Relative roles of genetic and epigenetic variation on the ecology and evolution of mangrove killifishes (*Kryptolebias* spp.)

Waldir M. Berbel-Filho, MSc. BSc.

Submitted to Swansea University in fulfilment of the requirements for the Degree of Doctor of Philosophy (October, 2019)

SUMMARY

The field of ecological epigenetics aims to understand the implications of epigenetic modifications in adaptation, inheritance and ultimately, evolution. Many questions remain open within ecological epigenetics, in particular, how epigenetic variation is influenced by genetic background, the extent of environmentally-induced epigenetic variants, as well as its degree of heritability. This thesis used the unique diversity of mating systems present in the killifish genus *Kryptolebias* to investigate how genetic and environmental variation shape epigenetic variation in animals.

Genetic and epigenetic structure was investigated in natural populations of *K. hermaphroditus* in northeast Brazil, with the species being confirmed as the second example of mixed-mating system in vertebrates. Cytosine methylation was largely influenced by genetic background. However, within-populations, when individuals were more genetically similar, DNA methylation was mostly affected by parasites.

Kryptolebias ocellatus, here confirmed as an outcrossing-only androdioecious species, showed deep genetic structure in southeast Brazil. Hybridisation between *K. ocellatus* and the predominantly selfing *K. hermaphroditus* was uncovered, representing the first example of hybridisation between species with different mating systems in vertebrates. Hybrids had intermediate patterns of cytosine methylation relatively to the parental species, with important biological processes being potentially misregulated.

Environmental enrichment was shown to affect brain cytosine methylation patterns in two inbred strains of *K. marmoratus*, however genetic background had a stronger effect than environmental variation. Commonly-affected epialleles between genotypes predominantly showed a genotype-by-environment reaction norm, suggesting that exclusively environmentally-induced epialleles may be rare. Intergenerationally, parental activity affected offspring activity, and a limited number of putative intergenerational epialleles were identified. This is the first example of behavioural parental effects induced by environmental enrichment in fish.

These findings show that genetic background has a prominent effect and must be take into account when evaluating the evolutionary potential of cytosine methylation variation. In addition, inheritance of environmentally-induced cytosine methylation epialleles may be limited, with other epigenetic mechanisms, such as microRNAs, being more likely to escape epigenetic reprogramming and transmit epigenetically-induced parental effects.

DECLARATION AND STATEMENTS

I, Waldir M. Berbel-Filho, certify that this work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed:

Date:

STATEMENT 1

This thesis is the result of my own investigations, except where otherwise stated. Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed:

Date:

STATEMENT 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, available online in the University's Open Access Repository and for the title and summary to be made available to outside organisations.

Signed:

Date:

AUTHORSHIP DECLARATION

The following authors and their respective institutions contributed to the publication of the work undertaken as part of this thesis:

	Name	Institution	Chapter (s)
Candidate	Waldir M. Berbel-Filho (WMB-F)	Swansea University	1, 4
Author 1	Sofia Consuegra (SC)	Swansea University	1, 4
Author 2	Carlos Garcia de Leaniz (CGL)	Swansea University	1, 4
Author 3	SergioMaia Queiroz Lima (SMQL)	Universidade Federaldo Rio Grande do Norte (Brazil)	1
Author 4	Palóma Moran (PM)	Universidad de Vigo (Spain)	1
Author 5	Joanne Cable (JC)	Cardiff University	1
Author 6	Deiene Rodriguéz-Barreto (DRB)	Swansea University	4
Author 7	Nikita Berry (NB)	Swansea University	4

Publication 1 (Chapter 1):

Berbel-Filho, W. M., de Leaniz, C. G., Móran, P., Cable, J., Lima, S. M. Q. and S. Consuegra. 2019.Local parasite pressures and host genotype modulate epigenetic diversity in a mixed-mating fish. Ecology and Evolution 9:8736-8748.

WMB-F, SC, and CGL conceived the work; SMQL planned the field work and conducted the sampling together with WMB-F, CGL, and SC; WMB-F did the microsatellite and parasite screening, with contributions from JC; WMB-F and PM performed the MS-AFLP analyses. WMB-F analysed the data with the contribution of SC, CGL, and PM. WMB-F and SC wrote the paper with contributions from all authors.

Publication 2 (Chapter 4):

Berbel-Filho, W. M., Rodriguéz-Barreto, D., Berry, N, de Leaniz, C. G. and S. Consuegra. 2019. Contrasting DNA methylation responses of inbred fish lines to different rearing environments. Epigenetics 14:939-948.

WMB-F and SC designed the experiment. WMB-F, NB, DRB performed the experiment. WMB-F and DRB analysed the epigenetic data. CGL helped analysing the data. WMB-F and SC wrote the manuscript with participation of all authors.

We, the undersigned, agree with the Authorship contribution stated above and agree with the inclusion of this work as part of the current thesis.

Candidate:

Author 1 (SC):

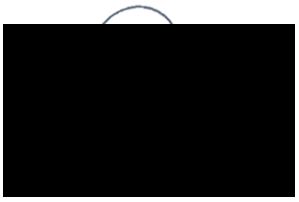
Author 2 (CGL):

Author 3 (SMQL):

Author 4 (PM):

Author 5 (JC):

Author 6 (DRB):



Author 7 (NB):

DISCLAIMER

Funding: This thesis was funded by the Brazilian government throughout the Science Without Borders PhD scholarship program (CNPq grant number 233161/2014-7) and by National Geographic/WAITT foundation grant scheme (grant number W461-16).

Ethics: Information and protocols for animal research were approved by Swansea University College of Science Ethics Review Committee. Project reference numbers are provided in appropriate chapters.

Chapter 1: Waldir M. Berbel-Filho (WMB-F), Sofia Consuegra (SC), and Carlos Garcia de Leaniz (CGL) conceived the work; Sergio M. Q. Lima (SMQL) planned the field work and conducted the sampling together with WMB-F, CGL, and SC; WMB-F did the microsatellite and parasite screening, with contributions from Joanne Cable; WMB-F and Paloma Móran performed the MS-AFLP analyses. WMB-F analysed the data with the contribution of SC, CGL, and PM. WMB-F and SC wrote the chapter with contributions from all authors.

Chapter 2: WMB-F, SC, and, SMQL conceived the idea. SC, CGL and John Avise secured funding. WMB-F, SMQL, Hélder Espírito-Santo and Mateus Lira carried out the fieldwork. WMB-F and Andrey Tatarenkov performed the laboratory and genetic analysis using mtDNA and microsatellites data. WMB-F and Carlos Rodríguez-López (CRL) performed the genotype-by-sequencing library. George Pacheco helped analysing the genetic data. WMB-F wrote the chapter with contributions from all authors.

