UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI BIOLOGIA

Corso di Laurea in Biologia

ELABORATO DI LAUREA

Analysis of morphological characters linked to survival and differential reproduction in hybrids between *Xiphophorus birchmanni* (Lechner, Radda, 1987) and *Xiphophorus malinche* (Rauchengerger, Kallmann, Morizot, 1990)

Tutor: Prof. Gil Guastoni Rosenthal

Dipartimento di Biologia

Laureando: Giuseppe Paganin

ANNO ACCADEMICO 2023/2024

INDEX

1. ABSTRACT

Ecological selection, sexual selection and genetic drift play a fundamental role in shaping morphological change over time. This work attempts to highlight the strengths of the three main evolutionary forces in an experimental model. Observing the effects of microevolution is particularly easy in hybrid populations, because they demonstrate greater phenotypic variability than parental species, so, a crossbreed beetwen two mexican swordtail fish, *Xiphophorus birchmanni* and *Xiphophorus malinche*, has been investigated. Hybrids manteined into indipendent stocktank at two altidtudinal sites for several generations show differentiation in many characters, in paticular for traits linked to body size, driven by temperature; contemporarily, for some characteristics such as the presence of melanistic pigmentation, an asymmetric distribution between the tanks is observed, due to random segregation of carrier individuals. These results suggest the persistence across generations of the effects of genetic drift, most evidently for qualitative traits, and, at the same time, the emergence and gaining greater importance of ecological selection, especially for quantitative traits.

2. INTRODUCTION

2.1 Morphological characters in relation to evolutionary forces

Charles Darwin, in his most important work, "On the origin of species by means of natural selection", introduced the concept of natural selection as the process that favors individuals with heritable traits better adapted to the environment (Darwin, 1859). In this same book, however, he also introduces the notion of sexual selection as the process that favors individuals with traits whose exclusive function is to increase the reproductive success. This theme will be further explored by Darwin himself in "The descent of man and selection in relation to sex", in which he distinguished two types of male traits shaped by sexual selection: armaments and ornaments (Darwin, 1871). While armaments are involved in direct intra-sexual competition between males, sometimes even violent, ornaments perform their function at the inter-sexual level, implying the female choice of partner. Phenotypic traits that are widely spread in a population, therefore, are shaped by these selective pressures (but not exclusively these two), which tend to exacerbate differences between isolated populations, if these are subjected to different conditions. The aim of this work is to test whether the morphological differences between hybrid populations of *Xiphophorus birchmanni* (Lechner, Radda, 1987) and *Xiphophorus malinche* (Rauchengerger, Kallmann, Morizot, 1990), maintained in different stocktanks at different altitudes, are statistically significant. In this case, the causes of these differences should be hypothesized, without forgetting that not all characters are the result of adaptive evolution, as ironically pointed out by Stephen J. Gould and Richard C. Lewontin in "The Spandrels of San Marco and the panglossian paradigm" (Gould, Lewontin, 1979); other stochastic processes are essential for understanding microevolutionary changes, such as the phenomenon called genetic drift, the random fluctuation of allele frequencies across generations, strongly affected by sample size.

2.2 Modelling the biological question

When approaching a scientific problem, it is necessary to develop an ideal model from which to draw theoretical hypotheses. This initial step is critical to designing an experiment that confirms or rejects assumptions. To test which selective pressure has the greatest influence in a microevolutionary process, we can create a system that allows us to distinguish the action of different evolutionary forces. We can imagine the initial situation of a uniform population, in which the variability of phenotypic traits is normally distributed. We can divide this population into a large number of subpopulations and keep them isolated from each other, so that their evolution over time is absolutely independent. If all subpopulations are subjected to the same environmental conditions (ideally), the morphological differences that will accumulate generation after generation will be solely the result of the two evolutionary forces that can act differentially: sexual selection and genetic drift. In isolated subpopulations, several social contexts can be established; this can lead to the creation of different mating systems, with the consequent diversification of characters, in particular those linked to the mate choice. The mechanisms of genetic drift, in particular the founder effect, can heavily influence this system, the more we divide the initial population into modest-sized subsets. Although behavioral observations can determine the presence of different mating pattern between subpopulations, it's also true that genetic drift phenomena can underlie the creation of different social systems, therefore the two evolutionary forces can be closely related and involved in morphological diversification; in this context, it becomes difficult to clearly distinguish their effects. Now, we can imagine to create two groups of independent subpopulations, and maintaining them in two different contexts: we will have a certain number of isolated subpopulations subjected to the same environmental conditions, different from the conditions of the other set of subpopulations. In this case we have divergent ecological selection at the level of the two groups. This ideal model allows us to test the robustness of ecological selection, respect to other evolutionary forces: the variability between subpopulations maintained in the same environment is caused only by sexual selection and genetic drift, while the variability between groups maintained in different environments is also caused by ecological selection (a particular type of natural selection). If, in an experimental model, we demonstrate that morphological differences are more significant between groups in different environments than between subpopulations in the same environment, we would have evidence that ecological selection has acted as the main evolutionary force. Otherwise, if we demonstrated the opposite, it would mean that sexual selection and genetic drift (unfortunately without being able to clearly distinguish them) have a greater importance in microevolution. If no significant differences were highlighted, it could reasonably be assumed that the different pressures are relaxed, and require a longer time to lead to an appreciable differentiation.

2.3 Experimental lines of evolving hybrids

Our simple ideal model could serve as a theoretical schematization for the experimental apparatus on which I carried out the analyses. The ELEH project (Experimental Lines of Evolving Hybrids or Experimentos Largos de Evolucion Hibrida) was born in 2013, under the direction of Professor Gil Guastoni Rosenthal and Dr. Rhonda Struminger, co-

founders and co-directors since 2005 of the research station of CICHAZ (Centre de Investigaciones Cientificas de las Huastecas “Aguazarca”) in Calnali, in the federal state of Hidalgo, Mexico. Twenty-four 2000L artificial mesocosm stock tanks were built near the CICHAZ field station, divided into three sites with eight tanks each, at three different altitudes: high altitude (STH i.e. Stock Tank High, 1514 m), intermediate (STM, 980 m) and low altitude (STL, 186 m); In each replicate, 10 males and 10 females F1 hybrids (generated by the crossbreed between *X. malinche* females collected from the Chicayotla locality on the Rio Xontla, and *X. birchmanni* males collected in the Rio Coacuilco) were introduced (Bovio, 2022). The populations thus formed were left to evolve independently in the following years, and continue to this day.
We can easily trace the experimental structure just described to the ideal model of the previous paragraph: the subpopulations are represented by 24 singles tanks and the groups in different contexts are represented by the set of tanks (8) at different altitudinal sites. The differentiated environmental conditions that guide ecological selection are attributable to altitude, and therefore, primarily, to temperature (net of any human errors in feeding fish). Through the use of statistical tools I tried to determine whether the major morphological differences are found between sites or between tanks.
For my work, I only considered four replicates at high (STH 3, 4, 7, 8) and four at low (STL 2, 4, 6, 8) altitude (I excluded STM essentially due to lack of time, including this site could be a good starting point for future work). This system has been used in recent years for various experiments, but particularly interesting for my research, as it also deals with the same topic as me, is certainly Richard Stephen Bovio's dissetation, dated 2022. He proved that “there is more variation in morphology between replicates within hybrid populations than between hybrid populations at different sites”. This pattern is expected when ecological selection is not the primary force governing morphological evolution, so he concludes that “genetic drift or arbitrary runaway selection may have a greater influence in early generation hybrids”. It is important to highlight one aspect of this statement: Bovio's work is concentrated exclusively on the early generation of hybrids, those immediately following the admixture event (carried out in 2015) which led to the creation of F1, and has analyzed samples up to 2019. The data I collected, however, refer to the 2023 sampling, May for STL and July for STH. This means that a relevant number of generations have passed and therefore, what was worth in the short term since hybridization, may change as time passes. My work can be considered as an "update" on the subsequent developments of his experimental model. Consequently, even if my results were reversed compared to his (i.e. if there was more variation between populations at different sites than between replicates within hybrid populations), the contradiction would only be apparent, as it would be the time factor that would allow an inversion of the trend, highlighting the long-term importance of ecological selection. The latter, however, is absolutely not relegated to a minority role by Bovio, in fact "Principal component analysis revealed the two natural hybrid populations located at lower elevations to share similar morphology with the lowland parental species X. *birchmanni*, and that the third natural hybrid population located at higher elevations shared more similarity with the highland parental species, *X. malinche*. Overall, I found in the lowland natural hybrid populations, environment drives morphology to resemble the lowland parental species".

2.4 *Xiphophorus* hybrids as a study model

Hybridization is generally disfavoured, according to two possible models. In the "extrinsic postzygotic isolation" (EPI, or Darwin model), when two species diverge, they acquire specific adaptations, so, intermediate hybrids will be less adapted to the environment than the parental species. For "Genetic incompatibility" (Bateson-Dobzhansky–Muller model), new mutations arising in diverging species can interact negatively in hybrids and, when alleles never occurred together interact, they prove to be incompatible (Powell et al., 2020). *Xiphophorus* hybrids are capable to overcoming both of these complications. Phenotypes generated by crossing may even be better adapted than parental phenotypes in intermediate habitats along the environmental gradient (Lewontin, Birch, 1966; Bovio, 2022). Furthermore, although hundreds genetic incompatibilities that could cause reduced fitness have been identified (Schumer et al., 2014), according to Rosenthal (2017b) "later-generation and backcross hybrids observed in nature show few detectable defects with respect to viability, and male hybrids are attractive for females at least as much as males of the two parental species". Therefore, in areas where the two species are in sympatry, fertile hybrid populations are found. The distribution of these fish in the Sierra Madre Oriental of Mexico is determined by temperature: populations of *X. malinche* are found at high altitudes with colder water (7-25°C), while *X. birchmanni* are found at lower altitudes in warmer waters (15-35°C) (Culumber et al., 2012). Hybridization occours at intermediate altitudes as a consequence of water pollution caused by human activities: females are unable to recognize the pheromonal signals, necessary to discern between conspecific and heterospecific males, probably due to the high levels of humic acid HA (Fisher et al., 2006). When this happens, genetic assets are reshuffled, generating new phenotypic combinations on which selection can act. Hybrids give us the advantage of being able to observe a wider range of phenotypes than the parental species (see figure 1); in particular, *Xiphophorus* hybrids express extensive variability in traits associated with social communication, female mating choice and aggressive interactions between males. They are, therefore, an excellent model not only for this work, but also for a notable number of experiments in the field of behavioral genetic.

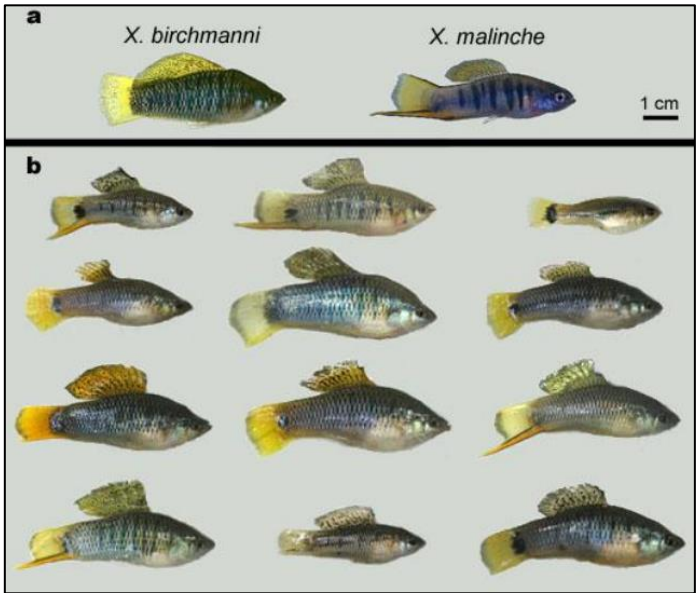


Fig. 1 Wild-caught, sexually mature male (a) Pure *X. birchmanni* and *X. malinche*. (b) *X. birchmanni/malinche* hybrids collected in the Rio Calnali at Chahauco. Hybrids display virtually all possible combinations of male traits and often show traits outside of the range of parental species (Fisher et al., 2006).

2.5 *X. birchmanni* x *X. malinche*'s traits

The parental species, belonging to the genus *Xiphophorus* (Heckel, 1848) of the Poeciliidae family (Bonaparte, 1831), are commonly known as "swordtails", together with about thirty species distributed along shallow rivers throughout Central America. In the figure 2., the phylogenetic relationships with some of the main species of the genus is represented; we can notice how *X. birchmanni* and *X. malinche* are really closely related. Like all the Pecilids, they are ovoviviparous with internal fertilization, possible thanks to the gonopodium. They do not exhibit parental care. Sexual dimorphism is marked, and the elaborate male characters I measured, are functional to courtship rituals and intrasexual competition. The two species differ in several characteristics; for example, *X. malinche*'s males possess pigmented caudal fin extensions called "swords", medium-sized dorsal fins, and more slender body depths. *X. birchmanni*'s males, instead do not have swords, have larger dorsal fins and larger body depths, as we can see in both figure 1 and 2. Hybrid males present a wide possibility of combinations for body size, presence of sword, size of the dorsal fin, ... The effects of the expression of unusual secondary sexual signals can confer an advantage in terms of sexual selection: knowing that parental females do not avoid hybrid males in nature, it can be hypothesized that they present traits that are preferred in mate choice, even compared to parental species (this could be another reason for the fact that hybridization is not disadvantaged) (Serena G., 2012).

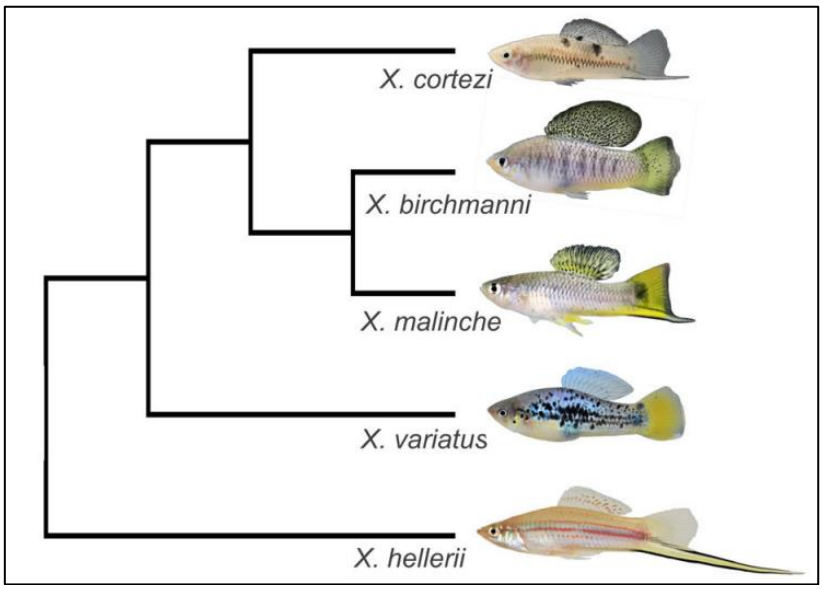


Fig. 2 Phylogenetic relationships between *X. birchmanni, X. malinche*, and other swordtail. Representative individuals from parental species populations are pictured. (Powell et al., 2021)

2.5.1 Body size and shape

Females have variable dimensions, usually between 4 and 5 cm in length. Males are also very variable, presenting a wide range of body sizes (4-6 cm) and this is due to the possible trade-off between investment in different characters: body size is linked to traits with an importance for the probability of survival and producing offspring, such as sword and dorsal fin, therefore growth can be limited by the need to adopt different strategies. Impacting factors, especially on reproductive tactics, can be environmental conditions, availability of resources and the social context, such as density-dependent effects. An extreme example illustrates how, in *X. nigrensis*, males with a smaller body size, called sneakers, allocate more resources to their ejaculates because, being disadvantaged compared to larger males in courtship, the best tactic for them is to try to prevail in sperm competition (Smith, Ryan, 2010). In fact, female preference is directed towards larger males, "wich can be accounted for by increased stimulation of a greater number of retinal photoreceptors" (Rosenthal, 2017a). Examples of this type are present not only among Poecilids, but also among other families, for example Blennidae and Gobiidae, as highlighted by research conducted by the University of Padova (Giacomello et al., 2007; Scaggiante et al.,1999). The two species studied in this work are morphologically distinct because the size of *X. malinche* is smaller (it is difficult to find individuals that exceed 6 cm in length) and the overall shape of the body in this species is more elongated. *X. birchmanni* is characterized by a stockier body shape that in males is even more pronounced by the presence of a cephalic fatty hump (Rosenthal, 2017b). To take body size into account, I measured the standard length, the body depth and the peduncle depth. The measurement guidelines are expressed in figure 3.

