# Deep learning reveals the role of copy number variation in the genetic architecture of a highly polymorphic sexual trait

Wouter van der Bijl, Jacelyn Shu, Versara S. Goberdhan, Linley Sherin, Maria Cortazar-Chinarro, Alberto Corral-Lopez & Judith E. Mank

## Abstract

The extraordinary variation in male guppy coloration has been a powerful model for studying the interplay of natural and sexual selection. Many guppy populations exhibit substantial Y-linkage of color traits, and this has hampered the identification of the genetic architecture underlying male guppy color, as well as clouded our understanding of how this exceptional level of diversity is maintained. Here we used a population with low levels of Y-linkage for color variation to identify the heritability and genetic basis of male color variation using convolutional neural networks for high-resolution phenotyping coupled with selection experiments and controlled pedigrees and whole-genome resequencing for Genome Wide Association Study (GWAS) of colour. Our phenotypic and genomic results converge to show that color patterning in guppies is a combination of many heritable features, each with a partially overlapping genetic architecture. Unusually, our GWAS results suggest that copy number variation (CNV) is responsible for much of the variation in colour in guppies, providing a potential mechanism for the maintenance of variation of this classic model trait.

## Introduction

Male guppies, *Poecilia reticulata*, exhibit remarkable variation in coloration, which has been extensively studied in the context of fitness (*1*) and female preference (*2*, *3*). The widespread variation in the extent and patterning of male color across wild populations (*4*, *5*) and the rapid response to changes in natural and artificial selection (*6*, *7*), coupled with high heritability estimates (*8*, *9*) all suggest that substantial genetic variation for this trait is maintained in guppy populations. Although female mate preference for rare male patterns likely acts to preserve some variation from the eroding forces of genetic drift and selection (*2*, *10*, *11*), the genetic architecture underlying extraordinary diversity in male guppy color remains mysterious, as does the maintenance of this diversity.

Part of the difficulty in determining the genetic basis of male guppy coloration lies in the multiple forms of variation. Males vary in both their overall patterning of both yellow-orange xanthophores and black melanophores, as well as in the extent and saturation of these colors. Therefore, the genetic basis of color variation should include loci involved in the migration of the embryonic neural crest cells that form the precursors of xanthophores and melanophores that constitute pattern variation (*12*), as well as genes involved in the synthesis and deposition of pigments. Importantly, orange carotenoids are not synthesized in de novo in vertebrates, but are collected from food sources (*13*) and modified prior to deposition.

Numerous studies have identified a significant Y chromosome component to male guppy coloration (*8*, *14–16*). This simplifies phenotypic observations of color pattern inheritance, as Y-linked patterns are inherited through patrilines from a father to all his sons, but complicates the identification of the genes underlying color variation, as the Y chromosome does not recombine and instead is inherited as a single haplotype. Here we take advantage of a population with low levels of Y-linkage for color variation (*9*) to identify the heritability and genetic basis of male color variation. We used convolutional neural networks for high-resolution phenotyping of 3229 fish coupled with selection experiments and controlled pedigrees to determine the heritability of specific color patterns as well as the overall extent of color as a percent of total body area. We then used our phenotype data for Genome Wide Association Study (GWAS) of colour to identify the genetic architecture of a suite of color parameters.

Our phenotypic and genomic results converge to show that colour patterning in guppies is a combination of many heritable features, each with a partially overlapping genetic architecture. Unusually, our GWAS results suggest that copy number variation (CNV) is responsible for much of the variation in colour in guppies.

## Artificial selection on orange coloration

Starting from an outbred laboratory population with large variation in color patterning, we created three replicate pairs of selection lines for reduced and increased amounts of orange coloration as a percent of total body area. Each male in the experiment (*n* = 3,229) had multiple photos taken from each side to track measurement error (*n* = 14,100). Using a series of neural networks, we performed image segmentation to separate the fish from the background, and to identify the orange and black coloration on the fish (Fig 1A, Supplementary Methods). In addition to orange, xanthophores also create the less frequent yellow color patches, which we include in our orange measurements. We then warped images to a common reference shape based on placed landmarks to enable fine-scale color mapping. Combined with a full pedigree, this dataset allowed us to trace the inheritance of color patterns across four generations (Fig 1B). While some body areas are almost always pigmented, color is rare in most places, giving each male a unique pattern.

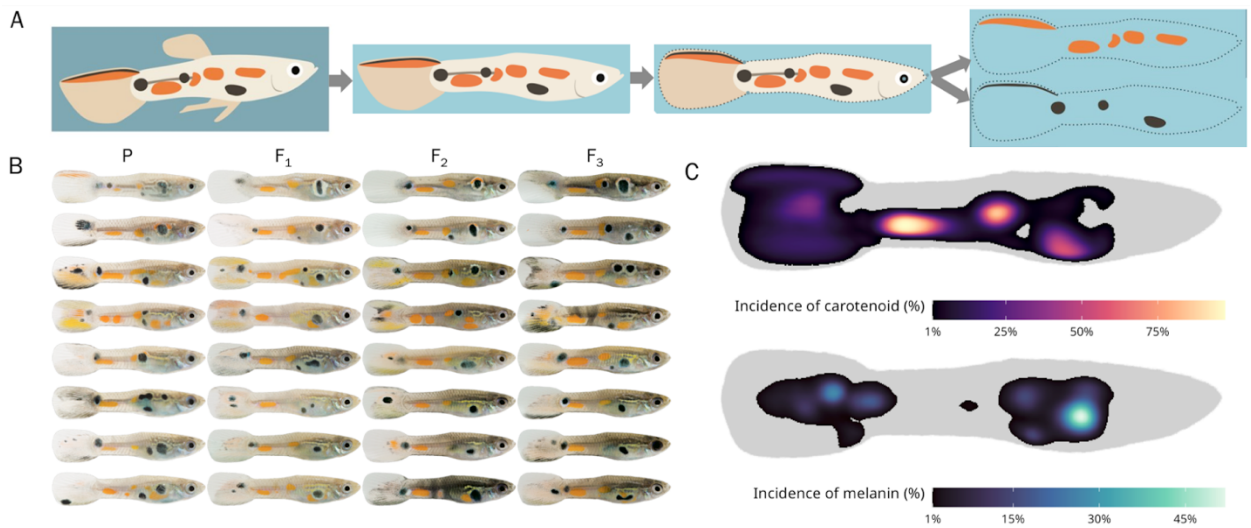


**Figure 1: Quantification of pattern variation**. A) Overview of the phenotyping pipeline. From each photo, the fish is extracted, the body is aligned to a common reference shape, and the position of orange and black color is extracted. B) Randomly selected examples of color inheritance along patrilines. Images on the same row belong to the same patriline, and columns are successive generations of the selection experiment. C) The incidence of color across the guppy body among all males. Body positions where the incidence of color is less than 1% are colored grey.

The amount percent of orange color responded rapidly and strongly to selection (Fig 2A). After just three generations, up-selected males had more than double the amount of orange color compared to their down-selected counterparts ($\chi^2 = 1236.8$, df = 25, $p < 0.001$). We modeled the inheritance of black and orange color along the pedigree (*17*), separating autosomal, X-linked and Y-linked genetic effects, as well as influences due to the maternal and tank environments. The strong response in orange color to

selection is consistent with its substantial heritability (narrow-sense $h^2$: 0.59, 95% credible interval: [0.50, 0.68], Fig S1). Orange and black ornaments share the same body plan, but selection on orange had nonetheless a negligible effect on the amount of black coloration ($\chi^2 = 28.63$, df = 25, $p = 0.28$). While the amount of black color is also substantially heritable ($h^2$: 0.67 [0.53, 0.80], Fig S2), the amount of the two colors is inherited largely independently as evidenced by their low genetic correlations ($r_{autosomal}$: 0.07 [-0.15, 0.26], Fig S3). While previous work on other populations (*8*, *14–16*) has identified strong Y-linkage of the amount of orange and black color, we find that most variation for total color is autosomal and, to a lesser extent, X-linked ($h^2_{Y\text{-linked}}$ orange: 0.03 [0.00, 0.09], black: 0.04 [0.00, 0.08], Figs S1&S2).

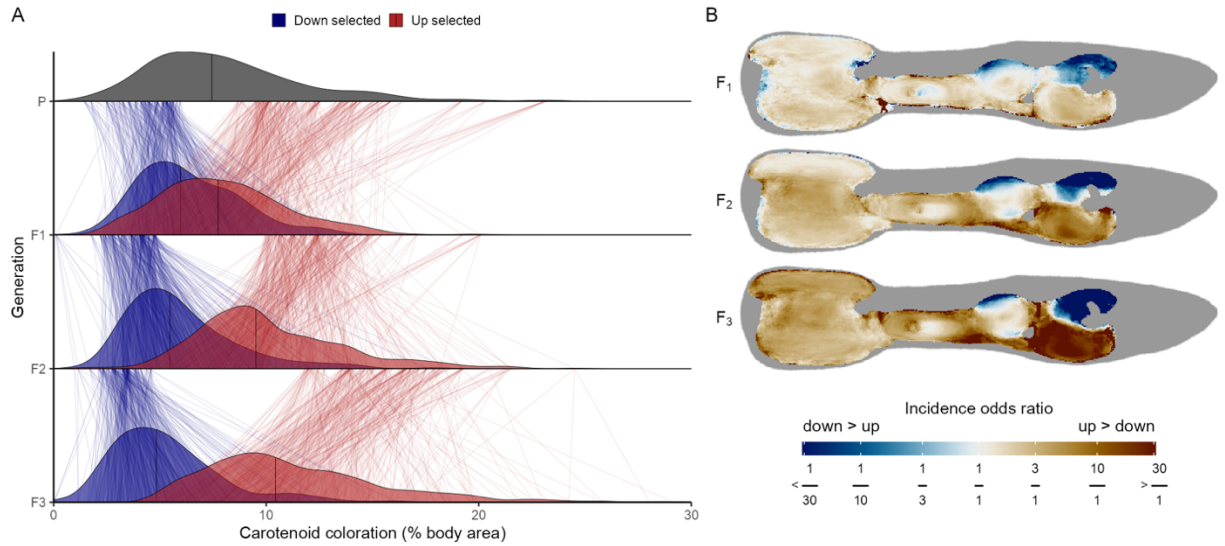


**Figure 2: Orange color responds rapidly to selection**. A) Density plots show the distribution of the percent of orange coloration per generation and selection regime. Thin lines between generations connect fathers and sons. Thick lines inside the densities show the median value of color area. B) Heatmaps illustrating the effect of selection across the body. Each heatmap cell is colored by the log odds ratio (as estimated by a generalized linear mixed model), illustrating the relative odds that a male has orange color at that location. Body positions where the incidence of orange color is less than 1% are colored grey.

Our selection regime affected the incidence of orange color at almost all positions on the body, with effect sizes increasing with each successive generation of selection (Fig 2B). The strength and direction of the effects were spatially heterogeneous. In the $F_3$ generation, up-selected males were approximately 5x more likely to have orange color at any particular position in the tail or caudal fin (posterior to the gonopodium). The strongest effects were located close to the pectoral fin and gills, where up-selected males were much more likely to have orange color on the ventral side, while down-selected males were more likely to be orange dorsally, with differences in the incidence of orange exceeding 100x in both cases (although this dorsal color was much less prevalent overall, Fig 1C).

Guppy males show off their colors during characteristic sigmoid display behavior, where they curve their body into an S-shape, quiver, and swim laterally in front of females (*18*, *19*). Using males from the $F_3$ generation, we evaluated the effects of our selection regime on male sexual behavior. In a standard no-choice test with a non-receptive female, males from the up-selected lines performed both more (short displays: $z = 2.99$, $p = 0.0028$, long displays: $z = 2.12$, $p = 0.034$) and longer ($z = 2.21$, $p = 0.027$) displays to the female (Fig S4). In addition, they spent more time following the female ($z = 2.40$, $p = 0.017$), and performed more attempts at coercive copulations ($z = 2.40$, $p = 0.016$), suggesting the more colorful males were more sexually active overall (Fig S5). Males did not show changes in body size or tail area (Fig S6), nor did selection alter the fecundity of females ($F_2$ offspring per brood, down-selected: 7.76, up-selected: 7.84, ratio [CI]: 0.98 [0.86, 1.10]) or the time between pairing and the first brood ($F_2$, difference in days [CI]: 1.187 [-1.06, 3.29]).

## Localized heritability, ornaments and pattern space

While the amounts of orange and black coloration are important to female choice (*5*, *20*), their patterning is largely inherited as discrete spots and stripes at fixed body positions (*9*, *14–16*). To understand the spatial structure of color inheritance, we modeled the heritability of the presence of color separately for each body location (Fig 3), confirming the fine-scale heritability of color patterning. In many places the localized heritability is high (> 0.8), much higher than the heritability of the total amount of color. When the different sources of heritability are separated into autosomal and sex-linked components, it is clear that the genetic architecture is spatially highly heterogeneous. For example, orange color in the top of the tail fin and black color ventrally of the dorsal fin were strongly X-linked, while both colors near the pectoral fin showed a large Y-linked component. This patchwork of genetic architecture assembles the complexity of color pattern.