Chapter 3: WMB-F, SC and CRL conceived the idea. WMB-F analysed the data with contribution of CRL. WMB-F wrote the chapter with contributions from CGL and SC.

Chapter 4: WMB-F and SC designed the experiment. WMB-F, Nikita Berry (NB), and Deiene Rodriguéz-Barreto (DRB) performed the experiment. WMB-F and DRB analysed the epigenetic data. CGL helped analysing the data. WMB-F wrote the chapter with participation of all authors.

Chapter 5: WMB-F and SC designed the experiment. WMB-F, Nikita Berry (NB), and Deiene Rodriguéz-Barreto (DRB) performed the experiment. Sofia Teixeira helped with the cortisol analysis. SC and CGL helped analysing the data. WMB-F wrote the chapter with participation of all authors.

ACKNOWLEDGEMENTS

I am very grateful to everyone who has helped directly or indirectly to the development of this thesis. Firstly, I would like to express my gratitude towards my supervisor Dr. Sonia Consuegra for all the help as a mentor and friend throughout my PhD years. Your support was truly essential to make this happen. In addition, I would like to thank to my second supervisor, Dr. Carlos Garcia de Leaniz to always provide his support, expertise and advice when needed.

This thesis would not have been possible without support and help of numerous people. I have a to express my immense gratitude to my third supervisor and friend Sergio Lima, who has contributed significantly to my formation as a person and scientist. A big thanks to my 'Kryptotrip' friends, Helder Espírito-Santo and Mateus Lira, who embarked into a big adventure into a tinny car looking for *Kryptolebias* across the 4.000km across the Brazilian coast. Many thanks to all students (Nikita, Mel, Kirei, Charlotte, Ed, Jasmine, Dylan) that volunteered their time and care to the Killifish lab. A special thanks to Dr. Priscila Fernandes for keeping the lab in order while I was stressed in the last months of my PhD. Many thanks to the very supportive group of fellow PhD and graduate students who shared their frustrations, happiness and leisure times together. We build strong connections during this time which made the PhD times more enjoyable and funnier than I thought.

Thanks to my Mum, Dad and family who always supported me on my scientific pursues even without understanding exactly what I was getting into. My dearest Mum and Dad, I know it has not been easy to have your son that far away, but I truly appreciate all the support you gave me, even the tiniest messages will be always remembered. I love you both.

Lastly, my immense gratitude to my project manager, bioinformatics consultant, psychologist, wife, and love, Larissa. Life is much funnier and enjoyable with you. Thanks for being the partner and friend you have always been. I truly believe this could not have be done without you and your wise advises to keep my anxious head in peace. Thanks for every help, every advice, and to embark into the crazy ideas and adventures I propose. Thanks for being you. I don't have words to express how important you have been to me. I feel safe to take any random path in life, as long as you are with me.

Thanks everyone (even the ones I forgot to mention or do not know they were important somehow) who has made this come true!

TABLE OF CONTENTS

I. GENERAL INTRODUCTION

I.I. Reviewing gene-based approach

The astonishing diversity of phenotypes observed in nature - from the marvellous colour patterns on butterfly wings to the complex mating behaviours of birds of paradise - has been long thought to be exclusively generated and inherited by what is written in the "code of life", namely the DNA sequence. The traditional view of DNA sequence as the sole unit of inheritance, also called "gene-centred" or "gene-based approach" (Bonduriansky and Day 2018; Sultan 2015), has deep and historical implications on the ways that the field of evolutionary biology has emerged and evolved, and also has other more complex cultural, historical and social legacies (Bonduriansky and Day 2018; Garver and Garver 1991; Kevles and Hood 1992). The following statements, for example, were made just after the announcement that the human genome had been sequenced in June 2000:

"It represents an immense step forward for humanity in deciphering the make-up of life itself." (Former Japanese Prime Minister Yoshiro Mori)

"Today we are learning the language in which God created life." (Former US President Bill Clinton)

"We've now got to the point in human history where for the first time we are going to hold in our hands the set of instructions to make a human being." (Dr John Sulston, UK Sanger Centre)

DNA sequence is undoubtably an important component of cellular machinery and key source of heritable information. However, mounting empirical evidence show genetically-identical individuals vary in phenotype across all levels of biological organization (e. g. gene expression, cellular differentiation, physiology, behaviour), both within (phenotypic plasticity) and across generations (known as transgenerational

phenotypic plasticity or parental effects) (Callahan et al. 1997; Des Marais et al. 2013; Fusco and Minelli 2010; Jablonka 2006; Scheiner et al. 2012; Sentis et al. 2018; Uller 2008; Verhoeven and Preite 2014; West-Eberhard 2003), emphasising the role of developmental and environmental variation on shaping phenotypes. Therefore, information about DNA sequence itself may not be enough to decipher "the make-up of life", or perhaps DNA alone does not possess all of the self-contained instructions on how to "make a human being" or any other organism.

Evolutionary biologists are now been increasingly stimulated by the enormous body of evidence challenging the deeply engrained gene-based ideas (Badyaev 2005; Dickins and Rahman 2012; Pigliucci 2007). In face of these new challenges, some evolutionary biologists are calling for an review/expansion of the concepts and mechanisms included in the Modern Synthesis (MS) (Huxley 1942; Pigliucci and Muller 2010), a review which has been named as the "Extended Evolutionary Synthesis" (EES) (Bonduriansky and Day 2018; Dickins and Rahman 2012; Laland et al. 2015). The EES proposes to expand the scope of the MS in a more integrative view of phenotypic variation, heredity and evolution, with genomes being one component an of a complex regulatory system in which organisms affect and are affected by environmental variation, and environmental signals interacting with genomes through non-genetic mechanisms (Sultan 2015). Much discussion is currently undergoing about whether the mechanisms being proposed by EES, such as phenotypic plasticity, niche construction and inclusive inheritance, were already included in scope of the MS (Laland et al. 2014; Svensson 2018) or even if those are relevant to the basic long-term process driving evolutionary change (e.g. natural selection, genetic drift, recombination) or simply represent "add-ons" mechanisms for generating variability (Charlesworth et al. 2017). This debate, together with the technological advances in

acquiring genomic data, have helped to redraw attention to molecular mechanisms underlying phenotypic plasticity, both with and among generations (Danchin et al. 2019; Laland et al. 2015; Sultan 2015).

