PNAS
www.pnas.org

Supplementary Information for

Polygenic sex determination produces modular sex polymorphism in an African cichlid fish

Emily C. Moore, Patrick J. Ciccotto, Erin N. Peterson, Melissa S. Lamm, R. Craig Albertson, and Reade B. Roberts

Emily C Moore and Reade B. Roberts.
E-mail: emily.christine.moore@gmail.com; rbrober2@ncsu.edu

This PDF file includes:

Other supplementary materials for this manuscript include the following:

Supporting Information Text

Expanded Materials and Methods

Animals. A wild-derived line of Metriaclima mbenjii (1) from Mbenji Island, Lake Malawi was maintained under Institutional Animal Care and Use Committee (IACUC) guidelines in the Roberts' Lab aquaculture facility at North Carolina State University in Raleigh, NC. Like several of the species within the genus, M. mbenjii has females of two pigmentation patterns, a "plain" (P) morph (figure 4.2 B) and an "orange blotch" (OB) morph (figure 4.2 C). This species breeds best with a single breeding pair, so we initially established a breeding group with an OB female and large, dominant male. Resultant F1 offspring were collected, raised to adulthood, and moved to 50- and 125-gallon aquaria to facilitate growth. Additional families were collected from F1 sibling matings and raised in the same manner. A total of 50 F1 from 3 families and 89 F2 from 3 families were collected over a span of three years for phenotyping, though not all individuals have data for all phenotypes. The gut length comparison included 42 F2 from 2 additional families.

Identification of segregating sex determiners. We used two methods to identify sex genotype in F1 and F2 individuals. For the female (ZW) system on chr. 5, we readily identified genotype by eye because it is linked to the OB pigmentation polymorphism; females that are heterozygous for the linked W and pax7ablotch alleles have disrupted "blotchy" dark melanophore patterning with orange/yellow xanthophore pigmentation in the remaining areas, whereas females with a ZZ genotype at the locus have a pax7aWT allele (which results in the "plain" morph, generally brown tinged with blue), and no dominant female sex determiner present (2, 3). We genotyped the male (XY) system on chr. 7 using established cichlid sex-linked microsatellite markers (3, 4).

Tagging and gross measurements. All individuals were anesthetized with buffered 100 mg/L tricaine methanesulfonate (MS-222). Following sedation, we measured fish for initial weight and standard length (from the base of the tail to the tip of the rostrum), photographed their ventral anogenital region under a dissecting microscope, and collected a small sample of tissue from the caudal fin. Condition was calculated as body mass divided by standard length. For 50 of 50 F1 and 62 of 64 F2, we also injected an 8mm x 1.4mm full duplex (FDX-B) passive integrated transponder (PIT) radio tag intraperitoneally (Oregon RFID) for future identification. Animals were allowed a few days to recover before returning to co-culture conditions with non-injected fish, and allowed at least one month to recover before further phenotyping.

Open field assay. Over two days, 34 individuals from the F1 families were assayed for behavioral response to a novel environment. Each fish was gently netted from the home system into a 50cm x 50cm opaque white arena filled with 15 liters of water from the standard recirculating system, lit at an angle to avoid water glare. An overhead camera captured video for 6 minutes with QuickTime Player 7.6.6 (Apple). Fish XY coordinates were tracked using C-Trax (v0.5.4); arena coordinates were added to the file and mistakes in tracking were fixed with the 'fixerrors' utility (v0.2.17) for C-trax in Matlab (vR2013b) (5). Summary values for position and speed in the arena were generated using custom R-scripts (R v3.3.1; (6)). All videos were taken between 10:00 and 16:00 to control for time of day.

Home tank behavior assay. We placed fish in a small recirculating aquaculture system with ten, 43 x 23 x 28 cm (20 liter) aquaria, each with a single small cave and white sand substrate and permanently mounted overhead video camera. Cohorts of up to ten individuals from all sex genotypes were acclimated to the mini-system for 24 hours, and then filmed between the hours of 11:00 and 16:00 two 30-minute assays, once with no disturbance to assess home tank behavior, and once after adding in a snail shell to assess response to a novel object. The videos were evaluated for time spent in the cave territory and number of grooming instances using the Observational Data Recorder (ODRec v2.0 beta), and summary statistics were generated with custom R-scripts (R v3.3.1).

