Evolution and genetic basis of pelvic-brooding, a derived reproductive strategy in Sulawesi ricefishes (Beloniformes: Adrianichthyidae)

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Jana M. Flury

aus

Langenthal, Schweiz

Bonn im Juli, 2022

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

Gutachter: Prof. Dr. Bernhard Misof
Gutachter: Prof. Dr. Arne W. Nolte

Tag der Promotion: 29.09.2022
Erscheinungsjahr: 2022

Summary

In Sulawesi ricefishes (Adrianichthyidae, Beloniformes) a derived reproductive strategy evolved. It is unique in the way that females of pelvic-brooding species carry the eggs externally for up to 18 days until the fry hatches. The eggs form a bundle and are connected to the female by attaching filaments which originate from the egg surface and are anchored within the females' gonoduct. Several other morphological adaptations are related to pelvic-brooding, like elongated pelvic fins and shortened ribs to form a ventral concavity. These traits are missing in transfer-brooding species, the more common and ancestral brooding strategy of the ricefishes. The most astonishing is, that pelvic-brooding appears in two polyphyletic and distantly related (> 16 my divergence) lineages of ricefishes on Sulawesi. Hence, various questions arise about its evolution: Did pelvic-brooding evolve twice or is it an ancient trait that was lost several times? Is the genetic basis independent in the two lineages or were the same mechanisms involved? In this thesis I give a general introduction to the topic in chapter 1. Here, I give an introduction in convergent evolution and the role of hybridization and introgression in this respect and introduce Sulawesi ricefishes as model to address fundamental questions in evolutionary biology. In chapter 2, I investigate the genetic architecture of pelvic-brooding traits in *Oryzias eversi* and in chapter 3, I study whether

gene flow into the stem lineage of *Oryzias* pelvic-brooding species could have fuelled its evolution. In chapter 4, I compare the sex determination of the pelvic-brooding *O. eversi* and its close relative, the transfer-brooding species *O. nigrimas,* to determine differences and similarities that might connect the sex determination with sexually dimorphic traits related to pelvic-brooding.

Chapter 2: The complexity of the genetic basis of adaptive traits has for a long time puzzled evolutionary biologists. There are several theories and some claim that a few loci with major impact are responsible for quantitative traits. In contrast, there is the infinitesimal model, saying that many loci with small effect are the basis of quantitative traits. In Sulawesi ricefishes, a reproductive strategy has evolved that is phenotypically complex. It entails elongated pelvic fins, shortened ribs and a plug that builds up in the females gonoduct to anchor an egg cluster. Using the Castle-Wright estimator, we estimated the number of loci responsible for the pelvic fin length, the rib length, and the size of the genital papilla. Additionally, we tested if these traits were correlated and used a landmark analysis, to see if different parts of the body were more correlated in comparison to other parts. We found that most traits related to pelvic-brooding were based on only a few loci and that the traits were correlated in both, pelvic-brooding and transfer-brooding species, and formed a unit. Further, geometric morphometrics revealed that body shapes were modular and certain parts were more strongly connected compared to others. This led to the conclusion that pelvic-brooding traits are based on only a few loci and that modularity had facilitated its evolution.

Chapter 3: Introgression has been proven to have the potential to serve as a fast-track for evolution by bringing in adaptive alleles from one population or species into the other. We tested, if there was gene flow into the more recent lineage of pelvic-brooding species of the Sulawesi ricefishes,

which could have fuelled the evolution of this trait. We used 1907 orthologous genes to detect species-tree – gene-tree incongruencies in a multispecies coalescence model and checked for signs of introgression using D-statistics in ~38 million SNPs. We found no gene flow between *A. oophorus* and *O. eversi* nor *O. sarasinorum*, but we detected a hybridization event between the ancestor of *O. eversi* and *O. sarasinorum* and the ancestor of the Lake Poso *Oryzias* (*O. nigrimas*, *O. nebulosus*, *O. orthognathus*) and *O. soerotoi*. Further, we confirm that pelvic-brooding was not introgressed, but evolved twice convergently. However, our results highlight that the introduced genetic variation from the hybridization between the ancestor of the Lake Poso *Oryzias* and the ancestor of todays pelvic-brooding *Oryzias* has potentially contributed to the evolution of pelvic-brooding in *Oryzias*.

Chapter 4: Sexually reproducing species usually have a genetic sex determination and they evolve under natural selection. For many groups of animals, these mechanisms are conserved, as for example in mammals and birds. However, a large variety of sex determination systems was documented for teleost fishes. In *O. latipes*, a genetic model organism, the sex determination was sustained for over 300 million years. With the radiation of the Sulawesi ricefishes starting 16 million years ago, several sex determination turnovers occurred. Here we wanted to know how similar the genetic architecture of sex determination was between a pelvic-brooding (*O. eversi*) and a closely related transfer-brooding species (*O. nigrimas*). We crossed a male *O. eversi* with a female *O. nigrimas* to receive interspecific F1 hybrids and further crossed F1 hybrids with each other to receive intercrosses. Backcrosses were bred by crossing F1 hybrid males with females of the respective parental species. The three F2 cross types were used to map the sex determining

regions. We found a significant sex determining region corresponding to the *O. latipes* chromosome (OLchr) 24 in the *O. eversi* backcrosses and the intercrosses. In the *O. nigrimas* backcrosses we found three peaks on regions corresponding to OLchr 4, 20 and 24, but these peaks did not reach significance level probably due to the low number of females in this cross. Interestingly, a recent study found that the quantitative trait locus for pelvic fin length in a F2 intercross between *O. eversi* and *O. dopingdopingensis* lies on OLchr 24, however another study has mapped the sex determining locus of *O. eversi* on OLchr 4. We discuss the findings of both studies and state, why they are not necessarily conflicting and what our implications are on the connection between a sex determination turnover and the evolution of pelvic-brooding in *Oryzias*.

Content

1. Introduction

1.1 Natural selection, adaptation and speciation

Like gravity, natural selection is a force on earth no living individual could or can escape. As a result of natural selection, organisms' characteristics serving a certain function changed over time, because they were advantageous (see Williams 1966; Futuyma 1986). Individuals carrying these (maybe only slightly) advantageous characteristics were more successful in reproduction, passing their genetic material onto the next generation more likely and more frequently (Brandon 1978). However, adaptation alone does not explain the emergence of new species (see Wagner 2014). This problem was tackled in the Modern synthesis. Four major factors were determined, that contribute to speciation: Mutation, recombination, directed selection and genetic drift (Dobzhansky 1937; Huxley 1942; Mayr 1942). Mutation and recombination bring new genetic variants into a population. Genetic variation is essentially the foundation for natural selection to work on. Directed selection then favours one over the other allele or haplotype, resulting in an augmented fraction of it in the next generation. However, also genetic drift and changing population sizes might alter allele frequencies in a population and contribute to speciation. Everything that leads to reproductive isolation between populations allows speciation, because it results in different allele frequencies in different subpopulations, serving as a new starting point for selection to act upon (Dobzhansky 1951; Baker 1959; Barton and Bengtsson 1986). Dobzhansky and Mayr went one step further and emphasized the importance of sexual isolation in the "biological species concept", based on which groups of animals only are identified as species if they reproduce exclusively with each other (Dobzhansky 1935; Mayr 1942). However, the gradualistic view underlying also the "tree of life" was disputed recently because of evidence mainly from

bacteria and microbes, but also from eukaryotes showing that the "tree" of life is rather a "network" of life: Horizontal gene transfer in microbes, incomplete lineage sorting and introgression lead to species tree – gene tree incongruencies, meaning that species boundaries are in some cases not absolute and do not need to be (Mallet et al. 2016; Shapiro et al. 2016; Fitzgerald and Rosenberg 2019; Novick and Doolittle 2019).

Further, the transmission of phenotypic traits from one into the next generation must not be exclusively genetic. One example is human height. Genetic factors determined by genome wide association studies (GWAS) were only able to explain a small proportion of the observed variation in human height (Lettre et al. 2008; Maher 2008). The other proportion might be determined by non-genetic components. Among these are for example parental-effects, ecological and cultural inheritance (reviewed in Danchin et al. 2011). These non-genetic effects were largely neglected in the theory of evolution, although they might play an important role in adaptation and speciation (Danchin et al. 2011).

Nevertheless, natural selection is still assumed to be the main driver of speciation (Shapiro et al. 2016). In sexually reproducing species, sexual selection can aid to manifest assortative mating and separate species or populations from each other (Phelan and Baker 1987; Coyne and Orr 2004). Individuals of different sexes choose to mate with certain individuals of the other sex, based on a variety of phenotypic cues (morphology, size, behaviour, odour, pheromones) that are attractive to them and individuals of the same sex compete against each other for access to the other sex (Darwin 1871). This can lead to the expression of extreme phenotypes and a preference for it in the other sex, which is known as the "Fisherian runaway" (Fisher 1930). In some groups, this

interplay of natural and sexual selection has enabled species to evolve in sympatry, which means in the exact same locality without geographic barriers (Poulton 1904; Jordan 1905; Mayr 1942; Gavrilets 2003; Coyne and Orr 2004; Bird et al. 2012). It has long puzzled evolutionary biologists how this could be possible without species boundaries breaking up due to interbreeding (reviewed in Gourbiere 2004), resulting in the disappearance of many species. Yet, some of the most rapid and species rich radiations have evolved in sympatry and in the presence of gene flow (Seehausen 2004; Papadopulos et al. 2011; Martin et al. 2013).

1.2 Hybridization as booster for evolution and speciation

Although for a long time considered to hinder speciation (Mayr 1947, 1963), or at least to be of minor importance in animals (Coyne and Orr 2004; Barton 2008), hybridization has now been shown to be an effective stimulator of speciation and particularly adaptive radiations (e.g. Seehausen 2004; Grant et al. 2005; Mallet 2007; Litsios and Salamin 2014, reviewed in Schumer et al. 2018). The introduction of genetic variants already tested by selection is much faster and safer compared to mutation (Urban et al. 2021) which is mostly expected to be neutral or deleterious (Ohta 1992). Although a single hybridization event probably has no long-term effect, if many hybridization events occur among closely related lineages, chances increase for some to bring in adaptive variation (Abbott et al. 2013) and lead to speciation (Mavárez et al. 2006; Grant and Grant 2014; Lavretsky et al. 2015; Elgvin et al. 2017; Lamichhaney et al. 2018). Hence, so called adaptive introgression, where advantageous genes are imported from one population or species into another, has been proven to be a powerful mechanism to adapt to a rapidly changing environment,

as for example during the current climate change (Hamilton and Miller 2016; Ma et al. 2019). Another example is the African malaria-transmitting mosquito species pair *Anopheles gambiae* and *A. coluzzi*, where the introgression of a large region of chromosome 2 containing insecticide-resistance genes was observed from *A. gambiae* to *A. coluzzi*. The frequency increased steadily and was interpreted as a reaction to an insecticide-treated bed net campaign in Mali (Norris et al. 2015). As an adaptation to less and less snow cover during winter, snowshoe hares possess an introgressed black-tailed jackrabbit allele leading to brown instead of white winter coats (Jones et al. 2018b). In *Heliconius* butterflies, adaptive *optix* alleles were introgressed from one species into the other in order to generate mimetic red patterns (Pardo-Diaz et al. 2012). Further, it has been shown in humans that adaptive alleles have introgressed from the Neanderthals, associated for example with pathogen-resistance (reviewed in Racimo et al. 2015). Besides adaptive alleles that introgressed by gene flow from one into the other population or species, genetic variation might also come from standing genetic variation or genetic variation that was introduced a long time ago by an ancient gene flow event. Standing genetic variation is defined as the presence of alternative forms of a gene (=alleles) at a given locus in a population (Barrett and Schluter 2008). Same as hybridization, polymorphisms derived from standing genetic variation are suggested to allow for a faster adaptive response to environmental change than mutations and its role has been validated in several genomic studies (Jones et al. 2012; O'Donnell et al. 2014; Reid et al. 2016; Lai et al. 2019). To discriminate between standing genetic variation and polymorphisms originating from ancient hybridization might not be easy, and future studies could identify observed standing genetic variation within a species to be the result of an ancient gene flow event. A few cases have been

documented where such polymorphisms have resulted in adaptive alleles and finally speciation, even a long time after the actual gene flow event (reviewed in Marques et al. 2019). For example, the polymorphisms in *LWS* opsin haplotypes in cichlids which affect adaptation to light condition and mate choice predate the origin of the Lake Victoria Superflock of cichlid fishes and has probably emerged as the result of a hybridization event between two ~1.5 million years divergent lineages (Terai et al. 2006; Brawand et al. 2014; Meier et al. 2017). In threespine sticklebacks (*Gasterosteus aculeatus* species complex) repeated adaptation of anadromous populations to freshwater has happened within the last 12,000 years, always involving the same genomic regions (Colosimo et al. 2004, 2005). However, the genomic variation at habitat-associated loci dates back to six million years ago, 10% of the loci even to more than 10 million years ago (Nelson and Cresko 2018).

1.3 Convergent evolution – how evolution repeats itself

Similar traits that appear in different species are usually called convergent or parallel. Parallel evolution was understood as the evolution of similar traits in closely related lineages, sharing the same evolutionary paths (Conte et al. 2012). Convergent evolution on the other hand, was defined as the appearance of similar traits that evolved in distantly related lineages (Waters and McCulloch 2021). A textbook example for convergent evolution is the marsupials from Australia. When mammals started to evolve, it was mostly the placentalia that radiated globally (Murphy et al. 2021). However, in Australia, the marsupials with external pouches became more prominent (Springer et al. 1997). There are many placental and marsupial species which are extremely similar and fill the same ecological niches, for example the marsupial mole and the mole, sugar glider and

flying squirrel, wombat and groundhog, quoll and wild cat (see Losos 2017). The most compelling example might be the thylacine, also known as the Tasmanian tiger or Tasmanian wolf. Unfortunately, it became extinct because it was eradicated by Tasmanian rangers about a century ago, but the resemblance to a wolf was astonishing. Also outside of Australia, many convergently evolved traits can be observed. The North American porcupine and the African crested porcupine have the same spiky coat, though they have evolved their spikes independently in two different lineages of rodents: the North American porcupine, also called urson, is more closely related to Guinea pigs than to the Old World porcupines (Voloch et al. 2013). Another astounding example is the eyeball of an octopus. It almost looks like an eyeball of a vertebrate, including humans, but our most recent common ancestor lived about 550 million years ago and had no eyes (Fernald 2006).

Large contribution to the topic of convergent evolution came from lizards of Caribbean islands, where the same compilation of ecotypes has evolved repeatedly on different islands (Losos 2004, 2009, 2011; Mahler et al. 2013). Also in the Great African lakes, ecotypes have evolved repeatedly (pescivores, algivores, scale-eaters, etc.) (Kocher et al. 1993; Rüber et al. 1999). To study evolution in such natural replicates provides many advantages. They tend to have a similar genetic background and evolved in isolated but comparable environments due to their shared recent evolutionary history (Elmer and Meyer 2011). If we find the genetic basis of parallel phenotypes, we might make assumptions about the independence of evolution under natural selection (Elmer and Meyer 2011).

However, the distinction between “parallel” and “convergent” evolution has recently been questioned. Examples which were classified as “parallel evolution” (similar traits in closely related

populations or species) are the reduction of armour plates in sticklebacks (Colosimo et al. 2005; Terekhanova et al. 2014), mimetic coloration in *Heliconius* butterflies (Brower 1996), migratory ecotypes of freshwater fishes (Veale and Russello 2017), flower coloration shifts in plant lineages (Smith and Rausher 2011), flight reduction in birds (Campagna et al. 2019) and in insects (Suzuki et al. 2019). On the other side, convergent evolution was only thought to appear in distantly related species (Waters and McCulloch 2021). However, to define what is closely and what is distantly related is challenging, because evolutionary relationships represent a continuum rather than distinct episodes (Arendt and Reznick 2008). It has further been shown that also in closely related lineages, similar phenotypes might evolve based on different genetic mutations. In *Timema* stick insects, similar camouflage colour-morphs have evolved based on mutations in different genetic regions (Villoutreix et al. 2020). In Hawaiian crickets, a flat-wing male phenotype has evolved several times within a very short period of time, based on non-overlapping genetic regions (Zhang et al. 2021). An example, showing that there are exceptions in both ways concerning parallel and convergent evolution, are beach mice. In many animals, one gene (*mc1r*) is responsible for a change in coat coloration suggesting a rather conserved function across several taxa (Ritland et al. 2001; Theron et al. 2001; Eizirik et al. 2003; Nachman et al. 2003; Mundy et al. 2004). However, a subspecies of beach mice from the Atlantic coast evolved a different genetic mechanism towards a higher coat coloration, since the derived *mc1r* allele which changes for instance the coat coloration for beach mice from Florida's Gulf Coast is missing (Hoekstra et al. 2006).

1.4 Ricefishes as a new model system in evolutionary biology

In ray-finned fishes, a vast diversity of reproductive tactics, modes of parental care and mechanisms of sex determination have evolved (reviewed in Mank and Avise 2006). Also here, convergent evolution is ubiquitous and has been studied in several species (e.g. viviparity (Blackburn 1999), parental care and pair ponding (Kidd et al. 2012; Fischer et al. 2019), sex determining regions (Böhne et al. 2019) or male alternative reproductive tactics (Knapp 2003; Fitzpatrick et al. 2005)). In ricefishes, a small group of closely related species became model organisms in developmental biology, genetics and toxicology: the *Oryzias latipes* species complex, also known as medaka sensu lato (Yamamoto 1975; Wittbrodt et al. 2002; Kasahara et al. 2007). Its relatives, however, were studied much less, but are highly interesting for evolutionary biologists as they are closely related but diverse in aspects of morphology, physiology and reproduction (Hilgers and Schwarzer 2019).

Sulawesi was colonized by their common ancestor about 16 million years ago and they evolved into two radiations: the *Oryzias* and the *Adrianichthys* (Mokodongan and Yamahira 2015). Most of the ricefishes are “transfer-brooding” which is the ancient mode of reproduction in this family (Parenti 2008; Mokodongan and Yamahira 2015). They only have a few eggs, but they spawn daily and lose the eggs a few hours after spawning (Yamamoto 1975; Wootton and Smith 2014). In this thesis, I focus on “pelvic-brooding” (Kottelat 1990) which is the more derived brooding strategy that appears in three species of ricefishes from Sulawesi. Interestingly, pelvic-brooding appears in species of both radiations: *Adrianichthys oophorus* on the one hand and in two sister species, *Oryzias eversi* and *Oryzias sarasinorum* on the other hand (Mokodongan and Yamahira 2015). In

pelvic-brooding species, females carry an egg-cluster in a ventral concavity formed by shorter ribs and the egg-cluster is covered by elongated pelvic fins (Spanke et al. 2021). The eggs are carried up to three weeks during which ovulation is delayed (Iwamatsu et al. 2007). It is hypothesized that the ventral concavity and the elongated pelvic fins might help to reduce drag during egg-carrying for the female (Spanke et al. 2021). While *O. eversi* females carry the eggs, a plug forms within their gonoduct (Iwamatsu et al. 2008). We could show that inflammatory reactions take place in the gonoduct which is slightly injured by the attaching filaments of the eggs and further, mammalian placenta genes were overexpressed in the plug, hinting towards a co-option of genes in very distantly related animals but in similar contexts (inflammatory response to build new tissue) (Hilgers et al. 2022).

1.5 Aims of study

In this thesis I focused on the evolution of pelvic-brooding and on factors that might have benefitted it. In chapter 2 I studied how and if traits related to pelvic-brooding correlate in *O. eversi* and *Oryzias nigrimas* and how many loci underlie each of these traits. I measured the traits based on photographs and x-ray images and estimated the number of loci using the Castle-Wright estimator. Further, I tested whether modularity exists in pelvic-brooding traits. On the one hand I used the correlations between the measured traits and on the other hand landmarks which I placed on the side of the body. If modularity was present, it would allow the respective traits or body parts to evolve together, “quasi-independently” from other body parts (Raff 1996). I hypothesize that if we find that traits related to pelvic-brooding are modular even in the transfer-brooding *O.*

nigrimas, this connectivity – if also reflected on the molecular level - might have facilitated the evolution of pelvic-brooding.

In chapter 3, I investigated gene flow into the pelvic-brooding lineage of the two sister species *O. eversi* and *O. sarasinorum*. I hypothesize that gene flow into the pelvic-brooding *Oryzias* lineage could 1) have brought in adaptive alleles from the other pelvic-brooding lineage in the *Adrianichthys* or 2) have elevated genetic variation in the ancestor of pelvic-brooding *Oryzias* which led to a faster and facilitated evolution of pelvic-brooding in *Oryzias*. We used orthologous genes to reconstruct a phylogeny and did an ancestral state reconstruction. Further, we used whole-genome sequences to find single nucleotide polymorphisms and used D-statistics to distinguish between species-tree - gene-tree incongruencies coming from incomplete lineage sorting or gene flow and compared genomic regions with high introgression signal to known quantitative trait loci (QTLs) of pelvic-brooding. If we find no gene flow between the two pelvic-brooding lineages, it strongly suggests that the trait evolved independently and convergently. If we find gene flow from other *Oryzias* into the stem lineage of pelvic-brooding *Oryzias* in regions associated with pelvic-brooding traits (QTLs), it might suggest that introduced genetic variation was recruited for the evolution of pelvic-brooding.

Chapter 4 deals with the genetic architecture of sex determination in the pelvic-brooding *O. eversi* and the closely related transfer-brooding species *O. nigrimas*. Genetic markers were designed and sequenced in F1 hybrids, intercrosses and backcrosses in both directions, and I identified genetic regions associated with sex using QTL mapping. I hypothesize that there are two different genetic architectures of sex determination if we observe biased sex ratios in the hybrid crosses. Further, if

the QTL regions for sex differ, it could imply that the new sex determining locus in the pelvic-brooding species has a connection to the new brooding strategy, since all traits associated with it are sexually dimorphic and a sex-linked basis of pelvic-brooding became more plausible

2. The genetic basis of a novel reproductive strategy in Sulawesi ricefishes: how modularity and a low number of loci shape pelvic brooding

This chapter is published in the following article (open access):

Jana M. Flury, Leon Hilgers, Fabian Herder, Tobias Spanke, Bernhard Misof, Daisy Wowor, Farnis Boneka, Letha Louisiana Wantania, Daniel F. Mokodongan, Christoph Mayer, Arne W. Nolte, Julia Schwarzer (2022) The genetic basis of a novel reproductive strategy in Sulawesi ricefishes: how modularity and a low number of loci shape pelvic brooding. Evolution (N.Y.) 76(5):1033 -1051, DOI: 10.1111/evo.14475

Original article is attached.

2.1 Higher evolvability due to modularity?

The way organisms' phenotypes are organized cannot be random and selection forces must have formed them into units which we are able to easily recognize (Wagner 1996). The correlations between such traits forming a unit are higher compared to non-functionally coupled traits (Olson and Miller 1958). Even though this early concept of modularity was already connected to functionality, modularity was first introduced by Rudy Raff (1996) in relation to development. He stated that "multicellular organisms are partitioned into quasi-independent processes and parts", which would develop independently of other structures, what he showed by transplanting limb buds to different locations where they still developed into proper limbs (Raff 1996). In the same year, Wagner (1996) started to investigate what he called "modular units of evolutionary transformation". Based on the principle of homology, he stated that modularity comes from differential integration of formerly presumably independent characters serving a common function. This implies, integrated characters are able to vary separately, being optimized by selection without interfering with each other, resulting in higher evolvability (the ability to respond to a selective challenge (Hansen 2003, e.g. Riedl 1977; Wagner and Altenberg 1996). This was further developed into the "model" of variational modularity, assuming that phenotypic traits are correlated due to pleiotropic effects of genes (a set of genes tends to have pleiotropic effects on the same set of traits, but few and weaker effects on other traits (Wagner and Altenberg 1996)). The hypothesis that pleiotropy enhances evolvability is contrasted by the idea that pleiotropy constraints the flexibility of traits to adapt (Orr 2000). Recently, the first hypothesis has received more consent, because pleiotropy has been found to be mostly restricted to functional, integrated

trait modules (Wagner and Zhang 2011), still leaving enough flexibility for traits of other modules to evolve independently (Wagner and Altenberg 1996). For example, it was found in Darwin finches that beak length and depth are regulated by different gene regulatory pathways (Abzhanov et al. 2004, 2006). This allows both traits to evolve separately in different directions, resulting in a successful radiation with beaks highly adapted to different food sources (Bowman 1963). Other examples supporting a positive effect of modularity on evolvability can be found across many taxa. In mammals, modularity of the backbone supposedly promoted the rapid evolution of locomotion and static body support (Jones et al. 2018a). In birds, the cranium was found to be highly modular, and the face and cranial vault seem to evolve faster than other regions, leading to a burst of rapid evolution (Felice and Goswami 2018). The power-amplified system of mantis shrimp consists of the engine, amplifier and tool. There are two main strategies exhibited by the mantis shrimps: “smashing” and “spearing” (Caldwell and Dingle 1976; Patek et al. 2004; Patek and Caldwell 2005; de Vries et al. 2012). There are other intermediate shapes between smashers and spearers, but it has been found that smashers have a lower modularity in their power-amplified system than non-smashers and a 10-fold slower evolutionary rate of morphological change (Claverie and Patek 2013). A higher evolutionary rate has also been found in ray-finned fishes, where the module comprising the trunk region evolves five times faster than the tail region and three times faster than the head region (Larouche et al. 2018). And an example showing that modularity may even help to overcome constraints has been documented in *Merianieae*, where modularity in other flower parts helped to overcome constraints in their tubular anthers by increased rates of evolution, allowing rapid adaptation to new pollinators (Dellinger et al. 2019).

The traits related to pelvic-brooding in *Oryzias* have probably evolved within the last ~2 million years. In the following publication we investigated whether the body is modular in a pelvic-brooding as well as in a transfer-brooding species, which represents the ancestral state of brooding. Further we looked for correlations between traits related to pelvic-brooding in the pelvic-brooding and the transfer-brooding species, and in their first- and second-generation hybrid offspring. Based on trait variations measured in the parental species and in the hybrids, we estimated the number of loci underlying pelvic-brooding traits.

2.2 Summary and personal contribution

To find the genetic basis of adaptive traits has been a major challenge for Evolutionary Biologists. Understanding how complex traits such as a novel reproductive strategy evolve on a molecular level would provide valuable insights in the genomics of adaptation and evolutionary processes shaping adaptive radiations. My direct supervisor Dr Julia Schwarzer and I conceptualized this study. Aim of the study was to gain knowledge on the genomic architecture of external morphological traits related to pelvic-brooding. We wanted to find out, how many loci underlie pelvic-brooding traits and if they are connected, i.e., correlated. Correlations between traits might be a signal for modularity, which means that they would be part of a unit that evolved together (Olson and Miller 1958; Wagner 1996; Klingenberg 2008). Such modules are “quasi-independent” (Raff 1996) and are expected to be based on a set of pleiotropic genes (reviewed in Wagner et al. 2007). If some traits related to pelvic-brooding like elongated pelvic fins and shortened ribs were correlated and form a module in *O. eversi* and their close relative and transfer-brooding species *O.*

nigrimas, we would hypothesize that the presence of modularity in these traits even in the ancestral state of transfer-brooding could be a hint that the evolution of pelvic-brooding was facilitated by its modularity. Further, we assumed a low number of partially pleiotropic loci to be responsible, because especially for the pelvic fins it has been shown that they have evolved into a variety of structures across teleost taxa (reviewed in Yamanoue et al. 2010). In sticklebacks, the reduction of the pelvic fin has been assigned to the disfunction of one gene (Shapiro et al. 2004; Coyle et al. 2007).

We crossed *O. eversi* and *O. nigrimas* to receive F1 generation hybrids; the F1 hybrids were then intercrossed and F1 males were backcrossed with females of both species. Crosses were established at the University of Plön and the Carl-von-Ossietzky University in Oldenburg by Prof. Dr Arne Nolte, and I continued to breed *O. eversi* backcrosses at the Zoological Research Museum Alexander Koenig in Bonn. I measured the traits related to pelvic-brooding in 36 *O. nigrimas*, 23 *O. eversi*, 23 interspecific F1 hybrids, 67 intercrosses, 26 *O. nigrimas* backcross, and 227 *O. eversi* backcrosses and placed landmarks on x-rays picturing the side of the fishes and on photographs of the ventral side. Using the measurements, I calculated the Castle-Wright estimator (Castle 1921; Lynch and Walsh 1998; Jones 2001) using a R-script I wrote and calculated the correlations between the traits for each line (*O. eversi*, *O. nigrimas*, F1, intercrosses, *O. eversi* backcrosses and *O. nigrimas* backcrosses). Finally, I used the landmark analysis of the whole body shape to calculate the covariance ratio (Adams 2016) in order to find correlations between sets of landmarks. I did all analyses in R and prepared the data for publication on Dryad.

We found that only a few loci are responsible for the traits of interest. Further, in *O. eversi* and *O. nigrimas* the traits were correlated. The modularity in the body plan is expressed in three modules in *O. nigrimas* and in four in *O. eversi*. Hence, pelvic-brooding seems to be based on few loci which are probably pleiotropic and therefore influence more than one trait at a time. Since modularity is already present in the transfer-brooding species which represents the ancestral state, we claim that the evolution of pelvic-brooding was facilitated by a modular organization of the body plan and potentially also a modular and pleiotropic genetic basis of a few genes. I wrote the initial draft of the manuscript, created all the figures, and prepared the supplement. The introductory figure (Figure 1) was created in Designer (Affinity) as well as Figure S1, all the photographs were taken by me except for the picture of *O. eversi* which was taken by a coauthor (Leon Hilgers). I created the figures showing statistical results (Figure 2 to 4, supplementary Figure S2 to S10) in R and optimized them using Designer (Affinity) and/or Inkscape.

3. Contribution of hybridization to the evolution of a derived reproductive strategy in ricefishes

This chapter is submitted to Genome Biology and Evolution and is published on bioRxiv (open access):

Jana M. Flury, Karen Meusemann, Sebastian Martin, Leon Hilgers, Tobias Spanke, Astrid Böhne, Fabian Herder, Daniel F. Mokodongan, Janine Altmüller, Daisy Wowor, Bernhard Misof, Arne W. Nolte, Julia Schwarzer (2022) Contribution of hybridization to the evolution of a derived reproductive strategy in ricefishes. bioRxiv (https://doi.org/10.1101/2022.07.05.498713)

Manuscript is attached in the form it was submitted to Genome Biology and Evolution and published on bioRxiv.

3.1 Positive effects of hybridization on adaptation and speciation in ricefishes from Sulawesi

Hybridization may have beneficial effects on speciation by bringing new adaptive alleles into a population (Barton 2008). Additionally, hybridization might trigger chromosomal rearrangements which itself can be advantageous (reviewed in Abbott et al. 2013): if adaptive loci get physically closer together by chromosomal fusions, chances are diminished that they will be separated by recombination. This allows the formation of regions of low recombination with the potential to form adaptive clusters (Noor et al. 2001; Guerrero and Kirkpatrick 2014; Liu et al. 2022). In general, this so called recombination suppression hypothesis suggests that if there are multiple beneficial alleles across several loci, linkage induced for example by inversion or fusion will decrease the probability of creating maladaptive recombinant genotypes (Dobzhansky 1948; Dobzhansky and Dobzhansky 1970; Rieseberg 2001; Kirkpatrick and Barton 2006; Feder and Nosil 2009; Yeaman and Whitlock 2011; Yeaman 2013). In Sulawesi ricefishes, several gene flow events were discovered between nowadays geographically close species (Mandagi et al. 2021) but also over great distances (>100 km) (Horoiwa et al. 2021). Further, the number of chromosomes is known for four Sulawesi ricefish species and varies between 2N = 36 and 2N = 42 which is why they were labelled as the “fused-chromosome” group (Naruse 1996). This is remarkable, because *O. latipes*, a genetic model organism also known as Medaka (Wittbrodt et al. 2002), has kept the same number of chromosomes (2N = 48) for over 300 million years (Naruse 1996; Kasahara et al. 2007).

We thus assume that hybridization was probably common within ricefishes and may have resulted in chromosomal fusions which could have played a role in the radiation of *Oryzias* on Sulawesi. Regarding the history of hybridization within the Sulawesi ricefishes it seems possible that pelvic-brooding has introgressed from one lineage into the other. Adaptive introgression was observed over a variety of taxa, can serve as a fast-track for evolution (e.g. Whitney et al. 2006; Barbato et al. 2017; Jones et al. 2018b; Edelman et al. 2019) and it can explain the appearance of similar traits in different species or populations (Grant et al. 2004; Witt and Huerta-Sánchez 2019; Jones et al. 2020). However, a recent study showed that there was no gene flow between pelvic-brooding lineages, rendering adaptive introgression of pelvic-brooding unlikely (Montenegro et al. 2022). Besides introgression, similar traits may evolve independently by de novo mutations. Typically, a similar environment may lead to convergent traits (Elmer and Meyer 2011). This is even more pronounced in extreme habitats, e.g. poecilids in sulfur springs (Greenway et al. 2020) or whales' echolocation in the deep ocean (Park et al. 2019). However, de novo mutations go to fixation rather slowly and are rarely beneficial (Ohta 1992). A more rapid way to evolve adaptations is to make use of standing genetic variation or the introduction of genetic variation that have already undergone selection. It is more likely for fast adaptions to be based on standing genetic variation, because they are present immediately when the environment changes and there is no need to wait for advantageous mutations (Barrett and Schluter 2008). Further, the probability of fixation is much higher for standing genetic variation, because the advantageous allele is present in several individuals of a population, whereas for a single new mutation in one individual, it is highly unlikely to reach fixation (Hermisson and Pennings 2005). Examples for convergent evolution based on

standing genetic variation are reduced armour plates in multiple stickleback species (Colosimo et al. 2005) and reduced pigmentation in cave fish (*Astyanax mexicanus*) (Gross et al. 2009).

3.2 Elevated genetic variation due to ancient gene flow

The maintenance of genetic variation within a population can be due to several factors. Local adaptation and a heterogeneous environment with balanced selection might play a role (Mitchell-Olds et al. 2007), but also gene flow contributes to high levels of genetic variation (Yeaman and Jarvis 2006). However, not only recent but also ancient hybridization between ancestral species or populations can serve as an important resource for genetic variation: it has been found that the genetic variants underlying several speciation events were introduced significantly earlier as the actual splitting time (reviewed in Marques et al. 2019). Therefore, hybridization could have contributed to the evolution of pelvic-brooding by introducing (maybe old) genetic variation, allowing the convergent evolution of pelvic-brooding, due to similar selection pressures in a comparable environment. In the following manuscript we tested for signals of gene flow into the more recent lineage of pelvic-brooding ricefishes from Sulawesi. We used D-statistics to distinguish between incomplete lineage sorting and hybridization. If we find hybridization, it could imply that the evolution of pelvic-brooding was facilitated by the increase of genetic variability due to gene flow.

3.3 Summary and personal contribution

Stephen Jay Gould claimed that if life on earth would start over, probably the outcome would be completely different (Gould 1989). However, convergent traits are omnipresent and show that

evolution sometimes repeats itself (e.g. Rüber et al. 1999; Fernald 2006; Losos 2011; Voloch et al. 2013). Reasons why traits evolve multiple times are manifold: for example, similar environments lead to similar phenotypes (Elmer and Meyer 2011) or beneficial genetic solutions are recruited in a different population or species, either due to introgression (e.g. The Heliconius Genome Consortium et al. 2012; Jones et al. 2018b) or standing genetic variation (Colosimo et al. 2005; Gross et al. 2009). In Sulawesi ricefishes, a derived reproductive strategy called "pelvic-brooding" occurs in two distantly related lineages (Kottelat 1990; Herder et al. 2012; Mokodongan and Yamahira 2015). One pelvic-brooding species derived from a more ancestral lineage is *Adrianichthys oophorus* and two species, *O. eversi* and *O. sarasinorum* evolved more recently. A recent study has found no gene flow between the two species, ruling out adaptive introgression as a reason for the appearance of the trait in two different lineages (Montenegro et al. 2022). However, gene flow from other species into the more recent lineage of *Oryzias* could still have facilitated the evolution of pelvic-brooding by raising the level of genetic variability.

I conceptualized the study together with my direct supervisor Dr Julia Schwarzer. For the first step, we used transcriptome data of twelve Sulawesi ricefish species and one custom made genome of *O. dopingdopingensis*. I used the transcriptome assemblies and the genome assembly as input for Orthograph v.0.7.1 (Petersen et al. 2017), a program which detects single-copy orthologous genes. I ran Orthograph with the help of Dr Karen Meusemann and built the reference data set based on the Actinoperygii orthologous gene set from OrthoDB v.9.1 (Waterhouse et al. 2013) together with Dr Alexandros Vasilikopoulos. I and Dr Karen Meusemann filtered the resulting single- copy protein-coding orthologous genes for outliers and masked uninformative sites according to Misof

et al. (2014). I created single gene trees in IQ-TREE (Nguyen et al. 2014). I checked the trees for paralogous sequences using PhyloTreePruner (Kocot et al. 2013) and removed gene trees with extremely long branches using TreeShrink (Mai and Mirarab 2018). For further analysis, we only kept genes present in all taxa. Using this data set, I ran a multispecies coalescence model with Astral and checked for hybrid edges using SNaQ (Solís-Lemus and Ané 2016; Solís-Lemus et al. 2017). Further, I used published whole-genome sequences of 18 ricefish species to filter for SNPs between the species (Ansai et al. 2021). Using the SNPs, I looked for a signal of introgression between the species using D-statistics (also known as ABBA-BABA) in Dsuite (Malinsky et al. 2021). Further, I ran a graph-based model in TreeMix (Pickrell and Pritchard 2012) to find hybrid edges.

We did not find any gene flow between the pelvic-brooding lineages. We could find gene flow within the Malili lake system *Oryzias*, a pattern which was already discovered by other studies (Mandagi et al. 2021). Interestingly, we discovered a hybridization event between the ancestor of the pelvic-brooding *Oryzias* and the ancestor of the Lake Poso *Oryzias*. Lake Poso is about two million years old (Rintelen et al. 2004; von Rintelen and Glaubrecht 2006) and is the sole locality where *A. oophorus* can be found today. Therefore, we have the following hypothesis: the introduced genetic variation from the ancient hybridization between Lake Poso *Oryzias* and pelvic-brooding *Oryzias* might have facilitated the convergent evolution of pelvic-brooding in a similar habitat as *A. oophorus*, perhaps in some sort of paleolake. Another possibility could be that genetic variants underlying pelvic-brooding in *A. oophorus* has introgressed into the ancestor of the Lake Poso *Oryzias* and was further transferred to the ancestor of the pelvic-brooding *Oryzias*. It has been shown in other examples, that old genetic variants might only become relevant a long time

after their introduction (reviewed in Marques et al. 2019). I and Dr Julia Schwarzer wrote the initial draft of the manuscript. I prepared the supplement and designed figures 1, 3 and 4. For figure 2, me and my supervisor Dr Julia Schwarzer contributed equally.

4. Sex determination in crosses between two Celebensis group medaka species

This chapter is prepared for submission to the Journal of Evolutionary Biology:

Jana M. Flury*, Kristin Tietje*, Julia Schwarzer, Arne W. Nolte (2022) Sex determination in crosses between two Celebensis group medaka species. Journal of Evolutionary Biology

* equal contribution

Manuscript is attached.

4.1 Sex determination systems and the sex determination cascade

In sexually reproducing species there are usually two sexes and they are in most cases genetically determined. The alleles defining the primary (ovaries and testis) and secondary (e.g. mammary glands, pheromones, colourful plumage, scales or coats, etc.) sex-specific traits of females and males are sometimes located on sex chromosomes, but not necessarily (Rice 1984; Mank 2009; Alexander et al. 2015). There are species with heteromorphic sex chromosomes, which means that one chromosome is degenerated compared to the other (Graves 2006; Bergero and Charlesworth 2009). This results from a lower recombination rate and favours linkage between sexually antagonistic alleles and sex-determining loci. Here, if the male carries the heteromorphic chromosomes, turning it into the heterogametic sex, it is called the XY system which is for example common in mammals (Ohno 1967, 1979; Cooper et al. 1975). If the female carries the heteromorphic chromosome, it is a ZW system, which is the case in butterflies and birds (Bachtrog et al. 2014). Besides the genetic sex determination, in some species the sex of the offspring depends on the environment. In crocodiles, most turtles and some fishes, the temperature determines whether the young will be male or female (Bull 1980; Crews 1993).

It has long been defined that after sex determination, animals go through sexual differentiation (Lillie 1939). However, it has recently been proposed that this division into two separate steps should be lifted because the process rather resembles a continuum, with several factors influencing each other in a network (Uller and Helanterä 2011; Heule et al. 2014). This was observed mainly in teleost fishes, where for example anemonefish change the sex during their lifetime (Moyer and Nakazono 1978; Fricke 1983) and it has been shown experimentally in Nile

tilapia and medaka that sex reversals are possible, also after genetic fixation of one sex at fertilization (Paul-Prasanth et al. 2013). Still, in most animals the sex is determined by a "master switch" and is followed by a cascade of activated down-stream genes, initiating the expression of sexually dimorphic traits (Heule et al. 2014). Well-known examples for such traits are colourful plumage in male birds, manes in lions, a cockscomb and many more. A bit less attention was given to female secondary sexual traits, such as a larger body size in African jacana and pronounced ornamentation in spotted hyenas (Clutton-Brock 2007). So far, there is no better explanation for these observations than sexual selection (Darwin 1871). It was for a long time defined as selection on traits used in male-male competition and female choice. However, with more and more evidence for intrasexual competition of females and male mate choice, the theory needed to be broadened (Clutton-Brock 2004). Most of the sexually dimorphic traits documented so far do not seem to fulfil an ecological function and it remained difficult for a long time to prove the opposite (Shine 1989). Nevertheless, it has been shown in a few examples that sexual dimorphism might be due to different food resources, habitat preferences or parental care (dwarf chameleons (Stuart-Fox and Moussali 2007), hummingbirds (Temeles et al. 2000), damselflies (Cooper 2010), seabirds (Weimerskirch et al. 2009), sailfin silversides (Wasiljew et al. 2021) and cichlids (Ronco et al. 2019).

4.2 High variability of sex determination in fishes and its potential impact on the evolution of pelvic-brooding

Even though sex determination was assumed to be conserved (Mank and Avise 2009), it has been found to be extremely variable in teleost fishes (Bachtrog et al. 2014). This might be due to the

homomorphic sex chromosomes in fishes (Ohno 1974), in contrast to heteromorphic sex chromosomes found in mammals and birds, where either the Y or the Z chromosome is degenerated (Takagi and Sasaki 1974; Graves 2006). For a turnover in sex determination to happen, either a former autosomal chromosome gains a mutation that from then on determines the sex, or a mutation within the sex determining locus changes the heterogametic sex into the homogametic sex. Up to seven such turnovers are hypothesized in the Celebensis group of Sulawesi ricefishes with potentially six different sex determining regions, containing known sex determining genes such as *sox7*, *pgm3/PGM*, *hsdl1* and *amh* (Ansai et al. 2022). Taken together with their radiation and the diversity of sexual dimorphism (Parenti 2008; Ansai et al. 2021), they serve as a perfect model to study the evolution of sex determining genes and sexually dimorphic phenotypic diversity. In the following manuscript we aimed to find the sex determining region in second generation hybrids of a pelvic-brooding (*O. eversi*) and a transfer-brooding (*O. nigrimas*) ricefish. We genotyped custom-designed markers and created a linkage map for every cross (backcross with *O. nigrimas*, backcross with *O. eversi* and intercross) and mapped the region associated with sex. If the genetic architecture of sex determination differs between the two species, it might be connected to the transition between transfer-brooding and pelvic-brooding, which comes with sexually dimorphic, adapted traits.

4.3 Summary and personal contribution

Fishes have a large variety of sex determination systems compared to other animals, such as mammals and birds that only have one way to determine sex (reviewed in Bachtrog et al. 2014).

Maybe this variation contributed to the huge diversity of fish species. In Sulawesi ricefishes, up to seven sex determination turnovers have been suggested (Myosho et al. 2015; Ansai et al. 2022). Here we wonder, if one of these turnover events might play a role in the evolution of pelvic-brooding. It has been shown in a cichlid species that a sex determination locus has evolved close to another locus important for sexual dimorphism (Streelman et al. 2003). Hence I wanted to know, if the sex determination mechanism is different between two closely related Sulawesi ricefishes, a transfer-brooding *O. nigrimas* and a pelvic-brooding *O. eversi* (Kottelat 1990; Herder et al. 2012). I conceptualized a genetic mapping study together with Prof. Dr Arne W. Nolte, Dr Julia Schwarzer and Dr Kristin Tietje. Prof. Dr Arne W. Nolte crossed a male *O. eversi* and a female *O. nigrimas* to produce F1 interspecific hybrids and crossed males and females of the F1 generation to receive F2 intercrosses. Moreover, F1 hybrid males were crossed back with females of the parental species, *O. nigrimas* and *O. eversi* to receive F2 backcrosses. I entered this project when the resulting juveniles had been raised to sexual maturity and sacrificed for genetic analysis. A total of 74 females and 192 males of the F2 intercrosses, 71 females and 103 males of the *O. eversi* backcrosses and 25 females and 102 males of the *O. nigrimas* backcrosses were genotyped using a customized Genotyping-in-Thousands by sequencing approach (Campbell et al. 2015). For this purpose, Dr Kristin Tietje identified genome fragments carrying single nucleotide polymorphisms (SNPs) that distinguish between the parental species O. eversi and O. nigrimas and that could be place on a *O. latipes* reference genome. Markes were chosen to be within *O. latipes* genes and evenly distributed across the *O. latipes* reference genome (~10 markers per chromosome, 2N=48) to obtain an efficient marker panel for genetic mapping studies. After testing the primers, Dr Kristin

Tietje and I established and ran the PCR based genotyping protocol with 178 markers and 567 fishes. PCR products for the complete marker panel were multiplex PCR amplified in a single reaction and labelled with short sequence tags as individual identifiers. PCR products for all individuals were pooled and jointly sequenced on an Illumina platform. After sequencing, I called individual multilocus genotypes from the raw reads using STACKS (Catchen et al. 2011) and created a linkage map separately for the intercrosses, the *O. nigrimas* backcrosses and the *O. eversi* backcrosses using JoinMap 4.0 (Van Ooijen 2006). I checked for synteny with *O. latipes* using Blast+ (Camacho et al. 2009) and mapped the sex determining locus for each line on the linkage map. I found a significant locus for the intercrosses and the *O. eversi* backcrosses on the linkage group referring to *O. latipes* chromosome 24. The sex determining locus of seven other Sulawesi ricefishes has been found on this chromosome. In the intercrosses, another locus on the linkage group referring to *O. latipes* chromosome 21 has been recorded, however non-significant. For the *O. nigrimas* backcrosses, no significant locus was detected. However, in this cross, only a low number of females was included due to a severe sex-bias. Therefore, I still considered the three observed peaks as informative. They were found on the linkage groups referring to *O. latipes* chromosome 4, 20 and 24, which is in line with candidate loci found for other related species (Myosho et al. 2015; Ansai et al. 2022).

I created the figures of the linkage maps in Joinmap 4.0 and edited them in Inkscape. I created the LOD plots in R and the synteny plot using Circos (Krzywinski et al. 2009). I wrote the initial draft of the manuscript.

5. Discussion and outlook

In this thesis I focussed on the evolution of a derived reproductive strategy in *Oryzias eversi,* a ricefish from Sulawesi, Indonesia. They underwent a probably costly transition from ancestral transfer-brooding to pelvic-brooding, resulting in extensive maternal brood care accompanied with several sexually dimorphic adaptations. Interestingly, this transition happened twice within the ricefishes from Sulawesi: in the distantly related lineages *Adrianichthys* and *Orzyias*. It is still unclear how such fundamental changes take place, from an evolutionary perspective as well as on a molecular level. In this chapter, I will discuss my results and elaborate on contributions of my findings to the field of evolutionary biology with a focus on the genomic basis of convergent adaptive traits. Finally, I will give an outlook of the next steps that could be done and how I think they should be done.

5.1 The genetic basis of convergent adaptive phenotypes

In the last 20 years, next generation sequencing gave biologists the opportunity to investigate questions in evolutionary biology from a completely new perspective. Many examples of convergent evolution were analysed on a molecular level, revealing common genetic backgrounds of similar structures in different species. For example, the same independent mutation in a key developmental gene has evolved four times into a “bony spines” phenotype in seahorses (Li et al. 2021). A mutation in *rhodopsin* across all fish species has facilitated the transition to a different light environment, resulting in at least 20 independent convergently evolved adaptations (Hill et al. 2019). However, the appearance of convergent phenotypes can also be less independent. Adaptive introgression has led to similar traits in closely related species that are based on the same

alleles (The Heliconius Genome Consortium et al. 2012; Racimo et al. 2015; Jones et al. 2018b; Graham et al. 2021). Repeated sorting of standing genetic variation can also lead to similar phenotypes in independent populations as was observed in the reduced armour of freshwater stickleback populations (Colosimo et al. 2005) and the reduced eye size and albinism in cavefishes (Ornelas-García et al. 2008; Gross 2016). In chapter 3, I showed that convergent phenotypes may evolve "quasi-independently", using introduced ancient genetic variation that was not directly causing the phenotype in the species of origin, however these polymorphisms paired with *de novo* mutations could be the basis of the convergent phenotype of pelvic-brooding in the recipient species. A case, where the evolution of a complex convergent trait in a distantly related species could be connected to genetic variation introduced by a third species not expressing the convergent trait, has to my best knowledge not been documented so far. Scenarios are known, where allelic polymorphisms were introduced by ancient hybridization and were later used in the in the descendant lineages resulting in convergent evolution (Meier et al. 2017; Irisarri et al. 2018; Nelson and Cresko 2018). But here, the convergently evolved traits are not coming from the same pool of polymorphisms, as *Adrianichthys* has split from *Oryzias* on Sulawesi over 16 million years ago. However, independent convergent evolution is abundant (e.g. Blackburn 1992; Blackledge and Gillespie 2004; Losos 2009; Fischer et al. 2019) and maybe the origin of alleles responsible for the traits is not clarified yet in most cases, but could in the future be traced back to introduced variation based on gene flow. Therefore, my findings in chapter 3 lay the ground for future studies on convergent evolution considering ancient gene flow as a source of variation which can be recruited for the evolution of convergent adaptive phenotypes. This adds to the previous

knowledge that hybridization - under certain circumstances - can be highly beneficial (e.g. Seehausen 2004; Mallet 2007; Litsios and Salamin 2014; Grant and Grant 2019). Ancient genetic variation and its impact on speciation has only recently been explored in more detail (e.g. Genner and Turner 2012; Irisarri et al. 2018) and holds huge potential to be a valuable resource for natural selection to act upon (reviewed in Marques et al. 2019).

5.2 Complexity of adaptive traits

Studying the complexity of an adaptive trait on a molecular level leads to a better understanding of its evolution. Many adaptive traits are based on single genes, like *mcr1* which caused change in coat coloration in several taxa (e.g. Theron et al. 2001; Eizirik et al. 2003; Nachman et al. 2003; Gross et al. 2009) or *pitx1* which was responsible for pelvic-fin reduction in sticklebacks (Coyle et al. 2007; Chan et al. 2010). However, also more complex genetic underpinnings have been observed, as for example for wing colour patterns in *Heliconius erato* where loci of major effect and minor effect complement each other (Papa et al. 2013). Theory as well as empiric evidence state that an exponential distribution of allelic effects is likely, which means a few loci of large effect and many loci of smaller effect contribute to the expression of a trait (Robertson 1967; Mackay 2001). In chapter 2, I showed that only a few loci responsible for traits related to pelvic-brooding were estimated to be involved. This finding was confirmed by a recent QTL study (Montenegro et al. 2022) which renders my applied Castle-Wright estimator a valid approximation. However, the complexity of a trait-complex cannot be judged based solely on the number of loci involved in each individual trait, but also based on their connectivity. Body parts serving a common

function are often organized in modules (Wagner 1996). This means, they can evolve "quasi-independently" from other body parts which enhances evolvability (Raff 1996; Wagner and Altenberg 1996; Hansen 2003). The genetic basis of such modules is thought to be a set of pleiotropic genes (Wagner et al. 2007). Although I found modularity in the pelvic-brooding as well as in the transfer-brooding species, I did not assume a simple pleiotropic basis, as the correlations were lost or reduced in the F1 and F2 generation hybrids. The results from Montenegro et al. (2022) support this hypothesis, because the loci associated with pelvic-brooding traits mapped to different chromosomes. However, maybe hormones or other upstream regulators led to the co-expression of the responsible genes found in the QTL study, as for example in *O. woworae* from Sulawesi, where an androgen-induced expression of sexually dimorphic red fins in males has been documented (Ansai et al. 2021).

If we think of a network of already connected loci in combination with new genetic variation in the edges of this network, either introduced by hybridization or by advantageous *de novo* mutations, this maybe served as the perfect playground for new phenotypes to evolve. *De novo* mutations are more likely to be deleterious than advantageous (Ohta 1992), hence, if several loci are interacting in a network, the weight of one mutation might not be detrimental enough to effectively harm the fitness of an individual. This robustness in complex regulatory gene networks has been shown to be crucial for the evolution of innovation (Ciliberti et al. 2007). Additionally, the introduction of genetic variation from other species is rarely harmful because it was already tested by natural selection (Abbott et al. 2013; Marques et al. 2019; Urban et al. 2021). Therefore, it seems likely that a balanced influence of both, *de novo* mutations and ancient polymorphisms probably

arranged in some sort of network contributed to the evolution of pelvic-brooding. Modularity in this case likely did not pose constraints as was suggested by Orr (2000), but enlarged the benefit of mutations by providing the necessary robustness for innovation to evolve (see Wagner 2008).

5.3 Influence of sex determination on adaptation

Although sex determination and sex differentiation were historically thought to be two separate processes (Lillie 1939), they should rather be understood as a continuum, best explained by a genetic cascade or a network (Uller and Helanterä 2011; Heule et al. 2014). In most cases, the cascade is started by a "master switch" (i.e. one gene or a set of genes) that initiates the expression of the rest of the sex determining genetic cascade (Bachtrog et al. 2014). There is often a connection between the genetic sex determination and sexually dimorphic (adaptive) traits (Mank 2009) which in regard to this thesis, could imply an evolutionary connection between the genetic architecture of sex and pelvic-brooding in *O. eversi*. I discussed potential sex determining loci in *O. eversi* in chapter 4, resulting in two probable locations of sex determining genetic factors on OLchr 4 (Ansai et al. 2022) and OLchr 24. Interestingly, the locus found in my study is located on the same chromosome (OLchr 24) as a quantitative trait locus (QTL) for pelvic-fin length, a sexually dimorphic trait related to pelvic-brooding (Montenegro et al. 2022). However, the number of adult individuals in the wild was already low at time of discovery and capture (Herder et al. 2012) which is why we cannot exclude an effect of biased sampling in the wild for both studies, resulting in two different sex determining loci in the different aquarium populations. Additionally, both populations were

bred in the aquarium for years, where also new sex determining loci may arise or are lost, like in zebrafish (Wilson et al. 2014).

The genetic architecture of sex determination in the closely related transfer-brooding species, *Oryzias nigrimas*, seems to be polygenic based on my analyses (potential QTL on OLchr 4, 20 and 24). Ansai et al. (2022) did not detect any locus related to sex for this species, suggesting that they used too few individuals (10 males, 10 females) to reliably map several loci. However, the genomic regions introgressed from the ancestor of the Lake Poso *Orzyias* (chapter 3) not only overlap with a QTL for pelvic fins on OLchr 24, but also with candidate sex determining genes (*sox7*, *pgm3/PGM* and *hsdl1*) found in other *Orzyias* from Sulawesi (Ansai et al. 2022). Hence, it seems like the cascade of genetic sex determination and the potential genetic network of traits related to pelvic-brooding in *O. eversi* (chapter 2) somehow overlap on OLchr 24, and they were potentially both influenced by ancient gene flow.

The results in chapter 4 support the growing awareness that polygenic sex determination is not necessarily evolutionary unstable or transitory (Kosswig 1964; Volff and Schartl 2001; Vandeputte et al. 2007; Roberts et al. 2016). Fishes are an extreme example of how variable sex determination can be and it was suggested in East African cichlids that their species diversity could be connected to highly variable sex determination systems (El Taher et al. 2021; Feller et al. 2021). However, it still needs to be investigated, whether variable sex determination systems are the cause or the consequence of species radiations (Feller et al. 2021). The same is true for the evolution of pelvic-brooding and sex determination in *O. eversi*: to see if one caused the other, we need to know the

genetic architecture of sex determination in more details as well as the genetic basis of pelvic brooding, preferably in all pelvic-brooding *Oryzias* species.

5.4 Outlook

5.4.1 Karyotyping

In chapter 4 we created three linkage maps based on three different F2 crosses of *O. eversi* and *O. nigrimas*. Linkage groups found in this study should represent real chromosomes. However, we only know the number of chromosomes of *O. nigrimas* (2N = 38) (Naruse 1996) and due to the variability in chromosome numbers in ricefishes from Sulawesi, we might have crossed two species with different chromosome numbers. Therefore, we will karyotype *O. eversi* to be able to set a meaningful threshold for the high-resolution linkage map we want to create.

5.4.2 High resolution linkage map and quantitative trait loci mapping

So far, we have an estimate of the number of responsible loci for traits related to pelvic-brooding (chapter 2) and there was a study published about the genetic basis of the pelvic fin length and the ventral concavity using hybrid crosses of *O. eversi* and *O. dopingdopingensis* (Montenegro et al. 2022). We plan a QTL study using the *O. eversi* backcrosses (male F1 hybrids between *O. eversi* and *O. nigrimas* backcrossed with *O. eversi* females), because in this cross we were able to breed a large number of females, which is necessary for a QTL study (Beavis 1994). We want to do RAD-sequencing (restriction-site associated DNA markers) to receive randomly distributed markers across the genome which we can genotype. In addition to pelvic fin length, we will measure the

distance between the pelvic fin insertions, the rib length as approximation for the concavity and the genital papilla size of the female *O. eversi* backcrosses and create a phenotype-genotype map which we can compare to the QTLs found in (Montenegro et al. 2022). Once we know the regions of interest, we can as well compare them to the introgressed regions found in chapter 3. Further, we will map sex again to see if it overlaps with QTL regions for traits related to pelvic-brooding when both are mapped in the same study.

5.4.3 Comparison to *Adrianichthys*, the other pelvic-brooding lineage

After we have identified the genomic regions associated with pelvic-brooding in *Oryzias*, it would be interesting to compare them to the regions associated with this brooding strategy in *Adrianichthys*, the second lineage of pelvic-brooding species from Sulawesi using a comparative genomic approach. Here we´ll map genomic regions associated with pelvic-brooding found in *O. eversi* onto the *A. oophorus* genome and also to transfer-brooding species. We then use this data to investigate similarities between both pelvic-brooding species that are different compared to transfer-brooding species. These genomic regions may serve potential candidate regions to study a common genetic basis of morphological traits associated to their brooding strategy in more depth.

References

References of all chapters, including manuscripts of chapter 3 and 4 in the attachment.

Abbott, R., D. Albach, S. Ansell, J. W. Arntzen, S. J. E. Baird, N. Bierne, J. Boughman, A. Brelsford, C. A. Buerkle, R. Buggs, R. K. Butlin, U. Dieckmann, F. Eroukhmanoff, A. Grill, S. H. Cahan, J. S. Hermansen, G. Hewitt, A. G. Hudson, C. Jiggins, J. Jones, B. Keller, T. Marczewski, J. Mallet, P. Martinez-Rodriguez, M. Möst, S. Mullen, R. Nichols, A. W. Nolte, C. Parisod, K. Pfennig, A. M. Rice, M. G. Ritchie, B. Seifert, C. M. Smadja, R. Stelkens, J. M. Szymura, R. Väinölä, J. B. W. Wolf, and D. Zinner. 2013. Hybridization and speciation. J. Evol. Biol. 26:229–246.

Aberer, A. J., D. Krompass, and A. Stamatakis. 2013. Pruning rogue taxa Improves phylogenetic accuracy: an efficient algorithm and webservice. Syst. Biol. 62:162–166.

Abzhanov, A., W. P. Kuo, C. Hartmann, B. R. Grant, P. R. Grant, and C. J. Tabin. 2006. The calmodulin pathway and evolution of elongated beak morphology in Darwin's finches. Nature 442:563–567.

Abzhanov, A., M. Protas, B. R. Grant, P. R. Grant, and C. J. Tabin. 2004. *Bmp4* and morphological variation of beaks in Darwin's finches. Science 305:1462–1465.

Adams, D. C. 2016. Evaluating modularity in morphometric data: challenges with the RV coefficient and a new test measure. Methods Ecol. Evol. 7:565–572.

Alexander, H. J., J. M. L. Richardson, S. Edmands, and B. R. Anholt. 2015. Sex without sex chromosomes: genetic architecture of multiple loci independently segregating to determine sex ratios in the copepod *Tigriopus californicus*. J. Evol. Biol. 28:2196–2207.

Allen, E. 1932. Sex and internal secretions: a survey of recent research. Williams and Wilkins Co, Baltimore.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403–410.

Ansai, S., K. Mochida, S. Fujimoto, D. F. Mokodongan, B. K. A. Sumarto, K. W. A. Masengi, R. K. Hadiaty, A. J. Nagano, A. Toyoda, K. Naruse, K. Yamahira, and J. Kitano. 2021. Genome editing reveals fitness effects of a gene for sexual dichromatism in Sulawesian fishes. Nat. Commun. 12:1350.

Ansai, S., J. Montenegro, K. W. A. Masengi, A. J. Nagano, K. Yamahira, and J. Kitano. 2022. Diversity of sex chromosomes in Sulawesian medaka fishes. bioRxiv. https://doi.org/10.1101/2022.02.28.482394

Apweiler, R., A. Bairoch, C. H. Wu, W. C. Barker, B. Boeckmann, S. Ferro, E. Gasteiger, H. Huang, R. Lopez, M. Magrane, M. J. Martin, D. A. Natale, C. O'Donovan, N. Redaschi, and L. L. Yeh. 2004. UniProt: the universal protein knowledgebase. Nucleic Acids Res. 32:D115 – 119.

Arendt, J., and D. Reznick. 2008. Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? Trends Ecol. Evol. 23:26–32.