Phenotypic plasticity has been shown to facilitate colonisation of new habitats (Yeh and Price 2004), allow rapid shifts between adaptive peaks (Agrawal et al. 1999), promote morphological diversity from pre-existing phenotypes (Levis et al. 2018) and influence population gene flow and ultimately speciation (Price et al. 2003). Parental effects, when parental phenotypes affect offspring phenotypes, are a transgenerational form of phenotypic plasticity also known to affect organism's evolutionary trajectories (Burgess and Marshall 2014). Although within-generation phenotypic plasticity seems to be mostly mediated by chemical modifications in the molecular machinery regulating gene expression, parental effects can also be transmitted by other non-genetic mechanisms. These include parental care, cultural transmission, intrauterine environments, which affect offspring fitness without being directly transmitted by the germline cells (Bonduriansky and Day 2009). Although phenotypically-plastic traits have been extensively reported across a wide range of taxa (West-Eberhard 2003), and the evidence for parental effects transmitted via germline cells have been increasingly reported (Chen et al. 2016; Illum et al. 2018; Posner et al. 2019; Rodgers et al. 2015; Rodriguez-Barreto et al. 2019), the knowledge about the molecular mechanisms in charge of generating and transmitting plastic phenotypes throughout generations is still incipient, as well as their stability and evolutionary role (Dickins and Rahman 2012; Duncan et al. 2014; Richards 2006).

I.II. Ecological epigenetics

The term 'epigenetics' (Greek prefix 'epi' meaning on top of/over) was first coined by Waddington (1942) referring to the interactions between genes and environments.

More recently however, the term has been narrowed to the individual level, referring to the chemical modifications of chromatin and/or transcribed DNA which affect gene activity and expression without modifying the underlying DNA sequence (Duncan et al. 2014; Jablonka and Raz 2009; Richards 2006).

Epigenetic mechanisms are essential for vital biological processes, such as cell differentiation, regulation of cell-cycle, genomic imprinting and dosage-compensation (Reik 2007). Several epigenetic modifications, such as histone modifications, DNA acetylation and microRNAs have been described, however DNA methylation is by far the best characterised and mostly extensively studied (Lea et al. 2017). At the molecular scale, DNA methylation consists of the modification of a DNA base with the addition of a methyl group (CH_3). In animals, it mostly occurs in CpG (i.e. cytosine followed by guanine) residues with the inclusion methyl group by methyltransferase enzymes to the fifth carbon of a cytosine (Jones 2012). This epigenetic modification affects the accessibility of transcription factors to DNA strand by the recruitment of methyl CpG binding proteins (Moore et al. 2013). The effect of DNA methylation on gene activity and expression levels varies across taxa and genomic contexts (Banta and Richards 2018; Best et al. 2018; Jones 2012; Moore et al. 2013). In mammals, DNA methylation on gene promoter is generally associated with silencing of gene expression (Jones 2012). Gene body methylation, however, has been shown to have multiple functions, such as repressing transcriptional noise (Horvath et al. 2019; Huh et al. 2013), and ensuring the inclusion of the first and last exons in a transcript (Sati et al. 2012). Recently, a conserved inverse correlation between first intron DNA methylation and gene expression was found across tissues and phylogenetically distant vertebrates (Anastasiadi et al. 2018). DNA methylation also interacts with other

epigenetic mechanisms, such as histone modifications and microRNAs, in a complex and integrative mode to regulate transcription levels (Moore et al. 2013).

Epigenetic modifications can have profound effects on the ways we currently understand ecology and evolution (Bossdorf et al. 2008). Within an ecological perspective, given its more dynamic and reversible nature compared to genetic mutations (Bonduriansky and Day 2018; Schmitz et al. 2011), epigenetic modifications (either environmentally or stochastically-induced) will result in changes on gene expression, promoting phenotypic plasticity in an heterogeneous environments (Angers et al. 2010; Bossdorf et al. 2010; Leung et al. 2016; Massicotte et al. 2011). Epigenetic modifications can also represent an additional source of phenotypic variation that can be particularly important for organisms with reduced capacity of genetically-based adaptation, such as asexual organisms or highly inbred taxa (Castonguay and Angers 2012; Massicotte and Angers 2012; Schrey et al. 2012; Verhoeven and Preite 2014). From an evolutionary perspective, given the potential of epigenetic modifications to be transgenerationally-transmitted (Chen et al. 2016; Posner et al. 2019; Richards 2006), epigenetic inheritance imposes a challenge to the notion that molecular inheritance transmitted via germline cells resides solely on DNA sequences (Danchin et al. 2011; Dickins and Rahman 2012).

The inherent implications of epigenetic modifications for ecology and evolution (Bossdorf et al. 2008), the recent advances in chromatin research (Allis and Jenuwein 2016) and the increased availability of 'omics' technologies has led to a substantial development of epigenetics research over the last 20 years (Lea et al. 2017). More specifically, the field of 'Ecological Epigenetics' (EE) is trying to integrate epigenetic information into ecological and evolutionary framework, aiming to understand what factors drive the natural patterns of epigenetic variation and their eco-evolutionary

consequences (Richards et al. 2017; Schrey et al. 2013). Despite the recent advances, a substantial number of conceptual publications listed key questions that remain open in the field of EE, highlighting their theoretical implications and calling for urgent experimental research addressing those questions, especially beyond the scope of model organisms (Bossdorf et al. 2008; Duncan et al. 2014; Jablonka and Raz 2009; Jeremias et al. 2018; Richards et al. 2017; Richards et al. 2010; Verhoeven et al. 2016). On the following subsections, I will highlight the theorical premises of some of the questions analysed by this thesis, and how addressing those contributed to a better understanding of the ecological and evolutionary potential of epigenetic modifications.

I.II.I. What is the interplay between genetic and epigenetic variation?

Epigenetic variants which are strictly under genetic control (e. g. allele-specific methylation) simply represent downstream effects of genetic variants, not adding any additional evolutionary value to epigenetics which cannot be already explained by genetics (Verhoeven et al. 2016). Despite the recent advances with acquiring genome-wide epigenetic data (Lea et al. 2017), the interplay between the genomes, epigenome and environmental changes remains poorly understood (Verhoeven et al. 2016). Richards (2006) proposed a classification of epialleles according to their degree of autonomy from the underlying genetic background, with most of epialleles being obligatory (fully dependent of genetic variation), followed by facilitated (partially dependent) and pure epialleles (autonomous) (Fig. I.1). While obligatory epialleles are downstream consequences a specific genetic allele, facilitated epialleles can act as potential intermediators between environmental conditions and genome responses, while the pure epialleles represent environmentally-driven responses (Schmitz et al. 2013). The degree of autonomy of epialleles also brings implications for the study of

phenotypic plasticity, as epigenetically-regulated plastic phenotypes would show different reaction norms according to their degree of autonomy of its epialleles (Sultan 2015). Obligatory epialleles should not show any response to environmental variation, while facilitated and pure epialleles would show genotype-by-environment (G x E) (direction of variation across environments changing in a genotypic-specific manners) and canalised plasticity (same direction of variation across environments regardless of the genotypes), respectively (Fig I.1).