Hormones. We used a non-lethal method for collecting hormones from holding water as follows (modified from (7)). From co-culture, each fish was placed in a blinded beaker with 300 mL of aquaculture system water for one hour. From the original 300 mL, 200 mL of collected water was run through a single-use C-18 chromatography column (Waters) pre-flushed with 4 mL ethanol and 4 mL reverse osmosis filtered, deionized water. The extracted hormones were eluted into 4 mL 100% molecular-grade ethanol, then pelleted using a SpeedVac vacuum centrifuge and stored at -80°C. After the samples for all individuals were collected, we chose 40 individuals to quantify for hormone levels, subsampling evenly across all four genotypes by family (i.e., the F1 family had 4 individuals represented from each genotypic class). Cortisol, 11-ketotestosterone, and estradiol were quantified using a colorimetric enzyme-linked immunoassay following manufacturer's instructions (EIA, Cayman chemicals).

Photography. Fish were anesthetized with buffered 100 mg/L MS-222 for all photographs. Whole fish ventral photographs were taken under controlled light conditions with a mirrorless digital camera (Olympus), and high-definition head photographs were taken using a dissecting microscope (Olympus SZX16/DP26). While anesthetized, an additional round of weight (g) and standard length (mm) measurements were taken.

Gut Length. To adequately control for plasticity in gastrointestinal development, we raised a separate cohort for gut length measurements. Two F2 families were raised in density-controlled aquaria with standardized measured feedings until 5 months, when they were humanely euthanized and measured for standard length, weight, and gut length. Gut length increases

allometrically with body size, so we controlled for overall size with the model Log10Gut Length (mm) Log10Body Mass (g) and used the residuals to test for differences by sex genotype.

Gonad histology. Once all assays were complete for a family, fish were humanely euthanized with 250 mg/L MS-222. Gonads were extracted and preserved in 10% neutral-buffered formalin, and processed for histological staining by wax embedding. Slides were stained with hematoxylin and eosin (H&E) to visualize gonadal tissue features.

Morphometrics and statistics. The left sides of anaesthetized fish were photographed and 20 landmarks were digitized from each image (Fig.1) in TPSdig v2.32 (8). Geometric morphometric analyses using these digitized landmarks were conducted in the R package geomorph (9). Landmarks were superimposed by a Generalized Procrustes Analysis (10) yielding Procrustes coordinates and centroid sizes for each fish to be analyzed in linear models. The two-dimensional Procrustes coordinates were set as the response variable and genotype was set as the factor. Log-transformed centroid size was included as a covariate to control for size-related changes in shape. A Procrustes ANOVA with random residual permutation procedures (Adams and Otarola-Castillo 2013) (RRPP; 1000 iterations) was run to quantify shape variation between the four genotypes using the procD.lm function (shape genotype * size; Table S1). Post-hoc pairwise comparisons of Procrustes distances between least squares means of the four genotypes were then conducted using the advanced.procD.lm function, using RRPP to test significance (11). Anogenital measurements were quantified from ventral photographs as described previously (12).

All other statistical analyses were performed using packages available in JMP v12 (SAS) and R v3.6.1(13). For the genital morphology and hormone data, sex genotype effects were modeled using ANOVA, and statistical groupings were determined by Tukey's pairwise comparison method (JMP). Hormone levels were adjusted from EIA concentration to account for the original volume of water in the test, and normalized to fish weight (reported as (pg/mL) [hormone concentration] / g [fish weight]). For behavioral tests, the amount of time spent in the cave territory was modeled with a lmer repeated measures mixed model including genotype and time as fixed effects and family as a random effect (lme4 v1.1-27.1; (14)). Counts of grooming, controlling for family as a random effect, were modeled with a generalized regression (glm) using a Poisson distribution to account for the skewed nature of count data (MASS v7.3-51.4; (15)). Comparisons between null mixed effects models and full mixed effects models were made with the anova function (lmerTest v3.1-3; (16)). Latency to groom was modeled with a Kaplan-Meier survival model function survfit (survival v3.1-8; (17)) and plotted as cumulative hazard (survminer v0.4.9; (18)).