Arnold, M. L. 2007. Evolution through genetic exchange. Oxford University Press.

Assaf, G., and G. J. Hannon. 2010. FASTX-toolkit.

Baack, E. J., and L. H. Rieseberg. 2007. A genomic view of introgression and hybrid speciation. Curr. Opin. Genet. Dev. 17:513-518.

Bachtrog, D., J. E. Mank, C. L. Peichel, M. Kirkpatrick, S. P. Otto, T.-L. Ashman, M. W. Hahn, J. Kitano, I. Mayrose, R. Ming, N. Perrin, L. Ross, N. Valenzuela, J. C. Vamosi, and T. T. of S. Consortium. 2014. Sex determination: why so many ways of doing it? PLoS Biol.

12:e1001899.
Baker, H. G. 1959. Reproductive methods as factors in speciation in flowering plants. Cold Spring Harb. Symp. Quant. Biol. 24:177–91.
Balounova, V., R. Gogela, R. Cegan, P. Cangren, J. Zluvova, J. Safar, V. Kovacova, R. Bergero, R. Hobza, B. Vyskot, B. Oxelman, D. Charlesworth, and B. Janousek. 2019. Evolution of sex determination and heterogamety changes in section Otites of the genus *Silene*. Sci. Rep. 9:1045.
Barbato, M., F. Hailer, P. Orozco-terWengel, J. Kijas, P. Mereu, P. Cabras, R. Mazza, M. Pirastru, and M. W. Bruford. 2017. Genomic signatures of adaptive introgression from European mouflon into domestic sheep. Sci. Rep. 7:7623.
Barrett, R. D. H., and D. Schluter. 2008. Adaptation from standing genetic variation. Trends Ecol. Evol. 23:38–44.
Barton, N., and B. O. Bengtsson. 1986. The barrier to genetic exchange between hybridising populations. Heredity (Edinb). 57:357–376.
Barton, N. H. 2008. The role of hybridization in evolution. Mol. Ecol. 10:551–568.
Beavis, W. D. 1994. The power and deceit of QTL experiments: lessons from comparative QTL studies. Proc. forty-ninth Annu. corn sorghum Ind. Res. Conf. 250.
Benson, G. 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 27:573–580.
Bergero, R., and D. Charlesworth. 2009. The evolution of restricted recombination in sex chromosomes. Trends Ecol. Evol. 24:94–102.
Bird, C. E., I. Fernandez-Silva, D. J. Skillings, and R. J. Toonen. 2012. Sympatric speciation in the post “Modern Synthesis” era of evolutionary biology. Evol. Biol. 39:158–180.
Blackburn, D. G. 1992. Convergent evolution of viviparity, matrotrophy, and specializations for fetal nutrition in reptiles and other vertebrates. Am. Zool. 32:313–321.
Blackburn, D. G. 1999. Viviparity and oviparity: evolution and reproductive strategies. Encycl. Reprod. 4:994–1003.
Blackledge, T. A., and R. G. Gillespie. 2004. Convergent evolution of behavior in an adaptive radiation of Hawaiian web-building spiders. Proc. Natl. Acad. Sci. U. S. A. 101:16228–33.
Blackmon, H., and J. P. Demuth. 2014. Estimating tempo and mode of Y chromosome turnover: explaining Y chromosome loss with the fragile Y hypothesis. Genetics 197:561–572.
Böhne, A., A. A.-T. Weber, J. Rajkov, M. Rechsteiner, A. Riss, B. Egger, and W. Salzburger. 2019. Repeated evolution versus common ancestry: sex chromosome evolution in the haplochromine cichlid *Pseudocrenilabrus philander*. Genome Biol. Evol. 11:439–458.
Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.
Bouckaert, R. R., and J. Heled. 2014. DensiTree 2: seeing trees through the forest. bioRxiv 012401.
Bowman, R. I. 1963. Morphological differentiation and adaptation in the Galápagos finches. Auk 80:202–207.
Brandon, R. N. 1978. Adaptation and evolutionary theory. Pp. 181–206 in Studies in history and

philosophy of science.

Brawand, D., C. E. Wagner, Y. I. Li, M. Malinsky, I. Keller, S. Fan, O. Simakov, A. Y. Ng, Z. W. Lim, E. Bezault, J. Turner-Maier, J. Johnson, R. Alcazar, H. J. Noh, P. Russell, B. Aken, J. Alföldi, C. Amemiya, N. Azzouzi, J.-F. Baroiller, F. Barloy-Hubler, A. Berlin, R. Bloomquist, K. L. Carleton, M. A. Conte, H. D'Cotta, O. Eshel, L. Gaffney, F. Galibert, H. F. Gante, S. Gnerre, L. Greuter, R. Guyon, N. S. Haddad, W. Haerty, R. M. Harris, H. A. Hofmann, T. Hourlier, G. Hulata, D. B. Jaffe, M. Lara, A. P. Lee, I. MacCallum, S. Mwaiko, M. Nikaido, H. Nishihara, C. Ozouf-Costaz, D. J. Penman, D. Przybylski, M. Rakotomanga, S. C. P. Renn, F. J. Ribeiro, M. Ron, W. Salzburger, L. Sanchez-Pulido, M. E. Santos, S. Searle, T. Sharpe, R. Swofford, F. J. Tan, L. Williams, S. Young, S. Yin, N. Okada, T. D. Kocher, E. A. Miska, E. S. Lander, B. Venkatesh, R. D. Fernald, A. Meyer, C. P. Ponting, J. T. Streelman, K. Lindblad-Toh, O. Seehausen, and F. Di Palma. 2014. The genomic substrate for adaptive radiation in African cichlid fish. Nature 513:375–381.

Broman, K., and S. Sen. 2009. A guide to QTL mapping with R/qtl. Springer.

Broman, K. W., H. Wu, Ś. Sen, and G. A. Churchill. 2003. R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889–890.

Brower, A. V. Z. 1996. Parallel race formation and the evolution of mimicry in *Heliconius* butterflies : a phylogenetic hypothesis from mitochondrial DNA sequences. Evolution (N. Y). 50:195–221.

Bull, J. J. 1980. Sex determination in reptiles. Q. Rev. Biol. 55:3–21.

Caldwell, R. L., and H. Dingle. 1976. Stomatopods. Sci. Am. 234:80–89.

Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, and T. L. Madden. 2009. BLAST+: architecture and applications. BMC Bioinformatics 10:421.

Campagna, L., K. G. McCracken, and I. J. Lovette. 2019. Gradual evolution towards flightlessness in steamer ducks. Evolution (N. Y). 73:1916–1926.

Campbell, N. R., S. A. Harmon, and S. R. Narum. 2015. Genotyping-in-Thousands by sequencing (GT-seq): a cost effective SNP genotyping method based on custom amplicon sequencing. Mol. Ecol. Resour. 15:855–867.

Castle, W. 1921. An improved method of estimating the number of genetic factors concerned in cases of blending inheritance. Science 54:223–223.

Castric, V., J. Bechsgaard, M. H. Schierup, and X. Vekemans. 2008. Repeated adaptive introgression at a gene under multiallelic balancing selection. PLoS Genet. 4.

Catchen, J. M., A. Amores, P. Hohenlohe, W. Cresko, and J. H. Postlethwait. 2011. Stacks: building and genotyping loci de novo from short-read sequences. G3 Genes|Genomes|Genetics 1:171–182.

Chan, Y. F., M. E. Marks, F. C. Jones, G. V. Jr, M. D. Shapiro, S. D. Brady, A. M. Southwick, D. M. Absher, J. Grimwood, J. Schmutz, R. M. Myers, D. Petrov, B. Jónsson, D. Schluter, M. A. Bell, and D. M. Kingsley. 2010. Adaptive evolution of pelvic reduction of a *pitx1* enhancer. Science (80-.). 327:302–305.

Ciliberti, S., O. C. Martin, and A. Wagner. 2007. Innovation and robustness in complex regulatory gene networks. Proc. Natl. Acad. Sci. 104:13591–13596.

Claverie, T., and S. N. Patek. 2013. Modularity and rates of evolutionary change in a power-amplified prey capture system. Evolution (N. Y). 67:3191–3207.

Clutton-Brock, T. H. 2007. Sexual selection in males and females. Science 318:1882–1885.

Clutton-Brock, T. H. 2004. What is sexual selection? Pp. 24–36 in P. M. Kappeler and C. P. van Schaik, eds. Sexual selection in primates: new and comparative perspectives. Cambridge University Press, Cambridge.

Colosimo, P. F., K. E. Hosemann, S. Balabhadra, G. Villarreal, M. Dickson, J. Grimwood, J. Schmutz, R. M. Myers, D. Schluter, and D. M. Kingsley. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. Science 307:1928–33.

Colosimo, P. F., C. L. Peichel, K. Nereng, B. K. Blackman, M. D. Shapiro, D. Schluter, and D. M. Kingsley. 2004. The genetic architecture of parallel armor plate reduction in threespine sticklebacks. PLoS Biol. 2:635–641.

Conte, G. L., M. E. Arnegard, C. L. Peichel, and D. Schluter. 2012. The probability of genetic parallelism and convergence in natural populations. Proc. R. Soc. B Biol. Sci. 279:5039–5047.

Cooke, S. J., D. P. Philipp, D. H. Wahl, and P. J. Weatherhead. 2006. Energetics of parental care in six syntopic centrarchid fishes. Oecologia 148:235–249.

Cooper, D. W., P. O. Johnston, C. E. Murtagh, and J. L. Vandeberg. 1975. Sex chromosome evolution and activity in mammals, particularly kangaroos. P. in P. J. Peacock and R. Brock, eds. The eukaryote chromosome. Australien National University Press, Canberra.

Cooper, I. A. 2010. Ecology of sexual dimorphism and clinal variation of coloration in a damselfly. Am. Nat. 176:566–572.

Coyle, S. M., F. A. Huntingford, and C. L. Peichel. 2007. Parallel evolution of *Pitx1* underlies pelvic reduction in Scottish threespine stickleback (Gasterosteus aculeatus). J. Hered. 98:581–586.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer Associates, Sunderland.

Crews, D. 1993. The organizational concept and vertebrates without sex chromosomes. Brain. Behav. Evol. 42:202–214.

Danchin, É., A. Charmantier, F. A. Champagne, A. Mesoudi, B. Pujol, and S. Blanchet. 2011. Beyond DNA: integrating inclusive inheritance into an extended theory of evolution. Nat. Rev. Genet. 12:475–486.

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, and R. Durbin. 2011. The variant call format and VCFtools. Bioinformatics 27:2156-2158.

Darwin, C. 1871. The descent of man and selection in relation to sex. Modern Library, New York.

Darwin, C. 1859. The origin of species by means of natural selection. John Murray, London.

de Vries, M. S., E. A. K. Murphy, and S. N. Patek. 2012. Strike mechanics of an ambush predator: the spearing mantis shrimp. J. Exp. Biol. 215:4374–4384.

Dellinger, A. S., S. Artuso, S. Pamperl, F. A. Michelangeli, D. S. Penneys, D. M. Fernández-Fernández, M. Alvear, F. Almeda, W. Scott Armbruster, Y. Staedler, and J. Schönenberger. 2019. Modularity increases rate of floral evolution and adaptive success for functionally specialized pollination systems. Commun. Biol. 2:453.

Dobzhansky, T. 1935. A critique of the species concept in biology. Philos. Sci. 2:344–355.

Dobzhansky, T. 1951. Genetics and the origin of species. 3rd ed. Columbia University Press, New York.

Dobzhansky, T. 1937. Genetics and the origin of species. Columbia University Press, New York.

Dobzhansky, T. 1948. Genetics of natural populations; experiments on chromosomes of Drosophila pseudoobscura from different geographic regions. Genetics 33:588–602.

Dobzhansky, T., and T. G. Dobzhansky. 1970. Genetics of the evolutionary process. Columbia University Press, New York.

Edelman, N. B., P. B. Frandsen, M. Miyagi, B. Clavijo, J. Davey, R. B. Dikow, G. García-accinelli, S. M. Van Belleghem, and N. Patterson. 2019. Genomic architecture and introgression shape a butterfly radiation. Science 366:594–599.

Eizirik, E., N. Yuhki, W. E. Johnson, M. Menotti-Raymond, S. S. Hannah, and S. J. O'Brien. 2003. Molecular genetics and evolution of melanism in the cat family. Curr. Biol. 13:448–453.

El Taher, A., F. Ronco, M. Matschiner, W. Salzburger, and A. Böhne. 2021. Dynamics of sex chromosome evolution in a rapid radiation of cichlid fishes. Sci. Adv. 7:eabe8215.

Elgvin, T. O., C. N. Trier, O. K. Tørresen, I. J. Hagen, S. Lien, A. J. Nederbragt, M. Ravinet, H. Jensen, and G.-P. Sætre. 2017. The genomic mosaicism of hybrid speciation. Sci. Adv. 3:e1602996.

Elmer, K. R., and A. Meyer. 2011. Adaptation in the age of ecological genomics: insights from parallelism and convergence. Trends Ecol. Evol. 26:298–306.

Feder, J. L., S. H. Berlocher, J. B. Roethele, H. Dambroski, J. J. Smith, W. L. Perry, V. Gavrilovic, K. E. Filchak, J. Rull, and M. Aluja. 2003. Allopatric genetic origins for sympatric host-plant shifts and race formation in *Rhagoletis*. Proc. Natl. Acad. Sci. 100:10314–10319.

Feder, J. L., and P. Nosil. 2009. Chromosomal inversions and species differences: when are genes affecting adaptive divergence and reproductive isolation expected to reside within inversions? Evolution (N. Y). 63:3061–3075.

Felice, R. N., and A. Goswami. 2018. Developmental origins of mosaic evolution in the avian cranium. Proc. Natl. Acad. Sci. U. S. A. 115:555–560.

Feller, A. F., V. Ogi, O. Seehausen, and J. I. Meier. 2021. Identification of a novel sex determining chromosome in cichlid fishes that acts as XY or ZW in different lineages. Hydrobiologia 848:3727–3745.

Fernald, R. D. 2006. Casting a genetic light on the evolution of eyes. Science 313:1914–1918.

Feron, R., M. Zahm, C. Cabau, C. Klopp, C. Roques, O. Bouchez, C. Eché, S. Valière, C. Donnadieu, P. Haffray, A. Bestin, R. Morvezen, H. Acloque, P. T. Euclide, M. Wen, E. Jouano, M. Schartl, J. H. Postlethwait, C. Schraidt, M. R. Christie, W. A. Larson, A. Herpin, and Y. Guiguen. 2020. Characterization of a Y-specific duplication/insertion of the anti-Mullerian hormone type II receptor gene based on a chromosome-scale genome assembly of yellow perch, *Perca flavescens*. Mol. Ecol. Resour. 20:531–543.

Fischer, E. K., J. P. Nowicki, and L. A. O'Connell. 2019. Evolution of affiliation: patterns of convergence from genomes to behaviour. Philos. Trans. R. Soc. B Biol. Sci. 374:20180242.

Fisher, R. A. 1930. The genetical theory of natural selection. Clarendon Press, Oxford.

Fitzgerald, D. M., and S. M. Rosenberg. 2019. What is mutation? a chapter in the series: how microbes "jeopardize" the modern synthesis. PLoS Genet. 15:1–14.

Fitzpatrick, M. J., Y. Ben-Shahar, H. M. Smid, L. E. M. Vet, G. E. Robinson, and M. B. Sokolowski. 2005. Candidate genes for behavioural ecology. Trends Ecol. Evol. 20:96–104.

Ford, C., K. Jones, P. Polani, and J. De Almeida. 1959. A sex-chromosome anomaly in a case of gonadal dysgenesis (Turner's syndrome). Lancet 1:711–713.

Fricke, H. W. 1983. Social control of sex: field experiments with the anemonefish *Amphiprion bicinctus*. Z. Tierpsychol. 61:71–77.

Futuyma, D. J. 1986. Evolutionary biology. 2nd ed. Sinauer Associates Inc., Sunderland, Mass.

Gamble, T., J. Coryell, T. Ezaz, J. Lynch, D. P. Scantlebury, and D. Zarkower. 2015. Restriction site-associated DNA sequencing (RAD-seq) reveals an extraordinary number of transitions among gecko sex-determining systems. Mol. Biol. Evol. 32:1296–1309.

Gavrilets, S. 2003. Perspective: models of speciation: what have we learned in 40 years. Evolution (N. Y). 57:2197–2215.

Genner, M. J., and G. F. Turner. 2012. Ancient hybridization and phenotypic novelty within Lake Malawi's cichlid fish radiation. Mol. Biol. Evol. 29:195–206.

Gompel, N., and B. Prud'homme. 2009. The causes of repeated genetic evolution. Dev. Biol. 332:36–47.

Gould, S. J. 1989. Wonderful life: the burgess shale and the nature of history. W. W. Norton & Co., New York.

Gourbiere, S. 2004. How do natural and sexual selection contribute to sympatric speciation? J. Evol. Biol. 17:1297–1309.

Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. Di Palma, B. W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman, and A. Regev. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol., doi: 10.1038/nbt.1883.

Graham, A. M., J. L. Peters, R. E. Wilson, V. Muñoz-Fuentes, A. J. Green, D. A. Dorfsman, T. H. Valqui, K. Winker, and K. G. Mccracken. 2021. Adaptive introgression of the beta-globin cluster in two Andean waterfowl. Heredity (Edinb). 127:107–123.

Grant, P. R., and B. R. Grant. 2019. Hybridization increases population variation during adaptive radiation. Proc. Natl. Acad. Sci. U. S. A. 116:23216–23224.

Grant, P. R., and B. R. Grant. 2014. Synergism of natural selection and introgression in the origin of a new species. Am. Nat. 183:671–681.

Grant, P. R., B. R. Grant, J. A. Markert, L. F. Keller, and K. Petren. 2004. Convergent evolution of Darwin's finches caused by introgressive hybridization and selection. Evolution (N. Y). 58:1588–1599.

Grant, P. R., B. R. Grant, and K. Petren. 2005. Hybridization in the recent past. Am. Nat. 166:56–67.

Graves, J. A. M. 2006. Sex chromosome specialization and degeneration in mammals. Cell 124:901–914.

Greenway, R., N. Barts, C. Henpita, A. P. Brown, L. A. Rodriguez, C. M. Rodríguez Peña, S. Arndt, G. Y. Lau, M. P. Murphy, L. Wu, D. Lin, M. Tobler, J. L. Kelley, and J. H. Shaw. 2020. Convergent

evolution of conserved mitochondrial pathways underlies repeated adaptation to extreme environments. Proc. Natl. Acad. Sci. U. S. A. 117:16424–16430.

Gross, J. B. 2016. Convergence and parallelism in *Astyanax* cave-dwelling fish. Pp. 105–119 in Evolutionary Biology. Cham.

Gross, J. B., R. Borowsky, and C. J. Tabin. 2009. A novel role for *Mc1r* in the parallel evolution of depigmentation in independent populations of the cavefish *Astyanax mexicanus*. PLoS Genet 5:1000326.

Guerrero, R. F., and M. Kirkpatrick. 2014. Local adaptation and the evolution of chromosome fusions. Evolution (N. Y). 68:2747–2756.

Gurevich, A., V. Saveliev, N. Vyahhi, and G. Tesler. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075.

Haas, B. J., A. Papanicolaou, M. Yassour, M. Grabherr, P. D. Blood, J. Bowden, M. B. Couger, D. Eccles, B. Li, M. Lieber, M. D. Macmanes, M. Ott, J. Orvis, N. Pochet, F. Strozzi, N. Weeks, R. Westerman, T. William, C. N. Dewey, R. Henschel, R. D. Leduc, N. Friedman, and A. Regev. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc., doi: 10.1038/nprot.2013.084.

Haley, C. S., and S. A. Knott. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity (Edinb). 69:315–324.

Hall, R. 2001. Cenozoic reconstructions of SE Asia and the SW Pacific: changing patterns of land and sea. Faunal Flor. Migr. Evol. SE Asia–Australasia 35–56.

Hamilton, J. A., and J. M. Miller. 2016. Adaptive introgression as a resource for management and genetic conservation in a changing climate. Conserv. Biol. 30:33–41.

Handsaker, Wysoker, Fennell, Ruan, Homer, Abecasis, and 1000 Genome Project Data Processing Subgrou. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079.

Hansen, T. F. 2003. Is modularity necessary for evolvability? remarks on the relationship between pleiotropy and evolvability. Biosystems 69:83–94.

Hattori, R. S., Y. Murai, M. Oura, S. Masuda, S. K. Majhi, T. Sakamoto, J. I. Fernandino, G. M. Somoza, M. Yokota, and C. A. Strüssmann. 2012. A Y-linked anti-Müllerian hormone duplication takes over a critical role in sex determination. Proc. Natl. Acad. Sci. U. S. A. 109:2955–9.

Herder, F., R. K. Hadiaty, and A. W. Nolte. 2012. Pelvic-fin brooding in a new species of riverine ricefish (Atherinomorpha: Beloniformes: Adrianichthyidae) from Tana Toraja, Central Sulawesi, Indonesia. Raffles Bull. Zool. 60:467–476.

Hermisson, J., and P. S. Pennings. 2005. Soft sweeps: molecular population genetics of adaptation from standing genetic variation. Genetics 169:2335–2352.

Heule, C., W. Salzburger, and A. Böhne. 2014. Genetics of sexual development: an evolutionary playground for fish. Genetics 196:579–591.

Hilgers, L., O. Roth, A. W. Nolte, A. Schüller, T. Spanke, J. M. Flury, I. V. Utama, J. Altmüller, D. Wowor, B. Misof, F. Herder, A. Böhne, and J. Schwarzer. 2022. Inflammation and convergent placenta gene co-option contributed to a novel reproductive tissue. Curr. Biol. 32:715-724.

Hilgers, L., and J. Schwarzer. 2019. The untapped potential of medaka and its wild relatives. Elife 8:1–14.

Hill, J., E. D. Enbody, M. E. Pettersson, C. G. Sprehn, D. Bekkevold, A. Folkvord, L. Laikre, G. Kleinau, P. Scheerer, and L. Andersson. 2019. Recurrent convergent evolution at amino acid residue 261 in fish rhodopsin. Proc. Natl. Acad. Sci. 116:18473–18478.

Hoekstra, H. E., R. J. Hirschmann, R. A. Bundey, P. A. Insel, and J. P. Crossland. 2006. A single amino acid mutation contributes to adaptive beach mouse color pattern. Science 313:101–104.

Horoiwa, M., I. F. Mandagi, N. Sutra, J. Montenegro, F. Y. Tantu, K. W. A. Masengi, A. J. Nagano, J. Kusumi, N. Yasuda, and K. Yamahira. 2021. Mitochondrial introgression by ancient admixture between two distant lacustrine fishes in Sulawesi Island. PLoS One 16:1–14.

Huxley, J. 1942. Evolution. The modern synthesis. George Alien & Unwin Ltd., London.

Irisarri, I., P. Singh, S. Koblmüller, J. Torres-Dowdall, F. Henning, P. Franchini, C. Fischer, A. R. Lemmon, E. M. Lemmon, G. G. Thallinger, C. Sturmbauer, and A. Meyer. 2018. Phylogenomics uncovers early hybridization and adaptive loci shaping the radiation of Lake Tanganyika cichlid fishes. Nat. Commun. 9:3159.

IUCN. 2019. The IUCN Red List of Threatened Species.

Iwamatsu, T., H. Kobayashi, M. Sato, and M. Yamashita. 2008. Reproductive role of attaching filaments on the egg envelope in *Xenopoecilus sarasinorum* (Adrianichthidae, Teleostei). J. Morphol. 269:745–750.

Iwamatsu, T., H. Kobayashi, Y. Shibata, M. Sato, N. Tsuji, and K. Takakura. 2007. Oviposition cycle in the oviparous fish *Xenopoecilus sarasinorum*. Zoolog. Sci. 24:1122–1127.

Jeffries, D. L., G. Lavanchy, R. Sermier, M. J. Sredl, I. Miura, A. Borzée, L. N. Barrow, D. Canestrelli, P.-A. Crochet, C. Dufresnes, J. Fu, W.-J. Ma, C. M. Garcia, K. Ghali, A. G. Nicieza, R. P. O'Donnell, N. Rodrigues, A. Romano, Í. Martínez-Solano, I. Stepanyan, S. Zumbach, A. Brelsford, and N. Perrin. 2018. A rapid rate of sex-chromosome turnover and non-random transitions in true frogs. Nat. Commun. 9:4088.

Jones, A. G., and J. C. Avise. 1997. Microsatellite analysis of maternity and the mating system in the Gulf pipefish Syngnathus scovelli, a species with male pregnancy and sex-role reversal. Mol. Ecol. 6:203–213.

Jones, C. D. 2001. Extension of the Castle - Wright effective factor estimator to sex linkage and haplodiploidy. J. Hered. 92:274–276.

Jones, F. C., M. G. Grabherr, Y. F. Chan, P. Russell, E. Mauceli, J. Johnson, R. Swofford, M. Pirun, M. C. Zody, S. White, E. Birney, S. Searle, J. Schmutz, J. Grimwood, M. C. Dickson, R. M. Myers, C. T. Miller, B. R. Summers, A. K. Knecht, S. D. Brady, H. Zhang, A. A. Pollen, T. Howes, C. Amemiya, J. Baldwin, T. Bloom, D. B. Jaffe, R. Nicol, J. Wilkinson, E. S. Lander, F. Di Palma, K. Lindblad-Toh, and D. M. Kingsley. 2012. The genomic basis of adaptive evolution in threespine sticklebacks. Nature 484:55–61.

Jones, K. E., L. Benitez, K. D. Angielczyk, and S. E. Pierce. 2018a. Adaptation and constraint in the evolution of the mammalian backbone. BMC Evol. Biol. 18:172.

Jones, M. R., L. S. Mills, P. C. Alves, C. M. Callahan, J. M. Alves, D. J. R. Lafferty, F. M. Jiggins, and J.

D. Jensen. 2018b. Adaptive introgression underlies polymorphic seasonal camouflage in snowshoe hares. Science 360:1355–1358.

Jones, M. R., L. S. Mills, J. D. Jensen, and J. M. Good. 2020. Convergent evolution of seasonal camouflage in response to reduced snow cover across the snowshoe hare range*. Evolution (N. Y). 74:2033–2045.

Jordan, D. S. 1905. The origin of species through isolation. Science 22:545–562.

Junier, T., and E. M. Zdobnov. 2010. The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. Bioinformatics 26:1669–1670.

Kalendar, R., D. Lee, and A. H. Schulman. 2011. Java web tools for PCR, in silico PCR, and oligonucleotide assembly and analysis. Genomics 98:137–144.

Kalyaanamoorthy, S., B. Q. Minh, T. K. F. Wong, A. von Haeseler, and L. S. Jermiin. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14:587–589.

Kasahara, M., K. Naruse, S. Sasaki, Y. Nakatani, W. Qu, B. Ahsan, T. Yamada, Y. Nagayasu, K. Doi, Y. Kasai, T. Jindo, D. Kobayashi, A. Shimada, A. Toyoda, Y. Kuroki, A. Fujiyama, T. Sasaki, A. Shimizu, S. Asakawa, N. Shimizu, S. I. Hashimoto, J. Yang, Y. Lee, K. Matsushima, S. Sugano, M. Sakaizumi, T. Narita, K. Ohishi, S. Haga, F. Ohta, H. Nomoto, K. Nogata, T. Morishita, T. Endo, T. Shin-I, H. Takeda, S. Morishita, and Y. Kohara. 2007. The medaka draft genome and insights into vertebrate genome evolution. Nature 447:714–719.

Katoh, K., and D. M. Standley. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30:772–780.

Kidd, M. R., N. Duftner, S. Koblmüller, C. Sturmbauer, and H. A. Hofmann. 2012. Repeated parallel evolution of parental care strategies within *Xenotilapia*, a genus of cichlid fishes from Lake Tanganyika. PLoS One 7:e31236.

Kim, H. S., B. Y. Lee, J. Han, C. B. Jeong, D. S. Hwang, M. C. Lee, H. M. Kang, D. H. Kim, D. Lee, J. Kim, I. Y. Choi, and J. S. Lee. 2018. The genome of the marine medaka *Oryzias melastigma*. Mol. Ecol. Resour. 18:656-665.

Kirkpatrick, M., and N. Barton. 2006. Chromosome inversions, local adaptation and speciation. Genetics 173:419–434.

Klingenberg, C. P. 2008. Morphological integration and developmental modularity. Annu. Rev. Ecol. Evol. Syst. 39:115–132.

Klüver, N., F. Pfennig, I. Pala, K. Storch, M. Schlieder, A. Froschauer, H. O. Gutzeit, and M. Schartl. 2007. Differential expression of anti-Müllerian hormone (*amh*) and anti-Müllerian hormone receptor type II (*amhrII*) in the teleost medaka. Dev. Dyn. 236:271–281.

Knapp, R. 2003. Endocrine mediation of vertebrate male alternative reproductive tactics: the next generation of studies. Integr. Comp. Biol. 43:658–668.

Knaus, B. J., and N. J. Grünwald. 2017. vcfr : a package to manipulate and visualize variant call format data in R. Mol. Ecol. Resour. 17:44–53.

Kocher, T. D., J. A. Conroy, K. R. McKaye, and J. R. Stauffer. 1993. Similar morphologies of cichlid fish in Lakes Tanganyika and Malawi are due to convergence. Mol. Phylogenet. Evol. 2:158–165.

Kocot, K. M., M. R. Citarella, L. L. Moroz, and K. M. Halanych. 2013. PhyloTreePruner: a phylogenetic tree-based approach for selection of orthologous sequences for phylogenomics. Evol. Bioinform. Online 9:429–35.

Kofler, R., P. Orozco-terWengel, N. De Maio, R. V. Pandey, V. Nolte, A. Futschik, C. Kosiol, and C. Schlötterer. 2011. PoPoolation: a toolbox for population genetic analysis of next generation sequencing data from pooled individuals. PLoS One 6:e15925.

Kosswig, C. 1964. Polygenic sex determination. Experientia 20:190–199.

Kottelat, M. 1990. Synopsis of the endangered Buntingi (Osteichthyes: Adranichtyidae and Oryziidae) of Lake Poso, Central Sulawesi, Indonesia, with a new reproductive guild and descriptions of three new species. Ichthyol. Explor. Freshwaters 1:46–67.

Krzywinski, M., J. Schein, I. Birol, J. Connors, R. Gascoyne, D. Horsman, S. J. Jones, and M. A. Marra. 2009. Circos: an information aesthetic for comparative genomics. Genome Res. 19:1639–45.

Kück, P. 2009. ALICUT: a Perlscript which cuts ALISCORE identified RSS. Department of Bioinformatics, Zoologisches Forschungs Museum A. Koenig (ZFMK), Bonn.

Kück, P., and G. C. Longo. 2014. FASconCAT-G: extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. Front. Zool., doi: 10.1186/s12983-014-0081-x.

Kück, P., K. Meusemann, J. Dambach, B. Thormann, B. M. von Reumont, J. W. Wägele, and B. Misof. 2010. Parametric and non-parametric masking of randomness in sequence alignments can be improved and leads to better resolved trees. Front. Zool. 7:10.

Lai, Y.-T., C. K. L. Yeung, K. E. Omland, E.-L. Pang, Y. Hao, B.-Y. Liao, H.-F. Cao, B.-W. Zhang, C.-F. Yeh, C.-M. Hung, H.-Y. Hung, M.-Y. Yang, W. Liang, Y.-C. Hsu, C.-T. Yao, L. Dong, K. Lin, and S.-H. Li. 2019. Standing genetic variation as the predominant source for adaptation of a songbird. Proc. Natl. Acad. Sci. 116:2152–2157.

Lamichhaney, S., J. Berglund, M. S. Almén, K. Maqbool, M. Grabherr, A. Martinez-Barrio, M. Promerová, C. J. Rubin, C. Wang, N. Zamani, B. R. Grant, P. R. Grant, M. T. Webster, and L. Andersson. 2015. Evolution of Darwin's finches and their beaks revealed by genome sequencing. Nature 518:371–375.

Lamichhaney, S., F. Han, M. T. Webster, L. Andersson, B. R. Grant, and P. R. Grant. 2018. Rapid hybrid speciation in Darwin's finches. Science 359:224–228.

Lander, E. S., and S. Botstein. 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185.

Langmead, B., and S. L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9:357–359.

Larouche, O., M. L. Zelditch, and R. Cloutier. 2018. Modularity promotes morphological divergence in ray-finned fishes. Sci. Rep. 8:7278.

Lavretsky, P., A. Engilis, J. M. Eadie, and J. L. Peters. 2015. Genetic admixture supports an ancient hybrid origin of the endangered Hawaiian duck. J. Evol. Biol. 28:1005–1015.

Lee, B.-Y., M.-S. Kim, B.-S. Choi, A. J. Nagano, D. Wai, T. Au, R. Shiu, S. Wu, Y. Takehana, and J.-S. Lee. 2019. Construction of high-resolution RAD-seq based linkage map, anchoring reference

genome, and QTL Mapping of the sex chromosome in the marine medaka *Oryzias melastigma*. G3 Genes|Genomes|Genetics 9:3537-3545.

Lee, B. Y., J. C. Park, M. S. Kim, B. S. Choi, D. H. Kim, J. S. Lim, S. Yum, U. K. Hwang, G. J. Nah, and J. S. Lee. 2020. The genome of the Java medaka (Oryzias javanicus): Potential for its use in marine molecular ecotoxicology. Mar. Pollut. Bull. 154:111118.

Lettre, G., A. U. Jackson, C. Gieger, F. R. Schumacher, S. I. Berndt, S. Sanna, S. Eyheramendy, B. F. Voight, J. L. Butler, C. Guiducci, T. Illig, R. Hackett, I. M. Heid, K. B. Jacobs, V. Lyssenko, M. Uda, M. Boehnke, S. J. Chanock, L. C. Groop, F. B. Hu, B. Isomaa, P. Kraft, L. Peltonen, V. Salomaa, D. Schlessinger, D. J. Hunter, R. B. Hayes, G. R. Abecasis, H.-E. Wichmann, K. L. Mohlke, and J. N. Hirschhorn. 2008. Identification of ten loci associated with height highlights new biological pathways in human growth. Nat. Genet. 40:584–591.

Li, C., M. Olave, Y. Hou, G. Qin, R. F. Schneider, Z. Gao, X. Tu, X. Wang, F. Qi, A. Nater, A. F. Kautt, S. Wan, Y. Zhang, Y. Liu, H. Zhang, B. Zhang, H. Zhang, M. Qu, S. Liu, Z. Chen, J. Zhong, H. Zhang, L. Meng, K. Wang, J. Yin, L. Huang, B. Venkatesh, A. Meyer, X. Lu, and Q. Lin. 2021. Genome sequences reveal global dispersal routes and suggest convergent genetic adaptations in seahorse evolution. Nat. Commun. 12:1094.

Li, H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27:2987–93.

Liew, W. C., R. Bartfai, Z. Lim, R. Sreenivasan, K. R. Siegfried, and L. Orban. 2012. Polygenic sex determination system in zebrafish. PLoS One 7:e34397.

Lillie, F. R. 1939. General biological introduction. Pp. 3–14 in E. Allen, C. Danforth, and A. Doisy, eds. Sex and internal secretions. Williams and Wilkins Co, Baltimore.

Litsios, G., and N. Salamin. 2014. Hybridisation and diversification in the adaptive radiation of clownfishes. BMC Evol. Biol. 14:245.

Liu, Z., M. Roesti, D. Marques, M. Hiltbrunner, V. Saladin, and C. L. Peichel. 2022. Chromosomal fusions facilitate adaptation to divergent environments in threespine stickleback. Mol. Biol. Evol. 39:msab358.

Losos, J. B. 2004. Adaptation and speciation in Greater Antillean anoles. P. 515 in U. Dieckmann, M. Doebeli, J. A. J. Metz, and D. Tautz, eds. Adaptive speciation. Cambridge University Press, Cambridge.

Losos, J. B. 2011. Convergence, adaptation, and constraint. Evolution (N. Y). 65:1827–1840.

Losos, J.B. 2017. Improbable destinies: fate, chance, and the future of evolution. Riverhead Books, New York.

Losos, J. B. 2009. Lizards in an evolutionary tree: ecology and adaptive radiation of anoles. University of California Press, Oakland, California.

Losos, J. B., T. R. Jackman, A. Larson, K. de Queiroz, and L. Rodríguez-Schettino. 1998. Contingency and determinism in replicated adaptive radiations of island lizards. Science 279:2115–2118.

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland.

Ma, Y., J. Wang, Q. Hu, J. Li, Y. Sun, L. Zhang, R. J. Abbott, J. Liu, and K. Mao. 2019. Ancient introgression drives adaptation to cooler and drier mountain habitats in a cypress species complex. Commun. Biol. 2:213.

Mackay, T. F. C. 2001. The genetic architecture of quantitative traits. Annu. Rev. Genet. 35:303–339.

Magoc, T., and S. L. Salzberg. 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27:2957–2963.

Maher, B. 2008. Personal genomes: the case of the missing heritability. Nature 456:18–21.

Mahler, D. L., T. Ingram, L. J. Revell, and J. B. Losos. 2013. Exceptional convergence on the macroevolutionary landscape in island lizard radiations. Science 341:292–296.

Mai, U., and S. Mirarab. 2018. TreeShrink: fast and accurate detection of outlier long branches in collections of phylogenetic trees. BMC Genomics 19:272.

Malinsky, M., R. J. Challis, A. M. Tyers, S. Schiffels, Y. Terai, B. P. Ngatunga, E. A. Miska, R. Durbin, M. J. Genner, and G. F. Turner. 2015. Genomic islands of speciation separate cichlid ecomorphs in an East African crater lake. Science 350:1493–1498.

Malinsky, M., M. Matschiner, and H. Svardal. 2021. Dsuite - Fast D -statistics and related admixture evidence from VCF files. Mol. Ecol. Resour. 21:584–595.

Mallet, J. 2007. Hybrid speciation. Nature 446:279–283.

Mallet, J. 2005. Hybridization as an invasion of the genome. Trends Ecol. Evol. 20:229-237.

Mallet, J., N. Besansky, and M. W. Hahn. 2016. How reticulated are species? BioEssays 38:140–149.

Malmstrøm, M., M. Matschiner, O. K. Tørresen, K. S. Jakobsen, and S. Jentoft. 2017. Whole genome sequencing data and de novo draft assemblies for 66 teleost species. Sci. Data 4:160132.

Mandagi, I. F., R. Kakioka, J. Montenegro, H. Kobayashi, K. W. A. Masengi, N. Inomata, A. J. Nagano, A. Toyoda, S. Ansai, M. Matsunami, R. Kimura, J. Kitano, J. Kusumi, and K. Yamahira. 2021. Species divergence and repeated ancient hybridization in a Sulawesian lake system. J. Evol. Biol. 1–14.

Mandagi, I. F., D. F. Mokodongan, R. Tanaka, and K. Yamahira. 2018. A new riverine ricefish of the genus Oryzias (Beloniformes, Adrianichthyidae) from Malili, Central Sulawesi, Indonesia. Copeia 106:297–304.

Mank, J. E. 2009. Sex chromosomes and the evolution of sexual dimorphism: lessons from the genome. Am. Nat. 173:141–150.

Mank, J. E., and J. C. Avise. 2009. Evolutionary diversity and turn-over of sex determination in teleost fishes. Sex. Dev. 3:60–67.

Mank, J. E., and J. C. Avise. 2006. The evolution of reproductive and genomic diversity in ray-finned fishes: insights from phylogeny and comparative analysis. J. Fish Biol. 69:1–27.

Manni, M., M. R. Berkeley, M. Seppey, F. A. Simão, and E. M. Zdobnov. 2021. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. Mol. Biol. Evol. 38:4647–4654.

Marín, I., and B. S. Baker. 1998. The evolutionary dynamics of sex determination. Science

281:1990–4.
Marques, D. A., J. I. Meier, and O. Seehausen. 2019. A combinatorial view on speciation and adaptive radiation. Trends Ecol. Evol. 34:531–544.
Martin, H., F. Carpentier, S. Gallina, C. Godé, E. Schmitt, A. Muyle, G. A. B. Marais, and P. Touzet. 2019. Evolution of young sex chromosomes in two dioecious sister plant species with distinct sex determination systems. Genome Biol. Evol. 11:350–361.
Martin, S. H., K. K. Dasmahapatra, N. J. Nadeau, C. Salazar, J. R. Walters, F. Simpson, M. Blaxter, A. Manica, J. Mallet, and C. D. Jiggins. 2013. Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. Genome Res. 23:1817–28.
Martin, S. H., J. W. Davey, and C. D. Jiggins. 2015. Evaluating the use of ABBA–BABA statistics to locate introgressed loci. Mol. Biol. Evol. 32:244–257.
Mavárez, J., C. A. Salazar, E. Bermingham, C. Salcedo, C. D. Jiggins, and M. Linares. 2006. Speciation by hybridization in *Heliconius* butterflies. Nature 441:868–871.
Mayr, E. 1963. Animal species and evolution. Harvard University Press, Cambridge.
Mayr, E. 1947. Ecological factors in speciation. Evolution (N. Y). 1:263–288.
Mayr, E. 1942. Systematics and the origin of species, from the viewpoint of a zoologist. Harvard University Press, Cambridge.
Meier, J. I., D. A. Marques, S. Mwaiko, C. E. Wagner, L. Excoffier, and O. Seehausen. 2017. Ancient hybridization fuels rapid cichlid fish adaptive radiations. Nat. Commun. 8:1–11.
Misof, B., S. Liu, K. Meusemann, R. S. Peters, A. Donath, C. Mayer, P. B. Frandsen, J. Ware, T. Flouri, R. G. Beutel, O. Niehuis, M. Petersen, F. Izquierdo-Carrasco, T. Wappler, J. Rust, A. J. Aberer, U. Aspöck, H. Aspöck, D. Bartel, A. Blanke, S. Berger, A. Böhm, T. R. Buckley, B. Calcott, J. Chen, F. Friedrich, M. Fukui, M. Fujita, C. Greve, P. Grobe, S. Gu, Y. Huang, L. S. Jermiin, A. Y. Kawahara, L. Krogmann, M. Kubiak, R. Lanfear, H. Letsch, Y. Li, Z. Li, J. Li, H. Lu, R. Machida, Y. Mashimo, P. Kapli, D. D. McKenna, G. Meng, Y. Nakagaki, J. L. Navarrete-Heredia, M. Ott, Y. Ou, G. Pass, L. Podsiadlowski, H. Pohl, B. M. Von Reumont, K. Schütte, K. Sekiya, S. Shimizu, A. Slipinski, A. Stamatakis, W. Song, X. Su, N. U. Szucsich, M. Tan, X. Tan, M. Tang, J. Tang, G. Timelthaler, S. Tomizuka, M. Trautwein, X. Tong, T. Uchifune, M. G. Walzl, B. M. Wiegmann, J. Wilbrandt, B. Wipfler, T. K. F. Wong, Q. Wu, G. Wu, Y. Xie, S. Yang, Q. Yang, D. K. Yeates, K. Yoshizawa, Q. Zhang, R. Zhang, W. Zhang, Y. Zhang, J. Zhao, C. Zhou, L. Zhou, T. Ziesmann, S. Zou, Y. Li, X. Xu, Y. Zhang, H. Yang, J. Wang, J. Wang, K. M. Kjer, and X. Zhou. 2014. Phylogenomics resolves the timing and pattern of insect evolution. Science 346:763–767.
Misof, B., and K. Misof. 2009. A Monte Carlo approach successfully identifies randomness in multiple sequence alignments: a more objective means of data exclusion. Syst. Biol. 58:21–34.
Mitchell-Olds, T., J. H. Willis, and D. B. Goldstein. 2007. Which evolutionary processes influence natural genetic variation for phenotypic traits? Nat. Rev. Genet. 8:845–856.
Mokodongan, D. F., and K. Yamahira. 2015. Origin and intra-island diversification of Sulawesi endemic Adrianichthyidae. Mol. Phylogenet. Evol. 93:150–160.
Montenegro, J., S. Fujimoto, S. Ansai, A. J. Nagano, M. Sato, Y. Maeda, R. Tanaka, K. W. A.

Masengi, R. Kimura, J. Kitano, and K. Yamahira. 2022. Genetic basis for the evolution of pelvic-fin brooding, a new mode of reproduction, in a Sulawesian fish. Mol. Ecol., doi: 10.1111/mec.16555.

Moore, E. C., and R. B. Roberts. 2013. Polygenic sex determination. Curr. Biol. 23:510–512.

Morgulis, A., E. M. Gertz, A. A. Schäffer, and R. Agarwala. 2006. A fast and symmetric DUST implementation to mask low-complexity DNA sequences. J. Comput. Biol. 13:1028–1040. Mary Ann Liebert, Inc. 2 Madison Avenue Larchmont, NY 10538 USA.

Moyer, J. T., and A. Nakazono. 1978. Protandrous hermaphroditism in six species of the anemonefish genus *Amphiprion* in Japan. Japanese J. Ichthyol. 25:101–106.

Mundy, N. I., N. S. Badcock, T. Hart, K. Scribner, K. Janssen, and N. J. Nadeau. 2004. Conserved genetic basis of a quantitative plumage trait involved in mate choice. Science 303:1870–1873.

Murphy, W. J., N. M. Foley, K. R. Bredemeyer, J. Gatesy, and M. S. Springer. 2021. Phylogenomics and the genetic architecture of the placental mammal radiation. Annu. Rev. Anim. Biosci. 9:1–25.

Myers, E. W., G. G. Sutton, A. L. Delcher, I. M. Dew, D. P. Fasulo, M. J. Flanigan, S. A. Kravitz, C. M. Mobarry, K. H. J. Reinert, K. A. Remington, E. L. Anson, R. A. Bolanos, H.-H. Chou, C. M. Jordan, A. L. Halpern, S. Lonardi, E. M. Beasley, R. C. Brandon, L. Chen, P. J. Dunn, Z. Lai, Y. Liang, D. R. Nusskern, M. Zhan, Q. Zhang, X. Zheng, G. M. Rubin, M. D. Adams, and J. C. Venter. 2000. A whole-genome assembly of *Drosophila*. Science 287:2196–2204.

Myosho, T., Y. Takehana, S. Hamaguchi, and M. Sakaizumi. 2015. Turnover of sex chromosomes in Celebensis group medaka fishes. G3 Genes|Genomes|Genetics 5:2685–2691.

Nachman, M. W., H. E. Hoekstra, and S. L. D'Agostino. 2003. The genetic basis of adaptive melanism in pocket mice. Proc. Natl. Acad. Sci. 100:5268–5273.

Naruse, K. 1996. Classification and phylogeny of fishes of the genus *Oryzias* and its relatives. Fish Biol. J. MEDAKA 8:1–9.

Naruse, K., S. Fukamachi, H. Mitani, M. Kondo, T. Matsuoka, S. Kondo, N. Hanamura, Y. Morita, K. Hasegawa, R. Nishigaki, A. Shimada, H. Wada, T. Kusakabe, N. Suzuki, M. Kinoshita, A. Kanamori, T. Terado, H. Kimura, M. Nonaka, and A. Shima. 2000. A detailed linkage map of medaka, *Oryzias latipes*: comparative genomics and genome evolution. Genetics 154:1773–84.

Nelson, T. C., and W. A. Cresko. 2018. Ancient genomic variation underlies repeated ecological adaptation in young stickleback populations. Evol. Lett. 2:9–21.

Nguyen, L.-T., H. A. Schmidt, A. von Haeseler, and B. Quang Minh. 2014. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32:268–274.

Noor, M. A. F., K. L. Grams, L. A. Bertucci, and J. Reiland. 2001. Chromosomal inversions and the reproductive isolation of species. Proc. Natl. Acad. Sci. 98:12084–12088.

Norris, L. C., B. J. Main, Y. Lee, T. C. Collier, A. Fofana, A. J. Cornel, and G. C. Lanzaro. 2015. Adaptive introgression in an African malaria mosquito coincident with the increased usage of insecticide-treated bed nets. Proc. Natl. Acad. Sci. U. S. A. 112:815–20.

Novick, A., and W. F. Doolittle. 2019. How microbes “jeopardize” the modern synthesis. PLoS Genet. 15:3–5.

O’Donnell, D. R., A. Parigi, J. A. Fish, I. Dworkin, and A. P. Wagner. 2014. The roles of standing genetic variation and evolutionary history in determining the evolvability of anti-predator strategies. PLoS One 9:e100163.

Ohno, S. 1979. Major sex-determining genes. Springer-Verlag, Berlin.

Ohno, S. 1974. Protochordata, cyclostomata, and pisces. P. 92 in M. S. Arsham, M. J. Barch, and H. J. Lawce, eds. Animal cytogenetics. Gebrüder Borntraeger, Berlin.

Ohno, S. 1967. Sex chromosomes and sex-linked genes. Springer-Verlag, Berlin.

Ohta, T. 1992. The nearly neutral theory of molecular evolution. Annu. Rev. Ecol. Syst. 23:263–286.

Olson, E. C., and R. L. Miller. 1958. Morphological integration. University of Chicago Press, Chicago.

Ornelas-García, C., O. Domínguez-Domínguez, and I. Doadrio. 2008. Evolutionary history of the fish genus *Astyanax* Baird & Girard (1854) (Actinopterygii, Characidae) in Mesoamerica reveals multiple morphological homoplasies. BMC Evol. Biol. 8:340.

Orr, H. A. 2000. Adaptation and the cost of complexity. Evolution (N. Y). 54:13–20.

Paciulli, L. M., and C. M. Cromer. 2018. Secondary sex characteristics. Pp. 1–9 in J. Vonk, and T. Shackelford, eds. Encyclopedia of Animal Cognition and Behavior. Springer International Publishing, Cham.

Palmer, D. H., T. F. Rogers, R. Dean, and A. E. Wright. 2019. How to identify sex chromosomes and their turnover. Mol. Ecol. 28:4709–4724.

Palmer, J., and J. Stajich. 2019. nextgenusfs/funannotate: funannotate v1.5.3.

Pan, Q., R. Feron, A. Yano, R. Guyomard, E. Jouanno, E. Vigouroux, M. Wen, J.-M. Busnel, J. Bobe, J.-P. Concordet, H. Parrinello, L. Journot, C. Klopp, J. Lluch, C. Roques, J. Postlethwait, M. Schartl, A. Herpin, and Y. Guiguen. 2019. Identification of the master sex determining gene in Northern pike (*Esox lucius*) reveals restricted sex chromosome differentiation. PLOS Genet. 15:e1008013.

Papa, R., D. D. Kapan, B. A. Counterman, K. Maldonado, D. P. Lindstrom, R. D. Reed, H. F. Nijhout, T. Hrbek, and W. O. McMillan. 2013. Multi-allelic major effect genes interact with minor effect QTLs to control adaptive color pattern variation in *Heliconius erato*. PLoS One 8:e57033.

Papadopulos, A. S. T., W. J. Baker, D. Crayn, R. K. Butlin, R. G. Kynast, I. Hutton, and V. Savolainen. 2011. Speciation with gene flow on Lord Howe Island. Proc. Natl. Acad. Sci. 108:13188–13193.

Paradis, E., and K. Schliep. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35:526–528.

Pardo-Diaz, C., C. Salazar, S. W. Baxter, C. Merot, W. Figueiredo-Ready, M. Joron, W. O. McMillan, and C. D. Jiggins. 2012. Adaptive Introgression across species boundaries in *Heliconius* butterflies. PLoS Genet. 8:e1002752.

Parenti, L. R. 2008. A phylogenetic analysis and taxonomic revision of rice fishes, *Oryzias* and

relatives (Beloniformes, Adrianichthydae). Zool. J. Linn. Soc. 154:494–610.
Park, T., B. Mennecart, L. Costeur, C. Grohé, and N. Cooper. 2019. Convergent evolution in toothed whale cochleae. BMC Evol. Biol. 19:195.
Patek, S. N., and R. L. Caldwell. 2005. Extreme impact and cavitation forces of a biological hammer: strike forces of the peacock mantis shrimp *Odontodactylus scyllarus*. J. Exp. Biol. 208:3655–3664.
Patek, S. N., W. L. Korff, and R. L. Caldwell. 2004. Deadly strike mechsanism of a mantis shrimp. Nature 428:819–820.
Pattengale, N. D., M. Alipour, O. R. P. Bininda-Emonds, B. M. E. Moret, and A. Stamatakis. 2010. How many bootstrap replicates are necessary? J. Comput. Biol. 17:337–354.
Paul-Prasanth, B., R. K. Bhandari, T. Kobayashi, R. Horiguchi, Y. Kobayashi, M. Nakamoto, Y. Shibata, F. Sakai, M. Nakamura, and Y. Nagahama. 2013. Estrogen oversees the maintenance of the female genetic program in terminally differentiated gonochorists. Sci. Rep. 3:2862.
Peichel, C. L., S. R. McCann, J. A. Ross, A. F. S. Naftaly, J. R. Urton, J. N. Cech, J. Grimwood, J. Schmutz, R. M. Myers, D. M. Kingsley, and M. A. White. 2020. Assembly of the threespine stickleback Y chromosome reveals convergent signatures of sex chromosome evolution. Genome Biol. 21:177.
Petersen, M., K. Meusemann, A. Donath, D. Dowling, S. Liu, R. S. Peters, L. Podsiadlowski, A. Vasilikopoulos, X. Zhou, B. Misof, and O. Niehuis. 2017. Orthograph: a versatile tool for mapping coding nucleotide sequences to clusters of orthologous genes. BMC Bioinformatics 18:1–10.
Phelan, P. L., and T. C. Baker. 1987. Evolution of male pheromones in moths: reproductive isolation through sexual selection? Science 235:205–207.
Pickrell, J. K., and J. K. Pritchard. 2012. Inference of population splits and mixtures from genome-wide allele frequency data. PLoS Genet. 8:e1002967.
Pokornà, M., and L. Kratochvìl. 2009. Phylogeny of sex-determining mechanisms in squamate reptiles: are sex chromosomes an evolutionary trap? Zool. J. Linn. Soc. 156:168–183.
Poulton, E. B. 1904. What is a species. Proc. Entomol. Soc. London 115:77–116.
Price, M. N., P. S. Dehal, and A. P. Arkin. 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol. Biol. Evol. 26:1641–1650.
Quinlan, A. R., and I. M. Hall. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842.
Racimo, F., S. Sankararaman, R. Nielsen, and E. Huerta-Sánchez. 2015. Evidence for archaic adaptive introgression in humans. Nat. Rev. Genet. 16:359–371.
Raff, R. A. 1996. The shape of life: genes, development, and the evolution of animal form. The University of Chicago Press, Chicago.
Reid, N. M., D. A. Proestou, B. W. Clark, W. C. Warren, J. K. Colbourne, J. R. Shaw, S. I. Karchner, M. E. Hahn, D. Nacci, M. F. Oleksiak, D. L. Crawford, and A. Whitehead. 2016. The genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish. Science 354:1305–1308.
Rice, W. R. 1984. Sex chromosomes and the evolution of sexual dimorphism. Evolution (N. Y).

38:735.
Riedl, R. 1977. A systems-analytical approach to macro-evolutionary phenomena. Q. Rev. Biol. 52:351–370.
Rieseberg, L. H. 2001. Chromosomal rearrangements and speciation. Trends Ecol. Evol. 16:351–358.
Rieseberg, L. H. 2009. Evolution: replacing genes and traits through hybridization. Curr. Biol. 19:R119-R122.
Rintelen, T. von, A. B. Wilson, A. Meyer, and M. Glaubrecht. 2004. Escalation and trophic specialization drive adaptive radiation of freshwater gastropods in ancient lakes on Sulawesi, Indonesia. Proc. R. Soc. London. Ser. B Biol. Sci. 271:2541–2549.
Ritland, K., C. Newton, and H. D. Marshall. 2001. Inheritance and population structure of the white-phased "Kermode" black bear. Curr. Biol. 11:1468–1472.
Roberts, N. B., S. A. Juntti, K. P. Coyle, B. L. Dumont, M. K. Stanley, A. Q. Ryan, R. D. Fernald, and R. B. Roberts. 2016. Polygenic sex determination in the cichlid fish *Astatotilapia burtoni*. BMC Genomics 17:835.
Robertson, A. 1967. The nature of quantitative genetic variation. Pp. 265–280 in A. Brink, ed. Heritage from Mendel. University of Wisconsin Press, Madison.
Ronco, F., M. Roesti, and W. Salzburger. 2019. A functional trade-off between trophic adaptation and parental care predicts sexual dimorphism in cichlid fish. Proc. R. Soc. B Biol. Sci. 286:20191050.
Rüber, L., E. Verheyen, and A. Meyer. 1999. Replicated evolution of trophic specializations in an endemic cichlid fish lineage from Lake Tanganyika. Proc. Natl. Acad. Sci. U. S. A. 96:10230–10235.
Sayyari, E., J. B. Whitfield, and S. Mirarab. 2018. DiscoVista: interpretable visualizations of gene tree discordance. Mol. Phylogenet. Evol. 122:110–115.
Schluter, D., and L. M. Nagel. 1995. Parallel speciation by natural selection. Am. Nat. 146:292–301.
Schultheis, C., A. Böhne, M. Schartl, J. N. Volff, and D. Galiana-Arnoux. 2009. Sex determination diversity and sex chromosome evolution in poeciliid fish. Sex. Dev. 3:68–77.
Schumer, M., G. G. Rosenthal, and P. Andolfatto. 2018. What do we mean when we talk about hybrid speciation? Heredity (Edinb). 120:379–382.
Sedlazeck, F. J., P. Rescheneder, and A. Von Haeseler. 2013. NextGenMap: fast and accurate read mapping in highly polymorphic genomes. Bioinformatics 29:2790-2791.
Seehausen, O. 2013. Conditions when hybridization might predispose populations for adaptive radiation. J. Evol. Biol. 26:279–281.
Seehausen, O. 2004. Hybridization and adaptive radiation. Trends Ecol. Evol. 19:198–207.
Shapiro, B. J., J. B. Leducq, and J. Mallet. 2016. What Is Speciation? PLoS Genet. 12:1–14.
Shapiro, M. D., M. E. Marks, C. L. Peichel, B. K. Blackman, K. S. Nereng, B. Jónsson, D. Schluter, and D. M. Kingsley. 2004. Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. Nature 428:717–723.
Shine, R. 1989. Ecological causes for the evolution of sexual dimorphism: a review of the

evidence. Q. Rev. Biol. 64:419–461.
Simão, F. A., R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, and E. M. Zdobnov. 2015. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31:3210-3212.
Smith, S. D., and M. D. Rausher. 2011. Gene loss and parallel evolution contribute to species difference in flower color. Mol. Biol. Evol. 28:2799–2810.
Solís-Lemus, C., and C. Ané. 2016. Inferring phylogenetic networks with maximum pseudolikelihood under incomplete lineage sorting. PLOS Genet. 12:e1005896.
Solís-Lemus, C., P. Bastide, and C. Ané. 2017. PhyloNetworks: a package for phylogenetic networks. Mol. Biol. Evol. 34:3292–3298.
Song, W., Y. Xie, M. Sun, X. Li, C. K. Fitzpatrick, F. Vaux, K. G. O'Malley, Q. Zhang, J. Qi, and Y. He. 2021. A duplicated *amh* is the master sex-determining gene for Sebastes rockfish in the Northwest Pacific. Open Biol. 11:210063.
Spanke, T., L. Hilgers, B. Wipfler, J. M. Flury, A. W. Nolte, I. V. Utama, B. Misof, F. Herder, and J. Schwarzer. 2021. Complex sexually dimorphic traits shape the parallel evolution of a novel reproductive strategy in Sulawesi ricefishes (Adrianichthyidae). BMC Ecol. Evol. 21:57.
Springer, M. S., J. A. W. Kirsch, and J. A. Case. 1997. The chronicle of marsupial evolution. Pp. 129–162 in T. J. Givnish and K. J. Sytsma, eds. Molecular evolution and adaptive radiation. Cambridge University Press, Cambridge.
Streelman, J. T., R. C. Albertson, and T. D. Kocher. 2003. Genome mapping of the orange blotch colour pattern in cichlid fishes. Mol. Ecol. 12:2465–2471.
Stuart-Fox, D., and A. Moussali. 2007. Sex-specific ecomorphological variation and the evolution of sexual dimorphism in dwarf chameleons (*Bradypodion* spp.). J. Evol. Biol. 20:1073–1081.
Sutra, N., J. Kusumi, J. Montenegro, H. Kobayashi, S. Fujimoto, K. W. A. Masengi, A. J. Nagano, A. Toyoda, M. Matsunami, R. Kimura, and K. Yamahira. 2019. Evidence for sympatric speciation in a Wallacean ancient lake. Evolution (N. Y). 73:1898–1915.
Suyama, M., D. Torrents, and P. Bork. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. 34:W609-W612.
Suzuki, T., N. Suzuki, and K. Tojo. 2019. Parallel evolution of an alpine type ecomorph in a scorpionfly: independent adaptation to high-altitude environments in multiple mountain locations. Mol. Ecol. 28:mec.15119.
Takagi, N., and M. Sasaki. 1974. A phylogenetic study of bird karyotypes. Chromosoma 46:91–120.
Takehana, Y., M. Zahm, C. Cabau, C. Klopp, C. Roques, O. Bouchez, C. Donnadieu, C. Barrachina, L. Journot, M. Kawaguchi, S. Yasumasu, S. Ansai, K. Naruse, K. Inoue, C. Shinzato, M. Schartl, Y. Guiguen, and A. Herpin. 2020. Genome sequence of the euryhaline javafish medaka, *Oryzias javanicus* : a small aquarium fish model for studies on adaptation to salinity. G3 Genes|Genomes|Genetics 10:907–915.
Temeles, E. J., I. L. Pan, J. L. Brennan, and J. N. Horwitt. 2000. Evidence for ecological causation of sexual dimorphism in a hummingbird. Science 289:441–443.
Terai, Y., O. Seehausen, T. Sasaki, K. Takahashi, S. Mizoiri, T. Sugawara, T. Sato, M. Watanabe, N.

Konijnendijk, H. D. J. Mrosso, H. Tachida, H. Imai, Y. Shichida, and N. Okada. 2006. Divergent selection on opsins drives Incipient speciation in Lake Victoria cichlids. PLoS Biol. 4:e433.