Associations between of genetic variants and DNA methylation variation (Bell et al. 2011; Dubin et al. 2015; Gertz et al. 2011; Gutierrez-Arcelus et al. 2013) suggest that autonomous (pure) epialleles may be limited (Dubin et al. 2015). However, there is currently very little data on the genetic-epigenetics relationship beyond model organisms (Verhoeven et al. 2016), especially due to the lack of whole-genome and methylome data on non-model organisms (Richards et al. 2017).

Genetic background is an important factor affecting not only specific epialleles but also at genome-wide DNA methylation levels, with associations between genotypes and epigenotypes being commonly found (Alonso et al. 2016; Herrera et al. 2016; Leung et al. 2016; Schrey et al. 2013). In addition, individual genetic heterozygosity levels also seem to influence global DNA methylation levels, with higher proportion of DNA methylation often reported for clonal and/or inbred individuals compared to their outcrossed counterparts (Richards et al. 2012; Schrey et al. 2012; Vergeer et al. 2012). Given the general reduced potential for genetically-based adaptation of inbred organisms, these findings made some authors suggest that epigenetic variation may be particularly important for clonal or inbred organisms, potentially working as an non-genetic mechanism for rapid phenotyic adjustment and adaptation (Castonguay and Angers 2012; Schrey et al. 2012; Sheldon et al. 2018; Verhoeven et al. 2010;

Verhoeven and Preite 2014). However, higher global DNA methylation levels commoly found of inbred individuals may be mediating inbreeding depression (Nakamura and Hosaka 2010). For example, phenotypic effects of inbreeding depression (i. e. reduced photosyntetic effiency, reduced biomass) disappeared after demethylation in inbred *Arabidopsis thaliana* (Vergeer et al. 2012). Additionally, gene-specific effects of inbreeding in controlled crosses of Chinook salmon (*Oncorhynchus tshawytscha*), suggests that methylation changes associated with inbreeding depression are targeted to specific genes, instead of genome-wide effects, in animals (Venney et al. 2016).

Understanding the intimate relationship between genomic characteristics and epigenetic variation is a crucial step on evaluating the origin, drivers, as well as, the evolutionary potential of epigenetic variation (Richards et al. 2017; Sultan 2015). As highlighted above, genomic background and diversity seem to influence DNA methylation variation, both at the allele or at global levels. However, detangling genetic from epigenetic variation has proven especially challenging, given the high levels of genetic variability commonly found of sexually-reproducing organisms in natural populations (Verhoeven and Preite 2014). With this thesis, I aimed to address questions regarding on how DNA methylation is influenced by genetic background (Chapters 1, 3 and 4), as well as the distribution and degree of autonomy of epialleles (Chapter 4).

I.II. II. How much epigenetic variation is environmentally-induced?

DNA methylation varies widely across taxa and genomic contexts (e. g. promoters, gene bodies, transposable elements) (Duncan et al. 2014; Jones 2012), but also intraspecifically, both at global DNA methylation estimates (Alonso et al. 2016;

Schrey et al. 2013) or at the gene-specific levels (Venney et al. 2016). Intraspecific variation suggests that DNA methylation can be influenced by the environment (Richards et al. 2017), be under genetic control (Dubin et al. 2015) and/or be induced by stochastic errors in maintaining DNA methylation patterns by DNA methyltransferases, which can be enhanced in stressful environmental conditions (Leung et al. 2016; Schmitz et al. 2011; Verhoeven et al. 2010). These effects are complementary, and their consequences on gene expression may contribute on generating epigenetically-mediated phenotypic variation (Leung et al. 2016).

One way to facilitate the study of environmentally-induced epialleles is to look for associations between epigenetic variation and environmental characteristics which are independent of the individuals genetic background (Verhoeven et al. 2016). Several studies have been using generically-uniform organisms to isolate the effect of genetic variants on epigenetic variation (Verhoeven and Preite 2014). DNA methylation patterns of clonal organisms are often associated with biotic and abiotic stressors, such as predation and parasites (Asselman et al. 2015; Verhoeven and van Gurp 2012; Verhoeven et al. 2010), salinity (Raj et al. 2011), water availability (Gao et al. 2010), pH (Massicotte and Angers 2012), among others (Verhoeven et al. 2010). Although suggestive, these studies cannot disentangle whether the putative environmentally induced epialleles were either inherited, repeatedly induced by the same environmental condition or emerged via stochastic epigenetic variation, given the high rate of epimutations (Schmitz et al. 2011).

Both genetic and epigenetic variation determine phenotypic variation, which interacts with environment, creating raw material for natural selection (Fig. I.2). However, many of factors interconnected into this complex network remain understudied, especially in non-model organisms in natural conditions (Richards et al. 2017). With

this thesis, I aimed to address the role of parasite pressures and genetic background on DNA methylation variation in wild populations (Chapter 1).

I.II.III. To what extent environmentally-induced epigenetic variation is transmitted to the offspring?

While the transgenerational stability of some epigenetic markers is well established, the evidence for environmentally-induced transgenerational effects and the extent of its transmission remains incipient (Bossdorf et al. 2008; Chen et al. 2016; Richards et al. 2017). Some factors can facilitate the transmission of environmentally-induced transgenerational effects. For example, the reset of epigenetic modifications during early gametogenesis and embryogenesis, known as epigenetic reprogramming, is far from complete in some taxa, leaving a window for inheritance environmentally-induced epialleles (Daxinger and Whitelaw 2012; Verhoeven et al. 2016).

Although transgenerational effects mediated by differential epigenetic modification in the germline have been reported, recent studies indicate that germline microRNAs modifications instead of DNA methylation as primary factors driving this type of parental effects in mammals (Chen et al. 2016). The cellular DNA methylation patterns in mammals are strongly affected by epigenetic reprogramming, with the few epialleles escaping it likely to be genetically-induced and stochastically rather than environmentally-induced (Shea et al. 2015), or not last in the subsequent generations (Radford et al. 2014). In fish, however, germline epigenetic reprogramming of DNA methylation seems to be less complete. For example, sperm DNA methylation patterns are used as a template for zebrafish (*Danio rerio*) embryos methylome (Jiang et al. 2013). Additional evidence in Atlantic salmon (*Salmo salar*) has shown that few DNA

methylation epialleles induced by domestication escape epigenetic reprogramming and are maintained in the offspring (Rodriguez-Barreto et al. 2019).