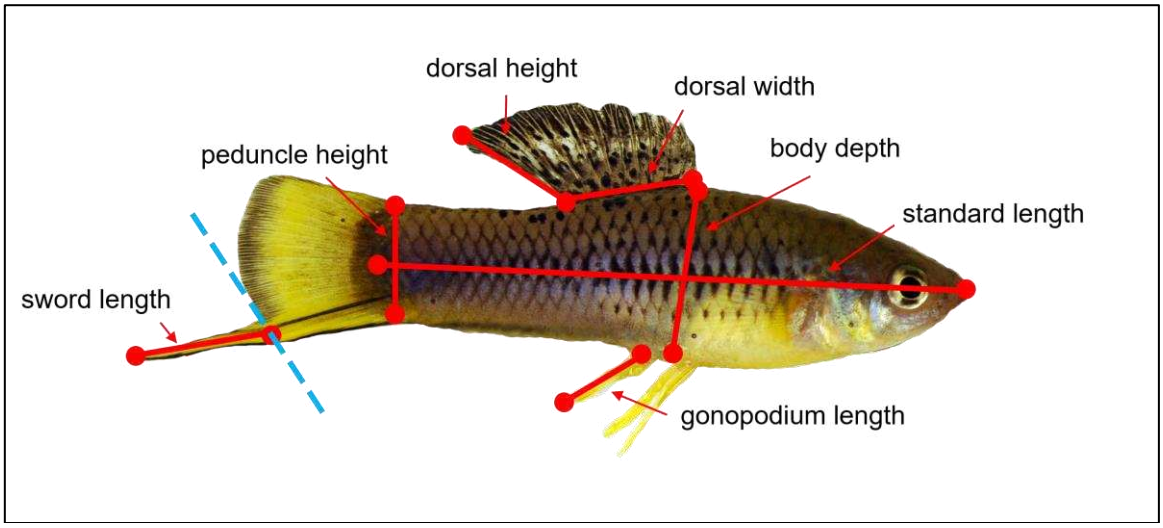


Fig. 3: Measurement guide of continuous traits: standard length, body depth, sword length, dorsal fin width, dorsal fin height, peduncle height, and gonopodium length. Blue dash line indicates cut-off for sworded (i.e. *X. malinche*-skewed) and non-sworded (*X. birchmanni*-skewed) individuals. Individuals that have an extension past the line are considered sworded individuals. (Bovio, 2022)

2.5.2 Sword

The sword is a prominent extension of the lower rays of the caudal fin, usually outlined along the edges by black melanistic stripes, produced by the upregulation of the msx genes, also implicated in gonopodium development, through a testosterone-induced signaling pathway (Zauner, 2003). In the species *Xiphophorus helleri*, females prefer partial swords containing complete black stripes over incomplete swords lacking black coloration, although a complete sword with all stripes is the most attractive (Trainor, Basolo, 2005). In general, females prefer to mate with males with longer swords, but experiments with computer-altered stimuli have shown that females are equally attracted by males with swords and males without swords of equivalent length; therefore, female preference could actually be directed towards males of large apparent size (Rosenthal, Evans, 1998). The latter would use the sword to appear larger, avoiding the costs of investment in body growth: while males that have a lot of available food can invest in both body and sword growth, males with a restricted diet stop growing and shift investments to the sword (Basolo, 1998). While lengthening a sword may be a cheaper solution than body growth in youth, it may still be costly in the long run. Males with longer swords expend more energy while swimming than males of the same size with experimentally shortened swords, as demonstrated in *X. montezumae*; the increased swimming costs, combined with greater visibility, may make males with long swords more susceptible to predators (Kruesi, Alcaraz, 2007). The high risk of predation favors early maturity and therefore more modest sexual ornaments (Kreusi et al., 2010). Given the trade-off between advantages and disadvantages, sword length is highly variable at the individual and interspecific levels. According to Rosenthal et al. (2002), "reduced preference for swords may result from increased choice costs due to predation risk, selection against mating with heterospecifics, or changes in the spatial and contrastive properties of the conspecific signal." In fact, the sword has been secondarily reduced or even lost several times in the phylogenetic tree, males of *X. birchamnni* are devoid of it. In this species, females have secondarily developed a disdain for the sword: males with sword are less attractive than unadorned males. Although preference for the sword must be the ancestral condition, using video animation of virtual males, Wong and Rosenthal (2006) found that when given the chance to choose between a conspecific without a sword and one with a digitally attached sword, females spent significantly more time associating with the former. How female preferences might be lost or reversed is an extremely intriguing field of study.

2.5.3 Dorsal fin

In *X. birchmanni*, Fisher and Rosenthal (2007), using video animations of conspecifics with manipulated fins, demonstrated that females prefer males that display small dorsal fins to those with large ones (Fig. 4). Since body size correlates positively with dorsal fin size, male traits and female preference are misaligned (Rosenthal, 2017a). Despite female avoidance of large dorsal fins, this trait may be advantageous if population density is high (many competitors are present during courtship), since males are less aggressive to other males with large dorsal fins. Dorsal fin display is voluntarily controlled, thus, males raise dorsal fins more frequently when courting in the presence of other males, while they keep them lowered when rivals are absent, mitigating their unattractiveness.

X. birchmanni is unusual in having an enlarged dorsal fin shaped like a trapezoidal sail. *X. malinche* also has a sail-shaped dorsal fin, although less pronounced.

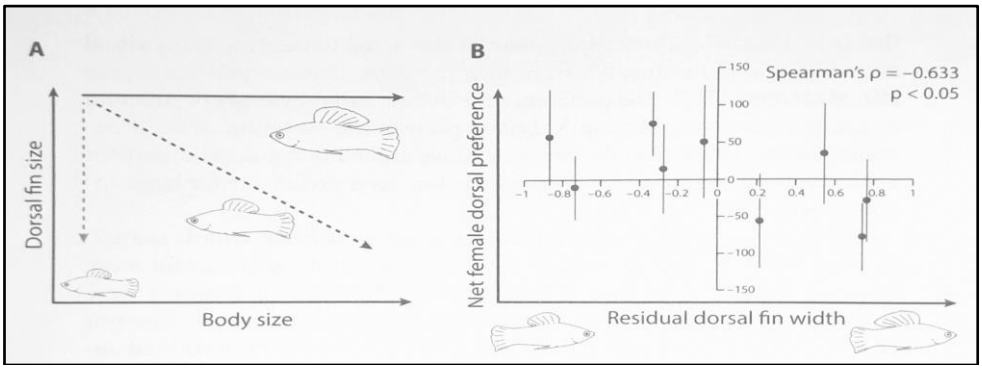


Fig. 4 Female preference for multiple traits in *Xiphophorus birchmanni* is misaligned with male trait variation. A- larger males have bigger dorsal fins, but females prefer males with large bodies (thick, solid arrow) and small dorsal fins (light gray, dashed arrow) resulting in a net preference (dashed arrow) oriented perpendicular to the main axis of male trait variation. B- as predicted by single choice experiments, females prefer males with proportionately small dorsal fins. (Rosenthal, 2017a)

2.5.4 Melanin-based traits and melanoma

According to a review by Culumber (2014), we can distinguish melanic traits into macromelanophores and micromelanophores. Macromelanophores (up to 500 lm) form dots or large spots of black pigment in a wide variety of shapes, producing a mottled appearance on the body (the "Carbomaculatus pattern", or Cm, Fig. 5). Particularly interesting is the "Spotted caudal pattern", or Sc, consisting of clusters of macromelanocytic cells at the junction of the caudal peduncle and the fin (Fig. 6 A, B, C). Micromelanophore, in contrast, are composed by smaller cells, including the caudal spot (Cb, extending across the caudal peduncle, Fig. 6 D), vertical bars, and a diversity of "tailspot" patterns. Furthermore, another difference from macromelanophore models is that their expression is influenced by the nervous system: the spots increase and become darker during social encounters. These characters are, therefore, implicated in communication and probably also in mate choice. I detected the presence of macromelanophore (Sc) or micromelanophore (Cb) pattern on the caudal fin, and the presence of macromelanophore on the body (Cm). As we said in paragraph 2.4, talking about genetic incompatibility, in an article by Schumer et al. (2014), an extensive work is reported, in which they performed a high-resolution genome scan for linkage disequilibrium between unlinked genomic regions in natural swordtail's hybrids, demonstrating that there are "hundreds of pairs of genomic regions contribute to reproductive isolation between these species, despite them being recently diverged". *X. birchmanni* x *X. malinche* hybrids develops melanoma from the spotted caudal pattern (Sc) (Powell et al. 2020). Sc occurs at low frequencies in *X. birchmanni* but is absent from *X.* malinche populations (Fig 7.); instead, hybrid populations show an higher frequency. Histological sections from hybrid individuals revealed penetration of melanocytes into the musculature and invasion of surrounding tissues (where they are

normally absent in *X. birchmanni*), which is indicative of a malignant melanoma. In the article, Powell et al., report to have not identified a single wild-caught *X. birchmanni* male with melanoma, on 1296 individuals collected from 2017 to 2019; So, they assert that "the presence of melanoma in hybrids, but not in the parental species, suggests that this melanoma is a hybrid incompatibility generated by interactions between alleles in the *X. birchmanni* and *X. malinche* genomes", and indicate the oncogene Xmrk, normally kept under control by a repressor in *X. birchmanni*, as responsible for the uncontrolled expression of melanotic tissue. To evaluate how melanoma impacts on viability, they measured swimming performance. They didn't find differences between phenotypes in ability to swim against a current, but, individuals with three-dimensional melanoma had slower escape responses, mainly due to degradation of the muscle tissue connected to the caudal fin. Furthermore, extended blotch makes them more visible for predators. Despite the evidence for reduced survival of spotted individuals in populations with high rates of melanoma, this trait can be found at high frequencies; so, researchers concluded that "there may be weaker effects of melanoma on overall fitness or, alternatively, other factors, such as mating advantages for individuals with large spots, may explain its maintenance". This last alternative hypothesis is also supported by the study on *X. cortezi* by Fernandez and Morris (2008), according to which females prefer males with an enhanced Sc.

Fig. 5: an example of Carbomaculatus pattern from STH

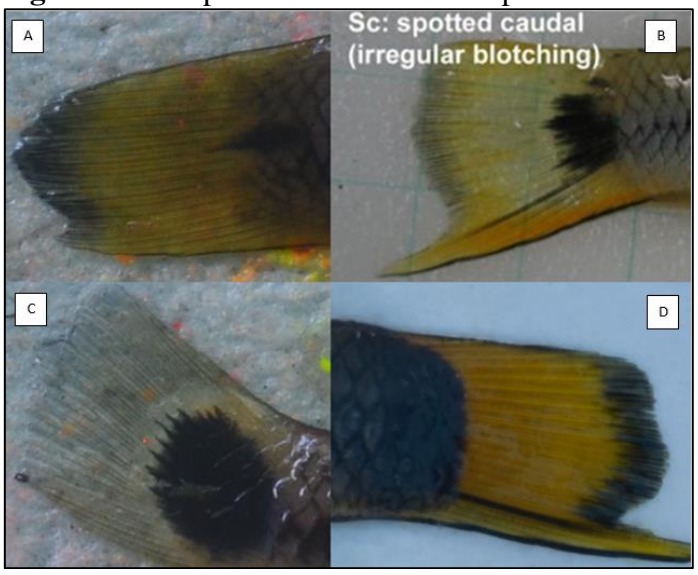


Fig. 6: A- Sc slightly pronounced from STL6; B- example of Sc from a guideline file, provided to me by prof. Rosenthal; C- the most pronounced case that I have been able to observe (probably classifible as cancer), from STL 6; D- example of Cb from STH 4.

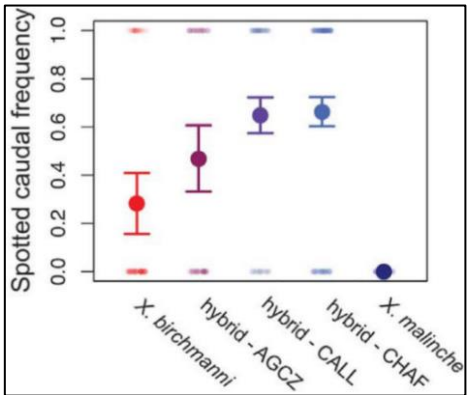


Fig. 7 Whereas X. birchmanni populations segregate for the presence of this spot, the trait is absent in *X. malinche* populations; hybrid populations have high frequencies of this trait. Hybrid phenotypes are shown from three populations on the Río Calnali. AGCZ, Aguazarca; CALL, Calnali low; CHAF, Chahuaco falls. (Powell et al. 2020)

3. METHODS

3.1 Processing untagged fish in the field

In each of the twenty-four tank involved in ELEH project, there are both fish already caught and marked (previously tagged) and fish never identified before (untagged). It is necessary, periodically, to capture as many specimens as possible (tagged and not), collect data and mark the untagged. The illustrated procedure is a summary of what is described in detail in the ELEH_master_protocol document (Bovio, Rosenthal, 2019).

Step 1: Anesthetize fish with MS22.
To anesthetize fish, an individual is placed in a cup with 200-300 mL of water, and the dose of MS-222 is incremented until it lose the equilibrium. One after another, the fish are placed in this solution, so that they can be easily handled.

Step 2: Tag
For each individual, a "scribe" record in a notebook all the useful data for identification and prepare the ID tags, notecards reporting site, tank, collection date, sex, side and color code. In the notebook, he also sign any general notes or mistakes/errors that occur (mislabeled fish in photo, tagging a tagged fish, more than one fish with the same code in a tank, etc.). A "Tattoo Artist" mark fish with three elastomers, applyed at specific positions indicated in Fig.8.

Step 3: Photograph
All the fish is photographed on left and right side, with the ID tags. For males, at least one photograph needs to have the dorsal fin, caudal fin, and gonopodium spread (Fig 9). For females, to spread the fins is not important, because the pricipal interest is the body size (Fig 10). Once collected in the field, photograph has been uploaded on a shared drive, making them available to the scientific community.

Step 4: Fin clip
Using forceps and dissecting scissors, a “surgeon” removes the top quarter of the caudal fin (Fig 8) and place it, with the ID tag notecard, in a 1.5-2.0 mL Eppendorf tube with EtOH.

This sample will be used to genotype individuals. After these steps, fish need recovery and to be returned to their original tank. Contextually, HOBO data (temperature, pH,…) are also collected.