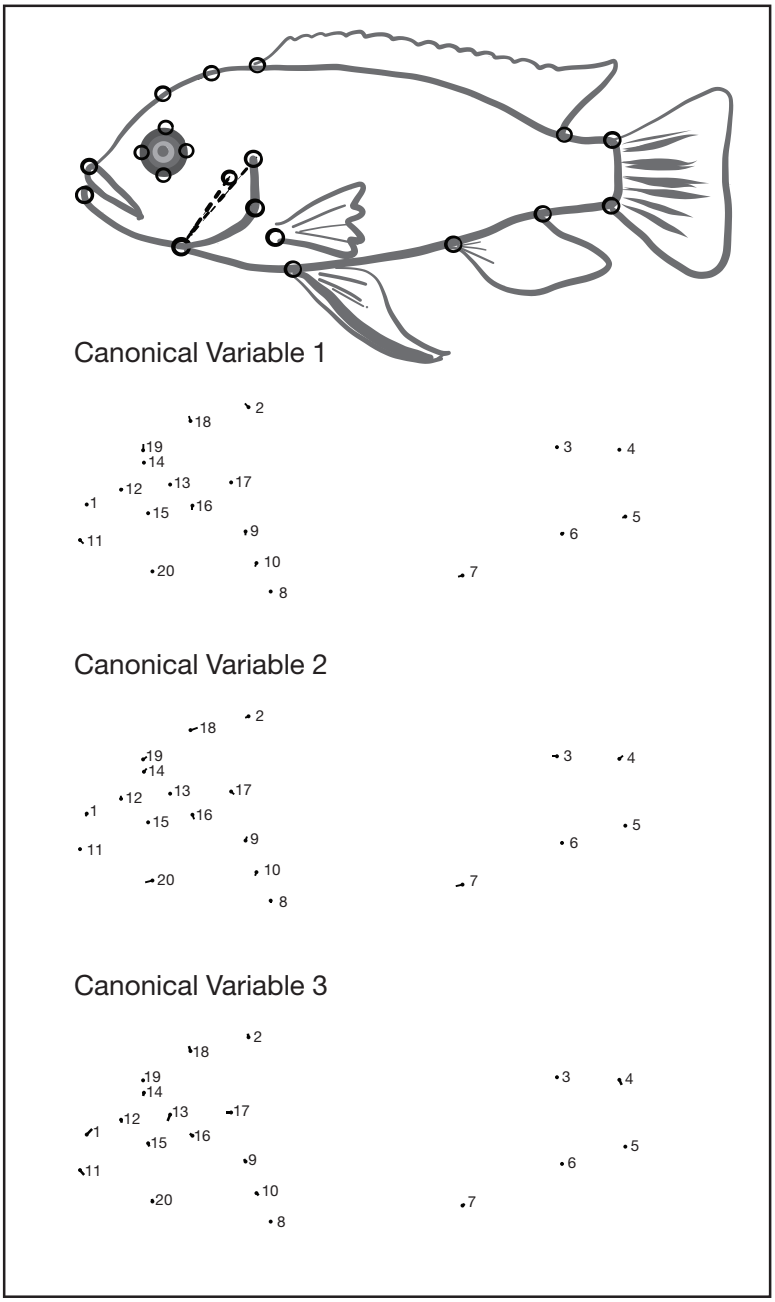

Fig. S1. Depictions of which morphometric landmark shifts comprise each canonical variable. Each dot indicates a standard position of a morphometric landmark when all groups are considered; all vectors directed off the landmarks indicate the specific shape change described by the multivariate canonical variable. Numbered points are as follows: 1) anterior tip of maxilla; 2) anterior insertion of dorsal fin; 3) posterior insertion of dorsal fin; 4) dorsal insertion of caudal fin; 5) ventral insertion of caudal fin; 6) posterior insertion of anal fin; 7) anterior insertion of anal fin; 8) insertion of pelvic fin; 9) dorsal insertion of pectoral fin; 10) ventral insertion of pectoral fin; 11) anterior tip of dentary; 12) anterior-most edge of eye; 13) posterior-most edge of eye; 14) dorsal-most edge of the eye; 15) ventral-most edge of eye; 16) dorsal tip of preoperculum; 17) dorsal tip of operculum; 18) intersection of the dorsal tip of the preoperculum with the dorsal profile of the body; 19) intersection of the dorsal-most edge of eye with the dorsal profile of the body; 20) intersection of the ventral-most edge eye with the ventral profile of the body