Within an evolutionary context, genetically-independent epigenetic modifications are particularly relevant if inherited by the offspring. Therefore, understanding the mechanisms of epigenetics modifications and to which extent those can be transmitted transgenerationally, is a crucial step to understand any potential additive role epigenetic modifications may have on heredity and evolution. With this thesis, I aimed to address questions regarding to which extent differentially DNA methylation patterns in the parents are transmitted to the offspring, both in natural (Chapter 3) and in lab-reared conditions (Chapter 5).

Although ecological epigenetics is now an active field within evolutionary biology (Verhoeven et al. 2016), several questions about the origin, variation and evolutionary consequences of epigenetic modifications are still open. In particular, the relationships between genome, epigenome and environmental variation are better understood in plant than in animal systems (Richards et al. 2017). This relative gap in knowledge is partly caused by need of model systems on which is possible to have within-genotype homogeneity (to detangle genetic from epigenetic variation) simultaneously with genotype diversity (to test the effects of common environmental pressures on different epigenomes). Mixed-mating organisms (alternating self-fertilisation hermaphroditism with outcrossing, following the definition of mating system for hermaphrodites in Barrett (2014)) are perfectly suitable to test associations between genetic and epigenetic variation with environmental pressures, because individuals with contrasting levels of genetic diversity and different genetic backgrounds naturally coexist. However, mixed-mating systems are much more commonly found in plants than in animals (Jarne and Auld 2006). This thesis aimed to expand the knowledge

about the ecological and evolutionary implications of epigenetic modifications in animals, using the unique diverse of mating systems present in the mangrove killifish species from the genus *Kryptolebias*, as models.

I.III. The genus *Kryptolebias*

Kryptolebias is a killifish genus (family Rivulidae) composed of seven known species (Fricke 2019; Vermeulen and Hrbek 2005), although this classification is likely to change as some taxonomic debate is still undergoing (Huber 2017; Tatarenkov et al. 2017). Phylogenetic analyses have indicated the presence of two distinct monophyletic clades within *Kryptolebias*, one composed of narrowly-distributed freshwater species living in temporary streams and pools in Central and South America (*K. campelloi* (Costa 1990), *K. sepia* Vermeulen & Hrbek 2005, *K. gracilis* Costa 2007, *K. brasiliensis* (Valenciennes 1821)). The other clade is composed of three species living on mangrove forests along the tropical and subtropical western Atlantic basin, 'the mangrove killifishes clade' (*K. marmoratus* (Poey 1880), *K. hermaphroditus* Costa 2011, *K. ocellatus* (Hensel 1868)) (Costa et al. 2010; Murphy et al. 1999; Tatarenkov et al. 2009; Vermeulen and Hrbek 2005).

Most of the rivulids (among approximately 350 species (Costa 2011b)) are relatively small (e.g. max. 200 mm of standard length), living in freshwater habitats, dioicous (composed by males and females) and sexually dimorphic (Loureiro et al. 2018). However, some remarkable reproductive adaptations and life-history strategies have evolved within this family. Some species are capable of internal fertilisation (Costa et al. 2016), while several species have an annual life-cycle, with eggs evolving drought resistance and diapause (embryo development suspended within the egg) (Costa 2011b). Perhaps more remarkably, in *Kryptolebias*, the 'mangrove killifishes clade' is composed by one of the few fish clades with androdioecious (males and

hermaphrodites; Weeks 2013) species: *K. marmoratus, K. hermaphroditus* and *K. ocellatus*, with the first two representing the only known vertebrate species capable of self-fertilisation (Avise and Tatarenkov 2015; Costa et al. 2010; Tatarenkov et al. 2009).

I.III. I. The mangrove killifishes

I.III. I.I. Geographical distribution and phylogenetic patterns

Several factors, such as its broad geographical distribution (encompassing the entire tropical and subtropical eastern Atlantic, 29°N to 23°S), morphological and ecological similarities, and taxonomic misidentifications, have contributed to a complex scenario about the taxonomic status and geographical distribution of the mangrove killifish species (Avise and Tatarenkov 2015; Costa 2011a; Costa 2016; Tatarenkov et al. 2017). Briefly, *K. ocellatus* was described first by Hensel (1868) using a single specimen from Rio de Janeiro, Brazil followed by *K. marmoratus* described by Poey (1880) using specimens from the USA or Cuba. Seegers (1984) suggested that in Rio de Janeiro there were two species, the hermaphroditic *K. ocellatus*, and a yet undescribed species, which he named *K. caudomarginatus* composed by males and hermaphrodites. Later, Costa (2011) established that *K. ocellatus* originally described by Hensel actually referred to the species with males, reclassifying *K. caudomarginatus* as a junior synonym of *K. ocellatus*, while the other syntopic species only composed of selfing hermaphrodites, actually represented an undescribed species, as supported by molecular data (Murphy et al. 1999; Tatarenkov et al. 2009), which was described as *K. hermaphroditus*.

Although this morphologically-based classification have caused historical confusion about the systematics of the mangrove killifish clade, its species are readily

distinguishable using genetic data. Phylogenetic studies have supported *K. ocellatus* is very divergent genetically from the other species in the group (see Chapter 2) and is likely to be the sister-species of the of the clade containing the selfing species *K. marmoratus* and *K. hermaphroditus*) (Murphy et al. 1999; Tatarenkov et al. 2009; Vermeulen and Hrbek 2005). As all the other known *Kryptolebias* species are phylogenetic distant, dioecious and inhabit freshwater habitats, the phylogenetic mapping of reproductive conditions in *Kryptolebias* seems to suggested that synchronous hermaphroditism have emerged in the common ancestor of all mangrove killifish species (*K. ocellatus, K. hermaphroditus* and *K. ocellatus*) while the selfing capacity evolving later in the common ancestor between *K. hermaphroditus* and *K. marmoratus* (Avise and Tatarenkov 2015; Costa et al. 2010) (Fig. I.3).