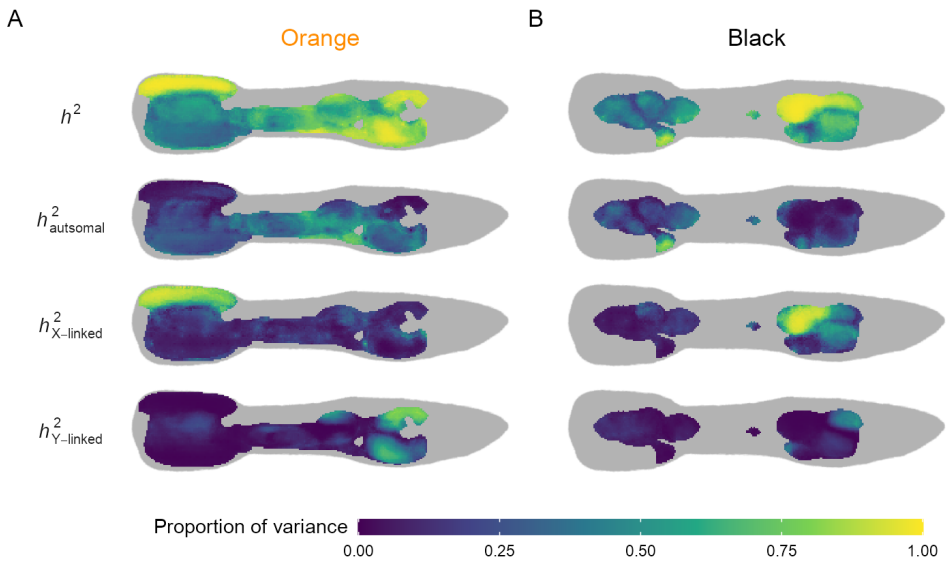


**Figure 3: The localized heritability of color is high, but the architecture is heterogeneous**. The narrow-sense heritability of the presence of orange (A) and black (B) color on the guppy body. For each heatmap cell, a Bayesian animal model was used to estimate the contribution of additive genetic and environmental factors. Heritabilities are expressed as a proportion of the total phenotypic variance. The autosomal, X-linked and Y-linked heritability add up to the total heritability. Body positions where the incidence of color is less than 1% are colored grey.

We defined 7 orange (Fig 4) and 8 black ornaments (Fig S7), based on the incidence of color (Fig 1C) and the localized genetic architecture (Fig 3). As expected, the incidence of all ornaments is heritable. In addition, their size is also heritable for all but one ornament. All orange ornaments responded to selection on the total amount of orange, either in their incidence, their size or both (Fig 4C & 4D). Concordant with the localized models, we find that the ornaments each have their own characteristic combination of autosomal, sex-linked and environmental causes. This is can also be seen from the low correlations between many different ornaments with autosomal or X-linked architectures. In contrast, we observed strong Y-linked genetic correlations (Figs S8 & S9), as the Y does not recombine. Within each ornament, however, the architecture of incidence and size are typically similar, which may indicate that additive genetic factors for ornament size also control its presence as a threshold trait.

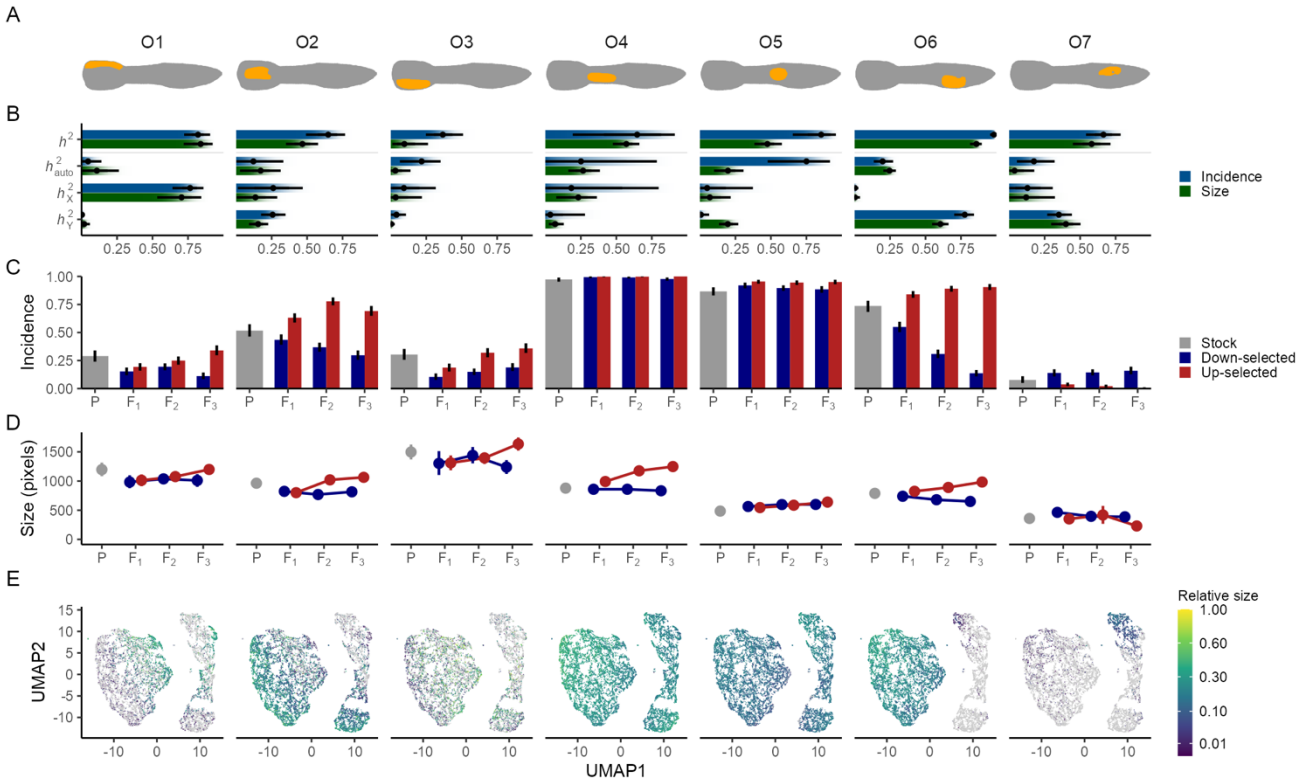


**Figure 4: Orange ornaments are heritable and their incidence and size responded to selection**. A) Pictograms of seven orange ornaments. B) Heritabilites of the incidence and size (when present) of each ornament. Dots and lines reflect point estimates and 95% credible intervals, and the gradient bars show the cumulative posterior density. C&D) Effect of selection on the incidence and size of orange ornaments. Error bars show 95% bootstrapped confidence intervals, and points represent means. X-axis show consecutive generations. E) Ornaments in orange pattern space. Points are the location of individuals in pattern space. Axes are the UMAP reduced representation of five-dimensional pattern space. Individuals are colored by the size of their relevant ornament, expressed as a fraction of the largest observed size, or colored grey if they lack the ornament. Note the use of a logarithmic scale for ornament size.

These localized approaches will underestimate the true pattern diversity in the population if separate color elements are co-located. For example, some males exhibit large patches of color in the tail that do not neatly fit the three ornaments defined there (Fig S10). Similarly, they do not capture potential heritable differences in spot shape or color characteristics such as hue and saturation. To describe the orange and black patterns free from assumptions about developmental mechanism, we developed a deep neural network embedder for each color which places the patterns in a 5-dimensional pattern space. Briefly, using triplet learning (*21*), these convolutional neural nets were trained to encode the heritable pattern variation (see Supplementary Methods). The networks are rewarded when they place patterns from related males closer together than those of unrelated males. As a result, patterns which are close in pattern space share heritable features, as can be observed when we map the presence and size of the ornaments onto this space (Fig 4E).

Orange pattern space is divided into two large clusters, differentiated by the presence of the highly heritable and Y-linked O6 ornament (Fig 4E), although we do not observe a further clustering of pattern space that would be suggestive of discrete color patterns. Complete Y-linkage of color patterns results in discrete patterns associated with each patriline that would be expected to form additional clusters. We did not observe this, rather, the many heritable features combine into a continuous space of pattern variation consistent with high levels of autosomal and X-linked variation that recombines. Our selection regime gradually moved the populations through pattern space over generations (Fig S11). Nonetheless, as there are many routes to becoming more or less orange, pattern variation is largely maintained within the selection lines among succeeding generations.

## Mapping color pattern to the genome

We selected 300 males from the $F_3$ generation for short-read whole genome resequencing, sampling across all families and across pattern space. We aligned these reads to the guppy female reference genome (*22*), which avoids reference bias to a particular Y haplotype, since guppy Y-chromosomes are highly diverse (*23*). After variant calling and filtering, we then performed multivariate GWAS separately for orange and black pattern space.

Consistent with the complex genetic architecture of the phenotypes, we found regions of strong association with orange and black across the genome (Fig 5). Highly significant associated SNPs span the entirety of linkage group (LG) 12, the sex chromosomes, likely due to strong linkage across this region. Given the low differentiation between the X and Y chromosome (*23*), most reads from the Y chromosome will nonetheless map to the X chromosome in the female reference genome, and SNPs on LG12 represent a combination of pseudo-autosomal, X-linked and Y-linked variants (*24*). Post-hoc analysis confirmed that ornaments with strongly sex-linked inheritance are associated most with LG12 (Fig S12) and that specific ornaments within the same biochemical color pathway are encoded through partially overlapping suites of loci, and that the pattern space GWAS (Fig 5) is largely a composite of these individual architectures.

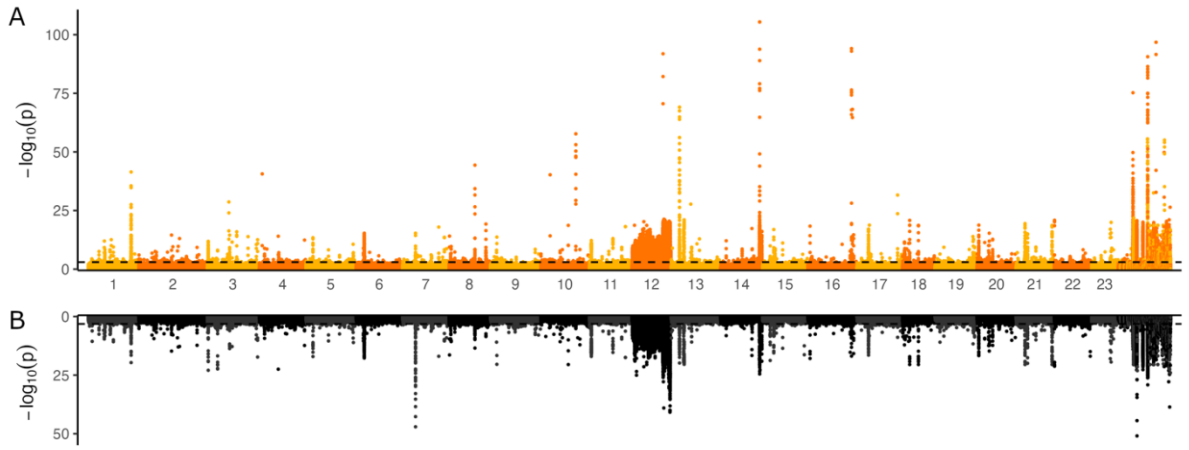


**Figure 5: Manhattan plots of genomic associations with orange and black color patterns***.* Results of multivariate GWAS on position in orange and black pattern space. Points represent SNPs and small indels, with their genomic location on the x-axis, and the p-value of the association on the y-axis. Numbers along the x axis denote linkage groups, with unmapped scaffolds plotted on the right-hand side. The dotted line represents the 1% FDR cut-off.

In addition, we observed peaks of associated SNPs on all autosomes. Strikingly, most of these regions showed elevated coverage and heterozygosity, suggesting reads from copy number variants (CNVs). Given this pattern, recent work that identified CNVs spanning autosomes and the Y chromosome in natural guppy populations (*25*), and the role of sex-linked variation in this system, we hypothesized that reads from duplications on sex-linked regions not represented in the female reference could map to autosomal regions, representing CNVs spanning multiple genomic locations. To elucidate the source of these SNPs, we compared the likelihood of autosomal, X- or Y-linked inheritance of the SNPs in these regions based on our pedigree (Fig 6).

The most strongly associated locus mapped to LG14, with many significant SNPs in a 35kb region (Fig 6A). The inheritance of most of these loci fits a pattern of Y-linkage, which suggests these variants are fixed differences between the autosomal region and its copy or copies on the Y. The gene located in this region (LOC103476393) is uncharacterized in the annotation, but has 72% sequence identity with *TeximYa* from the related Southern platyfish (*Xiphophorus maculatus*). Copies of *Texim* have also been found on the Y-chromosome of the closely related *Poecilia parae* (*26*), suggesting duplication of *Texim* to the Y-chromosome is common in this clade. The most strongly associated SNPs in this region are located downstream of this *Texim* copy, and almost perfectly predict the presence of the Y-linked ornament 06 (Fig 6C), and is more weakly associated with the presence of O2 and the size of O4 (Fig 6D). Coverage for most of this region is elevated five to ten-fold and heterozygous individuals have increased coverage compared to individuals that are homozygous reference (Fig 6C), indicating these SNPs may tag additional Y-linked CNVs.

