
310 recognise groups of related species. First, we attempted to find equivalents for characters and
311 character states frequently mentioned in species descriptions among the morphometrics. For
312 instance, the root lengths of the anchors in the attachment organ are frequently mentioned as
313 'large' or 'small', which can be reflected through the ratio of the two roots in each anchor
314 pair. If no measurement reflected the described features, we developed new discrete
315 variables. Detailed information on the continuous and proposed new discrete characters and
316 their character states can be found in Appendix 2.

317 *Phylogenetic analyses*

318 We assembled a four-locus concatenated multiple alignment from the sequences
319 generated during this study and sequences available at GenBank (see Appendix 3). The

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

320 alignment includes partial ITS1, 18S rDNA, 28S rDNA, and COI mtDNA sequence data.
321 Partial DNA sequence data (i.e. specimens with less than those four loci sequenced) were
322 included with the lacking fragments coded as missing data. We aligned the sequences of each
323 locus using the algorithm *L-INS-i* in *MAFFT* v.7.409 (Kato and Standley, 2013) as
324 recommended for ribosomal DNA by the *MAFFT* manual, and removed poorly aligned
325 positions and divergent regions using the options for less stringent parameters in *Gblocks*
326 v0.91b (Talavera and Castresana, 2007).

327 We estimated tree topologies under maximum parsimony (MP) through *TNT* v1.5
328 (Goloboff et al., 2008b; Goloboff and Catalano, 2016) using extended implied weighting
329 (Goloboff, 2014) in a range of values for the concavity constant K (20, 21, 23, 26, 30, 35, 41,
330 48, 56). Extended implied weighting reduces the impact of characters with missing data that
331 were weighted artificially high in the original implied weighting method (Goloboff, 1993).
332 For DNA sequence data, collectively weighting all sites in a partition (gene or codon) was
333 suggested as a more appropriate method (Goloboff et al., 2008a). Therefore, we explored
334 three weighting schemes proposed by Mirande (2019): all characters weighted separately
335 (SEP), all characters weighted according to the average homoplasy of the marker (BLK), and
336 characters in the protein-coding marker, i.e. COI, weighted according to the average
337 homoplasy of their position with other characters were weighted as in BLK (POS). Similar to
338 Mirande (2009), we selected the parameter combinations (k value and weighting scheme) that
339 produced the most stable topology to compute a strict consensus tree. The distortion
340 coefficient and subtree pruning and regrafting (SPR) distance were used as selection criteria
341 by computing the similarity of each consensus tree obtained under the different parameters to
342 the rest. MP tree searches involved rounds of tree fusing, sectorial searches, tree drifting, and
343 tree ratchet (Goloboff, 1999; Nixon, 1999) under default settings and each round was stopped
344 following three hits of the same optimum. Gaps were treated as missing data. As

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van Steeberge M, Koblmüller S, Sturmhuber C, Huysse T, Smeets K, Artois T, Vanhove MPM

345 bootstrapping and jackknifing are reported to be distorted by differently weighted characters
346 (Goloboff, 2003), branch support was estimated through symmetric resampling with a
347 probability of change of 0.33 and values expressed as differences in frequencies (GC:
348 ‘Groups present/Contradicted’). *Cichlidogyrus berrebi* Pariselle & Euzet, 1994, *C. kothiasi*
349 Pariselle & Euzet, 1994, and *C. pouyaudi* Pariselle & Euzet, 1994, parasites of tylochromine
350 cichlids, were used to root the phylogenetic trees due to the well-documented early diverging
351 phylogenetic position of these species (Mendlová et al., 2012; Messu Mandeng et al., 2015).

352 To infer the phylogenetic position of species of *Cichlidogyrus* without DNA sequence
353 data, we performed a second phylogenetic analysis including the morphometric
354 measurements of the attachment and reproductive organs as a separate block in *TNT* and
355 extended implied weighting using the same set of k values and the weighting scheme that
356 produced the most stable tree topologies for the DNA data. The homoplasy of each
357 morphometric character was estimated independently as recommended by Goloboff et al.
358 (2006) for continuous characters. All other settings in *TNT* remained the same as above. The
359 full *TNT* data matrix is provided in Supporting Information (Table S1).

360 For downstream analyses of character evolution, we also estimated phylogenies
361 through Bayesian inference (BI) and maximum likelihood (ML) methods using *MrBayes*
362 v3.2.6 (Ronquist and Huelsenbeck, 2003) on the CIPRES Science Gateway online server
363 (Miller et al., 2010) and *IQ-Tree* v1.6.12 (Nguyen et al., 2015) respectively. We used these
364 model-based approaches (BI and ML) to provide a consistent approach to the downstream
365 multivariate phylogenetic comparative analyses, which are themselves model-based. DNA
366 sequence data were partitioned by gene and, for the COI mtDNA, by codon position. We
367 selected the substitution models for each partition according to the Bayesian information
368 criterion (BIC) as rendered by *ModelFinder* in *IQ-TREE* (Kalyaanamoorthy et al., 2017)

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

369 using partition merging (Chernomor et al., 2016) (Appendix 4). For BI analyses, we selected
370 only models implemented in MrBayes (Appendix 4). We used two parallel runs and four
371 chains of Metropolis-coupled Markov chain Monte Carlo iterations. We ran 20 million
372 generations with a burn-in fraction of 0.25 and sampled the trees every 1000th generation. We
373 checked convergence criteria by assessing the average standard deviation of split frequencies
374 (< 0.01 in all datasets) and the effective sample size (> 200) using *Tracer* v1.7 (Rambaut et
375 al., 2018). For ML analyses, we estimated branch support values using both ultrafast
376 bootstrap approximation (Hoang et al., 2018) and Shimodaira-Hasegawa-like approximate
377 likelihood ratio tests (SH-aLRT) (Guindon et al., 2010) with 1000 replicates as recommended
378 by the *IQ-Tree* manual. We considered a BI posterior probability (PP) ≥ 0.95 , an ultrafast
379 bootstrap values ≥ 95 , and SH-aLRT statistic ≥ 80 as well-supported (Hoang et al., 2018). We
380 plotted the graphs and phylogenetic trees using the *R* packages *ggplot2* v3.3.5 (Wickham,
381 2016) and *ggtree* v3.13 (Yu et al., 2017, 2018).

382 To verify congruence of the final hypothesis of the parsimony analysis (molecular
383 data) with the BI and ML consensus trees, we analysed the congruence of the MP, BI, and
384 ML tree topologies using the Congruence Among Distance Matrices (CADM) tests
385 (Legendre and Lapointe, 2004; Campbell et al., 2011). We used the package *ape* v5.3
386 (Paradis and Schliep 2019) in *R* v4.1.0 (R Core Team, 2021) to calculate phylogenetic pair-
387 wise distance matrices and to conduct the CADM test.

388 *Clade affiliation and discriminative power of morphometrics: statistical classification,*
389 *literature review, and cluster analysis*

390 Beyond the clades supported in the combined MP analysis, we also used two
391 additional approaches to assess the phylogenetic position of species of *Cichlidogyrus* that
392 have not been sequenced to date. First, we characterised the morphology and host repertoires

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van

Steeberge M, Koblmüller S, Sturmhuber C, Huyse T, Smeets K, Artois T, Vanhove MPM

393 of the clades inferred from the molecular phylogeny (Fig. 2) based on character maps of the
394 continuous and discrete morphological characters and the host repertoires surveyed in the
395 literature. To map continuous characters on the consensus MP tree, we estimated ancestral
396 character states through the function *anc.ML* in the R package *phytools* v0.7-80 (Revell,
397 2012). To map the discrete characters, we estimated ancestral states under maximum
398 parsimony with all rates set to equal as we could not make any assumption on the transition
399 costs between character states. This analysis was implemented in the function
400 *asr_max_parsimony* in the R package *castor* v1.6.9 (Louca and Doebeli, 2018). All character
401 maps were plotted using the *ggplot2* and *ggtree*. Species of *Cichlidogyrus* not included in
402 molecular phylogenies were assigned to the clades based upon the measurements and their
403 character states while taking host repertoires (by tribe of cichlids or family of non-cichlids)
404 into consideration. These last criteria (natural and invasive host repertoires) were taken into
405 account as we expected the host environment to be a relevant factor in speciation processes of
406 these parasites (see McCoy, 2003; Huyse et al., 2005). We used the host-parasite list and
407 literature databases provided by Cruz-Laufer et al. (2021) to summarise the host repertoires.

408 Second, we predicted clade affiliation through a supervised machine learning
409 algorithm using support vector machines (SVMs). Machine learning can be used to assess of
410 the predictive power of morphometric measurements in taxonomic studies, in particular with
411 support vector machines (SVMs) (e.g. Zischke et al., 2016; Fang et al., 2018). In these
412 supervised machine learning approaches, a dataset is provided to train the learning algorithm
413 to classify individuals in distinct groups. The predictions generated by the algorithm are then
414 validated against a set of samples with a known affiliation. The performance can then provide
415 an insight into the predictive power of the input data. Subsequently, the optimised algorithm
416 can be applied to a test dataset of samples with unknown group affiliation. Here, we carried

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

417 out the SVM analysis with a radial basis kernel function as implemented in the method
418 *svmRadial* in the R package *caret* (Kuhn, 2008; Meyer et al., 2020). Missing data in the
419 measurements were imputed through k-nearest neighbour imputation, scaled, and centred as
420 implemented in the function *preProcess*. The C and σ parameters of the kernel function were
421 optimised through a grid-search with exponentially growing sequences as suggested by Hsu
422 et al. (2003) and through tenfold cross-validation with ten repetitions. As specimens
423 belonging to different clades were observed in unequal numbers (class imbalance), we
424 optimised the parameters based on Cohen's κ , a multiclass performance metric accounting for
425 class imbalance (Landis and Koch, 1977). We considered $\kappa < 0.2$ a *slight*, κ between 0.2 and
426 0.4 as *fair*, κ between 0.4 and 0.6 as *moderate*, κ between 0.6 and 0.8 as *substantial*, and κ
427 above 0.8 as *almost perfect* agreement (Landis and Koch, 1977). Prediction provided by the
428 optimised algorithm were compared to the clade affiliations provided the combined
429 molecular-morphological parsimony analyses and the literature survey. Finally, we inferred
430 the variable importance for SVM predictions through a pairwise receiver operating
431 characteristic (ROC) curve analysis as implemented in the function *filtervarImp*. We
432 considered an area under the curve (AUC) between below 0.7 *poor*, between 0.7 and 0.8
433 *acceptable*, between 0.8 and 0.9 *excellent*, and above 0.9 *outstanding*.

434 To further assess the systematic informativeness of the morphometrics, we
435 investigated the amount of possible redundancies by identifying clusters of multicollinear
436 morphometric measurements. We used a Pearson pairwise correlation matrix (Dormann et al.,
437 2013) and the *ward.d2* clustering algorithm (Murtagh and Legendre, 2014) to detect clustered
438 variables in the R package *ComplexHeatmap* v2.8.0 (Gu et al., 2016). If the absolute values
439 of the Pearson correlation coefficients ($|r|$) exceeded 0.7 (Dormann et al., 2013), we
440 considered measurements multicollinear. Hence, we split the variables into clusters along the
441 dendrograms provided by *ward.d2* until all coefficients within the same cluster exceeded this

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van

Steeberge M, Koblmüller S, Sturmbauer C, Huyse T, Smeets K, Artois T, Vanhove MPM

442 threshold. Heatmaps were plotted using the *R* package *ComplexHeatmap* v2.8.0 (Gu et al.,
443 2016).

444 *Character evolution of the attachment and reproductive organs and phylogenetic signal*

445 Using the morphometric measurements, we tested for possible patterns of a random walk
446 (H_1), stabilising selection (H_2), and adaptive radiation with a decelerating (H_3) and
447 accelerating divergence of characters (H_4) on the evolution of the attachment and
448 reproductive organs. We also tested if species from the East African lakes followed a
449 different evolutionary regime (H_5) as a result of the multiple host radiations in this region. To
450 detect these patterns, we employed multivariate phylogenetic comparative methods as
451 implemented in the *R* package *mvmorph* v1.1.4 (Clavel et al., 2015) to account for potential
452 interactions between characters. This software package addresses the sensitivity of previous
453 multivariate approaches (e.g. Khabbazian et al., 2016; Goolsby et al., 2017) to trait
454 covariation, data orientation, and trait dimensions (Adams and Collyer, 2018).

455 We fitted a range of multivariate models of continuous character evolution on a single
456 sample of 100 randomly selected tree topologies drawn from the post–burn-in phase of the BI
457 analysis of the molecular markers. Morphometric measurements were averaged by species.
458 Furthermore, we excluded all but one specimen per species from the phylogenetic trees to
459 avoid that specimens of the same species are assigned identical values, which might
460 otherwise artificially decrease the estimated variability of the measurements per taxon
461 affecting model performance. Specimens included in the subset trees, were chosen at random
462 and are highlighted in Appendix 3. The tested models include:

- 463 • The Brownian motion model (BM) (Felsenstein, 1973) simulates a ‘random walk’, i.e.
464 genetic drift or randomly fluctuating selection (see Hansen and Martins, 1996) (H_1). In

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

465 this case, the parasite morphology would not be affected by strong selection pressures and
466 would randomly diverge over time.

- 467 • The Ornstein-Uhlenbeck model (OU) (Butler and King, 2004) approximates a character
468 evolution under stabilising selection (H_2). Here, the parasite morphology experiences a
469 selection pressure towards a selective optimum due to a high host-dependence.
- 470 • The Early-Burst (EB) (Harmon et al., 2010) or accelerate-decelerate model (ACDC)
471 (Blomberg et al. 2003) models an early rapid evolution of characters followed by a slow-
472 down such as might occur during adaptive radiation events (H_3). Here, the parasite
473 morphology has rapidly diverged possibly mirroring the explosive speciation events
474 reported for some of the host lineages.
- 475 • The Late-Burst model (LB), a variation of the ACDC model with an accelerating
476 divergence of characters (Blomberg et al., 2003) (H_4), simulates evolution as expected
477 under a recent, ongoing radiation event. The speciation of the parasites would mirror the
478 host radiations but would still be ongoing.
- 479 • The multi-rate Brownian motion model (BMM) represents two different multi-selective
480 and multi-rate Brownian motion regimes for East African parasite lineages and other
481 lineages (O'Meara et al., 2006). Here, the parasite morphology would have undergone an
482 evolutionary rate shift because species infecting the host radiations in Eastern African
483 have developed differently than their congeners in the rest of Africa. The regimes were
484 defined using the function *paintSubTree* in the R package *phytools* v0.7-80 (Revell,
485 2012).

486 We fitted all models using the *mvglS* function (Clavel et al., 2019) in *mvmorph* as
487 recommended for independently evolving measurements with a natural orientation.
488 Therefore, we used a restricted maximum likelihood (REML) estimation with a leave-one-out
489 cross-validation of the penalized log-likelihood (PL-LOOCV), an estimated within-species

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van Steeberge M, Koblmüller S, Sturmbauer C, Huyse T, Smeets K, Artois T, Vanhove MPM

490 variation, an archetypal ridge estimator, and a diagonal unequal variance target matrix
491 (Clavel et al., 2019). To run the LB model with *mvgl*s, we set the upper bound of the EB
492 search algorithm to 10. For the BMM model, we defined a separate regime for the *EAR* clade
493 (Fig. 2), which encompasses species infecting cichlids of the East African radiation (Fig. 3b).
494 As traditional information criteria such as the Akaike information criterion (AIC) are not
495 applicable to penalised likelihoods (Clavel et al., 2019), we assessed the model performance
496 with the extended information criterion (EIC) (Ishiguro et al., 1997; Kitagawa and Konishi,
497 2010) using the function *EIC* in *mvmorph* with 100 bootstrap replicates. Smaller EIC values
498 were considered indicative of a better model performance. To test for phylogenetic signal in
499 the datasets, we also inferred Pagel's λ (Pagel, 1999) through the *lambda* option in *mvgl*s.

500 Finally, we modelled the character evolution of discrete characters proposed in previous
501 studies (see *Morphological and host range data collection*) under a univariate Pagel's λ
502 model (Pagel, 1999) using the *fitDiscrete* function in the *R* package *geiger* v2.0 (Pennell et
503 al., 2014) to assess the effects of data discretisation. We applied this function to the same set
504 of 100 randomly selected BI subset trees fitted to the morphometric data to account for
505 phylogenetic uncertainty. Model performance was assessed through the sample size–
506 corrected Akaike information criterion in relation to a white noise model (ΔAICc) that
507 assumes absolute phylogenetic independence (Pennell et al., 2014). We also produced
508 character maps for the discrete morphological characters (HC, VBS, AS) and the equivalent
509 sets of continuous characters using *asr_max_parsimony*, *anc.ML*, *ggplot2*, and *ggtree* as
510 mentioned above.

511 RESULTS

512 *Phylogenetic analyses: Molecular data*

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

513 We generated 60 sequences for 33 species and 64 specimens of *Cichlidogyrus* and
514 *Scutogyrus* including 47 28S, 14 combined 18S and ITS1 rDNA, and five COI mtDNA
515 sequences. For 21 of these species, these sequences are the first to be published representing
516 a 57% increase in taxon coverage to 42% of all currently described species, 45% of described
517 species of *Cichlidogyrus* and 71% of *Scutogyrus*. The alignments include 93 sequences/820
518 base pairs for 28S rDNA, 54 sequences/481 base pairs for 18S rDNA, 57 sequences/383 base
519 pairs for ITS rDNA, and 19 sequences/489 base pairs for COI mtDNA. The combined data
520 set includes sequences of 103 specimens belonging to 58 species (see Appendix 3), and has a
521 length of 2173 base pairs following the removal of poorly aligned positions and divergent
522 regions (missing sequence data were treated as gaps). We deposited sequences generated for
523 this study in GenBank (XXXXXX – XXXXXX). All sequences are listed in Appendix 3
524 including the respective GenBank accession numbers, the host species, and the sampling
525 locations. The concatenated alignments and MP, BI, and ML tree topologies can be accessed
526 at Treebase (www.treebase.org, accession number: XXXXXX). Substitution models used for
527 the different partitions are provided in Appendix 4.

528 The BLK weightings scheme produced the most stable tree topologies for the
529 molecular markers (distortion coefficients and SPR distances approximately 1.000) for all k
530 values. For the combined parsimony analysis, k = 48 produced the most stable tree topology
531 (distortion coefficient: 0.901; SPR distance: 0.743 with 46 SPR moves). Thus, the final
532 hypotheses (molecular and combined trees) were inferred from the strict consensus the trees
533 produced under these parameters (BLK; k = 48) (Fig. 2a; Fig. 4). Model selection for BI and
534 ML analyses resulted in the merging of all COI codon positions. The BI and ML
535 concatenated tree topologies were congruent with the final MP consensus tree (Kendall's W
536 = 0.76, $\chi^2 = 12066$, $p < 0.01$). Thus, we display support values of the ML analysis (UFBoot)
537 alongside the posterior probabilities in the BI phylogram (Fig. 2b).