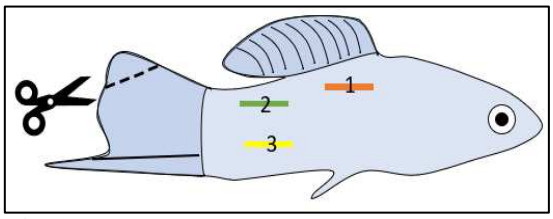


Fig 8. where to cut a fin clip and tag (rigth side). The ID for this fish would be R-OGY. Color Codes: P – Pink; R – Red; O – Orange; Y – Yellow; G – Green; B – Blue; V – Violet; W – White; K – Black. (Bovio, Rosenthal, 2019)

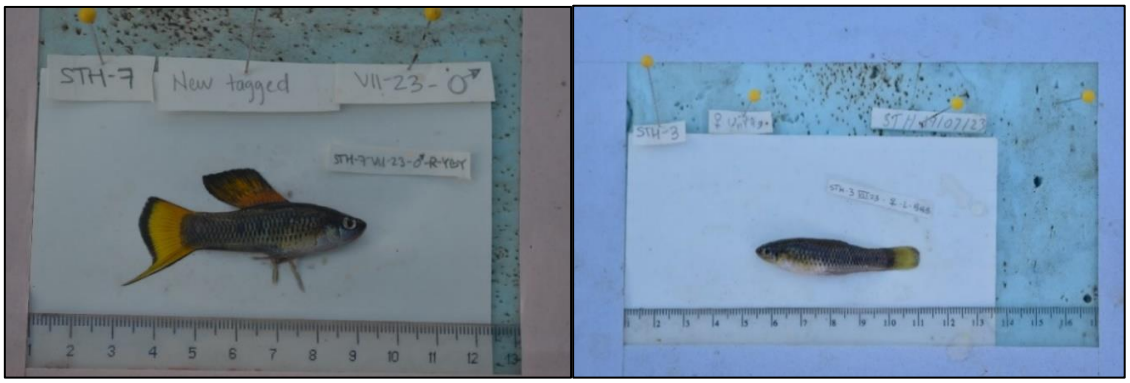


Fig. 9 and 10. Photographs of male and a female from STH 7 and 3

3.2 Data collection with ImageJ software

As regards the material I analysed (untegged fish from STH 3, 4, 7, 8 and from STL 2, 4, 6, 8), the photographs were divided into different folders, basing on the site and date of collection. I have further divided the folders by tank, to make it easier to search for a specific photograph. For both side of each individual (when photos were available), I measured (basing on guidelines in Fig. 3) the quantitative traits with the functions of the ImageJ software; I setted the scale, in millimeters, on the ruler in the photo and estimated: standard body length, body depth, peduncle depth, sword length, dorsal fin height, dorsal fin width and gonopodium length (last four, obviously, for males only). Then I condensed the results for the left and right sides, averaging between these two. For qualitative traits, on the other hand, I made binary decisions, based on previously visualized exemplary cases (Fig. 5 and 6), detecting presence or absence of: Spotted caudal (Sc)/Caudal blotch (Cb) patterns on caudal fin, Carbomaculatus (Cm) pattern on the body, melanistic stripes on upper and lower sword edge. The product of such processing is a single Excel spreadsheet that incorporates all the data collected. Table 1. shows the header and the lines relating to the STL2 tank of this file.

site	tank	sequence id	sex	Sc/Cb	Cm	upper edge	lower edge	standard_lenght	body_depth	peduncle_height	sword_lenght	dorsal_width	dorsal_height	gonopodium
STL	2	STL2V23FYYY	f	Cb	No			44,5935	13,671	7,3995				
STL	2	STL2V23FYYO	f	No	No			45,7285	14,2235	7,679				
STL	2	STL2V23FYYR	f	Cb	No			46,5395	14,0725	7,054				
STL	2	STL2V23FYRY	f	Cb	No			34,6945	12,591	6,3095				
STL	2	STL2V23FYOY	f	No	No			43,857	12,168	6,717				
STL	2	STL2V23FYRR	f	Cb	No			41,1315	10,5795	6,208				
STL	2	STL2V23FYOO	f	No	No			46,421	13,5835	7,457				
STL	2	STL2V23FOYY	f	No	No			45,794	13,023	7,2615				
STL	2	STL2V23FOOO	f	No	No			46,3405	13,2365	7,6755				
STL	2	STL2V23FOOY	f	No	No			41,6155	12,253	6,488				
STL	2	STL2V23FOYO	f	No	No			43,1785	12,581	6,9285				
STL	2	STL2V23FRRR	f	No	No			44,652	13,1355	7,7085				
STL	2	STL2V23FRRY	f	No	No			37,5995	11,1485	5,7195				
STL	2	STL2V23FRRO	f	No	No			43,1605	13,161	6,9915				
STL	2	STL2V23FORR	f	No	No			38,771	11,682	5,997				
STL	2	STL2V23FROR	f	No	No			35,134	10,1835	5,423				
STL	2	STL2V23FOOR	f	Cb	No			36,1225	10,905	5,5215				
STL	2	STL2V23MRRR	m	No	No	1	1	43,4425	12,863	8,236	8,935	9,57	9,2515	8,3325
STL	2	STL2V23MRRO	m	Cb	No	0	0	46,111	13,7535	7,8545	0	8,929	4,969	7,87
STL	2	STL2V23MROR	m	No	No	0	1	39,8525	11,9535	7,311	5,429	8,5325	8,3615	9,0995
STL	2	STL2V23MROO	m	Cb	No	0	1	40,1995	11,1095	6,6735	3,887	6,9795	7,469	8,7095
STL	2	STL2V23MORR	m	Cb	No	1	1	38,7395	11,23	7,71	5,779	8,5165	9,683	7,893
STL	2	STL2V23MORO	m	Cb	No	0	1	44,4335	13,324	8,581	6,697	9,446	9,8465	7,909
STL	2	STL2V23MOOR	m	No	No	1	1	36,8305	10,89	6,263	2,6925	9,2295	7,469	6,9195
STL	2	STL2V23MOOO	m	Cb	No	0	1	42,157	13,497	8,2035	5,1955	8,444	10,2845	7,919

3.3 Statistical analysis in Rstudio

The spreadsheet has been uploaded in R studio space to be used as a dataset for subsequent processing. All the elaboration obtained from the tests are integrally reported in chapter 6. “Appendix: all processing in R software”, in their original graphic form. Table 2 shows the number of individuals subjected to measurements. Although the total number of individuals included in the statistical analysis (388, of which 125 males and 263 females; 45 juveniles were excluded, to avoid that the incomplete development of morphological traits, especially sexual ones, could distort the results) constitutes a valid sample, we must highlight the disproportion between the number of high-altitude tanks (291) compared to those of low altitude (97). I suppose that this latter number represents only a part of all the individuals present in those tanks, but my work is necessarily based on the photographs that were available in the shared drive, so I could not analyze anything other than this material. The limited number of analyzed individuals from the low-altitude site is certainly one of the major limitations of my research. For qualitative traits, I performed Chi square test of independence respect to sites and tanks, to highlight whether there are correlations between these variables and the proportions observed in the population. For quantitative traits, I performed, as first step, Levene test to verify that I could assume homogeneity of variances. This was not possible in only one case, male body size, so I corrected the distortion using Welch’s ANOVA. In all other cases it was possible to perform Nested ANOVA and ANOVA, in order to highlight the presence of significant differences between sites and tanks. Subsequently, for those traits for which the analysis of variance showed significant differences between sites or between tanks, I also performed Nested ANCOVA and ANCOVA. Since many traits are positively correlated with body size, this tests allow us to understand whether the significance of ANOVA is directly driven by belonging to a site, or is an effect related to different growth rate. After that, I used two non-linear regression methods, logistic regression and linear discriminant anlysis, in order to estimate how much the analyzed characters can be good predictors of the site; this can indicate the degree of differentiation between the two sites: if the models are able to establish with a certain precision the belonging of the individuals, then we can conclude that the fish subjected to two different altitudinal conditions present different

characteristics. For these two steps, I used a subset of randomly selected individuals, in order to have the same number of individuals from both sites, thus avoiding distortions due to the sample size. Finally, I performed a Principal component analysis.

Tab. 2

2023	F	M	MJ	sex unknown J	TOT
STH3	36	15	9	3	63
STH4	54	17	9	0	80
STH7	60	26	4	0	90
STH8	57	26	13	0	96
STL2	17	8	0	0	25
STL4	8	11	0	6	25
STL6	15	10	0	0	25
STL8	16	12	0	1	29
TOT	263	125	35	10	
2023	F	M	MJ	sex unknown J	TOT
STH	207	84	35	3	329
STL	56	41	0	7	104
					433

4. RESULTS

4.1 Quantitative traits

Table 3 and 4 shows the mean values in millimeter of the measurements, divided by site and by tank. These are the values on which the ANOVA and ANCOVA tests are based to detect significant differences.

Tab. 3

MALE	body_length	body_depth	peduncle_depth	sword_length	dorsal_width	dorsal_height	gonopodium
3	44,02	12,81	8,14	6,82	10,78	10,10	9,10
4	42,60	12,43	8,12	4,37	10,29	9,09	8,23
7	43,11	12,54	7,99	4,87	10,18	9,07	8,17
8	41,55	12,28	7,79	5,83	10,36	10,24	9,58
STH	42,69	12,49	7,98	5,42	10,36	9,62	8,78
2	41,47	12,33	7,60	4,83	8,71	8,42	8,08
4	40,89	11,79	7,48	4,09	9,30	8,59	7,78
6	40,49	12,27	7,44	2,25	8,03	7,93	8,75
8	40,92	12,96	7,67	2,15	9,29	7,75	8,07
STL	40,91	12,35	7,55	3,21	8,87	8,15	8,16

Tab. 4

FEMALE	body_length	body_depth	peduncle_depth
3	47,15	12,45	7,43
4	44,47	12,16	7,18
7	43,93	11,87	7,09
8	44,40	12,78	7,35
STH	44,76	12,30	7,24
2	42,08	12,48	6,74
4	41,10	12,63	6,59
6	38,38	11,17	6,20
8	41,85	12,42	6,78
STL	40,88	12,14	6,58

4.1.1 Standard body length

Both for males and females, fish at higher altitude are longer on average (STH: 42.67mm vs STL: 40.91mm for male; STH: 44.76mm vs STL: 40.88mm for female; data from Tab.3) and show a wider range of variability (from 33 mm up to 55 mm as for STH among males and even up to 60 mm among females of STH3), compared to lower tank's fish (Fig. 11). If we look at the single tanks, we notice clear differences, especially among females: at higher altitude we have STH3 that reaches an average of 47.15 mm, while, at low altitude, the average for STL6 is 38.38 mm.

Males: since we can't assume the equality of variance (Levene: $F = 4.99$, $P < 0.05$), I have used Welch's ANOVA, which does not necessarily need this requirement to be met. There are significant differences between sites (Welch's ANOVA: $F = 4.40$, $P < 0.05$), but not within: there aren't significant differences neither between high-altitude tanks (Welch's ANOVA: $F = 1.16$, $P > 0.05$), nor between low-altitude tanks (Welch's ANOVA: $F = 0.12$, $P > 0.05$).

Females: Assuming the equality of variance (Levene: $F = 2.64$, $P > 0.05$), there are significant differences between sites, but not within (Nested ANOVA: between $F = 21.38$, $P < 0.0001$; within $F = 2.10$, $P > 0.05$).

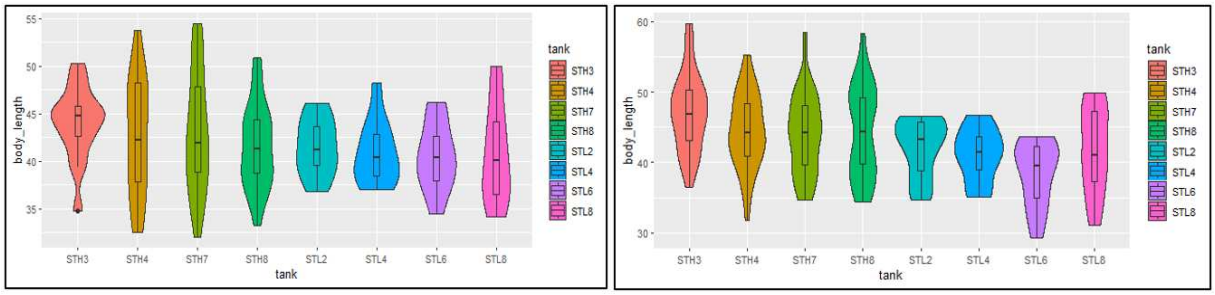


Fig. 11 Violin plot for males (right) and females (left)

4.1.2 Body depth

Both for males and females, there is little difference in average depth (STH: 12.49mm vs STL: 12.35mm for male; STH: 12.30mm vs STL: 12.14mm for female), but STH' s tank shows a wider range of variability than STL's (Fig. 12).

Males: Assuming the equality of variance (Levene: $F = 2.76$, $P > 0.05$), there aren't significant differences neither between sites, nor within (Nested ANOVA: between $F = 0.15$, $P > 0.05$; within $F = 0.54$, $P > 0.05$)

Females: Assuming the equality of variance (Levene: $F = 0.18$, $P > 0.05$), there aren't significant differences between sites, but there are within sites (Nested ANOVA: between $F = 0.41$, $P > 0.05$; within $F = 2.65$, $P < 0.05$). In particular, there is a significant difference between the high-altitude tanks (ANOVA: $F = 2.94$, $P < 0.05$), but not between the low-altitude ones (ANOVA: $F = 2.71$, $P > 0.05$). To verify whether the observed significance is due to real morphological changes or if it is rather linked to the difference in size, I performed the analysis of covariance, using the standard length as a covariable. As already indicated by Nested ANOVA, the differences between sites are not significant, while are within sites; the correlation between standard length and body depth is strong (Nested ANCOVA: between $F = 1.60$, $P > 0.5$; within $F = 7.44$, $P < 0.0001$; body length $F = 754.88$, $P < 0.0001$). Checking the high altitude tanks, we see that, as for the ANOVA, the difference is significant, and that size has an important effect (ANCOVA: between tank $F = 10.84$, $P < 0.0001$; body length $F = 547.59$, $P <$

0.0001). This last result was predictable because there is certainly a strong positive correlation between body length and depth.

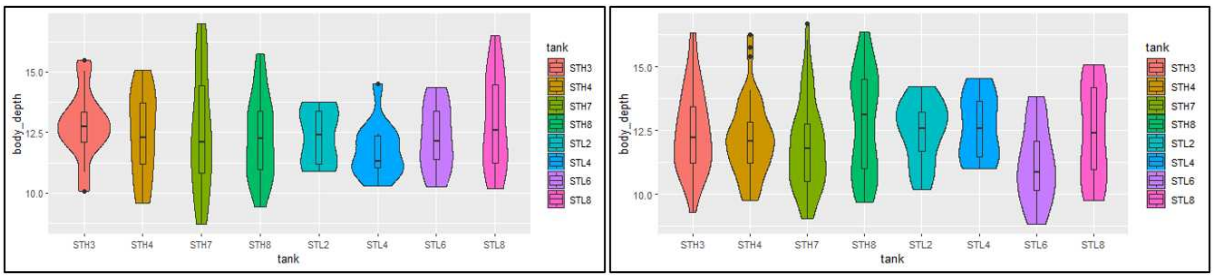


Fig. 12 Violin plot for males (right) and females (left)

4.2.3 Peduncle depth

Both for males and females, there is not a striking difference in average depth between sites (STH: 7,98mm vs STL: 7,55mm for male; STH: 7,24mm vs STL: 6,58mm for female), but those from the higher tanks are still deeper. Also in this case, STH shows a wider range of variability than STL (Fig. 13).

Males: Assuming the equality of variance (Levene: $F = 2.70$, $P > 0.05$), there aren't significant differences neither between sites, nor within (Nested ANOVA: between $F = 3.27$, $P > 0.05$; within $F = 0.21$, $P > 0.05$).

Females: Assuming the equality of variance (Levene: $F = 0.67$, $P > 0.05$), there are significant difference between sites, but not within (Nested ANOVA: between $F = 19.33$, $P < 0.0001$; within $F = 1.12$, $P > 0.05$). For the analysis of covariance, differences between sites are significant, and the effect of standard length must be taken into account (ANCOVA: between sites $F = 137.10$, $P < 0.0001$; body length $F = 1596.00$, $P < 0.0001$). This last result was predictable because there is certainly a strong positive correlation between body length and peduncle depth.