Terekhanova, N. V., M. D. Logacheva, A. A. Penin, T. V. Neretina, A. E. Barmintseva, G. A. Bazykin, A. S. Kondrashov, and N. S. Mugue. 2014. Fast evolution from precast bricks: genomics of young freshwater populations of threespine stickleback *Gasterosteus aculeatus*. PLoS Genet. 10:e1004696.

The Heliconius Genome Consortium, K. K. Dasmahapatra, J. R. Walters, A. D. Briscoe, J. W. Davey, A. Whibley, N. J. Nadeau, A. V. Zimin, D. S. T. Hughes, L. C. Ferguson, S. H. Martin, C. Salazar, J. J. Lewis, S. Adler, S.-J. Ahn, D. A. Baker, S. W. Baxter, N. L. Chamberlain, R. Chauhan, B. A. Counterman, T. Dalmay, L. E. Gilbert, K. Gordon, D. G. Heckel, H. M. Hines, K. J. Hoff, P. W. H. Holland, E. Jacquin-Joly, F. M. Jiggins, R. T. Jones, D. D. Kapan, P. Kersey, G. Lamas, D. Lawson, D. Mapleson, L. S. Maroja, A. Martin, S. Moxon, W. J. Palmer, R. Papa, A. Papanicolaou, Y. Pauchet, D. A. Ray, N. Rosser, S. L. Salzberg, M. A. Supple, A. Surridge, A. Tenger-Trolander, H. Vogel, P. A. Wilkinson, D. Wilson, J. A. Yorke, F. Yuan, A. L. Balmuth, C. Eland, K. Gharbi, M. Thomson, R. A. Gibbs, Y. Han, J. C. Jayaseelan, C. Kovar, T. Mathew, D. M. Muzny, F. Ongeri, L.-L. Pu, J. Qu, R. L. Thornton, K. C. Worley, Y.-Q. Wu, M. Linares, M. L. Blaxter, R. H. ffrench Constant, M. Joron, M. R. Kronforst, S. P. Mullen, R. D. Reed, S. E. Scherer, S. Richards, J. Mallet, W. O. McMillan, and C. D. Jiggins. 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. Nature 487:94.

Theron, E., K. Hawkins, E. Bermingham, R. E. Ricklefs, and N. I. Mundy. 2001. The molecular basis of an avian plumage polymorphism in the wild: a melanocortin-1-receptor point mutation is perfectly associated with the melanic plumage morph of the bananaquit, *Coereba flaveola*. Curr. Biol. 11:550–557.

Uller, T., and H. Helanterä. 2011. From the origin of sex-determining factors to the evolution of sex-determining systems. Q. Rev. Biol. 86:163–180.

Untergasser, A., I. Cutcutache, T. Koressaar, J. Ye, B. C. Faircloth, M. Remm, and S. G. Rozen. 2012. Primer3--new capabilities and interfaces. Nucleic Acids Res. 40:e115.

Urban, S., A. Nater, A. Meyer, and C. F. Kratochwil. 2021. Different sources of allelic variation drove repeated color pattern divergence in cichlid fishes. Mol. Biol. Evol. 38:465–477.

Utama, I. V., I. F. Mandagi, S. A. Lawelle, K. W. A. Masengi, K. Watanabe, N. Sawada, A. J. Nagano, J. Kusumi, and K. Yamahira. 2022. Deeply divergent freshwater fish species within a single river system in central Sulawesi. Mol. Phylogenet. Evol. 173:107519.

Van Belleghem, S. M., C. Vangestel, K. De Wolf, Z. De Corte, M. Möst, P. Rastas, L. De Meester, and F. Hendrickx. 2018. Evolution at two time frames: polymorphisms from an ancient singular divergence event fuel contemporary parallel evolution. PLOS Genet. 14:e1007796.

van der Valk, T., P. Pečnerová, D. Díez-del-Molino, A. Bergström, J. Oppenheimer, S. Hartmann, G. Xenikoudakis, J. A. Thomas, M. Dehasque, E. Sağlıcan, F. R. Fidan, I. Barnes, S. Liu, M. Somel, P. D. Heintzman, P. Nikolskiy, B. Shapiro, P. Skoglund, M. Hofreiter, A. M. Lister, A. Götherström, and L. Dalén. 2021. Million-year-old DNA sheds light on the genomic history of mammoths. Nature 591:265–269.

Van Ooijen, J. W. 2006. JoinMap 4: Software for the calculation of genetic linkage maps in

experimental populations. Kyazma BV, Wageningen, Netherlands.
Vandeputte, M., M. Dupont-Nivet, H. Chavanne, and B. Chatain. 2007. A polygenic hypothesis for sex determination in the European sea bass *Dicentrarchus labrax*. Genetics 176:1049–1057.
Veale, A. J., and M. A. Russello. 2017. Genomic changes associated with reproductive and migratory ecotypes in sockeye salmon (*Oncorhynchus nerka*). Genome Biol. Evol. 9:2921–2939.
Vicoso, B., and D. Bachtrog. 2015. Numerous transitions of sex chromosomes in diptera. PLOS Biol. 13:e1002078.
Villoutreix, R., C. F. de Carvalho, V. Soria-Carrasco, D. Lindtke, M. De-la-Mora, M. Muschick, J. L. Feder, T. L. Parchman, Z. Gompert, and P. Nosil. 2020. Large-scale mutation in the evolution of a gene complex for cryptic coloration. Science 369:460–466.
Volff, J.-N., and M. Schartl. 2001. Variability of genetic sex determination in poeciliid fishes. Genetica 111:101–110.
Voloch, C. M., J. F. Vilela, L. Loss-Oliveira, and C. G. Schrago. 2013. Phylogeny and chronology of the major lineages of New World hystricognath rodents: insights on the biogeography of the Eocene/Oligocene arrival of mammals in South America. BMC Res. Notes 6:160.
von Rintelen, T., and M. Glaubrecht. 2006. Rapid evolution of sessility in an endemic species flock of the freshwater bivalve *Corbicula* from ancient lakes on Sulawesi, Indonesia. Biol. Lett. 2:73–77.
Wagner, A. 2008. Robustness and evolvability: a paradox resolved. Proc. R. Soc. B Biol. Sci. 275:91–100.
Wagner, G. P. 1996. Homologues, natural kinds and the evolution of modularity. Am. Zool. 36:36–43.
Wagner, G. P. 2014. Homology, genes, and evolutionary innovation. Princeton University Press, Princeton.
Wagner, G. P., and L. Altenberg. 1996. Perspective: complex adaptations and the evolution of evolvability. Evolution (N. Y). 50:967–976.
Wagner, G. P., M. Pavlicev, and J. M. Cheverud. 2007. The road to modularity. Nat. Rev. Genet. 8:921–931.
Wagner, G. P., and J. Zhang. 2011. The pleiotropic structure of the genotype–phenotype map: the evolvability of complex organisms. Nat. Rev. Genet. 12:204–213.
Ward, B. J., and C. van Oosterhout. 2016. HybridCheck : software for the rapid detection, visualization and dating of recombinant regions in genome sequence data. Mol. Ecol. Resour. 16:534–539.
Wasiljew, B. D., J. Pfaender, B. Wipfler, M. Gabelaia, I. V. Utama, L. L. Wantania, and F. Herder. 2021. Sexual dimorphism in an adaptive radiation: does intersexual niche differentiation result in ecological character displacement? Ecol. Evol. 11:14615–14629.
Waterhouse, R. M., F. Tegenfeldt, J. Li, E. M. Zdobnov, and E. V. Kriventseva. 2013. OrthoDB: a hierarchical catalog of animal, fungal and bacterial orthologs. Nucleic Acids Res. 41:D358–D365.
Waters, J. M., and G. A. McCulloch. 2021. Reinventing the wheel? reassessing the roles of gene

flow, sorting and convergence in repeated evolution. Mol. Ecol. 30:4162–4172.
Weimerskirch, H., M. Le Corre, H. Gadenne, D. Pinaud, A. Kato, Y. Ropert-Coudert, and C.-A. Bost. 2009. Relationship between reversed sexual dimorphism, breeding investment and foraging ecology in a pelagic seabird, the masked booby. Oecologia 161:637–649.
Whitney, Randell, and Rieseberg. 2006. Adaptive introgression of herbivore resistance traits in the weedy sunflower Helianthus annuus. Am. Nat. 167.
Williams, G. C. 1966. Adaptation and natural selection. Princeton University Press, Princeton.
Wilson, C. A., S. K. High, B. M. McCluskey, A. Amores, Y. Yan, T. A. Titus, J. L. Anderson, P. Batzel, M. J. Carvan, M. Schartl, and J. H. Postlethwait. 2014. Wild sex in zebrafish: loss of the natural sex determinant in domesticated strains. Genetics 198:1291–1308.
Wilson, M. E. J., and S. J. Moss. 1999. Cenozoic palaeogeographic evolution of Sulawesi and Borneo. Palaeogeogr. Palaeoclimatol. Palaeoecol. 145:303–337.
Witt, K. E., and E. Huerta-Sánchez. 2019. Convergent evolution in human and domesticate adaptation to high-altitude environments. Philos. Trans. R. Soc. B 374:20180235.
Wittbrodt, J., A. Shima, and M. Schartl. 2002. Medaka - a model organism from the Far East. Nat. Rev. Genet. 3:53–64.
Wootton, R., and C. Smith. 2014. Reproductive biology of teleost fishes. John Wiley & Sons.
Yamamoto, T. 1975. Medaka: (Killifish). Biology and strains. Series of stock culture in biological field. Keigaku Publishing Company, Tokyo.
Yamanoue, Y., D. H. E. Setiamarga, and K. Matsuura. 2010. Pelvic fins in teleosts: structure, function and evolution. J. Fish Biol. 77:1173–1208.
Yeaman, S. 2013. Genomic rearrangements and the evolution of clusters of locally adaptive loci. Proc. Natl. Acad. Sci. 110:E1743–E1751.
Yeaman, S., and A. Jarvis. 2006. Regional heterogeneity and gene flow maintain variance in a quantitative trait within populations of lodgepole pine. Proc. R. Soc. B Biol. Sci. 273:1587–1593.
Yeaman, S., and M. C. Whitlock. 2011. The genetic architecture of adaptation under migration-selection balance. Evolution (N. Y). 65:1897–1911.
Young, W. 1961. Sex and internal secretions. Williams and Wilkins Co, Baltimore.
Yu, Y., C. Blair, and X. He. 2020. RASP 4: ancestral state reconstruction tool for multiple genes and characters. Mol. Biol. Evol. 37:604–606.
Zhang, C., M. Rabiee, E. Sayyari, and S. Mirarab. 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics 19:153.
Zhang, X., J. G. Rayner, M. Blaxter, and N. W. Bailey. 2021. Rapid parallel adaptation despite gene flow in silent crickets. Nat. Commun. 12:50.
Zhang, Z., S. Schwartz, L. Wagner, and W. Miller. 2000. A greedy algorithm for aligning DNA sequences. J. Comput. Biol. 7:203–214.

List of supplementary files

Supplementary files can be found on a USB drive on the back cover in print and in an attached folder in the online version.

i) Supplementary material to publication in chapter 2 of this thesis, published on Dryad

https://datadryad.org/stash/dataset/doi:10.5061/dryad.4mw6m90c5

Chap2_FLURY_2022_CW_Rscript.R: Rscript to calculate Castle-Wright estimator

Chap2_README_FLURY_2022_CW_Rscript.txt: Read-me file published with Rscript

Chap2_FLURY_2022_DATA.csv: Data used for the study

Chap2_README_FLURY_2022_DATA.txt Read-me file published with data

ii) Supplementary material to manuscript in chapter 3 of this thesis

Chap3_TabS1.xlsx: Taxon sampling

Chap3_TabS5.xlsx: Deleted non-corresponding genes

Chap3_TabS6.xlsx: Ortholog set reference genomes

Chap3_TabS7.xlsx: Statistics outlier genes

Chap3_TabS8_aa.xlsx: Masking results amino acid level

Chap3_TabS8_nt.xlsx: Masking results nucleotide level

Chap3_TabS9.txt: Substitution models for gene trees

iii) Supplementary material to manuscript in chapter 4 of this thesis

Chap4_TabS1.xlsx: Primer sequences used in custom GT-seq protocol

Chap4_TabS2.xlsx: List of samples used for GT-seq

Acknowledgements

Ein Doktortitel ist der höchste akademische Grad, den ein Mensch erreichen kann. Daher ist es in meinen Augen gerechtfertigt, dass er mit einigen Kosten verbunden ist. Das Ergebnis, allerdings, war hoffentlich die Mühe wert.

Ich bedanke mich bei meinen Gutachtern, Bernhard Misof und Arne W. Nolte und den beiden weiteren Promotionskommissionsmitgliedern Diana Imhof und Gerhard von der Emde, für das Bewerten und Lesen dieser Doktorarbeit.

Des Weiteren bedanke ich mich bei meiner direkten Betreuerin Julia Schwarzer. Es war mir eine Freude, deine erste Doktorandin zu sein und ich hoffe du kannst noch viele angehende WissenschaftlerInnen so begleiten und inspirieren wie mich.

In den weiteren Nennungen werde ich eine alphabetische Reihenfolge verwenden.

Ich bedanke mich bei Anja Bodenheim, Camilla Di Nizo, Malte Dittmann, Claudia Etzbauer, Timo Hartmann und dem Team der Tierpflege, Sandra Kukowka, Salome Mwaiko, Lars Podsiadlowski, Jana Thormann und Serkan Wesel für ihre Unterstützung im Molekular Labor, bei der Fotostation und bei der Fischzucht und -pflege.

Ich möchte mich bei all meinen KoautorInnen bedanken, bei den folgenden besonders. Ihr wart eine Inspiration in vielerlei Hinsicht, und es war sehr bereichernd mit euch zusammenzuarbeiten. Danke schön an Fabian Herder, Leon Hilgers, Karen Meusemann und Kristin Tietje.

An die Fischmenschen, Jan Möhring, Juliane Romahn, Alina Schüller, Tobias Spanke und Ben Wasiljew, vielen Dank für eure Freundschaft, professionellen und emotionalen Beistand. Ich werde unsere Mittagspausen schmerzlichst vermissen.

Ich möchte grossen Dank an meine Familie aussprechen, meine Eltern Karin und Manfred Flury, die mich in jeglicher Hinsicht unterstützt haben. Ausserdem möchte ich meiner Schwester und meinem Schwager danken, Daniel und Ronja Mader, die bei fast jedem Umzug geholfen haben. Ihr alle seid ein Geschenk.

Zum Schluss zu dir, lieber Dario. Dank deiner Leichtigkeit, Zuversicht und unerschütterlichem Glauben an mich hast du mich mit durch diese Zeit getragen. Danke für deine Unterstützung, deinen Witz und deine Ratschläge, die Gold wert waren.

Attachment

Jana M. Flury, Leon Hilgers, Fabian Herder, Tobias Spanke, Bernhard Misof, Daisy Wowor, Farnis Boneka, Letha Louisiana Wantania, Daniel F. Mokodongan, Christoph Mayer, Arne W. Nolte, Julia Schwarzer (2022) **The genetic basis of a novel reproductive strategy in Sulawesi ricefishes: how modularity and a low number of loci shape pelvic brooding.** Evolution (N.Y.) 76(5):1033 -1051.

Jana M. Flury, Karen Meusemann, Sebastian Martin, Leon Hilgers, Tobias Spanke, Astrid Böhne, Fabian Herder, Daniel F. Mokodongan, Janine Altmüller, Daisy Wowor, Bernhard Misof, Arne W. Nolte, Julia Schwarzer (2022) **Contribution of hybridization to the evolution of a derived reproductive strategy in ricefishes.** bioRxiv. https://doi.org/10.1101/2022.07.05.498713

Jana M. Flury*, Kristin Tietje*, Julia Schwarzer, Arne W. Nolte (2022) **Sex determination in crosses between two Celebensis group medaka species.** In prep for Journal of Evolutionary Biology

* equal contribution

doi:10.1111/evo.14475

The genetic basis of a novel reproductive strategy in Sulawesi ricefishes: How modularity and a low number of loci shape pelvic brooding

Jana M. Flury,[1,2] Leon Hilgers,[1,3] Fabian Herder,[1] Tobias Spanke,[1] Bernhard Misof,[1] Daisy Wowor,[4] Farnis Boneka,[5] Letha Louisiana Wantania,[1,5] Daniel F. Mokodongan,[4] Christoph Mayer,[1] Arne W. Nolte,[6] and Julia Schwarzer[1,7]

[1]*Leibniz Institute for the Analysis of Biodiversity Change, Zoological Research Museum Alexander Koenig, Bonn, Germany*

[2]*E-mail: j.flury@leibniz-lib.de*

[3]*LOEWE Centre for Translational Biodiversity Genomics, Frankfurt, Germany*

[4]*Museum Zoologicum Bogoriense, Research Center for Biosystematic and Evolution, National Research and Innovation Agency (BRIN), Cibinong, West Java, Indonesia*

[5]*Faculty of Fisheries and Marine Science, Sam Ratulangi University, Manado, Indonesia*

[6]*Carl von Ossietzky Universität, Oldenburg, Germany*

[7]*E-mail: j.schwarzer@leibniz-lib.de*

Received June 20, 2021

Accepted January 29, 2022

The evolution of complex phenotypes like reproductive strategies is challenging to understand, as they often depend on multiple adaptations that only jointly result in a specific functionality. Sulawesi ricefishes (*Adrianichthyidae*) evolved a reproductive strategy termed as pelvic brooding. In contrast to the more common transfer brooding, female pelvic brooders carry an egg bundle connected to their body for weeks until the fry hatches. To examine the genetic architecture of pelvic brooding, we crossed the pelvic brooding *Oryzias eversi* and the transfer brooding *Oryzias nigrimas* (species divergence time: ∼3.6 my). We hypothesize, that a low number of loci and modularity have facilitated the rapid evolution of pelvic brooding. Traits associated to pelvic brooding, like rib length, pelvic fin length, and morphology of the genital papilla, were correlated in the parental species but correlations were reduced or lost in their F1 and F2 hybrids. Using the Castle-Wright estimator, we found that generally few loci underlie the studied traits. Further, both parental species showed modularity in their body plans. In conclusion, morphological traits related to pelvic brooding were based on a few loci and the mid-body region likely could evolve independently from the remaining body parts. Both factors presumably facilitated the evolution of pelvic brooding.

KEY WORDS: Castle-Wright estimator, complex phenotypes, genetic architecture, modularity, pelvic brooding.

Investigating the genetic architecture of complex phenotypic traits is a crucial step in understanding the evolution of adaptations and species diversity (Moczek 2008; Moriyama and Koshiba-Takeuchi 2018). Reproductive strategies might serve as a good example of complex phenotypes, since they are often the result of complex adaptations that entail life history changes and often correlate with phenotypic and behavioral adaptations in both males and females. For example, parental care likely drove the convergent evolution of endothermy in birds and mammals (Farmer 2000). Endothermy allows the incubation of offspring (Ruben 1995) and relies on an increased resting metabolism, additional insulation, and an internal sensor, that

regulates metabolic heat production (Price and Dzialowski 2018). Similarly, lactation in mammals required the evolution of mammary glands, milk protein composition, and nursing behavior (Lefèvre et al. 2010). Besides mammals and birds, teleost fishes show a remarkable diversity of reproductive strategies. Parental care, a behavior enhancing offspring survival, evolved independently over 30 times (Mank et al. 2005; Wootton and Smith 2014) in teleost fishes ranging from nest guarding to live bearing (Wootton and Smith 2014). Most fish species are egg laying with external fertilization, whereas substantially fewer species (∼2%) are viviparous with internal fertilization and internal bearing (Blackburn 1999; Grier and Uribe 2005; Mank et al. 2005; Wootton and Smith 2014). Only a few species carry their eggs externally, which is called external bearing (Wootton and Smith 2014). An exceptional form of external bearing is known from ricefishes (*Beloniformes*: *Adrianichthyidae*). In all ricefishes, the externally fertilized eggs remain connected to the female's body after spawning (Yamamoto 1975). Attaching filaments originate from the surface of the chorion of the fertilized egg. They remain attached to the gonoduct until females of most species (so called "transfer- brooding" species) deposit the eggs in plants or on other substrate within a few hours after spawning (Yamamoto 1975; Wootton and Smith 2014). However, in at least three ricefish species (*Adrianichthys oophorus*, *Oryzias sarasinorum*, and *Oryzias eversi*) from two distantly related lineages (time to most recent common ancestor: ∼16 my, Mokodongan and Yamahira 2015), females carry the eggs up to 18 days until the fry hatches (Iwamatsu et al. 2007). This reproductive strategy was termed "pelvic brooding" by Kottelat (1990).

The reproductive cycle in female pelvic brooding ricefish species is much longer compared to daily spawning in some transfer brooding ricefish species, and can last up to 3 weeks during which ovulation is delayed (Iwamatsu et al. 2007). After spawning, the attaching filaments get entangled in the gonoduct and form a plug, a unique structure that holds the developing egg bundle in place (Iwamatsu et al. 2008). External morphological adaptations in female pelvic brooders led to sexual dimorphism (Spanke et al. 2021). Female pelvic brooders have elongated and thickened pelvic fins and shorter ribs creating a ventral concavity in which the egg bundle is carried (Kottelat 1990; Parenti 2008; Herder et al. 2012; Spanke et al. 2021). While several studies investigated comparative ricefish morphology (Iwamatsu et al. 2008; Parenti 2008; Kottelat 1990; Herder et al. 2012; Spanke et al. 2021), the morphological integration of pelvic brooding traits, the genetic architecture, and also the evolutionary origin of pelvic brooding remain unknown.

Our study aims to investigate the genetic architecture of pelvic brooding. Specifically, to shed light on modularity and potential evolutionary constraints acting on traits that are assumed to underlie pelvic brooding, we studied whether these traits were genetically correlated in the parental species as well as in their crosses and inferred how many genetic loci underlay these traits. We generated crosses between the pelvic brooding species *O. eversi* (Fig. 1A) and the transfer brooding species *O. nigrimas* (Fig. 1B). We focused on female external morphological traits that differ between the parental species and were shown to be connected with reproductive strategy such as the length of pelvic fins and the length of selected ribs forming the ventral concavity (Parenti 1986, 2005; Spanke et al. 2021). We further analyzed the location and position of pelvic fin insertions, overall body shape and size, and pigmentation of the genital papilla as they also differ between the parental species and likely are connected to reproductive strategy. *Oryzias nigrimas*, as well as other transfer brooding species, have a bilobed papilla, whereas females of pelvic brooding species have a single lobed papilla (Parenti 2008) (Fig. 1C). In *O. eversi*, the genital papilla is further heavily pigmented and undergoes tissue-specific changes in gene expression during egg carrying (Hilgers et al. 2021), indicating that it has a function in brooding.

We measured observed segregation variances of all the above-mentioned traits, determined trait cosegregation, and tested modularity of the body plan based on a geometric morphometrics approach. Using the segregation variances, we estimated the number of putative loci underlying phenotypic variation of focal traits with the Castle-Wright estimator (Castle 1921; Lynch and Walsh 1998; Jones 2001). As prior studies in fishes showed that adaptations in pelvic fin morphology are based on few loci (Shapiro et al. 2004; Coyle et al. 2007; Chan et al. 2010), we assume the same to be true for the pelvic fin traits in ricefishes. Further, we hypothesize that traits connected to reproductive strategy will be correlated and might have a pleiotropic genetic basis. We also assume that modularity exists in the body plan of *O. eversi* and *O. nigrimas*. Here, we use the definition of evolutionary modularity (Klingenberg 2008), which means that modules comprise genetically correlated (e.g., a set of pleiotropic genes, Wagner et al. 2007) and functionally coupled traits that are less integrated with other such modules (Wagner et al. 2007; Klingenberg 2008; Adams 2016). Traits belonging to different modules can, thus, evolve with less interference from the rest of the organism, allowing for greater evolvability (Wagner 1996; Wagner and Altenberg 1996; Yang 2001; Hansen 2003; West-Eberhard 2003; Schlosser and Wagner 2004).

Methods

BREEDING OF INDIVIDUALS

Populations of *O. eversi* and *O. nigrimas* were bred in aquaria at the Zoological Research Museum Alexander Koenig in Bonn,

Figure 1. The parental species *O. eversi* (A) and *O. nigrimas* (B) with eggs, the latter will only carry them for a few hours. Differences in the genital papilla (C) between the two species: single lobed and heavily pigmented in *O. eversi* and bilobed and not pigmented in *O. nigrimas*. We measured the total area of the genital papilla (GPA), only the "nipple-like" papilla area (PA) and the pigmented area (PI) of the genital papilla. The pelvic fins (D) vary in length, measured from the base of the second outer fin ray to the very tip of the fin (blue line). Landmarks were set on the ventral side of the females' body (E) to measure the distance between pelvic fin insertions (PF-PF) and to characterize the differences in distances between pelvic fin insertion and genital papilla that define the space and location of the egg cluster. (F) Position of landmarks defining body shape and measurement sections taken for standard length (SL) and rib length (R12). A detailed list of landmarks for both shape analyses can be found in Supporting information. Black scale bars are 1 mm and white scale bars 5 mm long.

the Max-Plank-Institute for Evolutionary Biology in Plön and Carl-von-Ossietzky University in Oldenburg since 2012. To generate experimental crosses, *O. eversi* males were crossed with *O. nigrimas* females. Interspecific F1 hybrid males were then backcrossed with *O. eversi* and *O. nigrimas* females, respectively. Intercrosses were bred by crossing interspecific F1 hybrid males with interspecific F1 hybrid females. In total, 36 *O. nigrimas*, 23 *O. eversi*, 23 interspecific F1 hybrids, 67 intercrosses, 26 *O. nigrimas* backcross, and 227 *O. eversi* backcross females were included in the present study (Supporting information Table S1). The large number of *O. eversi* backcrosses compared to the other two F2 crosses derived from the low number of females present in the other crosses, which exhibited a sex-bias toward males.

MEASUREMENTS OF EXTERNAL MORPHOLOGICAL TRAITS

We focused on morphological traits that are connected to female reproductive behavior in *O. eversi*, sexually dimorphic in this species and distinct from *O. nigrimas* and other transfer brooders (Spanke et al. 2021). That is why all traits used in this study were only measured in females (Fig. 1). All measurements and data are accessible on Dryad (https://doi.org/10.5061/dryad.4mw6m90c5). Total genital papilla area, papilla area, and the pigmented area of the papilla (Fig. 1C), pelvic fin length (Fig. 1D), and the distance between pelvic fins (Fig. 1E), were measured based on photographs scaled by a custom macroscript for ImageJ (Canon EOS 60D, Canon MP-E 65mm 1:2,8, macro ring flash, placed on a vertical stand, manual focus). In more detail, the pelvic fin length was measured from the base of the second lateral fin ray to the very end of the fin in ImageJ version 1.51u (Fig. 1D). The total genital papilla area was measured by outlining the outermost rim of the structure and the papilla area was measured by outlining the "nipple-like skin fold," as described in Yamamoto and Suzuki (1955). To measure the pigmented area of the papilla, we outlined the pigmented part (Fig. 1C). All areas were also measured in ImageJ, version 1.51u. Rib 12 length (R12) was chosen to be measured, because it is shortened in females of pelvic brooding species (Spanke et al. 2021). The landmarks on rib 12 were used to measure its length from vertebra to tip (Fig. 1F). We tested whether the measured traits differ between the parental species using Bonferroni corrected *t*-tests.

GEOMETRIC MORPHOMETRIC ANALYSIS OF OVERALL AND VENTRAL BODY SHAPE

We tested whether overall body shape differs between the two species and the crosses, because overall body shape might be associated with the reproductive strategy, as it was shown in livebearing poecilids (Zúñiga-Vega et al. 2007). However, other life-history traits are also related to body shape in fishes, for example, predator avoidance (Walker 1997; Price et al. 2015) and swimming performance (Langerhans and Reznick 2010). X-rays of the whole body were taken (Faxitron LX 60) (Fig. 1F) and 15 homologous skeletal points on the body, as well as on the vertebrae of ribs 11 and 12 and their tips (in total 19) were defined to place landmarks (Fig. 1F, Supporting information Fig. S1). Further, we investigated the ventral arrangement of pelvic fins, the genital papilla, and the anal fin insertion. The pelvic fins are known to be different in the two species with regard to length and thickness of fin rays (Parenti 2008; Spanke et al. 2021), but their positioning, the fin base width, and the distance between them might differ as a result of selection acting on traits related to pelvic brooding. A wider fin base and a larger distance between the pelvic fins might help to hold the egg bundle in the ventral concavity. Landmarks for the ventral side were placed on photographs on eight homologous points (Canon EOS 60D, Canon MP-E 65mm 1:2,8, macro ring flash, placed on a vertical stand, manual focus) (Fig. 1E, Supporting information Fig. S1).

Principal Component Analysis was done for the whole body landmarks and the ventral side using the function *PlotTangentSpace* from the R package ´geomorph´ (version 3.1.3) (Adams and Otárola-Castillo 2013). The resulting first principal components were used for the correlation analysis and the Castle-Wright estimator if they were significantly different in the parental lines, tested by a Bonferroni-corrected *t*-test. A canonical variate analysis (CVA) was additionally performed to highlight morphological differences between the groups (CVA, R package Morpho, version 2.8, Schlager 2017). Repeatability of the landmarks was tested with a subset of the specimens; detailed methods are described in the Supporting information.

TRAIT CORRELATIONS AND MODULARITY

We calculated trait correlations for each parental species and for each cross to evaluate whether correlations between different traits persist or get lost in the hybrid progeny. We tested 22 *O. nigrimas*, 21 *O. eversi*, 21 interspecific F1 hybrids, 45 intercross, 19 *O. nigrimas* backcross, and 153 *O. eversi* backcross specimens. Individuals with missing data in the analyzed traits were removed. The correlation matrix included the total genital papilla area, the papilla area, the pigmented papilla area, pelvic fin length, distance between pelvic fins, rib 12 length, and the first principal component of the landmark analysis of the whole body (PC1, Fig. 1F) and the ventral side (PC1_vs, Fig. 1E) since they are both significantly different in the parental species and explain most of the variation (see section Results) (R package "corrplot," version 0.84, Wei and Simko 2021). A Pearson rank correlation test (R package "Hmisc," version 4.4-0, Harell 2021) was used to test deviations from no correlation (null-hypothesis) and *p*-values were corrected for multiple testing using Bonferroni-correction. To compare *O. eversi* backcrosses to the other lines, 100 subsamples were collected with 50 randomly chosen *O. eversi* backcross

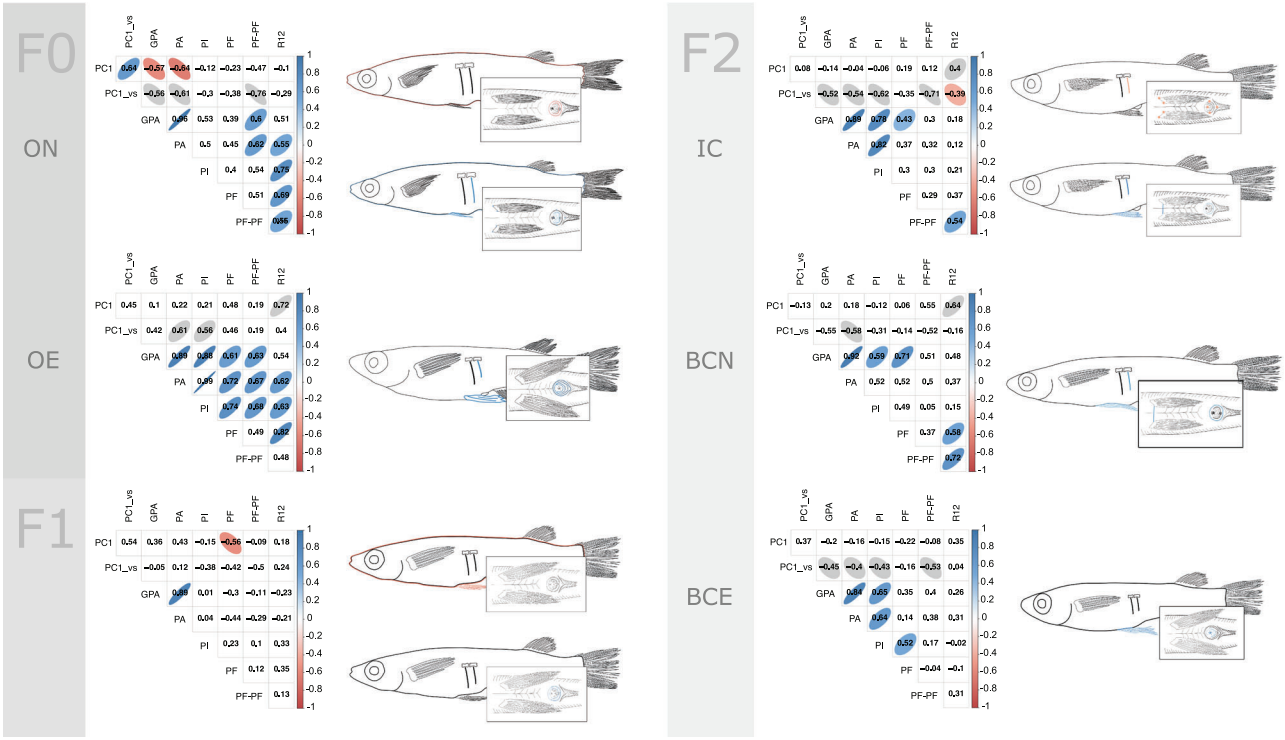


Figure 2. **Correlation plots of parental species and different crosses. All ellipses indicate a significant correlation (*p* value < 0.05). The numbers in the squares show the correlation coefficients, that is, the correlation between two traits. Red ellipses = negative correlation, blue ellipses = positive correlation, grey ellipses = significant correlation between traits of same measurement technique (landmarks). ON = *O. nigrimas*, OE = *O. eversi*, F1 = first generation hybrids (ON female, OE male), IC = intercrosses between F1, BCN = male F1 backcrossed with female ON, BCE = male F1 backcrossed with female OE.**

individuals. If the correlation was present in more than 90% of the subsamples, it was considered robust. All *O. eversi* backcrosses were used to produce Figure 2. We built categories of dependent variables in which traits are expected to correlate since their measurements are not independent. Measurements of the length of rib 12, genital papilla traits, and the distance between pelvic fins were taken from landmarks also present in the shape analyses. Therefore, correlations between PC1 and rib 12 length, between PC1_vs and the genital papilla traits and distance between pelvic fins were regarded as nonindependent. These five correlations were coloured in grey if significant (Fig. 2). To test the independence of total genital papilla area and genital papilla, we calculated the mean percentage of the area of the total genital papilla area covered by the papilla for the parental species and the crosses. If correlations are lost or reduced in interspecific F1 hybrids and intercrosses, a simple pleiotropic basis (based on one or few loci) seems rather unlikely.

We used a graphical modeling approach and plotted all correlations between the different traits in a network to visualize potential functionally interacting traits, that is, functional units (Wagner and Schwenk 2000) among all the measured traits associated with the phenotypic complex trait of pelvic brooding (Fig. 3) (Magwene 2001). In both parental species, traits were considered to be part of the unit if they had more correlations to traits within the unit than to traits outside of it.

In addition, to test for modularity on a broader scale we investigated trait integration based on overall body shape. Therefore, six different a priori modularity hypotheses were tested for landmarks defining the whole body shape of *O. eversi* (n = 20) and *O. nigrimas* (n = 30, Fig. 4). Modularity was assumed when traits had a higher covariance or correlation within a subset of trait variables, that is, a module, compared to trait variables among modules (Adams 2016). The hypotheses were formulated based on functional and anatomical criteria and were divided into three main groups (H1, H2, H3). H1 and H2 had two modules each and H3 had three, whereas H1 + a and H2 + a had three modules, and H3 + a had four: **H1**–module 1: head, module 2: rest of the body. **H2**–module 1: head and pectoral fins, module 2: rest of the body. **H3**–module 1: head, module 2: pectoral fins, module 3: rest of the body. For each group, we further divided one module into two, separated by a line between the rib landmarks and the anal fin insertion: **H1 + a**–module 1: head, module 2: mid-body region, module 3: caudal body with anal, dorsal, and caudal fins. **H2 + a**–module 1: head and pectoral fins, module 2: mid-body region, module 3: caudal body with anal, dorsal, and caudal fins. **H3 + a**–module 1: head, module 2: the pectoral fins,

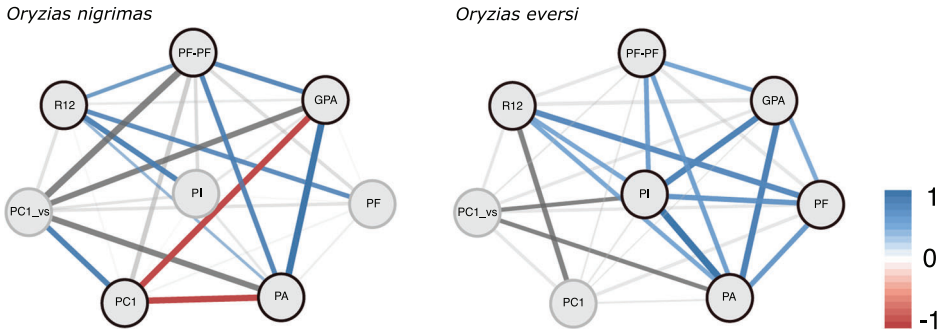


Figure 3. **Correlations between traits in *O. nigrimas* and in *O. eversi*. Colored lines indicate significant correlations, dark grey indicates significant but dependent correlations, light grey indicate nonsignificant. Traits framed in black are part of the functional morphological unit. In *O. nigrimas*, each trait of the functional morphological unit has at least two out of four possible correlations within the unit. PC1_vs, PI, and PF have only one significant, independent correlation and are, hence, not part of the unit. In *O. eversi*, each trait has at least three out of five possible correlations within the unit. PC1_vs and PC1 are not part of the unit because they only have dependent correlations. PF = pelvic fin length, PF-PF = distance between pelvic fins, R12 = rib 12 length, GPA = total genital papilla area, PA = papilla area, PI = pigmented area of the genital papilla, PC1_vs = first principal component of the ventral landmark analysis, PC1 = first principal component of body shape landmark analysis.**

module 3: mid-body region, module 4: caudal body with anal, dorsal, and caudal fins (Fig. 4).

The *CR* coefficients (covariance ratio) were used to determine modularity by calculating pairwise covariances between variables (Adams 2016). A *CR* coefficient of 1 means that the covariation is as large within as between modules. A *CR* coefficient over 1 indicates that covariation is larger between modules than within, which means that no modularity was observed. A *CR* coefficient below 1 means that covariation is larger within a module than between modules and implies that modularity is present. S_{11} and S_{22} are the covariance matrices within modules and S_{12} and S_{21} between modules.

$$\boldsymbol{S} = \begin{bmatrix} \boldsymbol{S}_{11} & \boldsymbol{S}_{12} \\ \boldsymbol{S}_{21} & \boldsymbol{S}_{22} \end{bmatrix}. \quad (1)$$

The covariance ratio is then calculated as follows:

$$\boldsymbol{CR} = \sqrt{\frac{\boldsymbol{trace}\,(\mathbf{S}_{12}\mathbf{S}_{21})}{\sqrt{\boldsymbol{trace}\,(\mathbf{S}^*_{11}\mathbf{S}^*_{11})\,\boldsymbol{trace}\,(\mathbf{S}^*_{22}\mathbf{S}^*_{22})}}}, \quad (2)$$

S^*_{11} and S^*_{22} are the covariance matrices within modules, but the diagonal elements are replaced with zeros.

To assess which hypothesis has the largest effect size, Z_{CR} was calculated. The more negative Z_{CR} is, the larger the effect size. CR_{obs} is the observed covariance ratio, $\hat{\mu}_r$is the expected value of CR if there is no modularity (null hypothesis), and sigma is the SE of the mean, found as the SD of the empirical sampling distribution (Adams and Collyer 2019).

$$Z_{CR} = \frac{CR_{obs} - \hat{\mu}_r}{\hat{\sigma}_r}. \quad (3)$$

Both *CR* and Z_{CR} were computed in R, using the package "geomorph" (version 3.1.3, Adams and Otárola-Castillo 2013).

CORRECTION FOR BODY SIZE DIFFERENCES

Crosses in this study differed in body size (Supporting information Fig. S2). This difference was partly species-specific and partly due to sampling of specimens at different ages. *Oryzias eversi*, *O. nigrimas*, and the interspecific F1 hybrids were used for breeding and were, thus, larger because they were sampled at a higher age. The second-generation hybrids were sampled as soon as they started to breed and were smaller than the parental species and the F1. We accounted for these differences in body size and their effect on the measured traits by using ratios (with the standard length) combined with calculating the square root of two-dimensional measurements. We also utilized this normalization in the variance-based analyses. An alternative would have been to use the residuals of a linear regression with the standard length as explanatory variable. However, taking the residuals for each group separately would result in a mean of zero per group, and calculating the regression over all groups would inflate the parental variances, since the measurements of the parental traits differ more from the linear regression obtained from all measurements than the measurements of the crosses. We, thus, normalized all measures by dividing pelvic fin length, distance between pelvic fins and rib 12 lengths by the standard length. For total genital papilla area, papilla area, and pigmented area, the square root was calculated (since they are two-dimensional structures) and then divided by the standard length. Trait values used for the Castle-Wright estimator should follow a normal distribution (Lynch and Walsh 1998), which was largely fulfilled. PC1, PC3, and PC1_vs were not divided by standard length since the

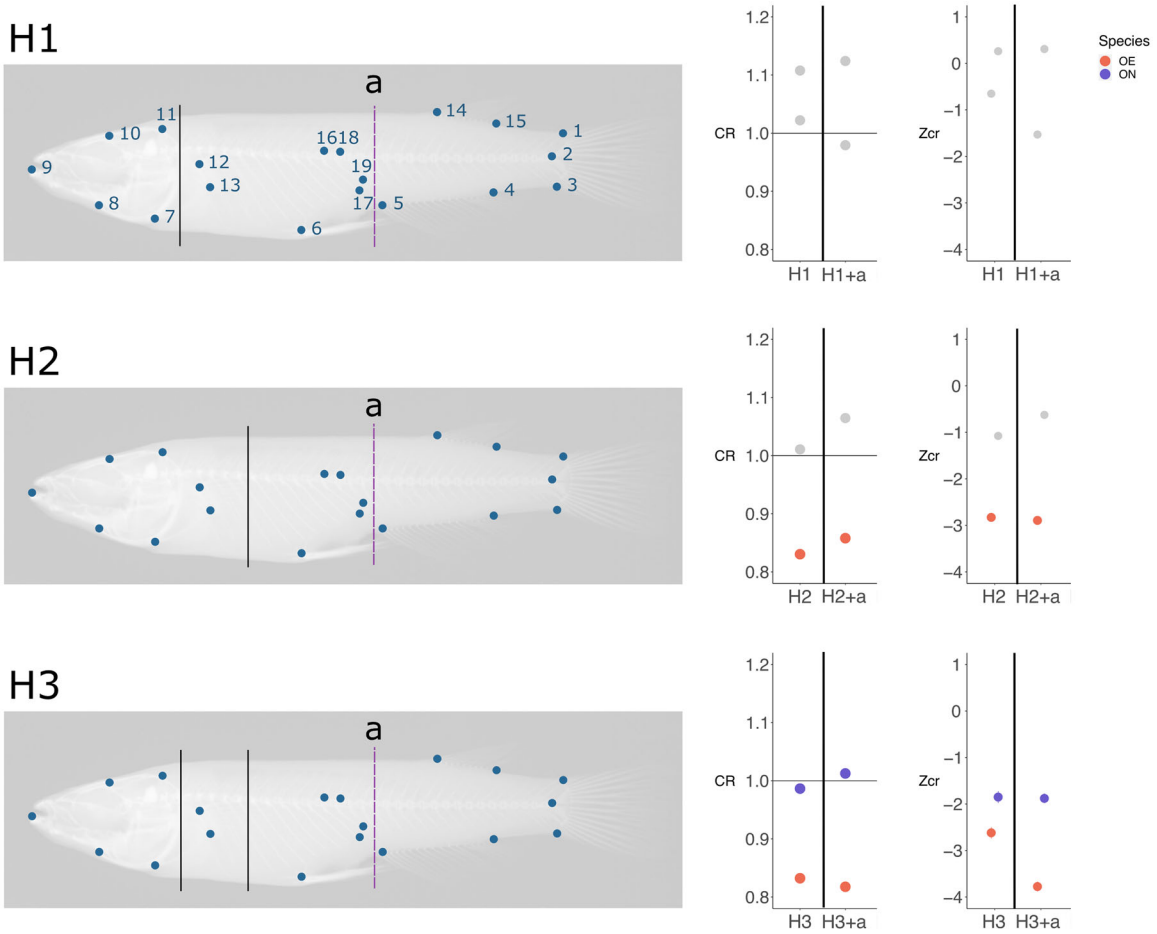


Figure 4. **Modularity hypotheses of the whole body shape using landmarks on the left and the resulting ratios of correlation and effect sizes on the right. The blue points are the landmarks from the body shape analysis. The black bars separate the landmarks assigned a priori to a module. The purple dashed line separates the mid-body from the tail and was tested in addition to the three main hypotheses. For each hypothesis, we calculated the *CR* ratio (significant values in color). *CR* below 1 indicates a valid modularity hypothesis and on the right of each *CR* plot are the effect sizes (Z_{CR}) for each module with standard deviation. The more negative Z_{CR} the larger the effect size.**

Procrustes superimposition implemented in "geomorph" already accounts for size differences.

ALLOMETRY

Ratios of each trait were modeled against standard length to account for allometric effects. If there was no correlation, that is, if ratios were constant for different standard lengths, this means that the measured body part had the same size or length compared to the standard length during the whole life of the fish (Haines 1942; Kozlowski 1996; Stamps et al. 1998; Sîrbulescu et al. 2017). If there was a correlation, the variation measured might be over- or underestimated since the fish used in this study vary in size. Bonferroni correction was used to account for multiple testing.

CASTLE-WRIGHT ESTIMATOR

Numbers of individuals per trait per line (Table 2) varied due to missing data in some measurements (e.g., broken fins or ribs, egg cluster covering the genital papilla). Numbers in *O. eversi* ranges from 21 to 23 individuals, in *O. nigrimas* from 23 to 34, in the F1 from 21 to 23, in the F2 intercrosses from 52 to 58, in *O. eversi* backcrosses from 160 to 218, and in *O. nigrimas* backcrosses from 20 to 24 (Table 2). We included only principal components, which were significantly differentiating the parental lines.

Additive versus dominance-additive model

The calculation of the estimator assumes genetic additivity. Therefore, the joint-scaling test was used to check an additivity- and a dominance-additivity model (Cavalli 1952; Mather and Jinks 1971; Hayman 1960a; Lynch and Walsh 1998). The simple additive model is:

$$\overline{z_i} = \mu_0 + \theta_{Si}\alpha_i^C + e_i, \quad (4)$$

Table 1. p values of allometric effects of all traits in all lines.

	OE	ON	F1	IC	BCE	BCN
PF	1	1	1	0.004	0.100	0.501
GPA	1	0.161	0.654	1	1	1
PA	0.698	0.055	1	1	1	1
PI	0.810	1	1	1	1	1
R12	1	1	1	1	1	1
PF-PF	1	1	1	1	1	1
PC1	0.071	1	1	0.416	0.086	1
PC1_vs	1	1	1	1	1	1

p values are Bonferroni-corrected. Null-hypothesis: ratio of measured traits to SL remains constant. ON = *O. nigrimas*, OE = *O. eversi*, F1 = first generation hybrids, IC = intercrosses between F1, BCN = ON backcrosses, BCE = OE backcrosses. PF = pelvic fin length, PF-PF = distance between pelvic fins, R12 = rib 12 length, GPA = total genital papilla area, PA = papilla area, PI = pigmented area of the genital papilla, PC1_vs = first principal component of the ventral landmark analysis, PC1 = first principal component of body shape landmark analysis.

where $\overline{z_i}$ is the mean of the ith line (here P1 = *O. nigrimas*, P2 = *O. eversi*, F1, F2, BC1 = *O. nigrimas* backcrosses, BC2 = *O. eversi* backcrosses), θ_{Si} its coefficient, and e_i the deviation of the observed mean from the predictions derived from the model. In matrix form, **M** denotes the matrix of coefficients and **a** is the vector of effects of μ_0 and α_1^c.

$$\overline{z} = Ma + e. \tag{5}$$

Assuming the simplest model with all gene action being additive, the coefficients for the effects μ_0 and α^C are given by:

$$\mathbf{M} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0.5 \\ 1 & -0.5 \end{pmatrix} \tag{6}$$

To estimate parameters $\hat{\mu}_0$ and $\hat{\alpha}^c$, the line means have to be weighted since they may vary regarding accuracy. The following formula was used:

$$\hat{\boldsymbol{a}} = \left(\boldsymbol{M}^T \boldsymbol{V}^{-1} \boldsymbol{M}\right)^{-1} \boldsymbol{M}^T \boldsymbol{V}^{-1} \bar{\boldsymbol{z}}, \tag{7}$$

where $\boldsymbol{V}$ is the covariance matrix, with the diagonal elements being the squared standard errors of the means. As **a** became weighted, predicted line means could be calculated and compared to measured line means in a X^2test. The number of degrees of freedom was DF = k-2, so when comparing six lines, DF = 4.

$$\boldsymbol{X}^2 = \sum_{i=1}^{k} \frac{(\bar{\boldsymbol{z}}_i - \hat{\boldsymbol{z}}_i)^2}{\boldsymbol{Var}\,(\bar{\boldsymbol{z}}_i)}. \tag{8}$$

To test the additive-dominant model, **M** was replaced with a matrix including coefficients for dominance:

$$\boldsymbol{M} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & -1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0.5 & 0 \\ 1 & -0.5 & 0 \end{pmatrix} \tag{9}$$

and in the end, the two X^2 tests were compared. DF = 1 in this case.

$$\Lambda = X_A^2 - X_{AD}^2. \tag{10}$$

If both models were rejected, epistatic effects would have to be included.

Segregation estimator

After checking genetic additivity, we proceeded by computing the segregational variance. Again, a joint-scaling test was applied, similar to the one used for the additivity and additivity-dominance models (Hayman 1960b; Lynch and Walsh 1998). Here, it was used on the line variances instead of the means. Therefore, **V** was a matrix with the sampling variance of the variance (v_j) in the diagonal elements and n_j is the number of individuals per line.

$$\frac{2v_j^2}{n_j + 2}. \tag{11}$$

M is the matrix of coefficients, S is the fraction of P1 genes in a line, H is the probability that a member of the line has one P1 and one P2 gene at a locus. The first column contains the values S–(H/2), which is the frequency of individuals having only P1 alleles. In the second column, the frequency of individuals

Table 2. Number of individuals, mean values, and variances for each trait in each line.

	OE			ON			F1			IC			BCE			BCN		
	mean	variance	n	mean	variance	n	mean	variance	n	mean	variance	n	mean	variance	n	mean	variance	n
PC1	−3.93E-02	2.81E-05	20	5.46E-02	2.53E-04	32	2.30E-02	2.70E-05	23	1.93E-02	5.23E-05	55	−1.71E-02	9.34E-05	218	4.91E-02	4.90E-05	24
PC3	9.31E-03	7.59E-05	20	−2.84E-03	1.22E-04	32	5.99E-03	2.65E-05	23	1.02E-03	6.48E-05	55	−1.80E-03	2.64E-04	218	4.34E-03	6.11E-05	24
PC1_vs	−1.29E-01	1.48E-03	27	9.43E-02	1.38E-03	28	−1.08E-02	3.92E-04	25	4.25E-02	1.60E-03	61	−1.49E-02	1.73E-03	181	5.06E-02	8.30E-04	24
GPA	0.0485	4.25E-05	23	0.0314	3.67E-05	27	0.0441	1.96E-05	23	0.0335	5.06E-05	54	0.0384	3.72E-05	168	0.0297	2.11E-05	21
PA	0.0297	3.68E-05	23	0.0243	4.67E-05	27	0.0348	2.46E-05	23	0.0217	2.83E-05	54	0.0238	2.65E-05	168	0.0203	3.19E-05	20
PI	0.0276	2.83E-05	22	0.0052	4.71E-05	25	0.0242	2.15E-05	21	0.0127	6.00E-05	53	0.0173	4.45E-05	166	0.0078	6.76E-05	20
PF	0.1755	7.05E-05	23	0.1145	6.08E-05	26	0.1473	2.97E-05	23	0.1334	3.32E-04	58	0.1495	4.82E-04	179	0.1184	1.62E-04	23
PF-PF	0.0311	4.35E-05	23	0.0199	1.48E-05	25	0.0243	1.94E-05	23	0.024	4.05E-05	53	0.0277	4.63E-05	162	0.022	3.02E-05	23
R12	0.0689	2.51E-05	22	0.1212	8.42E-05	33	0.1025	3.96E-05	21	0.0891	2.31E-04	58	0.0789	1.68E-04	216	0.1084	1.48E-04	24

To correct for body size PF, PF-PF, and r12 were divided by SL. The square root was taken for GPA, PA, and PI since they are two-dimensional and then divided by SL. Principal components were not divided by SL since the implemented Procrustes analysis already accounts for different body sizes. ON = *O. nigrimas*, OE = *O. eversi*, F1 = first generation hybrids, IC = intercrosses between F1, BCN = ON backcrosses, BCE = OE backcrosses. PF = pelvic fin length, PF-PF = distance between pelvic fins, R12 = rib 12 length, GPA = total genital papilla area, PA = papilla area, PI = pigmented area of the genital papilla, PC1_vs = first principal component of the ventral landmark analysis, PC1 = first principal component of body shape landmark analysis.

having only P2 alleles (1 –*S*–[*H*/2]) and in the third column, the frequency of individuals with one P1 and one P2 allele (*H*) is calculated (Lynch and Walsh 1998).

$$\boldsymbol{M} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0.5 & 0.5 & 0 \\ 0.5 & 0.5 & 1 \\ 0.75 & 0.25 & 0.5 \\ 0.25 & 0.75 & 0.5 \end{pmatrix}. \quad (12)$$

Values were again placed in this formula similar to above, with **v** being the vector of sampling variance of the variances.

$$\hat{a} = \left(M^T V^{-1} M\right)^{-1} M^T V^{-1} v. \quad (13)$$

Taking the estimated $\hat{a}$, sampling variances of the variance $\bar{v}$ were recalculated. This computation was iterated until $\hat{a}$ and the X^2value converged and

$$\overline{v} = Ma, \quad (14)$$

$$\boldsymbol{X}^2 = \sum_{j=1}^{k} \frac{(\boldsymbol{v}_j - \hat{\boldsymbol{v}}_j)^2}{2\hat{\boldsymbol{v}}_j^2 / (\boldsymbol{n}_j + 2)}. \quad (15)$$

The first value of the final vector of $\hat{a}$ is the segregation variance of parental species 1 (σ^2_{P1}), the second of parental species 2 (σ^2_{P2}), and the third of all offspring lines (σ^2_S).

Castle-Wright estimator

With the results from the segregational variance, the Castle-Wright estimator was calculated as:

$$\hat{\boldsymbol{n}}_e = \frac{\left(\boldsymbol{\mu}_{P1} - \boldsymbol{\mu}_{P2}\right)^2 - \boldsymbol{\sigma}^2_{P1} - \boldsymbol{\sigma}^2_{P2}}{8\boldsymbol{\sigma}^2_S}, \quad (16)$$

and its variance as:

$$\boldsymbol{Var}\ (\boldsymbol{n}_e) = \boldsymbol{n}_e^2 \left[\frac{4\left(\boldsymbol{\sigma}^2_{P1} + \boldsymbol{\sigma}^2_{P2}\right)}{(\boldsymbol{\mu}_{P1} - \boldsymbol{\mu}_{P2}} + \frac{\boldsymbol{Var}\left(\boldsymbol{\sigma}^2_S\right)}{\left(\boldsymbol{\sigma}^2_S\right)^2} \right]. \quad (17)$$

Zeng correction

Zeng (1992) introduced a correction to the Castle-Wright estimator accounting for linked loci and loci of unequal effect. The segregational variance is more generally expressed, leading to:

$$\boldsymbol{n}_{zeng} = \frac{2\bar{\boldsymbol{c}}\boldsymbol{n}_e + \boldsymbol{C}_\alpha (\boldsymbol{n}_e - 1)}{1 - \boldsymbol{n}_e (1 - 2\bar{\boldsymbol{c}})}, \quad (18)$$

where $\bar{c}$ can be obtained from chromosome number *M* by:

$$\bar{\boldsymbol{c}} = \frac{\boldsymbol{M} - 1}{2\boldsymbol{M}}, \quad (19)$$

and C_α being the squared coefficient of variation of effects, might be set to different values. $C_\alpha = 0$ means constant allelic effects, $C_\alpha = 0.25$ a normal distribution of allelic effects, $C_\alpha = 1$ negative exponential distribution, $C_\alpha = 4$ assumes a leptokurtic (L-shaped) distribution of allelic effects. The variance of n_{zeng} was calculated as follows:

$$\boldsymbol{Var}(\boldsymbol{n}_{zeng}) = \frac{4\bar{\boldsymbol{c}}^2 + (1 + \boldsymbol{C}_\alpha)^2 \boldsymbol{Var}\,(\boldsymbol{n}_e)}{[1 - \boldsymbol{n}_e\,(1 - 2\bar{\boldsymbol{c}})]^2}. \quad (20)$$

Results

PRINCIPAL COMPONENT AND CVA OF OVERALL AND VENTRAL BODY SHAPE

Principal component 1 (PC1) of the body shape explained 47.0% of the variation in shape and differentiated *O. eversi* from *O. nigrimas*. The *t*-test showed that the values along this axis differentiated the parental lines (*t*-test, *p*-value < 2.2e-16) (Supporting information Fig. S3). Principal component 2 explained 14.0% of the variation in shape but the values did not differ between the parental lines (*t*-test, *p*-value = 1), whereas values along the principal component 3 were significantly different between parental lines (*t*-test, *p*-value = 7.18e-04) and explained 10.8% of the variation (Supporting information Figs. S4 and S5). PC2 seemed to be dominated by a horizontal bend. This is frequent in geometric morphometric studies of fish, and might be due to fixation artifacts (Carpenter 1996; Cavalcanti et al. 1999) or different postures during landmark capture (Valentin et al. 2008). The artifact did not lead to a false distinction between the groups; since PC2 did not differ between the parental species and was not included in further analyses. All the other principal components explained less than 5.4% of the variation and did not differ significantly between the parental species. Therefore, PC1 and PC3 were included in the Castle-Wright estimator calculations. Main differences in body shape were a larger head relative to the body length in *O. eversi* compared to *O. nigrimas* and also shorter ribs and, thus, a more compressed middle part of the body in *O. eversi*. *Oryzias nigrimas* had a smaller head and a more torpedo-shaped body (also see Fig. 1A and B). Additionally, a CVA focusing on the difference between the parental lines and the crosses showed that 86.7% of the variation was explained by different head sizes and ventral compression (Supporting information Fig. S6).

PC1 of the ventral body shape (PC1_vs) explained 53.7% of the variation in shape and values were significantly different between *O. eversi* and *O. nigrimas* (*t*-test, *p* value < 2.2e-16) (Supporting information Fig. S7). PC2 and PC3 explained 15.3% and 10.3% of the variation, but the values of both did not separate the parental lines from each other (*t*-test, PC2_vs: *p* value = 1, PC3_vs: *p* value = 0.1081). Both PC2 and PC3 showed primarily a left-right asymmetry (Supporting information Fig. S8). PC2 and PC3 were excluded from further analysis since the values did not differ between the parental species. All other principal components, which explained less than 5.3% of the variation,

were excluded as well. The main differences in the ventral body shape were the different positions of the pelvic fins in *O. eversi* and *O. nigrimas*. The pelvic fins of *O. eversi* had a wider base compared to the *O. nigrimas* pelvic fins and were positioned closer to the anterior rim of the papilla. The papilla itself occupied a larger space between the pelvic fin and the anal fin insertion in *O. eversi* compared to *O. nigrimas* (Fig. 1C, Supporting information Fig. S7). The CVA showed the same results, and the first canonical axis explained 81.5% of the variation (Supporting information Fig. S9).

TRAIT CORRELATIONS AND MODULARITY

We detected significant correlations between more traits in the parental lines of *O. nigrimas* and *O. eversi* compared to all hybrid crosses (Fig. 2). The length of rib 12 was correlated with papilla area, pigment area, and pelvic fin length in both *O. eversi* and *O. nigrimas*. In the other crosses, we observed a correlation of rib 12 lengths with pelvic fin lengths only in *O. nigrimas* backcrosses. In *O. eversi*, pelvic fin length was positively correlated with all genital papilla structures, which means the longer the fin, the larger the papilla and the pigmented area. These correlations were not present in *O. nigrimas*, suggesting that pelvic fin length and genital papilla size varied independently from each other in *O. nigrimas*.

Total genital papilla area and papilla area were negatively correlated with general body shape (PC1) in *O. nigrimas*. Hence, the more *O. nigrimas*-like the body shape was, the smaller was the structure. This correlation was not found in *O. eversi*, in which the genital papilla attributes were correlated with other structures but not with the body shape. Similar to all other traits, genital papilla anatomy seemed to be disconnected from the general body shape in *O. eversi*.

In the F1 hybrids, PC1 was correlated with pelvic fin length. This means, the more *O. nigrimas*-like the body shape was, the shorter was the pelvic fin. This result was not surprising, since the pelvic fins of females *O. nigrimas* are significantly shorter in relation to standard length as the pelvic fins of females *O. eversi* (Spanke et al. 2021). The pelvic fin length and the body shape followed a *O. nigrimas–O. eversi* gradient in the F1 hybrid females. The total genital papilla area was positively correlated with pelvic fin length in *O. eversi*, intercrosses and *O. nigrimas* backcrosses (Fig. 2). This correlation did not exist in the *O. eversi* backcrosses.

Distance between pelvic fins and rib 12 length was correlated in intercrosses and *O. nigrimas* backcrosses. This correlation was also present in *O. nigrimas*. *Oryzias eversi* backcrosses had a positive correlation in pigment area and pelvic fin length, which means the longer the fins, the larger the pigmented area, a pattern also found in *O. eversi*. The total genital papilla area was always correlated with the papilla area. In *O. eversi* and the F2 generation, the percentage of the total area covered by the "nipple-like" skin fold is on average lower than 50% (OE: 38.6% [SD: 10.0%], IC: 43% [SD: 10.2%], BCE 39% [SD: 9.6%], BCN: 46.3% [SD: 13.8%]). In *O. nigrimas*, it is 59.3% (SD: 13.9%) and in F1 hybrids, it is 63.3% (SD: 9.5%).