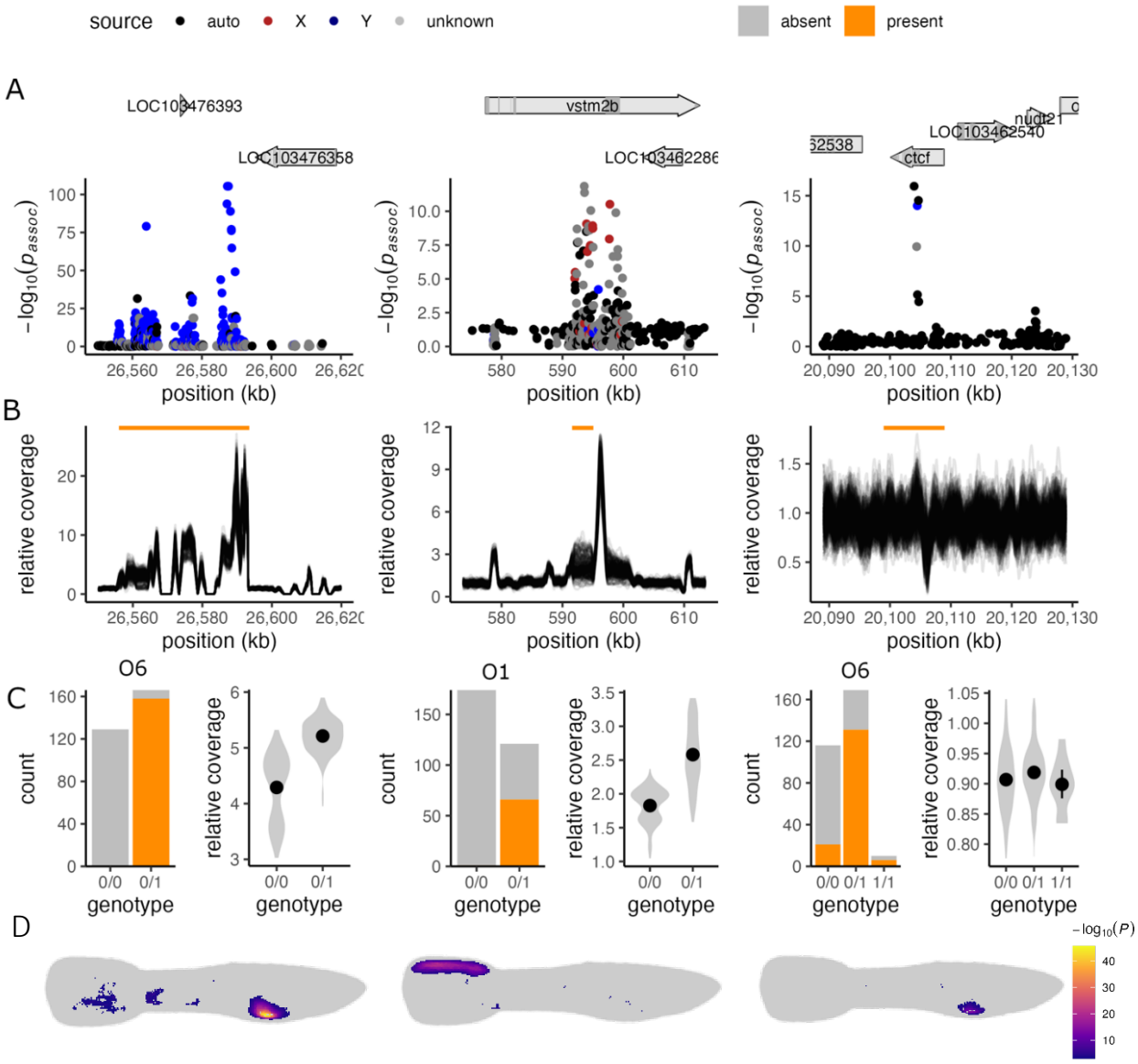


**Figure 6: Examples of a Y-linked, X-linked and autosomal GWAS peak**. A) P-values for the GWAS association with orange color pattern. Variants are colored by their likely source based on pedigree information, with ambiguous variants colored grey (AIC weight < 80%). Genes in the region are shown as arrows, with dark shaded areas representing exons. B) Relative coverage in the same regions as A. Each line represents an individual sample. Coverage is calculated as the average across 1kb sliding windows, with a distance of 100bp, and a relative coverage of 1 is the genome-wide average for that sample. C) Associations of the genotype at the top SNP with the absence/presence of the ornament (bars), and coverage (violin plots). The region across which coverage was calculated is indicated with the orange line in panel B. Points and error bars denote the mean and 95% bootstrapped confidence interval. D) Localized effect of the top SNP on the incidence of orange coloration across the body. P-values are adjusted for genome-wide relatedness, and values > 0.001 are not shown.

Other regions show a similar pattern of CNV, but the additional reads originate from the X-chromosome instead. For example, a region on LG3 contains significant SNPs of both autosomal and X-like

inheritance, and the top SNP is strongly predictive of the presence of the X-linked ornament O1. Also here, heterozygotes have increased coverage suggesting that X-linked CNVs are associated with this ornament. The duplicated region is intragenic, spanning a large exon of the gene *vstm2b*, a transmembrane protein with unknown function.

Across most peaks, similar to the two examples above, the genotype at the SNP is strongly related to the coverage. Since the SNPs mark both allelic variation and gene dose, we cannot differentiate the role of this sex-linked structural variation. It is notable however, that the genotype with increased coverage encodes the presence of ornaments in almost all cases, suggesting an increase in dose is associated with greater color.

Finally, a minority of GWAS peaks are not associated with sex-linked variation. For example, a significant locus on LG3 is inherited autosomally, and lacks increased coverage. Six SNPs show an association with orange pattern, and fall in the gene *ctcf*, which is involved in regulating gene expression and chromatin architecture. Individuals that were homozygous reference at this locus were less likely to present ornament O6 ($p < 0.001$). The limited detection of autosomal variants suggests that the autosomal contributions to pattern variation are highly polygenic and difficult to detect.

## Discussion

The remarkable variation in male guppy coloration has made them a model for studies of female preference (*2*, *3*) and the interaction of sexual and natural selection (*4*, *27*). Male patterns in many populations have been shown to exhibit strong patterns of Y-linkage (*8*, *14*, *16*). This Y-linkage complicates detailed analysis of the genetic architecture of male color variation, as fine-scale genetic mapping requires recombination. Here we used a population with low levels of Y-linkage for color variation (*9*) to dissect the detailed genetic architecture for male color variation, using a combination of high-resolution phenotyping via neural nets, controlled pedigrees, and GWAS.

Our population has shown rapid response to artificial selection for other complex traits (*28–30*), suggesting significant underlying genetic variation. Consistent with this, we observe substantial variation in male color (Fig 1) and rapid response to artificial selection for increased and decreased orange coloration in all our replicate selection lines (Fig 2). Male color has previously been shown to respond to artificial selection in other populations (*7*, *31*, *32*), and consistent with this, we observe high heritability (Figs 3 and 4) for both total color area and specific color ornaments.

One might expect that if black is epistatic to orange, selection to increase the percent of the fish with orange coloration would reduce the total black area. That is not the case, and we observe little change in black area throughout the selection experiment (Fig S7), more consistent with the repulsion of developing xanthophores and melanophores during color development (*33*).

We used then used whole genome sequencing for 300 individual males from the $F_3$ generation, sampled to maximize their distribution within color-space to perform GWAS for multiple color traits (Figs 5 and 6). Interestingly, genes previously identified as key to melanophore formation in guppies, *colony-stimulating factor 1 receptor a* (also called *fms*), *kita* (*34*, *35*) and *adenylate cyclase 5* (*adcy5*) (*36*), in guppies did not appear on our GWAS (Fig S13). This suggests that although these loci are fundamental and necessary in the development of color cells, variation in them does not contribute substantially to the variation in color we observe in our population.

In contrast to melanin, carotenoids are not synthesized in vertebrates de novo, but rather collected from dietary intake and modified prior to deposition (*37*). This mode of color development is hypothesized to make carotenoid-based coloration a key mediator of honest sexual signalling of male quality (*1*, *38*, *39*). Consistent with this, males from selection lines with greater orange area exhibited greater sexual vigor (Fig S4 & S5), though we did not observe any correlated changes in female fitness traits.

We might expect genes associated with carotenoid metabolism in other vertebrates to play a role in orange coloration in guppies. A key gene identified in red coloration of birds, cytochrome P450 locus CYP2J19 (*40*, *41*), lacks a reciprocal ortholog in teleost fish. Aldehyde dehydrogenase family 1 member A1, associated with red carotenoid coloration in lizards (*42*), is also not implicated in our GWAS.

Instead, our results suggest a far more complicated genetic architecture for color pattern and total color area. For both black and orange color, we observe specific ornaments encoded through a partially overlapping suites of loci (Fig S12). Moreover, the architecture is heavily influenced by CNVs, although it is challenging to differentiate between gene dose effects, with more copies of a gene associated with greater color, and allelic variation, with some copies associated with specific patterns. This aspect of the genetic architecture of guppy color is particularly interesting in that it might be expected to help maintain male phenotypic color variation more effectively than allelic variation in single-copy genes. It has been a long-standing mystery how the remarkable variation in male color is maintained against the eroding forces of genetic drift, selection and sweeps. Although female mate preferences for rare male

patterns can be expected to preserve male patterning variation (*2*, *10*, *11*), it is difficult to envisage how this would be sufficient to maintain the extensive variation observed in guppy populations, many of which formed with significant bottlenecks and founder effects.

The fact that we observe CNV associated with color spanning autosomes and the sex chromosome also may help to explain recent observations of variation in the extent of Y-linkage for color in wild populations (*43*, *44*). Although the size of the fully non-recombining region of the Y chromosome is relatively small (*24*, *45*), there is evidence that it can accumulate substantial recent duplications of genes from the remainder of the genome (*25*). If this process involves coloration loci and varies across populations, it is easy to envisage how the association between the Y chromosome and color variation could differ, even among closely related populations. It may also suggest that the low-levels of Y-linkage we observe in our population, originally collected from the high predation region of the Quare River, Trinidad, may be observed in other natural populations that either have lower rates of duplication to the Y, or lower levels of Y-duplicate retention.

Taken together, our results present a critical link between the extraordinary variation in male guppy coloration and the genome that encodes it. The fine-scale genetic architecture of color that we reveal presents an exciting opportunity for future molecular analyses of sexual selection dynamics and a critical point of comparison for future studies of the genetic architecture of color in other populations.

## Acknowledgements

We thank Lengxob Yong for sharing insights on guppy photography. We thank Ben Sandkam, Ben Furman, David Metzger, Lydia Fong, Yuying Lin, Christina Hodson and Linley Sherin for help with animal husbandry and valuable discussions. Funding was provided by a Canada 150 Research Chair, NSERC and an ERC grant (680951) to JEM.

## Author contributions

Conceptualization: WvdB, AC-L & JEM, Methodology: WvdB, Formal analysis: WvdB & VSG, Investigation: WvdB, JS, VSG, LS, MC-C, AC-L & JEM, Writing – Original Draft: WvdB & JEM, Writing – Review & Editing: WvdB, JS, VSG, LS, MC-C, AC-L & JEM, Visualization: WvdB, Supervision: JEM, Project administration: JS, Funding acquisition: JEM.

# Materials and methods

## Fish and animal husbandry

Our study population are descendants of guppies that were captured in Trinidad in 1998 and subsequently maintained in large, outbreeding stock populations to maintain genetic diversity, and males show large diversity in their ornamentation. 240 females for the parental generation (P) were randomly sampled as juveniles from our three stock populations and maintained as virgins until sexual maturity. 300 males for the parental generation were sampled as adults from the stock populations.

Fish were kept in 1.4 l tanks in three flow-through systems, with each experimental replicate in a separate system. Water was maintained at 25 °C, and lights were on a 12:12h cycle. Fish were fed daily with a mix of live brine shrimp and tropical fish micro pellets (Hikari, Japan), with immature fish fed an additional meal on weekdays of frozen daphnia and blood worms.

## Experimental procedure and selection

Each replicate was started using 100 males, evenly sampled across the three stock populations. We quantified the amount of carotenoid coloration for each male (see *Color pattern extraction and quantification*, below), and selected the 30 males with the largest and 30 males with the least total orange area to form a paired up- and down-selected line. For the P generation, these males were then randomly paired with females. To increase our statistical power to detect sex-linked variation, we paired 10 males of each line with two females each to generate half siblings, and 20 males with a single female, creating a partial half-sib breeding design. For subsequent generations, we selected the 30 males with greatest orange area in up-selected lines, and the 30 males with least orange for down-selected lines. Additionally, we increased the strength of selection by also selecting on females. Since females do not express carotenoid coloration, we used an animal model to estimate breeding values (see *Quantitative genetics,* below). We then selected the females with greatest and least breeding value (40 for each category), incorporating uncertainty by ranking across draws from the posterior distribution. To avoid a rapid loss of genetic variation, and limit the selective advantage of males paired with two females, we further limited our selection to include a maximum of three sons from each sire, and three daughters from each dam.

For the parental trios, females were kept in separate tanks, and the male was moved between females each week. We collected broods from the up- and down-selected lines until we had 90 broods per generation per replicate. 93.3% of paired males and 89.3% of paired females successfully sired offspring,

resulting in an average of 1.46 broods per sire, and 1.10 broods per dam. Having >1 brood per pair allowed us to estimate the magnitude of brood effects.

Broods were collected within a few days of birth to avoid cannibalism, and moved to their own tank. To maintain developmental rates, broods larger than 10 individuals were randomly reduced to 10. Broods were checked three times per week for signs of male maturation (first signs of gonopodium growth), and males were moved to a separate tank to avoid sibling mating. The tanks with male brood members were then queued for phenotyping after visual conformation of sexual maturity (fully developed color pattern), and a minimum age of 60 days. After phenotyping (see *Color pattern extraction and quantification*), males were housed in individual tanks to maintain their identities.

The three replicates were set up staggered in time, so that one replicate could be phenotyped while the others were breeding and maturing. Fish were only marked with their individual IDs, blinding the experimenters as to which line individuals belonged to throughout experimental procedures and phenotyping (except when setting up the next generation of breeding pairs).

### Animal photography and calibration

We (WvdB & JS) captured photos of male color patterns using a Canon EOS Rebel T7i (Canon, Japan) with Canon EF 100mm f/2.8 Macro USM lens (Canon, Japan). All photos were captured using identical camera settings, at an aperture of f/11, an exposure time of 1/320s, an ISO of 800 and an approximate distance of 35-40 cm, and we captured a color standard *in situ* before each photography session (ColorChecker Passport Photo 2, X-rite, USA). Photos were captured in RAW format, then color corrected in Adobe Lightroom (Adobe, USA), see *Supplemental methods*.