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van

Steeberge M, Koblmüller S, Sturmhuber C, Huyse T, Smeets K, Artois T, Vanhove MPM

538 We found 11 well-supported clades within *Cichlidogyrus* based on the MP analysis of
539 the molecular dataset (Fig. 2a). These clades were also supported the BI and ML analyses
540 suggesting a high stability of the phylogenies produced by the molecular dataset provided
541 here. The clades and their respective node support values (GC value/PP/UFBboot/SH-
542 aLRT) are (Fig. 2a): the 'EAR' clade (27/1/99/99) (#1), the 'CPO' clade (91/1/100/100) (#2),
543 the 'Hemi' clade (94/1/100/100) (#3), the 'Tilapiae' clade (99/1/100/100) (#4), the 'Oreo1'
544 clade (76/1/99/98) (#5), the 'Halli' clade (100/1/100/100) (#6), *Scutogyrus* (86/0.99/*/90)
545 (#7), the 'Cop' clade (48/1/100/100/*) (#8), the 'Bulb' clade (92/1/100/100) (#9), the 'Oreo2'
546 clade (81/1/97/94) (#10), and the basal 'Tylo' clade (100/1/100/100) (#11) (see
547 *Characterisation of species groups* for details of morphology and host range). Beyond the
548 well-supported monophyly of all species excluding the *Tylo* group (100/1/100/100), the MP
549 tree shows the *Bulb* clade as sister group to all the other clades (51/*/*/*). Furthermore, *C.*
550 *tilapiae* and *C. halli* are positioned as sister clades of *Hemi* and *Oreo1* respectively albeit
551 with moderate support (GC values of 11 and 9). The BI and MP analyses also provides partial
552 support for two main lineages containing *Oreo1*, *Cop*, *Scutogyrus*, and *Hemi* (0.99/*/*/*) or
553 *Cop*, *Scutogyrus*, and *Hemi* (*/*/87/*). Both main lineages would also include *C. tilapiae*
554 Paperna, 1960.

555 *Literature survey and character mapping*

556 For a more extensive classification of species of *Cichlidogyrus* into the proposed
557 species groups, we developed four new discrete characters for the reproductive organs: the
558 shape of the penis, its diameter, the shape of the accessory piece, and the shape of the
559 sclerotised vagina (for the respective character states, see Appendix 2). We mapped all
560 morphometric and the newly proposed discrete characters on the molecular MP phylogeny
561 (Fig. 5). Out of the 11 species groups (Fig. 2a), we found shared morphological

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

562 characteristics in eight. Species of the *Oreo2* and the *EAR* groups shared no apparent features
563 in their attachment and reproductive organs with species of the same group apart from the
564 characteristics mentioned in the genus diagnosis (Pariselle and Euzet, 2009). We were able to
565 affiliate all but 12 species without available DNA sequences to these clades (Appendix 5)
566 based on character states shared with species included in the molecular phylogeny. Some
567 morphological structures were particularly relevant as synapomorphies of the species groups
568 (Appendix 6; Appendix 7; also Fig. 5a and Fig. 6). Based on the number of species groups
569 defined by a structure, the male copulatory organ (MCO) was the most distinctive structure as
570 phenotypes were specific to eight species groups (and several subgroups) (Fig. 5b; Fig. 6;
571 Appendix 6). The first hook pair (U1) was enlarged in four species groups (Fig. 6) and one
572 subgroup of the morphologically diverse *EAR* group. The length of the auricles of the dorsal
573 bar (DB h) was distinct in two species groups (Fig. 6) and one subgroup of the *EAR* group.
574 Hook pairs 3–7 were enlarge in two groups (Fig. 6). For the detailed morphological
575 characterisations of all mentioned groups and a discussion on the host and geographical
576 ranges, see *Characterisation of species groups* as well as several character maps (Fig. 5) and
577 an overview of the morphological features (Fig. 6) (full species names are cited in Appendix
578 5 and reported alongside host names; for abbreviations of morphometric measures, see Fig.
579 1c).

580 *Characterisation of species groups*

581 In the following section we summarise the results of the literature survey. All continuous
582 and discrete characters and their states are explained in Appendix 2. Ancestral character
583 states for the different species groups (and subgroups) are listed in Appendix 6 and Appendix
584 7. The following characterisations also discuss morphological differences to species groups
585 with similar characteristics. We also refer to publications that have previously mentioned or
586 discussed shared characteristics of these species groups. For 12 species, we found no

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van

Steeberge M, Koblmüller S, Sturmhuber C, Huysse T, Smeets K, Artois T, Vanhove MPM

587 morphological similarities that could place them near any of the species groups characterised
588 above.

589 1. *EAR: Species infecting cichlids from the East African Radiation*. The *EAR* species
590 group comprises a large heterogeneous group of species that infect cichlids from East African
591 lineages, i.e. the East African Radiation (Schedel et al., 2019). The *EAR* clade consists of
592 multiple potential subgroups based on morphological features (Fig. 6) and node support
593 values from the phylogenetic analyses. The subgroups of the *EAR* group include species with
594 a long heel infecting non-haplochromine cichlids such as bathybatine, ectodine,
595 boulengerochromine, and possibly cyphotilapiine cichlids (*Heel*), species infecting
596 cyprichromine and trematocarine and potentially eretmodine cichlids (*CT*), and species
597 infecting tropheine cichlids (*Troph*) (see Fig. 6). The morphology of the attachment and
598 reproductive organs vary substantially between and within the subgroups (Fig. 5b). Hence,
599 we observed no characteristic features for the whole group with rather general and simple
600 ancestral features (e.g. small marginal hooks; straight, simple penis; no sclerotised vagina,
601 see Appendix 6 and Appendix 7). However, some subgroups display the following shared
602 morphological features.

603 • *Troph* (90*/99/80). These species share few similarities and were placed in a subgroup
604 due to the host species, which all belong to the tropheine radiation. Species without
605 available DNA sequences, infecting tropheines, and sharing morphological features with
606 species in this group, were also added to this subgroup. Dorsal anchor unlike ventral with
607 elongated inner root (except for *C. antoineparisellei*) and deeper indentation between the
608 roots. However, indentation generally shallow, sometimes with fused roots (see parasites
609 of *Interochromis loocki* including *C. antoineparisellei*, *C. buescheri*, *C.*
610 *schreyenbrichardorum*, *C. vealli* on the ventral and *C. antoineparisellei* and *C. buescheri*

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

611 on the dorsal anchor) (Fig. 6). Dorsal bar with well-developed but not elongated auricles.
612 Hooks generally small but can be slightly more developed (e.g. *C. irenae*, *C. gistelincki*,
613 *C. masilyai*, *C. salzburgeri*). The MCO consists of penis and accessory piece, both
614 variably shaped. Penis thin and arched (*C. antoineparisellei*, *C. buescheri*) or looped (*C.*
615 *schreyenbrichardorum*) or slightly broadened (*C. vealli*) (Fig. 5b; Fig. 6) for parasite of
616 *Interochromis loocki* or thin and slightly sinuous for *C. gistelincki*. Penis broadened to
617 wide and short (reminiscent of *C. halli*) for other species. Accessory piece simple,
618 elongated with few remarkable structures, sometimes with cap-like (*C. gistelincki*),
619 forked (*C. antoineparisellei*, *C. buescheri*, *C. masilyai*, *C. salzburgeri*), or hook-shaped
620 (*C. raeymaekersi*) distal end.

- 621 • *Heel* (97/1/99/97). Dorsal anchor unlike ventral with elongated inner root. Dorsal bar with
622 short auricles. Well-developed first hook pair. Male copulatory organ with short penis
623 sometimes with spirally screw thread-like thickened wall (*C. casuarinus*, *C. centesimus*,
624 *C. nshomboi*), characteristic elongated heel (Fig. 5a) of one third to twice the length of the
625 penis, and reduced filiform or sometimes missing (only *C. centesimus*) accessory piece
626 (Fig. 5b; Fig. 6). The species of the *Heel* subgroup have previously been grouped together
627 (Muterezi Bukinga et al., 2012), most recently by Rahmouni et al. (2018b).
- 628 • *CT* (97/1/100/90). Dorsal and ventral anchors similar in shape but can vary in size (see *C.*
629 *brunnensis*) with deep to shallow (*C. brunnensis*) indentation. Dorsal bar with short
630 auricles. Hooks generally short. Male copulatory organ consists of penis and accessory
631 piece and sometimes a marked heel (*C. brunnensis*). Penis medium-length, broadened,
632 and mostly straight (*C. brunnensis*, *C. evikae*, *C. jeanloujustinei*, *C. milangelnari*, *C.*
633 *sturmbaueri*) with thickened wall. Accessory piece with bifurcate distal end (*C.*
634 *brunnensis*), or in two-parts (*C. evikae*, *C. jeanloujustinei*, *C. milangelnari*). Similarities
635 of some species of this subgroup with species infecting cyphotilapiine and ectodine

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van

Steeberge M, Koblmüller S, Sturmbauer C, Huyse T, Smeets K, Artois T, Vanhove MPM

636 cichlids with elongated auricles (*C. adkoningsi*, *C. discophonum*, *C. glacicremoratus*, *C.*

637 *koblmuelleri*, *C. makasai*, *C. vandekerkhovei*) or elongated marginal hooks (*C.*

638 *rectangulus*, *C. sturmbaueri*) have previously been discussed (Rahmouni et al., 2017,

639 2018) but the phylogenetic analysis provided no support for a monophyletic group

640 including species of the *CT* subgroup and *C. vandekerkhovei* (Fig. 2a).

641 • *Other species.* The remaining species of the *EAR* species group shared no evident features
642 with any of the proposed subgroups. *Cichlidogyrus attenboroughi* appears to be the sister
643 taxon to the *Heel* group with moderate support (56/0.99/*/*) but shares no notable
644 characteristics with these species. Several species infecting cyphotilapiine and ectodine
645 cichlids display a broadened (*C. glacicremoratus*, *C. koblmuelleri*) or straight (*C.*
646 *adkoningsi*, *C. rectangulus*, *C. sturmbaueri*) penis with an accessory piece in two-parts
647 (*C. glacicremoratus*, *C. koblmuelleri*) or with a bifurcated distal end (*C. makasai*, *C.*
648 *rectangulus*, *C. sturmbaueri*, *C. vandekerkhovei*) similar to species of the *CT* group. Yet
649 many of these species also display elongated auricles at the dorsal bar (*C. adkoningsi*, *C.*
650 *discophonum*, *C. glacicremoratus*, *C. koblmuelleri*, *C. makasai*) similar to *C.*
651 *vandekerkhovei*, which appears to be unrelated to species of the *CT* group (Fig. 2a) but
652 also displays a bifurcate distal end. According to the phylogenetic analyses (Fig. 2a), *C.*
653 *consobrini* and *C. gillardinae* form a monophyletic group. However, we found few shared
654 characteristics except for a simple tubular penis with a small heel and a simple accessory
655 piece. These characteristics might place these two species close to *C. haplochromii* and *C.*
656 *longipenis*, both of which have been reported from haplochromine cichlids in Lake
657 Victoria (Pariselle and Euzet, 2009). Yet these simple features might also reflect ancestral
658 character states of the *EAR* group (Appendix 6; Appendix 7).

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

659 2. *CPO*: Species infecting coptodonine, pelmatolapiine, oreochromine and other cichlids.

660 Dorsal anchor with inner root longer than outer root and V-shaped indentation between the
661 roots but inner root of *C. arthracanthus* considerably longer than of other species. Roots of *C.*
662 *cubitus* and *C. louipaysani* fused. Ventral anchor similar shape to dorsal anchor but slightly
663 larger. However, roots of ventral anchor of *C. tiberianus* fused. Dorsal bar slightly arched
664 with large medium-sized to large auricles. Ventral bar always with membranous extensions.
665 First hook pair (U1) small and hook pairs 3–7 (U3–7) very long, i.e. more than triple the size
666 of first pair (Fig. 5a; Fig. 6) but U1 of *C. arthracanthus* large and U3–7 of *C. cubitus* and *C.*
667 *louipaysani* short (Fig. 6). The MCO consists of long, arched, or sometimes spiralled (*C.*
668 *arthracanthus*), tubular penis with a well-marked irregularly shaped heel, and a massive,
669 roughly S-shaped accessory piece (Fig. 5b; Fig. 6) that is frequently connected to the heel.
670 The accessory piece has an extension or thickening at the first turn in the proximal half and
671 frequently displays a folded back (*C. paganoi*, *C. vexus*), straight and pointy (*C. guirali*), or
672 hook-like (*C. bilongi*, *C. douellouae*, *C. ergensi*, *C. gallus*, *C. legendrei*, *C. microscutus*)
673 distal end, or sometimes additional terminations resulting in a furcate ending with two (*C.*
674 *aegypticus*, *C. agnesi*, *C. anthemocolpos*, *C. bouvii*, *C. cubitus*, *C. flexicolpos*, *C. lemoallei*,
675 *C. louipaysani*, *C. ouedraogoi*, *C. testificatus*, *C. thurstonae*, *C. tiberianus*) or three (*C.*
676 *bonhommei*, *C. hemi*, *C. kouassii*) digitations. However, the first turn is never V-shaped or
677 knee-like as in species of the *Hemi* clade and the hook-shaped termination is never sickle-like
678 such as in species of the *Cop* clade. Several species display an auxiliary plate (AuP) close to
679 the distal end of the MCO (Fig. 5a) including *C. aegypticus*, *C. agnesi*, *C. bilongi*, *C. gallus*,
680 *C. guirali* (two pieces), *C. microscutus*, *C. paganoi*, and *C. thurstonae*. A sclerotised vagina
681 is mostly present but has not been reported for *C. arthracanthus* (Paperna, 1960; Ergens,
682 1981). Other species with more developed hook pairs 3–7 include *Cichlidogyrus*
683 *bulbophallus*, *C. flagellum*, *C. lobus*, and *C. ranula*, some species of the *Hemi* clade, and all

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van

Steeberge M, Koblmüller S, Sturmhuber C, Huyse T, Smeets K, Artois T, Vanhove MPM

684 species of the *Heel* subgroup of the *EAR* clade. *Cichlidogyrus bulbophallus* and species of the
685 *Hemi* clade and the *Heel* subgroup have been likened to *C. arthracanthus* because of the
686 additional well-developed first hook pair (U1) (Geraerts et al., 2020). However, all of these
687 species differ from the confirmed clade members as follows.

- 688 • The dorsal and ventral anchors differ in shape and size but are similar for species of the
689 *CPO* clade.
- 690 • The penis is never long, arched, and tubular but short (*Heel*), bulbous (*C. bulbophallus*,
691 *C. flagellum*, *C. ranula*), or draws a loop (Fig. 6) or long spiral (*Hemi*).
- 692 • The accessory piece is not complex and S-shaped with an extension in the proximal half
693 but mostly reduced (*Heel*), simple (*C. bulbophallus*, *C. flagellum*, *C. ranula*), or has a
694 sharp knee-like bend (*Hemi*).

695 Previous phylogenetic studies (Mendlová et al., 2012; Messu Mandeng et al., 2015) have
696 placed *C. cubitus* alongside species of the *CPO* clade with long marginal hooks (“*tiberianus*”
697 group, see Pouyaud et al., 2006) but morphological similarities were not further discussed.

698 *3. Hemi: Species infecting hemichromine cichlids and non-cichlid fishes.* Dorsal
699 anchor with elongated narrow inner root and rounded outer root. Ventral anchor slightly
700 larger with short and robust inner root. Dorsal bar slightly arched with short auricles. Ventral
701 bar arched with small membranous extensions. First hook pair (U1) large and well-developed
702 (Fig. 5a, Fig. 6) except for *C. amieti*. Hook pairs 3–7 medium-sized to large with bases more
703 developed than for second hook pair. The MCO consists of thin and long penis that
704 frequently forms loops (*C. amieti*, *C. cf. bychowskii*, *C. dracolemma*, *C. inconsultans*, *C.*
705 *kmentovae*, *C. nandidae*) (Fig. 6) or spirals with one (*C. calycinus*, *C. teugelsi*) to multiple
706 turns (*C. euzeti*, *C. longicirrus*, *C. polyenso*, *C. sanseoi*) and re-joins the accessory piece in its

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

707 distal portion, and an accessory piece of two distinct portions. These portions can be shaped
708 like a V or a simple spiral with an expanded knee-like bend (*C. amieti*, *C. calycinus*, *C. cf.*
709 *bychowskii*, *C. dageti*, *C. dionchus*, *C. dracolemma*, *C. falcifer*, *C. kmentovae*, *C. nandidae*,
710 *C. teugelsi*) (see Dossou and Birgi, 1984; Pariselle and Euzet, 2004), or a large spiral followed
711 by a non-spiralled distal portion (*C. euzeti*, *C. longicirrus*, *C. polyenso*, *C. sanseoi*). In the
712 presence of the knee-like bend a heel is visible. The sclerotised vagina is long, sinuous
713 (except for *C. dageti*, *C. falcifer*, *C. kmentovae*) or spiralled (*C. longicirrus*, *C. sanseoi*).
714 Species of the *Cop* and *Tylo* groups share some characteristics with the species of the *Hemi*
715 group but have either a substantially shorter penis or do not display the characteristic knee-
716 like bend in the accessory piece. Previous studies have grouped species of the *Hemi* species
717 group together based on phylogenetic (Mendlová et al., 2012; Řehulková et al., 2013) and
718 morphological (Dossou and Birgi, 1984; Pariselle and Euzet, 2004; Jorissen et al., 2018a)
719 analyses.

720 4. *Tilapiae*: *The species complex of Cichlidogyrus tilapiae*. For a detailed
721 characterisation, see species descriptions and characterisations by Paperna (1960) and
722 Douëllou (1993). *C. tilapiae* have both been hypothesised to form a species complex
723 (Pouyaud et al., 2006).

724 5. *Oreol*: *Species infecting oreochromine and coptodonine cichlids*. Anchor pairs
725 similar in shape and size but dorsal anchors somewhat smaller. Anchors with short and
726 broadened outer root and long inner root. Dorsal bar arched with two broad and large
727 auricles. Ventral bar with thin mid-portion, and thickened towards the ends includes
728 triangular membranous extensions. Hooks small with small shaft. The MCO consists of long
729 tubular (*C. cirratus*, *C. giostrai*, *C. mbirizei*, *C. mvogoi*, *C. ornatus*) (Fig. 6) or short and
730 thickened (*C. acerbus*, *C. lagoonaris*, *C. nageus*, *C. njinei*, *C. slembroucki*) penis (Fig. 5b)
731 sometimes with bulbous extension at the base of the tube (*C. njinei*, *C. ornatus*, *C. giostrai*)

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van

Steeberge M, Koblmüller S, Sturmhuber C, Huyse T, Smeets K, Artois T, Vanhove MPM

732 and complicated roughly C-shaped (Fig. 6) accessory piece often with finger or hook-shaped
733 outgrowths and marked heel. Proximal end of accessory piece attached to penis bulb and
734 distal end holds the distal portion or a mid-portion of the penis. Vagina simple (*C. acerbus*,
735 *C. lagoonaris*, *C. slembroucki*), sinuous (*C. giostrai*, *C. njinei*, *C. ornatus*), or coiled (*C.*
736 *cirratus*, *C. mbirizei*, *C. mvogoi*) tube and only slightly sclerotised.