In *Oryzias nigrimas*, the length of rib 12 was correlated with papilla area, pigmented area, pelvic fin length, and distance between pelvic fins. Overall body shape was correlated with total genital papilla area, papilla area, and PC1_vs. The correlations between PC1_vs, the distance between pelvic fins and the genital papilla measurements were not included since they were not independent (Fig. 1E). Thus, there was one potential unit of functionally interacting traits in *O. nigrimas*, consisting of rib 12 length, distance between pelvic fins, total genital papilla area, papilla area and PC1, where each trait has at least two out of four possible correlations (Fig. 3). Pelvic fin length, pigmented area, and PC1_vs had only one significant correlation to another trait and were, therefore, not considered to belong to the unit. In *O. eversi*, length of rib 12, pelvic fin length, distance between pelvic fins, pigmented area, and papilla area formed a potential functional unit (Fig. 3). By analyzing modularity of body plans, we found support for the following hypotheses: In *O. eversi*, H2 (module 1: head and pectoral fins, module 2: rest of the body) and H3 (module 1: head, module 2: pectoral fins, module 3: rest of the body) were both supported with a $CR < 1$ and a p value ≤ 0.005, while H3 + a (module 1: head, module 2: pectoral fins, module 3: mid-body region, module 4: rest of the body) had the highest effect size (most negative Z_{CR}) (Supporting information Table S2, Fig. 4). In *O. nigrimas*, only H3 was supported ($CR < 1$ and p value $= 0.034$), in which head and the pectoral fins, each are separated modules.

ALLOMETRY

Most fish species never stop growing, so body size might act as an approximation for age. However, growth slows down with increasing age, and nutrition and environmental conditions influence the growth rate and body size as well (Kozlowski 1996; Lall and Tibbetts 2009). We tested whether the ratios of the measured traits to standard length remained constant the larger the fish got. In the intercrosses, an allometric effect in the pelvic fin length was observed (Table 1). Larger fishes had larger pelvic fins in relation to their body size. This could lead to an overestimation of the variation in pelvic fin length in the intercrosses. As segregation variances of four crosses are included in the Castle-Wright estimator, this effect might be minor, but could lead to an underestimation of the number of loci. For the other traits, no significant allometric effect was observed after Bonferroni-correction.

Table 3. Chi-square test for the additive (X^2A) and additive-dominance (X^2AD) model (Eq. [8]).

X^2 tests	PC1	PC3	PC1_vs	GPA	PA	PI	PF	PF-PF	R12
X^2 A	0.4052	3.7771	0.7903	0.8173	2.8560	0.1950	0.5180	0.0001	0.0025
X^2 AD	0.2339	3.7704	0.8532	1.2716	4.6536	0.5327	0.5311	0.0010	0.0016
Λ	0.1712	0.0067	−0.0628	−0.4543	−1.7975	−0.3378	−0.0131	−0.0009	0.0009

Lambda is the difference between the two models (Eq. [9]). Both X^2A and X^2AD have four degrees of freedom (k-2), so 9.49 would be the significance threshold for p value = 0.05, all traits are lower. Lamda (Λ) having one degree of freedom is also not significant in any trait (p-value 0.05 = 3.84). PF = pelvic fin length, PF-PF = distance between pelvic fins, R12 = rib 12 length, GPA = total genital papilla area, PA = papilla area, PI = pigmented area of the genital papilla, PC1_vs = first principal component of the ventral landmark analysis, PC1 = first principal component of body shape landmark analysis.

CASTLE-WRIGHT ESTIMATOR

All traits differed significantly between the two parental species (p values of t-tests: PC1 < 2e-16, PC3: 7.18e-04, PC1_vs < 2e-16, GPA < 2e-16, PA: 0.0418, PI < 2e-16, PF < 2e-16, PF-PF 4,32e-07, R12 < 2e-16). The mean values and variances of all traits for the parental species and all crosses are listed in Table 2. The additive gene action model was suitable for all traits (chi-square test, with DF = k−2, p value 0.05 = 9.49) (Table 3). Lamda (Λ) is nonsignificant in all lines (DF = 1, chi-square test p value 0.05 = 3.84), which means the additive-dominance model did not surpass the additive-model.

In the 20th iteration after chi-square values and **a** converged (formula 13–15), we observed high chi-square values in PC1, PC3, and pelvic fin length (Table 3). PC1 had a negative estimator, probably coming from a negative segregation variance (Var(S)) which could be the result of violations of the underlying assumptions (see Discussion). PC1_vs had the highest number of loci and variation. All traits except for PC1 had positive estimators. PC1_vs, total genital papilla area, papilla area, pigmented area, distance between pelvic fins and rib 12 length had low chi-square values and showed additive gene action (Tables 2 and 4). The estimator corrected by Zeng's equation predicted 1 to 2 loci for PC3, papilla area, pelvic fin length, and distance between pelvic fins, 1 to 3 for pigmented area and rib 12 length, 1 to 4 for total genital papilla area, and 1 to 28 for PC1_vs (Table 5).

Discussion

The Castle-Wright estimator suggests that the morphological traits measured in this study related to pelvic brooding are mainly controlled by few loci. One to four loci underlie all traits, except for the ventral body shape, which is determined by one to 28 loci (Table 5). As nearly no natural dataset is able to meet the assumptions for an accurate calculation of the Castle-Wright estimator (e.g., Magalhaes and Seehausen 2010; O'Quin et al. 2012; Haag et al. 2019), estimations have to be viewed as a minimum number of loci (Jones 2001). Nevertheless, the estimator has proven its validity in several cases (Zeng 1992; Wu et al. 1997; Gurganus et al. 1999; Otto and Jones 2000; Peichel et al. 2001; Magalhaes and Seehausen 2010; Nandamuri et al. 2017; Feller et al. 2020). All measured morphological traits correlate to some extent in the parental lines, *O. eversi* and *O. nigrimas*, but most correlations are reduced or lost in their first- and second-generation hybrids, which render a simple pleiotropic origin, that is, based on only one or few genes, of trait complexes unlikely (except for total genital papilla area and papilla area [Fig. 2]). Modularity of the body plan is more complex in *O. eversi* compared to *O. nigrimas* (Fig. 4). The mid-body region harboring the egg cluster forms its own module in the pelvic brooding species. In the transfer brooder *O. nigrimas*, the mid-body region and the tail are one module (Fig. 4). Assuming transfer brooding as ancestral state (based on the pyhlogeny of Mokodongan and Yamahira 2015), modularity may have contributed to the evolution of pelvic brooding by mitigating potential evolutionary constrains in the mid-body region.

GENETIC ARCHITECTURE OF PELVIC BROODING

Finding the number of genetic loci responsible for a trait is still difficult (Schielzeth and Husby 2014; San-Jose and Roulin 2017; Lind et al. 2018). It is nearly impossible to distinguish one QTL composed of several loci in tight linkage from a stand-alone locus or even gene controlling the quantitative trait (Slate 2005). As an approach to narrow down the expectation of how complex a trait truly is, an estimator might be applied (Jones 2001). Using the Castle-Wright estimator which was proven to be fairly accurate (Nandamuri et al. 2017), the number of underlying loci can be estimated by the segregational variance of a quantitative trait.

Traits with a proposed low number of loci (between 1–4) and a low chi-square value, indicating that all prerequisites for the Castle-Wright estimator are met, are the length of rib 12, all three genital papilla traits, and the distance between pelvic fins. However, papilla area has an almost significant allometric effect

Table 4. Segregation variance of parents and crosses after 20 iterations (Eqs. [13–15]) with standard error for every trait.

	Var(P1) 20th iteration	SE	Var(P2) 20th iteration	SE	Var(S) 20th iteration	SE	*X2*
PC1	9.16E-05	2.41E-10	1.53E-04	8.37E-10	−5.54E-05	4.98E-10	29.0245
PC3	3.75E-04	1.54E-09	9.70E-05	4.35E-10	−1.67E-04	7.94E-10	22.8975
PC1_vs	1.40E-03	6.68E-08	1.00E-03	5.36E-08	5.20E-04	1.16E-07	11.5198
GPA	3.30E-05	4.49E-11	2.87E-05	4.41E-11	1.29E-05	8.30E-11	6.5519
PA	2.89E-05	2.93E-11	4.01E-05	7.12E-11	−8.04E-06	5.50E-11	2.6679
PI	2.50E-05	3.88E-11	4.25E-05	9.30E-11	3.03E-05	1.20E-10	2.7978
PF	6.38E-05	2.53E-10	4.98E-05	1.56E-10	6.39E-04	4.48E-09	22.3379
PF-PF	4.20E-05	5.69E-11	1.39E-05	1.30E-11	1.72E-05	8.37E-11	1.8367
R12	2.53E-05	5.01E-11	7.46E-05	2.31E-10	2.37E-04	7.63E-10	3.033

High chi-square values here indicate a violation of the additive model. Maybe they are also violated due to partial non-normality of the data. PF = pelvic fin length, PF-PF = distance between pelvic fins, R12 = rib 12 length, GPA = total genital papilla area, PA = papilla area, PI = pigmented area of the genital papilla, PC1_vs = first principal component of the ventral landmark analysis, PC1 = first principal component of body shape landmark analysis.

Segregation variance of parents and crosses after 20 iterations with standard error for every trait. High Chi-Square values here indicate a violation of the additive model. Maybe they are also violated due to partial non-normality of the data.

Table 5. Castle-Wright estimators for all traits.

estimator	PC1	PC3	PC1_vs	GPA	PA	PI	PF	PF-PF	R12
$\hat{n}_e$	−19.3416	0.2430	11.3687	2.2516	0.6068	1.7953	0.7074	0.4997	1.3880
$Var(n_e)$	2.00E-04	1.08E-05	1.93E-03	1.55E-05	−3.06E-06	7.90E-06	5.27E-06	2.78E-06	3.80E-06
$n_{zeng}(C_\alpha=0)$	−18.4625	0.2320	10.8519	2.1492	0.5792	1.7137	0.6752	0.4769	1.3249
$Var(n_{zeng})(C_\alpha=0))$	0.0731	0.9526	16.7436	1.4033	1.0190	1.2808	1.0384	0.9989	1.1825
final number of estimated loci	-	1-2	1-28	1-4	1-2	1-3	1-3	1-2	1-3

A Zeng-correction (C_α) of 0 means constant allelic effects are assumed. A C_α of 0.25 assumes normally distributed allelic effects, C = 1 negative exponential distributed. In all cases, constant allelic effects (C_α = 0) showed the most probable results, all other values of C_α resulted in a negative estimator. Var() is the corresponding variance. The estimator is then calculated by adding the variance to the estimator and since there are no partial loci, the number is rounded up. PF = pelvic fin length, PF-PF = distance between pelvic fins, R12 = rib 12 length, GPA = total genital papilla area, PA = papilla area, PI = pigmented area of the genital papilla, PC1_vs = first principal component of the ventral landmark analysis, PC1 = first principal component of body shape landmark analysis.

in *O. nigrimas*, indicating the papilla area gets larger related to body size in older (and larger) individuals. This could lead to an underestimation of the mean and overestimation of the variance. Both would bring the value of the nominator down (formula 16). Therefore, the current results could be an overestimation of the number of loci, which is, however, estimated to only two. Pelvic fin length and PC3 are also underlain by only a low number of putative loci (Table 5), but chi-square values are high, indicating a violation of the assumptions of the estimator (Lynch and Walsh 1998). For PC1, the negative estimator and negative segregation variance are indications that the expectations are not met for a successful application of the Castle-Wright estimator for this trait. For pelvic fin length, we observe an allometric effect in the intercrosses and an almost significant effect in the *O. eversi* backcrosses. Since all four crosses are taken together for the segregation variance, one allometric effect per trait might not have a large influence on the overall estimation, but the variance could be overestimated. This could lead to an underestimation of the number of loci (see Eq. [16]). PC 1 of the ventral body shape, on the other hand, has a low chi-square value and an estimator of 11 with a variance of 17 loci, which is rather large. The low number of loci assumed for the pelvic fin length is in line with prior studies in sticklebacks where only one gene has been shown to regulate differential gene expression in the pelvic reduction (Shapiro et al. 2004; Coyle et al. 2007).

All traits we analyzed are sexually dimorphic in pelvic brooding species. Pelvic fins are shorter and ribs longer in the male *O. eversi* and resemble traits in transfer brooding species (Spanke et al. 2021). In the genetic model organism medaka (*Oryzias latipes*), the sexually dimorphic trait of anal fin length was mapped over 11 ontogenetic stages in males and females. It was found that the QTLs for male anal fin length are located near steroid-binding genes, like the androgen receptor beta, estrogen-related receptor alpha, and estrogen receptor beta (Kawajiri et al. 2014). Also in the Sulawesi transfer brooder *Oryzias woworae*, an androgen-dependent regulation of *csf1* determines the red coloration in male fins (Ansai et al. 2021). No QTL was mapped on the sex chromosome 24, even though it would be expected for sexually dimorphic traits (Rice 1984; Mank 2009). Another example for androgen-dependent sexual dimorphism can be found in anal fins of Poecilidae, more specifically in *Gambusi affinis*. Poecilids are internal fertilizers and a second growth phase of the male anal fin is induced by male gonadal hormones (Rosa-Molinar et al. 1994). Hence, traits measured in this study could be underlain by pleiotropic interactions, mainly hormonally controlled and/or X-linked. Considering the low recombination rate of sex chromosomes (Bergero and Charlesworth 2009), and since based on our data only papilla traits could have a single locus pleiotropic basis, an X-linked basis of trait complexes seems rather unlikely.

MODULARITY OF TRAITS ASSOCIATED WITH PELVIC BROODING

Correlated single phenotypic traits forming a phenotypic complex trait can be understood as a unit with functionally correlating traits (Wagner and Schwenk 2000) or described as a module (Klingenberg 2008; Felice et al. 2018). Modularity is a broadly applied concept that is used in many fields, like in macroevolution to identify trait evolution through time and across species phylogenies (Simpson 1944; Gould and Eldredge 1977). Per definition, a module has a large correlation within itself but not with other modules (Klingenberg 2008). Module-based trait evolution has been studied in several organisms, like wasps (Perrard 2020), mammals (Cheverud 1995; Marroig and Cheverud 2001; Goswami 2006; Zelditch et al. 2008; Porto et al. 2009; Verity Bennett and Goswami 2010; Kelly and Sears 2011), snakes (Rhoda et al. 2021), and fishes (Larouche et al. 2015; Ornelas-García et al. 2017; Larouche et al. 2018). In the present study, we found strong correlations between traits that are involved in pelvic brooding in *O. eversi* and hints toward the presence of functional units in subsets of the analyzed traits in both parental species (Figs. 2 and 3). Functional morphological units of *O. nigrimas* and *O. eversi* are similar, but are based on fewer correlated traits in *O. nigrimas* (Figs. 2 and 3). In *O. eversi*, an additional correlation with pelvic fin length and pigmentation of the papilla is present, and a correlation with PC1 is absent.

Analysis of modularity in the body plan of *O. eversi* and *O. nigrimas* reveals a modular structure in females of both species (Fig. 4). The best-supported hypothesis in *O. nigrimas* described the head and the pectoral fins as each forming a module, and they were separated from the trunk (Fig. 4: H3). In the pelvic brooder *O. eversi*, the trunk is further separated after the pelvic fin insertion from the posterior (Fig. 4: H3 vs. H3 + a). Hence, traits related to pelvic brooding situated in the mid-body region are correlated and the mid-body module is more decoupled from the rest of the body in the pelvic brooder, which indicates the possibility that this section gained increasing evolutionary independence from other body parts in other pelvic brooding lineages (Hansen 2003, 2006).

Several examples support a correlation between modularity and trait evolution. For example, modularity of the backbone has facilitated rapid evolution of locomotion and static body support in mammals (Jones et al. 2018). Also, the skull and forelimbs of cat-like carnivores are highly integrated compared to generalists, probably due to the high selection pressure on their hunting skills (Michaud et al. 2020). A study on ray-finned fishes showed that fins inserted along the trunk region exhibit a high level of integration, and this module also evolves more rapidly compared to the head and tail region (Larouche et al. 2018).

In our focal species, most of the correlations were lost in the first and weakened in the second-generation interspecific hybrids (Fig. 2); such loss of correlations in F1 hybrids was repeatedly shown in other organisms (e.g., Rieseberg and Ellstrand 1993; Matsubayashi et al. 2010; Selz et al. 2014; Atsumi et al. 2021). Despite this, a genetic basis of the observed trait correlations in the parental species appears likely, as samples from different generations and laboratories were used in our analyses. Furthermore, our results indicate that most traits are based on multiple loci. A potential explanation (for the observed break down of correlations) is that heterozygous loci in the F1 hybrids may not equally affect all traits within a module and, thus, lead to a loss of trait correlations. Under this scenario, lower levels of heterozygosity in F2 backcrosses would lead to a partial recovery of trait correlations, which is what we observe (Fig. 2). We, thus, assume that modularity in the analyzed traits is likely based on sets of pleiotropic genes and might at least be partly controlled by factors like timing or intensity of hormonal regulation (as we analyze female-only traits) and that both are affected by interspecific hybridization.

Exceptions are papilla traits, where correlations remained in all crosses, which hint toward a simple pleiotropic origin. In general, we assume that functional modularity is one mechanism for the formation of functional units in females of *O. eversi* but shared ancestry is probably the reason why similar correlations and a modular structure are present in both species. Further, most traits are not exclusively involved in reproductive behavior. In males and other ricefish species, the same traits are probably shaped more by the species' ecology. In comparison to other traits investigated, pelvic fins experience greater flexibility for modification. They have evolved into adhesive organs in snailfish, clingfish, gobies, and gastromyzontine loaches, or into fins allowing the fish to move on land like the mudskipper; pelvic fins were completely reduced in about 100 fish families (reviewed in Yamanoue et al. 2010).

The genetic basis of such functional units might be modular. Some theoretical approaches showed convergence of functional and genetic modularity, which was also supported by empirical data (Wagner 1996; Mezey et al. 2000; Klingenberg et al. 2004). However, in other studies, a lack of genetic integration even favors a larger variation to easily form the same functional traits (Wainwright et al. 2005; Young et al. 2007). Nevertheless, as the functional units were present in both parental lines, we do not only assume a functional linkage but also genetic modularity of correlated traits related to pelvic brooding in *O. eversi*.

DISTANCE BETWEEN PELVIC FIN IS NOT CORRELATED WITH PELVIC FIN LENGTH

In the two species studied, the width and positioning of pelvic fins and the distance from pelvic fin base to the genital papilla differed as well. In *O. eversi*, the fin base and the distance between the fins is larger compared to *O. nigrimas*. Additionally,

the angle of the fin insertion is flatter in *O. eversi*, resulting in a horizontally larger cover. Moreover, the fins are located closer to the anterior rim of the genital papilla in *O. eversi* (Supporting information Figs. S7-9). Both traits are part of the functional unit of *O. nigrimas* and *O. eversi* and we hypothesize that they matter in their respective reproductive strategies. The wider positioning of fins in pelvic brooding *O. eversi* likely provides more space for the (typically larger) egg cluster, whereas the closer proximity to the genital papilla and urogenital pore ensures an optimal position for the egg cluster to be covered by the elongated pelvic fins. Pelvic fins are also thickened in pelvic brooding ricefishes and we assume that they are crucial in protecting the eggs from mechanic damage, and hypothesize that they even improve swimming performance of the female during brooding (Spanke et al. 2021). Interestingly, the distance between pelvic fins and the pelvic fin length is not correlated in the two tested species and their crosses (Supporting information Fig. S10). While the distance between the fins gets larger, the length remains constant. Possibly, they reach an optimum at some length in *O. eversi* as well as *O. nigrimas*, but the distance can be further widened without a negative impact in both species.

POTENTIAL PLASTICITY IN PAPILLA TRAITS

The genital papilla is clearly different between the parental species (Fig. 1C): single-lobed and heavily pigmented in *O. eversi* and bilobed and much less or unpigmented in *O. nigrimas*. The function of the genital papilla with respect to brooding in ricefishes is, however, not well-understood and personal observations of *O. eversi* gave the impression that pigmentation and size of the papilla change, potentially over lifetime and depending on hormonal status of the female. In line with this, analyses of gene expression during pelvic brooding in *O. eversi* showed that gene expression in the papilla tissue differs between females that are currently carrying eggs compared to those that do not (Hilgers et al. 2021). However, the variance in the *O. eversi's* total genital papilla area, the papilla area, and the pigmented area still was not larger than in *O. nigrimas* compared to the mean (Table 2). We, thus, assume that our measurements did not over- or underestimate the variability of the papilla in *O. eversi*. Additionally, correlations between different papilla measurements were strong ($\geq$0.88) in *O. eversi*, which we would not expect if plasticity had a large impact.

Conclusions

Pelvic brooding is a complex phenotypic trait that entails many sex-specific morphological adaptations. Its evolution might have been facilitated by traits controlled by only a few loci which seem to be additive, leading to correlated and connected traits. In both, the transfer brooder and the pelvic brooder, evolutionary modularity may have allowed different body sections to evolve independently from others, and presumably facilitated the rapid evolution of pelvic brooding. Future studies using QTL mapping will test these predictions and provide further insight into the genetic architecture of pelvic brooding. In conclusion, ricefishes not only provide a promising model system to study how adaptation of ancestral modularity enables complex sex-specific adaptations, but also bear great potential to address other aspects on the evolution of complex reproductive strategies in the future.

AUTHOR CONTRIBUTIONS

JMF and JS planned and conducted the study. AWN developed and established the crossing experiments. JMF, TS, LH, JS continued breeding of crosses. FH and AWN provided samples of *O. nigrimas* and *O. eversi*. CM supervised the statistical work. JMF collected and analyzed the data. All authors contributed critically to writing of the manuscript.

ACKNOWLEDGMENTS

We would like to thank the Ministry of Research, Technology, and Higher Education of the Republic of Indonesia for providing a foreign research permit to FH and the team of authors (number 131/E5/E5.4/SIP/2019). Furthermore, we thank the animal caretakers Malte Dittmann, Timo Hartmann, Nils Behr, Serkan Wesel, and also Dr. Kristin Tietje for their help to establish and maintain the breeding lines. We are further grateful for the travel grants received from the Alexander-Koenig-Gesellschaft (AKG), Bonn, by JMF and TS. This research was funded by the ERC starting grant ("Evolmapping") and the Leibniz Association (SAW-Ricefish P91/2016).

LITERATURE CITED

Adams, D. C. 2016. Evaluating modularity in morphometric data: challenges with the *RV* coefficient and a new test measure. Methods Ecol Evol 7:565–572. https://doi.org/10.1111/2041-210X.12511

Adams, D. C., and M. L. Collyer. 2019. Comparing the strength of modular signal, and evaluating alternative modular hypotheses, using covariance ratio effect sizes with morphometric data. Evolution (N Y) 73:2352–2367. https://doi.org/10.1111/evo.13867

Adams, D. C., and E. Otárola-Castillo. 2013. geomorph: an r package for the collection and analysis of geometric morphometric shape data. Methods Ecol Evol 4:393–399. https://doi.org/10.1111/2041-210X.12035

Ansai, S., K. Mochida, S. Fujimoto, et al. 2021. Genome editing reveals fitness effects of a gene for sexual dichromatism in Sulawesian fishes. Nat Commun 12:1350. https://doi.org/10.1038/s41467-021-21697-0

Atsumi, K., M. Lagisz, and S. Nakagawa. 2021. Nonadditive genetic effects induce novel phenotypic distributions in male mating traits of F1 hybrids. Evolution (N Y) 75:1304–1315. https://doi.org/10.1111/evo.14224

Bergero, R., and D. Charlesworth. 2009. The evolution of restricted recombination in sex chromosomes. Trends Ecol Evol 24:94–102. https://doi.org/10.1016/J.TREE.2008.09.010

Blackburn, D. G. 1999. Viviparity and oviparity: evolution and reproductive strategies. Encycl Reprod 4:994–1003.

Carpenter, K. E. 1996 Morphometric pattern and feeding mode in emperor fishes (*Lethrinidae, Perciformes*). In: Advances in morphometrics. Boston, MA: Springer US, pp. 479–487.

Castle, W. 1921. An improved method of estimating the number of genetic factors concerned in cases of blending inheritance. Science 54:223.

Cavalcanti, M. J., L. R. Monteiro, P. Roberto, and D. Lopes. 1999. Landmark-based morphometric analysis in selected species of serranid fishes (*Perciformes: Teleostei*). Zoological Studies 38: 287–294.

Cavalli, L. L. 1952. An analysis of linkage in quantitative inheritance. In:Quantitative inheritance. London: H.M.S.O, pp. 135–144.

Chan, Y. F., M. E. Marks, F. C. Jones, et al. 2010. Adaptive evolution of pelvic reduction of a *Pitx1* enhancer. Science (80-) 327:302–305.

Cheverud, J. 1995. Morphological integration in the saddle-back tamarin (*Saguinus fuscicollis*) cranium. Am Nat 145:63–89. https://doi.org/10.1086/285728

Coyle, S. M., F. A. Huntingford, and C. L. Peichel. 2007. Parallel evolution of pitx1 underlies pelvic reduction in scottish threespine stickleback (*Gasterosteus aculeatus*). J Hered 98:581–586. https://doi.org/10.1093/jhered/esm066

Farmer, C. G. 2000. Parental care: the key to understanding endothermy and other convergent features in birds and mammals. Am Nat 155:326–334. https://doi.org/10.1086/303323

Felice, R. N., M. Randau, and A. Goswami. 2018. A fly in a tube: macroevolutionary expectations for integrated phenotypes. Evolution (N Y) 72:2580–2594. https://doi.org/10.1111/evo.13608

Feller, A. F., M. P. Haesler, C. L. Peichel, and O. Seehausen. 2020. Genetic architecture of a key reproductive isolation trait differs between sympatric and non-sympatric sister species of Lake Victoria cichlids. Proc R Soc B Biol Sci 287:20200270. https://doi.org/10.1098/rspb.2020.0270

Goswami, A. 2006. Morphological integration in the carnivoran skull. Evolution (N Y) 60:169–183.

Gould, S. J., and N. Eldredge. 1977. Punctuated equilibria: the tempo and mode of evolution reconsidered on JSTOR. Paleobiology 33:115–151.

Grier, H. J., and M. C. Uribe. 2005. Viviparous fishes. Neotropical Ichthyology 4. https://doi.org/10.1590/S1679-62252006000400012

Gurganus, M. C., S. V. Nuzhdin, J. W. Leips, and T. F. C. Mackay. 1999. High-resolution mapping of quantitative trait loci for sternopleural bristle number in *Drosophila melanogaster*. Genetics 152:1585–1604.

Haag, M., K. Wells, and W. Lamberson. 2019. Genetic basis of voluntary water consumption in two divergently selected strains of inbred mice. Vet Med Sci 5:569–573. https://doi.org/10.1002/vms3.192

Haines, R. W. 1942. The evolution of epiphyses and of endochondral bone. Biol Rev 17:267–292. https://doi.org/10.1111/j.1469-185X.1942.tb00440.x

Hansen, T. F. 2003. Is modularity necessary for evolvability?: Remarks on the relationship between pleiotropy and evolvability. Biosystems 69:83–94.

———. 2006. The evolution of genetic architecture. Annu Rev Ecol Evol Syst 37:123–157. https://doi.org/10.1146/annurev.ecolsys.37.091305.110224

Harell, F. E. J. 2021. Harrell Miscellaneous.

Hayman, B. I. 1960a. The separation of epistatic from additive and dominance variation in generation means. Genetica 31:133–146. https://doi.org/10.1007/BF01984430

———. 1960b. Maximum likelihood estimation of genetic components of variation. Biometrics 16:369. https://doi.org/10.2307/2527688

Herder, F., R. K. Hadiaty, and A. W. Nolte. 2012. Pelvic-fin brooding in a new species of riverine ricefish (*Atherinomorpha: Beloniformes: Adrianichthyidae*) from Tana Toraja, Central Sulawesi, Indonesia. Raffles Bull Zool 60:467–476.

Hilgers, L., O. Roth, A. W. Nolte, et al. 2021. Inflammation and convergent placenta gene co-option contributed to a novel reproductive tissue. Curr Biol https://doi.org/10.1016/J.CUB.2021.12.004

Iwamatsu, T., H. Kobayashi, M. Sato, and M. Yamashita. 2008. Reproductive role of attaching filaments on the egg envelope in *Xenopoecilus sarasinorum* (*Adrianichthidae, Teleostei*). J Morphol 269:745–750. https://doi.org/10.1002/jmor.10620

Iwamatsu, T., H. Kobayashi, Y. Shibata, et al. 2007. Oviposition cycle in the oviparous fish *Xenopoecilus sarasinorum*. Zoolog Sci 24:1122–1127. https://doi.org/10.2108/zsj.24.1122

Jones, C. D. 2001. Extension of the Castle-Wright effective factor estimator to sex linkage and haplodiploidy. J Hered 92:274–276. https://doi.org/10.1093/jhered/92.3.274

Jones, K. E., L. Benitez, K. D. Angielczyk, and S. E. Pierce. 2018. Adaptation and constraint in the evolution of the mammalian backbone. BMC Evol Biol 18:172. https://doi.org/10.1186/s12862-018-1282-2

Kawajiri, M., K. Yoshida, S. Fujimoto, et al. 2014. Ontogenetic stage-specific quantitative trait loci contribute to divergence in developmental trajectories of sexually dimorphic fins between medaka populations. Mol Ecol 23:5258–5275. https://doi.org/10.1111/mec.12933

Kelly, E. M., and K. E. Sears. 2011. Reduced phenotypic covariation in marsupial limbs and the implications for mammalian evolution. Biol J Linn Soc 102:22–36. https://doi.org/10.1111/j.1095-8312.2010.01561.x

Klingenberg, C. P. 2008. Morphological integration and developmental modularity. Annu Rev Ecol Evol Syst 39:115–132. https://doi.org/10.1146/annurev.ecolsys.37.091305.110054

Klingenberg, C. P., L. J. Leamy, and J. M. Cheverud. 2004. Integration and modularity of quantitative trait locus effects on geometric shape in the mouse mandible. Genetics 166:1909–1921. https://doi.org/10.1534/genetics.166.4.1909

Kottelat, M. 1990. Synopsis of the endangered *Buntingi* (*Osteichthyes*: *Adranichtyidae* and *Oryziidae*) of Lake Poso, Central Sulawesi, Indonesia, with a new reproductive guild and descriptions of three new species. Ichthyol Explor Freshwaters 1:46–67.

Kozlowski, J. 1996. Optimal allocation of resources explains interspecific life-history patterns in animals with indeterminate growth. Proc R Soc B Biol Sci 263:559–566. https://doi.org/10.1098/rspb.1996.0084

Lall, S. P., and S. M. Tibbetts. 2009. Nutrition, feeding, and behavior of fish. Vet Clin North Am Exot Anim Pract 12:361–372. https://doi.org/10.1016/J.CVEX.2009.01.005

Langerhans, R., and D. Reznick. 2010. Ecology and evolution of swimming performance in fishes: predicting evolution with biomechanics. In: Fish locomotion: an eco-ethological perspective. New Hampshire, USA: Science Publishers, pp. 248.

Larouche, O., R. Cloutier, and M. L. Zelditch. 2015. Head, body and fins: patterns of morphological integration and modularity in fishes. Evol Biol 42:296–311. https://doi.org/10.1007/s11692-015-9324-9

Larouche, O., M. L. Zelditch, and R. Cloutier. 2018. Modularity promotes morphological divergence in ray-finned fishes. Sci Rep 8:7278. https://doi.org/10.1038/s41598-018-25715-y

Lefèvre, C. M., J. A. Sharp, and K. R. Nicholas. 2010. Evolution of lactation: ancient origin and extreme adaptations of the lactation system. Annu Rev Genomics Hum Genet 11:219–238. https://doi.org/10.1146/annurev-genom-082509-141806

Lind, B. M., M. Menon, C. E. Bolte, et al. 2018. The genomics of local adaptation in trees: are we out of the woods yet? Tree Genet Genomes 14:29. https://doi.org/10.1007/s11295-017-1224-y

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. MA, USA: Sinauer Publisher.

Magalhaes, I. S., and O. Seehausen. 2010. Genetics of male nuptial colour divergence between sympatric sister species of a Lake Victoria cichlid fish. J Evol Biol 23:914–924. https://doi.org/10.1111/j.1420-9101.2010.01960.x

Magwene, P. M. 2001. New tools for studying integration and modularity. Evolution (N Y) 55:1734–1745. https://doi.org/10.1111/j.0014-3820.2001.tb00823.x

Mank, J. E. 2009. Sex chromosomes and the evolution of sexual dimorphism: lessons from the genome. Am Nat 173:141–150. https://doi.org/10.1086/595754

Mank, J. E., D. E. L. Promislow, and J. C. Avise. 2005. Phylogenetic perspectives in the evolution of parental care in ray-finned fishes. Evolution 59:1570–1578.

Marroig, G., and J. Cheverud. 2001. A comparison of phenotypic variation and covariation patterns and the role of phylogeny, ecology, and ontogeny during cranial evolution of New World monkeys. Evolution (N Y) 55:2576–2600.

Mather, K., and J. Jinks. 1971. Biometrical genetics. 2nd edn. New Fetter Lane, London: Chapman and Hall Ltd.

Matsubayashi, K. W., I. Ohshima, and P. Nosil. 2010. Ecological speciation in phytophagous insects. Entomol Exp Appl 134:1–27. https://doi.org/10.1111/j.1570-7458.2009.00916.x

Mezey, J. G., J. M. Cheverud, and G. P. Wagner. 2000. Is the genotype-phenotype map modular?: a statistical approach using mouse quantitative trait loci data. Genetics 156:305–311.

Michaud, M., G. Veron, and A. Fabre. 2020. Phenotypic integration in feliform carnivores: covariation patterns and disparity in hypercarnivores versus generalists. Evolution (N Y) 74:2681–2702. https://doi.org/10.1111/evo.14112

Moczek, A. P. 2008. On the origins of novelty in development and evolution. BioEssays 30:432–447. https://doi.org/10.1002/bies.20754

Mokodongan, D. F., and K. Yamahira. 2015. Origin and intra-island diversification of Sulawesi endemic *Adrianichthyidae*. Mol Phylogenet Evol 93:150–160. https://doi.org/10.1016/j.ympev.2015.07.024

Moriyama, Y., and K. Koshiba-Takeuchi. 2018. Significance of whole-genome duplications on the emergence of evolutionary novelties. Brief Funct Genomics 17:329–338. https://doi.org/10.1093/bfgp/ely007

Nandamuri, P., B. E. Dalton, and K. L. Carleton. 2017. Determination of the genetic architecture underlying short wavelength sensitivity in Lake Malawi cichlids. J Hered 108:379–390. https://doi.org/10.1093/jhered/esx020

O'Quin, C. T., A. C. Drilea, R. B. Roberts, and T. D. Kocher. 2012. A small number of genes underlie male pigmentation traits in Lake Malawi cichlid fishes. J Exp Zool Part B Mol Dev Evol 318:199–208. https://doi.org/10.1002/jez.b.22006

Ornelas-García, C. P., A. Bautista, F. Herder, and I. Doadrio. 2017. Functional modularity in lake-dwelling characin fishes of Mexico. PeerJ 2017:1–22. https://doi.org/10.7717/peerj.3851

Otto, S. P., and C. D. Jones. 2000. Detecting the undetected: estimating the total number of loci underlying a quantitative trait. Genetics 156:2093–2107.

Parenti, L. R. 2008. A phylogenetic analysis and taxonomic revision of rice fishes, *Oryzias* and relatives (*Beloniformes, Adrianichthydae*). Zool J Linn Soc 154: 494–610.

———. 1986. Homology of pelvic fin structures in female phallostethid fishes (*Atherinomorpha, Phallostethidae*). Copeia 1986: 305–310.

———. 2005. The phylogeny of atherinomorphs: evolution of a novel fish reproductive system. In: Viviparous fishes. Homestead, FL: New Life Publications, pp. 13–30.

Peichel, C. L., K. S. Nereng, K. A. Ohgi, et al. 2001. The genetic architecture of divergence between threespine stickleback species. Nature 414:901–905. https://doi.org/10.1038/414901a

Perrard, A. 2020. Wasp waist and flight: convergent evolution in wasps reveals a link between wings and body shapes. Am Nat 195:181–191. https://doi.org/10.1086/706914

Porto, A., F. B. de Oliveira, L. T. Shirai, et al. 2009. The evolution of modularity in the mammalian skull I: morphological integration patterns and magnitudes. Evol Biol 36:118–135. https://doi.org/10.1007/s11692-008-9038-3

Price, E. R., and E. M. Dzialowski. 2018. Development of endothermy in birds: patterns and mechanisms. J Comp Physiol B 188:373–391. https://doi.org/10.1007/s00360-017-1135-0

Price, S. A., S. T. Friedman, and P. C. Wainwright. 2015. How predation shaped fish: the impact of fin spines on body form evolution across teleosts. Proc R Soc B Biol Sci 282:20151428. https://doi.org/10.1098/rspb.2015.1428

Rhoda, D., P. D. Polly, C. Raxworthy, and M. Segall. 2021. Morphological integration and modularity in the hyperkinetic feeding system of aquatic-foraging snakes. Evolution (N Y) 75:56–72. https://doi.org/10.1111/evo.14130

Rice, W. R. 1984. Sex chromosomes and the evolution of sexual dimorphism. Evolution (N Y) 38:735. https://doi.org/10.2307/2408385

Rieseberg, L. H., and N. C. Ellstrand. 1993. What can molecular and morphological markers tell us about plant hybridization? CRC Crit Rev Plant Sci 12:213–241. https://doi.org/10.1080/07352689309701902

Rosa-Molinar, E., S. E. Hendricks, J. Rodriguez-Sierra, and B. Fritzsch. 1994. Development of the anal fin appendicular support in the western mosquitofish, *Gambusia affinis affinis* (Baird and Girard, 1854): a reinvestigation and reinterpretation. Cells Tissues Organs 151:20–35. https://doi.org/10.1159/000147639

Ruben, J. 1995. The evolution of endothermy in mammals and birds: from physiology to fossils. Annu Rev Physiol 57:69–95. https://doi.org/10.1146/annurev.ph.57.030195.000441

San-Jose, L. M., and A. Roulin. 2017. Genomics of coloration in natural animal populations. Philos Trans R Soc B Biol Sci 372:20160337. https://doi.org/10.1098/rstb.2016.0337

Schielzeth, H., and A. Husby. 2014. Challenges and prospects in genome-wide quantitative trait loci mapping of standing genetic variation in natural populations. Ann N Y Acad Sci 1320:35–57. https://doi.org/10.1111/NYAS.12397

Schlager, S. 2017. Morpho and Rvcg–shape analysis in R: R-packages for geometric morphometrics, shape analysis and surface manipulations. Stat Shape Deform Anal 2017:217–256. https://doi.org/10.1016/B978-0-12-810493-4.00011-0

Schlosser, G., and G. P. Wagner. 2004. Modularity in development and evolution. , Chicago, IL: Univ. Chicago Press.

Selz, O. M., K. Lucek, K. A. Young, and O. Seehausen. 2014. Relaxed trait covariance in interspecific cichlid hybrids predicts morphological diversity in adaptive radiations. J Evol Biol 27:11–24. https://doi.org/10.1111/jeb.12283

Shapiro, M. D., M. E. Marks, C. L. Peichel, et al. 2004. Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. Nature 428:717–723. https://doi.org/10.1038/nature02415

Simpson, G. 1944. Tempo and mode in evolution.New York, NY: Columbia Univ. Press.

Sîrbulescu, R. F., I. Ilieş, A. Meyer, and G. K. H. Zupanc. 2017. Additive neurogenesis supported by multiple stem cell populations mediates adult spinal cord development: a spatiotemporal statistical mapping analysis in a teleost model of indeterminate growth. Dev Neurobiol 77:1269–1307. https://doi.org/10.1002/dneu.22511

Slate, J. 2005. Quantitative trait locus mapping in natural populations: progress, caveats and future directions. Mol Ecol 14:363–379. https://doi.org/10.1111/j.1365-294X.2004.02378.x

Spanke, T., L. Hilgers, B. Wipfler, et al. 2021. Complex sexually dimorphic traits shape the parallel evolution of a novel reproductive strategy in Sulawesi ricefishes (*Adrianichthyidae*). BMC Ecol Evol 21:57. https://doi.org/10.1186/s12862-021-01791-z

Stamps, J. A., M. Mangel, and J. A. Phillips. 1998. A new look at relationships between size at maturity and asymptotic size. Am Nat 152:470–479. https://doi.org/10.1086/286183

Valentin, A. E., X. Penin, J. P. Chanut, et al. 2008. Arching effect on fish body shape in geometric morphometric studies. J Fish Biol 73:623–638. https://doi.org/10.1111/j.1095-8649.2008.01961.x

Verity Bennett, C., and A. Goswami. 2010. Does developmental strategy drive limb integration in marsupials and monotremes? Mamm Biol 76:79–83. https://doi.org/10.1016/j.mambio.2010.01.004

Wagner, G. P. 1996. Homologues, natural kinds and the evolution of modularity. Am Zool 36:36–43. https://doi.org/10.1093/icb/36.1.36

Wagner, G. P., and L. Altenberg. 1996. Perspective: complex adaptations and the evolution of evolvability. Evolution (N Y) 50:967–976. https://doi.org/10.1111/j.1558-5646.1996.tb02339.x

Wagner, G. P., M. Pavlicev, and J. M. Cheverud. 2007. The road to modularity. Nat Rev Genet 8:921–931. https://doi.org/10.1038/nrg2267

Wagner, G. P., and K. Schwenk. 2000. Evolutionarily stable configurations: functional integration and the evolution of phenotypic stability. In: Evolutionary *b*iology. Boston, MA: Springer US, pp. 155–217.

Wainwright, P. C., M. E. Alfaro, D. I. Bolnick, and C. D. Hulsey. 2005. Many-to-one mapping of form to function: a general principle in organismal design? Integr Comp Biol 45:256–262. https://doi.org/10.1093/icb/45.2.256

Walker, J. A. 1997. Ecological morphology of lacustrine threespine stickleback *Gasterosteus aculeatus* L. (Gasterosteidae) body shape. Biol J Linn Soc 61:3–50. https://doi.org/10.1111/j.1095-8312.1997.tb01777.x

Wei, T., and V. Simko. 2021. R package "corrplot": Visualization of a correlation matrix. Accessed October 2020. https://github.com/taiyun/corrplot

West-Eberhard, M. 2003. Developmental plasticity and evolution. Oxford, USA: Oxford Univ. Press.

Wootton, R., and C. Smith. 2014. Reproductive biology of teleost fishes. Hoboken, NJ: Wiley.

Wu, R., H. D. Bradshaw, and R. F. Stettler. 1997. Molecular genetics of growth and development in populus (Salicaceae). V. Mapping quantitative trait loci affecting leaf variation. Am J Bot 84:143–153. https://doi.org/10.2307/2446076

Yamamoto, T. 1975. Medaka: (Killifish). Biology and strains. Series of stock culture in biological field. Tokyo: Keigaku Publishing Company.

Yamamoto, T., and H. Suzuki. 1955. The manifestation of the urinogenital papillae of the medaka (*Oryzias Latipes*) by sex-hormones. Dev Growth Differ 2:133-144 file:///H:/_Arbeitsplatz/Desktop/YAMAMOTO_e. https://doi.org/10.1111/j.1440-169X.1955.tb00047.x

Yamanoue, Y., D. H. E. Setiamarga, and K. Matsuura. 2010. Pelvic fins in teleosts: structure, function and evolution. J Fish Biol 77:1173–1208. https://doi.org/10.1111/j.1095-8649.2010.02674.x

Yang, A. S. 2001. Modularity, evolvability, and adaptive radiations: a comparison of the hemi- and holometabolous insects. Evol Dev 3:59–72. https://doi.org/10.1046/j.1525-142x.2001.003002059.x

Young, R. L., T. S. Haselkorn, and A. V. Badyaev. 2007. Functional equivalence of morphologies enables morphological and ecological diversity. Evolution (N Y) 61:2480–2492. https://doi.org/10.1111/j.1558-5646.2007.00210.x

Zelditch, M. L., A. R. Wood, R. M. Bonett, and D. L. Swiderski. 2008. Modularity of the rodent mandible: integrating bones, muscles, and teeth. Evol Dev 10:756–768. https://doi.org/10.1111/j.1525-142X.2008.00290.x

Zeng, Z. B. 1992. Correcting the bias of Wright's estimates of the number of genes affecting a quantitative character: a further improved method. Genetics 131:987–1001.

Zúñiga-Vega, J. J., D. N. Reznick, and J. B. Johnson. 2007. Habitat predicts reproductive superfetation and body shape in the livebearing fish *Poeciliopsis turrubarensis*. Oikos 116:995–1005. https://doi.org/10.1111/j.0030-1299.2007.15763.x

Associate Editor: M. E. Maan
Handling Editor: M. L. Zelditch

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1: List of specimens used in this study, specifying the cross/species of the specimen, the abbreviation for the cross or species, year of death, a unique ID and the origin of the specimen.

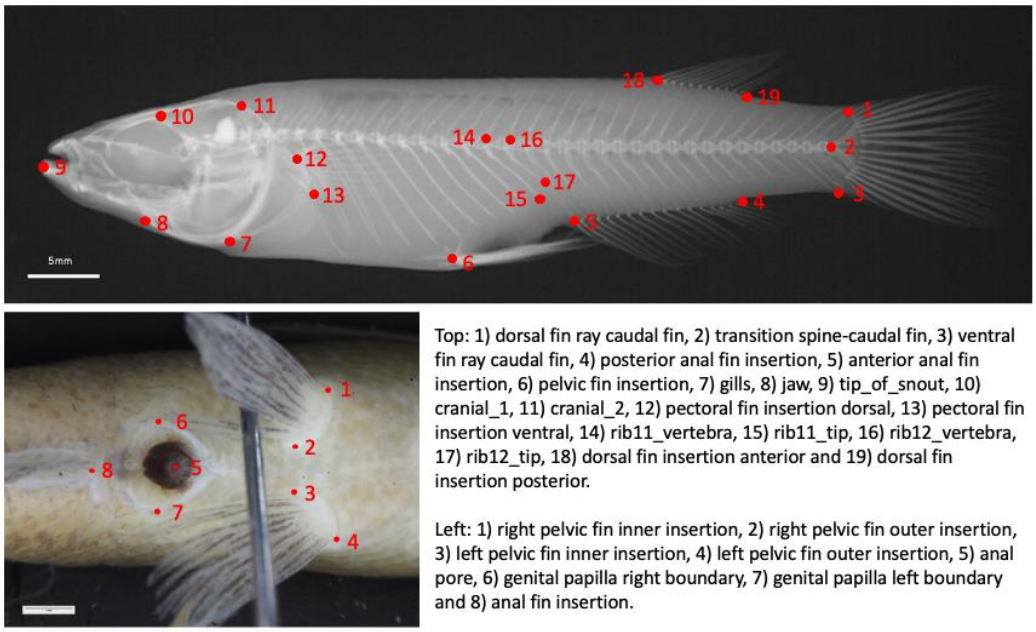
Table S2: Results of the body shape modularity analysis. CR = covariance ratio, CR_p = p-value of covariance ratio, Zcr = effect size of covariance ratio, stdZcr = standard deviation of effect size.

Directory

Supplementary Figures

Supplementary Tables

Supplementary Methods



Top: 1) dorsal fin ray caudal fin, 2) transition spine-caudal fin, 3) ventral fin ray caudal fin, 4) posterior anal fin insertion, 5) anterior anal fin insertion, 6) pelvic fin insertion, 7) gills, 8) jaw, 9) tip_of_snout, 10) cranial_1, 11) cranial_2, 12) pectoral fin insertion dorsal, 13) pectoral fin insertion ventral, 14) rib11_vertebra, 15) rib11_tip, 16) rib12_vertebra, 17) rib12_tip, 18) dorsal fin insertion anterior and 19) dorsal fin insertion posterior.

Left: 1) right pelvic fin inner insertion, 2) right pelvic fin outer insertion, 3) left pelvic fin inner insertion, 4) left pelvic fin outer insertion, 5) anal pore, 6) genital papilla right boundary, 7) genital papilla left boundary and 8) anal fin insertion.

Figure S1: Position of landmarks for Principal component analysis of body shape and the ventral side.

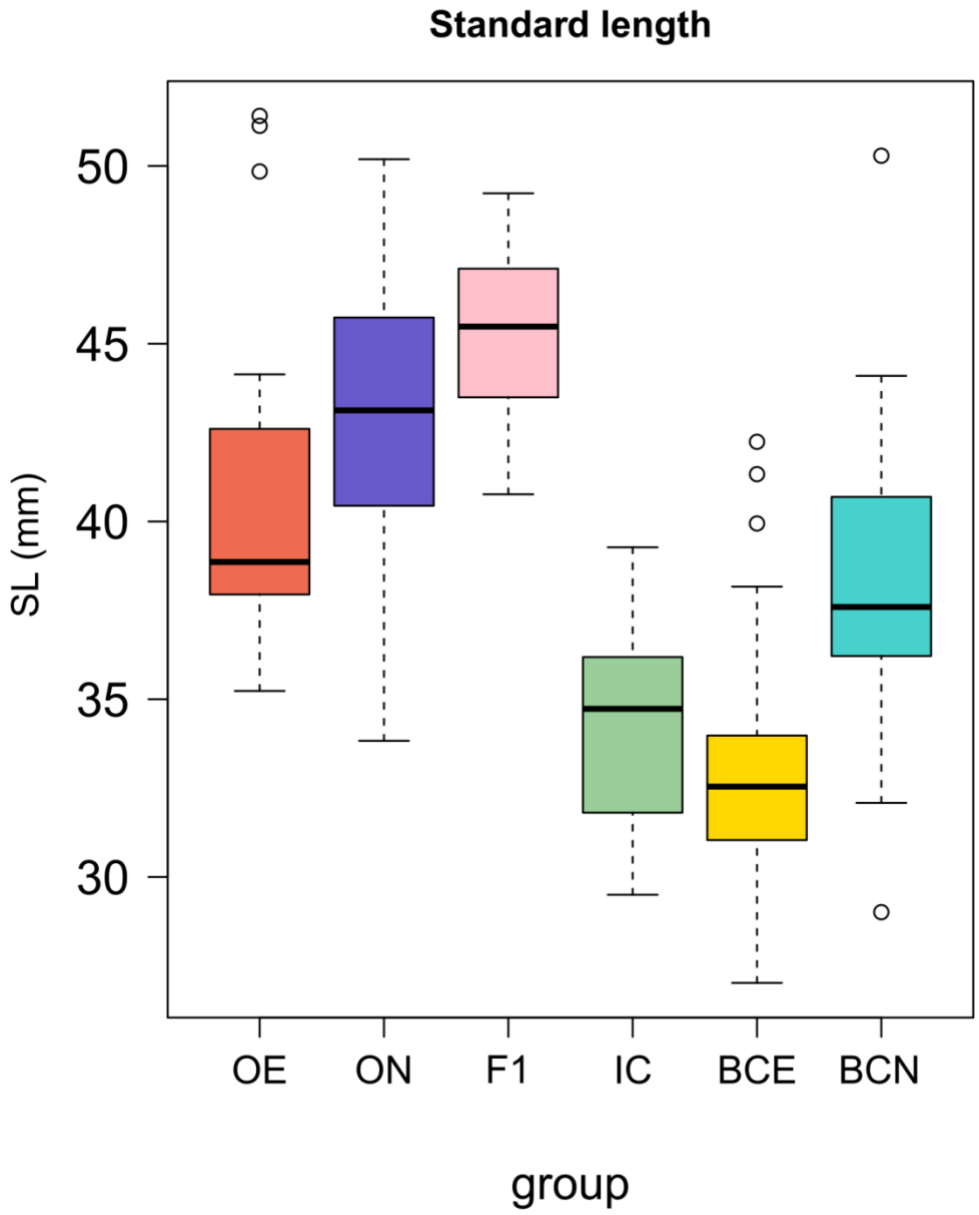


Figure S2: Standard length distribution of the parental species (OE = *O.eversi* and ON = *O. nigrimas*) and the crosses.

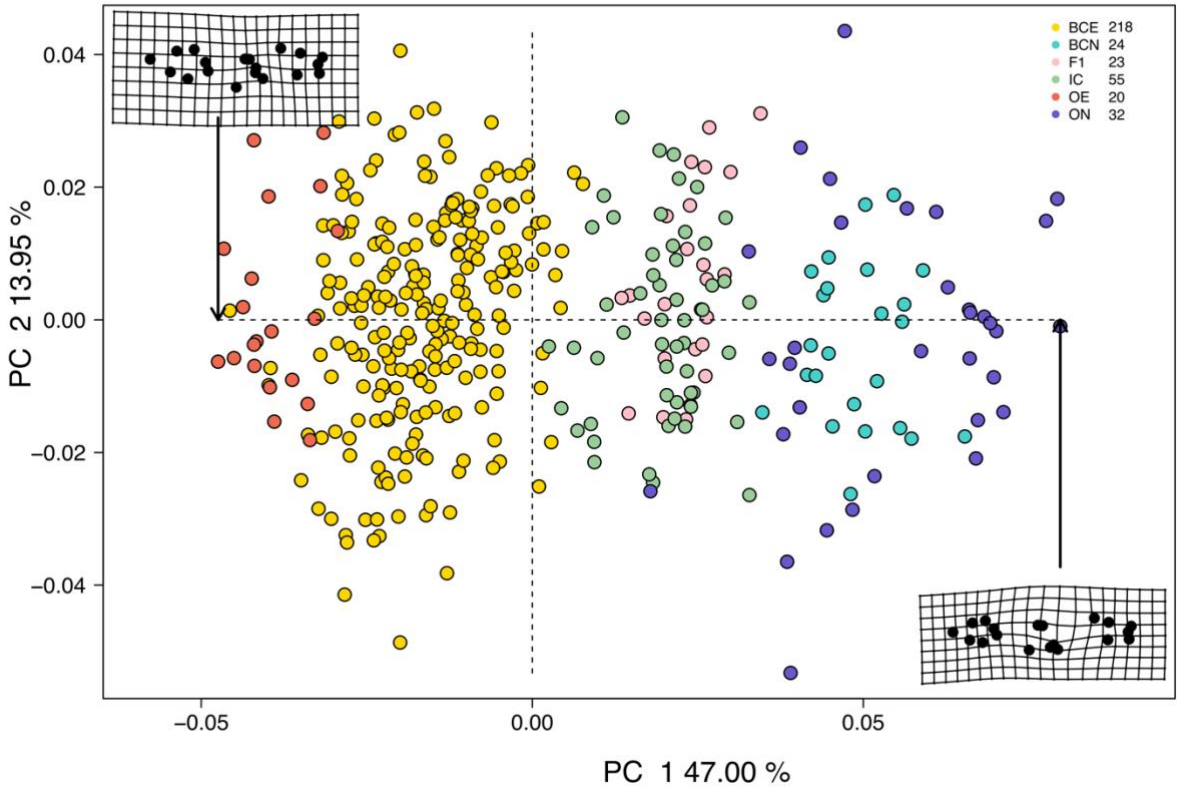


Figure S3: PCA plot of body shape. PC1 explains 47.0% percent of the variation and PC2 14.0%.

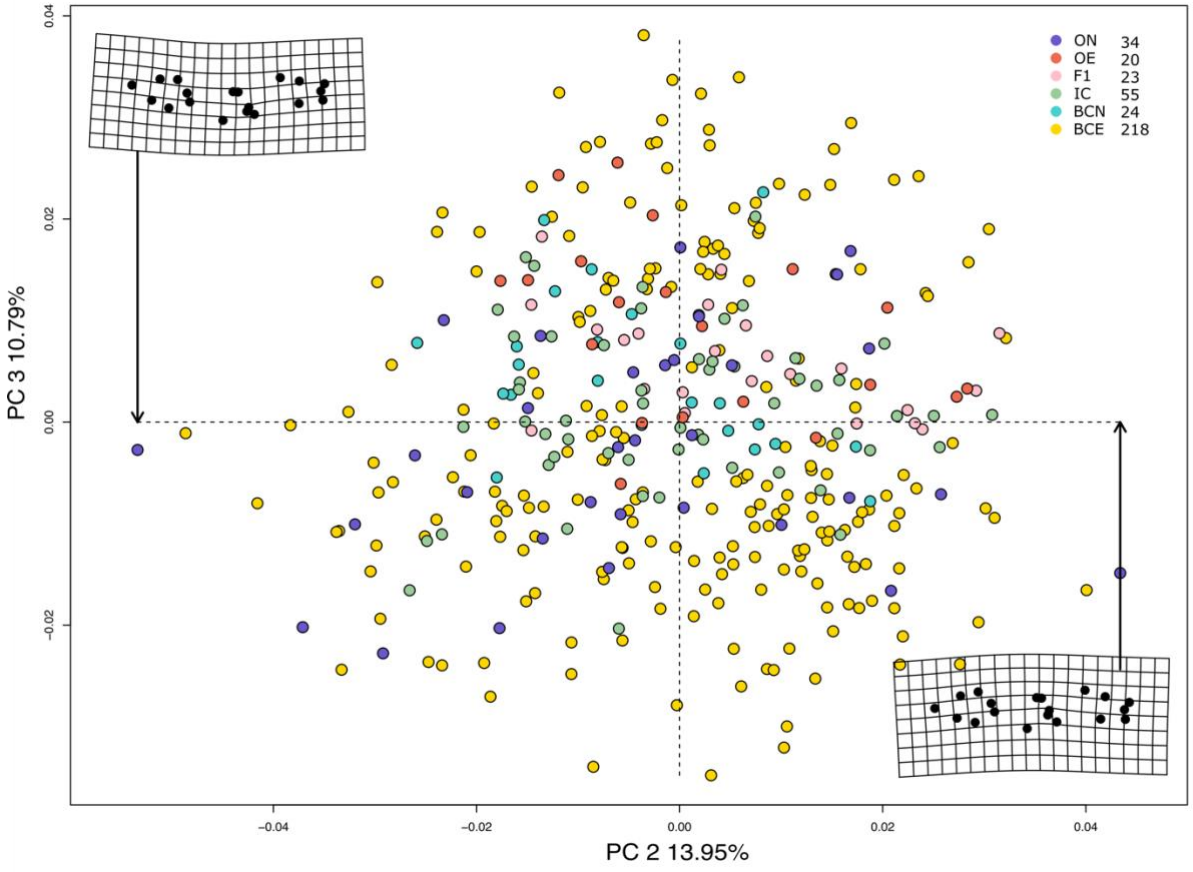


Figure S4: PCA plot of body shape. PC2 14.0% and PC3 10.8%.

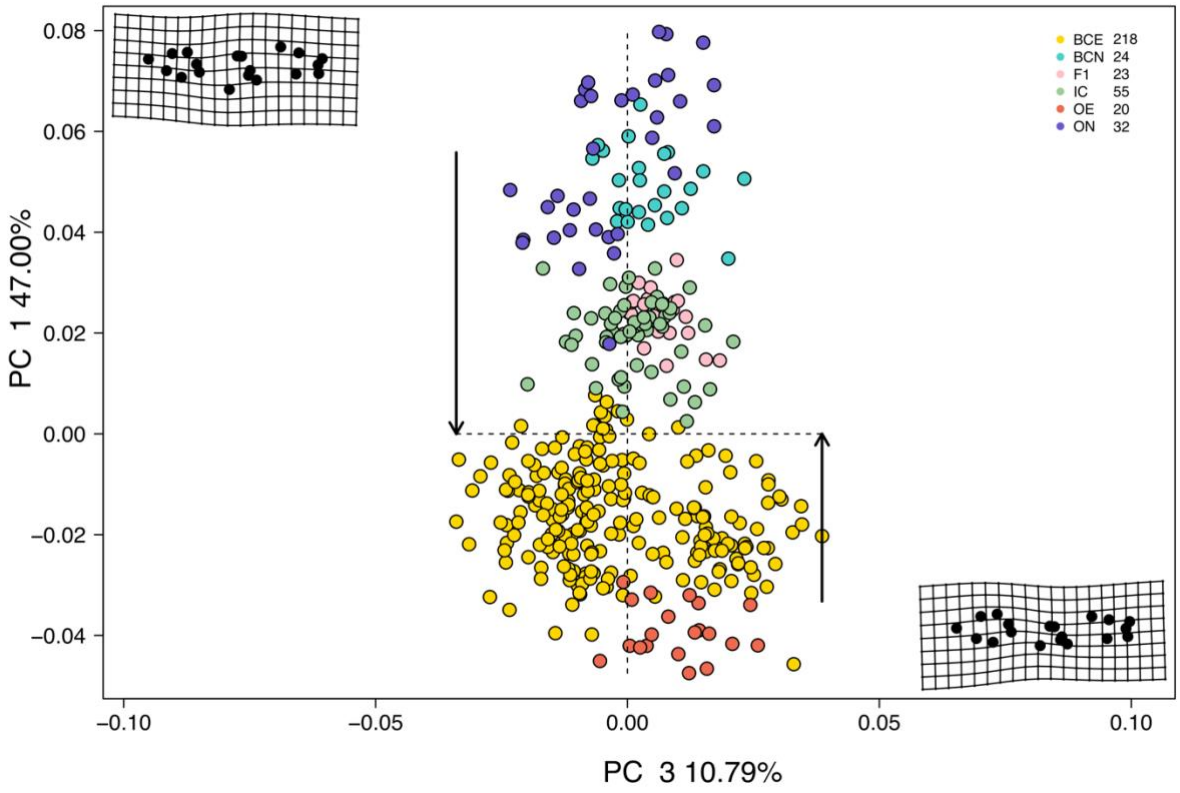


Figure S5: PCA plot of body shape. PC1 explains 47.0% percent of the variation and PC3 10.8%.