Regarding its geographical distribution, *K. ocellatus* is known to be endemic of mangrove forests of southeast Brazil, from Guanabara bay in Rio de Janeiro (22°S) (where is often syntopic with *K. hermaphroditus*, see Chapter 2) to Santa Catarina state (28ºS) (Costa 2016). *Kryptolebias hermaphroditus* until recently was only known by its type-locality, in Guaratiba, Rio de Janeiro (Costa 2011a). However, recent surveys have been expanding the geographical distribution of the species, as far as northeast Brazil and close to the Amazon river mouth (Guimarães-Costa et al. 2017; Lira et al. 2015). Given the morphological similarities between *K. hermaphroditus* and *K. marmoratus*, and the major gaps in sampling, the geographical range of each species was uncertain. Tatarenkov et al. (2017) using extensive geographical sampling, found deep genetic structure at mtDNA and microsatellite data within the selfing mangrove killifishes, with three major clades: a 'Northern' clade in Florida, northern Cuba, Bahamas, Belize and Honduras corresponding to *K. marmoratus*; a 'Southern' clade composed by populations along the Brazilian coast and a 'Central clade' present in

Panama, the Dutch Caribbean, Puerto Rico, Turks and Caicos, and southern Cuba, which might (or not) represent a different species, and it is phylogenetically more closely related to *K. hermaphroditus* (1% genetic distance at *cox1* mtDNA gene), despite its geographical proximity with *K. marmoratus* populations (3% genetic distance at *cox1* mtDNA gene) (Fig. I.3). Tatarenkov et al. (2018), found evidence for an unlikely hybridisation between individuals of *K. marmoratus* and *K. hermaphroditus* (Central clade) in San Salvador, Bahamas, suggesting that divergent *Kryptolebias* lineages, with different rates of selfing/outcrossing rates can still interbreed (see Chapter 2).

I.III. I. II. Ecology and behaviour

The mangrove killifishes are the only example of exclusively mangrove-associated fish in the western Atlantic, having their entire life cycle within mangrove forests, inhabiting a wide range of microhabitats, such as interment pools, crab burrows, mangrove leaf litter and fallen logs (Berbel-Filho et al. 2016; Lira et al. 2015; Taylor 2012) (Fig. I.4). Mangrove killifishes rarely share their microhabitats with other fish, particularly given the extreme environmental conditions of their preferred microhabitats (Taylor 2012). For example, *K. marmoratus*, the most extensively studied species, is known to tolerate extreme ranges of temperature and salinity (Taylor 2012), high levels of hydrogen sulphide (Cochrane et al. 2019), ammonia (Frick and Wright 2002) and hypoxic conditions (Regan et al. 2011). *Kryptolebias* species are usually predators from aquatic and terrestrial invertebrates (Taylor 2000), and selective (based on kinship) filial cannibalism has been reported in *K. marmoratus* (Wells and Wright 2017).

Kryptolebias marmoratus (and possibly the other mangrove killifish species), is also known for its amphibious behaviour (Turko and Wright 2015). Within this species,

emersion behaviour (jumping and moving out of water) seems to be an adaptive mechanisms to cope with several biotic and abiotic environmental pressures, such as coping with adverse aquatic conditions (Cochrane et al. 2019), increasing ammonia excretion (Frick and Wright 2002), reducing intra-specific aggression (Taylor 1990), escaping stressful thermal conditions (Gibson et al. 2015), increasing embryo survival (by enhancing embryo development and reducing embryo predation by conspecific cannibalistic individuals) (Wells et al. 2015; Wells and Wright 2017)) and increasing dispersal capacities (Bressman et al. 2018).

Behaviourally, *K. marmoratus* individuals are known to be more aggressive towards non-kin and non-familiar individuals (Edenbrow and Croft 2012). Different selfing lineages were already found to occupy the same microhabitats (e. g. crab burrow) (Ellison et al. 2012). As emersion can be used as a way to escape from intraspecific aggression, together with the highly-variable nature of mangrove microhabitats suggest a complex temporal, spatial and social scenario of population density/movement within mangroves, however not much information about its extent and how individual mangrove killifishes can cope with ever-changing complexity (temporal, spatial and social) scenarios is available (Taylor 2012).

I.III.I.III. Mixed-mating system and evolutionary implications

Since the discovery of self-fertilisation capacity in hermaphrodites of *K. marmoratus* by Harrington (1961), the species has been subjected to a wealthy amount of research, becoming an increasing common organism for evolutionary, ecological, physiological, ecotoxicological and medical research (Avise and Tatarenkov 2015; Earley et al. 2012; Lee et al. 2008; Taylor 2012; Wright 2012).

Another major breakthrough on the research of mangrove killifishes came with the discovery of mixed-mating system in the previously thought exclusively selfing species. Although males of *K. marmoratus* could be readily produced in the lab by incubating self-fertilised eggs at low temperatures (Harrington 1967) and their presence was known at rare frequency in nature, their ecological relevance was unclear, given the high levels homozygosity in the Floridian populations suggested that outcrossing in *K. marmoratus* was absent (Harrington and Kallman 1968). However, males have been subsequently uncovered in other populations, some with relatively high frequencies (e.g. >20% at Twin Cays, Belize (Davis et al. 1990)). Later, genetic analyses have confirmed male-mediated outcrossing events both in wild (Mackiewicz et al. 2006b; Taylor et al. 2001) and lab conditions (Mackiewicz et al. 2006a), supporting a mixed-mating strategy in *K. marmoratus* (Mackiewicz et al. 2006c). Despite the existence of outcrossing makes *K. marmoratus* by definition, a species with mixed-mating strategy, genetic data at the population level reveal that, save few exceptions, most of the populations of *K. marmoratus* are predominantly self-fertilising (outcrossing rate ≤ 0.2 (Jarne and Auld 2006; Tatarenkov et al. 2015; Tatarenkov et al. 2012; Tatarenkov et al. 2017; Tatarenkov et al. 2009).

The discovery of a mixed-mating system has made *K. marmoratus* an important vertebrate model to understand the environmental factors and molecular mechanisms in the evolution of mixed-mating systems. Both self-fertilisation and outcrossing have inherent benefits and costs. Selfing assures reproduction, propagates locally adapts genotypes and reduces the costs of biparental reproduction (i. e. finding a suitable mate, diluting genomic content) (Lively and Morran 2014), however over long-term creates inbred organisms, lacking the genetically-base adaptability to cope coevolving antagonists (e. g. parasites), which can be provided by occasional outcrossing (Lively

and Morran 2014; Morran et al. 2011). Ellison et al. (2011) found that *K. marmoratus* offspring originating from selfing harboured higher parasite load than its outcrossed counterparts, suggesting that outcrossing is an adaptive mechanism to cope with parasite pressure. The adaptive potential outcrossing as a generator of genetic diversity in *K. marmoratus* was later reinforced by the evidence that males prefer to associate with the most genetically-dissimilar hermaphrodites (Ellison et al. 2013). As outcrossing is likely to be male-mediated (Furness et al. 2015), and higher frequency of males has been associated with higher outcrossing rates at some populations (Mackiewicz et al. 2006b; Tatarenkov et al. 2015; Turner et al. 2006). The frequency of males seems to be caused by combinations of environmental (e.g. temperature) and genetic factors, with some selfing lines constantly showing higher proportion of males compared to others, even under similar environmental conditions (Turner et al. 2006). Ellison et al. (2015) found different DNA methylation patterns and gene expression levels between males and hermaphrodites in two genes related to temperature sex-determination in *K. marmoratus*.