737 6. *Halli*: *The species complex of Cichlidogyrus halli*. For detailed characterisation see
738 original species description by Price and Kirk (1967). *Cichlidogyrus halli* has both been
739 hypothesised to be a species complex (Douëllou, 1993; Jorissen et al., 2018b).

740 7. *Scutogyrus*: *Species infecting oreochromine and pematolapiine cichlids*. This
741 group has been diagnosed as the separate genus *Scutogyrus* including a detailed
742 morphological characterisation (Pariselle and Euzet, 1995). Concerning the haptor
743 morphology, species of *Scutogyrus* display well-developed hook pairs 3–7, elongated auricles
744 at the dorsal bar (Fig. 5a; Fig. 6) and a ventral bar that supports a large plate (Fig. 6). The
745 accessory piece presents a distal flap (Fig. 5b; Fig. 6).

746 8. *Cop*: *Species infecting coptodonine cichlids*. Anchor pairs dissimilar in shape.
747 Dorsal anchor with V-shaped indentation between the roots and an elongated inner root at
748 least four times longer (Fig. 5a) than the outer root. Ventral anchor with shallower
749 indentation, and relatively short roots including a considerably shorter inner root. Dorsal bar
750 thick with large and broad auricles. Ventral bar V-shaped and truncated at the ends and
751 thinner towards the mid-portion, displays membranous extensions. First hook pair (U1) well-
752 developed (Fig. 5a, Fig. 6). Hook pairs 3–7 (U3–7) small. The MCO consists of short tubular
753 penis with a marked heel and an accessory piece with a sickle-like terminal hook connected
754 to the basal bulb of the penis. No sclerotised vagina has been observed for species belonging

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

755 to this group except for *C. digitatus* in the original description (Dossou 1982). However, none
756 of the more recent reports of species of the *Cop* clade (Pariselle and Euzet, 1996, 2003)
757 remark on this report. Previous studies have grouped species of the *Hemi* clade, e.g. *C.*
758 *dionchus* (Pariselle and Euzet, 2003), and the *Tylo* clade, e.g. *C. kothiasi* (Pariselle et al.,
759 2013), with the species of the *Cop* clade because of the similar morphology of the hook pairs,
760 i.e. large U1 and small U3–7. However, none of these species displays an accessory piece
761 with a sickle-like terminal hook. Species of the *Cop* clade have previously been grouped
762 together (Pariselle and Euzet, 2003; Jorissen et al., 2018b) but without *C. berminensis*.

763 9. *Oreo2*: Two species infecting oreochromine and coptodonine cichlids. Both species
764 display simple features such as similar anchor pairs and small hook pairs. Most other features
765 differ considerably between the species. Dorsal and ventral anchors are similar in size and
766 shape, the roots are variable in length, i.e. a long inner root in *C. amphoratus* (see drawing
767 Pariselle and Euzet, 1996) or barely to no distinct roots in *C. sclerosus* (Douëllou, 1993).
768 Dorsal bar X-shaped with two auricles of variable shape and length. Ventral bar variable in
769 shape and size with membranous extensions (Mendlová et al., 2012). Hooks small with
770 developed but small shaft except for the second pair (Douëllou, 1993; Pariselle and Euzet,
771 1996). The MCO consists of arched, tubular penis (Fig. 5b) with large heel ('subspherical' in
772 *C. amphoratus* or shaped as a 'serrated plate' in *C. sclerosus*), and associated relatively
773 simple accessory piece (Douëllou, 1993; Pariselle and Euzet, 1996) (Fig 5b). Sclerotised
774 vagina present. *Cichlidogyrus amphoratus* presents a swollen middle portion of the penis
775 reminiscent of species of the *Bulb* clade (Pariselle and Euzet, 1996). However, in the latter
776 group, the dorsal anchors have long inner roots and the first hook pair is well developed (Fig.
777 6). The close relationship of these species has previously been reported but was not further
778 discussed (see Messu Mandeng et al., 2015).

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van

Steeberge M, Koblmüller S, Sturmhuber C, Huyse T, Smeets K, Artois T, Vanhove MPM

779 10. *Bulb: Species from Southern Africa with a bulbous penis*. Dorsal and ventral anchors
780 dissimilar. Dorsal anchors with deep indentation between short outer root and a long inner
781 root (Fig. 6). Dorsal bar short and stout with auricles medium-sized to large (average > 20
782 µm for *C. zambezensis*). First hook pair (U1) well-developed (Fig. 5a; Fig. 6) albeit
783 somewhat less in *C. zambezensis*. Hook pairs 3–7 small to slightly developed. Male
784 copulatory organ consists of a broad penis with a swollen, bulbous middle portion, a small
785 heel (Fig. 6), and a sharp tube-like (*C. philander*) or filiform (*C. maeander*, *C.*
786 *papernastrema*, *C. pseudozambezensis*, *C. zambezensis*) termination, and a simple accessory
787 piece of similar length, which is articulated to the base of the penis bulb (Price et al., 1969;
788 Douëllou, 1993; Jorissen et al., 2018b). *Cichlidogyrus amphoratus*, *C. bulbophallus*, *C.*
789 *flagellum*, *C. giostrai*, *C. karibae*, *C. lagoonaris*, *C. njinei*, *C. ornatus*, and *C. sanjeani* also
790 present a swollen portion in the penis but differ from the species of the group as follows.

- 791 • *C. amphoratus* has similar dorsal and ventral anchors with no well-developed inner root
792 of the dorsal anchor (Fig. 6).
- 793 • The first hooks (U1) of *C. amphoratus* (Fig. 6), *C. flagellum*, *C. giostrai*, *C. karibae*, *C.*
794 *lagoonaris*, *C. njinei*, *C. ornatus*, *C. ranula*, and *C. sanjeani* are considerably shorter and
795 less developed.
- 796 • The hooks U3–U7 of *C. bulbophallus* and *C. ranula* are considerably longer and resemble
797 the marginal hooks of the species of the *CPO* species group.
- 798 • The tubular distal portion of the penis of *C. amphoratus* (Fig. 6), *C. bulbophallus*, *C.*
799 *flagellum*, *C. giostrai*, *C. njinei*, and *C. ornatus* exceed the length of the swollen portion
800 of the penis considerably.

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

- 801 • *C. amphoratus*, *C. giostrai*, *C. lagoonaris*, *C. njinei*, *C. ornatus*, and *C. sanjeani* have
802 only been reported from Western Africa not Southern Africa (see Pariselle and Euzet
803 2009).
- 804 • *C. amphoratus* and *C. njinei* belong to the *Oreo2* and *Oreo1* species groups respectively.
- 805 11. *Tylo*: Species infecting tylochromine cichlids. Dorsal and ventral anchors dissimilar.
806 Roots of anchors frequently fused forming fenestrae (windows), e.g. for *Cichlidogyrus*
807 *berrebii*, *C. chrysopiformis*, *C. dijeto*, *C. kothiasi*, *C. pouyaudi* (Fig. 6), and *C. sergemorandi*
808 (Pariselle and Euzet, 1994; Pariselle et al., 2014; Rahmouni et al., 2018). Dorsal anchors with
809 short outer root and a long inner root, median portion of blade hollow in *C. berrebii*, *C.*
810 *pouyaudi* (Pariselle and Euzet, 1994), and *C. sigmocirrus* (Pariselle et al., 2014). Dorsal bar
811 with two generally short auricles fused to the bar (Fig. 6; Fig. 7a) is one of the most
812 distinctive features of species of this group (Pariselle et al., 2014; Rahmouni et al., 2018).
813 The hooks are generally small but the first pair is developed in *C. bixlerzavalai*, *C.*
814 *chrysopiformis*, *C. dijeto*, *C. kothiasi*, and *C. muzumanii* (Fig. 5a). The MCO consists of a
815 coiled tubular penis, a heel, and a flat ribbon-like (*C. berrebii*, *C. kothiasi*, *C. pouyaudi*, *C.*
816 *sigmocirrus*) (Pariselle and Euzet, 1994; Pariselle et al., 2014), drape-like (*C. chrysopiformis*,
817 *C. dijeto*, *C. mulimbwai*, *C. omari*, *C. sergemorandi*) (Muterezi Bukinga et al., 2012;
818 Pariselle et al., 2014; Jorissen et al., 2018a; Rahmouni et al., 2018), or reduced (*C.*
819 *bixlerzavalai*, *C. muzumanii*) (Muterezi Bukinga et al., 2012; Jorissen et al., 2018a) accessory
820 piece. Unlike other species of *Cichlidogyrus*, the penis and accessory piece of the MCO are
821 separated at the base but might be connected through a filament in *C. chrysopiformis* and *C.*
822 *sigmocirrus* (Pariselle et al., 2014). Species of the *Tylo* group have regularly been grouped
823 together since the descriptions by Pariselle and Euzet (1994) with most recent additions
824 provided by Jorissen et al. (2018a) and Rahmouni et al. (2018b).

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van

Steeberge M, Koblmüller S, Sturmbauer C, Huyse T, Smeets K, Artois T, Vanhove MPM

825 *Phylogenetic position of ‘unsequenced’ species and systematic informativeness: literature*

826 *survey, combined phylogeny, machine learning algorithm, and cluster analysis*

827 The addition of the morphometric data to the parsimony analysis resulted in an overall
828 less resolved and supported tree (Fig. 4). Overall, 30% (24/79) of the ‘unsequenced’ species
829 formed well-supported clades with sequenced species allowing us to affiliate these with the
830 respective clades from the molecular phylogeny. Phylogenetic positions inferred from the
831 combined MP consensus tree were in moderate accordance ($\kappa = 0.48$) with the classifications
832 in species groups inferred from the character maps. Meanwhile, the optimised SVMs ($C = 2^{-7}$,
833 $\sigma = 0.5$) predicted species group affiliation of the sequenced species with moderate success (κ
834 $= 0.68$) with 1035 observations included. For the unsequenced species (833 observations),
835 accordance with the parsimony analysis ($\kappa = 0.36$) and literature survey ($\kappa = 0.31$) was
836 comparatively low.

837 Variable importances of the optimised SVMs indicated an acceptable to excellent
838 discriminatory power on average (Fig. 7a) with exception of the surface of the auxiliary plate
839 (AuP) in the MCO that performed poorly (mean AUC < 0.6). Collapsing the AUC values by
840 species groups (Fig. 7b), indicates an uneven distribution with some groups such as *CPO*
841 being easier to discriminate against their congeners than others such as *Tylo*. Lastly, based on
842 the Pearson pairwise correlation matrix (Fig. 7c), we detected 15 clusters of multicollinear
843 measurements indicating that almost half of the measurements (14) are multicollinear.

844 *Character evolution of the attachment and reproductive organs: morphometrics and*
845 *discretisation*

846 For the attachment organ measurements, we detected a strong phylogenetic signal (λ
847 $= 0.91$; CI: 0.77–1.00). Brownian Motion (BM), Ornstein-Uhlenbeck (OU), and Early-Burst

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

848 (EB) models all performed equally (Fig. 4b). The Late Burst (LB) model performed worse.
849 Only the multi-rate BM model (BMM) performed better albeit with a slight overlap in EIC
850 estimates. For the reproductive organ measurements, the phylogenetic signal was
851 considerably weaker ($\lambda = 0.73$; CI: 0.43–1.00) and no model outperformed the BM model.

852 Out of the discrete characters, only the hook configuration (HC) and ventral bar shape
853 (VBS) produced strong phylogenetic signals (Fig. 4c). In contrast, we detected a relatively
854 weak phylogenetic signal for the anchor similarity (AS) and no signal for the index of
855 specificity (IS). Character maps (Fig. 9a) illustrate these phylogenetic pattern. Based on the
856 literature survey, we suggested a series of measurements that are equivalents of HC, VBS,
857 and AS. The HC summarises the lengths of the marginal hooks (U1–U7), VBS is described
858 the width of the ventral bar in relation to its size ('massive bar' in Mendlová et al., 2012) but
859 also includes information on the presence of membranous extensions, and AS likely refers to
860 a difference in lengths of the inner roots of the dorsal vs. the ventral anchor pair and possibly
861 a difference in size between the anchors pairs based on a drawing in Mendlová et al., (2012)
862 but the publication only mentions a 'difference in shape'. Character maps of these continuous
863 measurements are plotted next to the equivalent discrete characters (Fig. 9b).

864 DISCUSSION

865 We investigated the evolution of the attachment and reproductive organs of parasites
866 infecting a host lineage that has undergone multiple adaptive radiation events (Seehausen,
867 2006). Species of *Cichlidogyrus* (incl. the nested *Scutogyrus*) form a morphologically diverse
868 lineage of parasitic flatworms infecting the gills of African cichlid fishes. This diversity
869 alongside the model system status of the hosts led to the suggestion to use these parasites as a
870 model system for parasite macroevolution (Pariselle et al., 2003; Vanhove et al., 2016). In
871 this context, the present study is the most extensive reconstruction of intra-generic

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van Steeberge M, Koblmüller S, Sturmhuber C, Huysse T, Smeets K, Artois T, Vanhove MPM

872 evolutionary relationships of African metazoan parasites to date providing new
873 morphological, molecular, and host range data, which are made available in public databases.
874 We were able to identify and characterise 11 different species groups (Fig. 2a; Fig. 6)
875 amongst the 137 species that are currently known (Cruz-Laufer et al., 2021). For some of
876 these groups, member species have previously been hypothesised to be closely related, e.g.
877 species within the *Tylo*, *Hemi*, *CPO*, and ‘*Scutogyrus*’ groups. However, this study is the first
878 to summarise all these groups and to analyse the evolution of the attachment and reproductive
879 organ morphology using sequence data of 58 species and morphometric data of 136 species.
880 We here hypothesise the phylogenetic position of described species of *Cichlidogyrus* and
881 *Scutogyrus* without available molecular data through a synthesis of morphological,
882 molecular, and host range data. However, due to missing data, e.g. a limited coverage of the
883 four gene regions and a limited taxon coverage with the majority of species yet undiscovered
884 (Cruz-Laufer et al., 2021), the relationship between the proposed species groups remains
885 mostly unresolved with only moderate support for a few larger clades (Fig. 2a, b). In this
886 light, we opted against taking any nomenclatural act such as splitting the genus into several
887 new genera or synonymising *Scutogyrus* with *Cichlidogyrus* [for a more detailed discussion
888 of this matter and similar cases within Dactylogyridae, see Kmentová et al. (2021a)].

889 *Systematic informativeness of morphometric measurements and host range data*

890 To identify and classify species of *Cichlidogyrus*, previous publications have applied
891 a specific set of measurements of the sclerotised attachment and reproductive organs (Cruz-
892 Laufer et al., 2021). We investigated the systematic informativeness of the 29 measurements
893 most frequently provided in the literature (Fig. 1c). Generally, the measurements appear to
894 have a low systematic informativeness as a substantial portion of unsequenced species (70%)
895 could not be placed in any of the clades inferred from the molecular phylogeny through a

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

896 parsimony analysis combining molecular and morphometric data. Phylogenetic positions
897 inferred from the combined MP consensus tree were in moderate accordance with the
898 classifications based on character maps that considered additional discrete characters for the
899 reproductive organs proposed here (Fig. 5b). The limited systematic informativeness of the
900 measurements is also reflected by the performance of the classification analysis using support
901 vector machines (SVMs). The optimised machine learning algorithm predicted species group
902 affiliation within the sequenced species with moderate success ($\kappa = 0.48$). Yet for the
903 unsequenced species, accordance of the SVM-based classification with the parsimony
904 analysis ($\kappa = 0.36$) and literature survey ($\kappa = 0.31$) was comparatively low. One cause for this
905 performance might be the low number of systematically informative measurements in the
906 dataset as ten of these showed a moderate or low discriminatory power on average ($AUC <$
907 0.8 , see Fig. 7a). Furthermore, the discriminative power of the measurements varied
908 considerably between the different clades (Fig. 7b).

909 Notably, the phylogenetic signal was low for any of the evolutionary models applied
910 to the reproductive organ in the context of the phylogenetic comparative method. This lack of
911 information presents a stark contrast to the importance of the reproductive organ morphology
912 to the character map-based classification. The phenotype of the male copulatory organ was
913 crucial for characterising groups of related species in the form of several additional discrete
914 characters proposed here (Appendix 2). These characters allowed us to consider information
915 on the shape and diameter of the reproductive organs not included in the parsimony analysis
916 and SVM-based approach. Therefore, we could provide hypotheses for all but 12 species
917 regarding their phylogenetic position. This relative success highlights the value of
918 phylogenetic information conveyed by reproductive organ structures; e.g. Rahmouni et al.
919 (2017) noted that the heel of the male copulatory organ was shown to be similarly shaped in
920 possibly related Eastern African species. Similar conclusions have been drawn for other

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van

Steeberge M, Koblmüller S, Sturmbauer C, Huyse T, Smeets K, Artois T, Vanhove MPM

921 dactylogyrid monogeneans such as species of *Characidotrema* Paperna & Thurston, 1968
922 (Řehulková et al., 2019). Regarding the lack of a phylogenetic signal of reproductive organ
923 measurements, Pouyaud et al. (2006) proposed that this issue might be explained by a fast
924 evolutionary rate but we suggest that this observation might be rather linked to a limited
925 taxon coverage as more species have been described since. Instead, the low systematic
926 informativeness indicated here more likely reflects a lack of systematically informative
927 metrics as, for example, the number of reproductive organ measurements is low compared to
928 the attachment organ (5 vs. 24). The measurements that are currently employed, fail to
929 incorporate the systematic information offered by the attachment and reproductive organ
930 morphology.

931 The low predictive power of the morphometrics has encouraged the use of loosely
932 defined discrete characters in the past (Fig. 9). Several studies have categorised species
933 according to their attachment organ morphology (Vignon et al., 2011; Mendlová et al., 2012)
934 and host repertoires (Mendlová and Šimková, 2014). Here, we compared the systematic
935 informativeness of these discrete characters against their continuous counterparts (Fig. 9).
936 Two of the discrete characters lacked a phylogenetic signal (Fig. 8b) despite the moderate
937 discriminatory power of the continuous characters suggested by the machine learning
938 algorithm (Fig. 7a), a problem possibly caused by merging the information of multiple
939 continuous variables into a single discrete character (Fig. 9). Discretisation can also introduce
940 a researcher bias and produce misleading results. For instance, Mendlová et al. (2012)
941 proposed that similar anchor shapes were an ancestral state of the binary character ‘shape of
942 anchors’ (‘anchor similarity’ in the present study) based on a character map but the
943 phylogenetic informativeness of the ‘shape of anchors’ character was never investigated
944 further. In fact, the cluster analysis of the present study shows that dorsal anchor

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

945 measurements and the respective ventral anchor measurements are highly multicollinear (Fig.
946 7c) with the exception of the inner roots (DAd and VAd). The ventral bar shape was also
947 introduced as a variable for character mapping but no ancestral state could be inferred
948 (Mendlová et al., 2012). Yet species of *Scutogyrus* have a unique phenotype with an extended
949 plate associated to the bar structure (Pariselle and Euzet, 1995). These inconsistent results
950 between continuous characters and their discretised counterparts underline that interpretations
951 based on these discretised characters should be treated with caution. For instance, the hook
952 configuration character states introduced by Vignon et al. (2011) are founded on gaps
953 detected between the hook measurements of different species (see Pariselle and Euzet, 2009).
954 New species discoveries in the last decade demonstrate that these character states might not
955 be as clear-cut (see Rahmouni et al., 2017; Geraerts et al., 2020). Nonetheless, hook
956 measurements were highly relevant in a machine learning context (Fig. 7a) indicating that,
957 while an update to the coding of this character might be required, some character states might
958 still prove useful for characterising species groups of *Cichlidogyrus*, e.g. the monophyletic
959 *CPO* group species with well-developed hook pairs 3–7. If researchers wish to use discretised
960 morphological characters in the future, character states should be redefined through
961 quantitative character coding techniques such as proposed by Garcia-Cruz and Sosa (2006).
962 The resulting character states would be more independent from researcher bias, which could
963 increase the repeatability of such studies.

