

Evolutionary dynamics of pre- and postzygotic reproductive isolation in cichlid fishes

Rometsch, S.J.; Torres-Dowdall, J. and Meyer, A.

Philosophical Transactions of the Royal Society B

Supplementary information

Evaluation of premating isolation in cichlids (Figure 1)

We conducted a literature research in “Web of Science” using the keywords:

“cichlid*” AND “mating”; “cichlid*” AND “assortative mating”; “cichlid*” AND “mate choice”; “cichlid*” AND “random mating”; “cichlid*” AND “reproductive isolation”. The literature search was conducted on 04.02.2020 between 13:31 and 13:45 o'clock (CET). Citation reports were exported and recurrent papers that were returned from more than one keyword search were removed. This resulted in a total of 803 studies (Supplementary Table S1). After removing theoretical papers, review articles and studies only citing cichlid literature but using focal species outside of the family Cichlidae, 497 relevant studies remained. Additionally, we added 21 studies that investigated assortative mating in cichlids but were not included in any of the keyword searches. Next, we filtered those papers for studies specifically testing for assortative mating between populations or species. This reduced the number of relevant studies to 39 (see Supplementary Table S1 and S2 for information on filtering). While some of those papers quantified the degree of assortative mating by the number of spawning events, eggs laid or other count data, others used preference scores (calculated as the difference in the proportions of females' positive responses to male courtship between con- and heterospecifics; 1) to describe the strength of premating isolation. We extracted those data from the literature. Due to the discrepancy in reporting the strength of premating isolation, Figure 1 in the paper only considers the studies reporting count data (since count data and preference scores provide different measures of premating isolation they could not be combined in our analysis), which reflects what the majority of studies reported (29 out of 39

studies). To obtain values for the frequency of expected assortative and disassortative mating for count data, as required for our meta-analysis, we assumed a scenario of random mating. Additionally, we considered the number of potential mates. Accordingly, if females could choose among two males in ten trials, the count data for expected assortative and disassortative under random mating would equal five, respectively. For the meta-analysis investigating the effect of manipulations on assortative mating, expected values were derived from results of respective experiments prior to manipulation.

Supplementary Figure S1 depicts the results of the meta-analysis conducted for the preference scores.