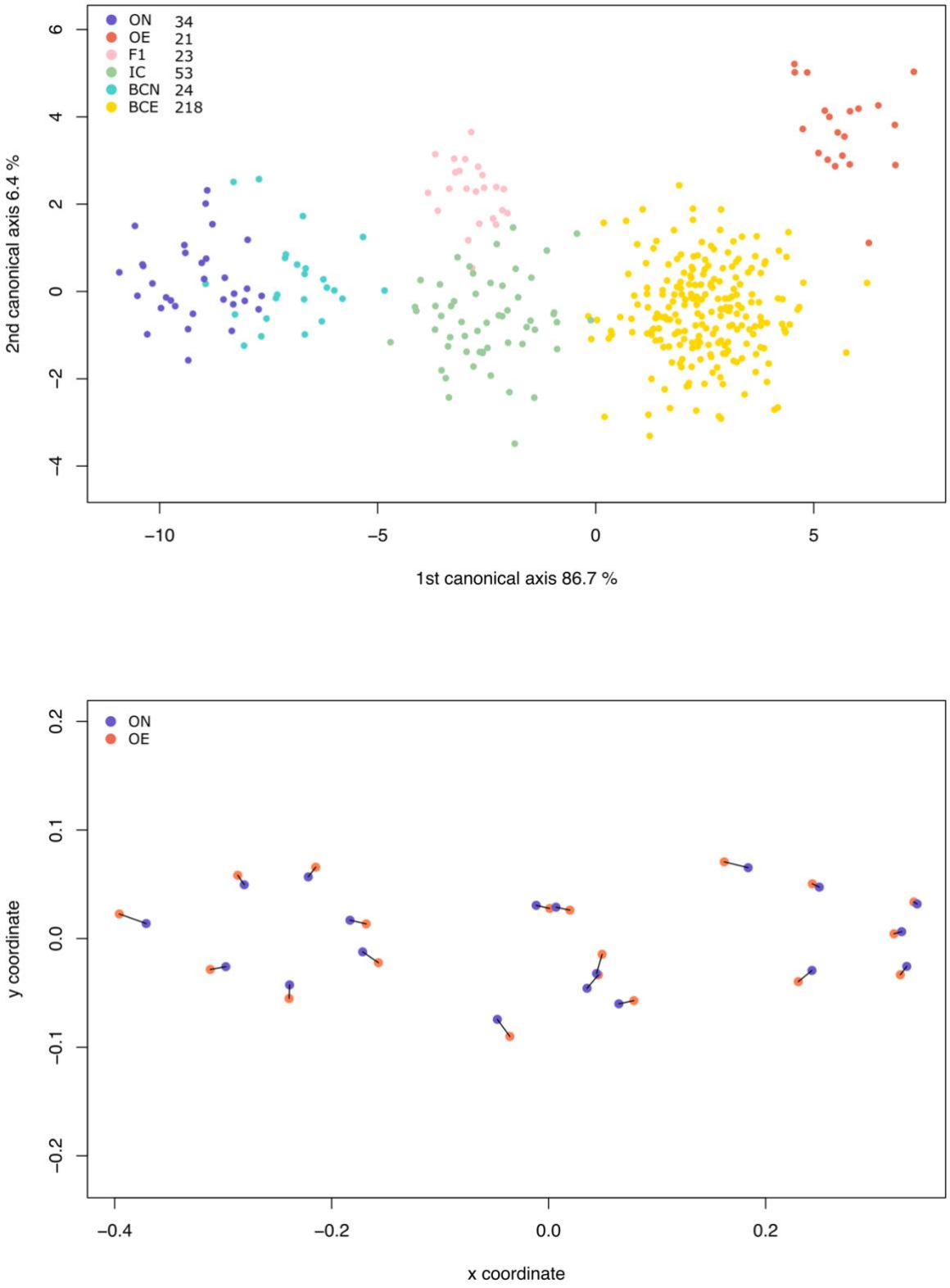


Figure S6: Canonical variate analysis of body shape. The results are similar to the principal component analysis, though here even 86.7% of the variation can be explained in the first canonical axis.

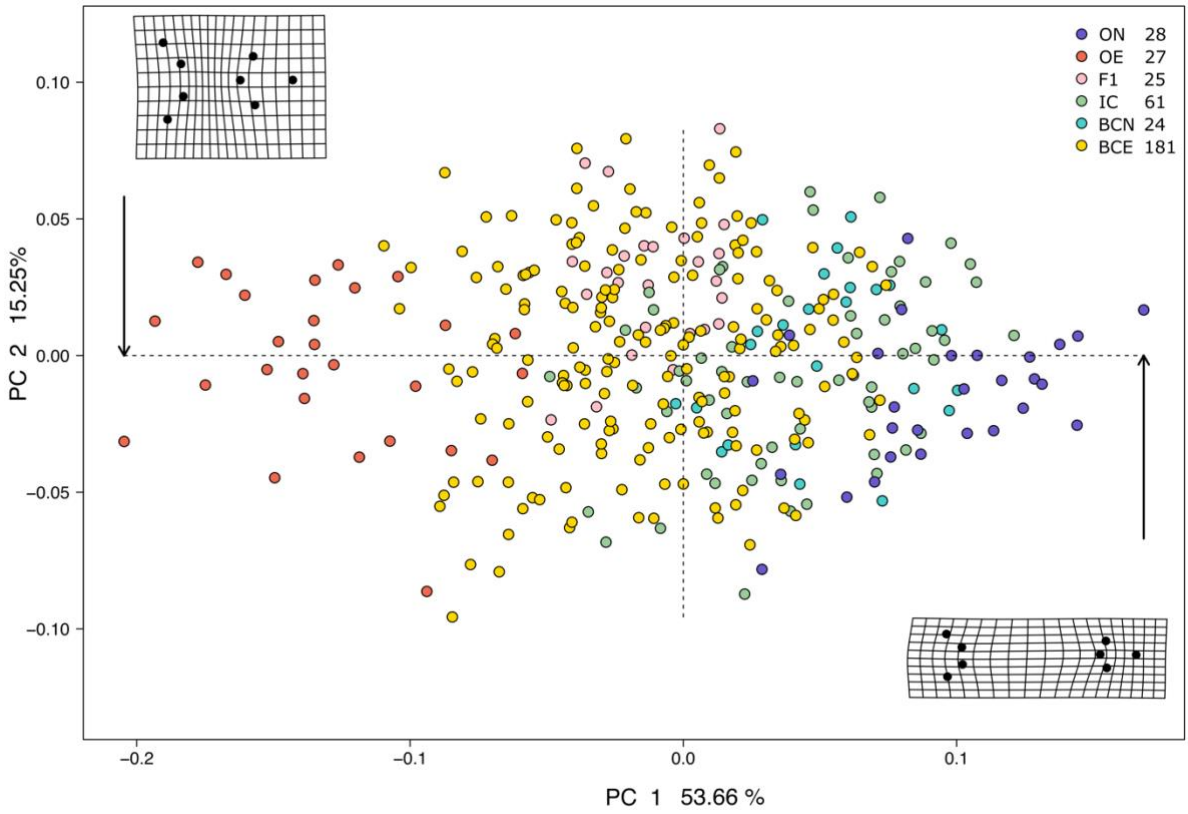
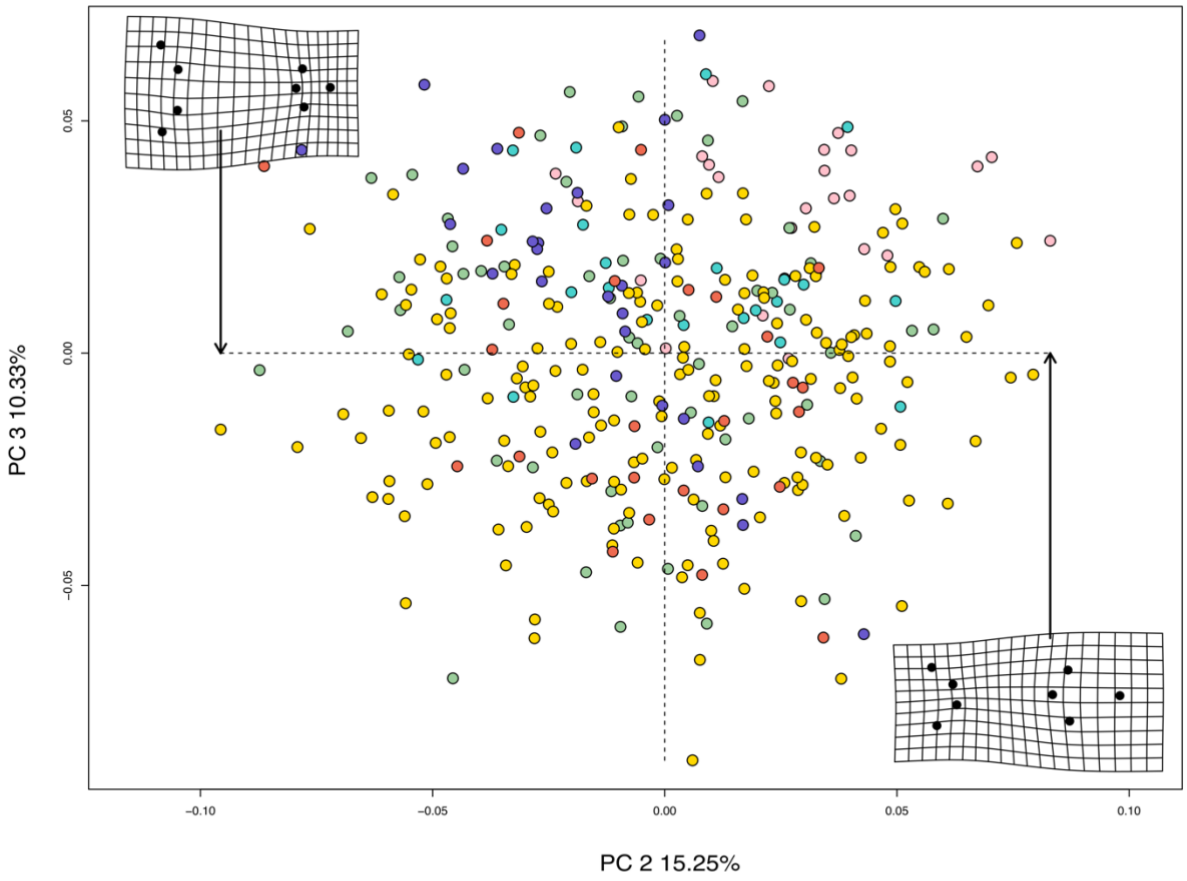


Figure S7: PCA plot of ventral body shape. PC1 explains 53.7% percent of the variation and PC2 15.3%.



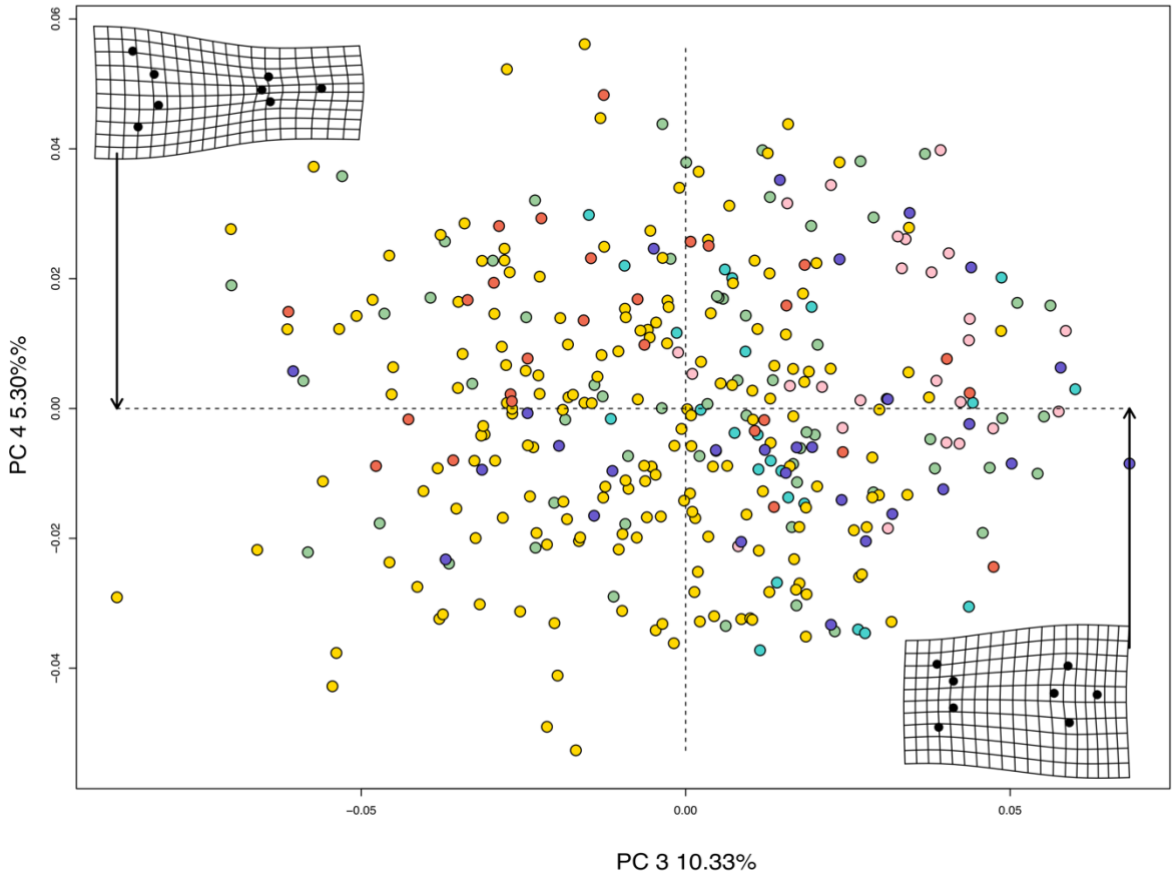


Figure S8: PCA plot of ventral body shape. PC2 explains 15.3%, PC3 10.3 % and PC4 5.3%.

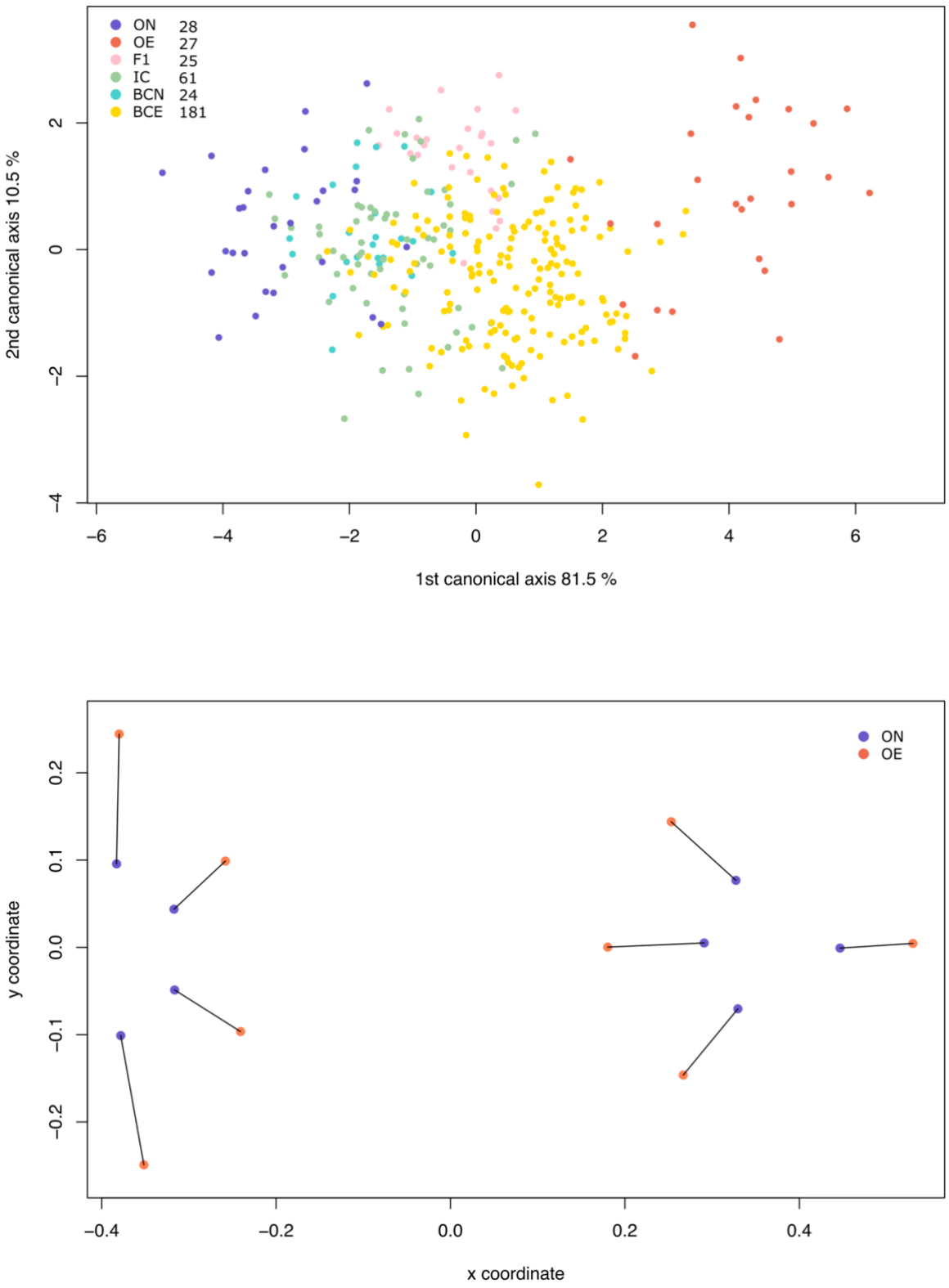


Figure S9: Canonical variate analysis of ventral body shape. The results are similar to the principal component analysis, though here even 81.5 % of the variation can be explained in the first canonical axis.

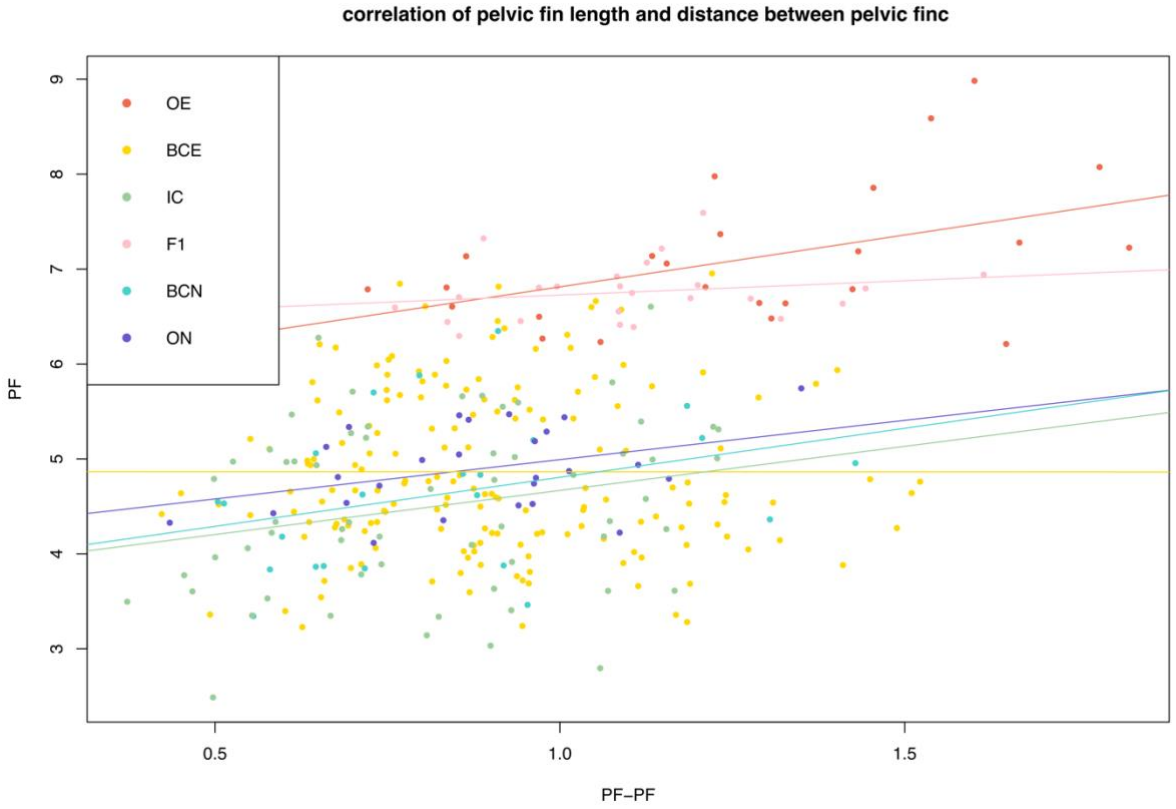


Figure S10: Scatterplot of pelvic fin length and pelvic fin insertion distance of the six tested lines. Lines are linear regressions and all are non-significant.

Table S1

Cross / Species	Abbreviation	Year_of_death	ID	origin
O. eversi	OE	2015	OE_2015_1	lab-raised
O. eversi	OE	2015	OE_2015_4	lab-raised
O. eversi	OE	2015	OE_2015_5	lab-raised
O. eversi	OE	2015	OE_2015_8	lab-raised
O. eversi	OE	2015	OE_2015_11	lab-raised
O. eversi	OE	2015	OE_2015_12	lab-raised
O. eversi	OE	2015	OE_2015_14	lab-raised
O. eversi	OE	2015	OE_2015_16	lab-raised
O. eversi	OE	2015	OE_2015_17	lab-raised
O. eversi	OE	2015	OE_2015_20	lab-raised
O. eversi	OE	2015	OE_2015_30	lab-raised
O. eversi	OE	2015	OE_2015_36	lab-raised
O. eversi	OE	2015	OE_2015_38	lab-raised
O. eversi	OE	2019	BCPOEf11_3	lab-raised
O. eversi	OE	2019	BCPOEf11_4	lab-raised
O. eversi	OE	2019	BCPOEf11_5	lab-raised
O. eversi	OE	2019	BCPOEf11_7	lab-raised
O. eversi	OE	2019	BCPOEf14_25	lab-raised
O. eversi	OE	2019	BCPOEf14_26	lab-raised
O. eversi	OE	2019	BCPOEf17_14	lab-raised
O. eversi	OE	2019	F1POEfT3_3	lab-raised
O. eversi	OE	2019	F1POEfT3_4	lab-raised
O. eversi	OE	2019	OEfRAD_1	lab-raised
O. eversi	OE	2019	OEfRAD_2	lab-raised
O. nigrimas	ON	2015	ON_2015_4	lab-raised
O. nigrimas	ON	2015	ON_2015_5	lab-raised
O. nigrimas	ON	2015	ON_2015_11	lab-raised
O. nigrimas	ON	2015	ON_2015_18	lab-raised
O. nigrimas	ON	2015	ON_2015_19	lab-raised
O. nigrimas	ON	2015	ON_2015_22	lab-raised
O. nigrimas	ON	2015	ON_2015_24	lab-raised
O. nigrimas	ON	2015	ON_2015_23	lab-raised
O. nigrimas	ON	2019	ONfRAD_1	lab-raised
O. nigrimas	ON	2019	ONfRAD_2	lab-raised
O. nigrimas	ON	2019	ONfRAD_3	lab-raised
O. nigrimas	ON	2019	OL19_ON_1	lab-raised
O. nigrimas	ON	2019	OL19_ON_2	lab-raised
O. nigrimas	ON	2019	OL19_ON_3	lab-raised
O. nigrimas	ON	2019	OL19_ON_4	lab-raised
O. nigrimas	ON	2019	OL19_ON_5	lab-raised
O. nigrimas	ON	2019	OL19_ON_6	lab-raised
O. nigrimas	ON	2019	OL19_ON_7	lab-raised

O. nigrimas	ON	2019	OL19_ON_8	lab-raised
O. nigrimas	ON	2019	OL19_ON_9	lab-raised
O. nigrimas	ON	2019	OL19_ON_10	lab-raised
O. nigrimas	ON		O_n_H7581.1	museum-specimen
O. nigrimas	ON		O_n_H7581.10	museum-specimen
O. nigrimas	ON		O_n_H7581.12	museum-specimen
O. nigrimas	ON		O_n_H7581.13	museum-specimen
O. nigrimas	ON		O_n_H7581.4	museum-specimen
O. nigrimas	ON		O_n_H7581.9	museum-specimen
O. nigrimas	ON	2019	JS19_416_1	wild-caught
O. nigrimas	ON	2019	JS19_416_2	wild-caught
O. nigrimas	ON	2019	JS19_416_3	wild-caught
O. nigrimas	ON	2019	JS19_416_4	wild-caught
O. nigrimas	ON	2019	JS19_416_5	wild-caught
O. nigrimas	ON	2019	JS19_417_1	wild-caught
O. nigrimas	ON	2019	JS19_417_2	wild-caught
O. nigrimas	ON	2019	JS19_417_3	wild-caught
O. nigrimas	ON	2019	JS19_417_4	wild-caught
O. nigrimas	ON	2019	JS19_417_5	wild-caught
F1	F1	2015	F1_2015_1	lab-raised
F1	F1	2015	F1_2015_2	lab-raised
F1	F1	2015	F1_2015_3	lab-raised
F1	F1	2015	F1_2015_5	lab-raised
F1	F1	2015	F1_2015_7	lab-raised
F1	F1	2015	F1_2015_8	lab-raised
F1	F1	2015	F1_2015_9	lab-raised
F1	F1	2015	F1_2015_11	lab-raised
F1	F1	2015	F1_2015_12	lab-raised
F1	F1	2019	OL19_F1_20	lab-raised
F1	F1	2019	OL19_F1_21	lab-raised
F1	F1	2019	OL19_F1_22	lab-raised
F1	F1	2019	OL19_F1_23	lab-raised
F1	F1	2019	OL19_F1_24	lab-raised
F1	F1	2019	OL19_F1_25	lab-raised
F1	F1	2019	OL19_F1_26	lab-raised
F1	F1	2019	OL19_F1_27	lab-raised
F1	F1	2019	OL19_F1_28	lab-raised
F1	F1	2019	OL19_F1_29	lab-raised
F1	F1	2019	OL19_F1_30	lab-raised
F1	F1	2019	OL19_F1_31	lab-raised

F1	F1	2019	OL19_F1_32	lab-raised
F1	F1	2019	OL19_F1_33	lab-raised
F1	F1	2019	OL19_F1_34	lab-raised
Intercross	IC	2015	IC_2015_6	lab-raised
Intercross	IC	2015	IC_2015_19	lab-raised
Intercross	IC	2015	IC_2015_22	lab-raised
Intercross	IC	2015	IC_2015_23	lab-raised
Intercross	IC	2015	IC_2015_26	lab-raised
Intercross	IC	2015	IC_2015_36	lab-raised
Intercross	IC	2015	IC_2015_55	lab-raised
Intercross	IC	2015	IC_2015_57	lab-raised
Intercross	IC	2015	IC_2015_64	lab-raised
Intercross	IC	2015	IC_2015_69	lab-raised
Intercross	IC	2015	IC_2015_70	lab-raised
Intercross	IC	2015	IC_2015_74	lab-raised
Intercross	IC	2015	IC_2015_78	lab-raised
Intercross	IC	2015	IC_2015_80	lab-raised
Intercross	IC	2015	IC_2015_82	lab-raised
Intercross	IC	2015	IC_2015_88	lab-raised
Intercross	IC	2015	IC_2015_91	lab-raised
Intercross	IC	2015	IC_2015_94	lab-raised
Intercross	IC	2015	IC_2015_95	lab-raised
Intercross	IC	2015	IC_2015_98	lab-raised
Intercross	IC	2015	IC_2015_100	lab-raised
Intercross	IC	2015	IC_2015_101	lab-raised
Intercross	IC	2015	IC_2015_103	lab-raised
Intercross	IC	2015	IC_2015_106	lab-raised
Intercross	IC	2015	IC_2015_110	lab-raised
Intercross	IC	2015	IC_2015_116	lab-raised
Intercross	IC	2015	IC_2015_117	lab-raised
Intercross	IC	2015	IC_2015_122	lab-raised
Intercross	IC	2015	IC_2015_126	lab-raised
Intercross	IC	2015	IC_2015_128	lab-raised
Intercross	IC	2015	IC_2015_130	lab-raised
Intercross	IC	2015	IC_2015_135	lab-raised
Intercross	IC	2015	IC_2015_149	lab-raised
Intercross	IC	2015	IC_2015_150	lab-raised
Intercross	IC	2015	IC_2015_159	lab-raised
Intercross	IC	2015	IC_2015_160	lab-raised
Intercross	IC	2015	IC_2015_164	lab-raised
Intercross	IC	2015	IC_2015_165	lab-raised
Intercross	IC	2015	IC_2015_169	lab-raised
Intercross	IC	2015	IC_2015_171	lab-raised
Intercross	IC	2015	IC_2015_172	lab-raised
Intercross	IC	2015	IC_2015_187	lab-raised

Intercross	IC	2019	F2 IC-OL2019-6	lab-raised
Intercross	IC	2019	F2 IC-OL2019-9	lab-raised
Intercross	IC	2019	F2 IC-OL2019-10	lab-raised
Intercross	IC	2019	F2 IC-OL2019-16	lab-raised
Intercross	IC	2019	F2 IC-OL2019-17	lab-raised
Intercross	IC	2019	F2 IC-OL2019-19	lab-raised
Intercross	IC	2019	F2 IC-OL2019-22	lab-raised
Intercross	IC	2019	F2 IC-OL2019-23	lab-raised
Intercross	IC	2019	F2 IC-OL2019-25	lab-raised
Intercross	IC	2019	F2 IC-OL2019-27	lab-raised
Intercross	IC	2019	F2 IC-OL2019-29	lab-raised
Intercross	IC	2019	F2 IC-OL2019-37	lab-raised
Intercross	IC	2019	F2 IC-OL2019-38	lab-raised
Intercross	IC	2019	F2 IC-OL2019-45	lab-raised
Intercross	IC	2019	F2 IC-OL2019-50	lab-raised
Intercross	IC	2019	F2 IC-OL2019-58	lab-raised
Intercross	IC	2019	F2 IC-OL2019-68	lab-raised
Intercross	IC	2019	F2 IC-OL2019-73	lab-raised
Intercross	IC	2019	F2 IC-OL2019-90	lab-raised
Intercross	IC	2019	F2 IC-OL2019-92	lab-raised
Intercross	IC	2019	F2 IC-OL2019-93	lab-raised
Intercross	IC	2019	F2 IC-OL2019-98	lab-raised
Intercross	IC	2019	F2 IC-OL2019-99	lab-raised
Intercross	IC	2019	F2 IC-OL2019-105	lab-raised
Intercross	IC	2019	F2 IC-OL2019-107	lab-raised
Intercross	IC	2019	F2 IC-OL2019-110	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_1	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_5	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_7	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_9	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_10	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_11	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_12	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_15	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_19	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_21	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_23	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_25	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_30	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_34	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_35	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_36	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_38	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_39	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_41	lab-raised

O. eversi backcross	BCE	2015	BCE_2015_44	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_49	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_56	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_59	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_61	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_62	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_74	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_75	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_76	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_77	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_78	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_84	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_85	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_93	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_94	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_97	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_98	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_99	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_100	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_102	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_104	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_108	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_111	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_114	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_115	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_117	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_118	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_120	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_122	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_128	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_129	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_133	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_137	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_138	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_140	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_141	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_142	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_143	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_150	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_152	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_153	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_154	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_155	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_158	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_162	lab-raised

O. eversi backcross	BCE	2015	BCE_2015_164	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_168	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_169	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_171	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_176	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_180	lab-raised
O. eversi backcross	BCE	2015	BCE_2015_182	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_7	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_8	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_11	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_12	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_16	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_18	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_19	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_23	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_25	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_29	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_30	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_31	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_33	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_35	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_37	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_38	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_39	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_40	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_43	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_46	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_49	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_51	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_52	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_53	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_57	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_60	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_61	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_62	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_63	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_64	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_65	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_66	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_68	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_71	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_72	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_74	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_75	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_76	lab-raised

O. eversi backcross	BCE	2018	BCE_2018_78	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_80	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_82	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_83	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_84	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_85	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_91	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_92	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_93	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_94	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_97	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_98	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_100	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_104	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_107	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_110	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_113	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_116	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_118	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_120	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_121	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_123	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_124	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_126	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_127	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_130	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_132	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_135	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_137	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_139	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_141	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_143	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_144	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_145	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_146	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_150	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_152	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_153	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_155	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_157	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_160	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_161	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_162	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_167	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_169	lab-raised

O. eversi backcross	BCE	2018	BCE_2018_170	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_175	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_177	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_178	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_179	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_181	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_183	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_184	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_185	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_187	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_190	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_192	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_196	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_197	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_198	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_199	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_200	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_201	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_203	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_205	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_206	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_207	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_208	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_209	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_210	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_211	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_212	lab-raised
O. eversi backcross	BCE	2018	BCE_2018_213	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_214	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_217	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_219	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_220	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_222	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_223	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_224	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_227	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_228	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_229	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_231	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_232	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_236	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_237	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_238	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_240	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_241	lab-raised

O. eversi backcross	BCE	2019	BCE_2019_246	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_247	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_248	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_249	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_251	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_252	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_254	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_256	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_257	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_260	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_262	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_264	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_265	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_267	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_268	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_270	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_273	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_279	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_280	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_291	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_292	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_293	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_295	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_296	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_297	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_299	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_300	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_306	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_308	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_310	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_311	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_312	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_317	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_318	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_319	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_320	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_323	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_324	lab-raised
O. eversi backcross	BCE	2019	BCE_2019_325	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_30	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_33	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_37	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_46	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_49	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_55	lab-raised

O. nigrimas backcross	BCN	2015	BCN_2015_69	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_73	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_76	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_79	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_90	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_112	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_118	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_128	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_142	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_157	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_161	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_162	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_165	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_167	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_173	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_183	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_236	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_247	lab-raised
O. nigrimas backcross	BCN	2015	BCN_2015_258	lab-raised

Table S2

Species	Hypothesis	CR	CR_p	Zcr	stdZcr
OE	H1	1.022125	0.254	-0.6506326	0.0708459
ON	H1	1.107771	0.601	0.2617138	0.06995404
OE	H1+a	0.9793011	0.077	-1.5295798	0.05681867
ON	H1+a	1.124128	0.602	0.3098917	0.06196888
OE	H2	0.8304398	0.004	-2.8278183	0.05588397
ON	H2	1.010778	0.164	-1.0762044	0.05877345
OE	H2+a	0.857989	0.001	-2.8930741	0.05968967
ON	H2+a	1.064609	0.287	-0.6234781	0.06136893
OE	H3	0.8323193	0.005	-2.6189632	0.1210259
ON	H3	0.9865468	0.034	-1.8503426	0.12722678
OE	H3+a	0.8176493	0.001	-3.7745855	0.10749192
ON	H3+a	1.012612	0.029	-1.8787785	0.10232962

Repeatability test for landmarks used in the whole body shape analysis and ventral body shape analysis

Methods:
For both analysis we chose 18 individuals, two of each parental species, two of the F1 generation and three of each F2 cross. They were randomly chosen within their respective group. We placed the same landmarks as described in the method sections "Geometric morphometric analysis of body shape" and "Geometric morphometrics of ventral body shape". The landmarks were placed twice on each set, with one day (body shape) and two days (ventral body shape) break in between. As all landmarks in this study, the landmarks were placed by the same observer (xxx). For the second placement of the landmarks, the order of the specimen was shuffled.
To test the repeatability, an ANOVA implemented in geomorph was run, with the coordinates as response variable and the specimen ID as predictor (Fruciano 2016). Using the formula (Fleiss and Shrout 1977; Sokal and Rohlf 1995; Arnqvist and Mårtensson 1998)

$$\boldsymbol{R} = \boldsymbol{S_A^2}/(\boldsymbol{S_W^2} + \boldsymbol{S_A^2}) \qquad (1)$$

$\boldsymbol{S_A^2}$ is the among-individuals variance and $\boldsymbol{S_W^2}$ is the within-individuals variance component. $\boldsymbol{S_A^2}$ is calculated by

$$\boldsymbol{S_A^2} = \boldsymbol{MS_{among}} - \boldsymbol{MS_{within}} \,/\, \boldsymbol{n} \qquad (2)$$

$\boldsymbol{S_W^2}$ is calculated by

$$\boldsymbol{S_W^2} = \boldsymbol{MS_{within}} \qquad (3)$$

$\boldsymbol{n}$ is the number of repeated measurements, $\boldsymbol{MS_{among}}$ is the among-groups ANOVA sum of squares and $\boldsymbol{MS_{within}}$ the within-group ANOVA sum of squares.

Results:
For the body shape landmarks, R is 98.5% (DF = 17, F = 68.592). For the ventral body shape the repeatability is a bit lower (R = 95.9%, DF = 17, F = 24.944) which still means less than 5% error and therefor high repeatability in both analyses.

References

Arnqvist, G., and T. Mårtensson. 1998. Measurement error in geometric morphometrics: Empirical strategies to assess and reduce its impact on measures of shape. Acta Zool. Acad. Sci. Hungaricae 44:73–96.

Fleiss, J. L., and P. E. Shrout. 1977. The effects of measurement errors on some multivariate procedures. Am. J. Public Health 67:1188–1191.

Fruciano, C. 2016. Measurement error in geometric morphometrics. Dev. Genes Evol. 226:139–158. Springer.

Sokal, R., and F. Rohlf. 1995. Biometry. The Principles and Practice of Statistics in Biological Research. WH Free Company, New York.

Contribution of hybridization to the evolution of a derived reproductive strategy in ricefishes

J. M. Flury[1], K. Meusemann[1], S. Martin[1], L. Hilgers[1], T. Spanke[1], A. Böhne[1], F. Herder[1], D. Mokodongan[2], J. Altmüller[3], D. Wowor[2], B. Misof[1], A.W. Nolte[4], J. Schwarzer[1]

[1] Leibniz-Institute for the Analysis of Biodiversity Change (LIB), Museum Koenig, Bonn, Germany

[2] Museum Zoologicum Bogoriense, Research Center for Biosystematics and Evolution, National Research and Innovation Agency (BRIN), Cibinong 16911, West Java, Indonesia

[3] Cologne Center for Genomics (CCG), Cologne University, Cologne, Germany

[4] Carl von Ossietzky Universität, Oldenburg, Germany

Corresponding author: Jana Flury (j.flury@leibniz-lib.de)

Abstract

Transitions between reproductive strategies are costly and involve major changes in life history, behaviour and morphology. Nevertheless, in Sulawesi ricefishes, pelvic-brooding evolved from transfer-brooding in two distantly related lineages within in the genera *Adrianichthys* and in *Oryzias*, respectively. Females of pelvic-brooding species carry their eggs attached to their belly until the fry hatches. Despite their phylogenetic distance, both pelvic-brooding lineages share a set of external morphological traits. A recent study found no direct gene-flow between pelvic-brooding lineages suggesting independent evolution of the derived reproductive strategy. It could, however, also be more complex. Pre-existing variation in an admixed population may enable the re-use of genetic variants when subjected to similar external selection pressure, resulting in similar phenotypes.

We thus used a multi-species coalescent (MSC) model and D-statistics to identify gene-tree – species-tree incongruencies, to evaluate the evolution of pelvic-brooding with respect to inter-specific gene-flow not only between pelvic-brooding lineages, but between pelvic-brooding lineages and other Sulawesi ricefish lineages. We found a general network-like evolution in Sulawesi ricefishes and as previously reported, no gene-flow between the pelvic-brooding lineages. Instead, we found hybridization between the ancestor of pelvic-brooding *Oryzias* and the common ancestor of the four *Oryzias* species from Lake Poso, home of the pelvic-brooding *Adrianichthys* lineage. Further, indications of introgression were located within two confidence intervals of quantitative trait loci (QTL) associated with pelvic-brooding in *O. eversi*. We thus hypothesize that a mix of *de novo* mutations and (ancient) standing genetic variation shaped the evolution of pelvic-brooding.

Keywords: ancient gene flow, convergent evolution, pelvic-brooding, reproductive strategy, Sulawesi ricefishes

Significance statement

The evolution of pelvic-brooding in *Oryzias eversi* (Beloniformes:Adrianichthyidae), was recently described to be independent from another pelvic-brooding ricefish lineage (*Adrianichthys*) from Sulawesi. We confirmed these results, and detected no gene flow between the two distantly related pelvic-brooding lineages. Instead, we found ancient gene flow from another *Orzyias* lineage into the pelvic-brooding *Oryzias* lineage. One of the previously described QTL for pelvic brooding overlaps with a region of high introgression signal. Therefore, we assume that not only *de novo* mutations contributed to the evolution of pelvic-brooding in *Orzyias*, but that introduced ancient genetic variation was likely also recruited for the evolution of this derived brooding strategy.

Introduction

The evolutionary as well as the genetic basis for the repeatability of evolution exemplified by convergent traits intrigued biologists since the early days (Darwin 1859). Here, similar selective regimes are predicted to result in similar adaptive phenotypes (Schluter & Nagel 1995; Elmer & Meyer 2011). In Sulawesi ricefishes (Beloniformes; Adrianichthyidae), a group of freshwater fishes endemic to the island of Sulawesi, Indonesia, an exceptional reproductive strategy evolved in two distantly related lineages (>15 my divergence time, Hilgers & Schwarzer, 2019; Mokodongan & Yamahira, 2015): This so called "pelvic-brooding" evolved in the genus *Adrianichthys* and in two closely related *Oryzias* species: *O. eversi* and *O. sarasinorum* (Kottelat, 1990; Mokodongan & Yamahira, 2015; Parenti, 2008, Fig. 1). Females of pelvic-brooding ricefish species carry a cluster of eggs attached to their gonoduct for up to three weeks, until the fry hatches (Kottelat 1990). They have elongated pelvic fins and shorter ribs, forming a concavity where the egg cluster is situated (Spanke et al. 2021; Kottelat 1990). In contrast, in the more common and ancestral brooding strategy transfer-brooding (Parenti 2008), females deposit the eggs after several hours (Yamamoto 1975; Wootton & Smith 2014). Changing reproductive strategies entails severe changes in life-history, but also morphological adaptations which in this case are related to prolonged carrying of eggs (Spanke et al. 2021; Parenti 2008; Herder et al. 2012). How such major transitions of reproductive strategies evolve, also on a molecular level, remains unclear.

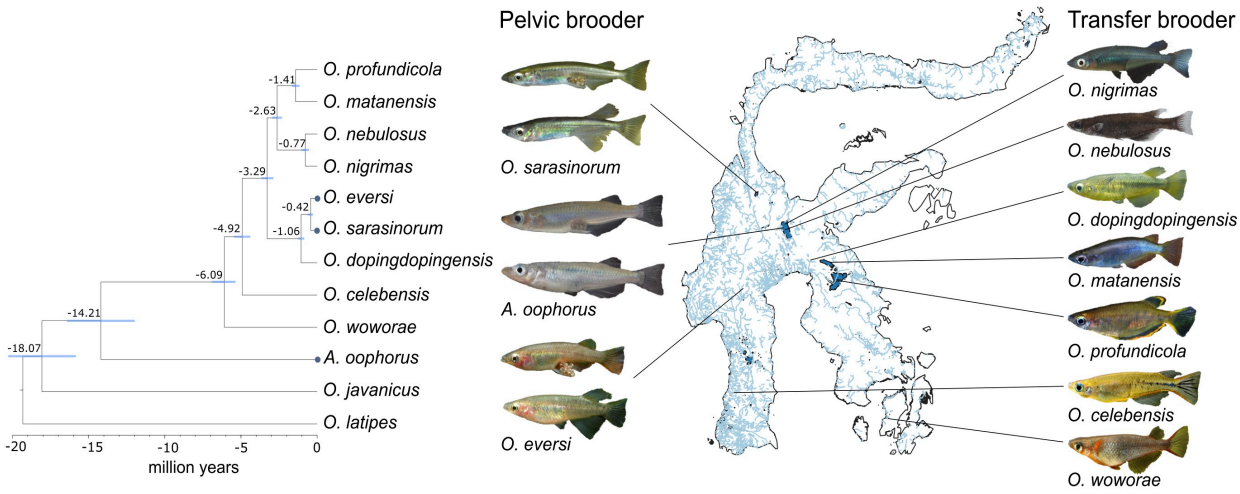


Fig 1: Dated phylogenetic tree adapted from Mokodongan & Yamahira (2015) on the left, pelvic-brooders are marked with blue dots. Map on the right with the distribution of pelvic-brooding (left, male and female showing the pronounced sexual dimorphism) and transfer-brooding species (right, only males). Each clade defined by Mokodongan & Yamahira 2015 for Sulawesi ricefishes, is represented by at least one species. Photos taken by Jan Möhring, Hans Evers and Andreas Wagnitz.

One genetic possibility how the transition from transfer- to pelvic-brooding could be realized are *de novo* mutations. However, they go to fixation rather slowly (Hermisson & Pennings 2005; Barrett & Schluter 2008) and are rarely beneficial (Ohta 1992). Alternatively, the evolution of similar phenotypic traits in different species can also be due to introgression or repeated sorting of shared standing genetic variation. Introgressed variation from one into another population or even species may facilitate the implementation of adaptive variants which are identical by descent (Waters & McCulloch 2021). This so-called adaptive introgression provides fast access to additional genetic variation, if species boundaries are weak (Mallet 2005; Whitney et al. 2006; Arnold 2007; Baack & Rieseberg 2007; Castric et al. 2008; Rieseberg 2009; The Heliconius Genome Consortium et al. 2012; Seehausen 2004). It was proposed that a high level of introgression led to the rapid radiation of *Heliconius* butterflies (Edelman et al. 2019), similar to Darwin's finches (Lamichhaney et al. 2015) and African Cichlids (Meier et al. 2017). Convergent evolution by repeated sorting (Waters & McCulloch

2021) involves alleles identical by descent, regardless of the potential source of variation, which may be recent (Colosimo et al. 2005; Terekhanova et al. 2014) or rather ancient (Nelson & Cresko 2018; Van Belleghem et al. 2018; van der Valk et al. 2021; Veale & Russello 2017; Brawand et al. 2014).

In the present study, we contrast a convergent evolution scenario of pelvic-brooding with the introgression of genetic variants, by using a phylogenomic approach. A recent study suggested no direct gene-flow between the two pelvic-brooding lineages (Montenegro et al. 2022). However, similar phenotypes may result from recombining pre-existing variation in an admixed population, where diverging species were exposed to similar external selection pressures and genetic variants were re-used (Schluter & Nagel 1995). Gene flow is abundant in Sulawesi ricefishes and has also been detected between pelvic-brooding *Oryzias* species (*O. sarasinorum* and *O. eversi*) (Horoiwa et al. 2021), within Lake Poso *Oryzias* (*O. nigrimas*, *O. nebulosus*, *O. orthognathus* and *O. soerotoi*) (Sutra et al. 2019) and the Malili Lake *Oryzias* (*O. matanensis, O. marmoratus* and *O. profundicola*) (Mandagi et al. 2021) which form the sister clades to the pelvic-brooding *Oryzias* (Fig. 1). Thus, gene flow from other sources into the pelvic-brooding lineages harbours the potential to increase genetic variation impacting the evolution of pelvic-brooding. We reconstructed 1907 gene-trees based on single-copy protein-coding orthologous genes of ten Sulawesi ricefishes, five mainland ricefishes and four outgroup species (two poeciliid and two killifish species each). We used these gene-trees as basis for a multi-species coalescence analysis to investigate gene-tree – species-tree incongruencies and to reconstruct a phylogenetic network. We further evaluated introgression on a genome-wide level on published genomes of 17 Sulawesi ricefishes and two mainland ricefish species.

Results

Filtering of single-copy orthologous genes

From the 8552 single-copy protein-coding genes derived from the orthologous gene set of Actinopterygii (ID 7898) from OrthoDB, 7331 genes were retained after removal of genes with identified outlier sequences. Randomly similar aligned sections were identified in 1806 genes, hence we masked them (16.5% of all base pairs). According to the TreeShrink analysis, 507 gene trees showed suspiciously long branches and we discarded them from the data set. Phylotreepruner, another method that uses a phylogenetic approach to identify potential wrongly identified orthologs returned no suspicious genes. For final analyses, we kept 1907 orthologous genes (2.415.561 bp) present in all 16 ricefish and four outgroup species.

Gene-tree – species-tree incongruencies and one moderately supported node in the ML tree

The topology found in the species tree resulting from Astral (Fig. 2A) is highly congruent with the one pulished by Mokodongan and Yamahira (2015) (Fig. 1). In our tree, however, *O. nigrimas* and *O. nebulosus* formed the sister clade to a clade comprising *O. eversi*, *O. sarasinorum* and *O. dopingdopingensis*. Matching this inconsistency, we found that these clades had many gene-tree – species-tree incongruencies (Fig. 2B). Even though the posterior support values were "1" in each branching event in the species tree from Astral (Fig.S1), quartet scores imply a high degree of uncertainty (Fig. 2C). We observed almost equal quartet frequencies for three splits (split 3, 4 and 7, Fig. 2C in red). At split 3, which indicates the pelvic-brooding *Oryzias* are sister to *O. dopingdopingensis,* one third of all gene trees supported a sistergroup relationship of the pelvic-brooding *Oryzias* lineage with the Lake Poso *Oryzias* and about one fifth supported that *O. dopingdopingensis* is most closely related to the Lake Poso *Oryzias*.

In split 4 – which was not inferred in the ML tree (Fig. 3), Lake Poso ricefishes were sistergroup to the Malili lake system ricefishes in one third of all gene trees and in one third the Malili lake system ricefishes were sistergroup to the pelvic-brooding species *O. eversi* and *O.*

sarasinorum and the transfer brooder *O. dopingdopingensis*. Regarding split 7, in one fourth of all gene trees *O. woworae* is sister to the other Sulawesi *Oryzias* instead of *O. celebensis*, and in another fourth of the gene trees, *O. woworae* and *O. celebensis* were sistergroups and the other Sulawesi *Oryzias* are sister to the rest of the ricefishes. In split 13, comprising also non-Sulawesi *Oryzias*, alternative quartetts occur with about one fourth of all gene trees supporting *O. latipes* and *O. mekongensis* as sistergroup to the Sulawesi ricefishes and one fifth *O. latipes* and *O. mekongensis* are sistergroup to a clade comprising *O. javanicus*, *O. melastigma* and *O. dancena*.

In the ML tree based on the concatenated supermatrix of 1907 orthologous genes, we found a third possible topology (Fig. 3). Here, *O. marmoratus* and *O. matanensis* are sistergroup to *O. eversi*, *O. sarasinorum* and *O. dopingdopingensis*. All splits obtained maximal bootstrap support, except for the sistergroup relationship between *O. marmoratus*, *O. matanensis* and *O. eversi*, *O. sarasinorum* and *O. dopingdopingensis*, which was only moderately supported. The ancestral state reconstruction supported that pelvic-brooding evolved convergently in *Oryzias* and *Adrianichthys* (Fig. 3).

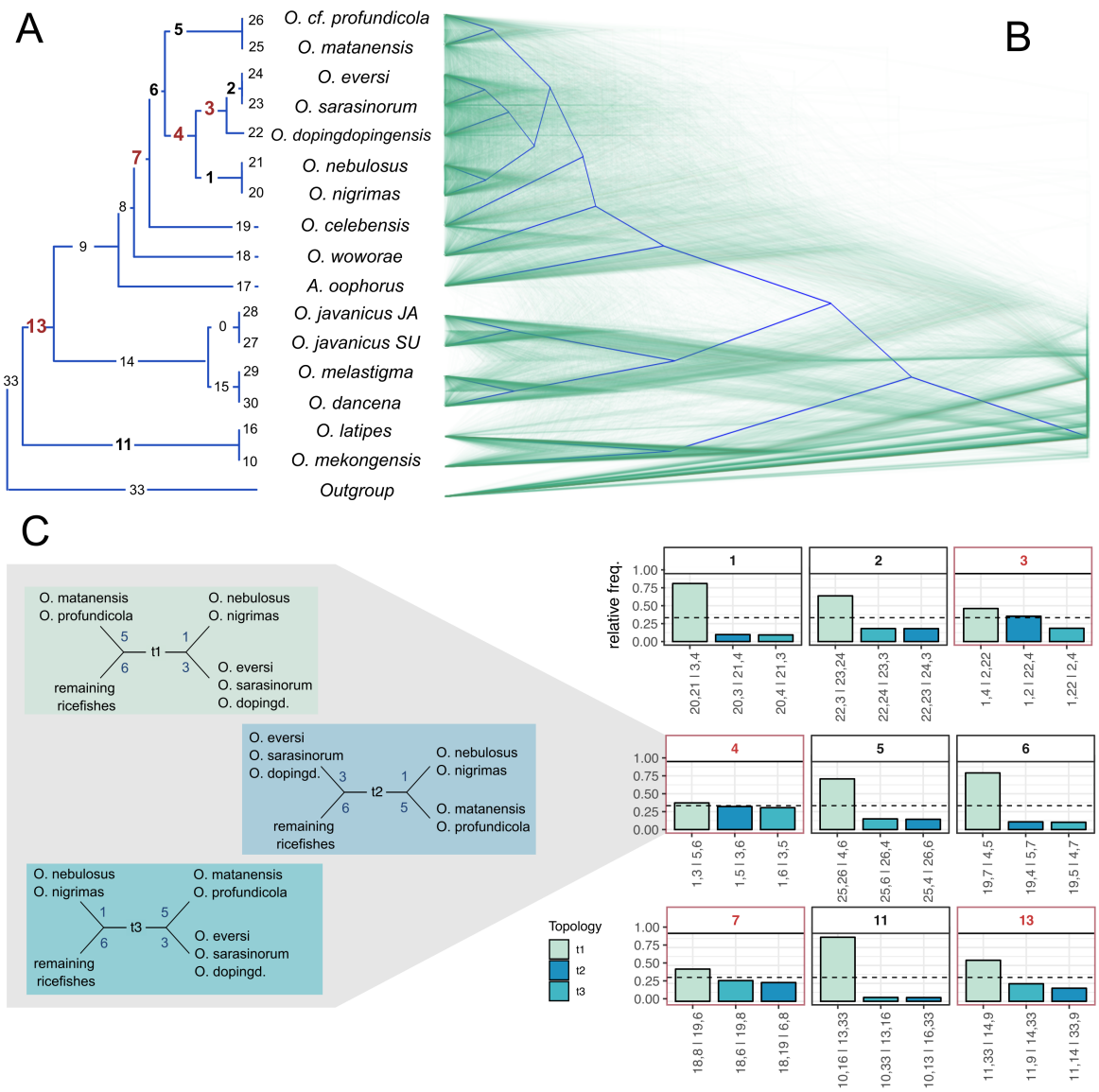


Fig 2: A) species tree from Astral created using Discovista. Numbers in black mean that ~100% of all gene trees follow the shown topology. Bold numbers refer to quartet scores with >60% of all gene trees following the shown topology. Red numbers show splits with the second most frequent topology found in >25% of all gene trees. B) Densitree based on 1907 single gene trees. Blurry areas indicate gene tree – species tree incongruencies. C) Barplots representing percentage of alternative topologies in corresponding nodes (indicated in red in A). Split 4 is highlighted as example depicting the three different possible topologies.

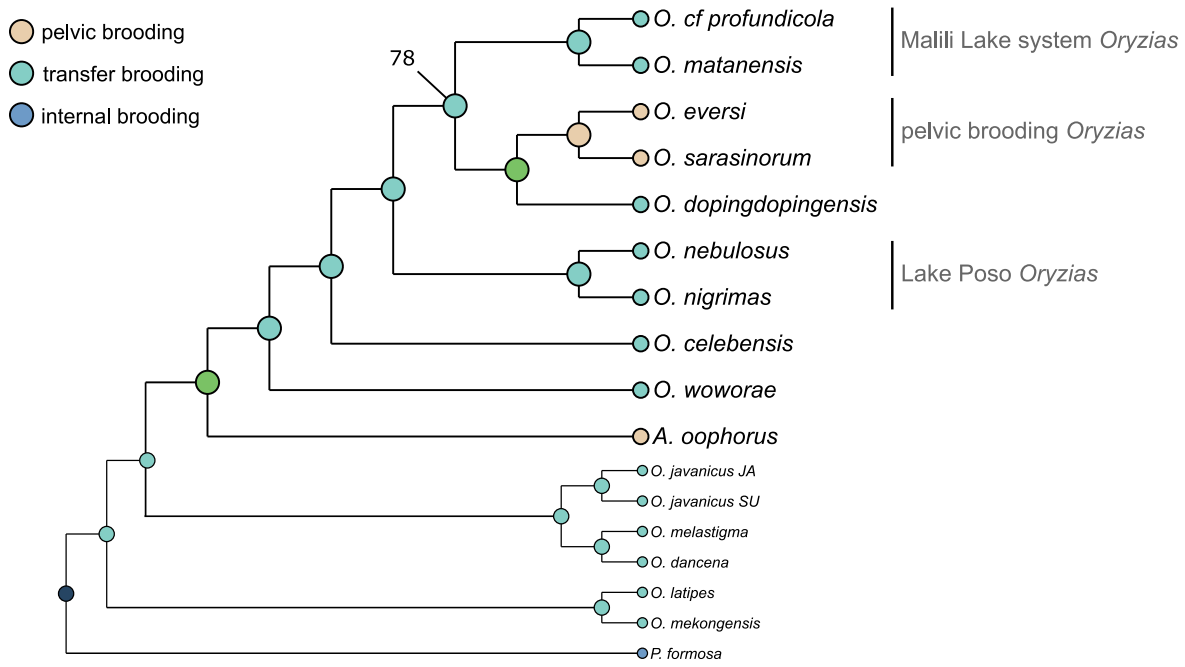


Fig 3: Maximum likelihood tree based on a supermatrix of 1907 concatenated single-copy protein-coding orthologous genes and ancestral state reconstruction. All splits had a bootstrap support of 100, except for the one depicted. Interrelationships between Malili Lake system *Oryzias* and pelvic-brooding *Oryzias* were not well resolved based on our data, indicating that other topologies are also supported (as shown in Figure 2, node 4).

Detected introgression events in the younger Oryzias clades

We found no introgression among the pelvic-brooding species *A. oophorus*, *O. eversi* and *O. sarasinorum* (Fig. 4), but between *O. marmoratus* from Mahalona and *O. matanensis* from Lake Matano (Fig. 4). Furthermore, we found a signal of hybridization between *O. sarasinorum* and Lake Poso species *O. nebulosus*, *O. nigrimas*, *O. orthognathus* and *O. soerotoi* as well as between *O. eversi* and the same Lake Poso species (Fig. 4). The strongest signals of introgression were located on chromosomes 9, 12_20_13 and 21 (on average >1 window with a f_d and f_{dM} belonging to the highgest 5% within 30 windows) of the *O. celebensis* reference genome (OCchr) (Fig. S4). Moderatley high signal of introgression was found on OCchr 11, 16 and 24 (on agerage 1 window with a f_d and f_{dM} belonging to the highgest 5% within 40 windows). For the confidence interval for QCon1 on OCchr 1_17_19 (Tab. S11), a locus associated with the concavity, in 1.46% for f_d and 0.41% for f_{dM} of the permutations, an equal or higher number of top 1% D-values was found. For the confidence interal of QEgg1, in 9.52% for f_d of the permutations a equal or more negative D-value was found. For QFin on OCchr 24, a proposed QTL was 16 windows away from a window with high introgression single (top 1%). A top 1% f_d value within a range of 15 windows from the proposed location of the QTL was found in 2.73% of the permutations. The other QTLs were on chromosomes with low introgression signal and could not be associated with a high introgression signal (Fig. S4, Tab. S11). The results of gene tree analyses with SNaQ were congruent with results of the D-statistic: all indicated introgression between the Lake Poso *Oryzias* and the pelvic brooders *O. eversi* and *O. sarasinorum* (Fig. 4, Fig. S2). Further, we observed a hybrid edge between the ancestor of *O. matanensis* and *O.* cf. *profundicola*. This was in line with the ambiguous node 4 in the quartet tree of Astral (Fig. 2) and the moderately supported node in the ML tree (Fig. 3). We also observed a hybrid edge between *O. dancena* and the ancestor of *O. celebensis* and the younger *Oryzias* lineages. However, the amount of similarity was rather low with 0.46% (Fig. 4). The genes with the highest D-values were found on OCchr 1_17_19, 8 and 24.

Using TreeMix (Fig. S3), we identified two introgression events supported with a light migration weight between *A. oophorus* and *O. melastigma* and *O. javanicus* and the most recent common ancestor of the pelvic-brooding *Oryzias* and the Malili Lake system *Oryzias*. With 2 edges (two introgression events), a migration event between the ancestor of the Poso ricefishes and the ancestor of the two *Oryzias* pelvic brooders was suggested, and also had a high migration weight (red arrow). With four edges, the migration event between *O. matanensis* and *O. marmoratus* "Mahalona" was observed. With a fifth edge, another event within the Lake Malili system *Oryzias* was hypothesized.