964 A discrete variable proposed to represent host specificity has also led to doubtful
965 conclusions. Mendlová and Šimková, (2014) used the index of specificity (IS) to summarise
966 the phylogenetic specificity of species of *Cichlidogyrus* (and *Scutogyrus*) into four categories
967 including species-, genus-, and tribe-specific species, and generalists (Fig. 3a). Their study
968 concluded that host specificity is related to parental care of the cichlid hosts. However, this
969 observation was made due to an inherent sampling bias towards hosts raising their offspring

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van

Steeberge M, Koblmüller S, Sturmbauer C, Huyse T, Smeets K, Artois T, Vanhove MPM

970 through mouthbrooding (Cruz-Laufer et al., 2021), a specific type of parental care. Mendlová
971 and Šimková (2014) also observed a correlation of the IS with the host phylogeny. However,
972 this observation might again reflect a study bias in cichlid parasitology, in this case towards
973 large economically relevant host species (Cruz-Laufer et al., 2021). Because of their
974 importance to fisheries, these fishes unlike others have been introduced to various new
975 habitats. Anthropogenic introductions might also increase the chances of co-introduction for
976 their parasites and create ecological opportunities to expand host repertoires (e.g. Jorissen et
977 al., 2020). To avoid these biases, future studies should treat host specificity as a complex
978 parameter encompassing structural, phylogenetic, and geographical aspects (Poulin et al.,
979 2011) as well as temporal fluctuations (Brooks et al., 2019). We suggest that these parameters
980 could be included, e.g., in a genus-wide comparative study in the future.

981 In general, the size of morphological structures should first be treated as continuous,
982 unmodified characters before further processing to avoid losing systematic information.
983 Discretisation should only be applied with sufficient statistical evidence to back up character
984 coding. We suggest increasing sampling and sequencing efforts to address data deficiency.

985 *Attachment organ evolution: rate heterogeneity, co-divergence, and host switching*

986 To infer evolutionary patterns from the attachment organ morphology, we applied a
987 penalised likelihood framework for multivariate phylogenetic comparative methods (PCMs)
988 to the continuous character data. The evolutionary models representing different evolutionary
989 hypotheses performed almost equally well (Fig. 4b). However, we found no significantly
990 increased performance of the Ornstein-Uhlenbeck (OU), Early Burst (EB), and Late Burst
991 (LB) models compared to the Brownian Motion (BM) model. Only the multi-rate BM
992 (BMM) model performed slightly but significantly better than the other models (Fig. 4b). The

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

993 OU, EB, and LB models present extensions of the Brownian Motion (BM) model (Butler and
994 King 2004; Harmon et al. 2010). Therefore, the character evolution of the attachment organs
995 measurements might simply follow a pattern associated with genetic drift or randomly
996 fluctuating natural selection (Losos, 2008). This pattern might also be reflected in
997 morphological features that are frequently shared by related species (Fig. 6). The relatedness
998 of the parasite species could determine the evolution of attachment organ as suggested by a
999 previous study (Vignon et al. 2011). Studies on a related lineage of monogeneans
1000 (*Ligophorus* Euzet & Suriano, 1977) also indicate a correlation of morphometrics and
1001 phylogenetic relationships (Rodríguez-González et al. 2017). However, the increased support
1002 for the BMM model suggests that the evolution of these parasites might have been shaped by
1003 different evolutionary regimes. This rate heterogeneity appears logical as host lineages
1004 outside the Great African lakes have not experienced the same explosive speciation as their
1005 East African relatives (see Brawand et al., 2015). In fact, purely single-rate models appear
1006 increasingly unlikely to capture the real evolutionary history of *Cichlidogyrus* because of its
1007 pan-African distribution (Cruz-Laufer et al., 2021) and ecological diversity of hosts (Burress,
1008 2014). In the present study, we only modelled a multi-rate BM model because a lack of
1009 software packages suited for high-dimensional dataset with pre-set hypothesis. Yet other
1010 multi-rate models including OU, EB, and LB processes could provide better fits. In
1011 particular, species of the *EAR* group (Fig. 2a) that infect cichlid radiations in East Africa (Fig.
1012 3) might mirror the explosive speciation of their hosts (see Vanhove et al., 2015). In recent
1013 years, multiple tools have been developed to investigate these more complex models but
1014 often these methods represent maximum likelihood (ML) approaches (e.g. Beaulieu et al.,
1015 2012; Ingram and Mahler, 2013; Puttick, 2018) that are sensitive to trait covariation, data
1016 orientation, and trait dimensions, particularly for high-dimensional datasets (Adams and
1017 Collyer, 2018). Other methods were developed to detect evolutionary rate shifts (e.g. Ingram

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van

Steeberge M, Koblmüller S, Sturmbauer C, Huyse T, Smeets K, Artois T, Vanhove MPM

1018 and Mahler, 2013; Uyeda and Harmon, 2014; Khabbazian et al., 2016) rather than to test

1019 hypotheses based on previous knowledge (such as for the *EAR* clade in the present case).

1020 Therefore, penalised log-likelihood approaches that address these issues remain currently

1021 limited to multi-rate BM models but future updates will most likely close this technological

1022 gap.

1023 Highly similar morphological characteristics can be found in multiple species groups

1024 (Fig. 5a; Fig. 6) such as a well-developed outer root of the dorsal anchor (e.g. species of the

1025 *Bulb* and *Cop* groups), large first hooks (e.g. species of *Tylo*, *EAR – Heel*, and *Hemi* groups),

1026 and long third to seventh hooks (e.g. species of the *EAR – Heel*, *CPO* groups). Character

1027 shifts observed within species groups appear to be adaptations to divergent host repertoires or

1028 environments (see Messu Mandeng et al., 2015). Host- and environmentally induced shifts in

1029 the attachment organ morphology have also been observed amongst other dactylogyrids, e.g.

1030 species of *Kapentagyris* (Kmentová et al., 2018) and *Thaparocleidus* (Šimková et al., 2013),

1031 and other parasite lineages, e.g. cymothoid isopods (Baillie et al., 2019). For *Cichlidogyrus*,

1032 host species infected by parasites from the same species group frequently belong to the same

1033 tribe or related tribes of cichlids (Fig. 3; Appendix 5). Consequently, the phylogenetic

1034 congruence amongst some clades of *Cichlidogyrus* and their hosts is strong (Vanhove et al.,

1035 2015). Examples for host-induced character shifts amongst species of *Cichlidogyrus* can be

1036 found across the genus:

- 1037 • *Cichlidogyrus amieti* has undergone a recent character shift as it displays small first
1038 hooks unlike other species of the *Hemi* clade, where they are well-developed. This change
1039 might result from a host switch (Messu Mandeng et al., 2015).
- 1040 • *Cichlidogyrus sclerosus*, a generalist infecting fishes belonging to various cichlid tribes
1041 (Fig. 3), shares almost no attachment organ features with its sister species *C. amphoratus*,

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

- 1042 a specialist infecting only coptodonine cichlids (Fig. 3), beyond those shared by all
1043 species of the genus, e.g. auricles associated with the dorsal bar of the attachment organ
1044 (see Pariselle and Euzet, 2009).
- 1045 • The diverse morphological features of species of the *EAR* group might relate to the
1046 ecological and morphological diversity of the hosts, e.g. related to feeding behaviour
1047 (Burress, 2014). The parasites, which we grouped into subgroups of the *EAR* clade, might
1048 have adapted to these ecologically diverse hosts resulting in divergent attachment and
1049 reproductive organ morphologies.
 - 1050 • Co-infections of different species of *Cichlidogyrus* can result in niche segregation on the
1051 gills of a host such as in species infecting Lake Victoria cichlids (Gobbin et al., 2020). At
1052 least some co-infecting species can be told apart based on their attachment organ
1053 morphology (see Gobbin et al., 2021). Thus, phenotypic differences might be linked to
1054 specific infection sites. Morphological adaptations to gill microhabitats have been
1055 reported for other monogeneans (Rohde, 1976; Ramasamy et al., 1985) but no study has
1056 so far investigated these mechanisms for species of *Cichlidogyrus*.

1057 The host repertoires of the parasites offer clues for evolutionary processes frequently
1058 associated with host-induced effects in the evolutionary history of *Cichlidogyrus* including
1059 co-divergence and host switching. For instance, co-divergence might be reflected in the host
1060 (Fig. 3; Appendix 5) and geographical repertoires of species of *Cichlidogyrus* (Vanhove et
1061 al., 2013) (e.g. *EAR* group species are exclusive to Eastern Africa and *Bulb* group species to
1062 Southern Africa). Furthermore, a small yet increasing number of natural and invasion-
1063 induced host switches are reported in the literature from cichlids to non-cichlids, e.g. in the
1064 *Hemi* clade (Birgi and Euzet, 1983; Birgi and Lambert, 1986), and from African cichlids to
1065 cichlids and non-cichlids in the Americas, the Levant, and Madagascar (*C. arthracanthus*, *C.*
1066 *sclerosus*, *C. tiberianus*, *C. tilapiae*, *C. thurstonae*, and *Scutogyrus longicornis*) (Paperna,

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van

Steeberge M, Koblmüller S, Sturmhuber C, Huyse T, Smeets K, Artois T, Vanhove MPM

1067 1960; Jiménez-García et al., 2001; Šimková et al., 2019). While the literature suggests that
1068 most African fish families have their own monogenean lineages (Carvalho Schaeffner, 2018)
1069 and the number of similar host switches might, hence, be limited, more host switches might
1070 be discovered in the future if more (also non-cichlid) host species are studied.

1071 Despite these host-induced character shifts, the evolutionary models employed for
1072 multivariate PCMs detected no changes in the evolutionary rate of the attachment organ
1073 measurements. Apart from the limitation regarding multi-rate models, mentioned above, we
1074 suggest that this observation might be a result of incomplete sampling, which has affected
1075 similar studies in the past. For instance, previous studies were limited to single structures, i.e.
1076 the anchors (Rodríguez-González et al., 2017), or to the much lower number of species of
1077 *Cichlidogyrus* described at the time (77 vs. 130 today), which comprised less phenotypical
1078 diversity than is currently known to exist (Vignon et al., 2011). In fact, the phylogenetic
1079 position of some species with features diverging from their respective species groups (e.g. *C.*
1080 *amieti* from other *Hemi* group species) were not known. Observed similarities of the
1081 attachment organ morphology in related species reflected the specificity of the parasites to
1082 the respective host lineage rather than only the relatedness of the monogeneans. While the
1083 present analysis includes more species, DNA sequence data of less than half of all described
1084 species are available at GenBank and many species are likely to be discovered in the future
1085 (Cruz-Laufer et al., 2021). This data deficiency could affect the performance of the models
1086 implemented through PCMs as, e.g., the lack of lineages might lower the resolution of
1087 phylogenetic inference and increase phylogenetic uncertainty. A more complete taxon
1088 coverage of morphological, molecular, and ecological data will provide more insight into the
1089 evolutionary patterns of the genus. Furthermore, a closer look into the evolution of single
1090 species groups might reveal different patterns than for a genus-wide study. For instance, the

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

1091 high variability in the attachment and reproductive organ morphology of the *EAR* group
1092 might indicate a late burst in the divergence of characters induced by phenotypic adaptation
1093 to the rapidly speciating host lineages. A slightly better performance of the multi-rate model
1094 provides reason to further investigate the evolutionary history of this group, as it might fall
1095 under a divergent evolutionary regime. Lastly, convergent evolution of morphological
1096 characters might also have affected PCM model performance. Convergence can mask
1097 evolutionary processes that have occurred in the past as phenotypes become more similar
1098 (Losos, 2008). Monogenean parasites infecting the haplochromine radiations in the Eastern
1099 African lakes might be particularly impacted by this effect because of the convergent
1100 evolution of their hosts. These fishes have evolved via replicated adaptive radiation events
1101 resulting in similar yet unrelated species occupying the same niches in different lakes
1102 (McGee et al., 2016). This setting makes their parasites also potential targets to study
1103 convergent evolution in parasites. Conversely, the Eastern African diversity of *Cichlidogyrus*
1104 has mostly been investigated in Lake Tanganyika, the oldest lake, limiting the current
1105 possibilities to study parasite evolution in the context of replicate host radiations in this
1106 region [but see Gobbin et al. (2021) for a study focusing on Lake Victoria].

1107 *Reproductive organ evolution*

1108 Unlike the attachment organ, the reproductive organs show no character changes
1109 correlated to host relatedness. Reproductive organ characters are evolutionarily stable within
1110 the different clades as evidenced by the role of the male copulatory organ in the
1111 morphological characterisation of the species groups. The evolution of the reproductive
1112 organs, unlike the attachment organ, has no apparent connection to the host range and seems
1113 to be determined mostly by phylogenetic constraints. Differences within species groups seem
1114 to reflect phylogenetic positions, e.g. the short penis of *C. falcifer* (*Hemi*) and simple
1115 accessory piece of *C. falcifer* and *C. cubitus* (*CPO*) are ancestral states (Fig. 5) of the

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van Steeberge M, Koblmüller S, Sturmbauer C, Huyse T, Smeets K, Artois T, Vanhove MPM

1116 respective clade. We detected a marginally better support for the BMM model (Fig. 8a)
1117 suggesting that the *EAR* group, with its host lineages that have resulted from adaptive
1118 radiation events, might have experienced an evolutionary regime slightly divergent from its
1119 congeners. However, the low phylogenetic signal in these data provides far less confidence in
1120 this result. Previous studies on the reproductive organs of monogenean flatworms have
1121 suggested that the complex surface of the reproductive organ might correlate with
1122 environmental factors, e.g. to avoid separation during copulation through the water flow
1123 (Kearn and Whittington, 2015). This hypothesis would suggest that the sclerotised structures
1124 evolve under selective pressure. However, while the screw thread-like thickening of the penis
1125 wall for some species (*C. casuarinus*, *C. centesimus*, *C. nshomboi*) (Fig. 6, *Heel* subgroup)
1126 might present such an example, we see no general trend to confirm this hypothesis. Neither
1127 generalist species such as *C. halli*, *C. tilapiae*, and *C. sclerosus* nor species infecting
1128 potentially more motile hosts living in fast-flowing waters, e.g. *C. consobrini*, *C. maeander*,
1129 and *C. ranula* infecting species of *Orthochromis*, display more complex MCOs than their
1130 congeners. Consequently, the diversification of the reproductive organs might be explained
1131 by reproductive isolation such as previously suggested for other dactylogyrid monogeneans
1132 (Šimková et al., 2002). A quantitative approach to measure reproductive organ shape could
1133 provide a more detailed answer to this question by capturing the shape variation currently
1134 only recorded as qualitative information (Fig. 5b) in taxonomic literature (see below for
1135 proposals).

1136 *Classification, prediction, and research prospects*

1137 We were able to demonstrate the existing yet limited predictive power of the
1138 measurements currently employed by most taxonomists through a phylogenetic analysis
1139 under maximum parsimony accompanied by a machine learning approach. To improve the

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

1140 accuracy of these predictions overall, datasets with a more complete taxon coverage and
1141 higher number of observations are needed. The results of the character map-based
1142 classification vs. the parsimony analysis and the statistical classification approaches have
1143 highlighted a discrepancy between the systematic informativeness of the attachment and
1144 reproductive organ morphology in taxonomic literature and the low number of
1145 phylogenetically informative measurements. Future studies should address this discrepancy
1146 by increasing the number of systematically informative metrics.

1147 Several possible solutions have been proposed to increase the number of metrics.
1148 First, some studies included novel measures such as the lengths of the shaft of the hooks, the
1149 curved length of the accessory piece, or the width of the penis tube and the bulbous portion of
1150 the penis (e.g. Geraerts et al., 2020) to describe species of *Cichlidogyrus* in more detail.
1151 Second, geomorphometric analyses have been employed for the anchors of some species of
1152 *Cichlidogyrus* (Rahmouni et al., 2021) as well as other monogeneans (Vignon and Sasal,
1153 2010; Rodríguez-González et al., 2017; Kmentová et al., 2020) to infer the most
1154 systematically informative structures. Third, 3D imaging techniques as previously developed
1155 for monogenean sclerotised structures (e.g. García-Vásquez et al., 2012) have been proposed
1156 to improve characterisations of species of *Cichlidogyrus* (Cruz-Laufer et al., 2021) using
1157 methods employed for shape analyses (Klingenberg, 2010; Dryden and Mardia, 2016). Key
1158 characters that could be inferred using this approach are for example elongation (long vs
1159 round), torsion (straight vs coiled), and complexity (few vs many extensions). However, all
1160 of these solutions remain currently unavailable for comparative analyses. Data resulting from
1161 the proposed additional measures, geomorphometric measurements, and 3D imaging
1162 techniques are available for only few morphological structures, species, and studies. Future
1163 research should increase the taxon coverage for proposed methods and improve
1164 quantification methods to address these limitations. Future datasets should encompass an

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van

Steeberge M, Koblmüller S, Sturmbauer C, Huyse T, Smeets K, Artois T, Vanhove MPM

1165 increasing number of DNA sequences and morphometric measurements from all known
1166 species as this study highlights the limitations of using morphological characters to infer the
1167 evolution of monogenean parasites. Nonetheless, raw data should regularly be published in
1168 supplements or data repositories (e.g. Kmentová et al., 2016b) to improve the accessibility of
1169 morphometric data. This approach could enhance information flow and academic capacity for
1170 research into species of monogenean parasites in the future (Cruz-Laufer et al., 2021).

1171 CONCLUSION

1172 Due to the species-richness and the model system status of their cichlid hosts, the
1173 parasite lineage including species of *Cichlidogyrus* and *Scutogyrus* could be an ideal
1174 macroevolutionary model system for speciation in parasitic organisms in the future. Notably,
1175 these species belong to one of the more extensively investigated lineages in close interaction
1176 with an adaptive radiation of other species, namely cichlid fishes. As adaptive radiations
1177 might explain a substantial part of the biodiversity on the planet, many species potentially
1178 interact with radiating lineages. Therefore, understanding evolutionary history in this context
1179 could be key for understanding the diversity of species interactions in general.