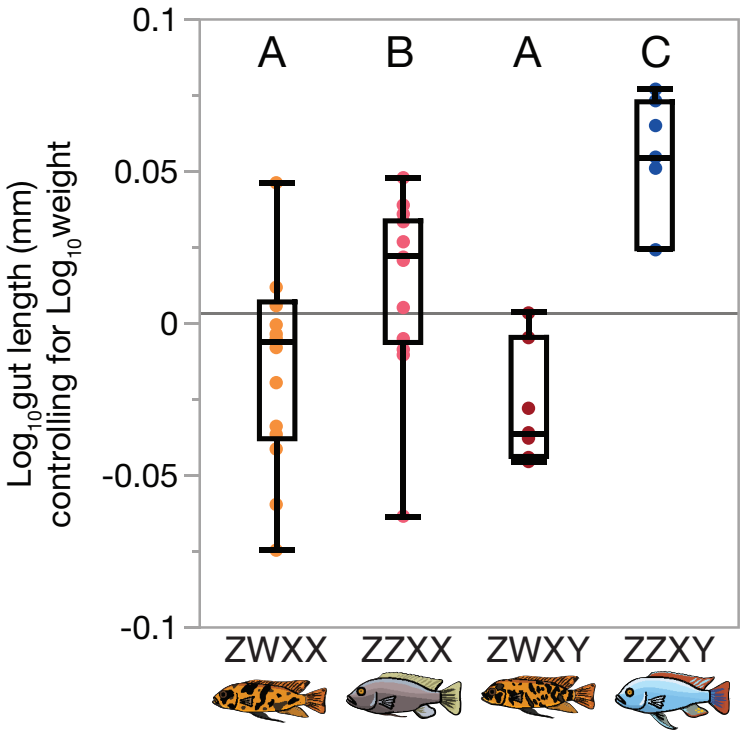

Fig. S2. Gut length differs by genotypic sex class. Residuals from a model where Log10 Gut Length (mm) Log10 weight show differences by sex genotype. Letters indicate statistical grouping determined by Tukey's HSD, $p > 0.05$