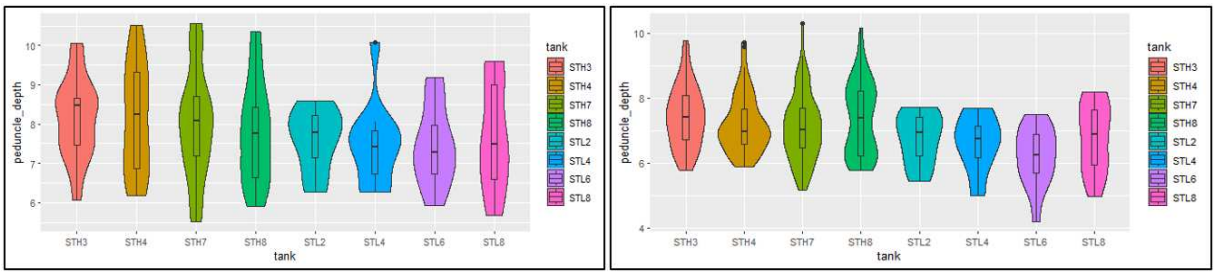


Fig.13 Violin plot for males (right) and females (left)

4.1.4 Sword extension length

This is certainly the trait that shows the widest variability, at all levels, as it was predictable (Fig. 14). Fish at higher altitude have longer swords on average, compared to lower tank's fish (STH: 5,42mm vs STL: 3,21mm). Within high site, STH3 shows the longest sword (6,82mm on average), followed by STH8 (5,83), while, at low site, STH6 (2,15mm on average) and STH8 (2,25) present a larger proportion of individuals without sword, compared to all the other tanks.

Assuming equality of variance (Levene: $F = 3e\text{-}4$, $P > 0.05$), there are significant differences, both between sites, and within (Nested ANOVA: between $F = 18.51$, $P < 0.0001$; within $F = 2.61$, $P < 0.05$). At the level of the tanks at the same altitude, there is

a significant difference between the high altitude tanks (ANOVA: $F = 4.89$, $P < 0.0001$), but not between the low altitude tanks (ANOVA: $F = 2.67$, $P > 0.05$). The analysis of covariance confirms the significance of the differences between sites; the interesting result is the non-significance of the effect of standard length (ANCOVA: site $F = 17.554$, $P < 0.0001$; body length $F = 3.76$, $P > 0.05$). Checking the high-altitude tanks, we see that, as for the ANOVA, the difference is significant, and that, even at this level, size does not have a significant effect (ANCOVA: between tank $F = 5.00$, $P < 0.0001$; body length $F = 3.64$, $P > 0.05$).

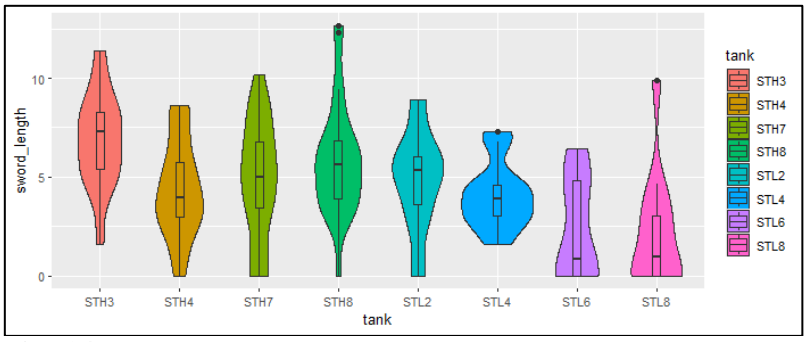


Fig. 14 violin plot

4.1.5 Dorsal height

Fish at higher altitude have bigger fins on average (STH: 9,62mm vs STL: 8,15mm), compared to lower tank's fish. Also in this case STH shows a wider range of variability than STL (Fig. 15). At the lower site, the distribution seems quite homogeneous, except for an outlier individual in STL6 (Fig. 16); it is peculiar, as all its characters are small in size, but not at the extreme level of the fin. On the other hand, its Spotted caudal pattern is by far the most pronounced that I have been able to observe (as I have already reported in Fig. 5C). This subject could present an extensive cancer and have been limited in its growth. Another plausible hypothesis is that, during the preparation for the photograph, the fin was not correctly extended and is actually partially folded under the body. Assuming the equality of variance (Levene: $F = 1.19$, $P > 0.05$), there are significant differences between sites, but not within (Nested ANOVA: between $F = 13.79$, $P < 0.001$; within $F = 0.56$, $P > 0.05$). For the analysis of covariance, differences between sites are significant, and the effect of standard length must be taken into account (Nested ANCOVA: site $F = 45.95$, $P < 0.0001$; tank $F = 2.13$, $P > 0.05$; body length $F = 272.45$, $P < .0001$). This last result was predictable because there is a positive correlation between body length and dorsal fin.

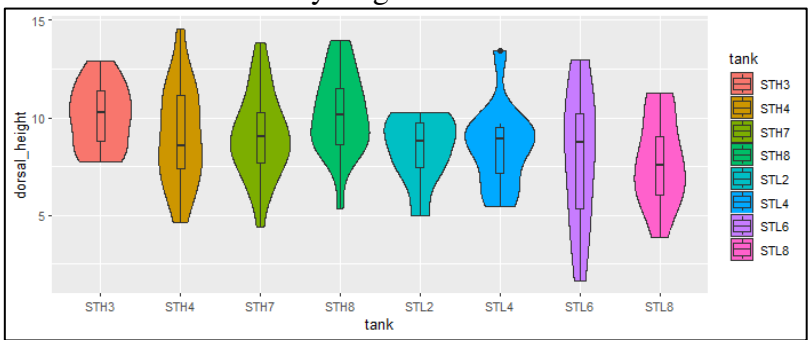


Fig. 15 violin plot

Fig.16 STL6V23MOYY (not genotyped)

4.1.6 Dorsal width

Fish at higher altitude have bigger fins on average (STH: 10,36mm vs STL: 8,87mm), compared to lower tank's fish. The violin plot shows us a distribution quite similar to what we have already observed for dorsal height, with the same outlier in STL6 (Fig. 15); obviously, dorsal height and dorsal width are extremely strongly correlated. Assuming equality of variance (Levene: $F = 0.42$, $P > 0.05$), there are significant differences between sites, but not between tanks (Nested ANOVA: between $F = 11.49$, $P < 0.001$; within $F = 0.99$, $P > 0.05$). For the analysis of covariance, differences between sites are significant, and the effect of standard length must be taken into account (ANCOVA: site $F = 18.16$, $P < 0.0001$; body length $F = 72.22$, $P < 0.0001$).

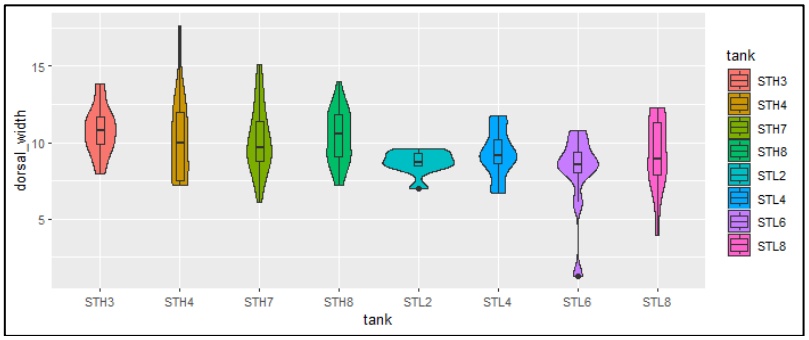


Fig.17 violin plot

4.1.7 Gonopodium length

Measuring the gonopodium was quite difficult, for several reasons: in many photographs it is not fully (or not at all) extended; in others, the brightness is not optimal, so it is almost transparent. This could have influenced the results. In the violin plot (Fig. 18), we see a rather heterogeneous distribution of measurements. The means between the two sites are not extremely far (STH: 8.78mm vs STL: 8.16mm), but significant differences within the two sites are also noticeable at first glance; the mean of STH reaches 9.58mm, while among the other tanks, only STH3 exceeds 9mm (9.10). The Kernel desity take on very different profiles within the lower site.
Assuming the equality of variance (Levene: $F = 1.50$, $P > 0.05$), there are significant differences, both between and within sites (Nested ANOVA: between $F = 8.16$, $P < 0.01$; within $F = 4.81$, $P < 0.001$). At the level of tanks at the same altitude, there is a significant difference at high-altitude (ANOVA: $F = 5.51$, $P < 0.0001$), but not at low-altitude (ANOVA: $F = 1.49$, $P > 0.05$). The analysis of covariance confirms the significance of the differences both between and within sites (Nested ANCOVA: site $F = 11.17$, $P < 0.01$; tank $F = 37.07$, $P < 0.0001$; body length $F = 7.77$, $P < 0.0001$); the effect of standard length must be taken into account. Checking the high altitude tanks,

we see that, as for the ANOVA, the difference is significant, and that, also at this level, size has an important effect (ANCOVA: between tank $F = 7.45$, $P < 0.0001$; body length $F = 42.03$, $P < 0.0001$).

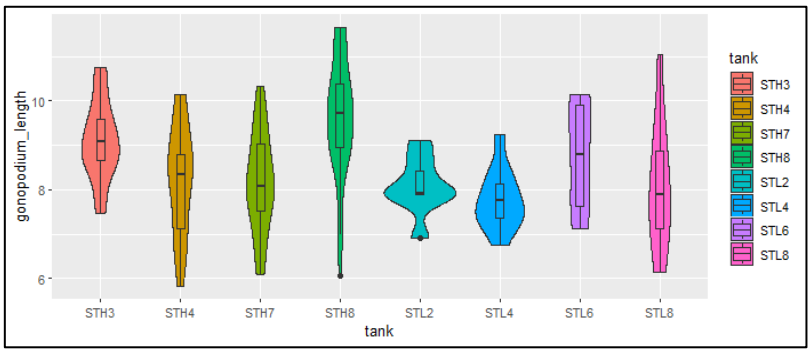


Fig.18 violin plot

4.2 Qualitative traits

Tables 5 (for males) and 6 (for females) shows the observed frequencies, in absolute and in percentage, in which the researched traits are present in the populations. The *X*-squared tests, used to investigate the relationship between the traits and the categorical variables "site" and "tank", are based on the proportions that can be obtained from these tables.

Tab. 5 and 6

MALE	Sc	Cb	None	Cm	upper edge	lower edge	TOT
STH3	0	14	1	0	14	15	15
STH4	0	13	4	0	13	14	17
STH7	3	15	8	3	12	22	26
STH8	0	23	3	0	14	25	26
STH	3	65	16	3	53	76	84
STL2	0	5	3	0	3	7	8
STL4	0	7	4	0	7	11	11
STL6	6	1	3	0	1	6	10
STL8	0	3	9	0	2	7	12
STL	6	16	19	0	13	31	41

MALE	Sc	Cb	None	Cm	upper edge	lower edge
STH3	0,0%	93,3%	6,7%	0,0%	93,3%	100,0%
STH4	0,0%	76,5%	23,5%	0,0%	76,5%	82,4%
STH7	11,5%	57,7%	30,8%	11,5%	46,2%	84,6%
STH8	0,0%	88,5%	11,5%	0,0%	53,8%	96,2%
STH	3,6%	77,4%	19,0%	3,6%	63,1%	90,5%
STL2	0,0%	62,5%	37,5%	0,0%	37,5%	87,5%
STL4	0,0%	63,6%	36,4%	0,0%	63,6%	100,0%
STL6	60,0%	10,0%	30,0%	0,0%	10,0%	60,0%
STL8	0,0%	25,0%	75,0%	0,0%	16,7%	58,3%
STL	14,6%	39,0%	46,3%	0,0%	31,7%	75,6%

FEMALE	Sc	Cb	None	Cm	TOT
STH3	0	24	12	0	36
STH4	0	34	20	0	54
STH7	7	28	25	11	60
STH8	0	33	24	0	57
STH	7	119	81	11	207
STL2	0	5	12	0	17
STL4	0	3	5	0	8
STL6	0	6	9	0	15
STL8	0	3	13	0	16
STL	0	17	39	0	56

FEMALE	Sc	Cb	None	Cm
STH3	0,0%	66,7%	33,3%	0,0%
STH4	0,0%	63,0%	37,0%	0,0%
STH7	11,7%	46,7%	41,7%	18,3%
STH8	0,0%	57,9%	42,1%	0,0%
STH	3,4%	57,5%	39,1%	5,3%
STL2	0,0%	29,4%	70,6%	0,0%
STL4	0,0%	37,5%	62,5%	0,0%
STL6	0,0%	40,0%	60,0%	0,0%
STL8	0,0%	18,8%	81,3%	0,0%
STL	0,0%	30,4%	69,6%	0,0%

4.2.1 Upper/lower sword edge's melanistic stripes

Upper sword edge is much more widespread in high altitude tanks (63.1% on average, from Table 5), compared to low altitude tanks (32%). We see that the percentages are very different even within the same site (for example: STH3- 93.3% vs STH7- 46.2%; STL4- 63.6% vs STL6- 10%). The site and upper sword edge variables are not independent (Pearson's Chi squared: $X = 9.67$, $P < 0.01$). Performing the test within the individual sites, both for STH (Pearson's Chi squared: $X = 11.36$, $P < 0.01$) and for STL (Pearson's Chi squared: $X = 8.73$, $P < 0.05$) the tank and upper sword edge variables are not independent. There is, therefore, a strong relationship between this character and the site/tank it belongs to. Lower sword edge, instead, is widely spread at all levels (STH- 90.5%, STL- 76%). The site and lower sword edge variables are independent (Pearson's Chi squared: $X = 3.81$, $P > 0.05$), even though we are very close to the 5% threshold (0.051). Running the test within the individual sites, both for STH (Pearson's Chi squared: $X = 4.89$, $P > 0.05$) and for STL (Pearson's Chi squared: $X = 7.42$, $P > 0.05$), the tank and lower sword edge variables are independent (even though we are quite close to the 5% threshold (0.059) for STL). There is, therefore, no strong relationship between this character and the site/tank it belongs to.

4.2.2 Sc/Cb/None

For males, Caudal blotch pattern is much more widespread in the high altitude tanks (77.4% on average), compared to the low altitude ones (39%). Spotted Caudal is limited to only two tanks, at both sites (STH7- 11.5%; STL6- 60%). The site and Sc/Cb/None variables are not independent (Pearson's Chi squared: $X = 16.63$, $P < 0.001$). Performing the test within the individual sites, we obtain a different result between the two: while for STH the tank and Sc/Cb/None variables are independent (Pearson's Chi squared: $X = 10.86$, $P > 0.05$), for STL they are not (Pearson's Chi squared: $X = 27.27$, $P < 0.001$). For females, as for males, we can note that Caudal blotch is more widespread in high altitude tanks (57.5% on average), compared to low altitude tanks (30.4%). Spotted Caudal is limited to a single high altitude tank, where it was also present for males (STH7- 11.5%). The variables site and Sc/Cb/None are not independent (Pearson's Chi squared: $X = 17.16$, $P < 0.001$). Performing the test within the individual sites, we obtain an inverted situation compared to that of the males: while for STH the variables tank and Sc/Cb/None are not independent (Pearson's Chi squared: $X = 19.78$, $P < 0.01$), for STL they are (Pearson's Chi squared: $X = 1.88$, $P > 0.05$).

4.2.3 Cm/None

The Carbomaculatus pattern is limited to a single tank, STH7, for both males (11.5%) and females (18.3%). Most of these individuals, in addition to Cm, also present Sc; however, this is not found in STL6, where Sc is widespread, but Cm is completely absent. It should be noted that, although they are both classified as Sc, the patterns present in STH7 and STL6, respectively, are rather different: in the first tank we have a typology like the one shown in Fig. 6A, with a small spot limited to the transition zone between the peduncle and the fin; in the second, instead, we have a typology like the one shown in Fig. 5, where the caudal fin is invaded by extensive spots, also present on the body, as if it were a posterior extension of Cm. Probably, two typologies of Spotted caudal traits should be distinguished.