1180 Similar to many parasite systems, the limited availability of phenotypic and genotypic
1181 data for the 137 species of *Cichlidogyrus* currently described has hampered investigations
1182 into the evolutionary history of the lineage. This study provides an example of how an
1183 increased availability of morphological and DNA sequence data can promote computational
1184 approaches such as parsimony analyses, phylogenetic comparative methods, and machine
1185 learning classification to answer basic questions of evolutionary research. The
1186 phylogenetically informative measurements (continuous and discrete, established and newly
1187 proposed) and characterisation of multiple clades within *Cichlidogyrus* have highlighted key

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

1188 information required to describe new species as well as the shortcomings of the
1189 morphological characters that are currently used. Features of the attachment organ used to
1190 characterise the proposed species groups remain key descriptors alongside the features of the
1191 reproductive organs. However, they only serve as predictors in a multivariate context. This
1192 conclusion presents a stark contrast to previous classification systems for this taxon, which
1193 used to predict species groups based on single discrete morphological characters.

1194 Using multivariate phylogenetic comparative methods, we were able to model
1195 evolutionary mechanisms of the attachment and reproductive organs suggested by previous
1196 studies. No support for alternative models suggests that the variation of the morphology of
1197 these organs across the genus has been shaped by different evolutionary regimes.
1198 Morphological patterns in different species groups further indicate that the attachment organ
1199 morphology can be shaped by host parameters. In contrast, the reproductive organs appear to
1200 follow a more random evolutionary regime. Other mechanisms might play a role in the
1201 evolution of this parasite lineage, e.g. within-host speciation as suggested by other studies.
1202 Therefore, more morphometric and DNA sequence data will be needed to provide a more
1203 detailed picture of the evolution of the genus. We encourage researchers to publish
1204 morphometric raw data regularly to improve the data accessibility for the scientific
1205 community, and to sequence systematically informative DNA regions of the remaining
1206 described or undiscovered species.

1207

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van Steeberge M, Koblmüller S, Sturmbauer C, Huyse T, Smeets K, Artois T, Vanhove MPM

1208 REFERENCES

- 1209 Adams, D.C., Collyer, M.L., 2018. Multivariate phylogenetic comparative methods:
1210 evaluations, comparisons, and recommendations. *Syst. Biol.* 67, 14–31.
- 1211 Altman, D.G., Royston, P., 2006. The cost of dichotomising continuous variables. *Br. Med. J.*
1212 332, 1080.
- 1213 Baillie, C., Welicky, R.L., Hadfield, K.A., Smit, N.J., Mariani, S., Beck, R.M.D., 2019.
1214 Hooked on you: shape of attachment structures in cymothoid isopods reflects parasitic
1215 strategy. *BMC Evol. Biol.* 19, 207.
- 1216 Barson, M., Přikrylová, I., Vanhove, M.P.M., Huyse, T., 2010. Parasite hybridization in
1217 African *Macrogyrodactylus* spp. (Monogenea, Platyhelminthes) signals historical host
1218 distribution. *Parasitology* 137, 1585–1595.
- 1219 Beaulieu, J.M., Jhwueng, D.C., Boettiger, C., O’Meara, B.C., 2012. Modeling stabilizing
1220 selection: expanding the Ornstein-Uhlenbeck model of adaptive evolution. *Evolution* 66,
1221 2369–2383.
- 1222 Birgi, E., Euzet, L., 1983. Monogènes parasites des poissons des eaux douces du Cameroun.
1223 Présence des genres *Cichlidogyrus* et *Dactylogyrus* chez *Aphyosemion*
1224 (*Cyprinodontidae*). *Bull. la Soc. Zool. Fr.* 108, 101–106.
- 1225 Birgi, E., Lambert, A., 1986. Présence chez un *Nandidae* (Téléostéen), *Polycentropsis*
1226 *abbreviata* Boulenger, 1901, du genre *Cichlidogyrus* (Monogenea, Monopisthocotylea,
1227 *Ancyrocephalidae*). *Ann. Parasitol. Hum. Comp.* 61, 521–528.
- 1228 Bitja-Nyom, A.R., Agnèse, J.F., Pariselle, A., Bilong-Bilong, C.F., Gilles, A., Snoeks, J.,
1229 2021. A systematic revision of the five-spotted *Hemichromis* complex (Cichliformes:

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

- 1230 Cichlidae) from West Africa and Lower Guinea, with the description of a new species
1231 from Cameroon. *Hydrobiologia* 5, 3779–3803.
- 1232 Blažek, R., Polačik, M., Smith, C., Honza, M., Meyer, A., Reichard, M., 2018. Success of
1233 cuckoo catfish brood parasitism reflects coevolutionary history and individual
1234 experience of their cichlid hosts. *Sci. Adv.* 4, eaar4380.
- 1235 Blomberg, S.P., Garland, T., Ives, A.R., 2003. Testing for phylogenetic signal in comparative
1236 data: behavioral traits are more labile. *Evolution* 57, 717–745.
- 1237 Brawand, D., Wagner, C.E., Li, Y.I., Malinsky, M., Keller, I., Fan, S., Simakov, O., Ng,
1238 A.Y., Lim, Z.W., Bezault, E., Turner-Maier, J., Johnson, J., Alcazar, R., Noh, H.J.,
1239 Russell, P., Aken, B., Alföldi, J., Amemiya, C., Azzouzi, N., Baroiller, J.F., Barloy-
1240 Hubler, F., Berlin, A., Bloomquist, R., Carleton, K.L., Conte, M.A., D’Cotta, H., Eshel,
1241 O., Gaffney, L., Galibert, F., Gante, H.F., Gnerre, S., Greuter, L., Guyon, R., Haddad,
1242 N.S., Haerty, W., Harris, R.M., Hofmann, H.A., Hourlier, T., Hulata, G., Jaffe, D.B.,
1243 Lara, M., Lee, A.P., MacCallum, I., Mwaiko, S., Nikaido, M., Nishihara, H., Ozouf-
1244 Costaz, C., Penman, D.J., Przybylski, D., Rakotomanga, M., Renn, S.C.P., Ribeiro, F.J.,
1245 Ron, M., Salzburger, W., Sanchez-Pulido, L., Santos, M.E., Searle, S., Sharpe, T.,
1246 Swofford, R., Tan, F.J., Williams, L., Young, S., Yin, S., Okada, N., Kocher, T.D.,
1247 Miska, E.A., Lander, E.S., Venkatesh, B., Fernald, R.D., Meyer, A., Ponting, C.P.,
1248 Streelman, J.T., Lindblad-Toh, K., Seehausen, O., Di Palma, F., 2015. The genomic
1249 substrate for adaptive radiation in African cichlid fish. *Nature* 513, 375–381.
- 1250 Brooks, D.R., Hoberg, E.P., Boeger, W.A., 2019. *The Stockholm Paradigm: climate change*
1251 *and emerging disease*. University of Chicago Press, Chicago, USA.
- 1252 Burress, E.D., 2014. *Cichlid fishes as models of ecological diversification: patterns,*
1253 *mechanisms, and consequences*. *Hydrobiologia* 748, 7–27.

- Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van Steeberge M, Koblmüller S, Sturmhuber C, Huysse T, Smeets K, Artois T, Vanhove MPM
- 1254 Butler, M.A., King, A.A., 2004. Phylogenetic comparative analysis: A modeling approach for
1255 adaptive evolution. *Am. Nat.* 164, 683–695.
- 1256 Campbell, V., Legendre, P., Lapointe, F.J., 2011. The performance of the Congruence
1257 Among Distance Matrices (CADM) test in phylogenetic analysis. *BMC Evol. Biol.* 11.
- 1258 Caña-Bozada, V., Llera-Herrera, R., Fajer-Ávila, E.J., Morales-Serna, F.N., 2021.
1259 Mitochondrial genome of *Scutogyrus longicornis* (Monogenea: Dactylogyridea), a
1260 parasite of Nile tilapia *Oreochromis niloticus*. *Parasitol. Int.* 81, 102281.
- 1261 Carvalho Schaeffner, V.C., 2018. Part 5: Host-parasite list. In: Scholz, T., Vanhove, M.P.M.,
1262 Smit, N., Jayasundera, Z., Gelnar, M. (Eds.), *A Guide to the parasites of African
1263 freshwater fishes*. CEBioS, Royal Belgian Institute of Natural Sciences, Brussels,
1264 Belgium, pp. 361–402.
- 1265 Chernomor, O., von Haeseler, A., Minh, B.Q., 2016. Terrace aware data structure for
1266 phylogenomic inference from supermatrices. *Syst. Biol.* 65, 997–1008.
- 1267 Clavel, J., Aristide, L., Morlon, H., 2019. A Penalized Likelihood Framework for High-
1268 Dimensional Phylogenetic Comparative Methods and an Application to New-World
1269 Monkeys Brain Evolution. *Syst. Biol.* 68, 93–116.
- 1270 Clavel, J., Escarguel, G., Merceron, G., 2015. mvMORPH: an R package for fitting
1271 multivariate evolutionary models to morphometric data. *Methods Ecol. Evol.* 6, 1311–
1272 1319.
- 1273 Cruz-Laufer, A.J., Artois, T., Smeets, K., Pariselle, A., Vanhove, M.P.M., 2021. The cichlid–
1274 Cichlidogyrus network: a blueprint for a model system of parasite evolution.
1275 *Hydrobiologia* 848, 3847–3863.

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

- 1276 de Meeûs, T., McCoy, K.D., Prugnolle, F., Chevillon, C., Durand, P., Hurtrez-Boussès, S.,
1277 Renaud, F., 2007. Population genetics and molecular epidemiology or how to *débusquer*
1278 *la bête*. *Infect. Genet. Evol.* 7, 308–332.
- 1279 Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G.,
1280 Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., McClean, C., Osborne, P.E.,
1281 Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D., Lautenbach, S., 2013.
1282 Collinearity: A review of methods to deal with it and a simulation study evaluating their
1283 performance. *Ecography* 36, 27–46.
- 1284 Dossou, C., 1982. Parasites de poissons d’eau douce du Bénin. III. Espèces nouvelles du
1285 genre *Cichlidogyrus* (Monogenea) parasites de Cichlidae. *Bull. Inst. Fondam. Afr. Noire*
1286 *Ser. A Sci. Nat.* 44, 295–322.
- 1287 Dossou, C., Birgi, E., 1984. Monogènes parasites d’*Hemichromis fasciatus* Peters, 1857
1288 (Teleostei, Cichlidae). *Ann. Sci. Nat. Zool.* 6, 101–109.
- 1289 Douëllou, L., 1993. Monogenean of the genus *Cichlidogyrus* Paperna, 1960 (Dactylogyridae:
1290 Ancyrocephalinae) from cichlid fishes of Lake Kariba (Zimbabwe) with description of
1291 five new species. *Syst. Parasitol.* 25, 159–186.
- 1292 Dryden, I.L., Mardia, K. V, 2016. *Statistical shape analysis: with applications in R*, 2nd ed. ,
1293 Wiley series in probability and statistics John Wiley & Sons, Chichester, UK.
- 1294 Dunz, A.R., Schlieven, U.K., 2013. Molecular phylogeny and revised classification of the
1295 haplotilapiine cichlid fishes formerly referred to as “*Tilapia*.” *Mol. Phylogenet. Evol.*
1296 68, 64–80.
- 1297 Ergens, R., 1981. Nine species of the genus *Cichlidogyrus* Paperna, 1960 (Monogenea:
1298 Ancyrocephalinae) from Egyptian fishes. *Folia Parasitol.* 28, 205–214.

- Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van Steeberge M, Koblmüller S, Sturmhuber C, Huyse T, Smeets K, Artois T, Vanhove MPM
- 1299 Fang, Z., Fan, J., Chen, X., Chen, Y., 2018. Beak identification of four dominant octopus
1300 species in the East China Sea based on traditional measurements and geometric
1301 morphometrics. *Fish. Sci.* 84, 975–985.
- 1302 Fannes, W., Vanhove, M.P.M., Huyse, T., Paladini, G., 2015. A scanning electron
1303 microscope technique for studying the sclerites of *Cichlidogyrus*. *Parasitol. Res.* 114,
1304 2031–2034.
- 1305 Felsenstein, J., 1973. Maximum likelihood estimation of evolutionary trees from continuous
1306 characters. *Am. J. Hum. Genet.* 25, 471–492.
- 1307 Firmat, C., Alibert, P., Mutin, G., Losseau, M., Pariselle, A., Sasal, P., 2016. A case of
1308 complete loss of gill parasites in the invasive cichlid *Oreochromis mossambicus*.
1309 *Parasitol. Res.* 115, 3657–3661.
- 1310 Forbes, A.A., Powell, T.H.Q., Stelinski, L.L., Smith, J.J., Feder, J.L., 2009. Sequential
1311 sympatric speciation across trophic levels. *Science* 323, 776–779.
- 1312 Garcia-Cruz, J., Sosa, V., 2006. Coding quantitative character data for phylogenetic analysis:
1313 a comparison of five methods. *Syst. Bot.* 31, 302–309.
- 1314 García-Vásquez, A., Shinn, A.P., Bron, J.E., 2012. Development of a light microscopy stain
1315 for the sclerites of *Gyrodactylus* von Nordmann, 1832 (Monogenea) and related genera.
1316 *Parasitol. Res.* 110, 1639–1648.
- 1317 Geraerts, M., Muterezi Bukinga, F., Vanhove, M.P.M., Pariselle, A., Chocha Manda, A.,
1318 Vreven, E.J., Huyse, T., Artois, T., 2020. Six new species of *Cichlidogyrus* Paperna,
1319 1960 (Platyhelminthes: Monogenea) from the gills of cichlids (Teleostei: Cichliformes)
1320 from the Lomami River Basin (DRC: Middle Congo). *Parasit. Vectors* 13, 187.

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

- 1321 Gillardin, C., Vanhove, M.P.M., Pariselle, A., Huyse, T., Volckaert, F.A.M., 2012.
1322 Ancyrocephalidae (Monogenea) of Lake Tanganyika: II: description of the first
1323 Cichlidogyrus spp. parasites from Tropheini fish hosts (Teleostei, Cichlidae). Parasitol.
1324 Res. 110, 305–313.
- 1325 Glor, R.E., 2010. Phylogenetic insights on adaptive radiation. Annu. Rev. Ecol. Evol. Syst.
1326 41, 251–270.
- 1327 Gobbin, T.P., Vanhove, M.P.M., Seehausen, O., Maan, M.E., 2020. Microhabitat
1328 distributions and species interactions of ectoparasites on the gills of cichlid fish in lake
1329 Victoria, Tanzania. Int. J. Parasitol.
- 1330 Gobbin, T.P., Vanhove, M.P.M., Seehausen, O., Maan, M.E., Pariselle, A., 2021. Four new
1331 species of *Cichlidogyrus* (Platyhelminthes, Monogenea, Dactylogyridae) from Lake
1332 Victoria haplochromine cichlid fishes, with the redescription of *C. bifurcatus* and *C.*
1333 *longipenis* [preprint]. bioRxiv 2021.01.29.428376.
- 1334 Goloboff, P.A., 1993. Estimating character weights during tree search. Cladistics 9, 83–91.
- 1335 Goloboff, P.A., 1999. Analyzing large data sets in reasonable times: solutions for composite
1336 optima. Cladistics 15, 415–428.
- 1337 Goloboff, P.A., 2003. Improvements to resampling measures of group support. Cladistics 19,
1338 324–332.
- 1339 Goloboff, P.A., 2014. Extended implied weighting. Cladistics 30, 260–272.
- 1340 Goloboff, P.A., Carpenter, J.M., Arias, J.S., Esquivel, D.R.M., 2008a. Weighting against
1341 homoplasy improves phylogenetic analysis of morphological data sets. Cladistics 24,
1342 758–773.
- 1343 Goloboff, P.A., Catalano, S.A., 2016. TNT version 1.5, including a full implementation of

- Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van Steeberge M, Koblmüller S, Sturmbauer C, Huyse T, Smeets K, Artois T, Vanhove MPM
1344 phylogenetic morphometrics. *Cladistics* 32, 221–238.
- 1345 Goloboff, P.A., Farris, J.S., Nixon, K.C., 2008b. TNT, a free program for phylogenetic
1346 analysis. *Cladistics* 24, 774–786.
- 1347 Goloboff, P.A., Mattoni, C.I., Quinteros, A.S., 2006. Continuous characters analyzed as such.
1348 *Cladistics* 22, 589–601.
- 1349 Goolsby, E.W., Bruggeman, J., Ané, C., 2017. Rphylopars: fast multivariate phylogenetic
1350 comparative methods for missing data and within-species variation. *Methods Ecol. Evol.*
1351 8, 22–27.
- 1352 Goren, M., 2014. *Tristramella sacra*. IUCN Red List Threat. Species 2014
1353 e.T61372A19010617.
- 1354 Grant, P.R., 1999. Ecology and evolution of Darwin’s finches, 2nd ed. Princeton Science
1355 Library, Princeton University Press, Princeton, USA.
- 1356 Gu, Z., Eils, R., Schlesner, M., 2016. Complex heatmaps reveal patterns and correlations in
1357 multidimensional genomic data. *Bioinformatics* 32, 2847–2849.
- 1358 Guerrero, A.M., Tye, A., 2009. Darwin’s finches as seed predators and dispersers. *Wilson J.*
1359 *Ornithol.* 121, 752–764.
- 1360 Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010.
1361 New algorithms and methods to estimate maximum-likelihood phylogenies: assessing
1362 the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321.
- 1363 Hansen, T.F., Martins, E.P., 1996. Translating between microevolutionary process and
1364 macroevolutionary patterns: the correlation structure of interspecific data. *Evolution* 50,
1365 1404–1417.