Meta-analysis was conducted using the software *OpenMEE*

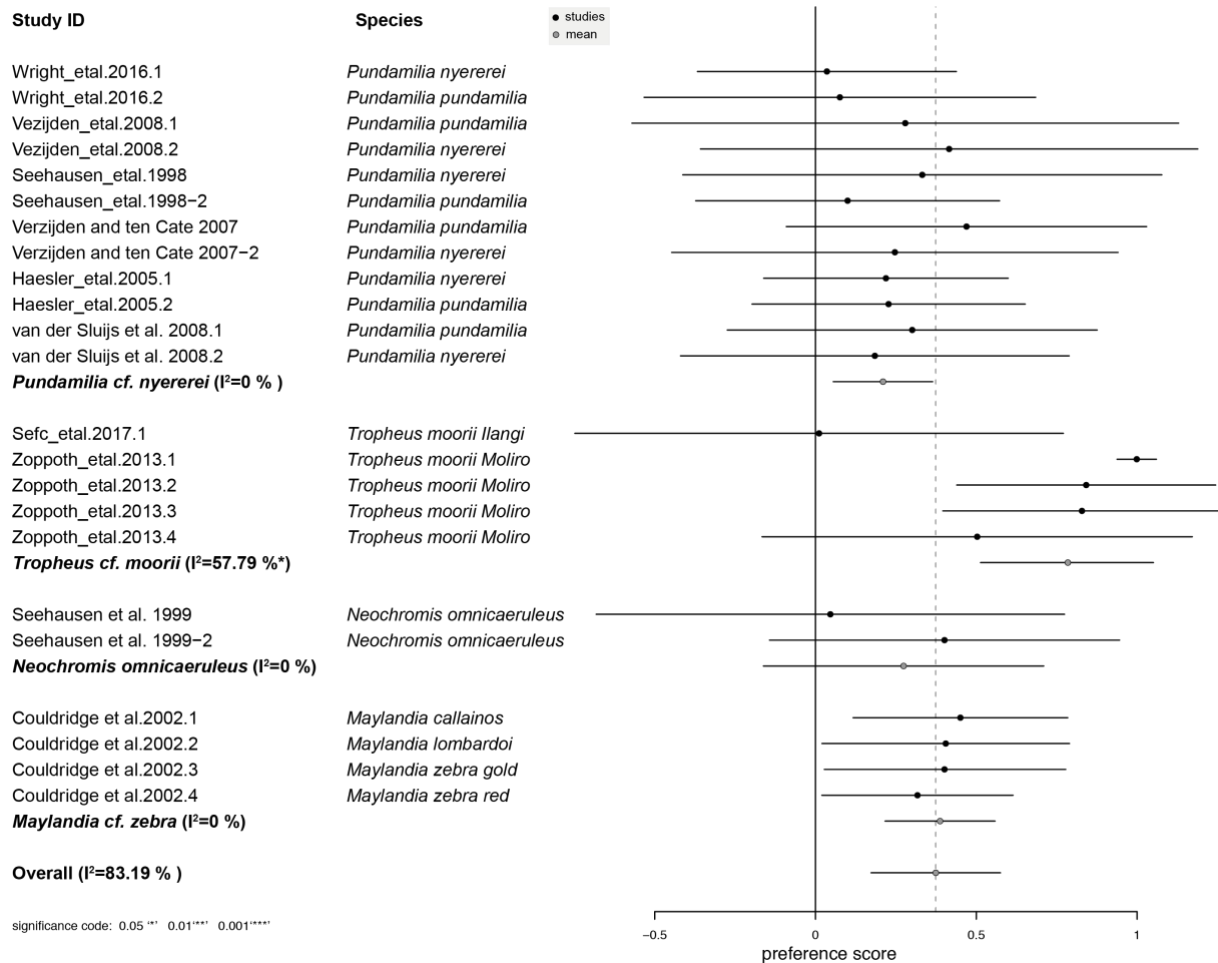
To analyze the data, we conducted a series of analyses at the meta-analysis level using the software *OpenMEE* (2). For the count data we estimated the odd ratios of mating with a conspecific versus mating with a heterospecific and its associated variance (see above). These two estimated values were used in downstream analyses. We conducted a phylogenetically controlled meta-analysis by modelling phylogenetic variance as a random effect (2). For this analysis, we constructed a simplified phylogenetic tree based on the full mitochondrial genome (Supplementary Table S4 and Supplementary Figure S4). A maximum likelihood (ML) analysis was performed in *PAUP* version 4.0a, build 167 (3), using the BioNJ method under the GRT model of nucleotide substitution. Closely related species were grouped together in this tree. Further, we aimed at improving our understanding of the different drivers of phylogenetic and among-study variance (τ^2_λ , τ^2). For this we used the subgroup analysis module in *OpenMEE*. As categories, we used lineage (*Pundamilia cf. nyererei*, *Maylandia cf. zebra*, *Tropheus cf. moori*, other haplochromines and Neotropical species; see Supplementary

Table S2 and Supplementary Figure S2) and the geographic setting of the current diversity (e.g., sympatric versus allopatric) and the taxonomic status of the studied taxa (population versus species; see Supplementary Table S2 and Supplementary Figure S3).