Photaria were placed on a pedestal covered with matte black felt, inside a 30 x 60 x 30 cm photo box of white plastic. The walls and top of the box were constructed of opaque white plastic, to minimize the influence of external lights. In the top, directly above the pedestal, we mounted 2 Neewer NL660 LED panels with white diffusers, with a color rendering index of >= 96 (Neewer, China).

We photographed each side of every fish, and for 99.9% of fish we had at least three photos, with the replicate photos allowing us to parse measurement error from asymmetry. We analyzed a total of 14,100 photos, or an average of 4.37 photos per fish.

## Color pattern extraction and quantification

After color correction, we used a custom pipeline to analyze the images. The pipeline consists of deep neural networks for image segmentation and landmark placement, and geometric morphometrics for image alignment. To be able to phenotype fish during the selection procedure, when we still had little data available to train the models, we adopted a semi-autonomous strategy. We used an interactive script to phenotype fish, which prompted the experimenter (WvdB) to visually check the output of each analysis step, and correct manually when the model prediction was incorrect. We then intermittently updated model weights with the growing number of manually annotated examples to improve model performance, gradually reducing the need for manual intervention throughout the experiment. The pipeline consisted of these five steps:

Step 1: We first performed image segmentation to separate the fish from the background. We excluded the gonopodium and dorsal fin from the extracted fish, since their size is not consistent on camera, they are difficult to segment due to their transparency, and they do not exhibit ornamentation in our population. We used a convolutional neural net with a Unet (*46*) architecture, implemented in the *unet* R package (*47*), with six layers in the encoder, 10 filters in the first convolution, a dropout of 0.25 and with binary cross-entropy as the loss function. Using the segmentation mask, we then cropped and rotated the fish to an image of dimension [800 x 300 x 4], where the fourth channel is the transparency mask.

Step 2: Next, we further segmented the fish image with two additional Unet models, to extract the carotenoid and melanic coloration respectively. Both used four layers in the encoder, 12 filters in the first convolution, a dropout of 0.5 and a Dice loss function.

Step 3: To be able to directly compare regions of the guppy body, we performed image alignment. We trained a convolutional neural net to place four landmarks on the fish: in the center of the eye, on the tip of the snout, and on the dorsal and ventral extremes of the boundary between the caudal peduncle and the tail fin. Instead of predicting the landmark coordinates directly, which worked poorly, the landmarks were represented as a stack of four images, each having a point (with Gaussian blur) at one landmark location. The model was then trained to predict these landmark images, taking the extracted fish from Step 1 as input. The model first has three convolutional layers, followed by three deconvolutional layers, to transform the three color channels into four heatmaps of landmark probability (Table S1). We then chose the highest activation on each output channel as the landmark coordinates.

Step 4: Treating the four landmarks from Step 3 as fixed, we then added sliding semi-landmarks algorithmically. In short, we first mirrored all images of the left side of the fish. We then found the boundary of the segmentation mask from Step 1, which is the outline of the fish. These pixels were ordered into a continuous connected curve by treating the connection between points as a Travelling Salesman Problem. Using the fixed landmarks from Step 3, we split this curve into three: a) the outline from the tip of the snout, along the back, to the dorsal attachment of the caudal fin, b) the outline of the caudal fin, dorsal to ventral, and c) from the ventral attachment of the caudal fin to the tip of the snout. Along each of these curves we placed 50 evenly spaced semi-landmarks.

Step 5: Using the fixed landmarks from Step 3 and sliding landmarks from Step 4, we use Generalized Procrustes analysis to generate a consensus shape using package *geomorph* (*48*). Using thin-plate spline transformation, we then warped our fish images to the consensus shape with package *Morpho* (*49*). This warping is applied to the segmented fish from Step 1, and the segmented ornamentation from Step 2. Since these are continuous transformations, we then rasterized the warped output by taking color averages along a grid, resulting in images of dimension [500 x 140 x 4].

By comparing the number of pixels in the extracted fish mask from Step 1, and the number of pixels in the extracted carotenoid and melanin masks from Step 2, we calculate the amount of coloration as a percentage of the body area. For use in our selection paradigm, we calculated the mean across multiple photos of each side of each fish, and then the mean across the sides.

For manual segmentation for Step 1, we annotated the fish' outline with splines using the *StereoMorph* package (*50*). For manual color segmentation in Step 2, we used the free and fuzzy select tools in the GNU Image Manipulation Program (GIMP). For melanin segmentation, we only selected distinct color patches of concentrated pigmentation, which is always present. We excluded the various facultatively expressed darker areas, sometimes referred to as "fuzzy black".

### Ornaments

We defined areas of the fish as ornaments manually, based on the visible spatial structure in the incidence of color and the localized genetic architecture. We then overlayed these areas on the extracted color patterns from each photo, and counted the number of pigmented pixels within each area. We scored an ornament as "absent" (ornament size = 0) if < 10% of the ornament area was pigmented. This limit was chosen based on the bimodal distribution of the number of pixels, as in some cases a few pixels from a neighboring ornament spilled over into other areas.

We determined the distribution of each male fish in pattern space using using triplet learning (*21*), where we trained convolutional neural nets (see Supplementary Methods).

## Male sexual behavior

We assessed the sexual behavior of male offspring from the F3 generation, using 126 males (63 from each selection direction, evenly split across the replicate lines). We isolated males upon early indications of sexual maturity and at around three months, tested their sexual behaviours using a no choice test with nonreceptive females, a standard for studying male sexual behaviour in poecilids. Behavioural trials took place in a 47 cm diameter arena, with fish transfer followed by a 5-minute acclimation to minimize stress and a 15-minute observation using a webcam at 1080p and 30 fps. We analysed the recorded videos to quantify male sexual behaviours. We utilized idTracker (*51*) for positional data during each trial. Trials where females exhibited receptivity or where fish showed signs of stress were excluded, leading to a final dataset of 95 individuals. Statistical analysis on behavioural differences between high and low orange males used generalized mixed models, with behavioural metrics as dependent variables, and selection direction, male body size, and female body size as fixed effects, and replicate line as random effect. Model adequacy was verified through residual plots and parameters were evaluated via Wald tests.

## Quantitative genetics

Estimating of quantitate genetic (QG) parameters was performed by MCMC sampling of animal models implemented in *Stan* (*52*) with *brms* (*53*, *54*). Additive genetic variance-covariance matrices were calculated from the pedigree using *nadiv* (*55*), with the *S* matrix for X-linked effects assuming no global dosage compensation (*45*). Y-linked genetic variances were estimated by including patriline as a group level effect. Similarly, we included dam identity to estimate maternal effects, and included brood identity to capture shared environmental effects between siblings, both as group level effects. We used default non-informative priors for all parameters.

For the QG heatmaps (Fig.3), we estimated separate animal models for each bin. Since estimating these parameters independently thousands of times using MCMC is prohibitively computationally expensive, we used an approach that makes use of the fact that the QG estimations will be similar between neighboring bins. To limit the number of models, we reduced the dimensions of the color pattern images by a factor of two (250 x 70 bins or pixels), and only ran models for bins with an incidence of at least 1%. We then performed model fitting in two rounds. First, we fit animal models for the bins with x- and y-coordinates divisible by five. For these models, we used four chains, with 500 warmup iterations

and 1500 sampling iterations per chain, and uninformative StudentT(3, 0, 5) priors on the intercept and variance components. In a second step, we then used the no output of these models to fill in the bins between this five-by-five grid. These models only used two chains of 500 warmup iterations and 500 sampling iterations, but were seeded with initial parameter values and step sizes that were the average of the closest nine models from Step 1, weighted by their inverse distance. Similarly, we set normal priors for the intercept and variance components with the mean (also weighted by inverse distance) and standard deviation of those parameters of the closest nine models.

For the pairwise correlations across ornaments, we used the Gaussian approximation and fitted the multivariate animal model on individual-level means using the restricted maximum likelihood method from *sommer* (*56*). Given the small magnitudes of maternal and brood effects from the full Bayesian animal models for each ornament, we did not include these effects here.

## Sampling and whole genome sequencing

To maximize power for the genetic mapping of color traits, we sampled 50 fish from each selection line (N = 300) for whole genome sequencing (WGS). To maximize genetic and phenotypic variation in our sample, and therefore power, we sampled males from each family with offspring in generation $F_3$, and sampled males with the most/least carotenoid coloration in up/down-selected lines respectively. This meant we sampled one son from each family, with some families contributing a second son until we reached the desired 50 males.

We sacrificed males in ice water, and stored their heads in 95% ethanol. DNA was extracted from half a head using Qiagen DNeasy blood & tissue kits (Qiagen, USA). Samples were then sequenced on four lanes of Illumina NovaSeq 6000 S4 PE150, to an average depth of 10x per individual. One library failed, and two individuals were excluded because their genomic relatedness did not match the pedigree, leaving 297 samples.

## Genome-wide association analysis

We first trimmed the adapters using *Trimmomatic*, and then mapped the reads to the reference using *bwa mem* using default mapping parameters. We marked PCR duplicates using GATK's *MarkDuplicates*. We then called variants using *bcftools mpileup* where we downgraded the mapping quality of reads containing mismatches with *–adjust-MQ 60*, and set a minimal mapping quality using *–min-MQ 30*, followed by the consensus caller *bcftools call*. We then performed variant filtering, and dropped sites with a QUAL score < 30, and sites where more than 20% of individuals had a depth < 5, or > 20% of

individuals had a GQ score < 10. Additionally, we filtered out sites with a major allele frequency < 10% or a missingness > 5%.

We performed multivariate GWAS by running GEMMA (*57*) on each of the 5 axes of pattern space, using the centered genomic relatedness matrix calculated from the SNPs mapping to the autosomes. We then calculated Z-scores and combined them to obtain multivariate p-values using CPASSOC (*58*). Coverage statistics were calculated using *samtools depth*.

The origin of SNPs at autosomal peaks with elevated coverage was estimated using animal models. We encoded the genotypes as doses of the alternative allele, and then fitted three models with the dose as the dependent variable. The three models included a random effect for the autosomal, X-linked or Y-linked covariance, respectively, as described in *Quantitive genetics*. For each model we then calculated Akaike's Information Criterion (AIC) and their AIC weight to compare between the three hypotheses.

## Software

Unless stated otherwise above, all analyses were performed in the R statistical environment (*59*).

## References


1. J. F. Stephenson, M. Stevens, J. Troscianko, J. Jokela, The Size, Symmetry, and Color Saturation of a Male Guppy's Ornaments Forecast His Resistance to Parasites. *Am. Nat.* **196**, 597–608 (2020).

2. T. Potter, J. Arendt, R. D. Bassar, B. Watson, P. Bentzen, J. Travis, D. N. Reznick, Female preference for rare males is maintained by indirect selection in Trinidadian guppies. *Science*. **380**, 309–312 (2023).

3. A. E. Houde, Mate choice based upon naturally occurring color-pattern variation in a guppy population. *Evolution*. **41**, 1–10 (1987).

4. J. A. Endler, Natural Selection on Color Patterns in Poecilia reticulata. *Evolution*. **34**, 76–91 (1980).

5. A. E. Houde, J. A. Endler, Correlated Evolution of Female Mating Preferences and Male Color Patterns in the Guppy Poecilia reticulata. *Science*. **248**, 1405–1408 (1990).

6. D. J. Kemp, D. N. Reznick, G. F. Grether, J. A. Endler, Predicting the direction of ornament evolution in Trinidadian guppies (Poecilia reticulata). *Proc. R. Soc. B Biol. Sci.* **276**, 4335–4343 (2009).

7. M. Herdegen-Radwan, S. Cattelan, J. Buda, J. Raubic, J. Radwan, What do orange spots reveal about male (and female) guppies? A test using correlated responses to selection. *Evolution*. **75**, 3037–3055 (2021).

8. R. Brooks, J. A. Endler, Direct and indirect sexual selection and quantitative genetics of male traits in guppies (Poecilia reticulata). *Evolution*. **55**, 1002–1015 (2001).

9. J. Morris, I. Darolti, W. van der Bijl, J. E. Mank, High-resolution characterization of male ornamentation and re-evaluation of sex linkage in guppies. *Proc. R. Soc. B Biol. Sci.* **287**, 20201677 (2020).

10. K. A. Hughes, A. E. Houde, A. C. Price, F. H. Rodd, Mating advantage for rare males in wild guppy populations. *Nature*. **503**, 108–110 (2013).

11. J. A. Farr, Male Rarity or Novelty, Female Choice Behavior, and Sexual Selection in the Guppy, Poecilia reticulata Peters (Pisces: Poeciliidae). *Evolution*. **31**, 162–168 (1977).

12. D. M. Parichy, J. E. Spiewak, Origins of adult pigmentation: diversity in pigment stem cell lineages and implications for pattern evolution. *Pigment Cell Melanoma Res.* **28**, 31–50 (2015).

13. M. B. Toomey, C. I. Marques, P. M. Araújo, D. Huang, S. Zhong, Y. Liu, G. D. Schreiner, C. A. Myers, P. Pereira, S. Afonso, P. Andrade, M. A. Gazda, R. J. Lopes, I. Viegas, R. E. Koch, M. E. Haynes, D. J. Smith, Y. Ogawa, D. Murphy, R. E. Kopec, D. M. Parichy, M. Carneiro, J. C. Corbo, A mechanism for red coloration in vertebrates. *Curr. Biol.* **32**, 4201-4214.e12 (2022).

14. Ø. Winge, One-sided masculine and sex-linked inheritance inLebistes reticulatus. *J. Genet.* **12**, 145–162 (1922).

15. Ø. Winge, E. Ditlevsen, Colour inheritance and sex determination in Lebistes. *Heredity*. **1**, 65–83 (1947).

16. C. P. Haskins, P. Young, R. E. Hewitt, E. F. Haskins, Stabilised heterozygosis of supergenes mediating certain Y-linked colour patterns in populations of Lebistes Reticulatus. *Heredity*. **25**, 575–589 (1970).