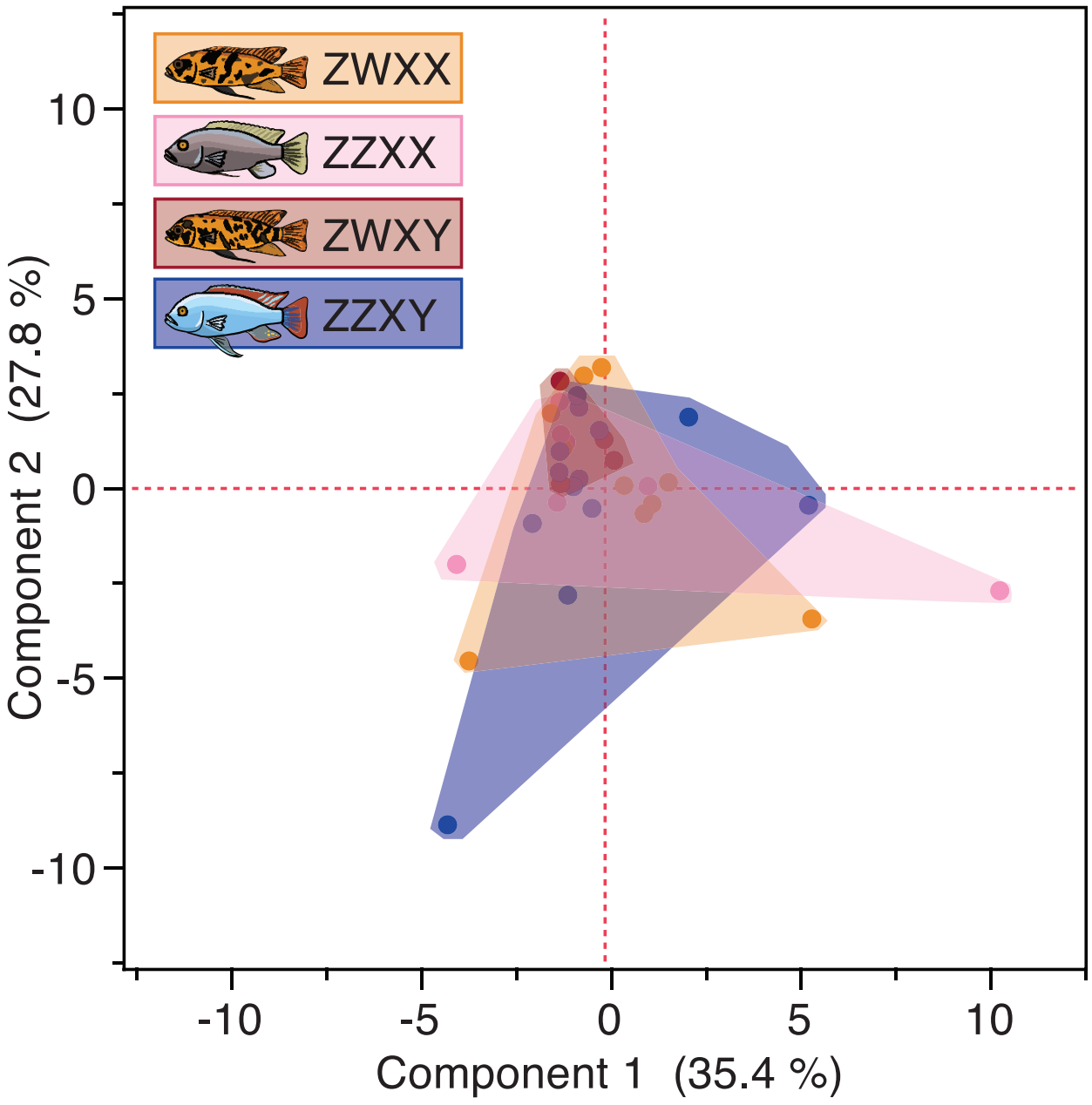

Fig. S3. Open field behaviors are not different by sex, as indicated by overlapping principle components 1 and 2 by genotype class