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

- 1366 Harmon, L.J., Losos, J.B., Jonathan Davies, T., Gillespie, R.G., Gittleman, J.L., Bryan
1367 Jennings, W., Kozak, K.H., McPeck, M.A., Moreno-Roark, F., Near, T.J., Purvis, A.,
1368 Ricklefs, R.E., Schluter, D., Schulte Ii, J.A., Seehausen, O., Sidlauskas, B.L., Torres-
1369 Carvajal, O., Weir, J.T., Mooers, A.Ø., 2010. Early bursts of body size and shape
1370 evolution are rare in comparative data. *Evolution* 64, 2385–2396.
- 1371 Hassouna, N., Michot, B., Bachellerie, J.-P., 1984. The complete nucleotide sequences of
1372 mouse 28S rRNA gene. Implications for the process of size increase of the large subunit
1373 rRNA in higher eukaryotes. *Nucleic Acids Res.* 12, 3563–3583.
- 1374 Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., Le Vinh, S., 2018. UFBoot2:
1375 Improving the Ultrafast Bootstrap Approximation. *Mol. Biol. Evol.* 35, 518–522.
- 1376 Hsu, C., Chang, C., Lin, C., 2003. A practical guide to support vector machines. Technical
1377 report. Department of Computer Science and Information Engineering, National Taiwan
1378 University, Taipei, Taiwan. Available at:
1379 <https://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf> (accessed Mar 6, 2021).
- 1380 Huyse, T., Poulin, R., Théron, A., 2005. Speciation in parasites: a population genetics
1381 approach. *Trends Parasitol.* 21, 469–475.
- 1382 Igeh, P.C., Avenant-Oldewage, A., 2020. Pathological effects of *Cichlidogyrus philander*
1383 Douëllou, 1993 (Monogenea, Ancyrocephalidae) on the gills of *Pseudocrenilabrus*
1384 *philander* (Weber, 1897) (Cichlidae). *J. Fish Dis.* 43, 177–184.
- 1385 Igeh, P.C., Dos Santos, Q.M., Avenant-Oldewage, A., 2017. Redescription of *Cichlidogyrus*
1386 *philander* (Monogenea, Ancyrocephalidae) using scanning electron microscopy (SEM)
1387 and molecular analysis. *Parasite* 24, 49.
- 1388 Ingram, T., Mahler, D.L., 2013. SURFACE: Detecting convergent evolution from

- Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van Steeberge M, Koblmüller S, Sturmbauer C, Huyse T, Smeets K, Artois T, Vanhove MPM
- 1389 comparative data by fitting Ornstein-Uhlenbeck models with stepwise Akaike
1390 Information Criterion. *Methods Ecol. Evol.* 4, 416–425.
- 1391 Ishiguro, M., Sakamoto, Y., Kitagawa, G., 1997. Bootstrapping log likelihood and EIC, an
1392 extension of AIC. *Ann. Inst. Stat. Math.* 49, 411–434.
- 1393 Jiménez-García, M.I., Vidal-Martínez, V.M., Lopez-Jiménez, S., Jiménez-García, M.I.,
1394 Vidal-Martínez, V.M., López-Jiménez, S., 2001. Monogeneans in introduced and native
1395 cichlids in México: Evidence for transfer. *J. Parasitol.* 87, 907.
- 1396 Jorissen, M.W.P., Huyse, T., Pariselle, A., Wamuini Lunkayilakio, S., Muterezi Bukinga, F.,
1397 Chocha Manda, A., Kapepula Kasembele, G., Vreven, E.J., Snoeks, J., Decru, E., Artois,
1398 T., Vanhove, M.P.M., 2020. Historical museum collections help detect parasite species
1399 jumps after tilapia introductions in the Congo Basin. *Biol. Invasions* 11, 1123.
- 1400 Jorissen, M.W.P., Pariselle, A., Huyse, T., Vreven, E.J., Snoeks, J., Decru, E., Kusters, T.,
1401 Wamuini Lunkayilakio, S., Muterezi Bukinga, F., Artois, T., Vanhove, M.P.M., 2018a.
1402 Six new dactylogyrid species (Platyhelminthes, Monogenea) from the gills of cichlids
1403 (Teleostei, Cichliformes) from the Lower Congo Basin. *Parasite* 25, 64.
- 1404 Jorissen, M.W.P., Pariselle, A., Huyse, T., Vreven, E.J., Snoeks, J., Volckaert, F.A.M.,
1405 Chocha Manda, A., Kapepula Kasembele, G., Artois, T., Vanhove, M.P.M., 2018b.
1406 Diversity and host specificity of monogenean gill parasites (Platyhelminthes) of cichlid
1407 fishes in the Bangweulu-Mweru ecoregion. *J. Helminthol.* 92, 417–437.
- 1408 Jorissen, M.W.P., Vanhove, M.P.M., Pariselle, A., Snoeks, J., Vreven, E.J., Šimková, A.,
1409 Wamuini Lunkayilakio, S., Chocha Manda, A., Kapepula Kasembele, G., Muterezi
1410 Bukinga, F., Artois, T., Huyse, T., 2021. Molecular footprint of parasite co-introduction
1411 with Nile tilapia in the Congo Basin [preprint]. *Research Square* rs.3.rs-995291/v1.

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

- 1412 Kaltz, O., Shykoff, J.A., 1998. Local adaptation in host–parasite systems. *Heredity* 81, 361–
1413 370.
- 1414 Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermini, L.S., 2017.
1415 ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14,
1416 587–589.
- 1417 Karvonen, A., Seehausen, O., 2012. The role of parasitism in adaptive radiations—When
1418 might parasites promote and when might they constrain ecological speciation? *Int. J.*
1419 *Ecol.* 2012.
- 1420 Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7:
1421 improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780.
- 1422 Kear, G., Whittington, I., 2015. Sperm transfer in monogenean (platyhelminth) parasites.
1423 *Acta Parasitol.* 60, 567–600.
- 1424 Khabbazian, M., Kriebel, R., Rohe, K., Ané, C., 2016. Fast and accurate detection of
1425 evolutionary shifts in Ornstein–Uhlenbeck models. *Methods Ecol. Evol.* 7, 811–824.
- 1426 Kitagawa, G., Konishi, S., 2010. Bias and variance reduction techniques for bootstrap
1427 information criteria. *Ann. Inst. Stat. Math.* 62, 209–234.
- 1428 Klingenberg, C.P., 2010. Evolution and development of shape: integrating quantitative
1429 approaches. *Nat. Rev. Genet.* 11, 623–635.
- 1430 Kmentová, N., Cruz-Laufer, A.J., Pariselle, A., Smeets, K., Artois, T., Vanhove, M.P.M.,
1431 2021a. Dactylogyridae 2021: seeing the forest through the (phylogenetic) trees.
1432 *EcoEvoRxiv* 10.32942/osf.io/5mh3d.
- 1433 Kmentová, N., Gelnar, M., Koblmüller, S., Vanhove, M.P.M., 2016a. First insights into the
1434 diversity of gill monogeneans of “*Gnathochromis*” and *Limnochromis* (Teleostei,

- Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van Steenberge M, Koblmüller S, Sturmbauer C, Huyse T, Smeets K, Artois T, Vanhove MPM
- 1435 Cichlidae) in Burundi: Do the parasites mirror host ecology and phylogenetic history?
1436 PeerJ 2016, e1629.
- 1437 Kmentová, N., Gelnar, M., Koblmüller, S., Vanhove, M.P.M., 2016b. Deep-water parasite
1438 diversity in Lake Tanganyika: Description of two new monogenean species from
1439 benthopelagic cichlid fishes. *Parasit. Vectors* 9, 1–9.
- 1440 Kmentová, N., Gelnar, M., Mendlová, M., Van Steenberge, M., Koblmüller, S., Vanhove,
1441 M.P.M., 2016c. Reduced host-specificity in a parasite infecting non-littoral Lake
1442 Tanganyika cichlids evidenced by intraspecific morphological and genetic diversity. *Sci.*
1443 *Rep.* 6, 39605.
- 1444 Kmentová, N., Hahn, C., Koblmüller, S., Zimmermann, H., Vorel, J., Artois, T., Gelnar, M.,
1445 Vanhove, M.P.M., 2021b. Contrasting host-parasite population structure: Morphology
1446 and mitogenomics of a parasitic flatworm on pelagic deepwater cichlid fishes from Lake
1447 Tanganyika. *Biology (Basel)*. 10, 797.
- 1448 Kmentová, N., Koblmüller, S., Van Steenberge, M., Artois, T., Muterezi Bukinga, F.,
1449 Mulimbwa N'sibula, T., Muzumani Risasi, D., Masilya Mulungula, P., Gelnar, M.,
1450 Vanhove, M.P.M., 2020. Failure to diverge in African Great Lakes: the case of
1451 *Dolicirroplectanum lacustre* gen. nov. comb. nov. (Monogenea, Diplectanidae) infecting
1452 latid hosts. *J. Great Lakes Res.* 46, 1113–1130.
- 1453 Kmentová, N., Van Steenberge, M., Raeymaekers, J.A.M., Koblmüller, S., Hablützel, P.I.,
1454 Bukinga, F.M., N'sibula, T.M., Mulungula, P.M., Nzigidahera, B., Ntakimazi, G.,
1455 Gelnar, M., Vanhove, M.P.M., 2018. Monogenean parasites of sardines in Lake
1456 Tanganyika: diversity, origin and intraspecific variability. *Contrib. Zool.* 87, 105–132.
- 1457 Kuhn, M., 2008. Building predictive models in R using the caret package. *J. Stat. Softw.* 28,

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

- 1458 1–26.
- 1459 Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data.
1460 *Biometrics* 33, 159.
- 1461 Landis, M.J., Freyman, W.A., Baldwin, B.G., 2018. Retracing the Hawaiian silversword
1462 radiation despite phylogenetic, biogeographic, and paleogeographic uncertainty.
1463 *Evolution* 72, 2343–2359.
- 1464 Legendre, P., Lapointe, F.-J., 2004. Assessing congruence among distance matrices: single-
1465 malt scotch whiskies revisited. *Aust. N. Z. J. Stat.* 46, 615–629.
- 1466 Lévêque, C., Oberdorff, T., Paugy, D., Stiassny, M.L.J., Tedesco, P.A., 2008. Global
1467 diversity of fish (Pisces) in freshwater. *Hydrobiologia* 595, 545–567.
- 1468 Littlewood, D.T.J., Rohde, K., Clough, K.A., 1997. Parasite speciation within or between
1469 host species? - Phylogenetic evidence from site-specific polystome monogeneans. *Int. J.*
1470 *Parasitol.* 27, 1289–1297.
- 1471 Llewellyn, J., 1963. Larvae and larval development of monogeneans. In: Dawes, B. (Ed.),
1472 *Advances in parasitology*. Academic Press, London and New York, pp. 287–326.
- 1473 Lockyer, A.E., Olson, P.D., Østergaard, P., Rollinson, D., Johnston, D.A., Attwood, S.W.,
1474 Southgate, V.R., Horak, P., Snyder, S.D., Le, T.H., Agatsuma, T., McManus, D.P.,
1475 Carmichael, A.C., Naem, S., Littlewood, D.T.J., 2003. The phylogeny of the
1476 Schistosomatidae based on three genes with emphasis on the interrelationships of
1477 *Schistosoma* Weinland, 1858. *Parasitology* 126, 203–224.
- 1478 Losos, J.B., 2008. Phylogenetic niche conservatism, phylogenetic signal and the relationship
1479 between phylogenetic relatedness and ecological similarity among species. *Ecol. Lett.*
1480 11, 995–1003.

- Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van Steeberge M, Koblmüller S, Sturmhuber C, Huyse T, Smeets K, Artois T, Vanhove MPM
- 1481 Losos, J.B., 2010. Adaptive radiation, ecological opportunity, and evolutionary determinism:
1482 American society of naturalists E. O. Wilson Award address. *Am. Nat.* 175, 623–639.
- 1483 Louca, S., Doebeli, M., 2018. Efficient comparative phylogenetics on large trees.
1484 *Bioinformatics* 34, 1053–1055.
- 1485 Mahler, D.L., Ingram, T., Revell, L.J., Losos, J.B., 2013. Exceptional convergence on the
1486 macroevolutionary landscape in island lizard radiations. *Science* 341, 292–295.
- 1487 Marchiori, N.C., Pariselle, A., Pereira, J., Agnèse, J.F., Durand, J.D., Vanhove, M.P.M.,
1488 2015. A comparative study of *Ligophorus uruguayense* and *L. saladensis* (Monogenea:
1489 Ancyrocephalidae) from *Mugil liza* (Teleostei: Mugilidae) in southern Brazil. *Folia*
1490 *Parasitol.* 62, 024.
- 1491 Matejusová, I., Gelnar, M., McBeath, A.J.A., Collins, C.M., Cunningham, C.O., 2001.
1492 Molecular markers for gyrodactylids (Gyrodactylidae: Monogenea) from five fish
1493 families (Teleostei). *Int. J. Parasitol.* 31, 738–745.
- 1494 McCoy, K.D., 2003. Sympatric speciation in parasites – what is sympatry? *Trends Parasitol.*
1495 19, 400–404.
- 1496 McGee, M.D., Neches, R.Y., Seehausen, O., 2016. Evaluating genomic divergence and
1497 parallelism in replicate ecomorphs from young and old cichlid adaptive radiations. *Mol.*
1498 *Ecol.* 25, 260–268.
- 1499 Meinilä, M., Kuusela, J., Zitiara, M.S., Lumme, J., 2004. Initial steps of speciation by
1500 geographic isolation and host switch in salmonid pathogen *Gyrodactylus salaris*
1501 (Monogenea: Gyrodactylidae). *Int. J. Parasitol.* 34, 515–526.
- 1502 Mendlová, M., Desdevises, Y., Civiňová, K., Pariselle, A., Šimková, A., 2012. Monogeneans

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

- 1503 of West African cichlid fish: evolution and cophylogenetic interactions. *PLoS One* 7.
- 1504 Mendlová, M., Šimková, A., 2014. Evolution of host specificity in monogeneans parasitizing
1505 African cichlid fish. *Parasit. Vectors* 7, 69.
- 1506 Messu Mandeng, F.D., Bilong Bilong, C.F., Pariselle, A., Vanhove, M.P.M., Bitja Nyom,
1507 A.R., Agnèse, J.F., 2015. A phylogeny of *Cichlidogyrus* spp. (Monogenea,
1508 Dactylogyridea) clarifies a host-switch between fish families and reveals an adaptive
1509 component to attachment organ morphology of this parasite genus. *Parasit. Vectors* 8,
1510 582.
- 1511 Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., 2020. e1071: Misc
1512 Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071),
1513 TU Wien: R package version 1.7-4. Vienna, Austria. Available at: [https://cran.r-](https://cran.r-project.org/web/packages/e1071/e1071.pdf)
1514 [project.org/web/packages/e1071/e1071.pdf](https://cran.r-project.org/web/packages/e1071/e1071.pdf) (accessed Oct 27, 2020).
- 1515 Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for
1516 inference of large phylogenetic trees. 2010 *Gatew. Comput. Environ. Work. GCE* 2010.
- 1517 Mirande, J.M., 2009. Weighted parsimony phylogeny of the family Characidae (Teleostei:
1518 Characiformes). *Cladistics* 25, 574–613.
- 1519 Mirande, J.M., 2019. Morphology, molecules and the phylogeny of Characidae (Teleostei,
1520 Characiformes). *Cladistics* 35, 282–300.
- 1521 Murtagh, F., Legendre, P., 2014. Ward's hierarchical agglomerative clustering method:
1522 Which algorithms implement Ward's criterion? *J. Classif.* 31, 274–295.
- 1523 Muterezi Bukinga, F., Vanhove, M.P.M., Van Steenberge, M., Pariselle, A., 2012.
1524 Ancyrocephalidae (Monogenea) of Lake Tanganyika: III: *Cichlidogyrus* infecting the
1525 world's biggest cichlid and the non-endemic tribes Haplochromini, Oreochromini and

- Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van Steeberge M, Koblmüller S, Sturmbauer C, Huyse T, Smeets K, Artois T, Vanhove MPM
1526 Tylochromini (Teleostei, Cichlidae). Parasitol. Res. 111, 2049–2061.
- 1527 Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and
1528 effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol.
1529 Biol. Evol. 32, 268–274.
- 1530 Nixon, K.C., 1999. The parsimony ratchet, a new method for rapid parsimony analysis.
1531 Cladistics 15, 407–414.
- 1532 O’Meara, B.C., Ané, C., Sanderson, M.J., Wainwright, P.C., 2006. Testing for different rates
1533 of continuous trait evolution using likelihood. Evolution 60, 922.
- 1534 Pagel, M., 1999. Inferring the historical patterns of biological evolution. Nature 401, 877–
1535 884.
- 1536 Paperna, I., 1960. Studies on monogenetic trematodes in Israel. 2. Monogenetic trematodes of
1537 cichlids. Bamidgeh 12, 20–33.
- 1538 Paperna, I., 1968. Monogenetic trematodes collected from fresh water fish in Ghana. Second
1539 report. Bamidgeh 17, 107–111.
- 1540 Paperna, I., 1979. Monogenea of inland water fish in Africa. Ann. Mus. R. Afr. Cent.
1541 Tervuren Belg. Ser. IN-8 Sci. Zool. 226, 1–131.
- 1542 Paperna, I., Thurston, J.P., 1969. Monogenetic trematodes collected from cichlid fish in
1543 Uganda; including the description of five new species of *Cichlidogyrus*. Rev. Zool. Bot.
1544 Afr. 79, 15–33.
- 1545 Parins-Fukuchi, C., 2018. Use of continuous traits can improve morphological phylogenetics.
1546 Syst. Biol. 67, 328–339.

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

- 1547 Pariselle, A., Bitja Nyom, A.R., Bilong Bilong, C.F., 2014. Four new species of
1548 *Cichlidogyrus* (Monogenea, Ancyrocephalidae) from *Sarotherodon mvogoi* and
1549 *Tylochromis sudanensis* (Teleostei, Cichlidae) in Cameroon. *Zootaxa* 3881, 258–266.
- 1550 Pariselle, A., Bitja Nyom, A.R., Bilong Bilong, C.F., Bilong, C.F., 2013. Checklist of the
1551 ancyrocephalids (Monogenea) parasitizing *Tilapia* species in Cameroon, with the
1552 description of three new species. *Zootaxa* 3599, 78–86.
- 1553 Pariselle, A., Boeger, W.A., Snoeks, J., Bilong Bilong, C.F., Morand, S., Vanhove, M.P.M.,
1554 2011. The monogenean parasite fauna of cichlids: a potential tool for host biogeography.
1555 *Int. J. Evol. Biol.* 2011, 471480.
- 1556 Pariselle, A., Euzet, L., 1994. Three new species of *Cichlidogyrus* Paperna, 1960
1557 (Monogenea, Ancyrocephalidae) parasitic on *Tylochromis jentinki* (Steindachner, 1895)
1558 (Pisces, Cichlidae) in West Africa. *Syst. Parasitol.* 29, 229–234.
- 1559 Pariselle, A., Euzet, L., 1995. *Scutogyrus* gen.n. (Monogenea: Ancyrocephalidae) for
1560 *Cichlidogyrus longicornis minus* Dossou, 1982, *C. l. longicornis*, and *C. l. gravivaginus*
1561 Paperna and Thurston, 1969, with description of three new species parasitic on African .
1562 *J. Helminthol. Soc. Washingt.* 62, 157–173.
- 1563 Pariselle, A., Euzet, L., 1996. *Cichlidogyrus* Paperna, 1960 (Monogenea, Ancyrocephalidae):
1564 gill parasites from West African Cichlidae of the subgenus *Coptodon* Regan, 1920
1565 (pisces), with descriptions of six new species. *Syst. Parasitol.* 34, 109–124.
- 1566 Pariselle, A., Euzet, L., 2003. Four new species of *Cichlidogyrus* (Monogenea:
1567 Ancyrocephalidae), gill parasites of *Tilapia cabrae* (Teleostei: Cichlidae), with
1568 discussion on relative length of haptor sclerites. *Folia Parasitol.* 50, 195–201.
- 1569 Pariselle, A., Euzet, L., 2004. Two new species of *Cichlidogyrus* Paperna, 1960 (Monogenea,

- Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van Steenberge M, Koblmüller S, Sturmbauer C, Huyse T, Smeets K, Artois T, Vanhove MPM
- 1570 Ancyrocephalidae) gill parasites on *Hemichromis fasciatus* (Pisces, Cichlidae) in Africa,
1571 with remarks on parasite geographical distribution. Parasite 11, 359–364.
- 1572 Pariselle, A., Euzet, L., 2009. Systematic revision of dactylogyridean parasites (Monogenea)
1573 from cichlid fishes in Africa, the Levant and Madagascar. Zoosystema 31, 849–898.
- 1574 Pariselle, A., Morand, S., Deveney, M.R., Pouyaud, L., 2003. Parasite species richness of
1575 closely related hosts: historical scenario and “genetic” hypothesis. In: Combes, C.,
1576 Jourdane, J., Ducreux-Modat, A., Pages, J.-R. (Eds.), Taxonomie, écologie et évolution
1577 des métazoaires parasites: Livre-hommage à Louis Euzet. Presses Universitaires de
1578 Perpignan, Perpignan, France, pp. 147–166.
- 1579 Pariselle, A., Muterezi Bukinga, F., Van Steenberge, M., Vanhove, M.P.M., 2015a.
1580 Ancyrocephalidae (Monogenea) of Lake Tanganyika: IV: *Cichlidogyrus* parasitizing
1581 species of Bathybatini (Teleostei, Cichlidae): reduced host-specificity in the deepwater
1582 realm? Hydrobiologia 748, 99–119.
- 1583 Pariselle, A., Van Steenberge, M., Snoeks, J., Volckaert, F.A.M., Huyse, T., Vanhove,
1584 M.P.M., 2015b. Ancyrocephalidae (Monogenea) of Lake Tanganyika: Does the
1585 *Cichlidogyrus* parasite fauna of *Interochromis loocki* (Teleostei, Cichlidae) reflect its
1586 host’s phylogenetic affinities? Contrib. Zool. 84, 25–38.
- 1587 Pennell, M.W., Eastman, J.M., Slater, G.J., Brown, J.W., Uyeda, J.C., Fitzjohn, R.G., Alfaro,
1588 M.E., Harmon, L.J., 2014. Geiger v2.0: an expanded suite of methods for fitting
1589 macroevolutionary models to phylogenetic trees. Bioinformatics 30, 2216–2218.
- 1590 Poulin, R., 2014. Parasite biodiversity revisited: frontiers and constraints. Int. J. Parasitol. 44,
1591 581–589.