In mixed-mating species, the balance between populational selfing/outcrossing affect populational genetic structure differently at different scales (Avise and Tatarenkov 2015). Selfing drastically reduces effective population size (N_e), which facilitates the formation of population structure, given the stronger effects of genetic drift and/or selection when Ne is small (Charlesworth and Willis 2009). In mixed-mating systems, low levels of genetic diversity are expected within selfing lines, while high genetic differentiation is expected between selfing lines, which were potentially originated via occasional outcrossing events and/or de novo mutations and maintained in the long-term via selfing. In *K. marmoratus*, several studies have been demonstrating this scenario, with low intraline genetic variation, but extensive interline genetic

differentiation, revealing strong population structure even at small spatial scales (Ellison et al. 2012; Tatarenkov et al. 2015; Tatarenkov et al. 2012; Tatarenkov et al. 2007).

Taken together, all this evidence suggests that in *K. marmoratus*, environmental factors (through epigenetic modifications) influence populational sex ratios, which p affects selfing/outcrossing rates, which has implications on offspring fitness and broader effects on populational genetic structure. However, save few exceptions, most of *K. marmoratus* populations seemed to be predominantly selfing with highly homozygous lines (Tatarenkov et al. 2017), suggesting that other non-genetic mechanisms, such as epigenetic modification may be contributing to fine-scales adjustments on gene expression to cope with environmental variability in absence of genetic variability. Fellous et al. (2018) found that *K. marmoratus* embryos have a longer epigenetic reprogramming period than other fish and mammals, with the authors that suggesting this could be a window for epigenetically-mediated phenotypic plasticity and epigenetic inheritance. However, the evolutionary potential of epigenetic variation and its potential inheritance is yet to be tested experimentally in this system.

As highlighted above, a great deal of information about the ecology, behaviour, genetic structure and evolutionary history of *K. marmoratus* is now available. However, the other South American *Kryptolebias*, remain largely understudied (Berbel-Filho et al. 2016; Costa et al. 2010; Lira et al. 2015). Until recently *K. hermaphroditus* was only known by extremely homozygous populations in Rio de Janeiro (Tatarenkov et al. 2011; Tatarenkov et al. 2009). Histological examinations and aquarium observations confirming its hermaphroditism (Costa et al. 2010) and the low levels of genetic diversity made Costa (2011a) classify the species as only composed by self-fertilising hermaphrodites. However, recent sampling surveys have expanded its distribution to

north and northeast Brazil (Costa 2016; Guimarães-Costa et al. 2017; Lira et al. 2015; Sarmento-Soares et al. 2014) and few males have been described in some populations (Berbel-Filho et al. 2016; Costa 2016). However, the question of whether *K. hermaphroditus* is exclusively selfing or mixed-mating is still to be answered, as well as the environmental and evolutionary drivers of variation on selfing/outcrossing rates and the effects of those on its population structure (Chapter 1). *Kryptolebias ocellatus* is also poorly studied, with Costa et al. (2010) histologically describing its simultaneous hermaphroditism nature and Tatarenkov et al. (2009) based on the selfing rates of two populations suggesting the species is outcrossing-only androdioecious, not undergoing selfing. Despite other sampling reports (Costa 2016), further details about the genetic structure, the possibility of self-fertilisation in other populations, as well its interactions with *K. hermaphroditus* on its syntopic populations have never been investigated (Chapter 2).

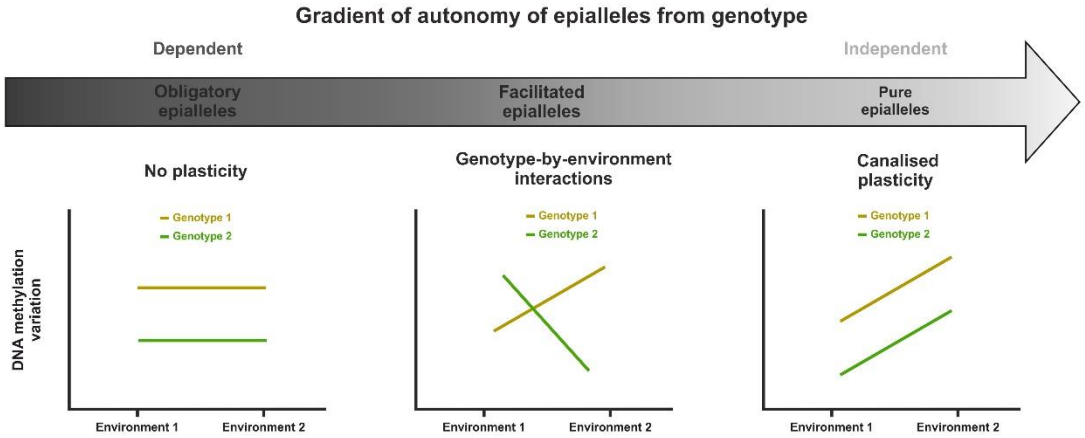


Figure I.1. Schematic representation of epiallele distribution according to its autonomy from the genetic background. Obligatory epialleles are fully dependent of its underlying genotype. Facilitated epialleles have their expression modulated by the interactions between genotypes and environments. Pure epialleles show similar direction of variation independently of its genetic background. Epiallele classification followed Richards (2006).

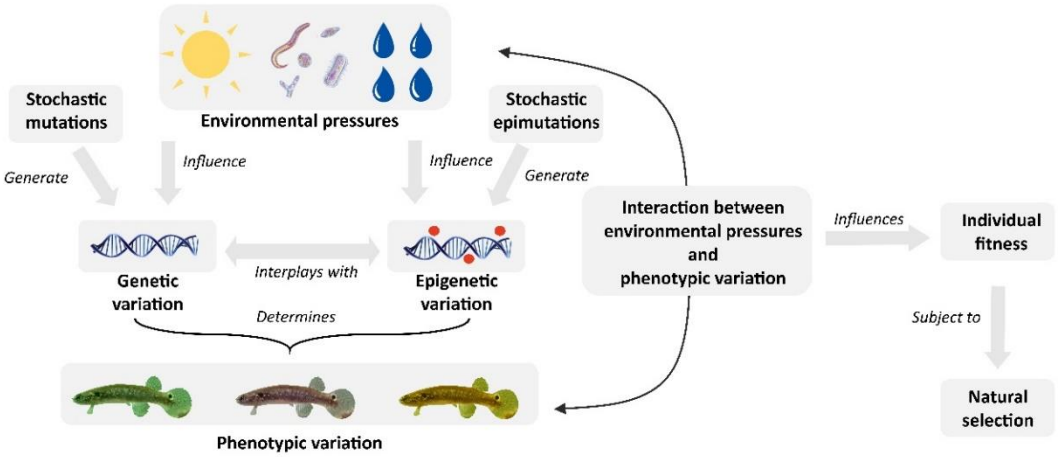


Figure I.2. Schematic representation of factors affecting genetic and epigenetic variation as well its outcomes and evolutionary forces.