For males, the site and Cm/None variables are independent (Pearson's Chi squared: X = 0.36, P > 0.05). Performing the test within individual sites, both for STH (Pearson's Chi squared: X = 6.94, P > 0.05) and for STL (Chi squared test for given probabilities: X = 0.85, P > 0.05), the tank and Cm variables are independent.
For females, the site and Cm/None variables are independent (Pearson's Chi squared: X = 1.92, P > 0.05). Running the test within the individual sites, we obtain a different result between the two: while for STH the variables tank and Cm are not independent (Pearson's Chi squared: X = 28.46, P < 0.0001), for STL they are (Chi squared test for given probabilities: X = 3.57, P > 0.05).

4.2.4 A peculiar case

Observing caudal patterns, I noticed some unusual cases and submitted them to the more expert eye of Professor Rosenthal. I first flagged the male in Fig. 19 and then detected at least 5 other similar cases (Fig. 21), in addition to some others of more uncertain interpretation, all coming from different tanks at the high altitude site. The peculiarity is in the area of insertion between the peduncle and the fin, where, instead of having the darker part as in Cb and Sc, we see a clearly delineated pigmentless half-moon area, followed by melanic pigmentation distributed along the direction of the central rays of the fin. Daniel Lee Powell, researcher at Stanford university whose papers I have cited in the bibliography, questioned by e-mail, states that he has not seen anything quite like this before, except in the dorsal fins of some F1 and F2 hybrids, and suggests that it is caused by increased micromelanophore density. Rosenthal, instead, hypothesized an epistatic interaction through a Turing mechanism (Kondo, Miura, 2010). In this case, it is possible to think that the interaction between the morphogenic molecules responsible for the formation of patterns on body surfaces according to the reaction-diffusion model (Turing, 1952), could mask the expression of melanophores in the area behind the peduncle. Powell himself, a week later, reported to us that he caught one putative trihybrid (X. malinche x X. birchmanni x X.variatus) coming from Calnali low natural site, downstream of the town, in which "the space between the crescent pattern and the rest of the caudal fin looks very much like our photos" (Fig. 20). It would be extremely interesting to follow, in the next few years, the evolution of this type of pattern. If it is indeed a recently appeared character, as we might hypothesize from the fact that it had not been observed until now, it could spread in the population. In that case, we could try to understand if it provides an advantage, perhaps in terms of sexual selection for female preference in mate choice, or in term of natural selection, by reduced incidence of melanoma. Unfortunately, only two of the individuals I reported are already genotyped.

Fig. 19 and 20

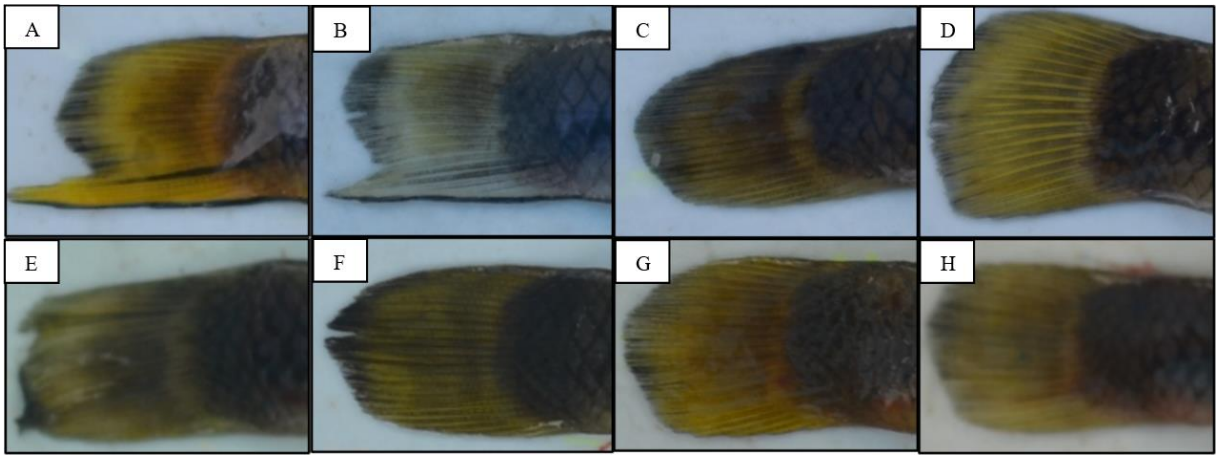


Fig. 21A- STH3VII23MBBO; **B**- STH3VII23MYBY; **C**- STH3VII23FAAO (genotyped); **D**- STH3VII23FOYY; **E**- STH4VII23FBBA; **F**- STH4VII23FBOO; **G**- STH8VII23FOYO; **H**- STH8VII23FAOA (genotyped)

4.3 Logistic regression model, Linear discriminant analysis and PCA

Logistic regression is a statistical tool used to classify samples according to different types of data (both quantitative and qualitative), and to assess which variables are useful predictors for classifying samples. The product is a sigmoid shape logistic function in the form: $S(t) = 1 / [1 + e^{(-t)}]$, where t represents the logs of the odds. The curve, that goes from 0 to 1 in ordinate, tells us the probability that a specific individual can be ascribed to one of the two possible categorical variables. In our case we want to predict the probability that a single fish belongs to the high altitude site rather than the low altitude site, based on the morphological traits collected. If we have an individual that presents typical characteristics of STH, we have a greater probability that it belongs to this site. Instead, if we have an individual that has intermediate characteristics between those that characterize the two sites, then it will have an equal probability of belonging to both (i.e. 50%). If the probabilities are greater than 50%, than we'll classify it as coming from STH, otherwise we'll classify it as coming from STL. If these classifications fit with reality, we will have built a good predictive model and the curve will have a sigmoid shape. In this case we can assume that the two sites are well differentiated, that is, individuals from the high site actually have different characteristics from those from the low site. For males, I used those characters that showed significant differences between sites, except body length to avoid the risk of unwanted effects due to the inequality of variances. Qualitative traits are recorded as 0 (absence) or 1 (presence). I obtained the coefficient t = - 7.76 + 0.11*(sword length) + 0.45*(dorsal width) - 0.29*(dorsal height) + 0.47*(gonopodium length) + 0.69*(melanistic spot) + 1.94*(upper sword edge). This can be interpreted in this way: an individual with a slightly longer sword, a slightly wider but slightly lower fin, a slightly longer gonopodium and with melanistic traits (Sc/Cb instead of None, presence of upper sword edge), has a higher probability of belonging to STH, rather than to STL. Although I have reported all the coefficients, only upper sword edge ($Z = 2.82$, $P < 0.01$) and dorsal width ($Z = 2.05$, $P < 0.5$), together with the intercept ($Z = -3.12$, $P < 0.01$), are significantly useful predictors of site membership, while other traits seem not

to be: sword length (Z = 0.77, P > 0.05), dorsal height (Z = -1.2 1, P > 0.05), gonopodium length (Z = 1.48, P > 0.05) and Sc/Cb/None (Z = 2.82, P < 0.01). The graph (Fig. 22 above), , shows the predicted probabilities that each individual belong to STH, along with their actual site membership: most of the individuals coming from STH (blue), are predicted to have a high probability of belonging to STH and most of the individuals coming from STL (red) are predicted to have a low probability of belonging to STH. There are some mis-predictions, but not many, so we can say we have built a discreet model; the two sites are quite easily distinguishable, so they have become quite different. Probably, with a greater sample size, It would have resembled more to a sigma-like shape. For females, I used all the measured traits. I obtained the coefficient t = - 4.87 + 0.11*(body length) - 1.94*(body depth) + 3.41*(peduncle depth) + 1.27*(melanistic spot), which we can interpret as: if a fish is a little taller, a little less stocky, have a wider peduncle, and present a melanistic trait, it is more likely to belong to STH. All the traits, apart from body length (Z = 0.87, P > 0.05), are significantly useful predictors of site membership: body depth (Z = -4.73, P < 0.0001), peduncle depth (Z = 3.63, P < 0.001), Sc/Cb/None (Z = 2.39, P < 0.5), together with the intercept (Z = -2.38, P < 0.05). In the graph (Fig. 22 below), most of the individuals coming from STH (red), are predicted to have a high probability of belonging to STH and most of the individuals coming from STL (red) are predicted to have a low probability of belonging to STH; as for males, mispredictions are few. Furthermore, perhaps due to the larger sample size, it resembles a sigma-like shape quite well. Linear discriminant analysis works for the same purpose: it tries to explain a categorical variable by the values of the independent variables. We are interested in maximizing the separability between the two groups, so we want to create a new axis and project the data onto it. For males I obtained these coefficients: sword length = 0.08 dorsal width = 0.32 dorsal height = -0.20 gonopodium length = 0.35 melanic spot = 0.53 upper edge of the sword = 1.63); they seem very similar to those obtained with logistic regression. I then used these coefficients to run a simulation in which I tested the ability of the model to predict site membership for each individual from a randomly selected sample of fish from both altitudes. The correct decision was made in 61 cases out of 82. This classification is highly significant compared to a random expectation (Pearson X Square: X = 17.62 , P < 0.0001). For females the coefficients are: standard length = 0.07, body depth = -1.12, pedicle depth = 1.79, melanic spot = 0.799. In this case the simulation made the right decision in 85 cases out of 112 and the classification is significant (Pearson X Square: X = 29.12, P < 0.0001). As final step, I performed a PCA (Principal component analysis), in order to make the situation graphically viewable. From the graphs for males and females (Fig. 23 and 24) a large amount of information can be obtained, the interpretation of which I will leave to the next chapter, after having exposed some reworkings of what was obtained with the previous analyses: the PCA provides us with a synthesis of all the results obtained, allowing us to recap the most important concepts.

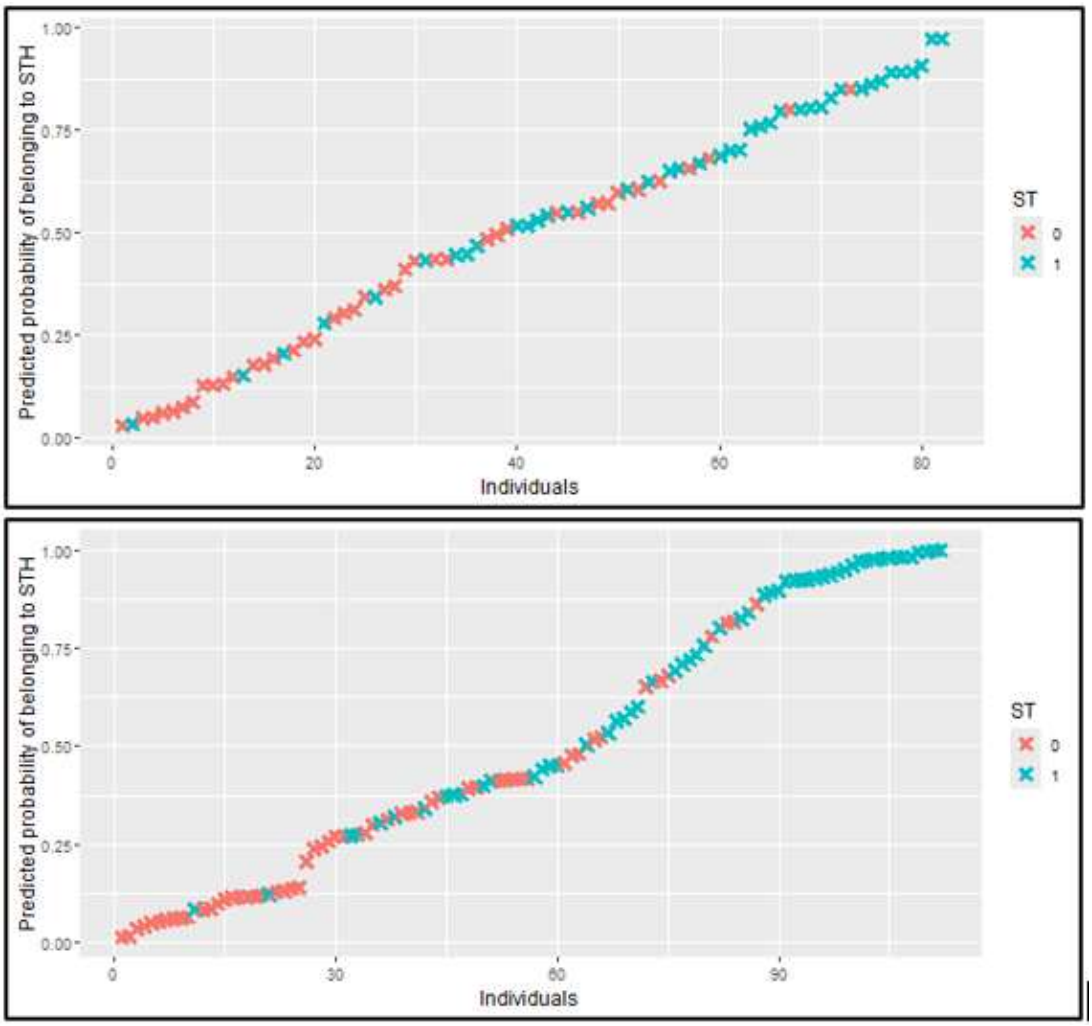
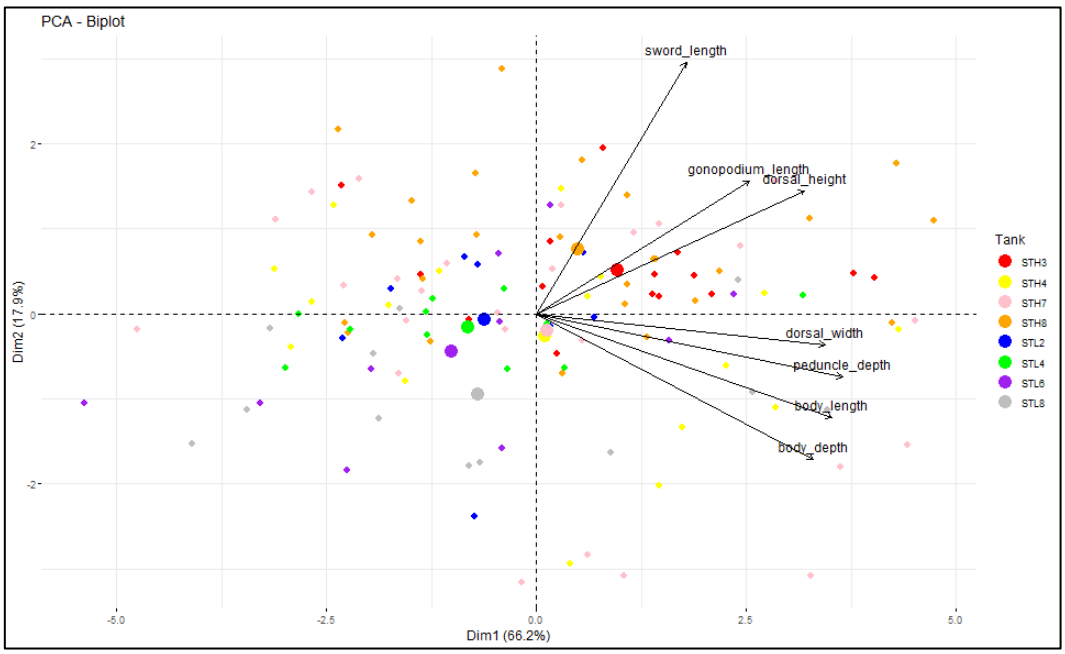