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

- 1592 Poulin, R., Hay, E., Jorge, F., 2019. Taxonomic and geographic bias in the genetic study of
1593 helminth parasites. *Int. J. Parasitol.* 49, 429–435.
- 1594 Poulin, R., Krasnov, B.R., Mouillot, D., 2011. Host specificity in phylogenetic and
1595 geographic space. *Trends Parasitol.* 27, 355–361.
- 1596 Poulin, R., Presswell, B., Jorge, F., 2020. The state of fish parasite discovery and taxonomy:
1597 a critical assessment and a look forward. *Int. J. Parasitol.* 50, 733–742.
- 1598 Pouyaud, L., Desmarais, E., Deveney, M., Pariselle, A., 2006. Phylogenetic relationships
1599 among monogenean gill parasites (Dactylogyridea, Ancyrocephalidae) infesting tilapiine
1600 hosts (Cichlidae): systematic and evolutionary implications. *Mol. Phylogenet. Evol.* 38,
1601 241–249.
- 1602 Price, C.E., Kirk, R.G., 1967. First description of a monogenetic trematode from Malawi.
1603 *Rev. Zool. Bot. Afr.* 76, 137–144.
- 1604 Price, C.E., Peebles, H.E., Bamford, T., 1969. The monogenean parasites of African fishes.
1605 IV. Two new species from South African hosts. *Rev. Zool. Bot. Afr.* 79, 117–124.
- 1606 Puttick, M.N., 2018. Mixed evidence for early bursts of morphological evolution in extant
1607 clades. *J. Evol. Biol.* 31, 502–515.
- 1608 R Core Team, 2021. R: a language and environment for statistical computing. R Foundation
1609 for Statistical Computing, Vienna, Austria. Available at: <https://www.r-project.org/>
1610 (accessed Jul 5, 2021).
- 1611 Rahmouni, C., Van Steenberge, M., Vanhove, M.P.M., Šimková, A., 2021. Intraspecific
1612 morphological variation in *Cichlidogyrus* (Monogenea) parasitizing two cichlid hosts
1613 from Lake Tanganyika exhibiting different dispersal capacities. *Hydrobiologia* 848,
1614 3833–3845.

- Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van Steeberge M, Koblmüller S, Sturmhuber C, Huyse T, Smeets K, Artois T, Vanhove MPM
- 1615 Rahmouni, C., Vanhove, M.P.M., Šimková, A., 2017. Underexplored diversity of gill
1616 monogeneans in cichlids from Lake Tanganyika: eight new species of *Cichlidogyrus*
1617 Paperna, 1960 (Monogenea: Dactylogyridae) from the northern basin of the lake, with
1618 remarks on the vagina and the heel of the male copulatory . *Parasit. Vectors* 10, 591.
- 1619 Rahmouni, C., Vanhove, M.P.M., Šimková, A., 2018. Seven new species of *Cichlidogyrus*
1620 Paperna, 1960 (Monogenea: Dactylogyridae) parasitizing the gills of Congolese cichlids
1621 from northern Lake Tanganyika. *PeerJ* 6, e5604.
- 1622 Ramasamy, P., Ramalingam, K., Hanna, R.E.B., Halton, D.W., 1985. Microhabitats of gill
1623 parasites (Monogenea and copepoda) of teleosts (*Scomberoides* spp.). *Int. J. Parasitol.*
1624 15, 385–397.
- 1625 Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior
1626 summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* 67, 901–904.
- 1627 Řehulková, E., Kičinjaová, M.L., Mahmoud, Z.N., Gelnar, M., Seifertová, M., 2019. Species
1628 of *Characidotrema* Paperna & Thurston, 1968 (Monogenea: Dactylogyridae) from
1629 fishes of the Alestidae (Characiformes) in Africa: new species, host-parasite associations
1630 and first insights into the phylogeny of the genus. *Parasit. Vectors* 12, 366.
- 1631 Řehulková, E., Mendlová, M., Šimková, A., 2013. Two new species of *Cichlidogyrus*
1632 (Monogenea: Dactylogyridae) parasitizing the gills of African cichlid fishes
1633 (Perciformes) from Senegal: morphometric and molecular characterization. *Parasitol.*
1634 *Res.* 112, 1399–1410.
- 1635 Revell, L.J., 2012. phytools: an R package for phylogenetic comparative biology (and other
1636 things). *Methods Ecol. Evol.* 3, 217–223.

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

- 1637 Rodríguez-González, A., Sarabeev, V., Balbuena, J.A., 2017. Evolutionary morphology in
1638 shape and size of haptor anchors in 14 *Ligophorus* spp. (Monogenea: Dactylogyridae).
1639 PLoS One 12, e0178367.
- 1640 Rohde, K., 1976. Monogenean gill parasites of *Scomberomorus commersoni* Lacépède and
1641 other mackerel on the Australian East Coast. *Zeitschrift für Parasitenkd.* 51, 49–69.
- 1642 Ronco, F., Büscher, H.H., Indermaur, A., Salzburger, W., 2020. The taxonomic diversity of
1643 the cichlid fish fauna of ancient Lake Tanganyika, East Africa. *J. Great Lakes Res.* 46,
1644 1067–1078.
- 1645 Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under
1646 mixed models. *Bioinformatics* 19, 1572–1574.
- 1647 Salzburger, W., 2018. Understanding explosive diversification through cichlid fish genomics.
1648 *Nat. Rev. Genet.* 19, 705–717.
- 1649 Salzburger, W., van Bocxlaer, B., Cohen, A.S., 2014. Ecology and evolution of the African
1650 Great Lakes and their faunas. *Annu. Rev. Ecol. Evol. Syst.* 45, 519–545.
- 1651 Schedel, F.D.B., Musilova, Z., Schlieven, U.K., 2019. East African cichlid lineages
1652 (Teleostei: Cichlidae) might be older than their ancient host lakes: new divergence
1653 estimates for the East African cichlid radiation. *BMC Evol. Biol.* 19, 1–25.
- 1654 Seehausen, O., 2006. African cichlid fish: a model system in adaptive radiation research.
1655 *Proc. R. Soc. B Biol. Sci.* 273, 1987–1998.
- 1656 Šimková, A., Ondračková, M., Gelnar, M., Morand, S., 2002. Morphology and coexistence of
1657 congeneric ectoparasite species: reinforcement of reproductive isolation? *Biol. J. Linn.*
1658 *Soc.* 76, 125–135.
- 1659 Šimková, A., Plaisance, L., Matějusková, I., Morand, S., Verneau, O., 2003. Phylogenetic

- Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van Steenberge M, Koblmüller S, Sturmbauer C, Huyse T, Smeets K, Artois T, Vanhove MPM
- 1660 relationships of the Dactylogyridae Bychowsky, 1933 (Monogenea: Dactylogyridea):
1661 The need for the systematic revision of the Ancyrocephalinae Bychowsky, 1937. Syst.
1662 Parasitol. 54, 1–11.
- 1663 Šimková, A., Řehulková, E., Rasoloariniaina, J.R., Jorissen, M.W.P., Scholz, T., Faltýnková,
1664 A., Mašová, Š., Vanhove, M.P.M., 2019. Transmission of parasites from introduced
1665 tilapias: a new threat to endemic Malagasy ichthyofauna. Biol. Invasions 21, 803–819.
- 1666 Šimková, A., Serbielle, C., Pariselle, A., Vanhove, M.P.M., Morand, S., 2013. Speciation in
1667 *Thaparocleidus* (Monogenea: Dactylogyridae) parasitizing Asian pangasiid catfishes.
1668 Biomed Res. Int. 2013, 353956.
- 1669 Takahashi, T., Koblmüller, S., 2011. The adaptive radiation of cichlid fish in Lake
1670 Tanganyika: a morphological perspective. Int. J. Evol. Biol. 2011, 1–14.
- 1671 Talavera, G., Castresana, J., 2007. Improvement of phylogenies after removing divergent and
1672 ambiguously aligned blocks from protein sequence alignments. Syst. Biol. 56, 564–577.
- 1673 Thurston, J.P., 1970. The incidence of Monogenea and parasitic Crustacea on the gills of fish
1674 in Uganda. Rev. Zool. Bot. Afr. 82, 111–130.
- 1675 Turner, G.F., Seehausen, O., Knight, M.E., Allender, C.J., Robinson, R.L., 2001. How many
1676 species of cichlid fishes are there in African lakes? Mol. Ecol. 10, 793–806.
- 1677 Uyeda, J.C., Harmon, L.J., 2014. A novel Bayesian method for inferring and interpreting the
1678 dynamics of adaptive landscapes from phylogenetic comparative data. Syst. Biol. 63,
1679 902–918.
- 1680 Van Steenberge, M., Pariselle, A., Huyse, T., Volckaert, F.A.M., Snoeks, J., Vanhove,
1681 M.P.M., 2015. Morphology, molecules, and monogenean parasites: an example of an

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

- 1682 integrative approach to cichlid biodiversity. PLoS One 10, e0124474.
- 1683 Van Steenberge, M., Vanhove, M.P.M., Muzumani Risasi, D., Mulimbwa N'Sibula, T.,
1684 Muterezi Bukinga, F., Pariselle, A., Gillardin, C., Vreven, E.J., Raeymaekers, J.A.M.,
1685 Huyse, T., Volckaert, F.A.M., Nshombo Muderhwa, V., Snoeks, J., 2011. A recent
1686 inventory of the fishes of the north-western and central western coast of Lake
1687 Tanganyika (Democratic Republic Congo). Acta Ichthyol. Piscat. 41, 201–214.
- 1688 Vanhove, M.P.M., Briscoe, A.G., Jorissen, M.W.P., Littlewood, D.T.J., Huyse, T., 2018. The
1689 first next-generation sequencing approach to the mitochondrial phylogeny of African
1690 monogenean parasites (Platyhelminthes: Gyrodactylidae and Dactylogyridae). BMC
1691 Genomics 19, 520.
- 1692 Vanhove, M.P.M., Hablützel, P.I., Pariselle, A., Šimková, A., Huyse, T., Raeymaekers,
1693 J.A.M., 2016. Cichlids: a host of opportunities for evolutionary parasitology. Trends
1694 Parasitol. 32, 820–832.
- 1695 Vanhove, M.P.M., Pariselle, A., Van Steenberge, M., Raeymaekers, J.A.M., Hablützel, P.I.,
1696 Gillardin, C., Hellemans, B., Breman, F.C., Koblmüller, S., Sturmbauer, C., Snoeks, J.,
1697 Volckaert, F.A.M., Huyse, T., 2015. Hidden biodiversity in an ancient lake:
1698 phylogenetic congruence between Lake Tanganyika trophic cichlids and their
1699 monogenean flatworm parasites. Sci. Rep. 5, 13669.
- 1700 Vanhove, M.P.M., Van Steenberge, M., Dessein, S., Volckaert, F.A.M., Snoeks, J., Huyse,
1701 T., Pariselle, A., 2013. Biogeographical implications of Zambezian *Cichlidogyrus*
1702 species (Platyhelminthes: Monogenea: Ancyrocephalidae) parasitizing Congolian
1703 cichlids. Zootaxa 3608, 398–400.
- 1704 Vanhove, M.P.M., Volckaert, F.A.M., Pariselle, A., 2011. Ancyrocephalidae (Monogenea) of

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van

Steeberge M, Koblmüller S, Sturmhuber C, Huyse T, Smeets K, Artois T, Vanhove MPM

1705 Lake Tanganyika: I: four new species of *Cichlidogyrus* from *Ophthalmotilapia ventralis*

1706 (Teleostei: Cichlidae), the first record of this parasite family in the basin. *Zoologia* 28,

1707 253–263.

1708 Vignon, M., Pariselle, A., Vanhove, M.P.M., 2011. Modularity in attachment organs of

1709 African *Cichlidogyrus* (Platyhelminthes: Monogenea: Ancyrocephalidae) reflects

1710 phylogeny rather than host specificity or geographic distribution. *Biol. J. Linn. Soc.* 102,

1711 694–706.

1712 Vignon, M., Sasal, P., 2010. The use of geometric morphometrics in understanding shape

1713 variability of sclerotized haptor structures of monogeneans (Platyhelminthes) with

1714 insights into biogeographic variability. *Parasitol. Int.* 59, 183–191.

1715 Weiblen, G.D., Bush, G.L., 2002. Speciation in fig pollinators and parasites. *Mol. Ecol.* 11,

1716 1573–1578.

1717 Wickham, H., 2016. *ggplot2: Elegant graphics for data analysis*, 2nd ed. , Use R!Springer,

1718 New York, USA.

1719 Wu, X.Y., Zhu, X.Q., Xie, M.Q., Li, A.X., 2006. The radiation of *Haliotrema* (Monogenea:

1720 Dactylogyridae: Ancyrocephalinae): molecular evidence and explanation inferred from

1721 LSU rDNA sequences. *Parasitology* 132, 659–668.

1722 Wu, X.Y., Zhu, X.Q., Xie, M.Q., Li, A.X., 2007. The evaluation for generic-level monophyly

1723 of Ancyrocephalinae (Monogenea, Dactylogyridae) using ribosomal DNA sequence

1724 data. *Mol. Phylogenet. Evol.* 44, 530–544.

1725 Yu, G., Lam, T.T.-Y., Zhu, H., Guan, Y., 2018. Two methods for mapping and visualizing

1726 associated data on phylogeny using ggtree. *Mol. Biol. Evol.* 35, 3041–3043.

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

- 1727 Yu, G., Smith, D.K., Zhu, H., Guan, Y., Lam, T.T.-Y., 2017. ggtree[□]: An R package for
1728 visualization and annotation of phylogenetic trees with their covariates and other
1729 associated data. *Methods Ecol. Evol.* 8, 28–36.
- 1730 Zhi, T., Xu, X., Chen, J., Zheng, Y., Zhang, S., Peng, J., Brown, C.L., Yang, T., 2018.
1731 Expression of immune-related genes of Nile tilapia *Oreochromis niloticus* after
1732 *Gyrodactylus cichlidarum* and *Cichlidogyrus sclerosus* infections demonstrating
1733 immunosuppression in coinfection. *Fish Shellfish Immunol.* 80, 397–404.
- 1734 Zischke, M.T., Litherland, L., Tilyard, B.R., Stratford, N.J., Jones, E.L., Wang, Y.-G., 2016.
1735 Otolith morphology of four mackerel species (*Scomberomorus* spp.) in Australia:
1736 species differentiation and prediction for fisheries monitoring and assessment. *Fish. Res.*
1737 176, 39–47.

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van

Steeberge M, Koblmüller S, Sturmhuber C, Huyse T, Smeets K, Artois T, Vanhove MPM

1738 LIST OF FIGURES

1739 **Figure 1.** Overview of morphology of species of *Cichlidogyrus* Paperna, 1960

1740 (Platyhelminthes: Monogenea, Dactylogyridae). (a) Multiple specimens attached to the gills

1741 of *Sarotherodon melanotheron* Rüppel, 1852 (Cichliformes: Cichlidae). (b) Microscopic

1742 image of *Cichlidogyrus agnesi* Pariselle & Euzet, 1995 with sclerotised structures of

1743 reproductive (male copulatory organ and vagina) and attachment organs indicated by arrows.

1744 (c) Overview of hard part morphology and most widely applied measurements of these

1745 structures in taxonomic literature. Abbreviations: a, anchor total length; b, anchor blade

1746 length; c, anchor shaft length; d, anchor guard length; e, anchor point length; 1–7, hook

1747 lengths; w, bar width; x, bar length; h, auricle length; y, distance between auricles; AuP,

1748 surface of auxiliary plate of male copulatory organ; AP, accessory piece length; Pe, penis

1749 length; He, heel length; l, length; w, width. Terminology and methodology of measurements

1750 according to Fannes et al. (2017).

1751 **Figure 2.** Phylograms of three monogenean flatworms belonging to *Cichlidogyrus* and

1752 *Scutogyrus* (Platyhelminthes, Monogenea, Dactylogyridae) based on three nuclear (18S, 28S,

1753 and ITS rDNA) and one mitochondrial (CO1 mtDNA) DNA sequence markers. Sequences of

1754 specimens in bold have been generated for this study. (a) Final hypothesis of analysis under

1755 maximum parsimony and extended implied weighting ($k = 48$, weighting scheme BLK) with

1756 node support estimated through symmetric resampling ($p = 0.33$). (b) Bayesian phylogram

1757 with Bayesian posterior probabilities (PP) followed by ultrafast bootstrap values (UFBoot)

1758 and Shimodaira-Hasegawa-like approximate likelihood ratios (SH-aLRT) inferred from

1759 maximum likelihood estimation indicated at nodes; asterisk (*) indicates low or moderate

1760 support below the threshold ($PP < 0.95$, $UFBoot < 95$, $SH-aLRT < 80$); black dots, internal

1761 nodes with strong support across all support values. Node labels (1-11), monophyletic clades

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

1762 considered strongly supported: species infecting (mostly) hemichromine cichlids (*Hemi*),
1763 species belonging to *Scutogyrus* (*Scutogyrus*), species infecting (mostly) coptodonine cichlids
1764 among others (*Cop*), the first species group infecting oreochromine cichlids among others
1765 (*Oreo1*), species infecting cichlids belonging to the East African Radiation (*EAR*), species
1766 infecting coptodonine, pematolapiine oreochromine, tilapiine, heterotilapiine, and
1767 gobiocichline cichlids (*CPO*), species from Southern Africa with a bulbous penis (*Bulb*), the
1768 second species group infecting mainly oreochromine cichlids among others (*Oreo2*), *C. halli*-
1769 clade (*Halli*), and species infecting tylochromine cichlids (*Tylo*). Abbreviation: *C.*
1770 *schreyenbrichard.*, *C. schreyenbrichardorum*.