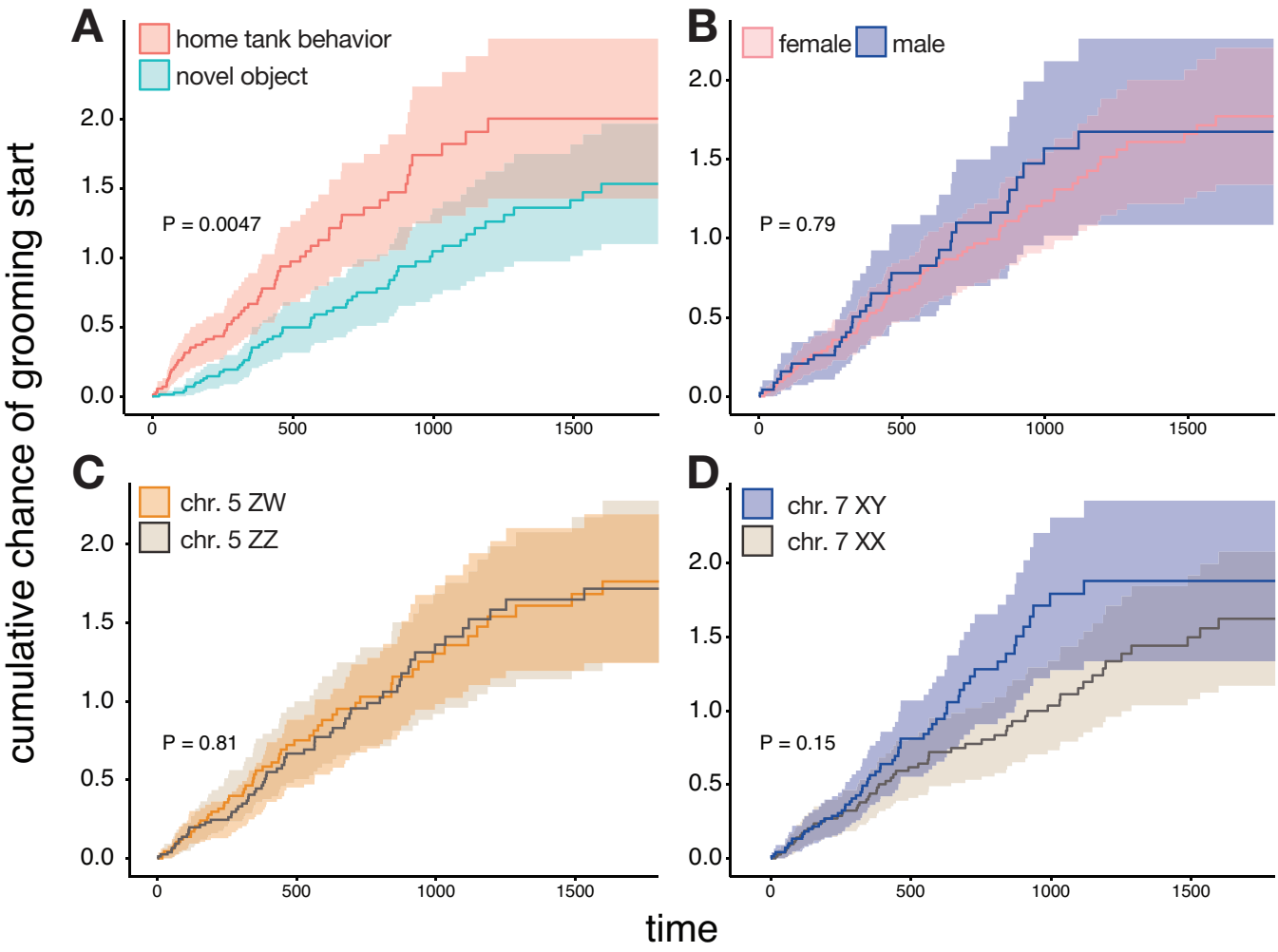

Fig. S4. Grooming latency. Cumulative hazard plots show the chance over time that individuals resumed territory grooming activity after the start of filming. P-values in panel result from Kaplan-Meier models of latency, with the assay (A), gonadal sex (B), chr. 5 sex genotype (C), and chr. 7 sex genotype (D) as model main effects

Table S1. Procrustes ANOVA of body shape and polygenic sex genotype. P values are from random residual permutation procedures (n = 1000). Significant effects in bold.

	df	SS	MS	F	P
Genotype	3	0.003	0.001	1.861	**0.004**
Log centroid size	1	0.001	0.001	2.042	**0.022**
Genotype*Log centroid size	3	0.001	0.0003	0.724	0.819
Residuals	53	0.028	0.001		
Total	60	0.034			

Table S2. Multivariate Procrustes distance metric by sex genotype.

Distances among groups

	ZWXX	ZWXY	ZZXX
ZWXY	0.0062	-	-
ZZXX	0.0127	0.0131	-
ZZXY	0.0127	0.0113	0.014

P-values from permutation tests (10000 rounds)

	ZWXX	ZWXY	ZZXX
ZWXY	0.8876	-	-
ZZXX	**0.0272**	0.1259	-
ZZXY	**0.0015**	0.0924	**0.0204**

Table S3. Territory occupancy models.

	home tank			shell		
Sex vs null	**P = 0.0002069**			**P = 0.00286**		
Genotype vs null	**P = 0.001314**			**P = 0.001323**		
Genotype vs sex	P = 0.3827			**P = 0.03374**		
	estimate	Z val	Pval	estimate	Z val	Pval
ZWXY	-13.327 ± 13.8	-0.965	0.33484	-37.5779 ± 14.8	-2.538	**0.011487**
ZZXX	7.059 ± 12.8	0.551	0.548819	-8.2461 ± 13.7	-0.6	0.548819
ZZXY	-36.669 ± 11.3	-3.254	**0.00123**	-42.1775 ± 12.1	-3.499	**0.000513**
time	-4.087 ± 0.52	-7.83	**3.73E-14**	-4.6171 ± 0.56	-8.211	**2.36E-15**

Table S4. Grooming count models.