Fig. 22: 0 (RED) = STL, **1** (BLUE) = STH



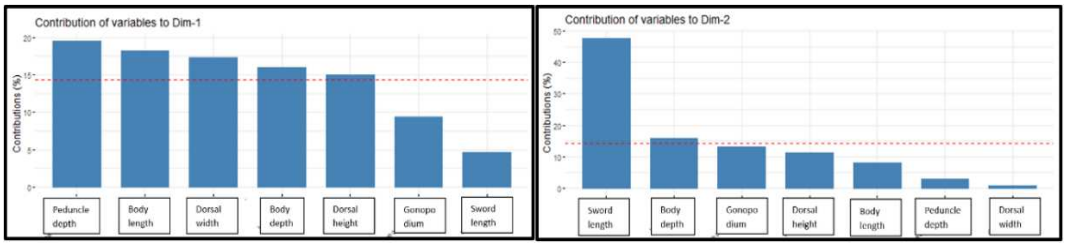


Fig. 23

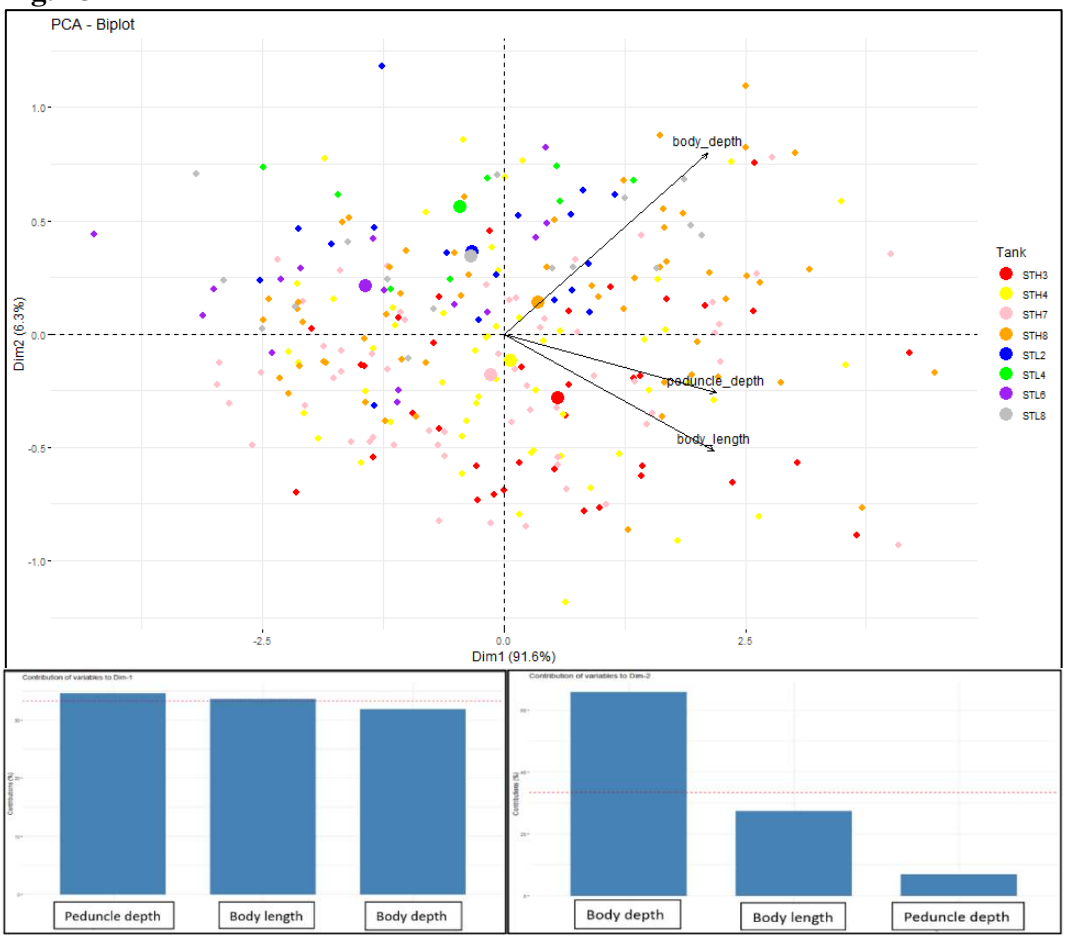


Fig. 24

5. DISCUSSION AND CONCLUSIONS

As stated in the introduction, my work has the main goal to verify whether there are significant differences between tanks and between sites. The first analyses conducted, ANOVA and Welch's ANOVA on quantitative characters, provided the following results:

- for 5 characters (male's and female's standard length, female's peduncle depth, male's dorsal width and height) I found significant differences beetwin sites but not within;
- for 2 characters (sword and gonopodium length) I found significant differences in both beetwin and within;
- for 1 character (female's body depth) I found significant differences within but not beetwen;
- for 2 characters (male's body and peduncle depth) I found no significant differences, neither beetwin nor within.

Where there are significant differences within (3 cases), they are always attributable to STH and not to STL. In fact, from the violin plots, it is easy to see that STH has greater variability than STL, because the kernel's density is usually more elongated in the vertical direction. Looking at the STH tanks, we see that the violin plot of STH3 very often differs in shape from the others, and sometimes even STH8. Furthermore, STH3 achieves higher average lengths for all characters. STH4 and STH7 are always quite similar. To test which tanks are responsible for differentiation within sites, post-hoc tests could be performed. The STL tanks are quite homogeneous, except, in some cases, STL6. STL2 and STL4 are always very similar. As we said in chapter 1, the differences within sites are attributable to genetic drift and sexual selection; unfortunately, with the processing I carried out, we are not able to distinguish them. If it is true that one of the tanks has different characteristics from the others, it's possible that a part of the population owning unusual characteristics had randomly segregated in it. Noting that STH3 has a smaller sample size than the other tanks (63 individuals), we can hypothesize that the best explanation is drift. This is a random event, so it could have occurred in some tanks only, and not in the others; probably no event of this type occurred in STL (except perhaps, with limited effects, for STL6), therefore no significant differences within sites are found. This confirms that the effect of genetic drift is present, as already demonstrated by Bovio. If, however, in his work, more differences were found within than between, I instead found the opposite: we have, overall, 7 cases of significant differences between sites and 3 within. Furthermore, the latter could actually be reduced to 2, given the particularity of female's body depth: the measurements may have been heavily influenced by the fact that many females are pregnant; this character undergoes large variations during the life of each individual, so it is not the ideal model for drawing theoretical conclusions. The two non-linear regression methods, Logistic regression and Linear discriminant analysis, also give us an indirect demonstration of the fact that the differences between the two sites are pronounced: the construction of predictive models capable of classifying single individuals with a reasonable precision is only possible if the two categorical variables have their own characteristics, sufficiently diversified to make them easily distinguishable. We know that the differences between different altitudes must be driven by different environmental conditions, mainly by temperature, therefore, we can say that we have observed a greater effect of ecological selection than that of drift. As we said in paragraph 2.3, this result is absolutely not in contrast with that of Bovio, as his research focused on early hybrid generations; on the contrary, we have demonstrated that evolutionary forces act according to different models and change over time: while genetic drift has a very pronounced immediate effect on the generations immediately following hybridization, ecological selection requires a greater number of generations to make its own effects evident. Once sufficient time has passed, ecological selection takes on a great strength, probably greater than genetic drift, which is however detectable in the persistence of some initial conditions, present since the division of the fish into the tanks. This persistence is much more evident when we look at qualitative traits: they seem to be more affected by the effects of genetic drift than quantitative traits. We have found that, generally, melanistic traits are more expressed at high altitudes, as shown by the percentages in Table 6. This finding, due to the greater incidence of Cb and sword's stripes, allows us to hypothesize a strong effect of ecological selection, which is certainly present at the level of differentiation of the two sites. If we look at the level of

the individual tanks within the same site, however, we notice that the situation is varied. Taking female's Cb as example, we see that all STH tanks have an incidence percentage greater than 50% and all STL tanks show a percentage lower than this. This is certainly due to divergent ecological selection. However, if we compare the percentages of tanks within the same site, we see that the differences are still pronounced (e.g. for female's Cb: STL6 = 40% vs STL8 = 18.8%, for male's Cb: STH3 = 93.3% vs STH7 = 57.7 %). But, it is the asymmetric distribution of Spotted caudal and Carbomaculatus that most highlights the effect of genetic drift: Sc is limited to two tanks only, one for each altitude, and with two very different forms of spot, as underlined in paragraph 4.2.3; Cm is even limited to a single tank, STH7. This distribution is certainly due to the random segregation of these characters in certain tanks and not in the others, following the arbitrary distribution of carrier individuals at the time of construction of the tanks themselves. The results of the X-squared tests highlight a rather varied situation. For Upper sword edge there is a strong dependence of this character both on the site and on the tank of belonging, therefore we cannot exclude genetic drift, nor even sexual selection, as the sword is strongly connected to the mate choice. This trait was found to be a useful predictor for both logistic regression and linear discriminant analysis, therefore, whatever the selective pressure most involved (or probably the combination of the same) is, upper edge's stripe is a strongly characterizing aspect for the categorical variable, on which it applies it is worth concentrating research efforts. On the contrary, Lower sword edge, proves to be independent of membership, as it is widely spread. I hypothesize that females of *X. malinche* have a strong preference for this trait and that, in the parental species, it has been subjected to sexual selection for a long enough time to become so widespread that it can be considered a "default" trait. Males completely devoid of sword's stripes are likely to be severely disadvantaged. It certainly would not have proven to be a useful predictor, so I did not use it in the two non-linear regressions. For Sc/Cb/None, as we had already deduced simply from the percentages, we have a pronounced effect of both ecological selection and drift: for both males and females we have dependence between the site and the character, as we expect from the former; at the level of the individual tanks, however, we have a jagged pattern, as we expect from the second one, in which males and females even demonstrate reversed situations. Unexpectedly, this trait proved to be a significant predictor only for females in the logistic regression. The X-squared tests for Cm (and also for Sc, if it were not included in the analysis together with Cb) are superfluous, the distribution alone is sufficient to demonstrate the predominant effect of random segregation.

To summarize, my research highlights the persistence over time of the effects of genetic drift, more evidently for qualitative characters, and at the same time, from the comparison with Bovio's work, the emergence and gaining greater importance of ecological selection with the passing of generations, especially for quantitative traits.

Another question we can focus on, is the direction in which phenotypes are pushed by ecological selection. We might have expected that the high-altitude population would tend to resemble *X. malinche* more, while the low-altitude population would tend towards the characteristics of *X. birchmanni* . In support of this prediction, as we have already mentioned in paragraph 2.3, Bovio found the environment drives morphology to resemble the lowland parental species in the lowland natural hybrid populations.

We know that *X. birchmanni* has greater body depth and a more pronounced dorsal fin, but it does not have a sword. Therefore, we might have expected to observe a greater

average length of the sword and a smaller size of the dorsal fin and body, at least in depth, at high altitudes. This expectation was satisfied for the sword, but not for the other characters; indeed, on the contrary, the average body dimensions, sometimes slightly (and not significantly according to the ANOVA tests) and sometimes more evidently, are greater at high altitude for all characters, as can be seen from tables 4 and 5. The explanation I suggest, is based on the presence of multiple selective pressures in the wild. We must, first of all, emphasize the fact that Bovio's statement is taken from observations on natural populations, while our results are obtaiend on artificial tanks under controlled conditions; we must not forget that our work can only highlight that microevolution driven by the temperature difference, caused by altitude. In nature, those individuals that resemble the parental species are actually favoured, because the latters respond better to all the selective pressures to which they are subjected, being adapted to their environment. In artificial stocktank, populations do not have to respond to all the complex stimuli (such as predation) present in a natural environment, but find themselves in a simplified situation; rather than adapting to the parental phenotype, it is convenient to follow Bergmann's rule (1847), according to which body mass is inversely proportional to temperature. This depends on the fact that animals with a smaller surface/volume ratio are favored by colder climates, because they disperse heat more slowly. Therefore, this observation does absolutely not compromise the role of ecological selection, on the contrary, it strengthens the hypothesis that it is the main evolutionary force that acted in our experimental context. Having stated that all traits have larger average dimensions at the high site, the subsequent question could be: do all quantitative traits follow Bergmann's rule, or do they have independent trajectories? This is equivalent to asking whether all the characters are so strongly correlated as to form a single cluster that moves together in evolution, changing morphology on the basis of a general increase in body size. Let us therefore try to evaluate the correlation between the different characters. In all the analyzed cases, ANCOVA and Nested ANCOVA confirm the significance of the differences previously identified, even after correction for the standard length covariate. The correlation between the latter and the trait under analysis was significant in all cases except one: the sword, both at the level of sites and STH's tanks, as reported in paragraph 4.1.4. Although it would be appropriate to perform linear regression analyses to confirm this assumption, we have an indication of the presence of two distinct selectable units: all the traits I analyzed except the sword, seem to be positively correlated, therefore they must evolve simultaneously in the same direction, participating in a general increase in body size. Sword, instead, does not seem to be necessarily related to this cluster, so it has the possibility to evolve independently, following its own path. The fact that it also increases in size at the highest site should not lead us to think that it is related to size like the others. Probably, swords are more sensitive to a greater number of factors, such as the trade-off of available resources, sexual selection, drift (individual variability must have been greater within parental species and this may increase the likelihood of differential distribution in tanks); consequently it is the trait least constrained by size. As anticipated in paragraph 4.3, I left the interpretation of PCA at this point, because it can provide us with a summary and can give graphic, rather intuitive, confirmation of what has been highlighted by the previous analyses. I constructed the graphs of figures 23 and 24 only with quantitative traits, because I wanted to use those characters that most highlight the action of ecological selection.

The points on the graph with warm colors represent individuals from high-altitude tanks, while cool colors represent low-altitude tanks. The points of greatest diameter represent the group means, while the vectors indicate in which direction each measure moves the individuals on the plane. The angle between two vectors indicates the correlation between two variables: a very small angle indicates that the two characters are positively correlated, 180° indicates negative correlation, 90° indicates no correlation. For males, we see that the group means of the tanks located on the same site are closer together than those of the other site. This provides the same information as ANOVA tests, meaning there are more differences between sites than within, supporting the model in which ecological selection is the primary evolutionary force. Those of STL are more crowded together (especially along the horizontal axis of the main dimension), while within STH we see that STH3 and STH8 are distanced from STH4 and STH7. STH3 is more spaced horizontally, while ST8 is more spaced vertically; given that the main differences must be observed along the X-axis, which explains 66.2% of the differences, STH3 is the most differentiated tank. This is consistent with the fact that we found significant differences within STH and not within STL. The body length, dorsal width, body depth and peduncle depth vectors are separated by small angles, so they are strongly correlated with each other.
The sword, respect to the size cluster, is placed at approximately 90°, therefore it is not correlated with those traits, as we have already highlighted with ANCOVA.
Curiously, dorsal height and gonopodium length are found in an intermediate position. I would have expected dorsal width and dorsal height to be strongly correlated, but the latter seems to move independently. This different direction of the dorsal height can also be read in the negative coefficient of the logistic regression and linear discriminant analyses, but was not highlighted by ANCOVA. If there were a correlation between gonopodium and sword, it could be explained by the involvement of the testosterone pathway and the msx genes. This result is quite intriguing, it would be interesting to investigate the link between dorsal height, gonopodium and the other characters.
For females, we see that dimension 1 accounts for 91.6%, so, distances on the horizontal axis are by far the most important to consider. In this sense we see that all the group means of the STH tanks are at higher values, while the STL tanks are at negative values. Although it is less important, we see that even vertically the two sites are spaced apart, with STH presenting the group means for the various tanks at lower values. Even for females, therefore, we can confirm a greater similarity between tanks at the same altitude. Here too, we see that STH3 and STH8 are quite far from STH4 and STH7, confirming what was seen among males. Among the low-altitude tanks, STL6 is far apart along the main axis from the other three, which are very close together. Despite this, ANOVA showed significant differences within sites only in body depth within STH. We can also see that body depth is a particular character from the angles that separate vectors: body length and peduncle depth are quite close together, while body depth seems not to be correlated with these two. In fact, even in the Logistic regression and LDA we see a negative coefficient and, as already underlined, it was responsible for the only case of significant differences within, but not between, sites. This non-correlation is present among females, but is not evident for males, so I can hypothesize that it is exclusively due to pregnancy. This suggests to not taking this female character into high consideration in research.