1771 **Figure 3.** Host repertoire of species of *Cichlidogyrus*. (a) Character map of index of
1772 specificity according to Mendlová and Šimková (2014). (b) Host range matrix of the species
1773 by tribe or subfamily of cichlid hosts and family of non-cichlid hosts (recent anthropogenic
1774 host range expansions italicised). APLO, Aplocheilidae; NOTH, Nothobranchiidae; POLY,
1775 Polycentridae; Pare, Paretroplinae; Ptyc, Ptychochrominae; Cich, Cichlasomatini; Chro,
1776 Chromidotilapiini; Tylo, Tylochromini; Hemi, Hemichromini; Gobi, Gobiocichlini; Copt,
1777 Coptodonini; Hete, Heterotilapiini; PelL, Pematolapiini; Oreo, Oreochromini; Tila, Tilapiini;
1778 Boul, Boulengerochromini; Bath, Bathybatini; Trem, Trematocarini; Bent, Benthochromini;
1779 Cypr, Cyprichromini; Ecto, Ectodini; Hapl, Haplochromini.

1780 **Figure 4.** Consensus tree of phylogenetic analysis under maximum parsimony and extended
1781 implied weighting (k = 48, BLK weighting scheme) combining molecular and morphometric
1782 data; red tip points indicate species with and blue without DNA sequence data; node support
1783 values constitute GC values inferred from symmetric resampling. Compared to the molecular
1784 tree (Fig. 2a), this tree is less resolved and less supported.

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van

Steeberge M, Koblmüller S, Sturmhuber C, Huyse T, Smeets K, Artois T, Vanhove MPM

1785 **Figure 5.** Character maps of morphological characters. (a) Scaled means of morphological
1786 continuous characters with abbreviations referring to measurements of sclerotised structures
1787 in the attachment and reproductive organ used to characterise species belonging to
1788 *Cichlidogyrus* (see Fig. 1). (b) Character maps of newly proposed discrete characters for the
1789 reproductive organs; indices in legend refer to character states suggested in Appendix 2. For
1790 details on all characters, see numbers in Appendix 2. Numbers on nodes refer to clades in Fig.
1791 2a.

1792 **Figure 6.** Morphology of attachment and reproductive organs of representative species
1793 belonging to proposed species groups and subgroups (Fig. 2) of *Cichlidogyrus* (incl.
1794 *Scutogyrus*) including cladogram and illustrations of sclerotised structures of attachment
1795 organ (anchors, bars, hooks) and the MCO of selected species (vagina morphology not shown
1796 as little unifying or contrasting morphological patterns were detected between groups, see
1797 “Characterisation of species groups”). Arrows indicate key features of each species group
1798 (blue) and subgroup of the *EAR* group (yellow), dashed lines indicate typical shapes. Species
1799 of the *EAR* group that are not displayed are labelled in grey. Scale = 30 μm .

1800 **Figure 7.** Systematic informativeness of morphometric measurements (a) Variable
1801 importance (AUC) according the optimised support vector machines (SVM) classifying
1802 specimens belonging to *Cichlidogyrus* into proposed species groups; colours indicate
1803 contribution by species group; AUC values > 0.8 are considered excellent discriminators;
1804 AUC values between 0.7 and 0.8 are considered acceptable discriminators; AUC values < 0.7
1805 are considered poor discriminators. (b) AUC values clustered by clade showing that the
1806 SVMs could distinguish species groups with different accuracy. (c). Cluster analysis of
1807 morphometric measurements of attachment and reproductive organs; clusters were detected
1808 using the *ward.d2* clustering algorithm. For abbreviation of measurements, see Fig. 1. Species

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

1809 groups were proposed based on phylogenetic analysis (Fig. 2).

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van

Steeberge M, Koblmüller S, Sturmhuber C, Huyse T, Smeets K, Artois T, Vanhove MPM

1810 **Figure 8.** Results of phylogenetic comparative analysis applied to one set of 100 randomly
1811 sampled tree topologies from the Bayesian post-burn in phase. Fitted continuous models
1812 include Brownian Motion (BM), Ornstein-Uhlenbeck (OU), Early-Burst (EB), Late-Burst
1813 (LB), and multi-rate Brownian Motion (BMM) models simulating a random walk (H_1),
1814 stabilising selection (H_2), adaptive radiation with a decelerating (H_3) and accelerating (H_4)
1815 divergence of characters, and divergent BM regimes in parts of the phylogeny (H_5). (a) Model
1816 fits were assessed through the Extended Information Criterion (EIC) for multivariate analysis.
1817 All models assuming a single evolutionary regime across the phylogeny performed similarly
1818 (LB performed worse). However, the BMM model suggesting a different regime for the EAR
1819 clade outperformed the latter. (b) Univariate PCMs for discrete characters including the hook
1820 configuration (HC) (Vignon et al. 2011), anchor similarity (AS), ventral bar shape (VBS)
1821 (Mendlová et al. 2012), and index of (host) specificity (IS) (Mendlová and Šimková 2014).
1822 HC and VBS show a strong phylogenetic signal, AS and IS shows no detectable phylogenetic
1823 signal. Model performance is assessed through difference in the sample size-corrected Akaike
1824 Information Criterion ($\Delta AICc$) compared to a white noise model under the assumption of a
1825 star-like phylogeny (absolute phylogenetic independence).

1826 **Figure 9.** Character maps of discrete characters for the attachment organ morphology as
1827 defined by Vignon et al. (2011) and Mendlová et al. (2012) vs. the respective set of
1828 continuous measurement suggested in this study. The character maps indicate that
1829 summarising multiple variables into a single discrete character might cause the information
1830 loss observed as a lack of a phylogenetic signal in Fig. 8b. Numbers on nodes refer to clades
1831 in Fig. 2a. Hook configuration A, B, and C assigned according to criteria by Vignon et al.
1832 (2011) with relative size of hook pair 1 vs pairs 3–7; haptor configuration D equates to the
1833 configuration of *C. arthracanthus* with large hooks 1 and 3–7. Anchor similarity either similar

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

1834 (1) or dissimilar (2). Ventral bar shape with (1) or without (2) membranous extensions,
1835 massive with extensions (3) or support large plate (4).LIST OF APPENDICES (IN PRINT)

1836 **Appendix 1.** Sampling localities and dates of specimens belonging to *Cichlidogyrus* and
1837 *Scutogyrus* collected for this study including a reference for the samples previously included
1838 in taxonomic studies. Freshwater ecoregions are assigned according to Thieme et al. (2005).

1839 **Appendix 2.** List of continuous and discrete characters and character states inferred from the
1840 taxonomic literature and used for character mapping in Fig. 1. All characters represent
1841 characteristics that are commonly used to described species of *Cichlidogyrus* and *Scutogyrus*.

1842 **Appendix 3.** Specimen data of cichlid parasites of the genera *Cichlidogyrus* and *Scutogyrus*
1843 used for phylogenetic analyses including host species, GenBank accession numbers, locality
1844 by country, and reference. Voucher/isolate ID and accession numbers in italics indicate
1845 specimens not included in subset trees used for phylogenetic comparative methods.

1846 **Appendix 4.** Substitution models of molecular evolution and partitions for Bayesian inference
1847 (BI) and maximum likelihood estimation (ML) of phylogeny of species of *Cichlidogyrus* and
1848 *Scutogyrus*. Models include the general time reversible model (GTR), the Kimura 1980 model
1849 (K80), the transitional model 3 with unequal base frequencies (TIM3e), the Tamura-Nei
1850 model (TN), and the three-parameter model 2 (TPM2) plus empirical base frequencies (+ F), a
1851 proportion of invariable sites (+ I), a discrete Γ model with four rate categories (Γ 4), or a
1852 FreeRate model with three categories (+ R3). For model specification see the IQ-TREE
1853 ModelFinder manual (Kalyaanamoorthy et al., 2017).

1854 **Appendix 5.** Species groups of *Cichlidogyrus* with species included, species potentially
1855 included, and the respective host ranges reported in the taxonomic literature including host
1856 species of candidate species.

Cruz-Laufer AJ, Pariselle A, Jorissen MWP, Muterezi Bukinga F, Al Assadi A, Van

Steeberge M, Koblmüller S, Sturmbauer C, Huyse T, Smeets K, Artois T, Vanhove MPM

1857 **Appendix 6.** Ancestral states of continuous characters inferred from a literature survey and

1858 used for character mapping. Character IDs refer to numbers in Appendix 2.

1859 **Appendix 7.** Ancestral states of discrete characters inferred from a literature survey and used

1860 for character mapping. Values represent probabilities for different character states. Character

1861 IDs refer to numbers in Appendix 2.

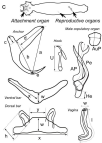
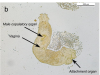
1862

EVOLUTION OF A SPECIES-RICH LINEAGE OF PARASITES

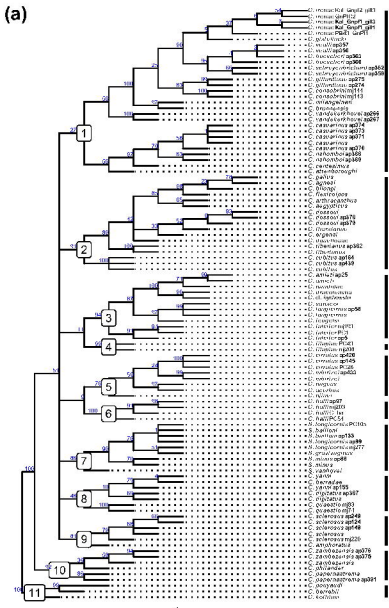
1863 SUPPORTING INFORMATION

1864 **File S1.** Input data matrix for use in *TNT* including morphometric and DNA sequence data.

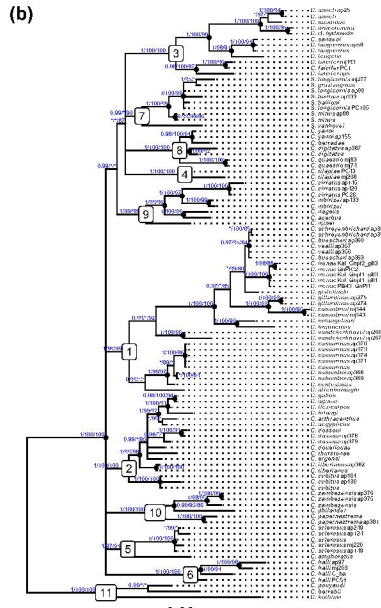
1865 **File S2.** Raw morphometric measurements for species of *Cichlidogyrus* and *Scutogyrus*.



(a)

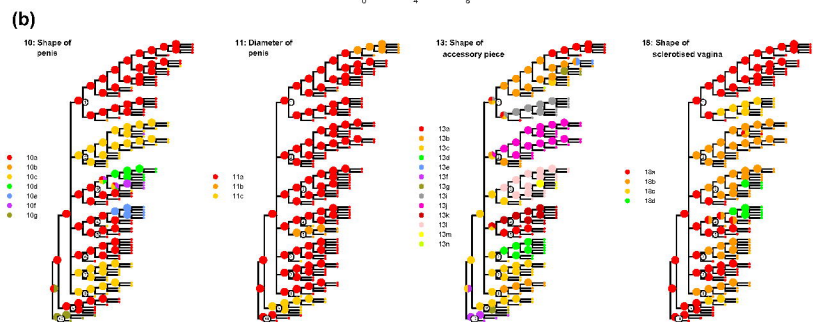
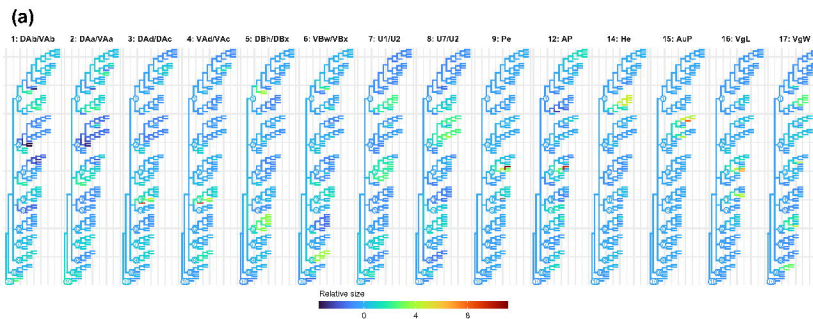


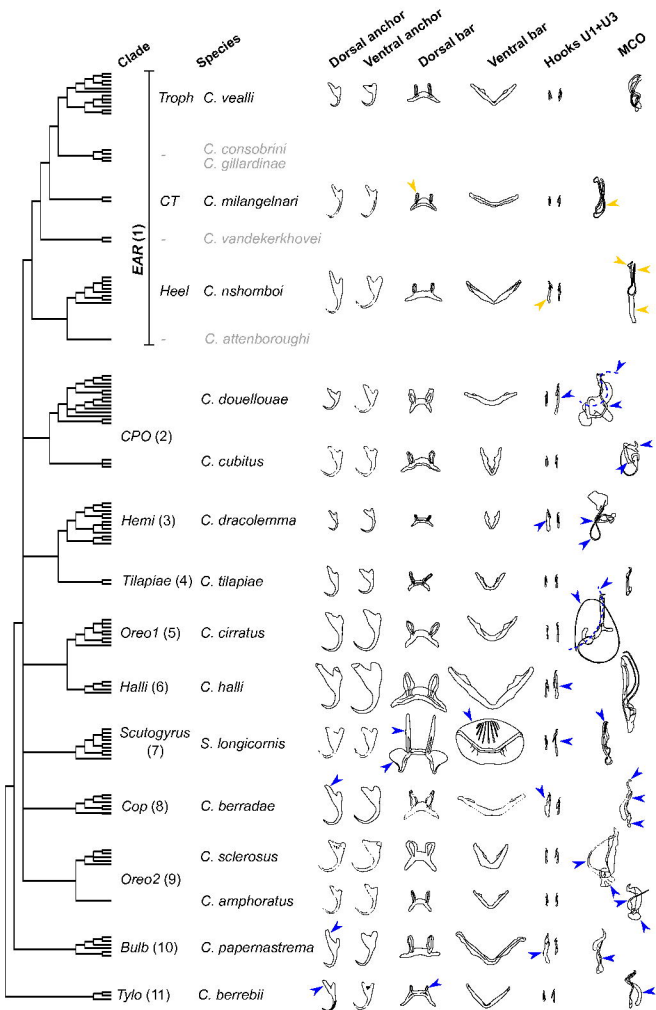
(b)

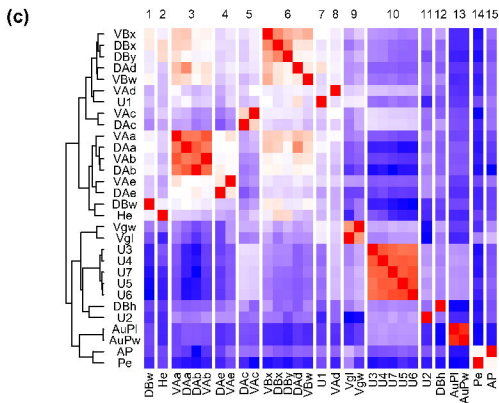
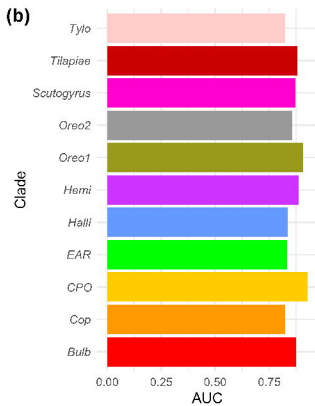
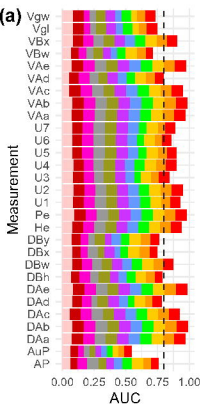


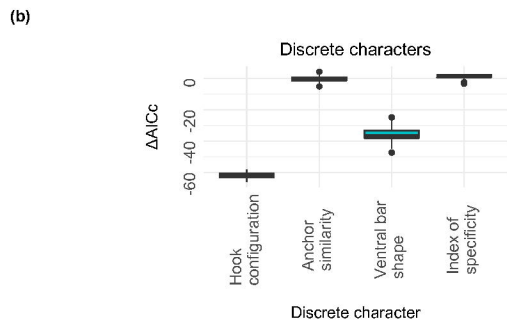
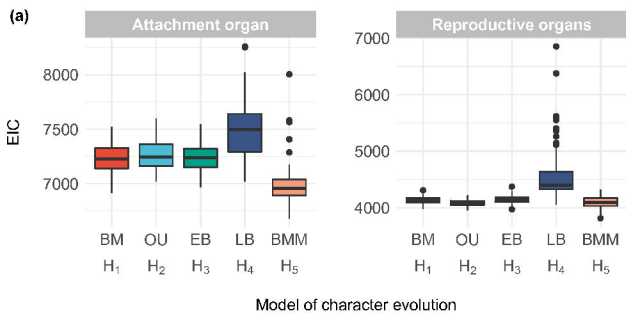
1

0.03



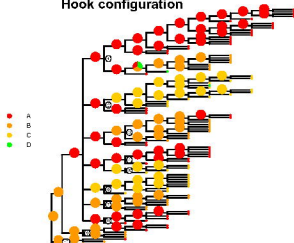






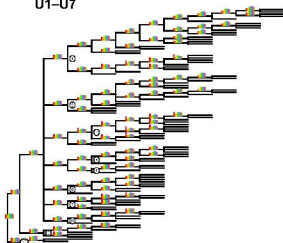
(a)

Hook configuration

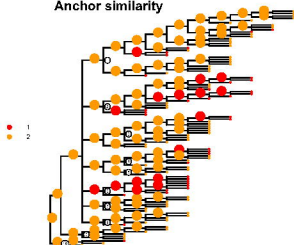


(b)

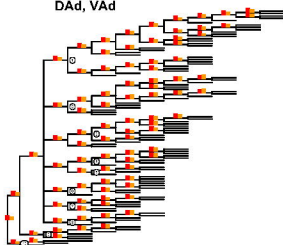
U1-U7



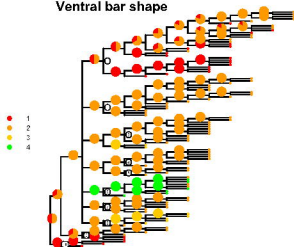
Anchor similarity



DAd, VAd



Ventral bar shape



VBw, VBx