17. L. E. B. Kruuk, Estimating genetic parameters in natural populations using the 'animal model.' *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **359**, 873–890 (2004).

18. N. R. Liley, Ethological Isolating Mechanisms in Four Sympatric Species of Poeciliid Fishes. *Behav. Suppl.* **1**, 1–197 (1966).

19. A. Houde, *Sex, Color, and Mate Choice in Guppies* (Princeton University Press, 1997).

20. R. Brooks, J. A. Endler, Female Guppies Agree to Differ: Phenotypic and Genetic Variation in Mate-Choice Behavior and the Consequences for Sexual Selection. *Evolution*. **55**, 1644–1655 (2001).

21. F. Schroff, D. Kalenichenko, J. Philbin, "FaceNet: A unified embedding for face recognition and clustering" in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2015), pp. 815–823.

22. A. Künstner, M. Hoffmann, B. A. Fraser, V. A. Kottler, E. Sharma, D. Weigel, C. Dreyer, The Genome of the Trinidadian Guppy, Poecilia reticulata, and Variation in the Guanapo Population. *PLOS ONE*. **11**, e0169087 (2016).

23. P. Almeida, B. A. Sandkam, J. Morris, I. Darolti, F. Breden, J. E. Mank, Divergence and Remarkable Diversity of the Y Chromosome in Guppies. *Mol. Biol. Evol.* **38**, 619–633 (2021).

24. A. E. Wright, I. Darolti, N. I. Bloch, V. Oostra, B. Sandkam, S. D. Buechel, N. Kolm, F. Breden, B. Vicoso, J. E. Mank, Convergent recombination suppression suggests role of sexual selection in guppy sex chromosome formation. *Nat. Commun.* **8**, 14251 (2017).

25. Y. Lin, I. Darolti, B. L. S. Furman, P. Almeida, B. A. Sandkam, F. Breden, A. E. Wright, J. E. Mank, Gene duplication to the Y chromosome in Trinidadian Guppies. *Mol. Ecol.* **31**, 1853–1863 (2022).

26. B. A. Sandkam, P. Almeida, I. Darolti, B. L. S. Furman, W. van der Bijl, J. Morris, G. R. Bourne, F. Breden, J. E. Mank, Extreme Y chromosome polymorphism corresponds to five male reproductive morphs of a freshwater fish. *Nat. Ecol. Evol.* **5**, 939–948 (2021).

27. J. A. Endler, "Natural and sexual selection on color patterns in poeciliid fishes" in *Evolutionary ecology of neotropical freswater fishes* (Springer, The Netherlands, 1984; https://link.springer.com/chapter/10.1007/978-94-015-7682-6_7), *Developments in Environmental Biology of Fishes*, pp. 95–111.

28. A. Kotrschal, B. Rogell, A. Bundsen, B. Svensson, S. Zajitschek, I. Brännström, S. Immler, A. A. Maklakov, N. Kolm, Artificial Selection on Relative Brain Size in the Guppy Reveals Costs and Benefits of Evolving a Larger Brain. *Curr. Biol.* **23**, 168–171 (2013).

29. S. Fong, B. Rogell, M. Amcoff, A. Kotrschal, W. van der Bijl, S. D. Buechel, N. Kolm, Rapid mosaic brain evolution under artificial selection for relative telencephalon size in the guppy (Poecilia reticulata). *Sci. Adv.* **7**, eabj4314 (2021).

30. A. Kotrschal, A. Szorkovszky, J. Herbert-Read, N. I. Bloch, M. Romenskyy, S. D. Buechel, A. F. Eslava, L. S. Alòs, H. Zeng, A. Le Foll, G. Braux, K. Pelckmans, J. E. Mank, D. Sumpter, N. Kolm, Rapid evolution of coordinated and collective movement in response to artificial selection. *Sci. Adv.* **6**, eaba3148 (2020).

31. A. E. Houde, Effect of artificial selection on male colour patterns on mating preference of female guppies. *Proc. R. Soc. Lond. B Biol. Sci.* **256**, 125–130 (1994).

32. G. L. Cole, J. A. Endler, Change in male coloration associated with artificial selection on foraging colour preference. *J. Evol. Biol.* **31**, 1227–1238 (2018).

33. M. Inaba, H. Yamanaka, S. Kondo, Pigment Pattern Formation by Contact-Dependent Depolarization. *Science*. **335**, 677–677 (2012).

34. V. A. Kottler, A. Fadeev, D. Weigel, C. Dreyer, Pigment Pattern Formation in the Guppy, Poecilia reticulata, Involves the Kita and Csf1ra Receptor Tyrosine Kinases. *Genetics*. **194**, 631–646 (2013).

35. D. M. Parichy, D. G. Ransom, B. Paw, L. I. Zon, S. L. Johnson, An orthologue of the kit-related gene fms is required for development of neural crest-derived xanthophores and a subpopulation of adult melanocytes in the zebrafish, Danio rerio. *Development*. **127**, 3031–3044 (2000).

36. V. A. Kottler, A. Künstner, I. Koch, M. Flötenmeyer, T. Langenecker, M. Hoffmann, E. Sharma, D. Weigel, C. Dreyer, Adenylate cyclase 5 is required for melanophore and male pattern development in the guppy (Poecilia reticulata). *Pigment Cell Melanoma Res.* **28**, 545–558 (2015).

37. A. H. Brush, Metabolism of carotenoid pigments in birds. *FASEB J.* **4**, 2969–2977 (1990).

38. G. E. Hill, J. D. Johnson, The Vitamin A–Redox Hypothesis: A Biochemical Basis for Honest Signaling via Carotenoid Pigmentation. *Am. Nat.* **180**, E127–E150 (2012).

39. G. F. Grether, J. Hudon, D. F. Millie, Carotenoid limitation of sexual coloration along an environmental gradient in guppies. *Proc. R. Soc. Lond. B Biol. Sci.* **266**, 1317–1322 (1999).

40. R. J. Lopes, J. D. Johnson, M. B. Toomey, M. S. Ferreira, P. M. Araujo, J. Melo-Ferreira, L. Andersson, G. E. Hill, J. C. Corbo, M. Carneiro, Genetic Basis for Red Coloration in Birds. *Curr. Biol.* **26**, 1427–1434 (2016).

41. N. I. Mundy, J. Stapley, C. Bennison, R. Tucker, H. Twyman, K.-W. Kim, T. Burke, T. R. Birkhead, S. Andersson, J. Slate, Red Carotenoid Coloration in the Zebra Finch Is Controlled by a Cytochrome P450 Gene Cluster. *Curr. Biol.* **26**, 1435–1440 (2016).

42. C. A. McLean, A. Lutz, K. J. Rankin, A. Elliott, A. Moussalli, D. Stuart-Fox, Red carotenoids and associated gene expression explain colour variation in frillneck lizards. *Proc. R. Soc. B Biol. Sci.* **286**, 20191172 (2019).

43. S. P. Gordon, A. López-Sepulcre, D. Rumbo, D. N. Reznick, Rapid Changes in the Sex Linkage of Male Coloration in Introduced Guppy Populations. *Am. Nat.* **189**, 196–200 (2017).

44. S. P. Gordon, A. López-Sepulcre, D. N. Reznick, PREDATION-ASSOCIATED DIFFERENCES IN SEX LINKAGE OF WILD GUPPY COLORATION. *Evolution*. **66**, 912–918 (2012).

45. I. Darolti, A. E. Wright, B. A. Sandkam, J. Morris, N. I. Bloch, M. Farré, R. C. Fuller, G. R. Bourne, D. M. Larkin, F. Breden, J. E. Mank, Extreme heterogeneity in sex chromosome differentiation and dosage compensation in livebearers. *Proc. Natl. Acad. Sci.* **116**, 19031–19036 (2019).

46. O. Ronneberger, P. Fischer, T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation" in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, A. F. Frangi, Eds. (Springer International Publishing, Cham, 2015), *Lecture Notes in Computer Science*, pp. 234–241.

47. D. Falbel, K. Zak, unet: U-Net: Convolutional Networks for Biomedical Image Segmentation (2023), (available at https://github.com/r-tensorflow/unet).

48. E. K. Baken, M. L. Collyer, A. Kaliontzopoulou, D. C. Adams, geomorph v4.0 and gmShiny: Enhanced analytics and a new graphical interface for a comprehensive morphometric experience. *Methods Ecol. Evol.* **12**, 2355–2363 (2021).

49. S. Schlager, "Morpho and Rvcg – Shape Analysis in R" in *Statistical Shape and Deformation Analysis*, G. Zheng, S. Li, G. Szekely, Eds. (Academic Press, 2017), pp. 217–256.

50. A. M. Olsen, M. W. Westneat, StereoMorph: an R package for the collection of 3D landmarks and curves using a stereo camera set-up. *Methods Ecol. Evol.* **6**, 351–356 (2015).

51. A. Pérez-Escudero, J. Vicente-Page, R. C. Hinz, S. Arganda, G. G. de Polavieja, idTracker: tracking individuals in a group by automatic identification of unmarked animals. *Nat. Methods*. **11**, 743–748 (2014).

52. B. Carpenter, A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, A. Riddel, Stan: A Probabilistic Programming Language. *J. Stat. Softw.* **76** (2017), doi:10.18637/jss.v076.i01.

53. P.-C. Bürkner, Advanced Bayesian Multilevel Modeling with the R Package brms. *R J.* **10**, 395–411 (2018).

54. P.-C. Bürkner, brms: An R Package for Bayesian Multilevel Models Using Stan. *J. Stat. Softw.* **80**, 1–28 (2017).

55. M. E. Wolak, nadiv: an R package to create relatedness matrices for estimating non-additive genetic variances in animal models. *Methods Ecol. Evol.* **3**, 792–796 (2012).

56. C.-P. Giovanny, Genome assisted prediction of quantitative traits using the R package sommer. *PLoS ONE*. **11**, 1–15 (2016).

57. X. Zhou, M. Stephens, Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **44**, 821–824 (2012).

58. X. Li, X. Zhu, "Cross-Phenotype Association Analysis Using Summary Statistics from GWAS" in *Statistical Human Genetics: Methods and Protocols*, R. C. Elston, Ed. (Springer, New York, NY, 2017; https://doi.org/10.1007/978-1-4939-7274-6_22), *Methods in Molecular Biology*, pp. 455–467.

59. R Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria, 2023; https://www.R-project.org/).

60. L. Yong, D. P. Croft, J. Troscianko, I. W. Ramnarine, A. J. Wilson, Sensory-based quantification of male colour patterns in Trinidadian guppies reveals no support for parallel phenotypic evolution in multivariate trait space. *Mol. Ecol.* **31**, 1337–1357 (2022).

## Supplemental methods

### *Learning of pattern space*

We devised a deep learning approach to map images of guppies onto a "pattern space". Our objective was to capture the heritable components of color patterns by placing images of related guppies more proximately in this pattern space compared to those of unrelated males. Our method is predicated on the "triplet loss learning" paradigm (*21*).

The deep learning architecture was designed to process images of dimensions 250x70x3, with the last dimension being color channels. Our model, a Convolutional Neural Network (CNN), had a series of layers that successively compress the images to lower dimensions. Specifically, we used separable convolutional layers with a kernel size of 3x3 and relu activation, followed by a max pooling layer, repeating this four times. Each convolutional layer had double the filters of the previous, starting at 8 filters. After the fourth pooling layer, we flattened the output to a dense layer of size 8, and then a second dense layer of the embedding dimension. Finally, since we wanted to achieve a dense representation, we did not use L1 normalization, but a batch normalization layer. Embeddings are obtained at this normalization layer.

The essence of our method hinges on the generation and use of "triplets". A triplet comprises an anchor image (a random male), a positive image (of a related male), and a negative image (of an unrelated male). Unrelated males had a coefficient of relatedness ($r$) of 0, while for related males $r$ was at least 1/8. During training, each image is passed through the CNN described above and the triplet loss function is applied. The objective was to reduce the distance (in the embedding space) between the anchor and positive images, while increasing the distance between the anchor and negative images, governed by a margin.

We trained for 50 epochs and a batch size of 128 images. Each epoch, we randomly generate 5,000 triplets. We randomly allocated 20% of individuals to the validation set, from which we generated 500 triplets per epoch to track model performance. Moreover, learning rate schedules were implemented to dynamically adjust the learning rate across epochs, with specified rates of 1e-4, 1e-5, and 1e-6 at epochs 0, 30, and 40 respectively. This scheduling ensured a faster convergence initially, with fine-tuning during the latter stages of training.

We chose the number of embedding dimensions by training the model at one to eight dimensions, then chose five dimensions, since the model accuracy for the validation sample stopped increasing above five. After training, we used the CNN to embed all 14,100 images.

### *Color calibration*

Fish were photographed while swimming in a photarium. For the P and $F_1$ generations, we used a glass photarium measuring 16 x 85 x 100 mm (internal length x width x height), with an approximate water level of 3 cm. For generations $F_2$ and $F_3$, we instead used a thin plastic 54 mm cuboid box with an internal sliding white wall. In this photarium (following (*60*)) the fish can be restricted to a very narrow area in the front by sliding the white wall forward. Since the change in photarium led to small changes in color characteristics, we performed color calibration to match the two setups. To this end, we photographed 15 fish with both methods, and extracted the fish from the images. We then generated look-up-tables (LUTs) by matching color quantiles from these images in Lab color space. Photographs from the new photorarium were then adjusted to match the old color profile using the LUTs prior to analysis.