Evaluation of intrinsic incompatibilities in cichlids (Figure 2)

We scanned the literature for studies reporting interspecific crosses and their fitness consequences for the resulting F1 hybrid offspring. We extracted those crosses from the previously published literature (Supplementary Table S3). Viability was scored as 0 if F1s were reported to be viable and as 1 if F1 inviability or cumulative fitness of 1 (see 4) were reported. Sex ratio distortion was scored as 0 if no skew in F1 sex ratio was reported and as 1 if F1 sex ratio deviated from 1:1. D-loop sequences obtained from NCBI GenBank were used to compute pairwise genetic distances for interspecific crosses. For some species without available D-loop sequences in GenBank, we used sequences of closely related species (Supplementary Table S3). All sequences were aligned using the online version of MAFFT 7 (5). Pairwise genetic distances were calculated as uncorrected p-distances using MEGA version X (6). For interspecific comparisons with data on reciprocal crosses, we only included both cross directions when differences in viability or sex ratio distortion were reported and otherwise just included one of the crosses. We then analyzed intrinsic hybrid incompatibilities using a generalized linear model with binomial error distribution in R using F1 viability or sex ratio distortion as response variable and genetic distance as explanatory variable.

Supplementary Figures:



Supplementary Figure S1. Meta-analysis using preference scores. Evaluation of premating isolation in cichlids using preference scores shows the same qualitative trend as when considering count data: females prefer conspecific over heterospecific males (Figure 1). However, variation is more pronounced. The only lineage differing in the pattern of premating isolation when comparing meta-analyses based on count data and preference scores was *Tropheus cf. moorii*. Preference scores (difference in the proportions of females' positive responses to male courtship between con- and heterospecific males) are depicted with 95% CI as a behavioural proxy for premating isolation. Preference scores of 1 indicate preference for conspecific males and scores of -1 indicate preference for heterospecific males. I^2 = percentage of variability that is due to heterogeneity across studies rather than sampling variance, study IDs correspond to studies from which we extracted corresponding data (see Supplementary Table S2).

Study ID

Machado et al._1
 Machado et al._2
 Elmer et al._1
 Elmer et al._2
 Baylis
 Romer et al._2
 Ready et al._1
 Ready et al._2
 Ready et al._3

Neotropical species (I²=76.83 %*)**

Rajkov et al._1
 Theis et al._2
 Theis et al._1
 Rajkov et al._3
 Theis et al._3
 Rajkov et al._2
 Tyers et al._4
 Tyers et al._2
 Nichols et al._1
 Nichols et al._2
 Tyers et al._1
 Gerlai_2
 Pauers et al._1
 Tyers et al._5
 Tyers et al._6
 Genner et al._1
 Genner et al._3
 Gerlai_1

other haplochromines (I²=86.13 % *)**

Jordan et al.-2
 Jordan et al.-1
 Kidd et al._1
 Knight et al._1
 Plenderleith et al._1
 Knight et al.5
 Blais et al._1
 Knight et al.1
 Knight et al.2
 Knight et al.3
 Knight et al.4
 Blais et al._4
 Jordan et al.-3
 Kidd et al._1-2
 Knight et al._2
 Knight et al._3

Maylandia cf. zebra (I²=53.3 % **)

Svensson et al._2
 Selz et al.-4
 Selz et al.-1
 Seehausen_1
 Selz et al.-2
 Seehausen_3
 Selz et al.-3
 Svensson et al._1
 Seehausen_2
 Seehausen_4

Pundamilia cf. nyererei (I²=8.23 %)

Egger_et al._2
 Egger_et al._5
 Secf et al._2
 Egger_et al._3
 Secf et al._1
 Egger_et al._6
 Secf et al._3
 Egger_et al._4
 Egger_et al._1