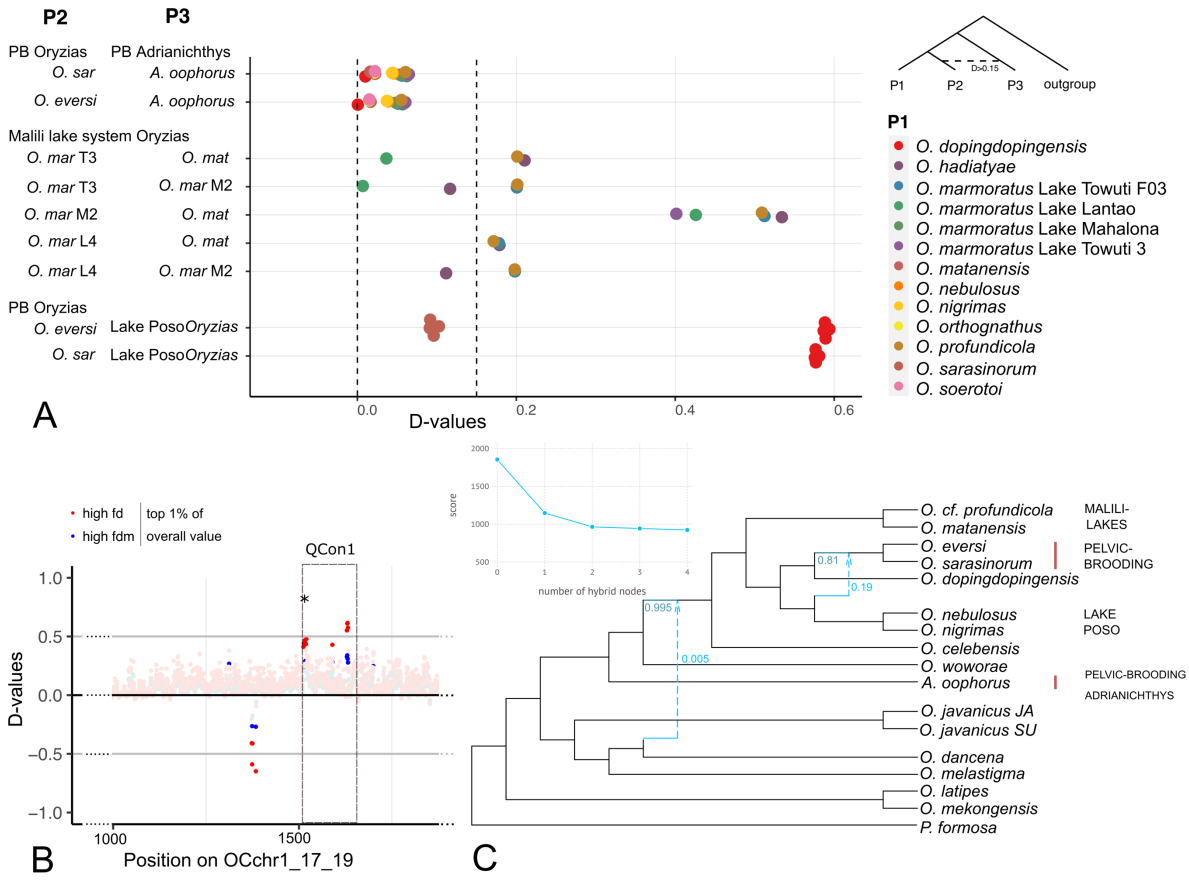


Fig 4: Results from Dsuite and SNaQ. In A) the Dsuite plot, the Lake Poso *Oryzias* refer to *O. nigrimas*, *O. nebulosus*, *O. orthognathus* and *O. soerotoi*. For the Malili lake *Oryzias*, O. mar T3 is *O. marmoratus* from Towuti, M2 from Mahalona, L4 from Lantoa. B) shows the region on the *O. celebensis* reference genome chr 1_17_19 with high introgression signal (top 1% of all D-values in color) overlapping with a QTL for the extent of the ventral concavity in *O. eversi*. Both A) Dsuite and C) SNaQ congruently show hybridization between ancestor of Lake Poso *Oryzias* and ancestor of pelvic-brooding *Oryzias*. In the SNaQ tree, another hybrid edge is indicated between *O. dancena* and *O. celebensis* and more recent lineages, though the proportion of genes inherited by this hybrid node from one of its hybrid parents is less than 0.5%.

Discussion

No Hybridization between *Adrianichthys* and *Oryzias*, but abundant gene-flow within *Oryzias* from Sulawesi

After analyzing 1907 single-copy protein-coding genes and ~ 38 mio genomic SNPs, we did not detect any sign of direct gene-flow between the pelvic-brooding lineages, which is in line with previously published results (Montenegro et al. 2022). Slightly different phenotypic expression of pelvic-brooding traits and a differing extent of sexual dimorphism between species of the two pelvic-brooding lineages, *Adrianichthys* and *Oryzias* (Spanke et al. 2021) provide additional support for an independent origin of pelvic-brooding in each lineage. We found, however, gene flow from Lake Poso *Oryzias* into the pelvic-brooding *Oryzias* lineage. Several ancient and recent gene flow events occurred, e.g., locally within lake systems as was shown in the Malili lake system including *O. marmoratus*, *O. matanensis* and *O. profundicola* (Mandagi et al. 2021). Gene flow over large spatial distances was observed in *O. sarasinorum* and *O. eversi* which nowadays occur 190km apart from each other (Horoiwa et al. 2021). Gene flow over both, large spatial as well as phylogenetic distance, was detected between *O. soerotoi* in Lake Tiu and the ancestor of *O. nebulosus* and *O. orthognathus* in Lake Poso (Horoiwa et al. 2021). All this evidence indicates that ricefish evolution is better described in a network-like evolutionary structure (Fig. 2), rather than by a simple bifurcating tree.

The quasi-independent origin of a complex reproductive strategy

The QTLs of two traits associated with pelvic-brooding (QCon1 and QFin; Montenegro et al. 2022) lie within regions with high D-values, suggesting that introduced genetic variants from the Lake Poso *Oryzias* might have played a role in the evolution of pelvic-brooding (Fig. S4). However, for two other QTLs the underlying genetic region seemed to have a different genotype compared to the Lake Poso ricefishes and *O. dopingdopingensis* (Qcon2 and QEgg1, expressed by highly negative D-values) and likely evolved *de novo* in the pelvic-brooding *Oryzias*. The introgression event occurred after the splitting of *O. dopingdopingensis* and the

pelvic-brooding Oryzias (~1.4–1.8 mya based on Ansai et al., 2021; Mokodongan & Yamahira, 2015). Hence, a potential gene flow event between the Lake Poso and pelvic-brooding stem lineage dates around the origin of Lake Poso, which is assumed to be not older than 1–2 mya (Rintelen et al. 2004; von Rintelen & Glaubrecht 2006). Intriguingly, all *Adrianichthys* species are endemic to Lake Poso, which would allow an indirect transmission of pelvic-brooding alleles into the stem lineage of the pelvic-brooding *Oryzias*. This assumption is contradicted by the lack of gene flow detected between *Adrianichthys* and any of the *Oryzias* species, though old hybridization events might be hard to detect. Based on our data, we found that a small percentage (0.46%; Fig.S1) of the *O. dancena* genome has introgressed into the ancestor of *O. celebensis* and the younger *Oryzias* lineages hinting towards that older hybridization events occurred between different ricefish lineages. Ancient hybridization events have the potential to serve as resource for adaptive genetic variation. Available ecological niche space allows selection to act upon genetic diversity originating from differently combined old alleles, generating adapted and divergent phenotypes (Seehausen 2013, 2004). For example, in the apple maggot *Rhagoletis pomonella*, introgressed genetic regions were found, deriving from an ancient hybridization with Mexican Altiplano highland fruit flies about 1.6 million years ago. However, the emergence of new species using this old variation to adapt to new hosts happened more recently (Feder et al. 2003). Under this combinatorial view, reassembling old variants into new combinations promotes rapid speciation and adaptive radiation (reviewed in Marques et al. 2019). To further disentangle the genomic background of pelvic-brooding in the two distantly related ricefish lineages, in-depth comparative genomic studies are needed.

Habitat dependence – why did pelvic-brooding evolve twice?

Irrespective of the genetic background, strong selective pressures must have shaped the evolution from ancestral transfer-brooding to the extended care (of pelvic-brooding), which is presumably costly in terms of increased predation risk and increased energetic costs of the care-giver (Wootton & Smith 2014; Cooke et al. 2006). Especially when assuming convergent

evolution the role of the environment was proven to be strong for the adaptive value of replicated trait differentiation (Losos et al. 1998; Gompel & Prud'homme 2009; Arendt & Reznick 2008). Present day macro-habitats of pelvic-brooding ricefishes range from lakes to small karst-ponds (reviewed in Parenti, 2008; e.g. Herder et al., 2012; Mandagi, Mokodongan, Tanaka, & Yamahira, 2018); in Lake Poso, pelvic-brooding *Adrianichthys* mainly occupy open-water habitats, leading to the hypothesis that it evolved in adaptation to the absence of suitable spawning substrates in pelagic habitats (Herder et al. 2012). *Oryzias eversi*, however, was described from a small karst pond where potential spawning substrates are abundant (Herder et al. 2012). Thus, present day habitats make it rather complicated to define a general selective regime that might have favoured the evolution of pelvic-brooding. Furthermore, the paleogeographical history of Sulawesis water bodies is barely known and complicated (Wilson & Moss 1999; Hall 2001), leaving plenty of room for speculations, like presumable ancient connections between waterbodies or the existence of a paleolake in central Sulawesi (Utama et al. 2022). Thus, even an ancestral syntopic distribution of *A. oophorus*, the ancestor of *O. eversi* and *O. sarasinorum* and ancestors of the lake Poso *Oryzias* seems possible. Given the costs and direct fitness effects of switching reproductive strategies and the rather fast convergent evolution of pelvic-brooding in two distinct lineages it appears likely that similar environmental and selection pressures in combination with gene flow fueled its evolution.

Conclusions

Several gene flow events within Sulawesi ricefishes indicate that the Sulawesi ricefish radiation did not follow a tree-like evolution. In this study, we found no introgression between the two distantly related pelvic-brooding lineages, but detected gene flow into the stem lineage of pelvic-brooding *Oryzias* from Lake Poso *Oryzias*. The presence of elevated D-values within the confidence intervals of QTLs for pelvic fin length and the ventral concavity raise the possibility that genetic variants were recruited for the evolution of pelvic-brooding via hybridziation with a closely related species.

Material & Methods

Taxon sampling

Fish were collected in Sulawesi between 2011 and 2013 and some species were bred in the aquarium until they were sacrificed (for details see supplementary Tab. S1). We sequenced the transcriptomes of twelve ricefish species (*Adrianichthys oophorus*, *Oryzias mekongensis*, *O. dancena*, *O.* cf. *profundicola*, *O. nebulosus*, *O. javanicus*, *O. sarasinorum*, *O. eversi*, *O. nigrimas*, *O. woworae*, *O. celebensis*, *O. matanensis*) and the genome of *O. dopingdopingensis*. Furthermore, a published genome of *O. javanicus* (GenBank Bioproject: PRJNA505405, Biosample: SAMN10417210) and the official gene sets (OGS, used reference genomes in Tab. S6) of two ricefish species (*O. latipes* and *O. melastigma*) and four outgroup species (*Poecilia formosa*, *Xiphophorus maculatus*, *Nothobranchius furzeri* and *Austrofundulus limaneus*) derived from the orthologous gene sets of Actinopterygii (TaxID 7898) from OrthoDB (Tab. S1) were included in our analyses.

Sequencing and orthology prediction

We extracted RNA from complete fish and prepared non-stranded Truseq mRNA libraries which were sequenced on an Illumina Hiseq2000 at the CCG in Cologne (we provide a shortened version of material and methods here, details about all used methods are in the supplement). We used the Qiagen DNeasy Blood &Tissue Kit to extract DNA of *O. dopingdopingensis*. A genomic short read TruSeq DNA PCR free library was prepared by Macrogen sequencing company; finally 15,698,917,340 bases and 103,966,340 reads were sequenced.

The raw transcriptomic reads were quality-filtered using trim-fast.pl from the PoPoolation pipeline (Kofler et al. 2011) and assembled using Trinity v2.8.4. (Grabherr et al. 2011; Haas et al. 2013). A *de novo* whole-genome assembly of *O. dopingdopingensis* short reads was generated according to Böhne et al. (2019) and Malmstrøm, Matschiner, Tørresen, Jakobsen, & Jentoft (2017) and annotated using gene predictions of *O. latipes* and *O.*

melastigma, a protein set from *O. javanicus* and a set of orthologous genes generated in this study. Busco completeness scores of the transcriptomes ranged from 60.7% to 80.9% (Actinopterygii Odb9) (Tab. S3) and the annotated genome assembly had 97% complete Busco completeness score.

We created the reference ortholog set using the orthologous gene set of Actinopterygii (TaxID 7898) downloaded from OrthoDB (Waterhouse et al. 2013). We used Orthograph v0.7.1 (Petersen et al. 2017) to find orthologous genes in our transcriptome assemblies and the genome assembly. The orthologous genes were aligned using MAFFT v7.221 with the L-INS-I algorithm (Katoh & Standley 2013) on the amino-acid level and outliers in genes were identified and the respective genes removed according to Misof et al. (2014). We created nucleotide alignments with the amino acid alignments as a control with a modified version of Pal2Nal v14 (Suyama et al. 2006; Misof et al. 2014) and masked ambiguously aligned regions (maximum number of pairwise sequence comparisons for each MSA (option -r), and the -e option for gap-rich data sets, leaving remaining settings to defaults) and subsequently removed all MSA sections, which were indicated as randomly similar aligned with ALISCORE v2.0 and ALICUT v2.3 (Misof & Misof 2009; Kück et al. 2010; Misof et al. 2014).

Phylogenomic analyses

We used IQ-TREE v1.6.12 to create trees for every gene separately (Nguyen et al. 2014) (for details on the methods described in this paragraph, see supplement). For each gene, a substitution model was estimated (Tab. S9). We used Astral v5.7.3 to run a multi-species coalescent model (Zhang et al. 2018). The gene trees were tested with Pyhlotreepruner (Kocot et al. 2013) and TreeShrink (Mai & Mirarab 2018). Phylotreepruner checks for paralogous sequences and TreeShrink for extremely long branches. The results of Astral were displayed using DiscoVista (Sayyari et al. 2018). The gene trees were made ultrametric using the function "chronos" in R v4.1.2 (package "ape" v5.6-2 (Paradis & Schliep 2019) to use them for Densitree2.01 (Bouckaert & Heled 2014).

The genes were ordered according to their position on the reference genome of *O. latipes* (RefSeq GCF_002234675.1, Bioprject PRJNA325079) and in this order concatenated using FASconCAT-G (Kück & Longo 2014). The resulting supermatrix was partitioned and for each partition a suitable model was searched in IQ-TREE, taking codon position into account. We ran 20 single tree searches and 50 non-parametric, slow bootstrap replicates.

Finally, we did an ancestral state reconstruction in RASP4.2 (Yu et al. 2020). We assigned a state of pelvic-brooding to *O. eversi*, *O. sarasinorum* and *A. oophorus*, a state of transfer brooding to the other ricefishes and a state of live-bearing to the outgroup *poecilia formosa*. We used the ML tree as a basis and ran the S-DIVA model.

Analyses of gene-flow pattern based on genome-wide SNPs and orthologous genes

We downloaded whole genome data from GenBank of 26 individuals of 19 ricefish species including four specimen of *O. marmoratus*, three *O. wolasi* specimen, and four *O. woworae* specimen from different localities (Bioproject PRJDB10385), *O. javanicus* (Bioproject PRJNA505405, accession number SRR8467745) and *O. melastigma* (Bioproject PRJNA556761, accession number SRR12442554) and mapped them on the *O. celebensis* reference genome from Genbank (GCA_014656515.1, Bioproject PRJDB10371, Ansai et al., 2021) (methods used in this paragraph are described in detail in the supplement). Filtering resulted in 38183142 SNPs and we created a vcf file which we used for Dsuite v0.4r41, a program to calculate D-statistics based on a vcf file (Malinsky et al. 2021). To locate regions of elevated D-values, we used DInvestigate, a follow-up analysis included in Dsuite. It calculates different D-statistics for windows across the genome. We left window size and steps at default. We used f_d and f_{dM}, two statistics specifically designed to detect introgressed loci. Where f_d is distributed on the interval [- Infinity, 1] (Martin et al. 2015), f_{dM} is distributed on the inveral [-1,1] under the null hypothesis of no introgression is symmetrically distributed around zero (Malinsky et al. 2015). To track down chromosomes with the strongest signal of introgression, windows with top 5% (Fig. S4) values for genome-wide f_d and f_{dM} were counted

and corrected for chromosome length. Additionally, we checked whether windows with the top 5% and top 1% of genome-wide f_d and f_{dM} values fall within the confidence intervals of QTLs associated with pelvic-brooding (Fig. S5). To test if D-values were significantly higher (if positive) or lower (if negative) than expected by chance in the QTL intervals, we distributed the windows with highest D-values (top 1%) randomly over the whole genome 10'000 times and checked, how often we find the same or higher D-values compared to the real QTL intervals. Further we used Treemix v1.13, a graph-model based program, to asses hybrid edges based on the SNP dataset with all missing data removed (Pickrell & Pritchard 2012). To check for signatures of hybridization within the orthologous genes, we used SNaQ which additionally takes different evolutionary rates of branches into account (Solís-Lemus et al. 2017; Solís-Lemus & Ané 2016). We used HybridCheck to find elevated D-values in our orthologous gene set (Ward & van Oosterhout 2016). The regions of elevated D-values were compared to the quantitative trait loci (QTLs) associated with pelvic fin length, the extent of the concavity and duration of egg carrying found in (Montenegro et al. 2022).

Acknowledgements

We thank Rob Waterhouse for the advice on using OrthoDB. Further, we are thankful for the help of Alexandros Vasilikopoulos with compiling the ortholog set. We thank Michael Matschiner and Thore Koppetsch for their insights in the pitfalls of different evolutionary rates in hybridization analysis and their recommendations. We are very grateful for the permission to use beautiful photos taken by Jan Möhring, Andreas Wagnitz and Hans Evers for this publication. Calculations for the genome assembly were performed at sciCORE (http://scicore.unibas.ch/) scientific computing center at University of Basel. Further, we used the annotation scripts from the Sigenae platform for the genome annotation. This work was supported by the Leibniz Association, grant P91/ 2016.

References

Ansai S et al. 2021. Genome editing reveals fitness effects of a gene for sexual dichromatism in Sulawesian fishes. Nat. Commun. 12:1350. doi: 10.1038/s41467-021-21697-0.
Arendt J, Reznick D. 2008. Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? Trends Ecol. Evol. 23:26–32. doi: 10.1016/j.tree.2007.09.011.
Arnold ML. 2007. *Evolution through genetic exchange*. Oxford University Press.
Baack EJ, Rieseberg LH. 2007. A genomic view of introgression and hybrid speciation. Curr. Opin. Genet. Dev. 17. doi: 10.1016/j.gde.2007.09.001.
Barrett RDH, Schluter D. 2008. Adaptation from standing genetic variation. Trends Ecol. Evol. 23:38–44. doi: 10.1016/j.tree.2007.09.008.
Van Belleghem SM et al. 2018. Evolution at two time frames: polymorphisms from an ancient singular divergence event fuel contemporary parallel evolution Schierup, MH, editor. PLOS Genet. 14:e1007796. doi: 10.1371/journal.pgen.1007796.
Böhne A et al. 2019. Repeated evolution versus common ancestry: sex chromosome evolution in the haplochromine cichlid Pseudocrenilabrus philander A Katz, L, editor. Genome Biol. Evol. 11:439–458. doi: 10.1093/gbe/evz003.
Bouckaert RR, Heled J. 2014. DensiTree 2: seeing trees through the forest. bioRxiv. 012401. doi: 10.1101/012401.
Brawand D et al. 2014. The genomic substrate for adaptive radiation in African cichlid fish. Nature. 513:375–381. doi: 10.1038/nature13726.
Castric V, Bechsgaard J, Schierup MH, Vekemans X. 2008. Repeated adaptive introgression at a gene under multiallelic balancing selection. PLoS Genet. 4. doi: 10.1371/journal.pgen.1000168.
Colosimo PF et al. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. Science. 307:1928–33. doi: 10.1126/science.1107239.
Cooke SJ, Philipp DP, Wahl DH, Weatherhead PJ. 2006. Energetics of parental care in six syntopic centrarchid fishes. Oecologia. 148:235–249. doi: 10.1007/s00442-006-0375-6.
Darwin C. 1859. *The origin of species by means of natural selection*. John Murray: London.
Edelman NB et al. 2019. Genomic architecture and introgression shape a butterfly radiation. Science. 366:594–599.
Elmer KR, Meyer A. 2011. Adaptation in the age of ecological genomics: insights from parallelism and convergence. Trends Ecol. Evol. 26:298–306. doi: 10.1016/j.tree.2011.02.008.
Feder JL et al. 2003. Allopatric genetic origins for sympatric host-plant shifts and race formation in *Rhagoletis*. Proc. Natl. Acad. Sci. 100:10314–10319. doi: 10.1073/pnas.1730757100.
Gompel N, Prud'homme B. 2009. The causes of repeated genetic evolution. Dev. Biol. 332:36–47. doi: 10.1016/J.YDBIO.2009.04.040.
Grabherr MG et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. doi: 10.1038/nbt.1883.
Haas BJ et al. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. doi: 10.1038/nprot.2013.084.
Hall R. 2001. Cenozoic reconstructions of SE Asia and the SW Pacific: changing patterns of land and sea. Faunal Flor. Migr. Evol. SE Asia–Australasia. 35–56.
Herder F, Hadiaty RK, Nolte AW. 2012. Pelvic-fin brooding in a new species of riverine ricefish (Atherinomorpha: Beloniformes: Adrianichthyidae) from Tana Toraja, Central Sulawesi, Indonesia. Raffles Bull. Zool. 60:467–476.
Hermisson J, Pennings PS. 2005. Soft sweeps: molecular population genetics of adaptation

from standing genetic variation. Genetics. 169:2335–2352. doi: 10.1534/genetics.104.036947.
Hilgers L, Schwarzer J. 2019. The untapped potential of medaka and its wild relatives. Elife. 8:1–14. doi: 10.7554/eLife.46994.
Horoiwa M et al. 2021. Mitochondrial introgression by ancient admixture between two distant lacustrine fishes in Sulawesi Island. PLoS One. 16:1–14. doi: 10.1371/journal.pone.0245316.
Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30:772–780. doi: 10.1093/molbev/mst010.
Kocot KM, Citarella MR, Moroz LL, Halanych KM. 2013. PhyloTreePruner: a phylogenetic tree-based approach for selection of orthologous sequences for phylogenomics. Evol. Bioinform. Online. 9:429–35. doi: 10.4137/EBO.S12813.
Kofler R et al. 2011. PoPoolation: a toolbox for population genetic analysis of next generation sequencing data from pooled individuals Kayser, M, editor. PLoS One. 6:e15925. doi: 10.1371/journal.pone.0015925.
Kottelat M. 1990. Synopsis of the endangered Buntingi (Osteichthyes: Adranichtyidae and Oryziidae) of Lake Poso, Central Sulawesi, Indonesia, with a new reproductive guild and descriptions of three new species. Ichthyol. Explor. Freshwaters. 1:46–67.
Kück P et al. 2010. Parametric and non-parametric masking of randomness in sequence alignments can be improved and leads to better resolved trees. Front. Zool. 7:10. doi: 10.1186/1742-9994-7-10.
Kück P, Longo GC. 2014. FASconCAT-G: extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. Front. Zool. doi: 10.1186/s12983-014-0081-x.
Lamichhaney S et al. 2015. Evolution of Darwin's finches and their beaks revealed by genome sequencing. Nature. 518:371–375. doi: 10.1038/nature14181.
Losos JB, Jackman TR, Larson A, Queiroz K de, Rodríguez-Schettino L. 1998. Contingency and determinism in replicated adaptive radiations of island lizards. Science. 279:2115–2118. doi: 10.1126/science.279.5359.2115.
Mai U, Mirarab S. 2018. TreeShrink: fast and accurate detection of outlier long branches in collections of phylogenetic trees. BMC Genomics. 19:272. doi: 10.1186/s12864-018-4620-2.
Malinsky M et al. 2015. Genomic islands of speciation separate cichlid ecomorphs in an East African crater lake. Science. 350:1493–1498. doi: 10.1126/science.aac9927.
Malinsky M, Matschiner M, Svardal H. 2021. Dsuite - Fast D -statistics and related admixture evidence from VCF files. Mol. Ecol. Resour. 21:584–595. doi: 10.1111/1755-0998.13265.
Mallet J. 2005. Hybridization as an invasion of the genome. Trends Ecol. Evol. 20. doi: 10.1016/j.tree.2005.02.010.
Malmstrøm M, Matschiner M, Tørresen OK, Jakobsen KS, Jentoft S. 2017. Whole genome sequencing data and de novo draft assemblies for 66 teleost species. Sci. Data. 4:160132. doi: 10.1038/sdata.2016.132.
Mandagi IF et al. 2021. Species divergence and repeated ancient hybridization in a Sulawesian lake system. J. Evol. Biol. 1–14. doi: 10.1111/jeb.13932.
Mandagi IF, Mokodongan DF, Tanaka R, Yamahira K. 2018. A new riverine ricefish of the genus *Oryzias* (Beloniformes, Adrianichthyidae) from Malili, Central Sulawesi, Indonesia. Copeia. 106:297–304. doi: 10.1643/ci-17-704.
Marques DA, Meier JI, Seehausen O. 2019. A combinatorial view on speciation and adaptive radiation. Trends Ecol. Evol. 34:531–544. doi: 10.1016/J.TREE.2019.02.008.
Martin SH, Davey JW, Jiggins CD. 2015. Evaluating the use of ABBA–BABA statistics to locate introgressed loci. Mol. Biol. Evol. 32:244–257. doi: 10.1093/molbev/msu269.
Meier JI et al. 2017. Ancient hybridization fuels rapid cichlid fish adaptive radiations. Nat. Commun. 8:1–11. doi: 10.1038/ncomms14363.
Misof B et al. 2014. Phylogenomics resolves the timing and pattern of insect evolution.

Science. 346:763–767. doi: 10.1126/science.1257570.
Misof B, Misof K. 2009. A Monte Carlo approach successfully identifies randomness in multiple sequence alignments: a more objective means of data exclusion. Syst. Biol. 58:21–34. doi: 10.1093/sysbio/syp006.
Mokodongan DF, Yamahira K. 2015. Origin and intra-island diversification of Sulawesi endemic Adrianichthyidae. Mol. Phylogenet. Evol. 93:150–160. doi: 10.1016/j.ympev.2015.07.024.
Montenegro J et al. 2022. Genetic basis for the evolution of pelvic-fin brooding, a new mode of reproduction, in a Sulawesian fish. Mol. Ecol. doi: 10.1111/mec.16555.
Nelson TC, Cresko WA. 2018. Ancient genomic variation underlies repeated ecological adaptation in young stickleback populations. Evol. Lett. 2:9–21. doi: 10.1002/evl3.37.
Nguyen L-T, Schmidt HA, von Haeseler A, Quang Minh B. 2014. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32:268–274. doi: 10.1093/molbev/msu300.
Ohta T. 1992. The nearly neutral theory of molecular evolution. Annu. Rev. Ecol. Syst. 23:263–286. doi: 10.1146/annurev.es.23.110192.001403.
Paradis E, Schliep K. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R Schwartz, R, editor. Bioinformatics. 35:526–528. doi: 10.1093/bioinformatics/bty633.
Parenti LR. 2008. A phylogenetic analysis and taxonomic revision of rice fishes, *Oryzias* and relatives (Beloniformes, Adrianichthydae). Zool. J. Linn. Soc. 154:494–610.
Petersen M et al. 2017. Orthograph: a versatile tool for mapping coding nucleotide sequences to clusters of orthologous genes. BMC Bioinformatics. 18:1–10. doi: 10.1186/s12859-017-1529-8.
Pickrell JK, Pritchard JK. 2012. Inference of population splits and mixtures from genome-wide allele frequency data Tang, H, editor. PLoS Genet. 8:e1002967. doi: 10.1371/journal.pgen.1002967.
Rieseberg LH. 2009. Evolution: replacing genes and traits through hybridization. Curr. Biol. 19. doi: 10.1016/j.cub.2008.12.016.
von Rintelen T, Glaubrecht M. 2006. Rapid evolution of sessility in an endemic species flock of the freshwater bivalve *Corbicula* from ancient lakes on Sulawesi, Indonesia. Biol. Lett. 2:73–77. doi: 10.1098/rsbl.2005.0410.
Rintelen T von, Wilson AB, Meyer A, Glaubrecht M. 2004. Escalation and trophic specialization drive adaptive radiation of freshwater gastropods in ancient lakes on Sulawesi, Indonesia. Proc. R. Soc. London. Ser. B Biol. Sci. 271:2541–2549. doi: 10.1098/rspb.2004.2842.
Sayyari E, Whitfield JB, Mirarab S. 2018. DiscoVista: interpretable visualizations of gene tree discordance. Mol. Phylogenet. Evol. 122:110–115. doi: 10.1016/J.YMPEV.2018.01.019.
Schluter D, Nagel LM. 1995. Parallel speciation by natural selection. Am. Nat. 146:292–301. doi: 10.1086/285799.
Seehausen O. 2013. Conditions when hybridization might predispose populations for adaptive radiation. J. Evol. Biol. 26:279–281. doi: 10.1111/jeb.12026.
Seehausen O. 2004. Hybridization and adaptive radiation. Trends Ecol. Evol. 19:198–207. doi: 10.1016/j.tree.2004.01.003.
Solís-Lemus C, Ané C. 2016. Inferring phylogenetic networks with maximum pseudolikelihood under incomplete lineage sorting. PLOS Genet. 12:e1005896. doi: 10.1371/journal.pgen.1005896.
Solís-Lemus C, Bastide P, Ané C. 2017. PhyloNetworks: a package for phylogenetic networks. Mol. Biol. Evol. 34:3292–3298. doi: 10.1093/molbev/msx235.
Spanke T et al. 2021. Complex sexually dimorphic traits shape the parallel evolution of a novel reproductive strategy in Sulawesi ricefishes (Adrianichthyidae). BMC Ecol. Evol.

21:57. doi: 10.1186/s12862-021-01791-z.
Sutra N et al. 2019. Evidence for sympatric speciation in a Wallacean ancient lake. Evolution (N. Y). 73:1898–1915. doi: 10.1111/evo.13821.
Suyama M, Torrents D, Bork P. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. doi: 10.1093/nar/gkl315.
Terekhanova N V. et al. 2014. Fast evolution from precast bricks: genomics of young freshwater populations of threespine stickleback *Gasterosteus aculeatus* Peichel, CL, editor. PLoS Genet. 10:e1004696. doi: 10.1371/journal.pgen.1004696.
The Heliconius Genome Consortium et al. 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. Nature. 487:94. doi: 10.1038/NATURE11041.
Utama I V. et al. 2022. Deeply divergent freshwater fish species within a single river system in central Sulawesi. Mol. Phylogenet. Evol. 173:107519. doi: 10.1016/J.YMPEV.2022.107519.
van der Valk T et al. 2021. Million-year-old DNA sheds light on the genomic history of mammoths. Nature. 591:265–269. doi: 10.1038/s41586-021-03224-9.
Veale AJ, Russello MA. 2017. Genomic changes associated with reproductive and migratory ecotypes in sockeye salmon (*Oncorhynchus nerka*). Genome Biol. Evol. 9:2921–2939. doi: 10.1093/gbe/evx215.
Ward BJ, van Oosterhout C. 2016. HybridCheck : software for the rapid detection, visualization and dating of recombinant regions in genome sequence data. Mol. Ecol. Resour. 16:534–539. doi: 10.1111/1755-0998.12469.
Waterhouse RM, Tegenfeldt F, Li J, Zdobnov EM, Kriventseva E V. 2013. OrthoDB: a hierarchical catalog of animal, fungal and bacterial orthologs. Nucleic Acids Res. 41:D358–D365. doi: 10.1093/nar/gks1116.
Waters JM, McCulloch GA. 2021. Reinventing the wheel? reassessing the roles of gene flow, sorting and convergence in repeated evolution. Mol. Ecol. 30:4162–4172. doi: 10.1111/mec.16018.
Whitney, Randell, Rieseberg. 2006. Adaptive introgression of herbivore resistance traits in the weedy sunflower *Helianthus annuus*. Am. Nat. 167. doi: 10.2307/3844738.
Wilson MEJ, Moss SJ. 1999. Cenozoic palaeogeographic evolution of Sulawesi and Borneo. Palaeogeogr. Palaeoclimatol. Palaeoecol. 145:303–337. doi: 10.1016/S0031-0182(98)00127-8.
Wootton R, Smith C. 2014. *Reproductive biology of teleost fishes*. John Wiley & Sons.
Yamamoto T. 1975. *Medaka: (Killifish). Biology and strains. Series of stock culture in biological field*. Keigaku Publishing Company: Tokyo.
Yu Y, Blair C, He X. 2020. RASP 4: ancestral state reconstruction tool for multiple genes and characters Yoder, A, editor. Mol. Biol. Evol. 37:604–606. doi: 10.1093/molbev/msz257.
Zhang C, Rabiee M, Sayyari E, Mirarab S. 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics. 19:153. doi: 10.1186/s12859-018-2129-y.

Data Availability Statement:

Genetic data

Raw sequence reads are deposited in the SRA (see supplementary material, Tab. S2).

Genome annotation data is available on DataDryad (link will be added after acceptance).

Sample metadata

Related metadata can be found in supplementary material (Tab. S1).

Benefit-Sharing Statement:

Benefits Generated: A research collaboration was developed with scientists from the countries providing genetic samples, all collaborators are included as co-authors, the results of research have been shared with the provider communities and the broader scientific community (see above). More broadly, our group is committed to international scientific partnerships, as well as share knowledge about the establishing and maintaining of scientific collections.

Author contributions:

JMF, JS, KM designed research. AWN and FH were leading the field expeditions in Sulawesi where the samples were collected. AWN provided the transcriptome sequences. LH did the transcriptome assemblies. JMF and KM performed research and analyzed data. AB did the genome assembly. SM annotated the genome and supported JMF with the NCBI submission process. JMF and JS wrote initial draft of the manuscript. All authors contributed to writing the final manuscript.

Supplement – directory

Additional tables

Additional figures

Table S2: Genbank accession numbers of used transcriptomic and genomic sequences.

Sample name	Species	Bioproject accession	Biosample accession
SN7640252_21073_A8	Adrianichthys oophorus	PRJNA827986	SAMN27626413
SN7640151_8336_Oeversi	Oryzias eversi	PRJNA827986	SAMN27626414
SN7640151_8337_Onigrimas	Oryzias nigrimas	PRJNA827986	SAMN27626415
SN7640151_8339_Owoworae	Oryzias woworae	PRJNA827986	SAMN27626416
SN7640151_8340_Ocelebensis	Oryzias celebensis	PRJNA827986	SAMN27626417
SN7640151_8341_Omatanensis	Oryzias matanensis	PRJNA827986	SAMN27626418
SN7640176_9758_Omeko	Oryzias mekongensis	PRJNA827986	SAMN27626419
SN7640252_21070_Ospec11	Oryzias sp.	PRJNA827986	SAMN27626420
SN7640252_21071_Oneb50	Oryzias nebulosus	PRJNA827986	SAMN27626421
SN7640252_21075_Oceleb121	Oryzias javanicus	PRJNA827986	SAMN27626422
SN7640252_21076_Ospec16	Oryzias sarasinorum	PRJNA827986	SAMN27626423
SN7640252_21077_Odan11	Oryzias dancena	PRJNA827986	SAMN27626424
O_dop_M_JS19_2	Oryzias dopingdopingensis	PRJNA838810	SAMN28464198

Table S3: Busco scores for transcriptomic and genomic assemblies used in Orthograph.

Sample	C	S	D	F	M	n
A8	74.6%	72.0%	2.6%	10.3%	15.1%	4584
O_celebensis	77.1%	74.8%	2.3%	8.6%	14.3%	4584
O_dan11	80.9%	78.1%	2.8%	8.5%	10.6%	4584
O_eversi	72.0%	69.9%	2.1%	12.1%	15.9%	4584
O_celeb121	60.7%	59.1%	1.6%	14.5%	24.8%	4584
O_matanensis	76.3%	73.8%	2.5%	10.5%	13.2%	4584
O_meko	80.3%	77.6%	2.7%	10.1%	9.6%	4584
O_neb50	74.7%	72.3%	2.4%	11.9%	13.4%	4584
O_nigrimas	76.7%	74.6%	2.1%	9.1%	14.2%	4584
O_spec11	70.5%	68.5%	2.0%	8.8%	20.7%	4584
O_spec16	63.1%	61.1%	2.0%	13.6%	23.3%	4584
O_woworae	73.3%	70.5%	2.8%	11.7%	15.0%	4584
O_dopingdopingensis *	90.5%	89.7%	0.8%	3.7%	5.8%	3640

*genome assembly

Table S4: Statistics by Quast about *O. dopingdopingensis* genome assembly done in Celera.

N50:	87692 bp
total length:	613401338 bp
number of contigs:	17142
largest contig:	789960 bp

Table S7: Orthograph statistics about found orthologous sequences within the transcriptomic or genomic assemblies.

Sample	Number of hits	Total number of aa	Number of stop	N50	average length	median length	max length	min length
A8	6151	2669391	21	539	433	360	3372	29
O_celeb121	5903	2125137	11	437	360	312	2638	21
O_celebensis	6095	2731016	15	561	448	367	3218	10
O_dan	6544	3041487	14	578	464	372	4639	30
O_eversi	6055	2644670	17	549	436	355	3449	30
O_javanicus_JAPAN_cds	6587	3333663	5	612	506	413	5079	31
O_matanensis	6315	2790909	28	550	441	361	2929	29
O_meko	6422	2906866	11	557	452	372	4640	30
O_neb50	6304	2721169	23	535	431	357	3148	6
O_nigrimas	6102	2686015	22	549	440	364	2898	30
O_spec11	5652	2307383	23	504	408	340	4640	12
O_spec16	5984	2180109	27	443	364	323	2664	15
O_woworae	6124	2655760	23	542	433	352	2999	30
O_dopingdopingensis_CDS	6874	3027352	13	529	440	361.5	4232	30

Table S11: results from comparison of DInvestigate results within QTL intervals (Montenegro et al. 2022) and 10'000 permutations of randomly distributed top 1% of D-values (n=162) on whole genome, to see if pattern could be observed by chance. No neg. values means that there were no high negative D-values (top 1% of most positive and most negative D-values) in the real interval, no pos. values means that there were no high positive (top 1%) D-values in the real interval. Percentages say how many percent of the permutations had the same or a higher amount of high D-values within the QTL intervals. For QFin, we tested in how many percent of the permutations we observe a high D-value within the ±15 windows interval, what was observed in the real interval.

	criteria	f_d (red)	f_dM (blue)
QCon1 chr 1_17_19	top 1% positive D-value	1.46%	0.41%
	top 1% negative D-value	no neg. values	no neg. values
QCon2 chr 6_23	top 1% positive D-value	77.42%	79.94%
	top 1% negative D-value	40.77%	no neg. values
QCon3 chr 13	top 1% positive D-value	93.12%	100.00%
	top 1% negative D-value	no neg. values	no neg. values
QEgg1 chr 8	top 1% positive D-value	no pos. values	no pos. values
	top 1% negative D-value	9.52%	29.64%
QEgg2 chr 15	top 1% positive D-value	73.53%	76.23%
	top 1% negative D-value	no neg. values	no neg. values
QFin1 chr 24	top 1% positive D-value	48.42%	82.70%
	top 1% negative D-value	no neg. values	no neg. values
QFin1 chr 24 (within +/- 15 windows interval from QTL)	top 1% positive D-value	2.73%	23.62%
	top 1% negative D-value	no neg. values	no neg. values

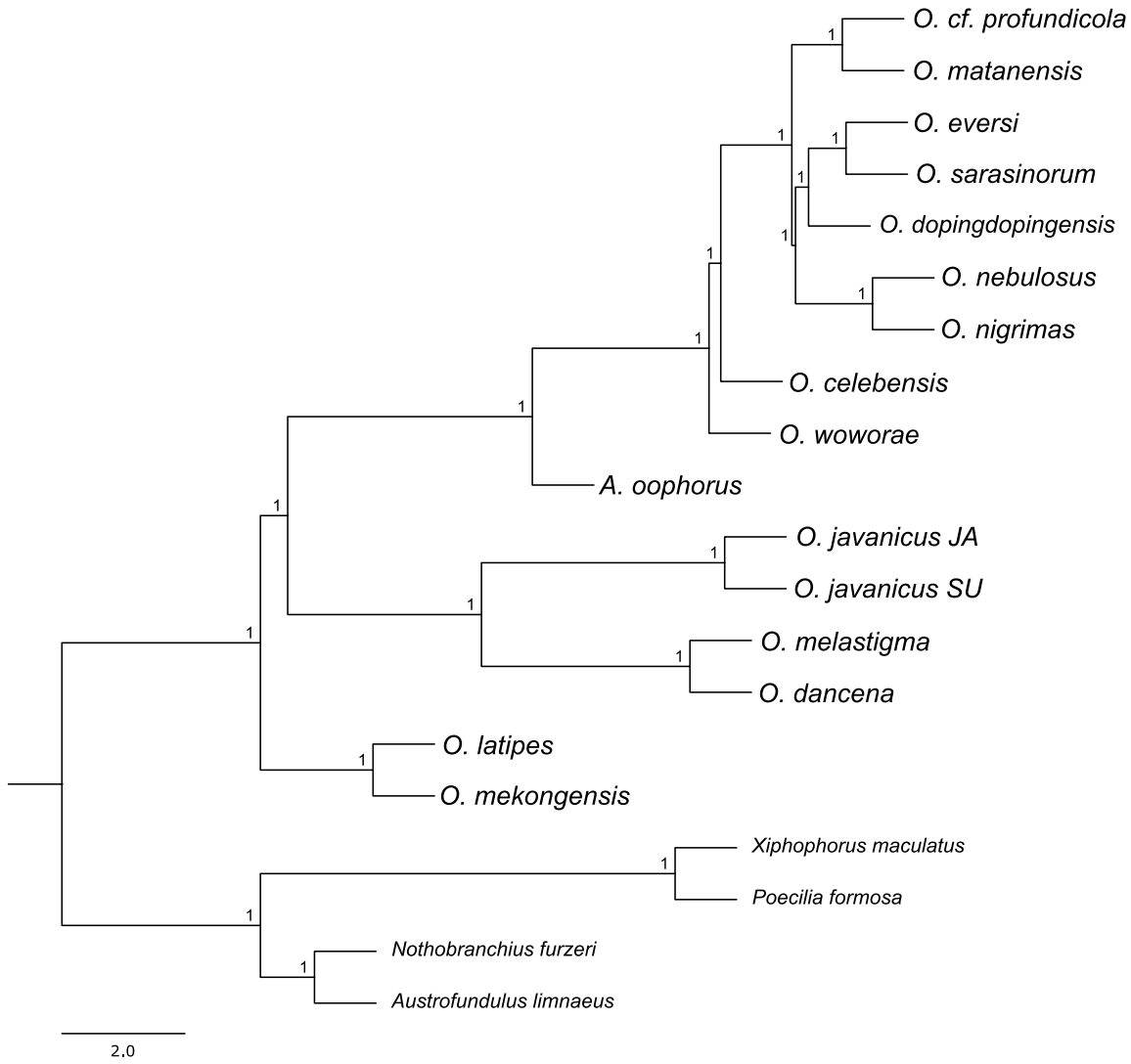


Fig S1: Posterior probabilities for species tree calculated in Astral based on 1907 gene trees.

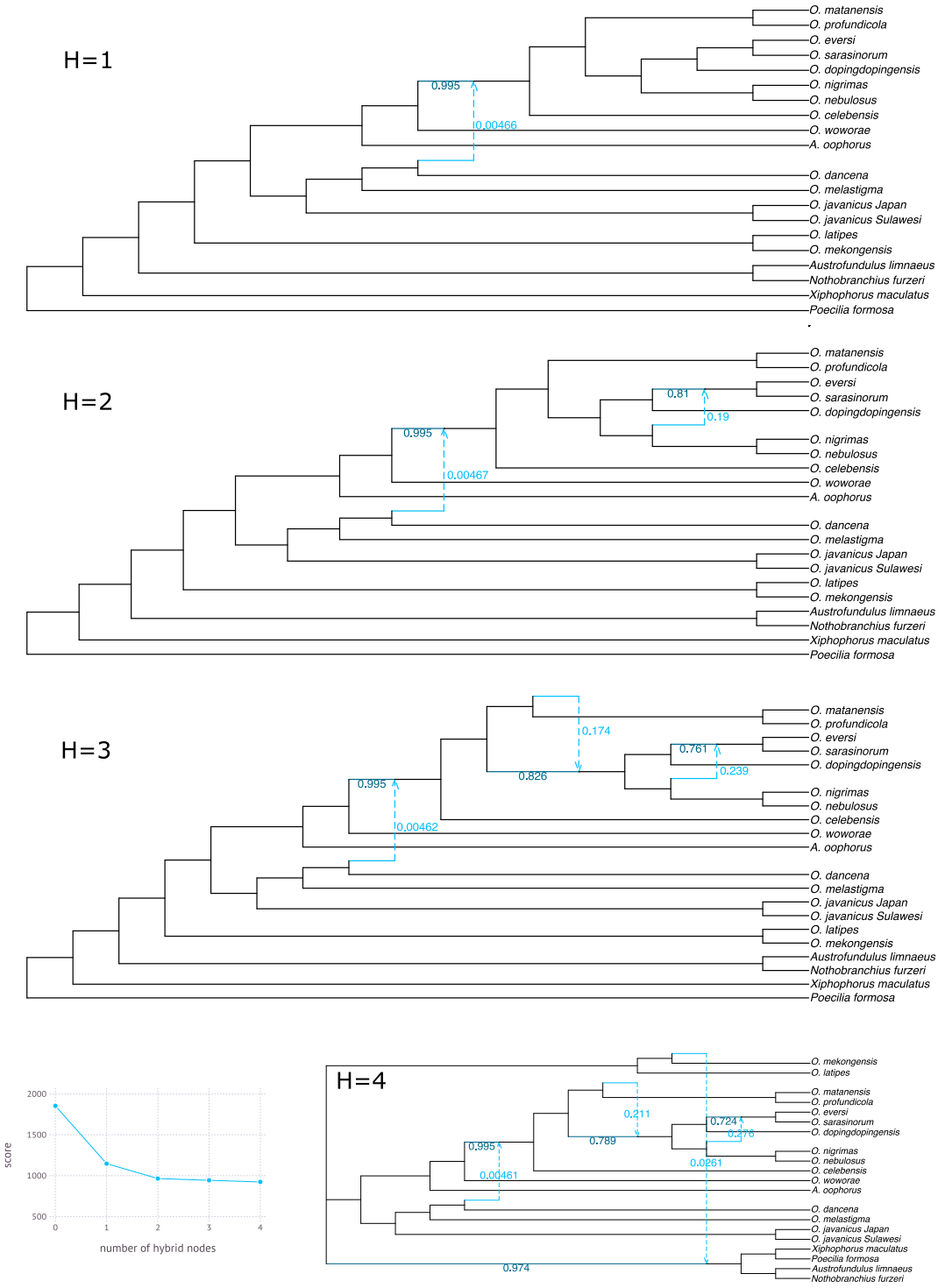


Fig S2: Result from SNaQ for 1-4 hybridization events. Hybridization between *O. mekongensis* and outgroup rather unlikely and probably false positive result.

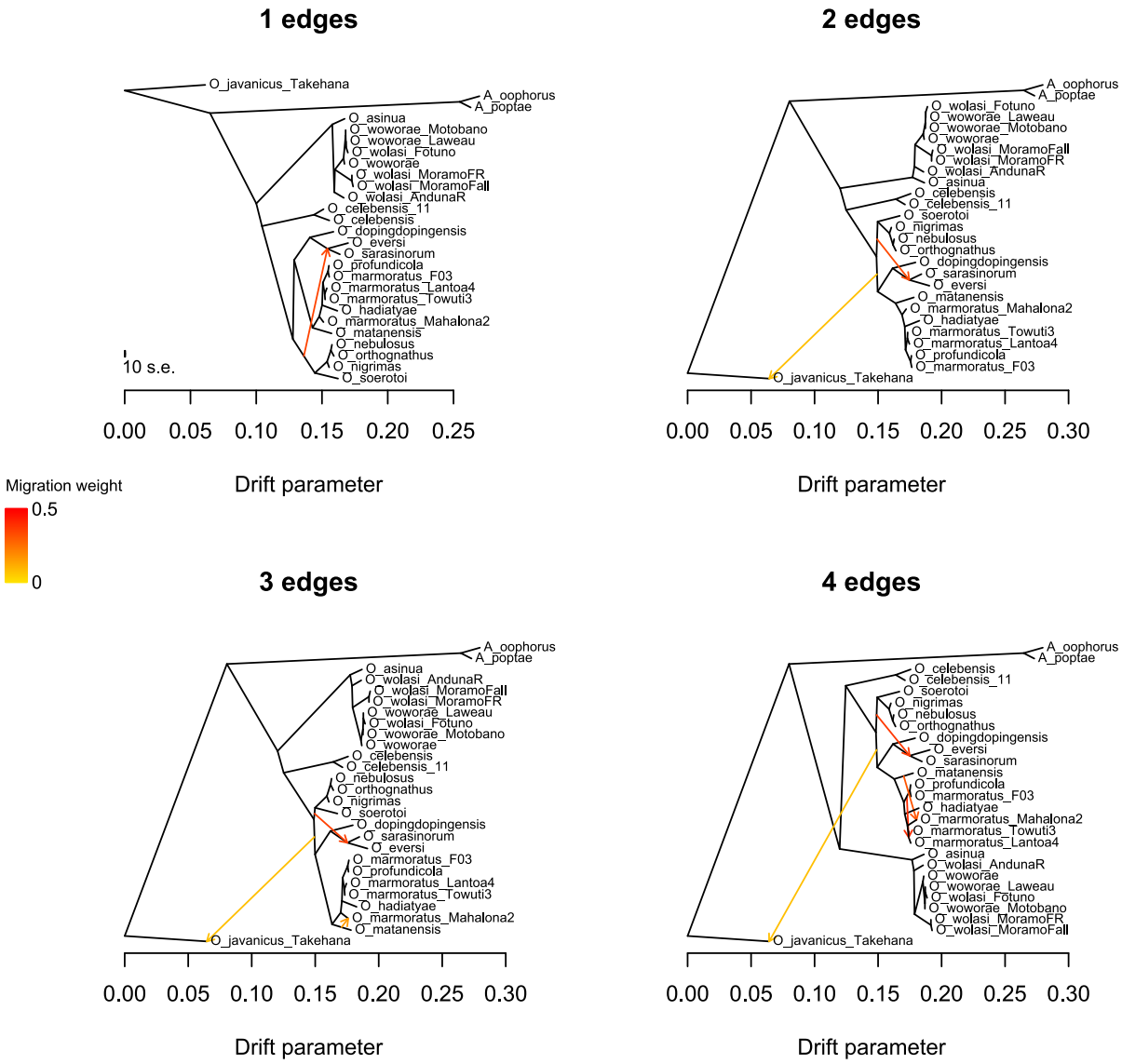
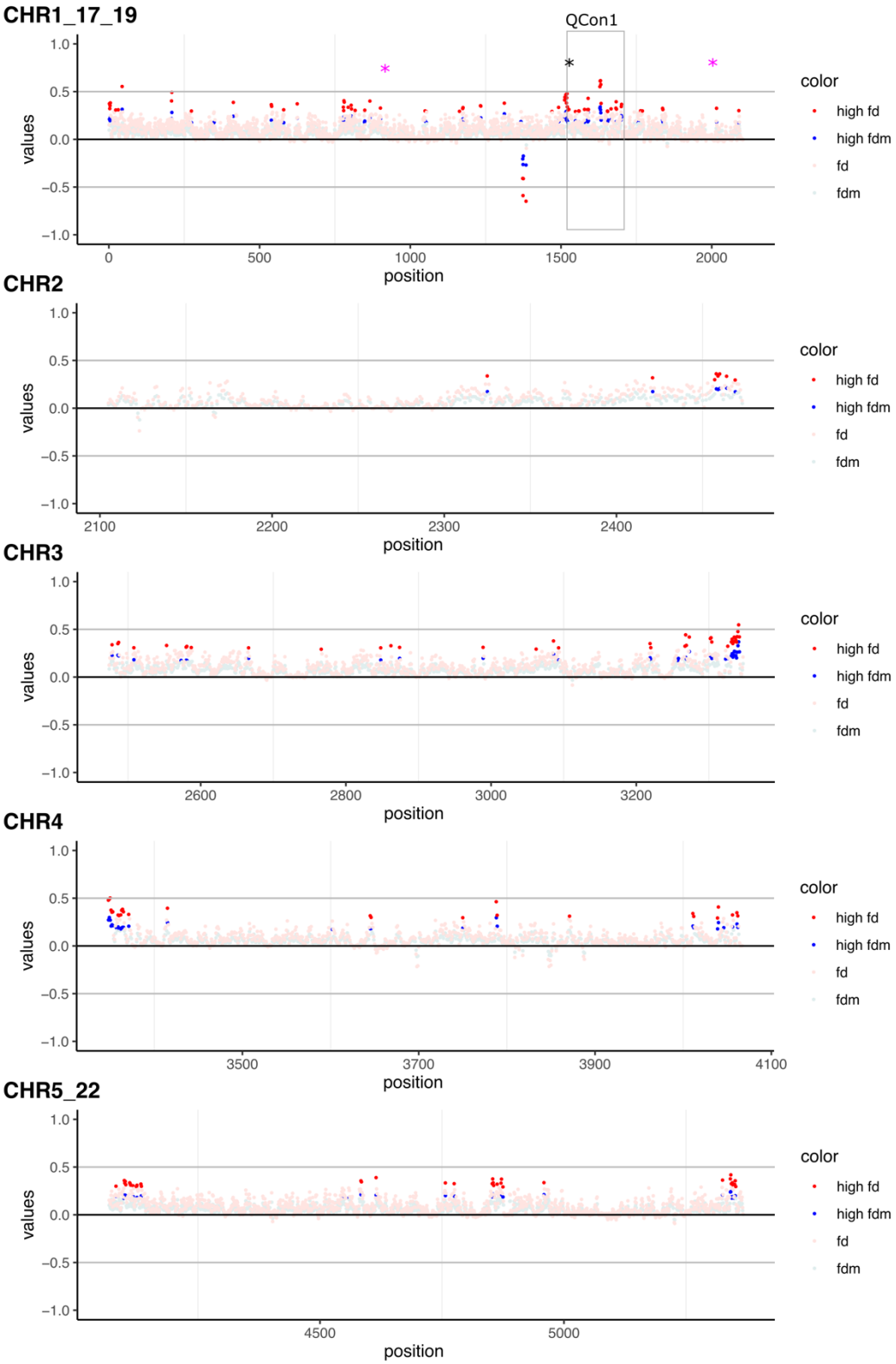
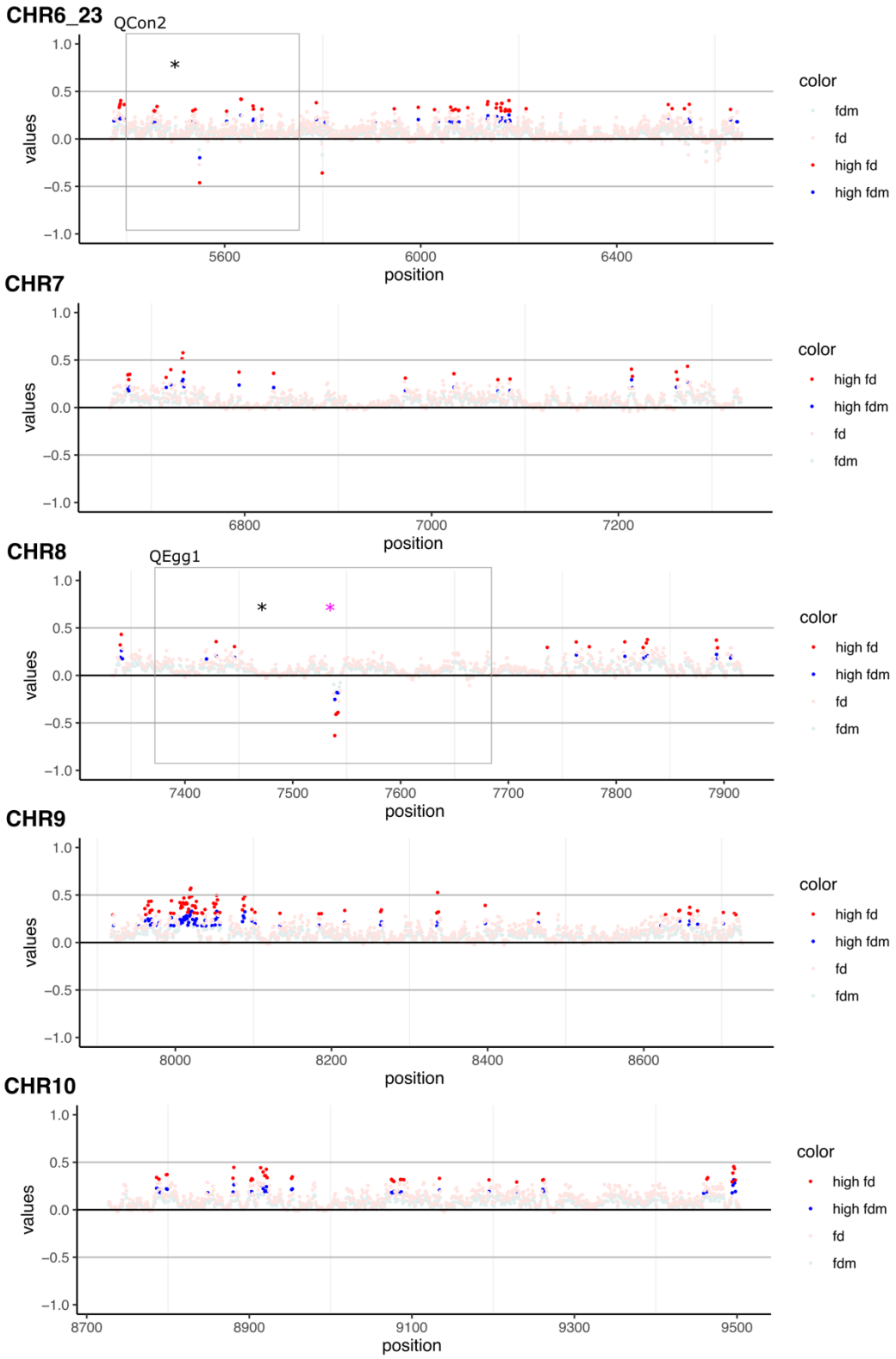
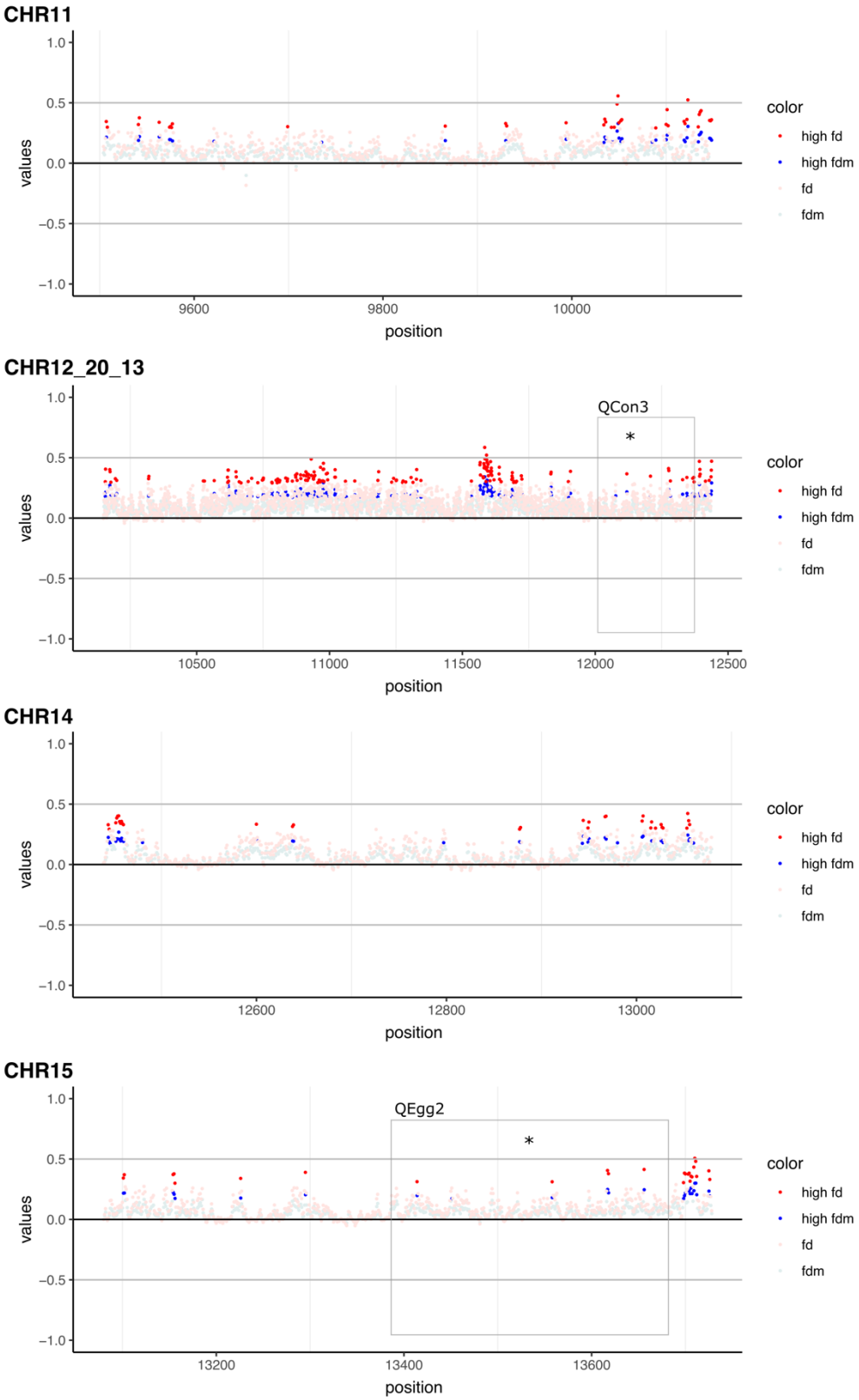


Fig S3: Results from TreeMix: same introgression events were found as in the Dsuite analysis, 1) between the *Oryzias* pelvic brooders and the *Oryzias* Lake Poso and 2) between *O. matanensis* and *O. marmoratus* Mahalona.