Supplemental figures

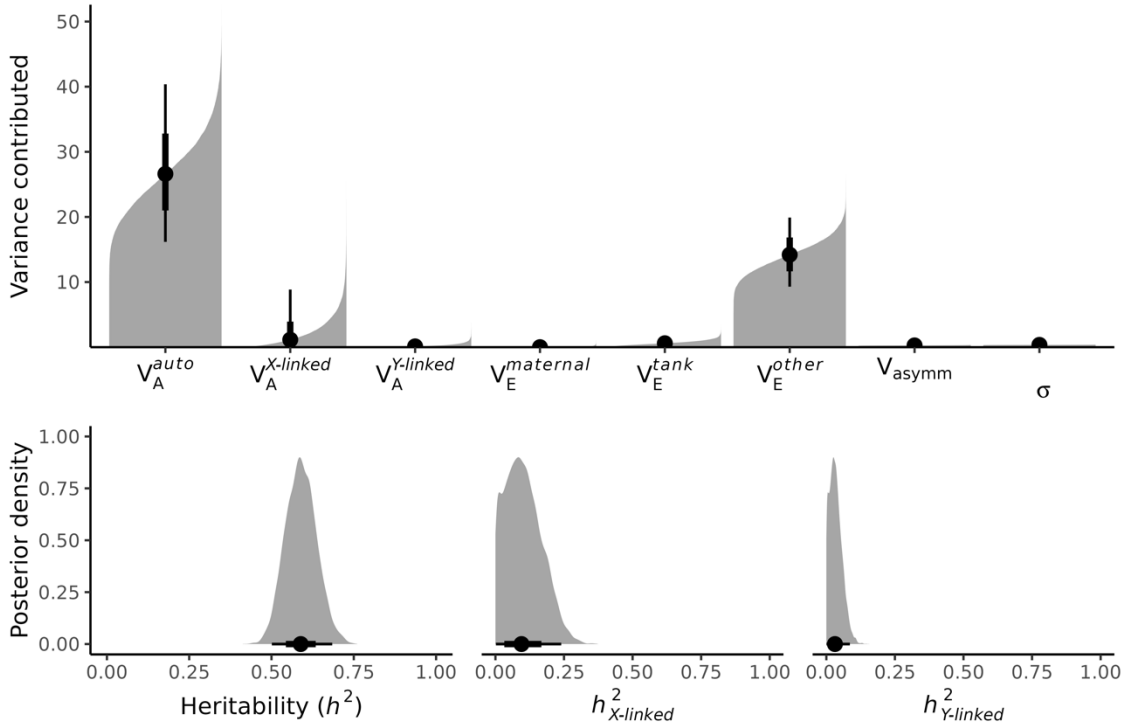


**Figure S1:** Variance components of total amount of orange color, as estimated by a Bayesian animal model. Top panel shows cumulative posterior density bargraphs, bottom panel shows posterior ditributions. Posterior summaries are shown as posterior means (point) and 66% (thick lines) and 95% (thin lines) credible intervals.

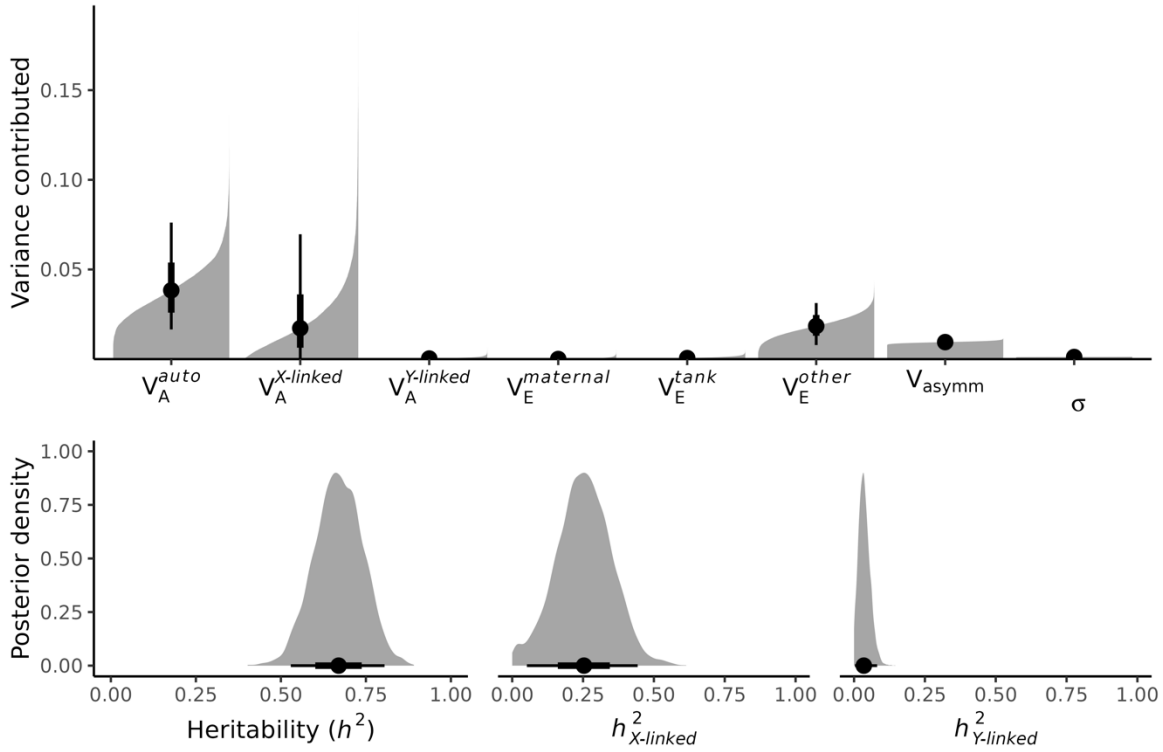


**Figure S2:** Variance components of total amount of black color, as estimated by a Bayesian animal model. Top panel shows cumulative posterior density bargraphs, bottom panel shows posterior ditributions. Posterior summaries are shown as posterior means (point) and 66% (thick lines) and 95% (thin lines) credible intervals.

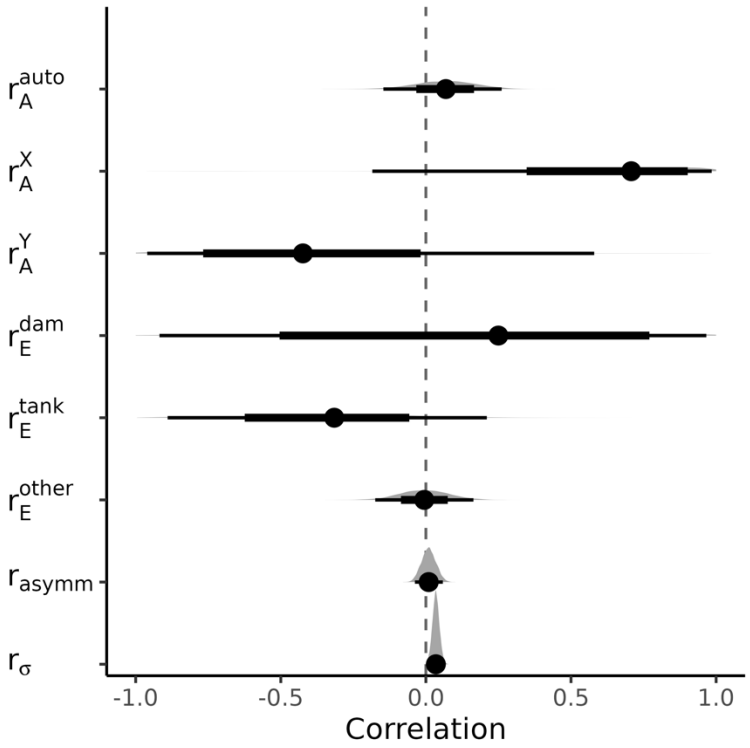


**Figure S3:** Posterior distributions of the correlation of orange and black variance components (Figs S1&S2). Posterior summaries are shown as posterior means (point) and 66% (thick lines) and 95% (thin lines) credible intervals.

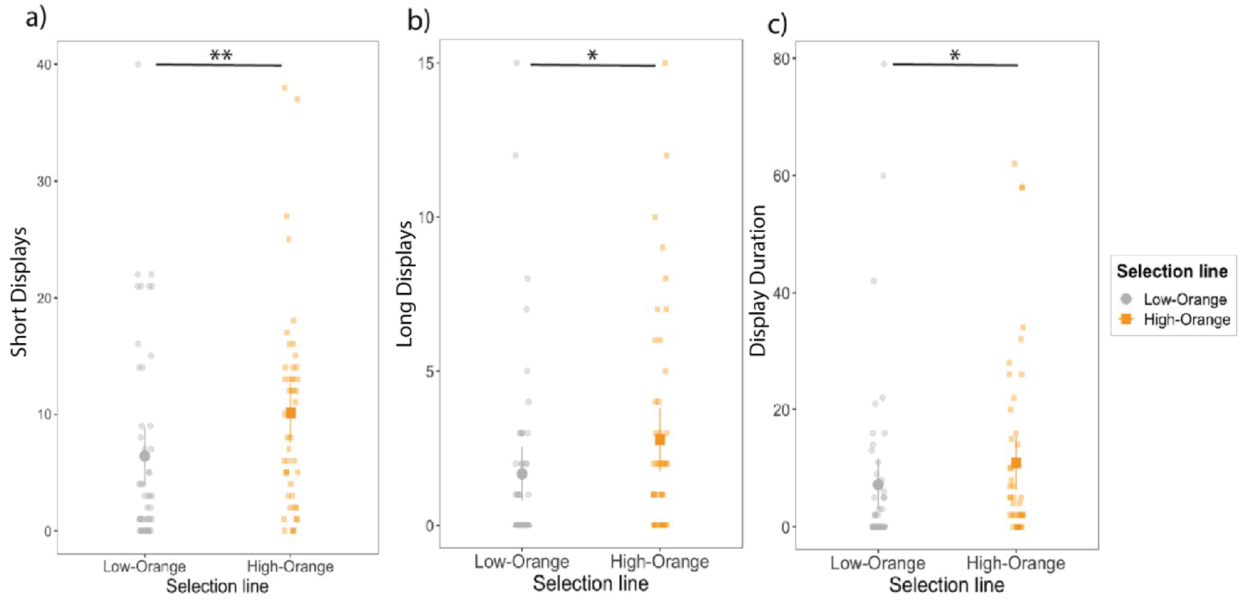


**Fig S4:** The effect of artificial selection for orange colouration on display sexual behaviour of male guppies. Comparison of a) frequency of short displays; b) frequency of long displays; and c) duration of display in male guppies following three generations of artificial selection for high (n = 49) and low (n = 46) proportion of orange in their body and tail. Low orange selected males in grey circles, high orange selected males with orange squares. Larger circles and squares indicate averages for each treatment and exposure time with 95% CI bars. All display sexual behaviours quantified showed significant differences between the two selection lines with high orange performing more. Stars indicate significance, with $p < 0.01$ **, $p < 0.05$ *

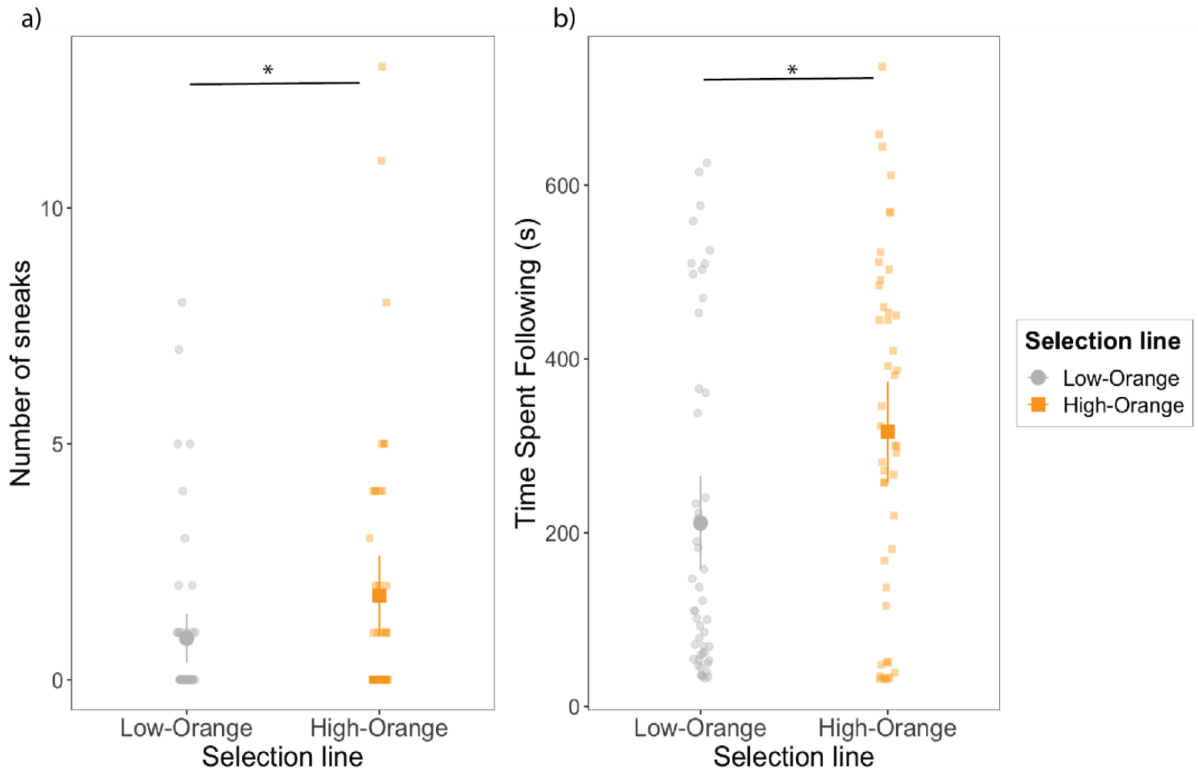


**Fig S5:** The effect of artificial selection for orange colouration on coercive sexual behaviours of male guppies. Comparison of a) frequency of sneak attempts and b) time spent following in male guppies following three generations of artificial selection for high (n = 49) and low (n = 46) proportion of orange in their body and tail. Low orange selected males in grey circles, high orange selected males with orange squares. Larger circles and squares indicate averages for each treatment and exposure time with 95% CI bars. All sexual behaviours presented a significant difference between the two selection lines with high orange performing more. A star indicates significance $p < 0.05$ *.