Tropheus cf. moorii (I²=77.61 % *)**

Overall (I²=83.81 % *)**

significance code: 0.05 *** 0.01**** 0.001*****

Species

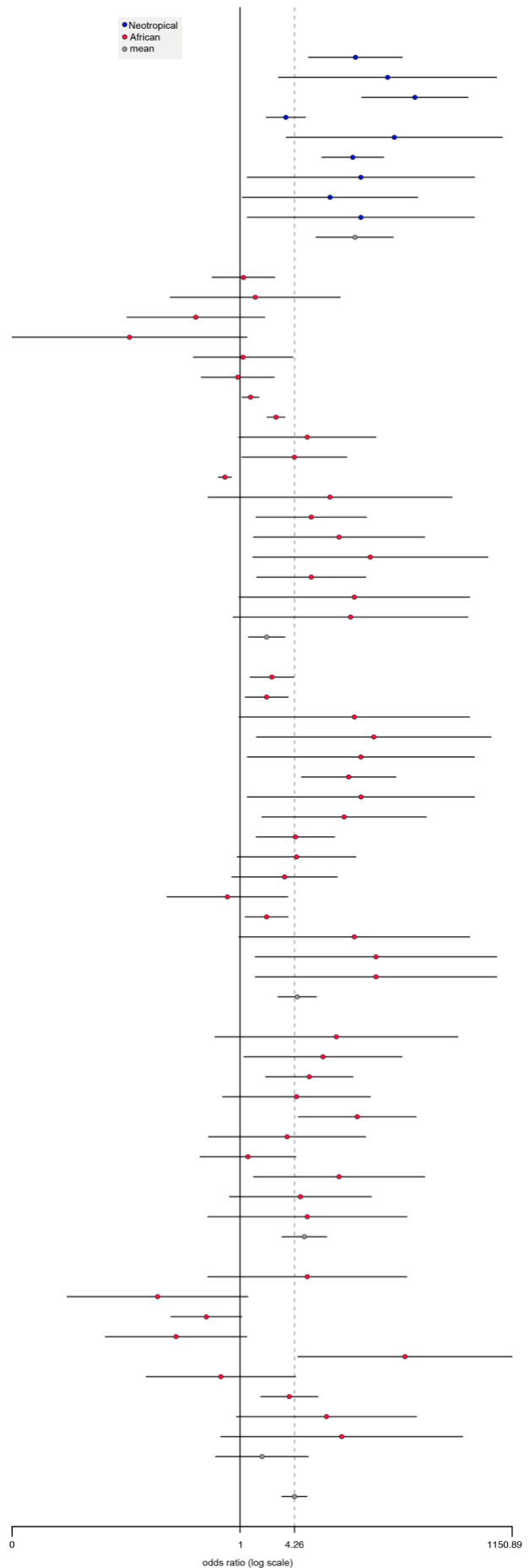
Amphilophus citrinellus
Amphilophus citrinellus
Amphilophus sagittae
Amphilophus xiloaensis
Amphilophus zaliosus
Apistogramma cacatuoides
Apistogramma caetei
Apistogramma caetei
Apistogramma caetei

Astatotilapia burtoni Kalambo lake
Astatotilapia burtoni Kalambo lake
Astatotilapia burtoni Kalambo river
Astatotilapia burtoni Lufubu river
Astatotilapia burtoni Lunza river
Astatotilapia burtoni Ndole lake
Astatotilapia calliptera "Chilingali"
Astatotilapia calliptera "Chizumulu"
Astatotilapia calliptera "Ruvuma"
Astatotilapia calliptera "Salima"
Astatotilapia calliptera "Salima"
Cynotilapia afra
Labeotropheus fuelleborni Chipoka
Lethrinops chilingalo
Lethrinops lethrinus
Rhamphochromis chilingali
Rhamphochromis chilingali
Sciaenochromis fryeri

Maylandia barlowi
Maylandia benetos
Maylandia benetos
Maylandia callainos
Maylandia emmiltos
Maylandia emmiltos
Maylandia emmiltos
Maylandia zebra Ruarwe
Maylandia zebra Chisumulu
Maylandia zebra Mara Rocks
Maylandia zebra Nkhata Bay
Maylandia zebra
Maylandia zebra
Maylandia zebra
Metriaclima zebra gold
Metriaclima zebra gold

Pundamilia "nyererei-like"
Pundamilia nyererei
Pundamilia nyererei Makobe
Pundamilia nyererei Makobe
Pundamilia nyererei Python
Pundamilia nyererei Python
Pundamilia nyererei Senga
Pundamilia "pundamilia-like"
Pundamilia "pundamilia" Makobe
Pundamilia pundamilia Python