CHR1_17_19
QCon1
values
position
color
high fd
high fdm
fd
fdm
CHR2
CHR3
CHR4
CHR5_22


CHR6_23
QCon2
values
position
color
fdm
fd
high fd
high fdm
CHR7
CHR8
QEgg1
CHR9
CHR10


CHR11
values
position
color
high fd
high fdm
fd
fdm
CHR12_20_13
QCon3
*
CHR14
CHR15
QEgg2
*

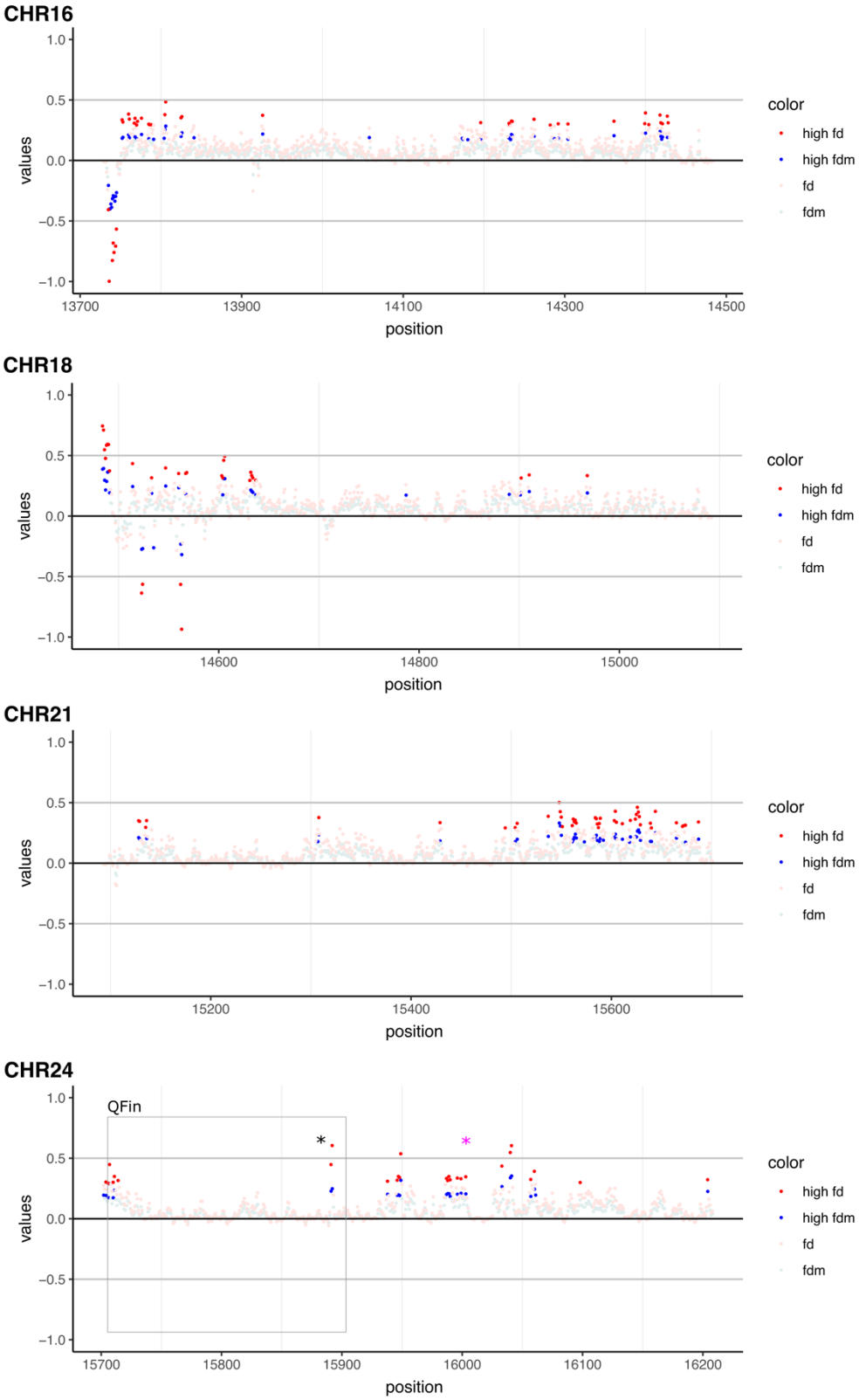
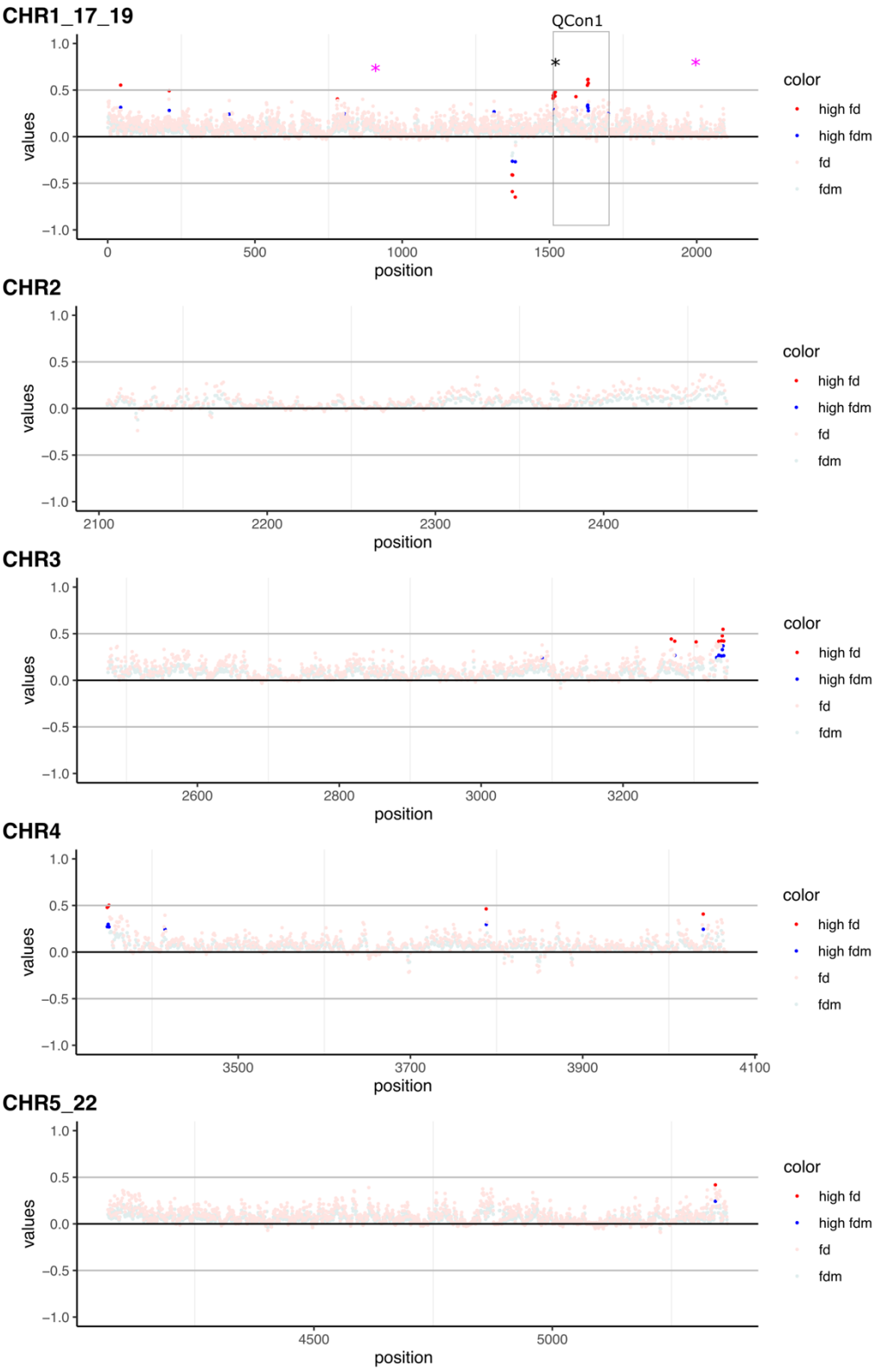
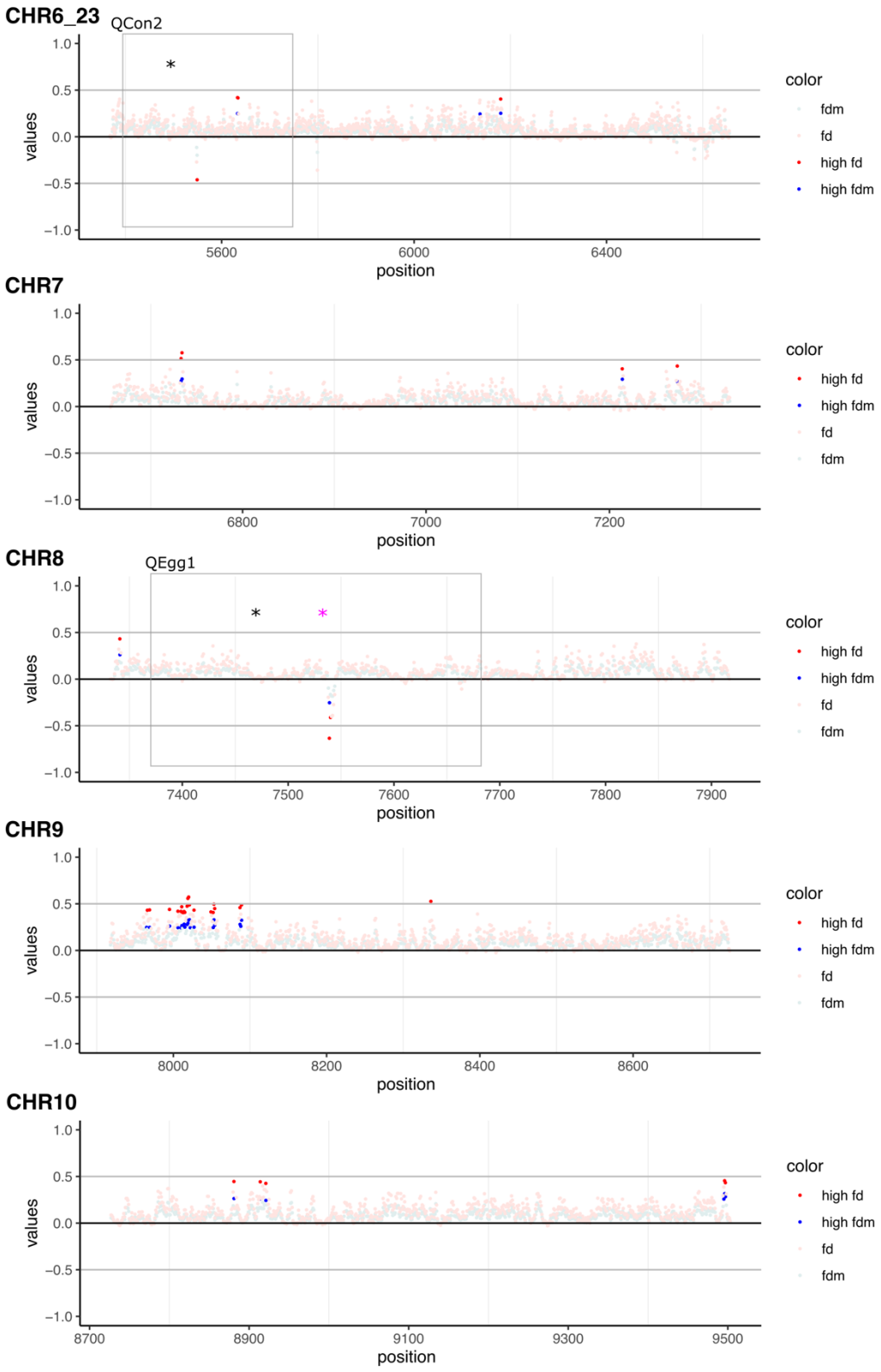
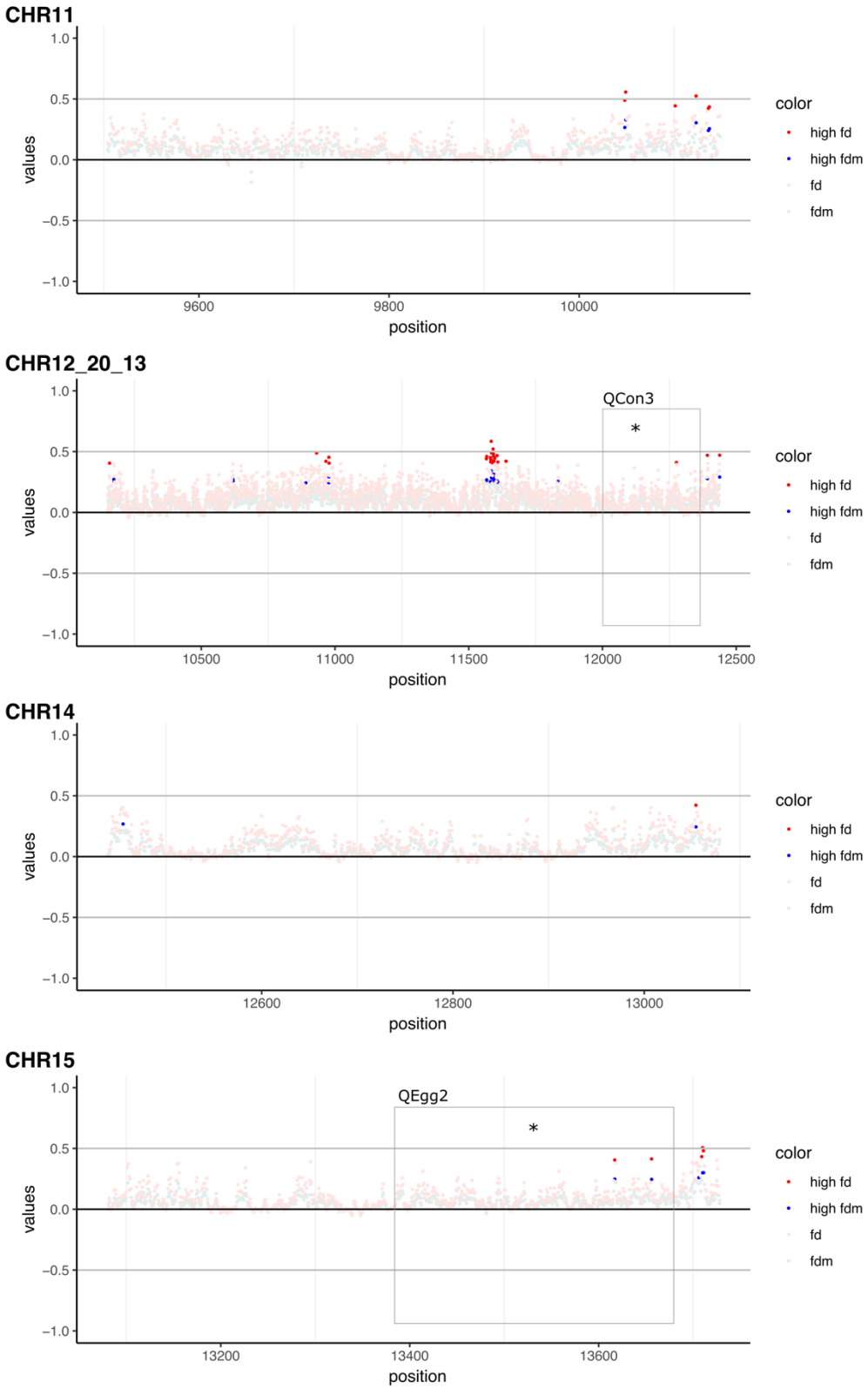

CHR16
values
position
color
high fd
high fdm
fd
fdm
CHR18
CHR21
CHR24
QFin

Fig S4: Results from DInvestigate: On the y-axis are the D-values for the sliding-windows (size 50 kbp, red top 5% values of f_d, blue top 5% values of f_{dM}). On the x-axis are the positions on the respective *O. celebensis* reference genome, naming follows the synteny with the *O. latipes* reference genom (how positions refer to base pairs can be read in Table S11). Grey boxes mark the confidence intervals of the QTLs found in Montenegro et al., 2022, black stars the QTLs (QCon1-3, QEgg1-2, QFin). Pink stars mark the position of the genes with elevated D-values found in HybridCheck.


CHR1_17_19
QCon1
values
position
color
high fd
high fdm
fd
fdm
CHR2
CHR3
CHR4
CHR5_22


CHR6_23
QCon2
*
values
position
color
fdm
fd
high fd
high fdm
CHR7
CHR8
QEgg1
CHR9
CHR10


CHR11
values
position
color
high fd
high fdm
fd
fdm
CHR12_20_13
QCon3
*
CHR14
CHR15
QEgg2

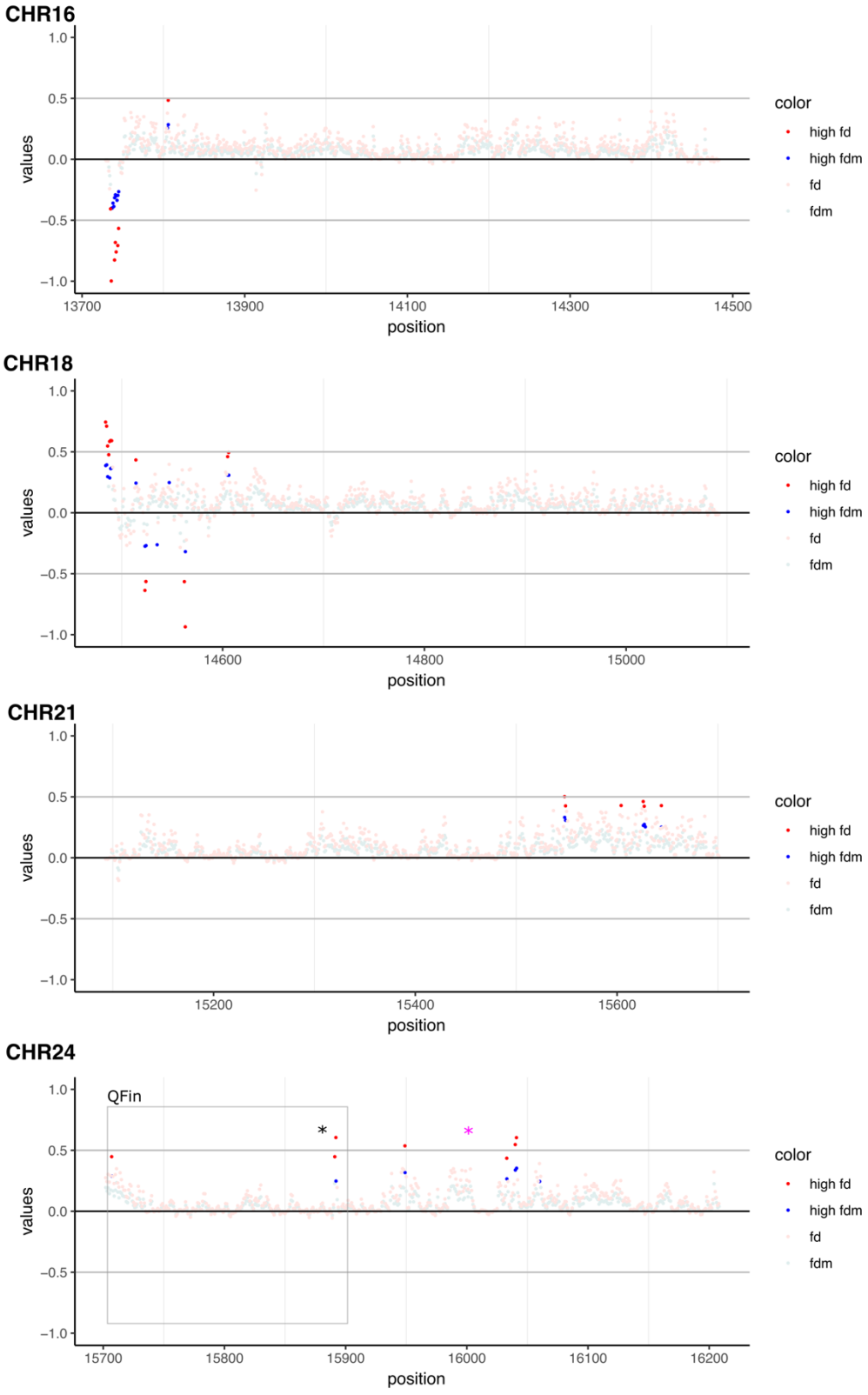

CHR16
values
position
color
high fd
high fdm
fd
fdm
CHR18
values
position
color
high fd
high fdm
fd
fdm
CHR21
values
position
color
high fd
high fdm
fd
fdm
CHR24
QFin
values
position
color
high fd
high fdm
fd
fdm

Fig S5: same figure as S4, but with only top 1% of D-values in red and blue. Sliding-windows have a size of 50 kbp, in red top 1% values of f_d, in blue top 1% values of f_{dM}). On the x-axis are the positions on the respective *O. celebensis* reference genome, naming follows the synteny with the *O. latipes* reference genome (how positions refer to base pairs can be read in Table S11). Grey boxes mark the confidence intervals of the QTLs found in Montenegro et al., 2022, black stars the QTLs (QCon1-3, QEgg1-2, QFin). Pink stars mark the position of the genes with elevated D-values found in HybridCheck.

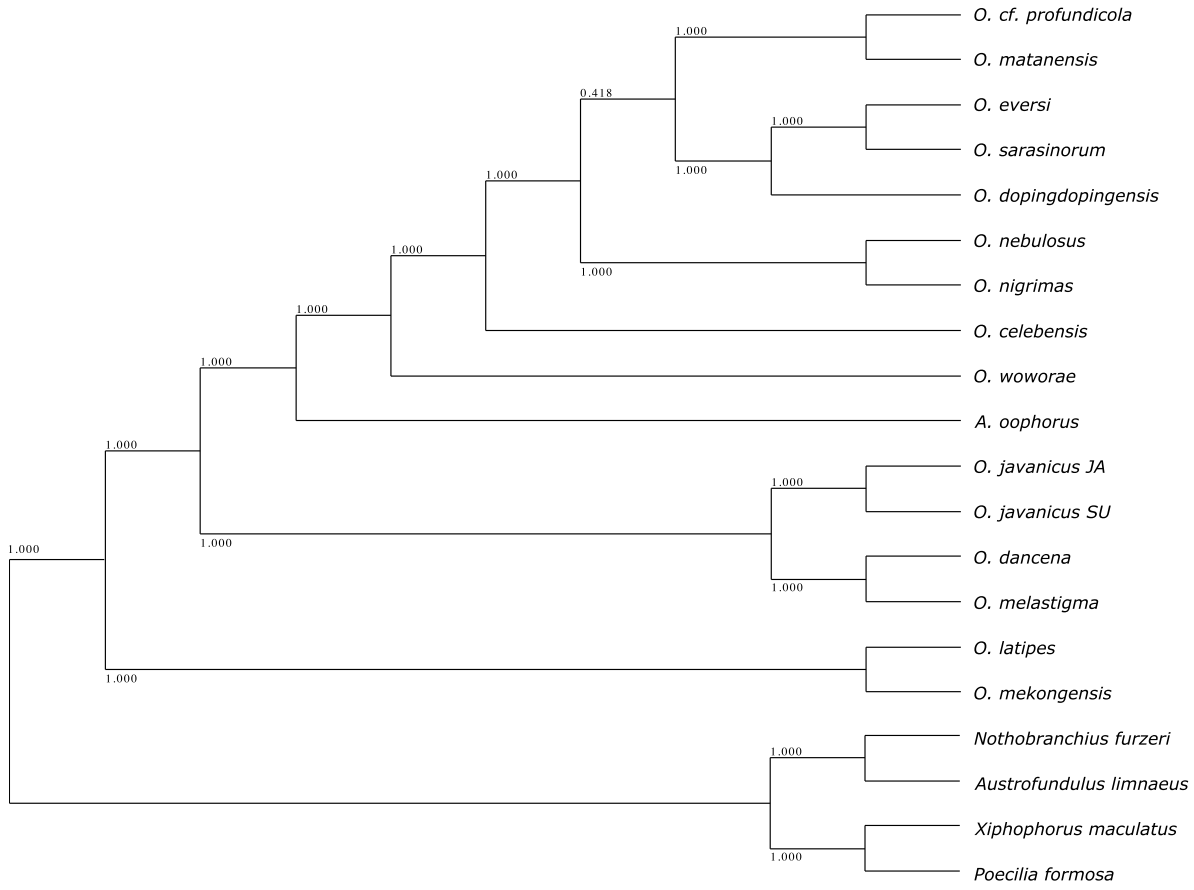


Fig S6: Results from FastTree: Cladogram with “fast-global” bootstrap support.

Material and Methods

RNA extraction and library preparation

The dead fish was put into a plastic tube and 1ml TRIzol per 100mg of tissue was added. A polytron was used to homogenize the fish thoroughly and the lyzed tissue was incubated at room temperature for 5 minutes. An aliquot of 1ml was transferred to a new tube and the remaining sample was put in the -80°C freezer. To the 1ml aliquot, we added 200µl of chloroform. The tube was shaken by hand for 15 seconds and incubated for 3 minutes at room temperature before centrifuging for 15 minutes at 12000g at 2-8°C. The aqueous phase was transferred to a new tube. Isopropanol (half the volume of originally added TRIzol) was added, mixed and incubated at room temperature for 15 minutes. The tube was placed into the centrifuge for 10 minutes at 12000g at 2-8°C. The RNA formed a gel like precipitate. The supernatant was removed and the pellet washed with 1 ml 70% EtOH. After a short mix the tube was put into the centrifuge for 5 minutes at 7500g at 2-8°C. The pellet was air-dried and diluted in RNAse free water (400µl) and stored at -20°C.

Post isolation cleanup of crude RNA extract

An equal volume of LiCL (5M) was added to the crude RNA extract, mixed and incubated for one hour at -20°C. The tube was centrifuged for 30 minutes at 16000g. The supernatant was removed and the pellet was washed with 70% EtOH until it appeared white and brittle. The tube was again centrifuged to collect particles and the EtOH was carefully removed. This wash step was repeated. After the pellet was air-dried and 455 µl of water and 5 µl of RNAse inhibitor was added. 10 µl of EDTA o.5 M were added, mixed and incubated for 15 minutes at 65°C. The solution was loaded on a microcon spin column and centrifuged at 14000g until all liquid has passed. The spin column was refilled with 300µl of RNAse free water, mixed using a filtered tip and centrifuged again. The last step was repeated. RNA was eluted in 200 µl of water and RNAse inhibitor was added.

DNA extraction and library preparation

DNA was extracted from a tissue sample of *O. dopingdopingensis* using the Qiagen DNeasy Blood & Tissue kit. Libraries were prepared using TruSeq DNA PCR free (350) at macrogen sequencing company and sequenced on an Illumina Hiseq2000.

Transcriptome sequencing, trimming and assembly

The Truseq mRNA libraries were prepared at the CCG in Cologne and sequenced on Illumina Hiseq2000. Raw data was trimmed and quality-filtered using trim-fast.pl from the PoPoolation pipeline with default options (Kofler et al. 2011). Transcriptomes of all species were assembled *de novo* using Trinity v2.8.4 (Grabherr et al. 2011; Haas et al. 2013). We ran Trinity in standard (not strand-specific) mode with in silico read normalization (max. read coverage = 50), two-fold minimal kmer coverage and a minimal contig length of 250 bp. BUSCO v3.0.2 (Simão et al. 2015) was used to generate estimates of transcriptome completeness, redundancy and fragmentation by searching for 4,584 Actinopterygii single copy orthologs (odb9).

Short read sequencing, assembly and contamination check

A genomic short read TruSeq DNA PCR free library was prepared by Macrogen sequencing company. De novo whole-genome assemblies were generated from the Illumina raw sequencing data following the approach described in Böhne et al. (Böhne et al. 2019) and Malmstrøm et al. (Malmstrøm et al. 2017) using CeleraAssembler v.8.3 (Myers et al. 2000) and FLASh v.1.2.11 (Magoc & Salzberg 2011). Assembly quality and read coverage were evaluated with QUAST v.502 (Gurevich et al. 2013) (Table S4). The completeness of the assemblies was assessed with BUSCO v.406 (Manni et al. 2021) using the BUSCO test library of 3640 conserved actinopterygian genes.

Annotation of genome assembly

Identification and masking of repetitive elements in the genome sequence of *O. dopingdopingensis* was performed with the following bioinformatic tool case. Nucleotides were masked using the DUST algorithm with dustmasker (version 1.0.0, part of blast+ 2.9.0

(Altschul et al. 1990; Camacho et al. 2009) (Kuzio et al., unpublished but described in Morgulis, Gertz, Schäffer, & Agarwala, 2006). Tandem Repeats were identified with Tandem Repeat Finder (trf version 4.09) (Benson 1999). A species-specific *de novo* repeat library was built with RepeatModeler v1.0.11 (http://www.repeatmasker.org/RepeatModeler/). Repeat Elements were located in the genome sequence using RepeatMasker (version 4.1.0) (http://www.repeatmasker.org) with the *de novo* and *Danio rerio* libraries. The information from all four repeat analyses was merged and the genome was softmasked with bedtools (2.29.2) (Quinlan & Hall 2010) PMID: 20110278; PMCID: PMC2832824.]. All steps of masking repetitive regions were performed with scripts provided by the sigenae platform, following the workflow from (Feron et al. 2020).

For the identification of genes the masked genome was annotated with funannotate (Palmer & Stajich 2019). The sequences were sorted by length with the 'funannotate sort' function, followed by a gene prediction with 'funannotate predict'. No training based on RNA-Seq data was performed since it was not available for this species. Additional external evidence from transcripts and proteins was added. As transcript evidence, gene predictions from *Oryzias latipes* (NCBI Bioproject:PRJNA183868; Assembly: GCF_002234675.1) (Kasahara et al. 2007) and *Oryzias melastigma* (NCBI Bioproject: PRJNA401159 ; Assembly: ASM292280v2) (Kim et al. 2018) were used. As protein evidence, a protein set from Oryzias javanicus (NCBI Bioprject : PRJNA505405 ; Assembly: GCA_003999625.1) (Lee et al. 2020), manually annotated reference sequences from UniProt Knowledgebase (UniProtKB) (Release 2020_02 (22-Apr-2020) UniProtKB/Swiss-Prot with 562,253 entries) (Apweiler et al. 2004) and a set of orthologous sequences generated in this study. Furthermore, the *de novo* gene predictors were trained with the Busco dataset of actinopterygii_odb10. Gene prediction resulted in a total of 56658 genes.

Ortholog set

We generated a reference set consisting of 8552 single-copy protein-coding genes derived from OrthoDB v.9.1 (Waterhouse et al. 2013) available for the following species: *Austrofundulus limnaeus, Centrocoris variegatus, Fundulus heteroclitus, Kryptolebias marmoratus, Nothobranchius furzeri, Oryzias latipes, O. melastigma, Poecilia formosa, P. latipinna ,P. mexicana, P. reticulata* and *Xiphophorus maculatus* (NCBI Accession numbers in Table S6). The hierarchical split was set to Actinopterygii (ID 7898). We used the script "make-ogs-corresponding.pl" to check for inconsistencies between the amino acid sequences and the corresponding nucleotide sequences and removed 50 problematic genes (Tab. S5).

Identification of orthologs for transcripts and genome and alignment of single-copy genes

Ortholog identification among 16 ricefish species and four outgroups (table) was carried out with Orthograph v0.7.1 (Petersen et al. 2017). Forward search for candidate transcript was left at default. Best reciprocal hit: Ortholog candidate genes needed at least one hit in either *O. latipes* or *O. melastigma* and we allowed concatenation of hits if they met the criteria and did not overlap. Max-blast-searches were set to 50, blast-max-hits were also set to 50. "U" in the amino acid sequences was changed to "X" to avoid issues in downstream analysis. The results of the orthology prediction were summarized for all species using a custom perl script coming with the orthograph package. Sequences of only those orthologs with all species present were aligned using MAFFT v7.221 with the L-INS-I algorithm on amino acid level (Katoh & Standley 2013). Outliers were identified according to Misof et al. 2014 and we subsequently removed all orthologs with identified outlier sequences from further analysis. We used the amino-acid alignments as blue print to generate corresponding nucleotide alignments with a modified version of Pal2Nal v14 (Suyama et al. 2006; Misof et al. 2014). To check each amino acid alignment for ambiguously aligned regions, we ran ALISCORE v2.0 with the maximal number of possible sequence selected pairs to analyze (-r) (Misof & Misof 2009; Kück et al. 2010; Misof et al. 2014). Sites which needed masking were cut out using ALICUT v2.3 (Kück

2009) from the amino acid alignments and correspondingly also from the nucleotide alignments. For further analyses we only proceeded with the data set on nucleotide level

Multispecies coalescent tree

For the masked data set a substitution model was estimated for each gene alignment on nucleotide level using IQ-TREE v1.6.12 with all available nt models (Nguyen et al. 2014; Kalyaanamoorthy et al. 2017). Using the most appropriate evolutionary model, we calculated ten trees with random seed for each gene alignment and chose the best-scoring tree according to AICc (corrected Akaike Information Criterion). In addition, for each gene, branch support was estimated using the fast mode with IQ-TREE from 3000 BS replicates and plotted onto the best scoring gene tree. We collapsed splits with a bootstrap support below 10%. Each tree was rooted with the outgroup taxa using Newick utilities (Junier & Zdobnov 2010). Using ASTRAL v5.7.3, we calculated a species tree from the gene trees under the multi-species coalescent model (-t 8 printing alternative quartet support, -t 3 printing local posterior probabilites (Zhang et al. 2018). We scanned the gene trees for paralogous sequences using PhyloTreePruner (Kocot et al. 2013). The analysis resulted in no suspicious genes. Further, we ran TreeShrink, to find gene trees with extremely long branches (Mai & Mirarab 2018). We found 492 gene trees with suspiciously long branches which we therefore removed from further analyses.

Discovista

To visualize phylogenetic discordance we used the software package DiscoVista (Sayyari et al. 2018). Each species was assigned as one clade and the four outgroup species as the base. Documentation on how to run the program can be found on https://github.com/esayyari/DiscoVista. Here we used the gene trees with all splits with a BS below 10% collapsed.

Densitree

We collapsed splits with lower than 30% bootstraps support using Newick Utilities. The trees needed to be ultrametric, which was done in R v4.1.2 using the package “ape” v5.6-2

(Paradis & Schliep 2019) and the function "chronos". All ultrametric trees were added to one file and imported into DensiTree2.01 (Bouckaert & Heled 2014).

Ordering genes according to the *Oryzias latipes* reference genome

To take linkage into account we assigned genes identified as single copy orthologs to *Oryzias latipes* linkage groups based on the *O. latipes* reference (GCF_002234675.1, Bioproject PRJNA325079) using blast+ vs 2.9.0 with evalue=0.00001, -outfmt 7, -max_target_seqs 10. Alignments were then concatenated according to linkage group information into a supermatrix (from LG 1 in ascending order) using FASconCAT-G (Kück & Longo 2014).

Phylogenetic reconstruction using maximum likelihood

We used the concatenated supermatrix to run FastTree v2.1.11 (Price et al. 2009) with the generalized time-reversible model with the neighbour-joining option and 3000 "fast-global" bootstraps (Fig. S5). For the ML tree reconstruction the same supermatrix and its partition scheme based on the substitution models found analysing single gene trees were used. We ran 20 single tree searches, 10 with a randomized starting tree, 10 with a parsimony starting tree using IQ-TREE v1.6.12 (Nguyen et al. 2014). Statistical support was derived from 50 non-parametric, slow bootstrap replicates which was subsequently mapped on the tree with the best log-likelihood score. The number of unique tree topologies were checked with Uniquetree version 1.9 (Wong, available upon request).

We checked for the convergence of bootstrap replicates a posteriori with RAxML version 8.2.11 (options: -autoMRE -B 0.01 --bootstop-perms=10000) starting with random seeds (Pattengale et al. 2010) 10 times independently. Bootstrap replicates converged always after 50 replicates. We checked the data for rogue taxa using RogueNaRok version 1.0 (Aberer et al. 2013) providing for each data set the best ML tree and otherwise default settings. Our data set was found to be free from rogue taxa.

Introgression analysis

We used published shortreads from Genbank of *O. javanicus* (Bioproject PRJNA505405, accession number SRR8467745) (Takehana et al. 2020), *O. melastigma* (Bioproject PRJNA556761, accession number SRR12442554) and 26 other *Adrianichthyidae* (Bioproject PRJDB10385) (Ansai et al. 2021) to assess introgression on a genomic level. The short reads were mapped on the *O. celebensis* reference genome (DRA010635 on DDBJ) (Ansai et al. 2021) using bowtie2 v2.3.5.1 (Langmead & Salzberg 2012) and sorted using samtools/1.10 (Li 2011). Further, we used the functions fixmate and markdup of samtools to eliminate PCR duplicates. SNPs were called using mpileup from bcftools/1.10.2 (Li 2011) and the vcf file was checked using the R package vcfR (Knaus & Grünwald 2017) and filtered using bcftools (QUAL>25, DP>30, MQ>25, only SNPs). The filtered vcf file was analyzed in Dsuite (Malinsky et al. 2021) using the function Dtrios which automatically tests all possible combinations. All comparisons with D-values above 0.15 were considered to be realistic introgression events. We used *O. javanicus* as outgroup. A shortened version of the same vcf file used in Dsuite was used as input for DInvestigate containing the following species: *Oryzias eversi* and *O. sarasinorum* were marked as one population (pelvic brooders), the Poso *Oryzias* (*O. nebulosus*, *O. nigrimas*, *O. orthognathus*, *O. soerotoi*) were defined as one population and *O. dopingdopingensis* as one. Window steps and size were set at default. We used the corrected D-values f_d (Martin et al. 2015) and f_{dM} (Malinsky et al. 2015) for whole genome analysis. We considered D-values as high if they belonged to the top 5% of detected D-values. To evaluate which chromosomes have highest introgression signal, the number of high D-values was divided by number of windows for each chromosome, for both f_d and f_{dM}. Further, we sampled as many random windows from all chromosomes as the number of windows within the confidence intervals of the QTLs found in (Montenegro et al. 2022), to see if the pattern of D-values within the confidence intervals cannot be observed by chance. The random sampling was done 10.000 times per confidence interval, and we looked at the highest D-value and the sum of D-values in the random sample, also for both f_d and f_{dM}.

With SNaQ we created a phylogenetic network to identify hybridization based on the gene trees derived from IQ-TREE. SNaQ is part of the PhyloNetworks package (Solís-Lemus et al. 2017) . We increased the number of hybrid edges until the log likelihood score did not anymore change significantly.

We prepared four multiple sequence alignments for HybridCheck, with *O. dopingdopingensis* as P1, either *O. sarasinorum* or *O. eversi* as P2, either *O. nebulosus* or *O. nigrimas* as P3 and *A. oophorus* as P1. We divided each alignment into 50 steps, where ABBA and BABA sites were counted and compared. Regions which were significantly deviated in all four comparisons were used and the gene was located on the *O. latipes* reference genome.

For the TreeMix (v1.13) (Pickrell & Pritchard 2012) analysis we used the >38 mio SNPs data set with all missing data removed and we used the Stacks (v2.60) (Catchen et al. 2011) function --treemix to prepare the input file. We ran TreeMix with k=1000 to account for loci in linkage. We tested four hypotheses with 1 to 4 introgression events (= edges) (Fig. S2).

Ancestral state reconstruction

We used RASP4.2 (Yu et al. 2020) to infer the ancestral state reconstruction of pelvic-brooding on the ML-tree. *Poecilia formosa* had status "A" as live-bearer, transfer brooders had status "B" and the pelvic brooders *A. oophorus*, *O. sarasinorum* and *O. eversi* had status "C". Each species could only have one status. We ran the S-DIVA model.

References

Aberer AJ, Krompass D, Stamatakis A. 2013. Pruning rogue taxa Improves phylogenetic accuracy: an efficient algorithm and webservice. Syst. Biol. 62:162–166. doi: 10.1093/sysbio/sys078.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403–410. doi: 10.1016/S0022-2836(05)80360-2.

Ansai S et al. 2021. Genome editing reveals fitness effects of a gene for sexual dichromatism in Sulawesian fishes. Nat. Commun. 12:1350. doi: 10.1038/s41467-021-21697-0.

Apweiler R et al. 2004. UniProt: the Universal Protein knowledgebase. Nucleic Acids Res. 32:115D – 119. doi: 10.1093/nar/gkh131.

Benson G. 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 27:573–580. doi: 10.1093/nar/27.2.573.

Böhne A et al. 2019. Repeated evolution versus common ancestry: sex chromosome evolution in the haplochromine cichlid Pseudocrenilabrus philander A Katz, L, editor. Genome Biol. Evol. 11:439–458. doi: 10.1093/gbe/evz003.

Bouckaert RR, Heled J. 2014. DensiTree 2: seeing trees through the forest. bioRxiv. 012401. doi: 10.1101/012401.

Camacho C et al. 2009. BLAST+: architecture and applications. BMC Bioinformatics. 10:421. doi: 10.1186/1471-2105-10-421.

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH. 2011. Stacks: building and genotyping loci *de novo* from short-read sequences. G3 Genes|Genomes|Genetics. 1:171–182. doi: 10.1534/g3.111.000240.

Feron R et al. 2020. Characterization of a Y-specific duplication/insertion of the anti-Mullerian hormone type II receptor gene based on a chromosome-scale genome assembly of yellow perch, Perca flavescens. Mol. Ecol. Resour. 20:531–543. doi: 10.1111/1755-0998.13133.

Grabherr MG et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. doi: 10.1038/nbt.1883.

Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 29:1072–1075. doi: 10.1093/bioinformatics/btt086.

Haas BJ et al. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. doi: 10.1038/nprot.2013.084.

Junier T, Zdobnov EM. 2010. The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. Bioinformatics. 26:1669–1670. doi: 10.1093/bioinformatics/btq243.

Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods. 14:587–589. doi: 10.1038/nmeth.4285.

Kasahara M et al. 2007. The medaka draft genome and insights into vertebrate genome evolution. Nature. 447:714–719. doi: 10.1038/nature05846.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30:772–780. doi: 10.1093/molbev/mst010.

Kim HS et al. 2018. The genome of the marine medaka Oryzias melastigma. Mol. Ecol. Resour. doi: 10.1111/1755-0998.12769.

Knaus BJ, Grünwald NJ. 2017. vcfr : a package to manipulate and visualize variant call format data in R. Mol. Ecol. Resour. 17:44–53. doi: 10.1111/1755-0998.12549.

Kocot KM, Citarella MR, Moroz LL, Halanych KM. 2013. PhyloTreePruner: a phylogenetic tree-based approach for selection of orthologous sequences for phylogenomics. Evol.

Bioinform. Online. 9:429–35. doi: 10.4137/EBO.S12813.
Kofler R et al. 2011. PoPoolation: a toolbox for population genetic analysis of next generation sequencing data from pooled individuals Kayser, M, editor. PLoS One. 6:e15925. doi: 10.1371/journal.pone.0015925.
Kück P. 2009. ALICUT: a Perlscript which cuts ALISCORE identified RSS. https://github.com/PatrickKueck/AliCUT.
Kück P et al. 2010. Parametric and non-parametric masking of randomness in sequence alignments can be improved and leads to better resolved trees. Front. Zool. 7:10. doi: 10.1186/1742-9994-7-10.
Kück P, Longo GC. 2014. FASconCAT-G: extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. Front. Zool. doi: 10.1186/s12983-014-0081-x.
Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods. 9:357–359. doi: 10.1038/nmeth.1923.
Lee BY et al. 2020. The genome of the Java medaka (Oryzias javanicus): Potential for its use in marine molecular ecotoxicology. Mar. Pollut. Bull. 154:111118. doi: 10.1016/j.marpolbul.2020.111118.
Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics. 27:2987–93. doi: 10.1093/bioinformatics/btr509.
Magoc T, Salzberg SL. 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics. 27:2957–2963. doi: 10.1093/bioinformatics/btr507.
Mai U, Mirarab S. 2018. TreeShrink: fast and accurate detection of outlier long branches in collections of phylogenetic trees. BMC Genomics. 19:272. doi: 10.1186/s12864-018-4620-2.
Malinsky M et al. 2015. Genomic islands of speciation separate cichlid ecomorphs in an East African crater lake. Science. 350:1493–1498. doi: 10.1126/science.aac9927.
Malinsky M, Matschiner M, Svardal H. 2021. Dsuite - Fast D -statistics and related admixture evidence from VCF files. Mol. Ecol. Resour. 21:584–595. doi: 10.1111/1755-0998.13265.
Malmstrøm M, Matschiner M, Tørresen OK, Jakobsen KS, Jentoft S. 2017. Whole genome sequencing data and de novo draft assemblies for 66 teleost species. Sci. Data. 4:160132. doi: 10.1038/sdata.2016.132.
Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. 2021. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes Kelley, J, editor. Mol. Biol. Evol. 38:4647–4654. doi: 10.1093/molbev/msab199.
Martin SH, Davey JW, Jiggins CD. 2015. Evaluating the use of ABBA–BABA statistics to locate introgressed loci. Mol. Biol. Evol. 32:244–257. doi: 10.1093/molbev/msu269.
Misof B et al. 2014. Phylogenomics resolves the timing and pattern of insect evolution. Science. 346:763–767. doi: 10.1126/science.1257570.
Misof B, Misof K. 2009. A Monte Carlo approach successfully identifies randomness in multiple sequence alignments: a more objective means of data exclusion. Syst. Biol. 58:21–34. doi: 10.1093/sysbio/syp006.
Montenegro J et al. 2022. Genetic basis for the evolution of pelvic-fin brooding, a new mode of reproduction, in a Sulawesian fish. Mol. Ecol. doi: 10.1111/mec.16555.
Morgulis A, Gertz EM, Schäffer AA, Agarwala R. 2006. A fast and symmetric DUST implementation to mask low-complexity DNA sequences. J. Comput. Biol. 13:1028–1040. doi: 10.1089/cmb.2006.13.1028.
Myers EW et al. 2000. A whole-genome assembly of *Drosophila*. Science (80-.). 287:2196–2204. doi: 10.1126/science.287.5461.2196.
Nguyen L-T, Schmidt HA, von Haeseler A, Quang Minh B. 2014. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol.

Evol. 32:268–274. doi: 10.1093/molbev/msu300.
Palmer J, Stajich J. 2019. nextgenusfs/funannotate: funannotate v1.5.3. doi: https://doi.org/10.5281/zenodo.2604804.
Paradis E, Schliep K. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R Schwartz, R, editor. Bioinformatics. 35:526–528. doi: 10.1093/bioinformatics/bty633.
Pattengale ND, Alipour M, Bininda-Emonds ORP, Moret BME, Stamatakis A. 2010. How many bootstrap replicates are necessary? J. Comput. Biol. 17:337–354. doi: 10.1089/cmb.2009.0179.
Petersen M et al. 2017. Orthograph: a versatile tool for mapping coding nucleotide sequences to clusters of orthologous genes. BMC Bioinformatics. 18:1–10. doi: 10.1186/s12859-017-1529-8.
Pickrell JK, Pritchard JK. 2012. Inference of population splits and mixtures from genome-wide allele frequency data Tang, H, editor. PLoS Genet. 8:e1002967. doi: 10.1371/journal.pgen.1002967.
Price MN, Dehal PS, Arkin AP. 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol. Biol. Evol. 26:1641–1650. doi: 10.1093/molbev/msp077.
Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 26:841–842. doi: 10.1093/bioinformatics/btq033.
Sayyari E, Whitfield JB, Mirarab S. 2018. DiscoVista: interpretable visualizations of gene tree discordance. Mol. Phylogenet. Evol. 122:110–115. doi: 10.1016/J.YMPEV.2018.01.019.
Simão FA, Waterhouse RM, Ioannidis P, Kriventseva E V., Zdobnov EM. 2015. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. doi: 10.1093/bioinformatics/btv351.
Solís-Lemus C, Bastide P, Ané C. 2017. PhyloNetworks: a package for phylogenetic networks. Mol. Biol. Evol. 34:3292–3298. doi: 10.1093/molbev/msx235.
Suyama M, Torrents D, Bork P. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. doi: 10.1093/nar/gkl315.
Takehana Y et al. 2020. Genome Sequence of the Euryhaline Javafish Medaka, *Oryzias javanicus* : A Small Aquarium Fish Model for Studies on Adaptation to Salinity. G3 Genes|Genomes|Genetics. 10:907–915. doi: 10.1534/g3.119.400725.
Waterhouse RM, Tegenfeldt F, Li J, Zdobnov EM, Kriventseva E V. 2013. OrthoDB: a hierarchical catalog of animal, fungal and bacterial orthologs. Nucleic Acids Res. 41:D358–D365. doi: 10.1093/nar/gks1116.
Yu Y, Blair C, He X. 2020. RASP 4: ancestral state reconstruction tool for multiple genes and characters Yoder, A, editor. Mol. Biol. Evol. 37:604–606. doi: 10.1093/molbev/msz257.
Zhang C, Rabiee M, Sayyari E, Mirarab S. 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics. 19:153. doi: 10.1186/s12859-018-2129-y.

Sex determination in crosses between two Celebensis group medaka species

Jana M. Flury*[1], Kristin Tietje*[2], J. Schwarzer[1], A. W. Nolte[2]

*equal contribution

1 LIB Leibniz Institute for the Analysis of Biodiversity Change, Bonn, Germany

2 Carl-von-Ossietzky University, Oldenburg, Germany

Abstract

In contrast to mammals and birds, fishes show a large variation in genetic sex determination. Why this flexibility in sex determination exists and how it contributes to adaptation and speciation remains elusive. The ricefishes from Sulawesi serve as the perfect system to study the connection between the evolution of different sex determining genetic architectures and adaptive traits: first, so far seven sex determination turnovers were documented in this family and second, a pronounced sexual dimorphism related to a derived brooding strategy in females can be observed. We hypothesize that the derived brooding strategy evolved together with a new genetic architecture of sex determination. We crossed a pelvic-brooding (*O. eversi*) with a transfer-brooding species (*O. nigrimas*) to map sex in their F2 generation (intercrosses and backcrosses with both parental species). We found a significant QTL for sex in the intercrosses and the *O. eversi* backcrosses on a linkage group corresponding to *O. latipes* chromosome 24 (OLchr 24). For the *O. nigrimas* backcrosses, we found three potential QTL on OLchr 4, OLchr 20 and OLchr 24, indicating a polygenic sex determining genetic architecture in the transfer brooding species. Our results add to previous studies, where a sex determining region for *O. eversi* was found on OLchr 4, that sex determination might as well be polygenic/polymorphic in the pelvic-brooding species. Nonetheless, the genetic architectures differ between the two

species, leading to the conclusion that the variation in sex determination might have aided the evolution of an adaptive brooding strategy.

Keywords: Sex determination, medaka, sex chromosome, genotyping-in-thousands, pelvic-brooding

Author contributions:

Conceptualization AWN, KT, JS, JMF; Methodology AWN, KT, JMF; Breeding of crosses AWN; Formal analysis JMF; Investigation KT, JMF; Visualization JMF; Data curation JMF, AWN; writing original draft JMF; review & editing all authors

Introduction

The conspicuous dimorphism between males and females of many species is relevant to many fields of biology and the genetic architecture of sex receives a lot of attention by evolutionary biologists (e.g. Marín and Baker 1998; Graves 2006; Mank 2009; Bachtrog et al. 2014). Animals go through two steps to develop into sexually mature individuals: The first is referred to as sex determination, the second sexual differentiation (Lillie, 1939). However, this clear distinction has been diluted because the process should rather be understood as a continuum of several factors influencing each other (Uller & Helanterä, 2011; Heule *et al.*, 2014). Sex determination begins with a set of genes that act as a "master switch" (Heule *et al.*, 2014) and initiate the development of either testis or ovaries (primary sexual characters) (Ford *et al.*, 1959). Sex determining genetic factors may be located on deeply conserved and degenerated sex chromosomes in some lineages such as mammals, insects and birds where sex determining genes maintain their function over long evolutionary times (Bachtrog *et al.*, 2014). On the other hand, a rapid evolutionary turnover of the genetic basis of sex determination has been reported

for fishes (Myosho *et al.*, 2015; El Taher *et al.*, 2021), amphibians (Jeffries *et al.*, 2018), reptiles (Pokornà & Kratochvìl, 2009; Gamble *et al.*, 2015), insects (Blackmon & Demuth, 2014; Vicoso & Bachtrog, 2015) and plants (Balounova *et al.*, 2019; Martin *et al.*, 2019) in which the initial sex determining factors may differ among closely related species. Sex determining genes activate downstream genes regulating for example the hormonal secretions of the gonads (Allen, 1932; Young, 1961) and finally the formation of the sexually dimorphic phenotype by expressing for instance a different body size, ornamentation or color (Paciulli & Cromer, 2018). Most sexually dimorphic traits are the result of sexual selection, and in the *Descent of Man* Darwin (1871) hypothesized that the exaggerated male ornaments he observed in all ranges from insects to mammals are not enhancing survival but are meant to attract females. Sexual selection has also led to the evolution of secondary sexual characters in females, although they have received much less attention (Clutton-Brock, 2007). A conspicuous example are pipefishes. Here, sex-roles are reversed and in some species, females evolved a striking ornamentation, probably due to sexual selection (Jones & Avise, 1997). Ecological adaptation can also drive sexual dimorphism. Examples from dwarf chameleons (Stuart-Fox & Moussali, 2007), hummingbirds (Temeles *et al.*, 2000), damselflies (Cooper, 2010), seabirds (Weimerskirch *et al.*, 2009) and sailfin silversides (Wasiljew *et al.*, 2021) show that males and females can express sexual dimorphism associated with different food resources, habitat preferences or parental care. Therefore, sexually dimorphic characteristics evolve in response to different selection forces that modify parts of the genetic cascade of sexual differentiation in both males and females. Especially in lineages that typically lack well differentiated sex chromosomes, both the key sex determining genes and the initiation of the genetic cascade that follows must change every time the genetic architecture of sex turns over (Heule *et al.*, 2014). Accordingly, the genetic basis for the development of diverged sexual phenotypes may

be evolutionarily stable or subject to frequent evolutionary change in different lineages of animals (reviewed in Palmer et al. 2019).

Knowledge about the genetic processes that are involved in generating phenotypic diversity and evolutionary transitions related to sexual dimorphism is still incomplete. It is conceivable that novel female secondary sexual traits may develop in response to „downstream genes" that are part of a conserved genetic cascade or, alternatively, that changes in the genetic master switch and changes in the genetic cascade could have a direct impact on the downstream genes in lineages where the genetic architecture of sex determination shows rapid evolutionary turnover. To study the roles of key sex determining factors or downstream genes in the evolution of sexually dimorphic traits, it is necessary to identify the suite of genes that are involved in the development of secondary sexual traits. Progress in this field requires study systems in which the evolution of secondary sexual traits can be reconstructed and that are accessible for genetic analysis.

Novel sexually dimorphic traits exaggerated in females and most likely connected with a unique reproductive strategy have been documented for endemic ricefishes in Sulawesi, Indonesia (Kottelat, 1990; Spanke *et al.*, 2021). They occur in two genera, *Oryzias* and *Adrianichthys* that form a monophyletic group and have colonized the island about 16 million years ago (Mokodongan & Yamahira, 2015). The ancestral mode of reproduction involves that eggs are laid and fixed to the urogenital pore for some time before the female deposits them at a suitable substrate and leaves them to themselves (Parenti, 2008). Interestingly, *Adrianichthys oophorus* and the sistertaxa *Oryzias eversi* and *Oryzias sarasinorum* evolved sexually dimorphic adaptations connected with a reproductive strategy called "pelvic-brooding" (Kottelat, 1990). In contrast to transfer-brooding species, the eggs of pelvic-brooding species are fixated in the oviduct of females by attaching filaments originating from the egg surface (Iwamatsu *et al.*,

2008) until the fry hatch. The females' ribs are shortened to form a ventral concavity and their pelvic fins are elongated compared to conspecific males and transfer brooding species (Spanke *et al.*, 2021), potentially to protect the eggs from physical impact, or to improve hydrodynamics of the mother (Spanke et al, in prep). The *Oryzias* system offers the opportunity to study the evolution of sexual dimorphism in closely related species in which a transition from transfer- to pelvic-brooding and associated morphological changes in the female occurred. In contrast to pelvic-brooding *Oryzias* and *Adrianichthys* that have evolved independently from one another after their initial divergence (Montenegro *et al.*, 2022), transfer-brooding and pelvic-brooding species of the genus *Oryzias* apparently share a history of hybridization and introgression (*O. eversi* and *Oryzias nigrimas*, see Flury *et al.,* 2022) which might have affected their evolution. Moreover, in contrast to mammals and birds, where sex determination is deeply conserved, even closely related fishes may display alternative sex determination mechanisms (e.g. Schultheis et al. 2009; Myosho et al. 2015; El Taher et al. 2021). Ansai and colleagues (2022) have demonstrated that this is true for the ricefishes from Sulawesi and identified candidate regions that are likely involved in the development of key female reproductive traits for *O. eversi,* like the pelvic fin length, the form of the concavity and duration of egg carrying. An analysis of the evolutionary changes associated with a novel reproductive strategy in pelvic-brooding ricefishes requires that 1) the functional context of novel female traits is known (compare to chapter 2, Spanke et al. 2021) and 2) that changes to the genetic architecture associated with novel female traits are identified. This study focusses on the second aspect and adds to the recent studies by Ansai et al. (2022) and Montenegro et al. (2022) to investigate the sex determining genetic architecture between closely related species which underwent a major transition from substrate brooding to pelvic-brooding.

Ansai et al. (2022) used an intraspecific genome-wide association study to identify a new sex determining region for the pelvic-brooding *O. eversi*. The region corresponds to *O. latipes* chromosome 4 (OLchr 4). For the sister species *O. sarasinorum*, two loci seem to be involved, one corresponding to OLchr 16 and one on OLchr 22. This contrasts with the majority of ricefishes from Sulawesi that share an ancestral sex determining locus that corresponds to OLchr 24 (present in *O. asinua*, *O. wolasi*, *O. woworae*, *O. celebensis*, *O. dopingdopingensis*, *O. matanensis*, *O. hadiatyae* (Myosho *et al.*, 2015; Ansai *et al.*, 2022)). While particularly relevant due to their possible gene flow with the lineage that gave rise to *O. eversi* and *O. sarasinorum* in the past, the sex determining locus of *O. nigrimas* and other *Oryzias* from lake Poso could not yet be determined. Montenegro et al. (2022) have presented a detailed QTL analysis of interspecies crosses between *O. eversi* and *O. dopingdopingensis* that provides a first inventory of key loci potentially involved in the evolution of pelvic-brooding. They suggest candidate loci that are associated with the development of female traits, one of which localizes on the same chromosome as the ancestral sex determining chromosome 24. Flury *et al.* (2022) showed that said QTL on chromosome 24 overlaps with a region of high introgression signal from the ancestor of the Lake Poso *Oryzias*. Besides, the chromosome holds a high number of the top 5% of D-values observed genome-wide, even though the overall introgression signal is low on this chromosome. This could imply that selection was strong on this chromosome, maybe due to its sex determining function connected to an adaptive complex trait.

In this study, we perform a QTL analysis to identify sex determining factors in an alternative interspecies cross between *O. eversi* and *O. nigrimas*. Our results partially contrast and complement the findings presented by previous authors. We focus on *O. eversi* and *O. nigrimas* because of their genetic exchange in the past and hypothesize (1) that the study species have two different genetic architectures underlying the sex determination and (2) that the sexually

dimorphic traits in the pelvic-brooding species may have evolved together with a new sex determining genetic architecture. We used a customized Genotyping-in-Thousands by sequencing (GT-seq) method (Campbell *et al.*, 2015) resulting in 101-115 genotyped markers in three different second generation cross types (n=567) between *O. eversi* and *O. nigrimas* to create a linkage map and locate genetic regions associated with sex.

Material & Methods

Breeding of fish

Oryzias eversi and *Oryzias nigrimas* individuals were collected from their respective type localities in Sulawesi, Indonesia in 2010 (comp. Herder *et al.*, 2012). Stocks were bred in aquaria at the Max-Planck-Institute for Evolutionary Biology in Plön and the Carl-von-Ossietzky University Oldenburg since 2011. Groups of fish were kept in aquaria decorated with plants at water temperatures between 23°C (*O. eversi*) and 28°C (*O. nigrimas*) and fed with a mixed diet of commercial fish food, *Artemia nauplii* and frozen invertebrates. Juveniles of *O. eversi* were obtained by transferring egg carrying females to separate aquaria filled with plants while eggs of *O. nigrimas* were collected from bundles of synthetic wool into which the females had deposited the eggs. To generate experimental crosses, *O. eversi* males were crossed with *O. nigrimas* females. Three mapping populations were obtained by crossing F1 males with *O. eversi* and *O. nigrimas* females respectively (referred to as *O. eversi* backcrosses and *O. nigrimas* backcrosses). F2 crosses were bred by crossing F1 males with F1 females (referred to as intercrosses here). We noted, that the eggs of F1 females were not deposited to the substrate but also did not attach firmly to the female urogenital pore as in pure *O. eversi*. Accordingly, eggs were lost in the breeding tank unless the eggs were carefully withdrawn from egg carrying females and transferred to glass dishes where they hatched. Juvenile fish were raised until

sexual maturity was reached, sacrificed with an overdose of MS222 and preserved in 70% ethanol for further analysis.

DNA extraction and mapping data set

Two different kits were used to extract the DNA from fin clips. The DNeasy 96 Blood & Tissue kit of Qiagen, following the manufacturer's protocol and the 5-Prime kit of ArchievePure DNA Purification were used following the manufacturer's protocols. For the purpose of mapping the genetic basis of sex, we obtained four pure *O. eversi* and four pure *O. nigrimas* (two of each sex) as reference samples. 174 *O. eversi* backcrosses, 127 *O. nigrimas* backcrosses and 266 F2-intercrosses (F1 x F1) were included for genotyping. The sex of each individual in the dataset was determined by examination of the external morphology of the reproductive fish. This included dorsal and anal fin length (long with extended filaments vs. short with straight edges), coloration (black vs. silvery) and body shape (rounded vs. slender) to identify individuals with clearly assignable sexes for genetic mapping.