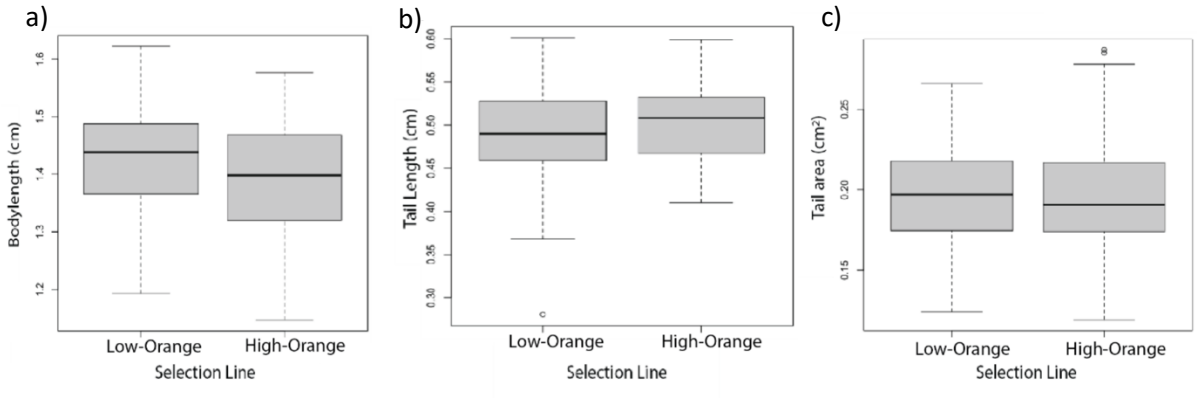


**Fig S6:** Morphological traits in guppies artificially selected for total area of orange. Comparison of a) standard length b) tail length and c) tail area in male guppies following three generations of artificial selection for high (n = 49) and low (n = 46) proportion of orange in their body and tail. Orange colouration was significantly different between individuals from these selection lines used for behavioural experiments (see Table 3.2). Horizontal lines indicate medians, boxes indicate the interquartile range, and whiskers indicate all points within 1.5 times the interquartile range.

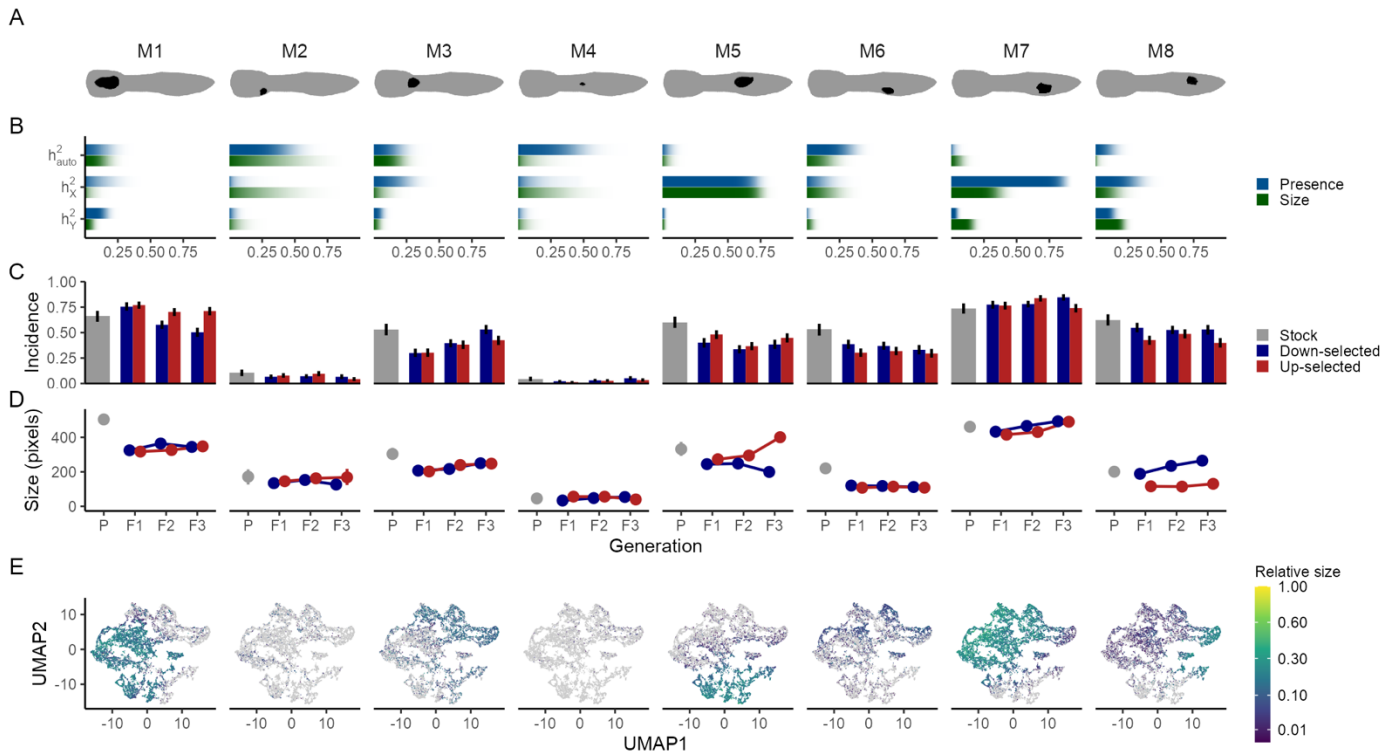


**Fig S7: Black ornaments are heritable but responded only weakly to selection on orange color**. A) Pictograms of eight black ornaments. B) Heritabilites of the incidence and size (when present) of each ornament. Dots and lines reflect point estimates and 95% credible intervals, and the gradient bars show the cumulative posterior density. C&D) Effect of selection on the incidence and size of black ornaments. Error bars show 95% bootstrapped confidence intervals, and points represent means. X-axis show consecutive generations. E) Ornaments in black pattern space. Points are the location of individuals in pattern space. Axes are the UMAP reduced representation of five-dimensional pattern space. Individuals are colored by the size of their relevant ornament, expressed as a fraction of the largest observed size, or colored grey if they lack the ornament. Note the use of a logarithmic scale for ornament size.

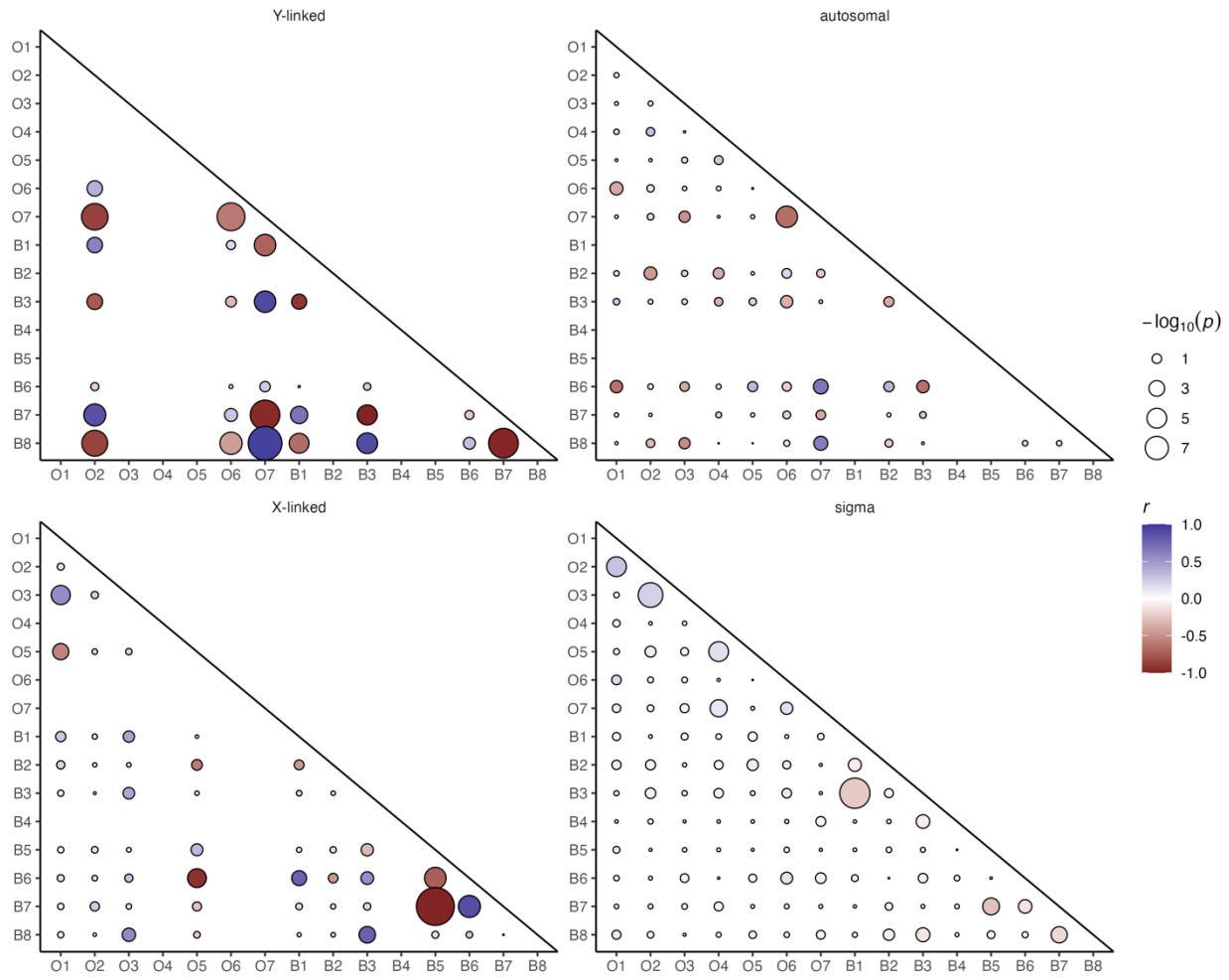


**Fig S8:** Pairwise correlations of the presence of ornaments, separated in the Y-linked, X-linked and autosomal components, and the environment (sigma). Correlations are only shown if the variance component was significantly greater than zero for both ornaments.

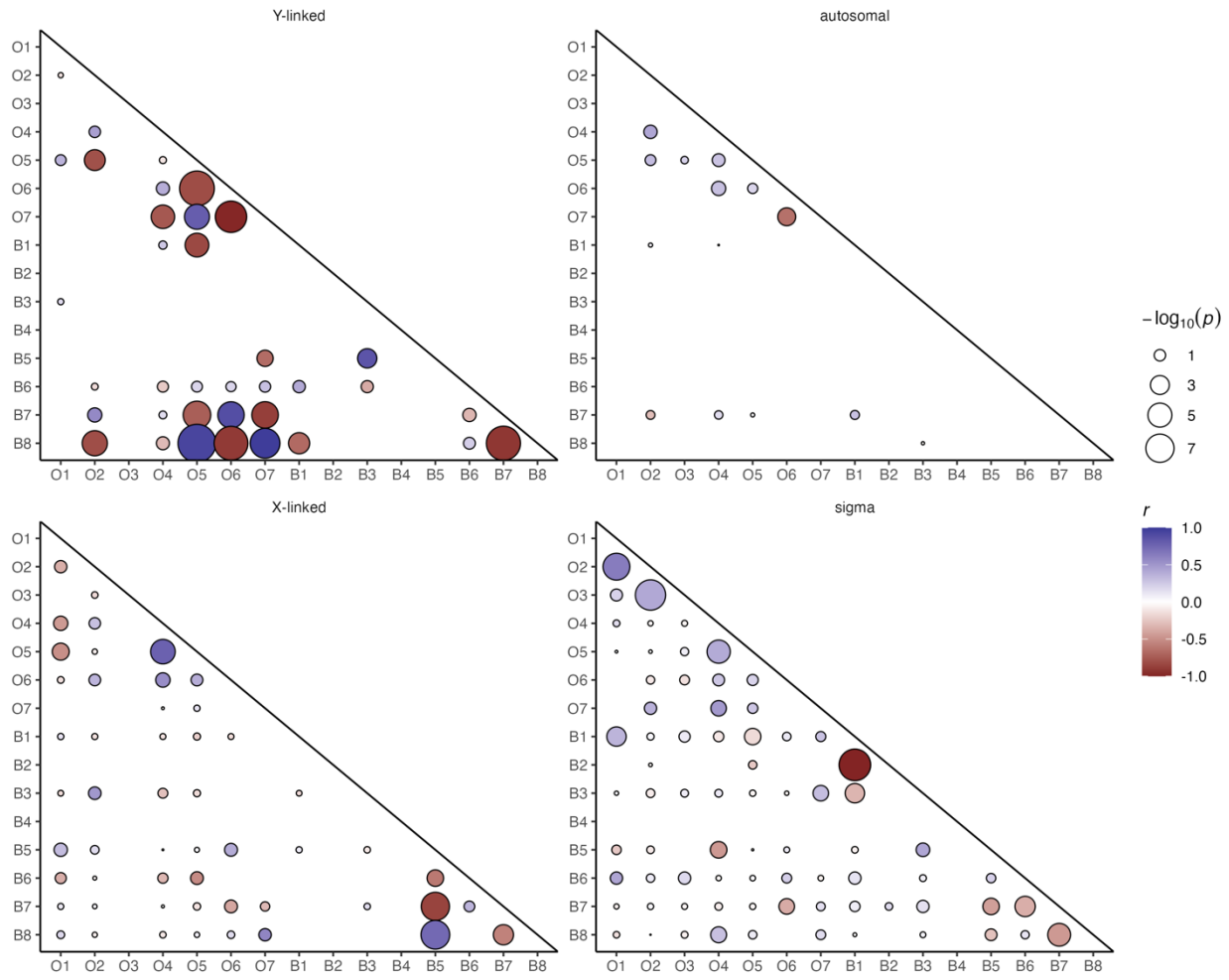


**Fig S8:** Pairwise correlations of the size of ornaments, separated in the Y-linked, X-linked and autosomal components, and the environment (sigma). Correlations are only shown if the variance component was significantly greater than zero for both ornaments. Models only included individuals with both ornaments.