Tropheus moorii blue "Nakaku"
Tropheus moorii Chimba
Tropheus moorii Mbita
Tropheus moorii Mbita
Tropheus moorii Moliro
Tropheus moorii Moliro
Tropheus moorii Nakaku
Tropheus moorii Nakaku
Tropheus moorii red "Moliro"



Supplementary Figure S2. Details of Meta-analysis considering different lineages. Detailed meta-analysis investigating if premating isolation differs among lineages (Figure 1B). Log odds ratios depicted with 95% CI indicate likelihood of premating isolation. Ratios of 1 indicate mating with con- versus heterospecifics is equally likely, values higher than one express the fold increase in likelihood of mating assortatively and values lower than one express the fold increase in mating disassortatively. I^2 = percentage of variability that is due to heterogeneity across studies rather than sampling variance, study IDs correspond to studies from which we extracted corresponding data (see Supplementary Table S2).

Study ID

Machado et al._1
 Elmer et al._1
 Elmer et al._2
 Seehausen_1
 Seehausen_3
 Seehausen_2
 Seehausen_4

sympatric populations (I²=75.74 %*)**

Machado et al._2
 Baylis
 Romer et al._2
 Gerlai_2
 Jordan et al.-2
 Jordan et al.-1
 Kidd et al._1
 Knight et al._1
 Plenderleith et al._1
 Jordan et al.-3
 Kidd et al._1-2
 Knight et al._2
 Knight et al._3
 Svensson et al._2
 Svensson et al._1
 Gerlai_1

sympatric species (I²=68.2 % *)**

Ready et al._1
 Ready et al._2
 Ready et al._3
 Rajkov et al._1
 Theis et al._2
 Theis et al._1
 Rajkov et al._3
 Theis et al._3
 Rajkov et al._2
 Tyers et al._4
 Tyers et al._2
 Nichols et al._1
 Nichols et al._2
 Tyers et al._1
 Pauers et al._1
 Knight et al.5
 Knight et al.1
 Knight et al.2
 Knight et al.3
 Knight et al.4
 Selz et al.-1
 Selz et al.-2
 Selz et al.-3
 Egger_et_al._2
 Egger_et_al._5
 Secf et al._2
 Egger_et_al._3
 Secf et al._1
 Egger_et_al._6
 Secf et al._3
 Egger_et_al._4
 Egger_et_al._1

allopatric populations (I²=84.6 % *)**

Tyers et al._5
 Tyers et al._6
 Blais et al._1
 Blais et al._4
 Selz et al.-4
 Genner et al._1
 Genner et al._3

allopatric species (I²=39.92 %)