Primer Design

We ran the Genotyping-in-Thousands protocol (Campbell *et al.*, 2015) to obtain a marker panel for cost efficient genotyping of QTL mapping families. The method amplifies targeted markers in a multiplex PCR for hundreds or thousands of individuals that are jointly sequenced in an Illumina amplicon sequencing run. In a first PCR step, all loci in the panel are amplified for each individual sample in a multiplex PCR reaction. In a second PCR step, unique barcode pairs that assign all reads to individual samples and sequences required for the next generation sequencing procedure were attached to PCR products from step one. Finally, all PCR Products were pooled to obtain a library that is ready for next generation sequencing. We designed 240 primer pairs for loci that carry a SNP distinguishing *O. eversi* and *O. nigrimas*. Loci were selected such that they matched an *O. latipes* annotated gene and that they were distributed evenly

over the 24 *O. latipes* chromosomes (~10 per chromosome). To find the respective SNPs, de novo transcriptome assemblies of *O. nigrimas* and *O. eversi* were mapped on the *O. latipes* genome available on NCBI (Bioproject PRJNA325079, GenBank accession number: GCA_002234675.1) using NextGenMap (Sedlazeck *et al.*, 2013). SNP calling was performed using samtools mpileup (Handsaker *et al.*, 2009; Li, 2011). BCFtools and VCFtools were used to transform and filter files (Handsaker *et al.*, 2009; Danecek *et al.*, 2011). To ensure medaka genes were also present in *O. eversi* and *O. nigrimas*, the annotated medaka genes were blasted against both the *O. eversi* and *O. nigrimas* de novo transcriptome assemblies (raw data available on NCBI, Bioproject PRJNA827986, assemblies will be published on Dryad after acceptance) using blastn (Zhang *et al.*, 2000). To get evenly spaced markers, the chromosome length was divided by 30 (the targeted number of markers, a surplus of the final number of markers chosen) and markers in the calculated distance where highlighted. The highlighted markers' SNP calling output were checked for indicative SNPs and 250 nucleotides before and after were checked for ambiguous positions to avoid designing a primer sequence targeting such regions. Using the mapped genome and transcriptome assemblies the coverage around the chosen SNPs was checked. If the coverage dropped within 1000 bp in both directions, the marker was dropped. All markers that passed these filtering steps were used for primer design with Primer3 (Untergasser *et al.*, 2012). The designed primers were checked for possible primer-dimer formation in FastPCR 6.6 (Kalendar *et al.*, 2011). Moreover, primer sequences were mapped against the de novo genome assemblies using NextGenMap to see whether they bind to repetitive elements. One primer pair was discarded, because it mapped more than once. Finally, locus specific primer pairs were modified by adding sequence at the 5'-end to allow for the GT seq protocol (Campbell et al. 2015) which involves two PCR steps and adds index

sequences and sequences required for the sequencing procedure (see Table S1 for primers used in this study). Primer Oligos were ordered from Metabion.

Multiplex PCR I

In this step, the selected markers were amplified in a multiplex PCR for each individual DNA sample. The Qiagen Multiplex PCR kit was used with 178 self-designed Primer pairs (Table S1) and 576 samples (Table S2). Per sample 5 µl 2x QIAGEN Multiplex PCR Master Mix, 1 µl 5x Q-Solution, 2 µl H_2O, 1µl of Primer Mix with a concentration of 200 nmol/L per Primer and 1 µl template DNA with a concentration between 20 and 100 ng/µl. In the thermal cycler, initial heat activation took 15 minutes at 95°C, followed by 6 cycles with 30 seconds 94°C for denaturation, 3 minutes at 60° for annealing and 15 seconds at 72° for extension. Next, we ran 25 cycles using 30 seconds at 94°C for denaturation, 3 minutes at 68°C for annealing and 30 seconds at 72°C for extension. We added a final extension for 10 minutes at 72°C.

Clean-up

The multiplex PCR product was cleaned-up using ExoSAP-IT. ExoSAP-IT is used to degrade remaining primers and nucleotides. We added 4µl of ExoSAP-IT to each post-PCR sample for 30 minutes at 37°C. To inactivate ExoSAP-IT, the plate was placed in a thermal cycler at 80°C for 15 minutes and stored at -20°C until further usage. After the ExoSAP-IT treatment, each sample had a volume of 12 µl. Performance of the protocol turned out to be best when PCR Products were further cleaned using Ampure magnetic Beads. 1.6x volumes of magnetic beads were added to each sample. Reagent and sample were mixed by pipetting up and down for at least 10 times and the samples were incubated for 5 minutes at room temperature. The plate was placed on a Agencourt SPRIPlate 96 Super Magnet Plate for 2 minutes to separate the beads from the solution. The supernatant was aspirated from the reaction plate and discarded without drawing out beads. While still on the magnetic plate, 150µl of fresh 70% ethanol was

added to each well. After 30 seconds, the ethanol was carefully removed and the wash step was repeated. After removing all ethanol, the beads were left to dry. When dry, the plate was removed from the magnetic rack and 40µl of H20 molecular grade was added and mixed by pipetting up and down for at least 10 times. After incubation of 2 minutes at room temperature, the plate was put back on the magnetic rack and left for 1 minute for the beads to separate from the solution. Finally, the 40µl eluate was removed and transferred to a new 96 plate.

Indexing PCR II

In this step, the barcodes labelling individual PCR Products and sequences required for next generation sequencing were attached to the amplified marker fragments. For each PCR reaction (step II), we used 1 µl of 1:10 diluted PCR I product, 5 µl 2x QIAGEN Multiplex PCR Master Mix, 1 µl of i7 Primer (2 µM/L), 1 µl of i5 primer (2 µM/L), 1 µl 5x Q-Solution and 1µl H20. In the thermal cycler, initial heat activation was done for 15 minutes at 95°C, followed by 6 cycles with 30 seconds at 94°C for denaturation, 3 minutes at 60° for annealing and 15 seconds at 72° for extension. In contrast to the first PCR, only 15 cycles were run for the second step, using 30 seconds at 94°C for denaturation, 3 minutes at 68°C for annealing and 30 seconds at 72°C for extension. After a final extension for 10 minutes at 72°C samples were cooled and frozen before further steps.

PCR product cleanup from agarose - Crush and Soak

The target PCR products were extracted from agarose to reduce the amount of PCR artefacts carried over into the Illumina sequencing procedure. PCR II products sere separated in a 4% Phor Agaraose Gel that is efficient in separating fragments of low size (<200 bp) for one hour at 147V. DNA was excised from the gel at the desired length and the resulting gel pieces were put into 2ml tubes and the gel was crushed using a metal spatula. 600 µl of TE low (10mM Tris + 0.1 mM EDTA) were added to each tube and incubated at room temperature while rotating

slowly on a rotator for 24 hours. After incubation for 24 hours the samples were centrifuged at 20,000 g for one minute. 350 µl of the supernatant was transferred to a new 2ml tube. A second volume of 350 µl TE low was added to the crushed gel and the tube was vortexed. The tubes were again centrifuged at 20,000 g and another 350 µl of supernatant was withdrawn to obtain a final volume of 700 µl supernatant. PCR products were precipitated using 70 µl 3 mol pH 5.5 sodium acetate, 1 µl glycogen and 770 µl of isopropanol. The samples were mixed well, incubated at room temperature for 15 minutes and then centrifuged for ten minutes at 20,000 g at 4°C. The supernatant was removed carefully, pellets were air-dried and resuspended in 20 µl molecular grade H_2O. At the end, all samples were pooled and represent a ready-to-use library that was sent off for Illumina sequencing.

SNP calling

The library was run on two lanes to reach a higher coverage (Illumina platform NovaSeq 6000). Reads were demultiplexed by the sequencing center based on the list of index sequences used by us. All reads were quality filtered using fastq_quality_filter from FASTX-Toolkit v 0.0.14 (Assaf & Hannon, 2010) (minimal quality of 10 in 100% or reads and minimal quality of 30 in 95% of the reads) and trimmed using Trimmomatic 0.36 (Bolger *et al.*, 2014). After trimming, the reads were mapped on a de novo genome assembly of *O. eversi* (will be submitted to Dryad after acceptance) using NextGenMap v 0.5.5. The Stacks pipeline using ref_map.pl was used to call the genotypes and VCFtools v 0.1.15 was used to filter the SNPs for only keeping variants successfully genotyped in >80% of the individuals, with a minimal quality score of 15 and minimum depth of 20. After this filtering, individuals with more than 20% missing data were removed. To prepare the VCF file for JoinMap 4.1, BCFtools v 1.6 was used to only keep variants that were different and homozygous in reference samples of *O. eversi* and *O.nigrimas*. We

prepared separate VCF files for each cross (backcrosses and intercrosses) and the parental species containing the genotypes at each locus.

Linkage map construction

For the construction of the linkage map in JoinMap 4.1 (Van Ooijen, 2006), SNPs with signs for extreme segregation distortion were removed (p-value < 0.001). Identical loci and individuals were also removed. The linkage map of the intercrosses was based on 266 individuals (71 females, 192 males), *O. eversi* backcrosses 174 (71 females, 103 males), *O. nigrimas* backcrosses 127 (25 females, 102 males). To build the linkage maps, the Kosambi mapping function was used (LOD threshold 1.0, recombination threshold 0.4, goodness-of-fit threshold 5.0, no fixed order).

QTL mapping of sex

The R-package "R/qtl" (Broman *et al.*, 2003) was used for the QTL mapping. We used the calc.genoprob function with a fixed step-size of 1 cM and assumed genotyping error of 0.01 to calculate conditional genotypes. We tested both, maximum likelihood via the EM algorithm (Lander & Botstein, 1989) and the Haley-Knott regression (Haley & Knott, 1992). To get a genome-wide LOD significance threshold, we used a 1000 permutation test for the Haley-Knott regression using a 0.05 alpha value (Broman & Sen, 2009).

Synteny with O. latipes *and assessment of map quality*

Assuming a well-conserved synteny among all ricefishes with the model species *O. latipes*, we based our marker development on known genetic distances from *O. latipes* and used the available high-quality genome of that species to assess the quality of the genetic maps we have reconstructed. Unambiguous positions of the markers used in this study on the medaka genome were identified using searches with blast+ 2.9.0 (Camacho *et al.*, 2009). SNP markers for the genetic mapping panel were chosen only if they have a known position on the *O. latipes*

reference genome and alleles were species-specific for *O. eversi* and *O. nigrimas*. Moreover, markers were chosen to be evenly spaced across the genome. This marker panel permitted a direct comparison among the maps generated here with the high-quality reference genome to assess the quality of our maps and infer synteny relationships. We used Circos v0.69 (Krzywinski *et al.*, 2009) to visualize maps and conservation of synteny.

Results

Mapping population and sex ratios

Juveniles from all types of crosses developed into clearly discernible males and females, but sex ratios were biased conspicuously towards males. This bias became more pronounced as the overall genetic background in the crosses shifted from mostly *O. eversi* towards *O. nigrimas*. In *O. eversi* backcrosses (overall 75% *O. eversi* ancestry) the sex ration was 1,6 males: 1 female (total offspring n 185), in intercrosses (50% *O. eversi* ancestry) it was 3,9: 1 (n 189) and in *O. nigrimas* backcrosses (25% *O. eversi* ancestry) it was 7,3: 1 (n 259). The small number of available females in the latter cross limited the genetic mapping analysis.

Sex determination

We mapped a sex QTL in the intercrosses and in the *O. eversi* backcrosses (LOD IC 5.5, LOD BCE 16.9) (Fig. 1). In both crosses, the sex determination was found on a linkage group corresponding to the *O. latipes* chromosome 24. In the intercrosses an additional peak below significance level was found at *O. latipes* chromosome 21. None of the peaks in the *O. nigrimas* backcrosses reached statistical significance, but three clearly visible peaks below significance were found at *O. latipes* chromosomes 4, 20 and 24.

Linkage maps

The linkage map of the intercrosses was based on 110 markers from 74 females and 192 males. It has a length of 756.86 cM and an average marker distance of 6.88 cM (Fig. 2). The *O. eversi* backcrosses' linkage map had a length of 777.84 cM with an average marker distance of 6.76 cM and was based on 115 markers from 71 females and 103 males. The shortest map was the *O. nigrimas* backcrosses' map with 616.89 CM and an average marker distance of 6.10 cM based on 101 markers from 25 females and 102 males (Fig. 3, Table 1). In the intercrosses an LOD = 4 was most suitable. It resulted in 21 groups, one group only contained one marker and in one containing three markers no linkage was found. In the *O. eversi* backcrosses also an LOD = 4 was used, what resulted in 22 groups, 2 with only 1 locus. In the *O. nigrimas* backcrosses an LOD = 5 was most reasonable, resulting in 22 groups, with 6 groups having only 1 marker. High density linkage maps of *O. latipes* and *O. melastigma* have 1354.5 cM respectively 1783.97 cM while having both 24 chromosomes (Naruse *et al.*, 2000; Lee *et al.*, 2019).

Synteny with O. latipes *and comparison to other Sulawesi ricefish linkage maps*

As the linkage map of the intercrosses was based on the largest number of individuals and recombination events, we compared their map to the *O. latipes* map (Naruse *et al.*, 2000) (Fig. 4). As the preparation of the markers aimed to represent loci on the medaka chromosomes, we could reconstruct 19 out of 24 chromosomes in our linkage map, and 4 of the found LGs contained loci from more than one medaka chromosome. LG13 also contained a marker from medaka chromosome 22. LG19 contained markers from medaka chromosomes 17 and 2. LG20 contained markers from medaka chromosome 12 and 11, and LG23 also contained markers from medaka chromosome 6. Only medaka chromosome 3 was not represented, as markers corresponding to it had only insufficient linkage. Comparing the linkage map from this study (Fig. 2) to the linkage maps based on F2 intercrosses between *O. eversi* and *O.*

dopingdopingensis (Montenegro *et al.*, 2022) and on F2 intercrosses between *O. celebensis* and *O. woworae* (Ansai *et al.*, 2021), we find that OLchr 6 and 23 are always fused. Chromosomes 17 and 19 are also fused, in our study additionally a marker mapping on chromosome 2 is part of this linkage group. In Ansai *et al.* (2021), chromosome 17 and 19 are fused with chromosome 1. Chromosome 20 is fused with different other chromosomes in all studies: In crosses between *O. eversi* and *O. dopingdopingensis* chromosome 20 is fused with 12, in crosses between *O. eversi* and *O. nigrimas* with chromosomes 11 and 12, in crosses between *O. celebensis* and *O. woworae* with 12 and 13. A marker mapped on chromosome 22 is further linked with chromosome 13 in this study, in Ansai *et al.* (2021) chromosome 22 is fused with chromosome 5.

Discussion

Based on linkage maps that are highly congruent with published maps of Sulawesi ricefishes (Ansai *et al.*, 2021; Montenegro *et al.*, 2022), as well as the medaka genome (Kasahara *et al.*, 2007), we analysed the genetic architecture of sex determination in crosses between pelvic-brooding *O. eversi* and transfer-brooding *O. nigrimas*. Our study differs from, but not necessarily contradicts, published results, and add first data on candidate regions that determine sexes in *O. nigrimas*. We found a single significant QTL for sex in the *O. eversi* backcrosses and the intercrosses on a linkage group corresponding to chromosome 24 in *O. latipes* which differs from the study by Ansai *et al.* (2022), who identified OLchr 4 as single major factor determining sex in *O. eversi*. In the *O. nigrimas* backcrosses, we found three potential QT loci on chromosomes 4, 20 and 24, however below significance threshold. This parallels the study of Ansai *et al.* (2022) who did not detect a significant region associated with sex for pure *O. nigrimas*. Nevertheless, we found conspicuous peaks that, if confirmed, may suggest the

action of multiple loci. This would also contribute to the difficulties to map strong genetic factors in a comparatively smaller number of females than in the other cross types. These results support that the master switch determining sexes must have changed together with the evolution of a new reproductive strategy and novel secondary sexually dimorphic traits in *O. eversi*.

A role of O. latipes *chr 24 in sex determination in O. eversi?*

We found the same sex determining locus in the *O. eversi* backcrosses and in the intercrosses on a linkage group corresponding to *O. latipes* chromosome 24. This locus is the sex determination locus of seven other Sulawesi ricefishes (Myosho *et al.*, 2015; Ansai *et al.*, 2022) (Fig. 1) and was recently associated with a QTL of pelvic fin length. However, the QTL was not located within the genomic region showing a high male-female divergence (Ansai *et al.*, 2022; Montenegro *et al.*, 2022). Intriguingly, we found evidence that *O. latipes* chromosome 24 could be subject to ancient gene flow among the Lake Poso *Oryzias* and the lineage that gave rise to *O. eversi* (Flury *et al.,* 2022). Our QTL analysis does not map candidate loci finely enough to draw conclusions about introgression at the causative loci, but we did note a signal of introgression in the likely candidate genes *sox7*, *pgm3/PGM* and *hsdl1* detected by Ansai *et al.* (2022) (see Flury *et al.,* 2022, positions 16016-16075 on CHR24). Hence our analysis together with the results presented in Flury *et al.* (2022) add evidence that introgression might have affected the evolution of the sex determining genetic architectures between *O. eversi* and *O. nigrimas*, which warrants further study.

On the other hand, there is a suggested sex determining locus of *O. eversi* on the linkage group corresponding to chromosome 4 (Ansai *et al.*, 2022) close to a candidate sex determination gene *amh*. It encodes anti-Müllerian hormone which is involved in gonad formation maintenance in both sexes in medaka fishes (Klüver *et al.*, 2007), and it has also been

documented as sex determination locus in other teleost fishes (Hattori *et al.*, 2012; Pan *et al.*, 2019; Peichel *et al.*, 2020; Song *et al.*, 2021). In our study, we did not observe a signal for this locus in the *O. eversi* backcrosses or the intercrosses, but in the *O. nigrimas* backcrosses this locus on *O. latipes* chromosome 4 had a peak (Fig. 1). The presence of two different sex determining loci in the two studies could be explained by effects of hybridization in the present study. As a sex determination locus on chromosome 24 has been known for several other Sulawesi ricefishes (Myosho *et al.*, 2015; Ansai *et al.*, 2022) it could be that the functionality of the potential new master switch on chromosome 4 was impeded due to dysfunction of the respective alleles in the hybrid genetic background, leading to the reactivation of the ancestral sex determining locus. However, this is not very likely, as we obtained a particularly strong and clear signal for a QTL in the *O. eversi* backcrosses with a mostly *O. eversi* genetic background in which no signs for the action of a locus on OLchr 4 could be detected. We note that, based on the comparison of our linkage maps (Fig. 2, Fig. 3) and linkage maps of other studies (Ansai *et al.*, 2021; Montenegro *et al.*, 2022), we did not observe chromosomal rearrangements involving chromosome 4 or 24 within the *Oryzias* from Sulawesi, which could affect patterns of segregation or have other effects in hybrid crosses. A second possibility is the evolution of new sex determining loci in the aquarium, as was already observed in zebrafish (Liew *et al.*, 2012; Wilson *et al.*, 2014) or that a polymorphic/polygenic sex determination already occurs in natural populations. The type locality of *O. eversi* in Tana Toraja, Central Sulawesi, is highly threatened by anthropogenic disturbance and the number of adults was already low at the time of discovery and capture of our aquarium stocks. Therefore, probably only a few individuals were caught for breeding in the aquarium for both studies and our strain of *O. eversi* has been kept in captivity since the species was discovered and described (Herder *et al.*, 2012). Accordingly, our strain or the strain used by Ansai *et al.* (2022) would have diverged in their

genetic architecture in isolated and genetically limited aquarium populations. This might have been facilitated by a polymorphic/polygenic sex determination in the wild population that could have been differentially sampled in the two studies. To identify all loci involved in sex determination in *O. eversi*, a broader sample from the wild would be necessary.

Potentially polygenic sex determination in O. nigrimas

We found three potential QTLs for sex in the *O. nigrimas* backcrosses. Due to shifted sex ratios in the hybrid crosses, we did not have a large number of females available in this cross type (1 female to 10 males) which might have led to lower statistical power. Adding on the results of Ansai *et al.* (2022), where no locus was clearly identified, we assume that sex determination in *O. nigrimas* might be polygenic. Polygenic sex determination was thought to be unstable and evolutionarily transitory, however it has been documented for several species (Moore & Roberts, 2013; Alexander *et al.*, 2015), including fishes (Kosswig, 1964; Volff & Schartl, 2001; Vandeputte *et al.*, 2007; Roberts *et al.*, 2016), that it can very well be a stable evolutionary strategy. However, the maintenance of polygenic sex determination is poorly understood. Future studies should include a larger number of samples from the wild, to better cover the variation in sex determination and be able to detect QTLs of smaller effect in *O. nigrimas*.

Conclusion

We do not think that the study of Ansai *et al.* (2022) or the analysis we present here are necessarily conflicting. The study of Ansai *et al.* (2022) is convincing, because they used pure *O. eversi* and *O. nigrimas* samples which excludes possible effects of interspecific hybridization. However, the strength of this study is that we included several different hybrid crosses, from a 75% *O. eversi* genetic background to 75% *O. nigrimas* genetic background and found congruent and strong results for the sex determining region of *O. eversi* in the *O. eversi* backcrosses, and

the same locus was detected in the intercrosses with 50% *O. eversi* genetic background. We consider it more likely that all three studies (including Montenegro *et al.,* 2022) identified components of the sex determining architecture, albeit with the limitation that it remains unclear if a single genetic architecture exist or if *O.eversi* (and *O.nigrimas*) may be polymorphic/polygenic. Sadly, the population of *O. eversi* in the Tilanga pool (type locality) is highly threatened by habitat destruction and pollution caused by humans (IUCN, 2019), and it is likely that the population already underwent severe bottlenecks. Therefore, the influence of genetic drift on their genetic sex determination might have overridden the influence of natural selection, which would make it difficult to draw conclusions on the context of the evolution of pelvic-brooding and their genetic architecture of sex determination. However, future studies might shed light on the connection between sex determination and pelvic-brooding traits in *O. eversi* by mapping them in the same crosses simultaneously.

Acknowledgements

We would like to thank Malte Dittmann for his great work in the molecular lab and fish care. Further we thank Simon Wietke and Raphael Schween, who contributed to establishing laboratory protocols. This work was supported by an ERC Starting Grant “Evolmapping” to A.N., the Max-Planck Institute for Evolutionary Biology (Plön, Germany) and the Carl von Ossietzky University (Oldenburg, Germany) and by a Leibniz Association grant P91/ 2016 to J.S. and the Museum Koenig (Bonn, Germany).

References

Alexander, H.J., Richardson, J.M.L., Edmands, S. & Anholt, B.R. 2015. Sex without sex chromosomes: genetic architecture of multiple loci independently segregating to determine sex ratios in the copepod *Tigriopus californicus*. *J. Evol. Biol.* **28**: 2196–2207.

Allen, E. 1932. *Sex and internal secretions: a survey of recent research.* Williams and Wilkins Co, Baltimore.

Ansai, S., Mochida, K., Fujimoto, S., Mokodongan, D.F., Sumarto, B.K.A., Masengi, K.W.A., *et al.* 2021. Genome editing reveals fitness effects of a gene for sexual dichromatism in Sulawesian fishes. *Nat. Commun.* **12**: 1350.

Ansai, S., Montenegro, J., Masengi, K.W.A., Nagano, A.J., Yamahira, K. & Kitano, J. 2022. Diversity of sex chromosomes in Sulawesian medaka fishes. *bioRxiv*. https://doi.org/10.1101/2022.02.28.482394

Assaf, G. & Hannon, G.J. 2010. FASTX-toolkit.

Bachtrog, D., Mank, J.E., Peichel, C.L., Kirkpatrick, M., Otto, S.P., Ashman, T.-L., *et al.* 2014. Sex determination: why so many ways of doing it? *PLoS Biol.* **12**: e1001899.

Balounova, V., Gogela, R., Cegan, R., Cangren, P., Zluvova, J., Safar, J., *et al.* 2019. Evolution of sex determination and heterogamety changes in section *Otites* of the genus *Silene*. *Sci. Rep.* **9**: 1045.

Blackmon, H. & Demuth, J.P. 2014. Estimating tempo and mode of Y chromosome turnover: explaining Y chromosome loss with the fragile Y hypothesis. *Genetics* **197**: 561–572.

Bolger, A.M., Lohse, M. & Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Broman, K. & Sen, S. 2009. *A guide to QTL mapping with R/qtl*. Springer.

Broman, K.W., Wu, H., Sen, Ś. & Churchill, G.A. 2003. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19**: 889–890.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., *et al.* 2009. BLAST+: architecture and applications. *BMC Bioinformatics* **10**: 421.

Campbell, N.R., Harmon, S.A. & Narum, S.R. 2015. Genotyping-in-Thousands by sequencing (GT-seq): a cost effective SNP genotyping method based on custom amplicon sequencing. *Mol. Ecol. Resour.* **15**: 855–867.

Clutton-Brock, T.H. 2007. Sexual selection in males and females. *Science* **318**: 1882–1885.

Cooper, I.A. 2010. Ecology of sexual dimorphism and clinal variation of coloration in a damselfly. *Am. Nat.* **176**: 566–572.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., *et al.* 2011. The variant call format and VCFtools. *Bioinformatics* **27**: 2156-2158.

Darwin, C. 1871. *The descent of man and selection in relation to sex*. Modern Library, New York.

El Taher, A., Ronco, F., Matschiner, M., Salzburger, W. & Böhne, A. 2021. Dynamics of sex chromosome evolution in a rapid radiation of cichlid fishes. *Sci. Adv.* **7**: eabe8215.

Flury, J.M., Meusemann, K., Martin, S., Hilgers, L., Spanke, T., Böhne, A., Herder, F., Mokodongan, D. F., Altmüller, J., Wowor, D., Misof, B., Nolte, A. W., Schwarzer, J. 2022. Contribution of hybridization to the evolution of a derived reproductive strategy in ricefishes. bioRxiv. https://doi.org/10.1101/2022.07.05.498713

Ford, C., Jones, K., Polani, P. & Almeida, J. De. 1959. A sex-chromosome anomaly in a case of gonadal dysgenesis (Turner's syndrome). *Lancet* **1**: 711–713.

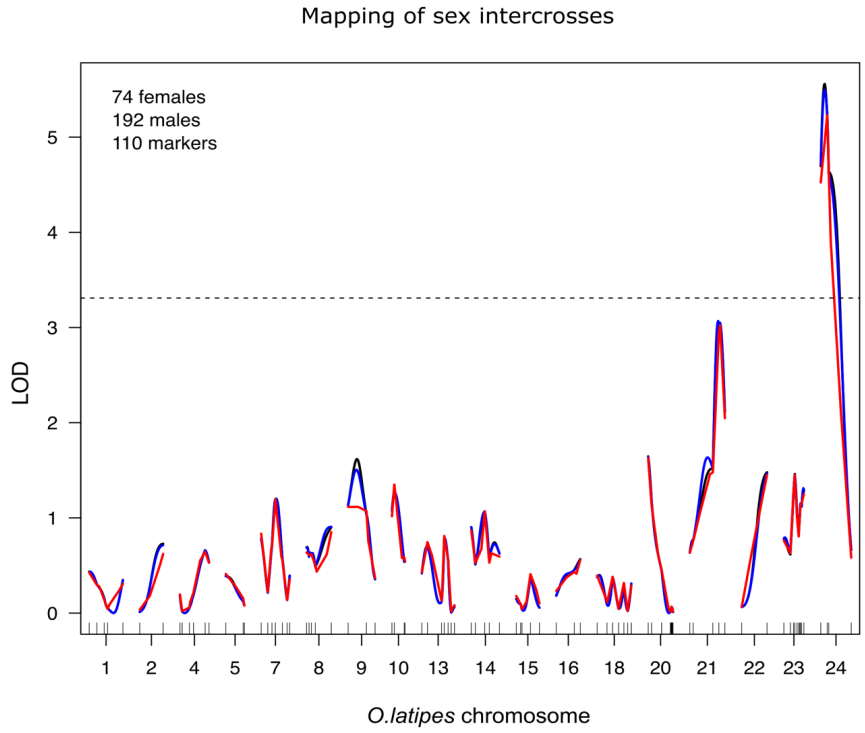
Gamble, T., Coryell, J., Ezaz, T., Lynch, J., Scantlebury, D.P. & Zarkower, D. 2015. Restriction site-associated DNA sequencing (RAD-seq) reveals an extraordinary number of

transitions among gecko sex-determining systems. *Mol. Biol. Evol.* **32**: 1296–1309.
Graves, J.A.M. 2006. Sex chromosome specialization and degeneration in mammals. *Cell* **124**: 901–914.
Haley, C.S. & Knott, S.A. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity (Edinb).* **69**: 315–324.
Handsaker, Wysoker, Fennell, Ruan, Homer, Abecasis, *et al.* 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**: 2078–2079.
Hattori, R.S., Murai, Y., Oura, M., Masuda, S., Majhi, S.K., Sakamoto, T., *et al.* 2012. A Y-linked anti-Müllerian hormone duplication takes over a critical role in sex determination. *Proc. Natl. Acad. Sci. U. S. A.* **109**: 2955–9.
Herder, F., Hadiaty, R.K. & Nolte, A.W. 2012. Pelvic-fin brooding in a new species of riverine ricefish (Atherinomorpha: Beloniformes: Adrianichthyidae) from Tana Toraja, Central Sulawesi, Indonesia. *Raffles Bull. Zool.* **60**: 467–476.
Heule, C., Salzburger, W. & Böhne, A. 2014. Genetics of sexual development: an evolutionary playground for fish. *Genetics* **196**: 579–591.
IUCN. 2019. The IUCN Red List of Threatened Species.
Iwamatsu, T., Kobayashi, H., Sato, M. & Yamashita, M. 2008. Reproductive role of attaching filaments on the egg envelope in *Xenopoecilus sarasinorum* (Adrianichthidae, Teleostei). *J. Morphol.* **269**: 745–750.
Jeffries, D.L., Lavanchy, G., Sermier, R., Sredl, M.J., Miura, I., Borzée, A., *et al.* 2018. A rapid rate of sex-chromosome turnover and non-random transitions in true frogs. *Nat. Commun.* **9**: 4088.
Jones, A.G. & Avise, J.C. 1997. Microsatellite analysis of maternity and the mating system in the Gulf pipefish *Syngnathus scovelli*, a species with male pregnancy and sex-role reversal. *Mol. Ecol.* **6**: 203–213.
Kalendar, R., Lee, D. & Schulman, A.H. 2011. Java web tools for PCR, in silico PCR, and oligonucleotide assembly and analysis. *Genomics* **98**: 137–144.
Kasahara, M., Naruse, K., Sasaki, S., Nakatani, Y., Qu, W., Ahsan, B., *et al.* 2007. The medaka draft genome and insights into vertebrate genome evolution. *Nature* **447**: 714–719.
Klüver, N., Pfennig, F., Pala, I., Storch, K., Schlieder, M., Froschauer, A., *et al.* 2007. Differential expression of anti-Müllerian hormone (amh) and anti-Müllerian hormone receptor type II (amhrII) in the teleost medaka. *Dev. Dyn.* **236**: 271–281.
Kosswig, C. 1964. Polygenic sex determination. *Experientia* **20**: 190–199.
Kottelat, M. 1990. Synopsis of the endangered Buntingi (Osteichthyes: Adranichtyidae and Oryziidae) of Lake Poso, Central Sulawesi, Indonesia, with a new reproductive guild and descriptions of three new species. *Ichthyol. Explor. Freshwaters* **1**: 46–67.
Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., *et al.* 2009. Circos: an information aesthetic for comparative genomics. *Genome Res.* **19**: 1639–45.
Lander, E.S. & Botstein, S. 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**: 185.
Lee, B.-Y., Kim, M.-S., Choi, B.-S., Nagano, A.J., Wai, D., Au, T., *et al.* 2019. Construction of high-resolution RAD-seq based linkage map, anchoring reference genome, and QTL Mapping of the sex chromosome in the marine medaka *oryzias melastigma*. *G3 Genes|Genomics|Genetics* **9**: 3537-3545.
Li, H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**: 2987–93.
Liew, W.C., Bartfai, R., Lim, Z., Sreenivasan, R., Siegfried, K.R. & Orban, L. 2012. Polygenic sex

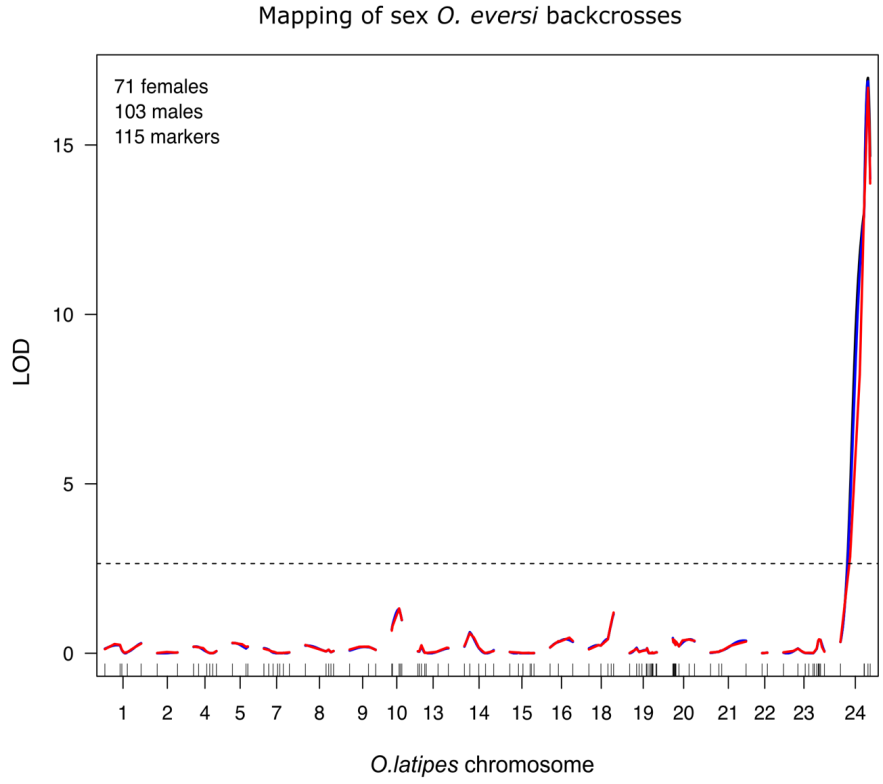
determination system in zebrafish. *PLoS One* **7**: e34397.
Lillie, F.R. 1939. General biological introduction. In: *Sex and internal secretions* (E. Allen, C. Danforth, & A. Doisy, eds), pp. 3–14. Williams and Wilkins Co, Baltimore.
Mank, J.E. 2009. Sex chromosomes and the evolution of sexual dimorphism: lessons from the genome. *Am. Nat.* **173**: 141–150. The University of Chicago Press.
Marín, I. & Baker, B.S. 1998. The evolutionary dynamics of sex determination. *Science* **281**: 1990–4.
Martin, H., Carpentier, F., Gallina, S., Godé, C., Schmitt, E., Muyle, A., *et al.* 2019. Evolution of young sex chromosomes in two dioecious sister plant species with distinct sex determination systems. *Genome Biol. Evol.* **11**: 350–361.
Mokodongan, D.F. & Yamahira, K. 2015. Origin and intra-island diversification of Sulawesi endemic Adrianichthyidae. *Mol. Phylogenet. Evol.* **93**: 150–160.
Montenegro, J., Fujimoto, S., Ansai, S., Nagano, A.J., Sato, M., Maeda, Y., *et al.* 2022. Genetic basis for the evolution of pelvic-fin brooding, a new mode of reproduction, in a Sulawesian fish. *Mol. Ecol.*, doi: 10.1111/mec.16555.
Moore, E.C. & Roberts, R.B. 2013. Polygenic sex determination. *Curr. Biol.* **23**: 510–512.
Myosho, T., Takehana, Y., Hamaguchi, S. & Sakaizumi, M. 2015. Turnover of sex chromosomes in Celebensis group medaka fishes. *G3 Genes|Genomes|Genetics* **5**: 2685–2691.
Naruse, K., Fukamachi, S., Mitani, H., Kondo, M., Matsuoka, T., Kondo, S., *et al.* 2000. A detailed linkage map of medaka, *Oryzias latipes*: comparative genomics and genome evolution. *Genetics* **154**: 1773–84.
Paciulli, L.M. & Cromer, C.M. 2018. Secondary sex characteristics. In: *Encyclopedia of Animal Cognition and Behavior* (J. Vonk & T. Shackelford, eds), pp. 1–9. Springer International Publishing, Cham.
Palmer, D.H., Rogers, T.F., Dean, R. & Wright, A.E. 2019. How to identify sex chromosomes and their turnover. *Mol. Ecol.* **28**: 4709–4724.
Pan, Q., Feron, R., Yano, A., Guyomard, R., Jouanno, E., Vigouroux, E., *et al.* 2019. Identification of the master sex determining gene in Northern pike (*Esox lucius*) reveals restricted sex chromosome differentiation. *PLOS Genet.* **15**: e1008013.
Parenti, L.R. 2008. A phylogenetic analysis and taxonomic revision of rice fishes, *Oryzias* and relatives (Beloniformes, Adrianichthydae). *Zool. J. Linn. Soc.* **154**: 494–610.
Peichel, C.L., McCann, S.R., Ross, J.A., Naftaly, A.F.S., Urton, J.R., Cech, J.N., *et al.* 2020. Assembly of the threespine stickleback Y chromosome reveals convergent signatures of sex chromosome evolution. *Genome Biol.* **21**: 177.
Pokornà, M. & Kratochvìl, L. 2009. Phylogeny of sex-determining mechanisms in squamate reptiles: are sex chromosomes an evolutionary trap? *Zool. J. Linn. Soc.* **156**: 168–183.
Roberts, N.B., Juntti, S.A., Coyle, K.P., Dumont, B.L., Stanley, M.K., Ryan, A.Q., *et al.* 2016. Polygenic sex determination in the cichlid fish *Astatotilapia burtoni*. *BMC Genomics* **17**: 835.
Schultheis, C., Böhne, A., Schartl, M., Volff, J.N. & Galiana-Arnoux, D. 2009. Sex determination diversity and sex chromosome evolution in poeciliid fish. *Sex. Dev.* **3**: 68–77.
Sedlazeck, F.J., Rescheneder, P. & Von Haeseler, A. 2013. NextGenMap: fast and accurate read mapping in highly polymorphic genomes. *Bioinformatics* **29**: 2790-2791.
Song, W., Xie, Y., Sun, M., Li, X., Fitzpatrick, C.K., Vaux, F., *et al.* 2021. A duplicated *amh* is the master sex-determining gene for Sebastes rockfish in the Northwest Pacific. *Open Biol.* **11**: 210063.
Spanke, T., Hilgers, L., Wipfler, B., Flury, J.M., Nolte, A.W., Utama, I. V., *et al.* 2021. Complex sexually dimorphic traits shape the parallel evolution of a novel reproductive strategy in

Sulawesi ricefishes (Adrianichthyidae). *BMC Ecol. Evol.* **21**: 57.
Stuart-Fox, D. & Moussali, A. 2007. Sex-specific ecomorphological variation and the evolution of sexual dimorphism in dwarf chameleons (*Bradypodion* spp.). *J. Evol. Biol.* **20**: 1073–1081.
Temeles, E.J., Pan, I.L., Brennan, J.L. & Horwitt, J.N. 2000. Evidence for ecological causation of sexual dimorphism in a hummingbird. *Science* **289**: 441–443.
Uller, T. & Helanterä, H. 2011. From the origin of sex-determining factors to the evolution of sex-determining systems. *Q. Rev. Biol.* **86**: 163–180.
Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., *et al.* 2012. Primer3--new capabilities and interfaces. *Nucleic Acids Res.* **40**: e115.
Van Ooijen, J.W. 2006. JoinMap 4: Software for the calculation of genetic linkage maps in experimental populations. Kyazma BV, Wageningen, Netherlands.
Vandeputte, M., Dupont-Nivet, M., Chavanne, H. & Chatain, B. 2007. A polygenic hypothesis for sex determination in the European sea bass *Dicentrarchus labrax*. *Genetics* **176**: 1049–1057.
Vicoso, B. & Bachtrog, D. 2015. Numerous transitions of sex chromosomes in diptera. *PLOS Biol.* **13**: e1002078.
Volff, J.-N. & Schartl, M. 2001. Variability of genetic sex determination in poeciliid fishes. *Genetica* **111**: 101–110.
Wasiljew, B.D., Pfaender, J., Wipfler, B., Gabelaia, M., Utama, I.V., Wantania, L.L., *et al.* 2021. Sexual dimorphism in an adaptive radiation: does intersexual niche differentiation result in ecological character displacement? *Ecol. Evol.* **11**: 14615–14629.
Weimerskirch, H., Le Corre, M., Gadenne, H., Pinaud, D., Kato, A., Ropert-Coudert, Y., *et al.* 2009. Relationship between reversed sexual dimorphism, breeding investment and foraging ecology in a pelagic seabird, the masked booby. *Oecologia* **161**: 637–649.
Wilson, C.A., High, S.K., McCluskey, B.M., Amores, A., Yan, Y., Titus, T.A., *et al.* 2014. Wild sex in zebrafish: loss of the natural sex determinant in domesticated strains. *Genetics* **198**: 1291–1308.
Young, W. 1961. *Sex and internal secretions*. Williams and Wilkins Co, Baltimore.
Zhang, Z., Schwartz, S., Wagner, L. & Miller, W. 2000. A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* **7**: 203–214.

A


Mapping of sex intercrosses
74 females
192 males
110 markers
LOD
0
1
2
3
4
5
1 2 4 5 7 8 9 10 13 14 15 16 18 20 21 22 23 24
O.latipes chromosome

B


Mapping of sex O. eversi backcrosses
71 females
103 males
115 markers
LOD
0
5
10
15
1 2 4 5 7 8 9 10 13 14 15 16 18 19 20 21 22 23 24
O.latipes chromosome

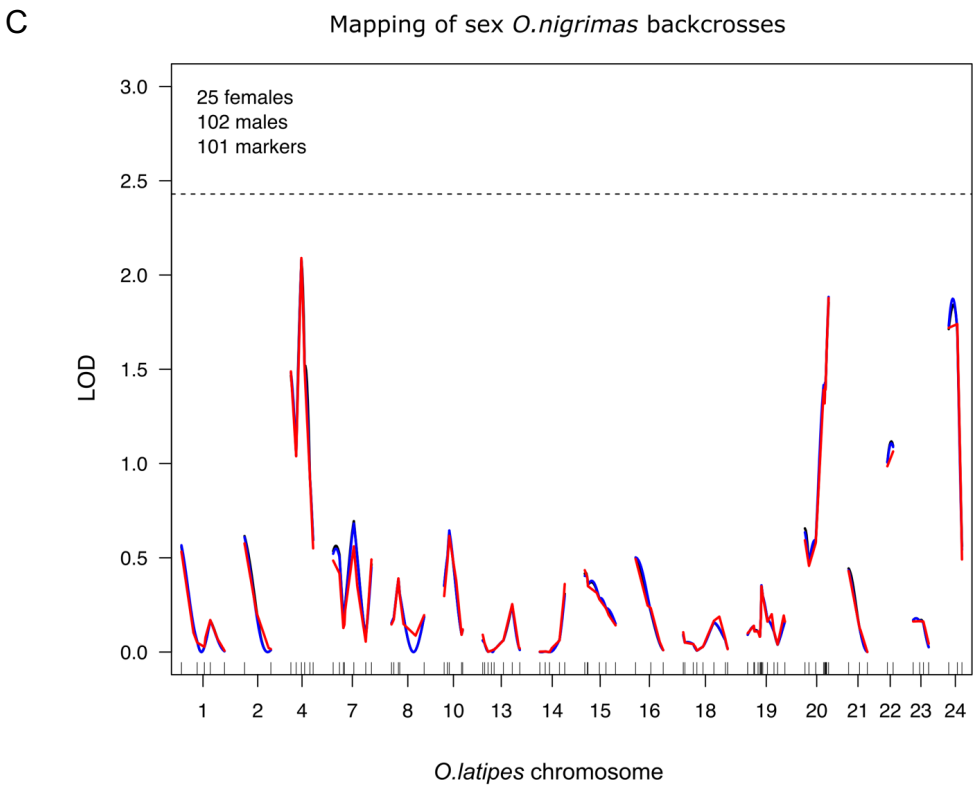


Fig. 1: Three LOD scores plots for loci associated with sex. In the intercrosses (A), a locus on chromosome 24 reached significance level, but also a locus on chromosome 21 has a high LOD score. In the *O. eversi* backcrosses (B), an extremely high LOD score for a locus on chromosome 24 was found and in *O. nigrimas* backcrosses (C), high peaks were located on chromosome 4, 20 and 24.

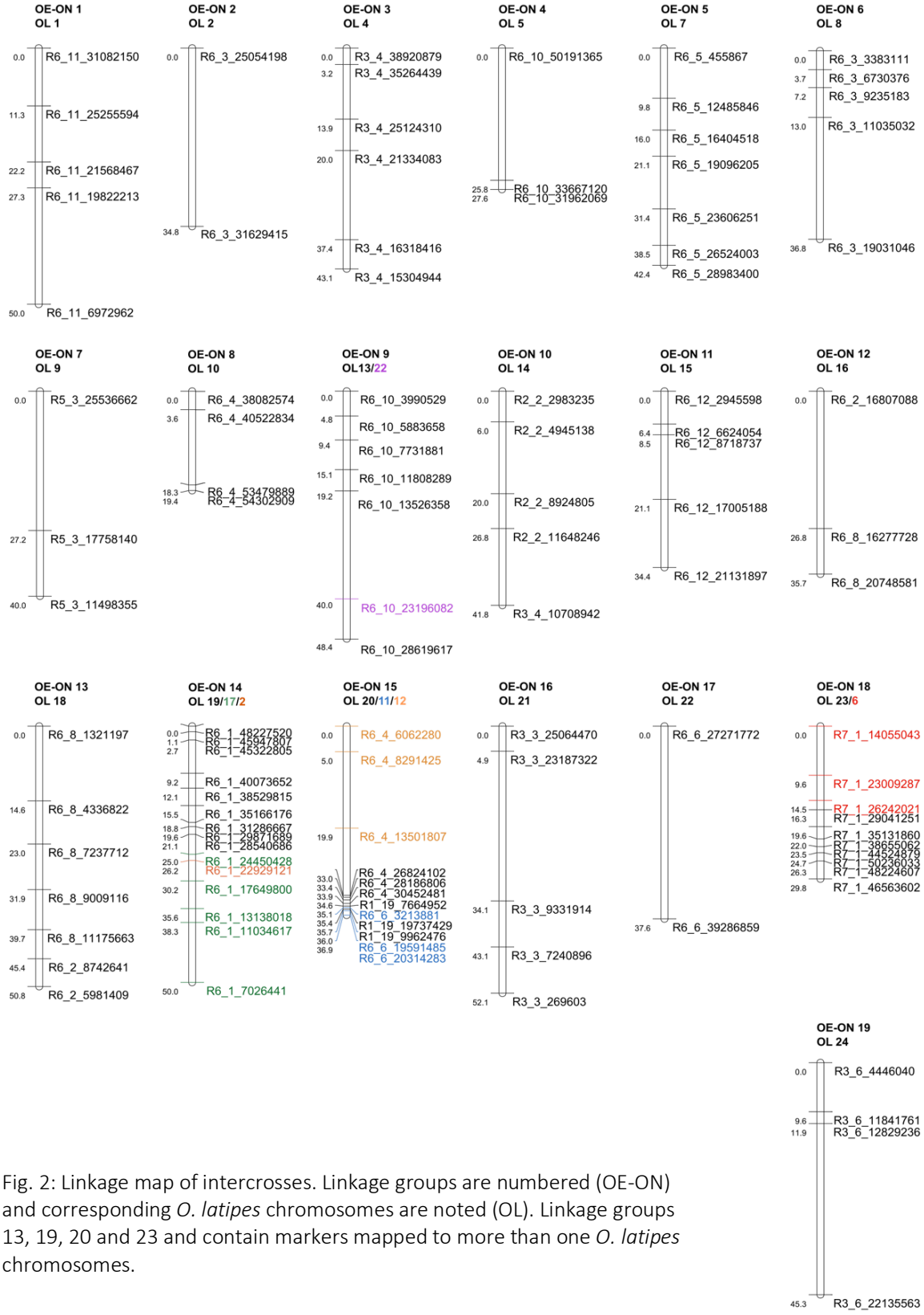
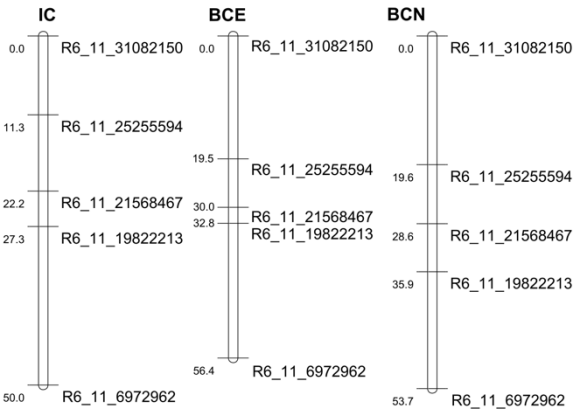
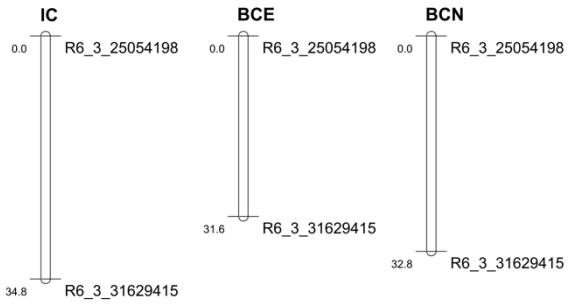


Fig. 2: Linkage map of intercrosses. Linkage groups are numbered (OE-ON) and corresponding *O. latipes* chromosomes are noted (OL). Linkage groups 13, 19, 20 and 23 and contain markers mapped to more than one *O. latipes* chromosomes.

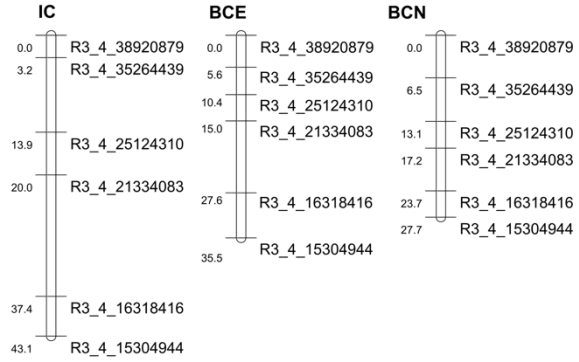
OE-ON 1
OL 1



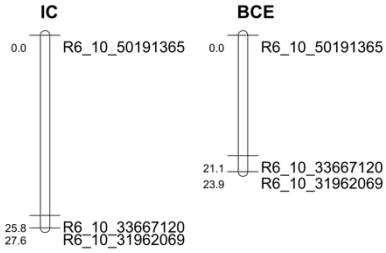
OE-ON 2
OL 2



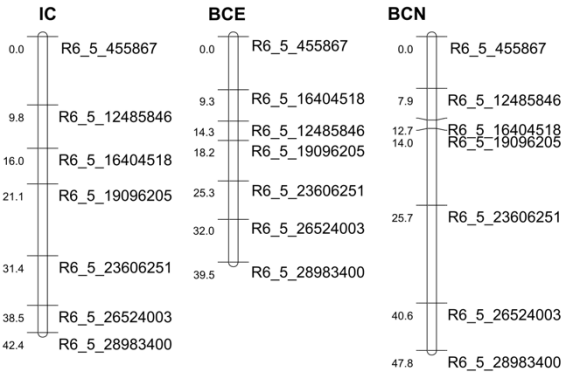
OE-ON 3
OL 4



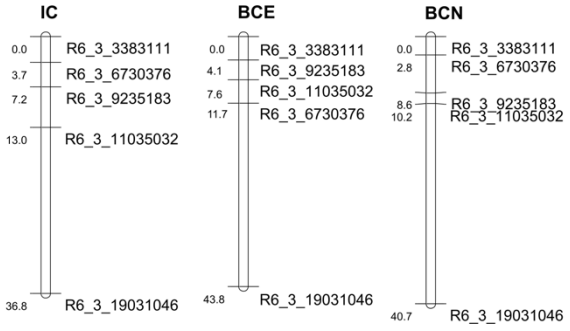
OE-ON 4
OL 5



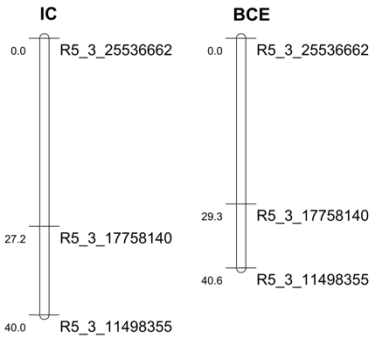
OE-ON 5
OL7



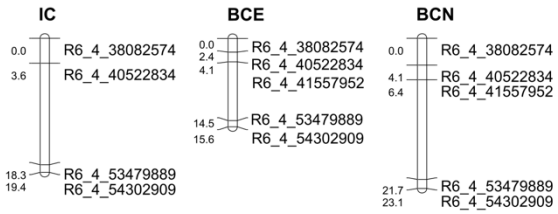
OE-ON 6
OL 8



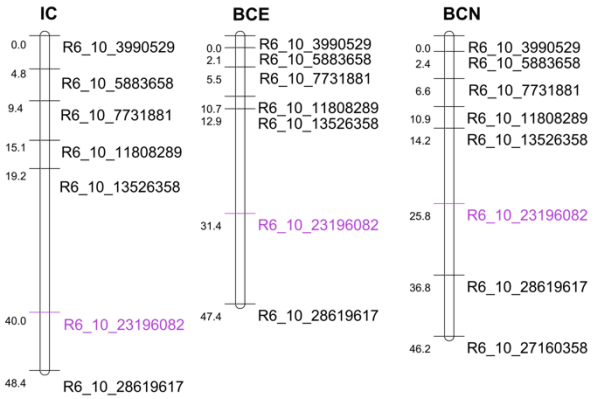
OE-ON 7
OL 9



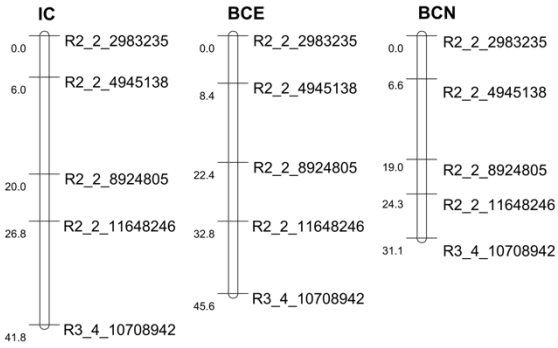
OE-ON 8
OL 10



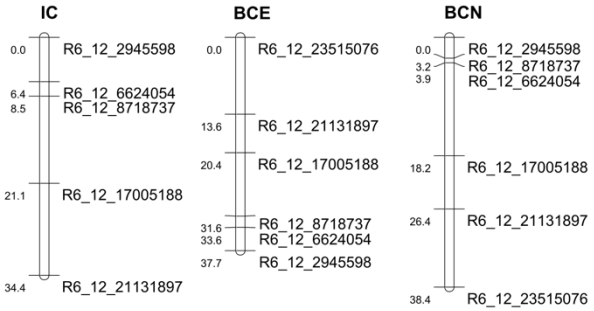
OE-ON 9
OL 13/22



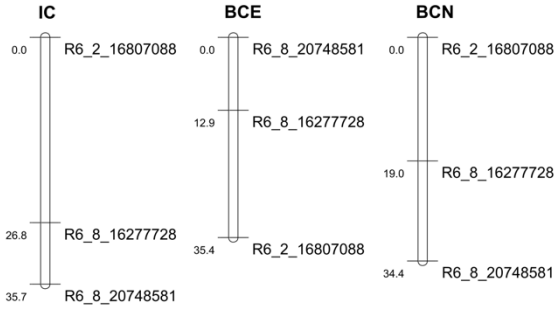
OE-ON 10
OL 14



OE-ON 11
OL 15



OE-ON 12
OL 16




OE-ON 13
OL 18
IC
0.0 R6_8_1321197
14.6 R6_8_4336822
23.0 R6_8_7237712
31.9 R6_8_9009116
39.7 R6_8_11175663
45.4 R6_2_8742641
50.8 R6_2_5981409
BCE
0.0 R6_8_1321197
18.7 R6_8_7237712
29.7 R6_8_9009116
34.7 R6_8_11175663
38.4 R6_2_8742641
BCN
0.0 R6_8_1321197
2.8 R6_8_1673931
17.1 R6_8_4336822
30.8 R6_8_7237712
38.4 R6_8_9009116
42.9 R6_8_11175663
53.4 R6_2_8742641
55.4 R6_2_5981409
OE-ON 14
OL 19/17/2
IC
0.0 R6_1_48227520
1.1 R6_1_45947807
2.7 R6_1_45322805
9.2 R6_1_40073652
12.1 R6_1_38529815
15.5 R6_1_35166176
18.8 R6_1_31286667
19.6 R6_1_29871689
21.1 R6_1_28540686
25.0 R6_1_24450428
26.2 R6_1_22929121
30.2 R6_1_17649800
35.6 R6_1_13138018
38.3 R6_1_11034617
50.0 R6_1_7026441
BCE
0.0 R6_1_45947807
1.2 R6_1_45322805
5.9 R6_1_38529815
6.9 R6_1_31286667
7.1 R6_1_35166176
7.9 R6_1_29871689
9.5 R6_1_28540686
12.1 R6_1_40073652
15.1 R6_1_24450428
16.3 R6_1_22929121
23.1 R6_1_17649800
27.4 R6_1_13138018
31.4 R6_1_11034617
42.2 R6_1_7026441
BCN
0.0 R6_1_45322805
7.7 R6_1_40073652
8.5 R6_1_38529815
13.0 R6_1_35166176
15.4 R6_1_31286667
16.4 R6_1_29871689
17.2 R6_1_24450428
17.4 R6_1_28540686
18.9 R6_1_22929121
24.6 R6_1_17649800
32.8 R6_1_13138018
37.3 R6_1_11034617
46.4 R6_1_7026441
OE-ON 15
OL 20/11/12
IC
0.0 R6_4_6062280
5.0 R6_4_8291425
19.9 R6_4_13501807
33.0 R6_4_26824102
33.4 R6_4_28186806
33.9 R6_4_30452481
34.6 R1_19_7664952
35.1 R6_6_3213881
35.4 R1_19_19737429
35.7 R1_19_9962476
36.0 R6_6_19591485
36.9 R6_6_20314283
BCE
0.0 R6_4_6062280
8.4 R6_4_8291425
24.2 R6_4_13501807
28.9 R6_4_28186806
29.6 R6_4_26824102
30.3 R6_4_30452481
31.4 R1_19_9962476
31.6 R6_6_20314283
31.9 R6_6_19591485
32.2 R6_6_3213881
32.9 R6_6_23755850
33.7 R1_19_7664952
BCN
0.0 R6_4_6062280
5.0 R6_4_8291425
13.5 R6_4_13501807
23.3 R6_4_28186806
24.6 R6_4_30452481
25.0 R6_4_26824102
25.6 R6_6_3213881
26.4 R1_19_9962476
27.1 R1_19_7664952
29.5 R6_6_19591485
OE-ON 16
OL 21
IC
0.0 R3_3_269603
9.0 R3_3_7240896
18.0 R3_3_9331914
47.2 R3_3_23187322
52.1 R3_3_25064470
BCE
0.0 R3_3_269603
13.0 R3_3_7240896
17.4 R3_3_9331914
55.2 R3_3_23187322
BCN
0.0 R3_3_269603
13.7 R3_3_7240896
23.3 R3_3_9331914
7
0.0 R3_3_23187322
6.3 R3_3_25064470
OE-ON 17
OL 22
IC
0.0 R6_6_39286859
37.6 R6_6_27271772
BCE
0.0 R6_6_39286859
8.2 R6_6_45574819
BCN
0.0 R6_6_39286859
7.2 R6_6_45574819
OE-ON 18
OL 23/6
IC
0.0 R7_1_14055043
9.6 R7_1_23009287
14.5 R7_1_26242021
16.3 R7_1_29041251
19.6 R7_1_35131860
22.0 R7_1_38655062
23.5 R7_1_44524879
24.7 R7_1_50236033
26.3 R7_1_48224607
29.8 R7_1_46563602
BCE
0.0 R7_1_3117179
22.5 R7_1_8059825
33.9 R7_1_14055043
39.5 R7_1_23009287
46.0 R7_1_26242021
48.0 R7_1_29041251
51.8 R7_1_35131860
54.1 R7_1_38655062
55.2 R7_1_44524879
55.8 R7_1_50236033
57.7 R7_1_46563602
63.7 R7_1_48224607
BCN
0.0 R7_1_14055043
8.1 R7_1_26242021
12.8 R7_1_38655062
19.3 R7_1_48224607

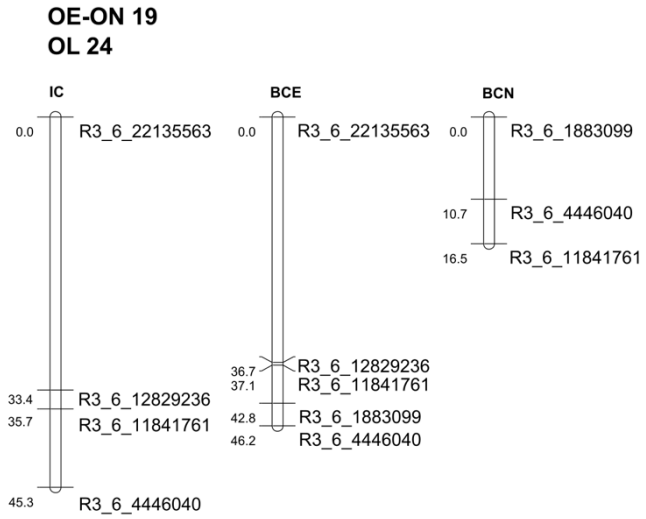


Fig. 3: Linkage map of all three crosses. Linkage groups are numbered (OE-ON) and corresponding *O. latipes* chromosomes are noted (OL).

Fig. 4: Synteny of intercross linkage map (in pink) with *O. latipes* chromosomes (in blue). Marker names derived from our *O. eversi* draft genome (unpublished).

Table 1: Information about linkage map of intercrosses, containing group ID for mapping plots, number of markers per linkage group, length of linkage group and number of corresponding *O. latipes* chromosome.

Linkage group	number of markers	length (in cM)	Chromosome O. latipes
OE-ON1	5	49.989	1
OE-ON2	2	34.791	2
OE-ON3	6	43.08	4
OE-ON4	3	27.551	5
OE-ON5	7	42.397	7
OE-ON6	5	36.753	8
OE-ON7	9	40.034	9
OE-ON8	4	19.43	10
OE-ON9	7	48.448	13, 22
OE-ON10	5	41.839	14
OE-ON11	5	34.383	15
OE-ON12	3	35.71	16
OE-ON13	7	50.833	18
OE-ON14	15	49.999	2, 17, 19
OE-ON15	12	36.869	11, 12, 20
OE-ON16	5	52.061	21
OE-ON17	22	37.588	22
OE-ON18	10	29.799	6, 23
OE-ON19	4	45.301	24