	home tank			shell		
Sex vs null	**P = 1.004e-15**			**P = 0.002358**		
Genotype vs null	**P <2.2e-16**			**P <2.2e-16**		
Genotype vs sex	**P = 0.001711**			**P <2.2e-16**		
	estimate	Z val	Pval	estimate	Z val	Pval
ZWXY	-0.0758 ± 0.059	-1.289	0.197417	0.490587 ± 0.063	-2.538	**0.011487**
ZZXX	-0.201306 ± 0.057	-3.536	**0.000407**	-0.199897 ± 0.073	-0.6	0.548819
ZZXY	-0.23093 ± 0.045	5.183	**2.18E-07**	0.233542 ± 0.056	-3.499	**0.000513**
time	0.036598 ± 0.002	16.479	**<2.2e-16**	0.062976 ± 0.003	-8.211	**2.36E-15**

Table S5. Primary sex by sex genotype in Metriaclima mbenjii x Aulonocara koningsi F2 hybrids, from a cross of ZWXX x ZZXY F1 hybrids.

Genotype	Female	Male
ZZXX	41	0
ZWXX	44	0
ZWXY	0	35
ZZXY	0	47

SI Dataset S1 (Data_S1.xlsx)
Individual metadata and phenotypes

References

1. JR Stauffer Jr., NJ Bowers, KA Kellogg, KR McKaye, A revision of the blue-black pseudotropheus zebra (teleostei: Cichlidae) complex from lake malawi, africa, with a description of a new genus and ten new species. *Proc. Acad. Nat. Sci. Phila.* p. 189–230 (1997).
2. RB Roberts, JR Ser, TD Kocher, Sexual conflict resolved by invasion of a novel sex determiner in lake malawi cichlid fishes. *Science* **326**, 998–1001 (2009).
3. JR Ser, RB Roberts, TD Kocher, Multiple interacting loci control sex determination in lake malawi cichlid fish. *Evolution* **64**, 486–501 (2010).
4. EN Peterson, ME Cline, EC Moore, NB Roberts, RB Roberts, Genetic sex determination in astatotilapia calliptera, a prototype species for the lake malawi cichlid radiation. *Naturwissenschaften* **104**, 41 (2017).
5. K Branson, AA Robie, J Bender, P Perona, MH Dickinson, High-throughput ethomics in large groups of drosophila. *Nat Methods* **6**, 451–7 (2009).
6. ZV Johnson, et al., Exploratory behaviour is associated with microhabitat and evolutionary radiation in lake malawi cichlids. *Animal Behav.* **160**, 121–134 (2020).
7. SC Wong, M Dykstra, JM Campbell, RL Earley, Measuring water-borne cortisol in convict cichlids (amatitlania nigrofasciata): Is the procedure a stressor? *Behaviour* **145**, 1283–1305 (2008).
8. FJ Rohlf, tpsdig, digitize landmarks and outlines, version 2.05. (2005).
9. DC Adams, E Otárola-Castillo, E Paradis, geomorph: an r package for the collection and analysis of geometric morphometric shape data. *Methods Ecol. Evol.* **4**, 393–399 (2013).
10. JF Rohlf, D Slice, Extensions of the procrustes method for the optimal superimposition of landmarks. *Syst. Zool.* **39**, 40–59 (1990).
11. ML Collyer, DJ Sekora, DC Adams, A method for analysis of phenotypic change for phenotypes described by high-dimensional data. *Hered. (Edinb)* **115**, 357–65 (2015).
12. EC Moore, RB Roberts, Genital morphology and allometry differ by species and sex in malawi cichlid fishes. *Hydrobiologia* **791**, 127–143 (2016).
13. RCoreTeam, R: A language and environment for statistical computing. r foundation for statistical computing, vienna, austria (2017).
14. D Bates, M Mächler, B Bolker, S Walker, Fitting linear mixed-effects models usinglme4. *J. Stat. Softw.* **67** (2015).
15. W Venables, B Ripley, *Modern Applied Statistics with S.* (Springer, New York), Fourth edition edition, (2002).
16. A Kuznetsova, PB Brockhoff, RHB Christensen, lmertest package: Tests in linear mixed effects models. *J. Stat. Softw.* **82** (2017).
17. T Therneau, A package for survival analysis in r. version 2.38 (2015).
18. A Kassambara, M Kosinski, P Biecek, survminer: Drawing survival curves using 'ggplot2'. r package version 0.4.9. (2021).