Overall (I²=83.81 % *)**

Species

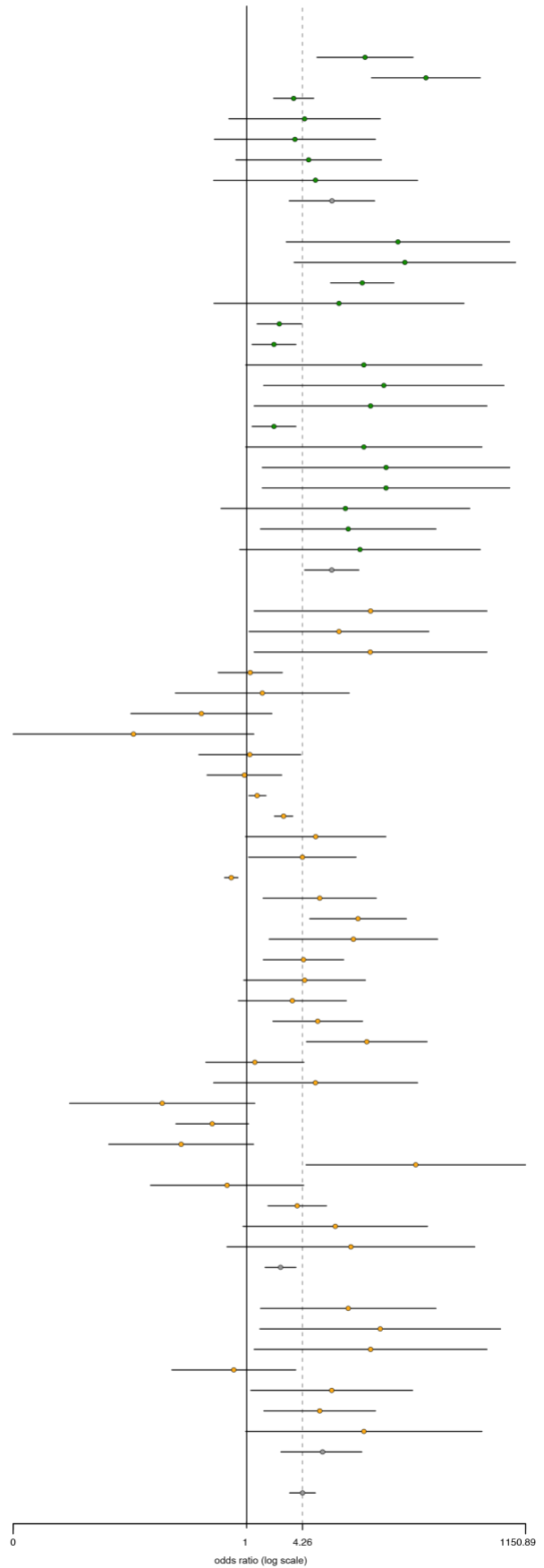
Amphilophus citrinellus
Amphilophus sagittae
Amphilophus xiloaensis
Pundamilia nyererei Makobe
Pundamilia nyererei Python
Pundamilia pundamilia Makobe
Pundamilia pundamilia Python

Amphilophus citrinellus
Amphilophus zaliosus
Apistogramma cacatuoides
Cynotilapia afra
Maylandia barlowi
Maylandia benetos
Maylandia benetos
Maylandia callainos
Maylandia emmitos
Maylandia zebra
Maylandia zebra
Maylandia zebra gold
Maylandia zebra
Pundamilia "nyererei-like"
Pundamilia "pundamilia-like"
Sciaenochromis fryeri

Apistogramma caetei
Apistogramma caetei
Apistogramma caetei
Astatotilapia burtoni Kalambo lake
Astatotilapia burtoni Kalambo lake
Astatotilapia burtoni Kalambo river
Astatotilapia burtoni Lufubu river
Astatotilapia burtoni Lunza river
Astatotilapia burtoni Ndole lake
Astatotilapia calliptera "Chilingali"
Astatotilapia calliptera "Chizumulu"
Astatotilapia calliptera "Ruvuma"
Astatotilapia calliptera "Salima"
Astatotilapia calliptera "Salima"
Labeotropheus fuelleborni Chipoka
Maylandia emmitos
Maylandia zebra Ruarwe
Maylandia zebra Chizumulu
Maylandia zebra Mara Rocks
Maylandia zebra Nkhata Bay
Pundamilia nyererei Makobe
Pundamilia nyererei Python
Pundamilia nyererei Senga
Tropheus moorii blue Nakaku
Tropheus moorii Chimba
Tropheus moorii Mbita
Tropheus moorii Mbita
Tropheus moorii Moliro
Tropheus moorii Moliro
Tropheus moorii Nakaku
Tropheus moorii Nakaku
Tropheus moorii red Moliro