To end the work with a single statement, the experimental model demonstrates the importance of altitude, and therefore temperature, as an agent of ecological selection, especially in modeling body morphology. However, we must remember that, in a natural system, the selective pressures are multiple and populations must respond to different needs simultaneously, refining the phenotypic expression in a more complex way; our experimental system cannot take into consideration all the variables present in the wild, but knowing deeply the mechanisms of each process can provide valuable information for creating more "realistic" models, which include the interactions between them.

6. BIBLIOGRAPHY

- Basolo A. L. (1998), "Shift in investment between sexually selected traits: tarnishing of the silver spoon", Animal behaviour, https://doi.org/10.1006/anbe.1997.0634

- Bovio R. S. (2022), "Microevolution in natural and experimental hybrid populations of swordtailfish", Dissertation, Texas A&M University

- Bovio R. S., Rosenthal G. G. (2019), ELEH_master_protocol document, uploaded on ELEH Drive Workspace

- Culumber Z., Shepard D., Coleman S., Rosenthal G., Tobler M. (2012), "Physiological adaptation along environmental gradients and replicated hybrid zone structure in swordtails (teleostei: *Xiphophorus*)", Journal of evolutionary biology, PMID: 22827312, DOI: 10.1111/j.1420-9101.2012.02562.x

- Culumber Z. W. (2014), "Pigmentation in *Xiphophorus*: an emerging system in ecological and evolutionary genetics", Zebrafish, DOI: 10.1089/zeb.2013.0939

- Darwin C. (1859), "On the origins of species by means of natural selection, or the preservation of favoured races in the struggle for life", London: John Murray.

- Darwin C. (1871), "The descent of man, and selection in relation to sex", London: John Murray.

- Fernandez A. A., Morris M. R. (2008), "Mate choice for more melanin as a mechanism to maintain a functional oncogene", PNAS, DOI: 10.1073/pnas.0803851105

- Fisher H. S., Wong B. B., Rosenthal G. G. (2006), "Alteration of the chemical environment disrupts communication in a freshwater fish", Proceedings of the Royal Society B: Biological Sciences, doi:10.1098/rspb.2005.3406

- Fisher H. S., Rosenthal G.G. (2007), "Male swordtails court with an audience in mind", Biology letters, doi:10.1098/rsbl.2006.0556

- Giacomello E., Neat F.C., Rasotto M.B. (2007), "Mechanisms enabling sperm economy in blenniid fishes", Behavioral Ecology and Sociobiology, DOI 10.1007/s00265-007-0491-2

- Gould S. J., Lewontin R. C. (1979), "The spandrels of San Marco and the panglossian paradigm: a critique of the adaptationist programme", Proceedings of the Royal Society B: Biological Sciences, doi:10.1098/rspb.1979.0086

-Heckel, J. J. (1848). "Eine neue Gattung von Poecilien mit rochenartigem Anklammerungs-Organe", Sitzungsberichte der Kaiserlichen Akademie der

Wissenschaften. Mathematisch-Naturwissenschaftliche
- Kondo S., Miura T. (2010), "Reaction-diffusion model as a framework for understanding biological pattern formation", Science, DOI: 10.1126/science.1179047

- Kruesi K., Alcaraz G. (2007), "Does a sexually selected trait represent a burden in locomotion?", Journal of Fish Biology, DOI:10.1111/j.1095-8649.2007.01379.x

- Kruesi K., Rosenthal G. G., Alcaraz G. (2010), "Growth and male ornamentation in *Xiphophorus montezumae"*, Marine and Freshwater Behaviour and Physiology, DOI: 10.1080/10236244.2011.598644

- Lewontin R. C., Birch L. (1966), "Hybridization as a source of variation for adaptation to new environments", Evolution, PMID: 28562982, DOI: 10.1111/j.1558-646.1966.tb03369.x

- Powell D. L., García-Olazábal M., Keegan M., Reilly P., Du K., Díaz-Loyo A. P., Banerjee S., Blakkan D., Reich D., Andolfatto P., Rosenthal G. G., Schartl M., Schumer M. (2020), "Natural hybridization reveals incompatible alleles that cause melanoma in swordtail fish", Science, doi:10.1126/science.aba5216.

- Powell D. L., Moran B., Kim B., Banerjee S. M., Aguillon S. M., Fascinetto-Zago P., Langdon Q., Schumer M. (2021), "Two new hybrid populations expand the swordtail hybridization model system", Evolution, doi:10.1111/evo.14337.

- Rosenthal G.G. (2017a), "Mate choice- the evolution of sexual decision making from microbes to humans", Princeton university press, pp. 88

- Rosenthal G.G. (2017b), "Swordtails and Platyfishes", Reference Module in Life Sciences, Elsevier, ISBN: 978-0-12-809633-8, http://dx.doi.org/10.1016/B978-0-12-809633-8.01213-9

- Rosenthal G. G., Evans C. S. (1998), "Female preference for swords in *Xiphophorus helleri* reflects a bias for large apparent size", PNAS, DOI: 10.1073/pnas.95.8.4431

- Rosenthal G. G., Wagner W. E., Ryan M. J. (2002), "Secondary reduction of preference for the sword ornament in the pygmy swordtail *Xiphophorus nigrensis* (Pisces: Poeciliidae)", Animal behaviour, doi:10.1006/anbe.2001.1887

- Scaggiante M., Mazzoldi C., Petersen C.W., Rasotto M.B. (1999), "Sperm competition and mode of fertilization in the grass goby *Zosterisessor ophiocephalus* (Teleostei: Gobiidae)", Journal of experimental zoology, https://hdl.handle.net/11577/2459157

- Serena G. (2012), "Aspetti cognitivi della scelta sessuale in *Poecilia reticulata*", PhD in psychological sciences- psychobiology address, University of Padova, https://hdl.handle.net/11577/3425817

- Schumer M., Cui R., Powell D. L., Dresner R., Rosenthal G. G., Andolfatto P. (2014), "High-resolution mapping reveals hundreds of genetic incompatibilities in hybridizing fish species", Elife, PMID: 24898754, DOI: 10.7554/eLife.02535

- Smith C. C., Ryan M. J. (2010), "Evolution of sperm quality but not quantity in the internally fertilized fish *Xiphophorus nigrensis"*, Journal of evolutionary biology, DOI: 10.1111/j.1420-9101.2010.02041.x

- Trainor B. C., Basolo, A. L. (2005), "Location, location, location: stripe position effects on female sword preference", Animal Behaviour, doi:10.1016/j.anbehav.2005.04.007

- Turing A. (1952), "The Chemical Basis of Morphogenesis", Philosophical Transactions of the Royal Society of London B, DOI:10.1098/rstb.1952.0012

- Wong B. B. M., G. G. Rosenthal (2006), “Female Disdain for Swords in a Swordtail Fish”, The american naturalist, DOI: 10.1086/498278

- Zauner H., Begemann G., Marí-Beffa M., Meyer A. (2003), “Differential regulation of msx genes in the development of the gonopodium, an intromittent organ, and of the "sword", a sexually selected trait of swordtail fishes (*Xiphophorus*)”, Evolution & Development, DOI: 10.1046/j.1525-142x.2003.03053.x

7. APPENDIX: ALL PROCESSING IN R SOFTWARE

1) Sc/Cb/None, male

```
> observed

    STH STL
Cb  64  16
No  16  19
Sc   4   6
> X$expected

      STH   STL
Cb 53.76 26.24
No 23.52 11.48
Sc  6.72  3.28
> X

        Pearson's Chi-squared test

data:  observed
X-squared = 16.633, df = 2, p-value = 0.0002444

> X$residuals

          STH       STL
Cb  1.396594 -1.999024
No -1.550598  2.219458
Sc -1.049263  1.501869
```

1

```
> observed

    3  4  7  8
Cb 14 12 15 23
No  1  4  8  3
Sc  0  1  3  0
> X$expected

             3         4         7         8
Cb 11.4285714 12.9523810 19.809524 19.809524
No  2.8571429  3.2380952  4.952381  4.952381
Sc  0.7142857  0.8095238  1.238095  1.238095
> X

        Pearson's Chi-squared test

data:  observed
X-squared = 10.866, df = 6, p-value = 0.0926

> X$residuals

             3         4         7         8
Cb  0.7606388 -0.2646281 -1.0806002  0.7168338
No -1.0987005  0.4234049  1.3694736 -0.8773190
Sc -0.8451543  0.2117024  1.5834538 -1.1126973
```

2

```
> observed

    2 4 6 8
Cb 5 7 1 3
No 3 4 3 9
Sc 0 0 6 0
> X$expected

             2         4         6         8
Cb 3.121951 4.292683 3.902439 4.682927
No 3.707317 5.097561 4.634146 5.560976
Sc 1.170732 1.609756 1.463415 1.756098
> X

        Pearson's Chi-squared test

data:  observed
X-squared = 27.275, df = 6, p-value = 0.0001286

> X$residuals

             2         4         6         8
Cb  1.0629034  1.3066965 -1.4692478 -0.7776901
No -0.3673536 -0.4861244 -0.7591124  1.4583433
Sc -1.0820036 -1.2687616  3.7501219 -1.3251783
```

3

1-Chi square test between sites; 2-Chi square test between STH tanks; 3-Chi square test between STL tanks

2) Sc/Cb/None, female

```
> observed

    STH STL
Cb 119  17
No  81  39
Sc   7   0
> X$expected

          STH       STL
Cb 107.041825 28.958175
No  94.448669 25.551331
Sc   5.509506  1.490494
> X

        Pearson's Chi-squared test

data:  tabobs
X-squared = 17.161, df = 2, p-value = 0.0001877

> X$residuals

          STH       STL
Cb  1.1558145 -2.2221804
No -1.3838257  2.6605569
Sc  0.6350004 -1.2208580
```

4

```
> observed

    3  4  7  8
Cb 24 34 28 33
No 12 20 25 24
Sc  0  0  7  0
> X$expected

             3         4         7         8
Cb 20.695652 31.043478 34.492754 32.768116
No 14.086957 21.130435 23.478261 22.304348
Sc  1.217391  1.826087  2.028986  1.927536
> X

        Pearson's Chi-squared test

data:  observed
X-squared = 19.78, df = 6, p-value = 0.00303

> X$residuals

             3          4          7          8
Cb  0.72635043  0.53063500 -1.10551589  0.04050841
No -0.55603844 -0.24591855  0.31405611  0.35903931
Sc -1.10335457 -1.35132785  3.48984033 -1.38835739
```

5

```
> observed

    2  4  6  8
Cb  5  3  6  3
No 12  5  9 13
> X$expected

             2         4         6         8
Cb  5.160714 2.428571  4.553571  4.857143
No 11.839286 5.571429 10.446429 11.142857
> X

        Pearson's Chi-squared test

data:  observed
X-squared = 1.8796, df = 3, p-value = 0.5978

> X$residuals

              2          4          6          8
Cb -0.07074562  0.36667940  0.67783020 -0.84266484
No  0.04670805 -0.24209101 -0.44752064  0.55634864
```

6

4-Chi square test between sites; 5-Chi square test between STH tank; 6-Chi square test between STL tank

3) At\Cm, male

```
> observed

       STH STL
  At/Cm   3   0
  No     81  41
> X$expected

          STH    STL
  At/Cm  2.016  0.984
  No    81.984 40.016
> X

        Pearson's Chi-squared test with Yates'

data:  observed
X-squared = 0.36297, df = 1, p-value = 0.5469

> X$residuals

              STH        STL
  At/Cm  0.6930265 -0.9919677
  No    -0.1086752  0.1555530
```

7

```
> observed

        3  4  7  8
  At/Cm 0  0  3  0
  No    15 17 23 26
> X$expected

                3         4         7         8
  At/Cm  0.5357143  0.6071429  0.9285714  0.9285714
  No    14.4642857 16.3928571 25.0714286 25.0714286
> X

        Pearson's Chi-squared test

data:  observed
X-squared = 6.9402, df = 3, p-value = 0.07383

> X$residuals

                3         4         7         8
  At/Cm -0.7319251 -0.7791937  2.1496230 -0.9636241
  No     0.1408590  0.1499559 -0.4136951  0.1854495
```

8

```
> observed

     2  4  6  8
  No  8 11 10 12
> X$expected
[1] 10.25 10.25 10.25 10.25
> X

        Chi-squared test for given probabilities

data:  observed
X-squared = 0.85366, df = 3, p-value = 0.8366

> X$residuals
[1] -0.70278193  0.23426064 -0.07808688  0.54660817
```

9

7-Chi square test between sites; 8-Chi square test between STH tank; 9-Chi square test between STL tank

4) At/Cm, None female

```
> observed

       STH STL
  At/Cm  11   0
  No    196  56
> X$expected

             STH       STL
  At/Cm   8.657795  2.342205
  No    198.342205 53.657795
> X

        Pearson's Chi-squared test with Yates'

data:  observed
X-squared = 1.9213, df = 1, p-value = 0.1657

> X$residuals

              STH        STL
  At/Cm  0.7960151 -1.5304265
  No    -0.1663096  0.3197485
```

10

```
> observed

        3  4  7  8
  At/Cm 0  0 11  0
  No    36 54 49 57
> X$expected

                3         4         7         8
  At/Cm  1.913043  2.869565  3.188406  3.028986
  No    34.086957 51.130435 56.811594 53.971014
> X

        Pearson's Chi-squared test

data:  observed
X-squared = 28.462, df = 3, p-value = 2.904e-06

> X$residuals

                3         4         7         8
  At/Cm -1.3831281 -1.6939791  4.3747464 -1.7403981
  No     0.3276655  0.4013067 -1.0363852  0.4123034
```

11

```
> observed

     2  4  6  8
  No 17  8 15 16
> X$expected
[1] 14 14 14 14
> X

        Chi-squared test for given probabilities

data:  observed
X-squared = 3.5714, df = 3, p-value = 0.3116

> X$residuals
[1]  0.8017837 -1.6035675  0.2672612  0.5345225
```

12

10-Chi square test between sites; 11-Chi square test between STH tank; 12-Chi square test between STL tank

5)Upper sword edge, male

```
> observed

    STH STL
  0  31  28
  1  53  13
> X$expected

       STH    STL
  0 39.648 19.352
  1 44.352 21.648
> X

        Pearson's Chi-squared test with Yates'

data:  observed
X-squared = 9.6688, df = 1, p-value = 0.001874

> X$residuals

          STH       STL
  0 -1.373425  1.965861
  1  1.298551 -1.858689
```

13

```
> observed

    3  4  7  8
  0  1  4 14 12
  1 14 13 12 14
> X$expected

          3        4         7         8
  0 5.535714  6.27381  9.595238  9.595238
  1 9.464286 10.72619 16.404762 16.404762
> X

        Pearson's Chi-squared test

data:  observed
X-squared = 11.356, df = 3, p-value = 0.009948

> X$residuals

                3          4          7          8
  0 -1.9277862 -0.9077963  1.4219835  0.7763261
  1  1.4743540  0.6942747 -1.0875206 -0.5937274
```

14

```
> observed

    2  4  6  8
  0 5  4  9 10
  1 3  7  1  2
> X$expected

          2        4        6        8
  0 5.463415 7.512195 6.829268 8.195122
  1 2.536585 3.487805 3.170732 3.804878
> X

        Pearson's Chi-squared test

data:  observed
X-squared = 8.7326, df = 3, p-value = 0.03307

> X$residuals

                2          4          6          8
  0 -0.1982612 -1.2814309  0.8306519  0.6304783
  1  0.2909678  1.8806265 -1.2190637 -0.9252892
```

15

13-Chi square test between sites; 14-Chi square test between STH tank; 15-Chi square test between STL tank