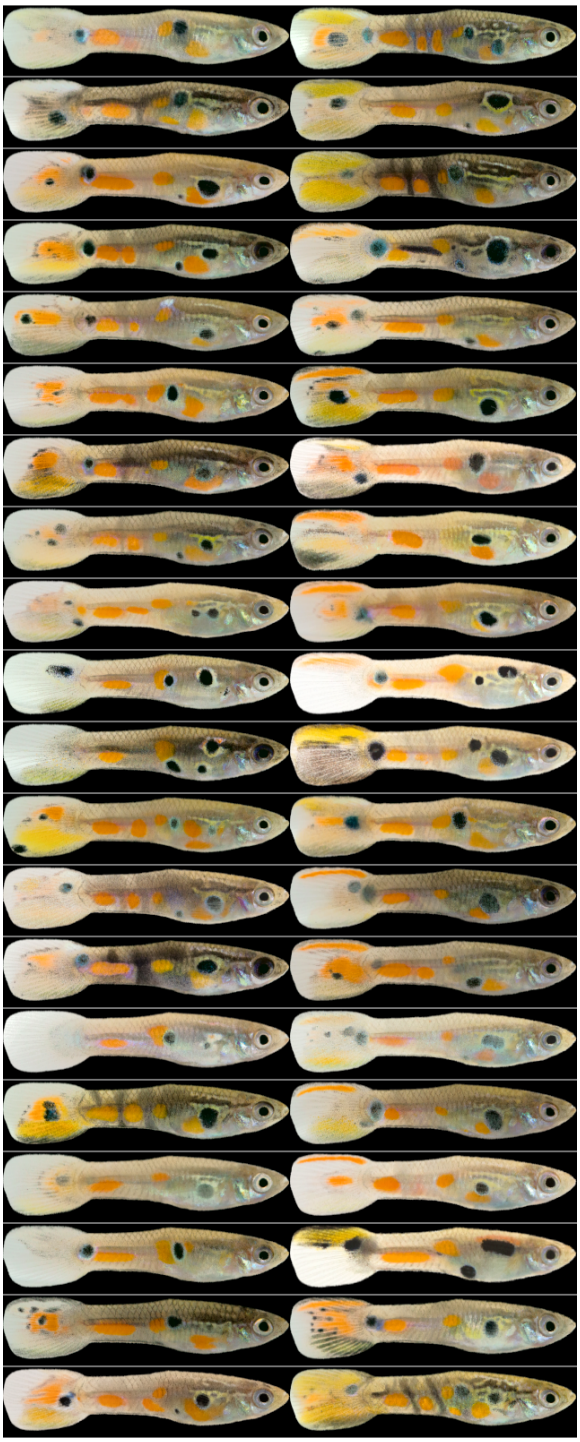

**Fig S10:** Ornament diversity in the tail fin. Displayed are randomly selected examples of fish with an absent (left) or present (right) ornament O1 (top of the tail fin). Note that the fish with a present ornament O1 have variety of tail ornamentation, possibly indicating spatially overlapping ornaments.

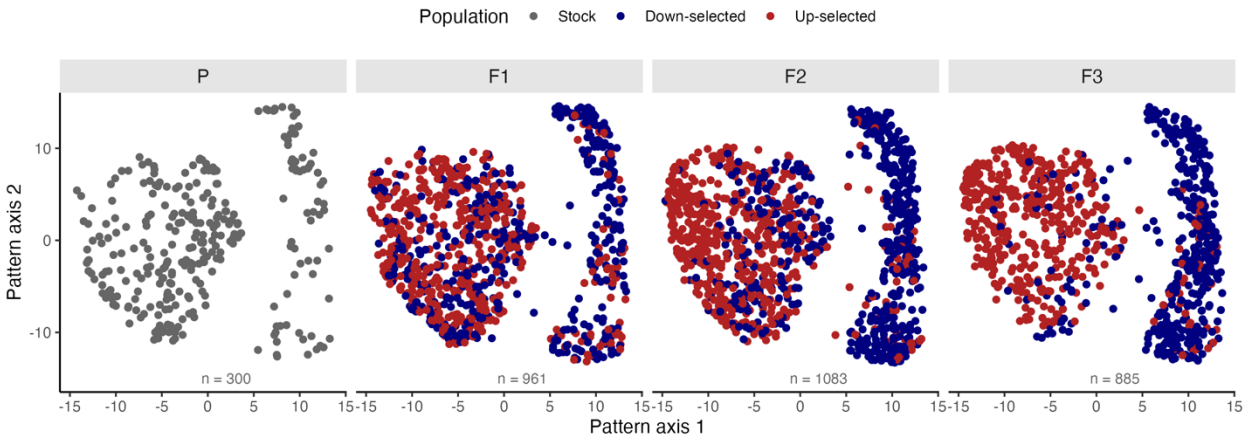


**Fig S11:** Selection moved populations through orange pattern space. Axes show the UMAP reduced representation of orange pattern space, with points indicating individuals, colored by selection direction, and panel showing the four generations.

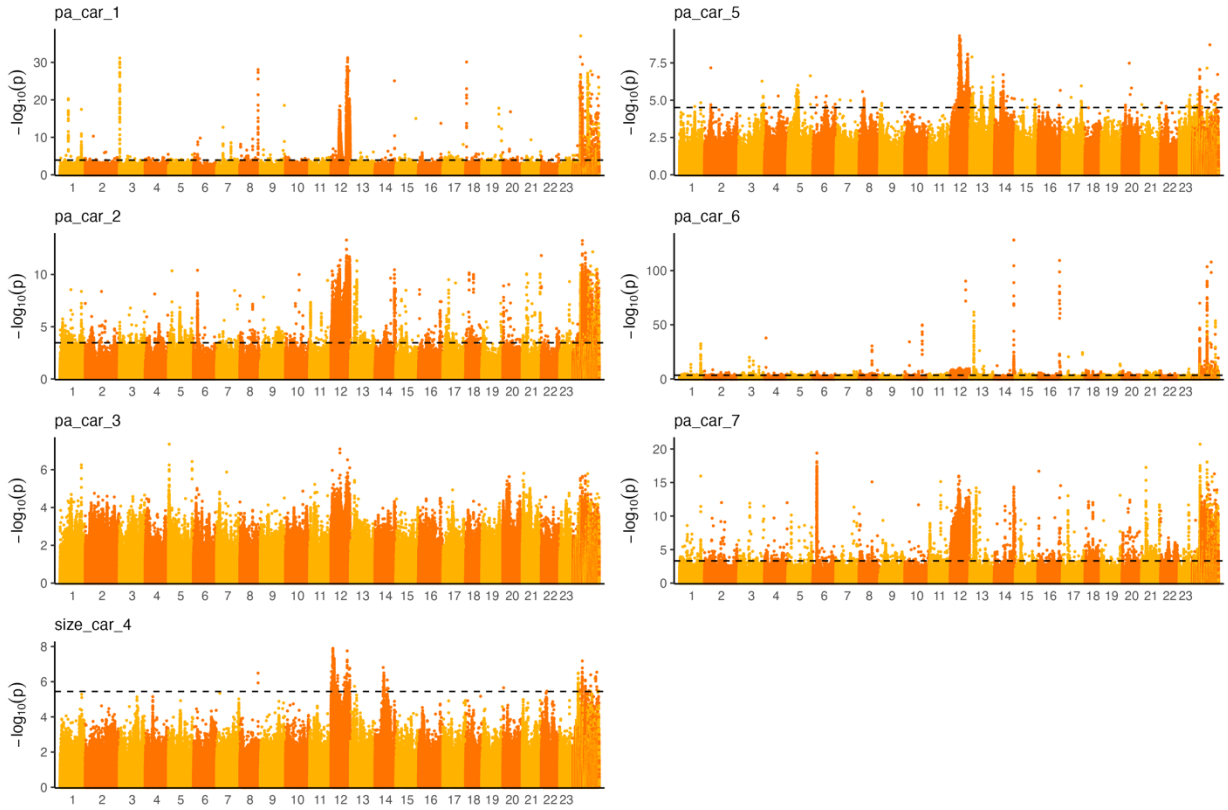


**Fig S12:** Post-hoc GWAS results for orange ornaments, as in Fig 5. Pa_car_x denotes the presence/absence of ornament Ox, while size_car_4 shows the GWAS for the size of ornament O4 (since

it is almost always present). Dotted lines show the cut-off for an FDR of 5%, but these are not corrected for testing across multiple phenotypes, and should not be used for inference.

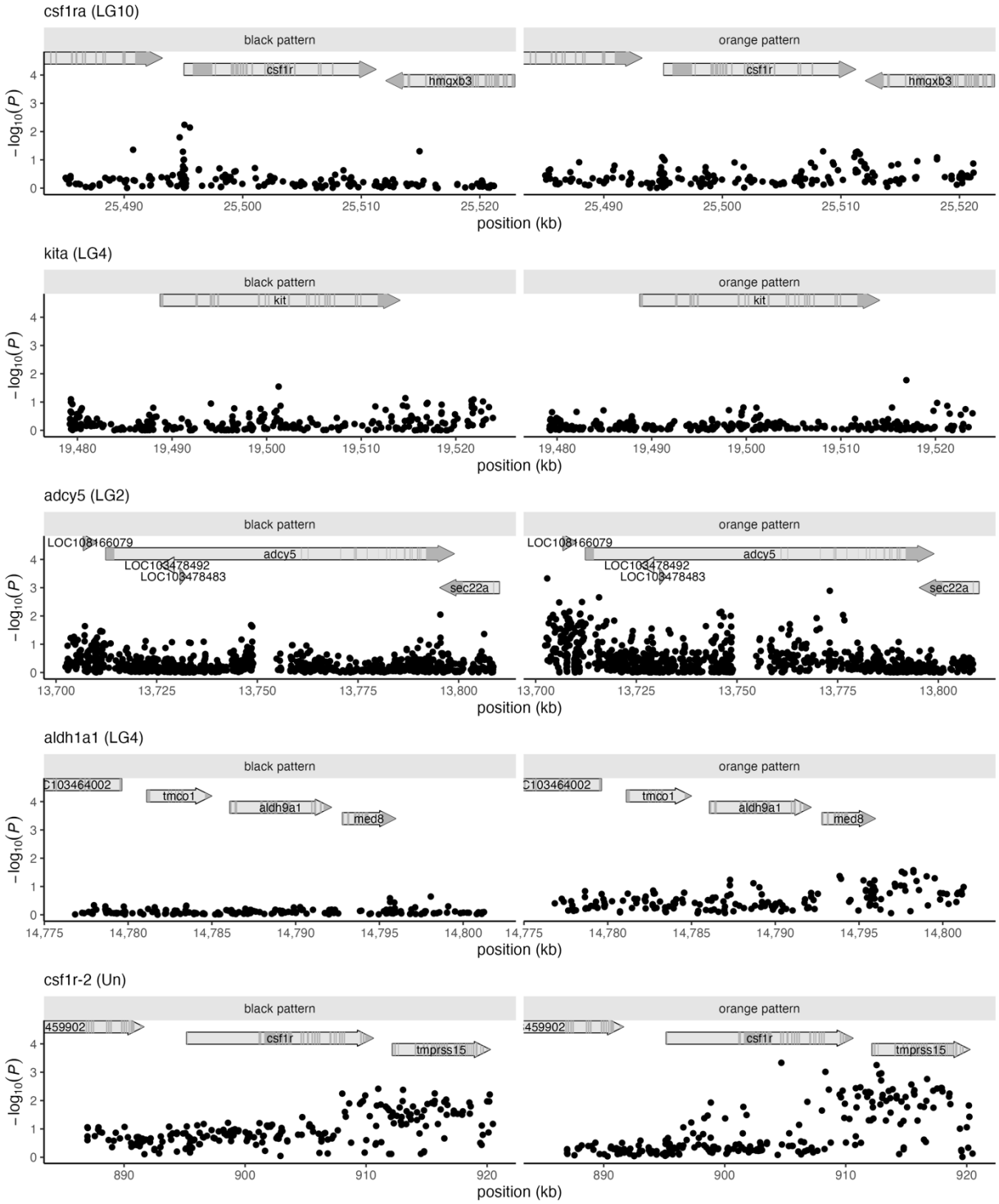


**Figure S13**: No relationship between black and orange color pattern and genes previously indicated in carotenoid and melanic pigmentation. Panels show GWAS results for black and orange patter space.

Supplemental tables

Table S1: Architecture of the convolutional neural net used in step 3 of the pipeline.

| Layer | Activation | Filters | Output shape | Spatial dropout |
|---|---|---|---|---|
| Convolution | Relu | 32 | [400, 144, 32] | 0.4 |
| Convolution | Relu | 64 | [400, 144, 64] | 0.4 |
| Max pooling | | | [200, 72, 64] | |
| Convolution | Relu | 128 | [200, 72, 128] | 0.4 |
| Deconvolution | Relu | 128 | [200, 72, 128] | 0.4 |
| Upsampling | | | [400, 144, 128] | |
| Deconvolution | Relu | 64 | [400, 144, 64] | 0.4 |
| Deconvolution | Linear | 4 | [400, 144, 4] | |