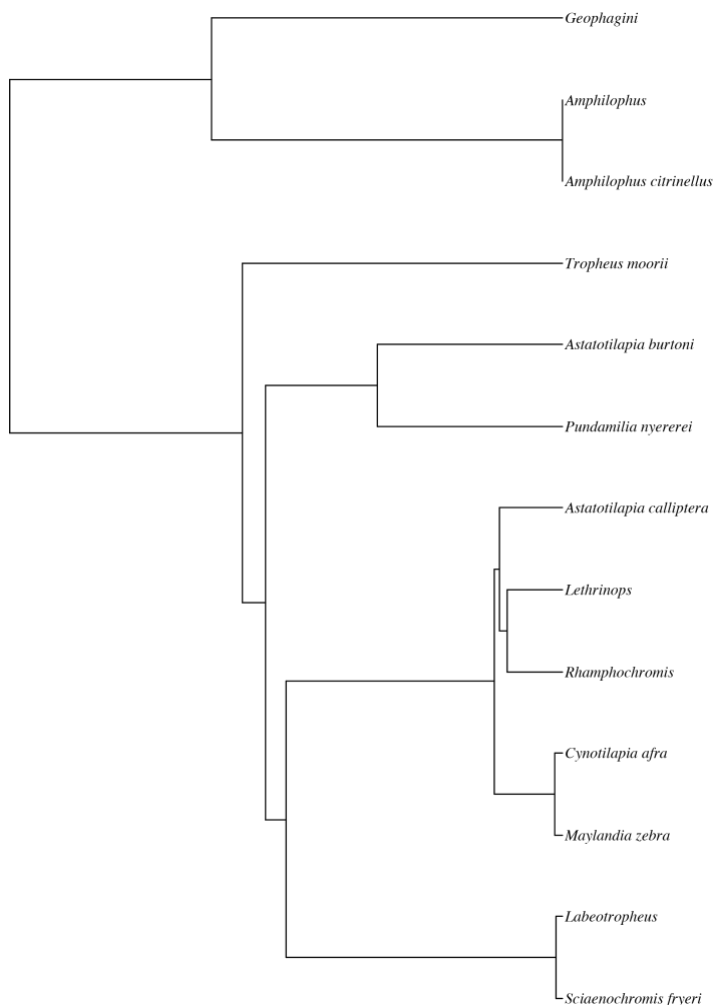
Lethrinops chilingali
Lethrinops lethrinus
Maylandia emmitos
Maylandia zebra
Pundamilia nyererei
Rhamphochromis chilingali
Rhamphochromis longiceps

● sympatric
 ● allopatric
 ○ mean



significance code: 0.05*** 0.01**** 0.001*****

Supplementary Figure S3. Details of Meta-analysis considering geographic settings. Detailed meta-analysis investigating if premating isolation differs among geographic settings and among populations and species (Figure 1C). Log odds ratios depicted with 95% CI indicate likelihood of premating isolation. Ratios of 1 indicate mating with con- versus heterospecifics is equally likely, values higher than one express the fold increase in likelihood of mating assortatively and values lower than one express the fold increase in mating disassortatively. I^2 = percentage of variability that is due to heterogeneity across studies rather than sampling variance, study IDs correspond to studies from which we extracted corresponding data (see Supplementary Table S2).



Supplementary Figure S4. Ultrametric tree based on mitochondrial genomes. A simplified phylogenetic tree was constructed based on full mitochondrial genomes (Supplementary Table S4). A chronogram was fitted to the tree using a strict clock model with the function `chronos` in package `ape` in R(7).

References:

1. Seehausen O, van Alphen JJ. The effect of male coloration on female mate choice in closely related Lake Victoria cichlids (*Haplochromis nyererei* complex). *Behavioral Ecology and Sociobiology*. 1998;42(1):1-8.
2. Wallace BC, Lajeunesse MJ, Dietz G, Dahabreh IJ, Trikalinos TA, Schmid CH, et al. Open MEE: Intuitive, open-source software for meta-analysis in ecology and evolutionary biology. *Methods in Ecology and Evolution*. 2017;8(8):941-7.
3. Swofford D. PAUP*. Phylogenetic analysis using parsimony (* and other methods). version. 2002;4(0):b4a.
4. Stelkens RB, Young KA, Seehausen O. The accumulation of reproductive incompatibilities in African cichlid fish. *Evolution: International Journal of Organic Evolution*. 2010;64(3):617-33.
5. Rozewicki J, Li S, Amada KM, Standley DM, Katoh K. MAFFT-DASH: integrated protein sequence and structural alignment. *Nucleic Acids Research*. 2019;47(W1):W5-W10.
6. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular biology and evolution*. 2018;35(6):1547-9.
7. Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*. 2019;35(3):526-8.