6) Lower sword edge, male

```
> observed
   STH STL
0   8  10
1  76  31
> X$expected
      STH    STL
0 12.096  5.904
1 71.904 35.096
> X

        Pearson's Chi-squared test with Yates'

data:  observed
X-squared = 3.8076, df = 1, p-value = 0.05102

> X$residuals
          STH        STL
0 -1.1777119  1.6857252
1  0.4830404 -0.6914028
```

16

```
> observed
    3  4  7  8
0   0  3  4  1
1  15 14 22 25
> X$expected
          3        4        7        8
0  1.428571  1.619048  2.47619  2.47619
1 13.571429 15.380952 23.52381 23.52381
> X

        Pearson's Chi-squared test

data:  observed
X-squared = 4.8899, df = 3, p-value = 0.18

> X$residuals
          3         4         7         8
0 -1.1952286  1.0852977  0.9683641 -0.9381027
1  0.3877834 -0.3521171 -0.3141788  0.3043607
```

17

```
> observed
    2  4  6  8
0   1  0  4  5
1   7 11  6  7
> X$expected
         2        4        6        8
0 1.95122 2.682927 2.439024 2.926829
1 6.04878 8.317073 7.560976 9.073171
> X

        Pearson's Chi-squared test

data:  observed
X-squared = 7.4252, df = 3, p-value = 0.05951

> X$residuals
          2         4         6         8
0 -0.6809695 -1.6379642  0.9995121  1.2118151
1  0.3867647  0.9303012 -0.5676847 -0.6882648
```

18

16-Chi square test between sites; 17-Chi square test between STH tank; 18-Chi square test between STL tank

7) Body length, male

```
Levene's Test for Homogeneity of Variance (center = mean)
      Df F value  Pr(>F)
group  1  4.9924 0.02727 *
```

19

```
        One-way analysis of means (not assuming equal variances)

data:  tuttiM$body_length and tuttiM$site
F = 4.401, num df = 1.00, denom df = 104.41, p-value = 0.03833
```

```
        One-way analysis of means (not assuming equal variances)

data:  tuttiMH_welch_tra_tank_in_site_$body_length and tuttiMH_welch_tra_tank_in_site_$stocktank
F = 1.1614, num df = 3.000, denom df = 40.204, p-value = 0.3364
```

```
        One-way analysis of means (not assuming equal variances)

data:  tuttiML_welch_tra_tank_in_site_$body_length and tuttiML_welch_tra_tank_in_site_$stocktank
F = 0.12132, num df = 3.000, denom df = 20.128, p-value = 0.9465
```

20

19- Levene's test; 20- Welch Anova between sites and between tanks at the same altitude

8) Body length, female

```
Levene's Test for Homogeneity of Variance (center = mean)
      Df F value Pr(>F)
group  1   2.639 0.1055
```

21

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
tuttiF$site	1	663	663.4	21.375	6.01e-06
tuttiF$site:tuttiF$tank	6	392	65.3	2.105	0.0532
Residuals	255	7914	31.0		

22

21-Levene test; 22-Nested Anova

9) Body depth, male

```
Levene's Test for Homogeneity of Variance (center = mean)
      Df F value  Pr(>F)
group  1  2.7571 0.09937 .
     123
```

23

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
tuttiM$site	1	0.5	0.491	0.148	0.701
tuttiM$site:tuttiM$tank	6	10.7	1.786	0.539	0.778
Residuals	117	388.0	3.316		

24

23-Levene test; 24-Nested Anova

10) Body depth, female

```
Levene's Test for Homogeneity of Variance (center = mean)
      Df F value Pr(>F)
group  1  0.1783 0.6732
```
25

```
                    Df Sum Sq Mean Sq F value Pr(>F)
tuttiF$site          1    1.2   1.185   0.414 0.5207
tuttiF$site:tuttiF$tank  6   45.7   7.608   2.655 0.0163
Residuals          255  730.9   2.866
```
26

```
            Df Sum Sq Mean Sq F value Pr(>F)
tank         3  19.28   6.427   2.712 0.0543
Residuals   52 123.22   2.370
```
27

```
            Df Sum Sq Mean Sq F value Pr(>F)
tank         3   26.4   8.790   2.937 0.0344
Residuals  203  607.6   2.993
```
28

```
            Df Sum Sq Mean Sq F value Pr(>F)
site         1    1.2     1.2     1.4  0.238
body_length  1  556.4   556.4   657.1 <2e-16
Residuals  260  220.1     0.8
```
29

```
            Df Sum Sq Mean Sq F value   Pr(>F)
tank         3   26.4     8.8   10.84 1.23e-06
body_length  1  443.9   443.9  547.59  < 2e-16
Residuals  202  163.7     0.8
```
30

```
            Df Sum Sq Mean Sq F value   Pr(>F)
site         1    1.2     1.2   1.608    0.206
body_length  1  556.4   556.4 754.888  < 2e-16
site:tank    6   32.9     5.5   7.445 2.34e-07
Residuals  254  187.2     0.7
```
31

25-Levene test; 26-Nested Anova; 27-Anova between STL tank; 28-Anova between STH tank; 29-Ancova between sites; 30-Ancova between STH tank; 31- Nested Ancova

11) Peduncle depth, male

```
Levene's Test for Homogeneity of Variance (center = mean)
      Df F value Pr(>F)
group  1  2.7084 0.1024
```
32

```
                    Df Sum Sq Mean Sq F value Pr(>F)
tuttiM$site          1   5.09   5.092   3.268 0.0732
tuttiM$site:tuttiM$tank  6   2.01   0.336   0.215 0.9713
Residuals          117 182.29   1.558
```
33

32-Levene test; 33-Nested Anova

12) Peduncle depth, female

```
Levene's Test for Homogeneity of Variance (center = mean)
      Df F value Pr(>F)
group  1  0.6747 0.4122
```
34

```
                    Df Sum Sq Mean Sq F value   Pr(>F)
tuttiF$site          1  19.12  19.120  19.334 1.61e-05
tuttiF$site:tuttiF$tank  6   6.66   1.110   1.123    0.349
Residuals          255 252.17   0.989
```
35

```
            Df Sum Sq Mean Sq F value Pr(>F)
site         1  19.12   19.12   137.1 <2e-16
body_length  1 222.58  222.58  1596.0 <2e-16
Residuals  260  36.26    0.14
```
36

```
            Df Sum Sq Mean Sq  F value   Pr(>F)
site         1  19.12   19.12  147.051  < 2e-16
body_length  1 222.58  222.58 1711.848  < 2e-16
site:tank    6   3.23    0.54    4.145 0.000542
Residuals  254  33.03    0.13
```
37

34-Levene test; 35-Nested Anova; 36-Ancova between STH sites; 37-Nested Ancova

13) Sword length, male

```
Levene's Test for Homogeneity of Variance (center = mean)
      Df F value Pr(>F)
group  1   3e-04 0.9853
```

38

```
                Df Sum Sq Mean Sq F value   Pr(>F)
tuttiM$site      1  133.4  133.43  18.515 3.52e-05
tuttiM$site:tuttiM$tank  6  112.7   18.78   2.606   0.0209
Residuals      117  843.2    7.21
```

39

```
            Df Sum Sq Mean Sq F value Pr(>F)
tank         3  52.12  17.375   2.666 0.0619
Residuals   37 241.10   6.516
```

40

```
            Df Sum Sq Mean Sq F value   Pr(>F)
tank         7  248.4   35.48   4.894 7.15e-05
Residuals  117  848.3    7.25
```

41

```
            Df Sum Sq Mean Sq F value   Pr(>F)
site         1  133.4  133.43  17.554 5.31e-05
body_length  1   28.6   28.56   3.758   0.0549
Residuals  122  927.3    7.60
```

42

```
            Df Sum Sq Mean Sq F value   Pr(>F)
tank         7  248.4   35.48   5.004 5.59e-05
body_length  1   25.8   25.79   3.637    0.059
Residuals  116  822.5    7.09
```

43

```
            Df Sum Sq Mean Sq F value   Pr(>F)
site         1  133.4  133.43  18.916 2.95e-05
body_length  1   28.6   28.56   4.049   0.0465
site:tank    6  109.1   18.18   2.577   0.0222
Residuals  116  818.3    7.05
```

44

38-Levene test; 39-Nested Anova; 40-Anova between STL tank; 41-Anova between STH sites; 42-Ancova between sites; 43-Ancova between STH sites; 44-Nested Ancova

14) Dorsal width, male

```
Levene's Test for Homogeneity of Variance (center = mean)
      Df F value Pr(>F)
group  1  1.1898 0.2775
```

45

```
                Df Sum Sq Mean Sq F value   Pr(>F)
tuttiM$site      1   61.4   61.42  13.786 0.000315
tuttiM$site:tuttiM$tank  6   15.0    2.51   0.563 0.759144
Residuals      117  521.3    4.46
```

46

```
            Df Sum Sq Mean Sq F value   Pr(>F)
site         1   61.4    61.4   43.53 1.13e-09
body_length  1  364.2   364.2  258.09  < 2e-16
Residuals  122  172.1     1.4
```

47

```
            Df Sum Sq Mean Sq F value   Pr(>F)
site         1   61.4    61.4  45.953 5.34e-10
body_length  1  364.2   364.2 272.453  < 2e-16
site:tank    6   17.1     2.8   2.131   0.0549
Residuals  116  155.0     1.3
```

48

45-Levene test; 46-Nested Anova; 47-Ancova between sites; 48-Nested Ancova

15) Dorsal height, male

```
Levene's Test for Homogeneity of Variance (center = mean)
      Df F value Pr(>F)
group  1   0.423 0.5167
```

49

```
                Df Sum Sq Mean Sq F value   Pr(>F)
tuttiM$site      1   59.8   59.83  11.495 0.000952
tuttiM$site:tuttiM$tank  6   31.1    5.18   0.995 0.431825
Residuals      117  609.0    5.21
```

50

```
            Df Sum Sq Mean Sq F value   Pr(>F)
site         1   59.8   59.83   18.16 4.03e-05
body_length  1  238.0  238.00   72.22 5.62e-14
Residuals  122  402.1    3.30
```

51

```
            Df Sum Sq Mean Sq F value   Pr(>F)
site         1   59.8   59.83  19.367 2.41e-05
body_length  1  238.0  238.00  77.040 1.71e-14
site:tank    6   43.7    7.29   2.358   0.0347
Residuals  116  358.4    3.09
```

52

49-Levene test; 50-Nested Anova; 51-Ancova between sites; 52-Nested Ancova

16) Gonopodium, male

```
Levene's Test for Homogeneity of Variance (center = mean)
      Df F value Pr(>F)
group  1  1.5074 0.2219
```
53

```
                      Df Sum Sq Mean Sq F value  Pr(>F)
tuttiM$site            1  10.80  10.797   8.157 0.00508
tuttiM$site:tuttiM$tank  6  38.23   6.372   4.814 0.00020
Residuals            117 154.86   1.324
```
54

```
           Df Sum Sq Mean Sq F value Pr(>F)
tank        3   5.18   1.726   1.488  0.234
Residuals  37  42.91   1.160
```
55

```
           Df Sum Sq Mean Sq F value   Pr(>F)
tank        7   56.1   8.014   5.514 1.69e-05
Residuals 117  170.1   1.453
```
56

```
            Df Sum Sq Mean Sq F value   Pr(>F)
site         1  10.80   10.80   8.377   0.0045
body_length  1  35.84   35.84  27.807 5.89e-07
Residuals  122 157.25    1.29
```
57

```
           Df Sum Sq Mean Sq F value   Pr(>F)
tank        7  56.10    8.01   7.447 2.19e-07
body_length 1  45.23   45.23  42.028 2.28e-09
Residuals 116 124.83    1.08
```
58

```
            Df Sum Sq Mean Sq F value   Pr(>F)
site         1  10.80   10.80  11.167  0.00112
body_length  1  35.84   35.84  37.070 1.52e-08
site:tank    6  45.09    7.52   7.773 4.87e-07
Residuals  116 112.16    0.97
```
59

53-Levene test; 54-Nested Anova; 55-Anova between STL tank; 56-Anova between STH tank; 57-Ancova between sites; 58-Ancova between STH tank; 59-Nested Ancova

17) sword length, dorsal width, dorsal heigth, gonopodium length, Sc/Cb/None and upper sword edge, male

```
Call:
glm(formula = site ~ ., family = "binomial", data = mlog3)

Coefficients:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)        -7.7638     2.4868  -3.122  0.00180 **
sword_length        0.1135     0.1462   0.776  0.43747
dorsal_width        0.4482     0.2184   2.052  0.04015 *
dorsal_height      -0.2876     0.2382  -1.207  0.22725
gonopodium_length   0.4729     0.3186   1.484  0.13774
`Sc/Cb/None`        0.6854     0.4498   1.524  0.12758
upper_sword_edge    1.9358     0.6853   2.825  0.00473 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 113.676  on 81  degrees of freedom
Residual deviance:  86.799  on 75  degrees of freedom
AIC: 100.8

Number of Fisher Scoring iterations: 4
```
60

```
Call:
lda(site ~ ., data = mlog3)

Prior probabilities of groups:
  0   1
0.5 0.5

Group means:
  sword_length dorsal_width dorsal_height gonopodium_length `Sc/Cb/None` upper_sword_edge
0     3.214841     8.870122      8.148463          8.158902    0.6829268        0.3170732
1     5.167037    10.321829      9.709317          8.825195    0.8048780        0.7073171

Coefficients of linear discriminants:
                      LD1
sword_length       0.07627395
dorsal_width       0.32176630
dorsal_height     -0.20419123
gonopodium_length  0.35014614
`Sc/Cb/None`       0.53011670
upper_sword_edge   1.63136449
> p
 [1] 1 1 1 1 1 1 1 1 1 1 0 1 1 0 1 1 1 0 0 1 0 1 0 0 1 0 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 0 0 0 1 0 0 0
[50] 1 0 0 0 1 1 1 1 1 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0
Levels: 0 1
> table(p,mlog3$site)

p    0  1
  0 30 10
  1 11 31
```
61

60-logistic regression; 61-linear discriminant analysis

18) body length, body depth, peduncle depth, Sc/Cb/None, female

```
Call:
glm(formula = site ~ ., family = "binomial", data = flog3)

Coefficients:
               Estimate Std. Error z value Pr(>|z|)
(Intercept)     -4.8711     2.0498  -2.376 0.017486 *
body_length      0.1073     0.1227   0.875 0.381658
body_depth      -1.9383     0.4098  -4.729 2.25e-06 ***
peduncle_depth   3.4106     0.9391   3.632 0.000282 ***
Sc_Cb_None       1.2688     0.5312   2.389 0.016909 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 155.26  on 111  degrees of freedom
Residual deviance: 100.21  on 107  degrees of freedom
AIC: 110.21

Number of Fisher Scoring iterations: 5
```

62

```
Call:
lda(site ~ ., data = flog3)

Prior probabilities of groups:
  0   1
0.5 0.5

Group means:
  body_length body_depth peduncle_depth Sc_Cb_None
0    40.88159   12.13633       6.584705  0.3035714
1    44.08734   12.09998       7.177643  0.5535714

Coefficients of linear discriminants:
                    LD1
body_length     0.06970338
body_depth     -1.11594457
peduncle_depth  1.78799311
Sc_Cb_None      0.78687023
> p=predict(LDA, newdata=flog3[ ,c(1:4)])$class
> p
  [1] 0 1 1 1 0 1 1 1 1 1 1 0 1 0 1 1 1 1 1 0 1 0 1 1 1 1 1 1 1 1 1 0 0 0 0 1 1 0 1 0 1 1 1 1 1
 [49] 0 0 1 0 1 0 0 0 0 0 0 0 0 1 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 0
 [97] 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0
Levels: 0 1
> table(p,flog3$site)

p    0  1
  0 48 19
  1  8 37
```

63

62-logistic regression; 63-linear discriminant analysis