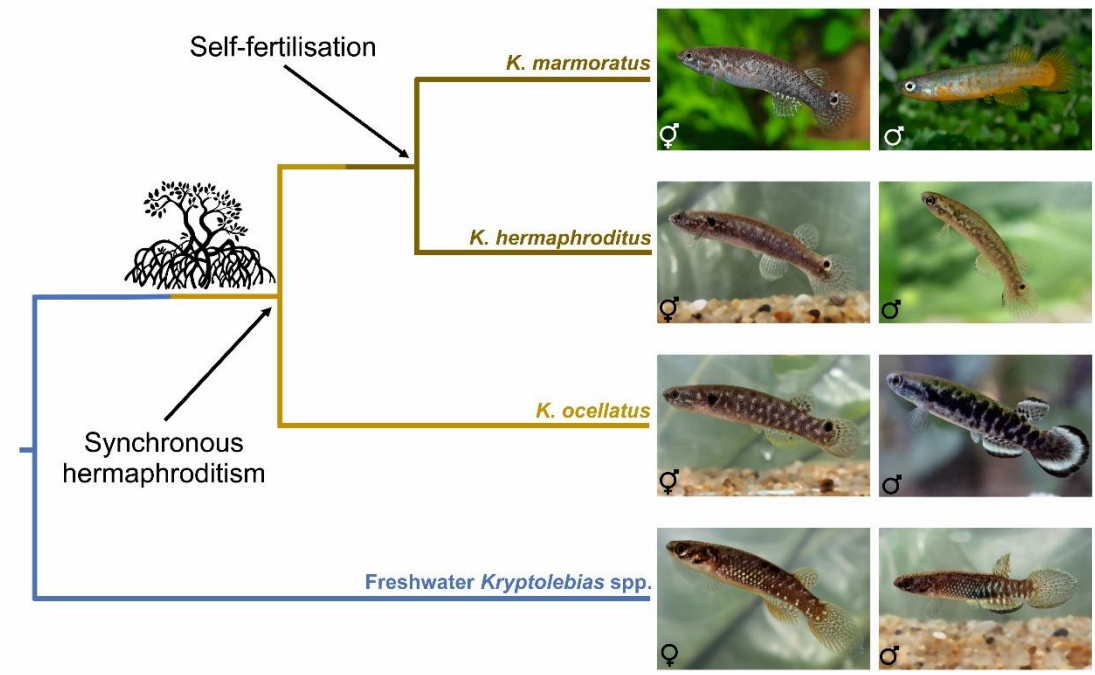


Figure I.3. Phylogenetic tree for *Kryptolebias*, with mapping of the reproductive traits and its respective species based on Costa et al. (2010). Freshwater *Kryptolebias* are represented by *K. gracilis* Costa 2007. Pictures from *K. marmoratus* provided courtesy of Frans Vermeulen.

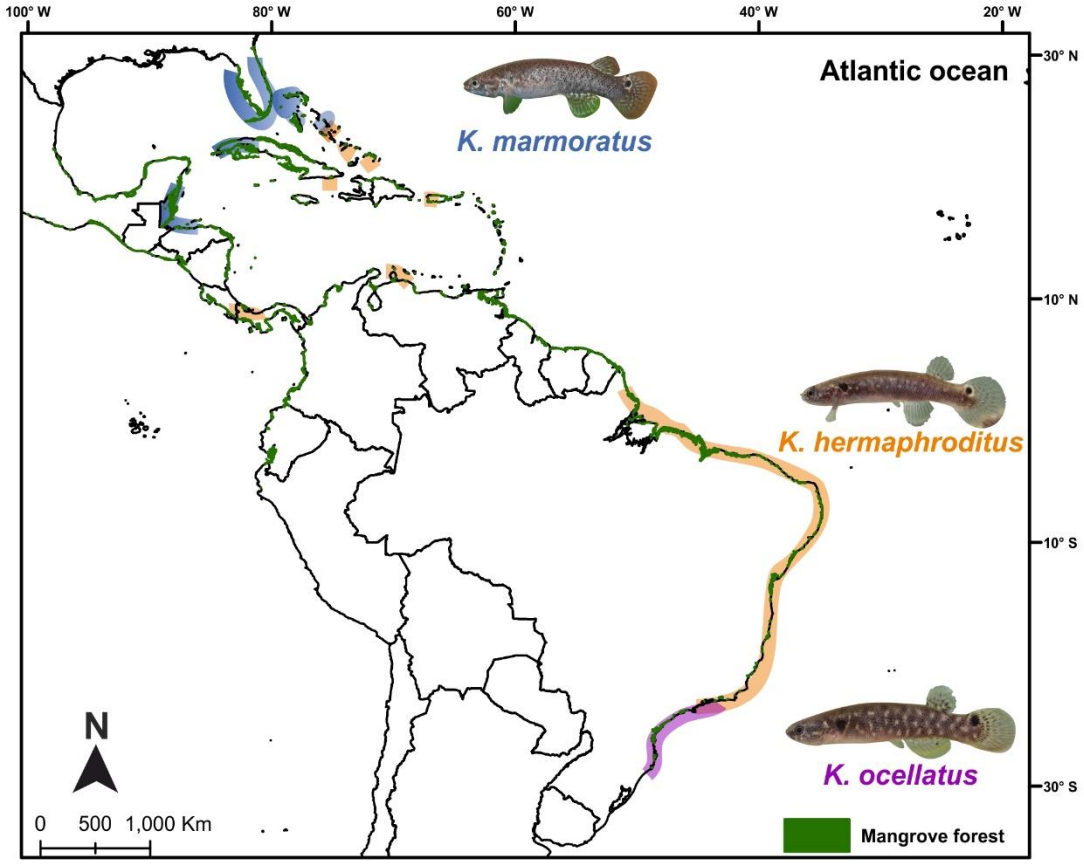


Figure I.4. Estimated geographical distribution of the three mangrove killifish species based on recent sampling surveys by Sergio Lima (personal communication unpublished data on *K. hermaphroditus*) and published data (Costa 2016; Guimarães-Costa et al. 2017; Tatarenkov et al. 2018; Tatarenkov et al. 2017).

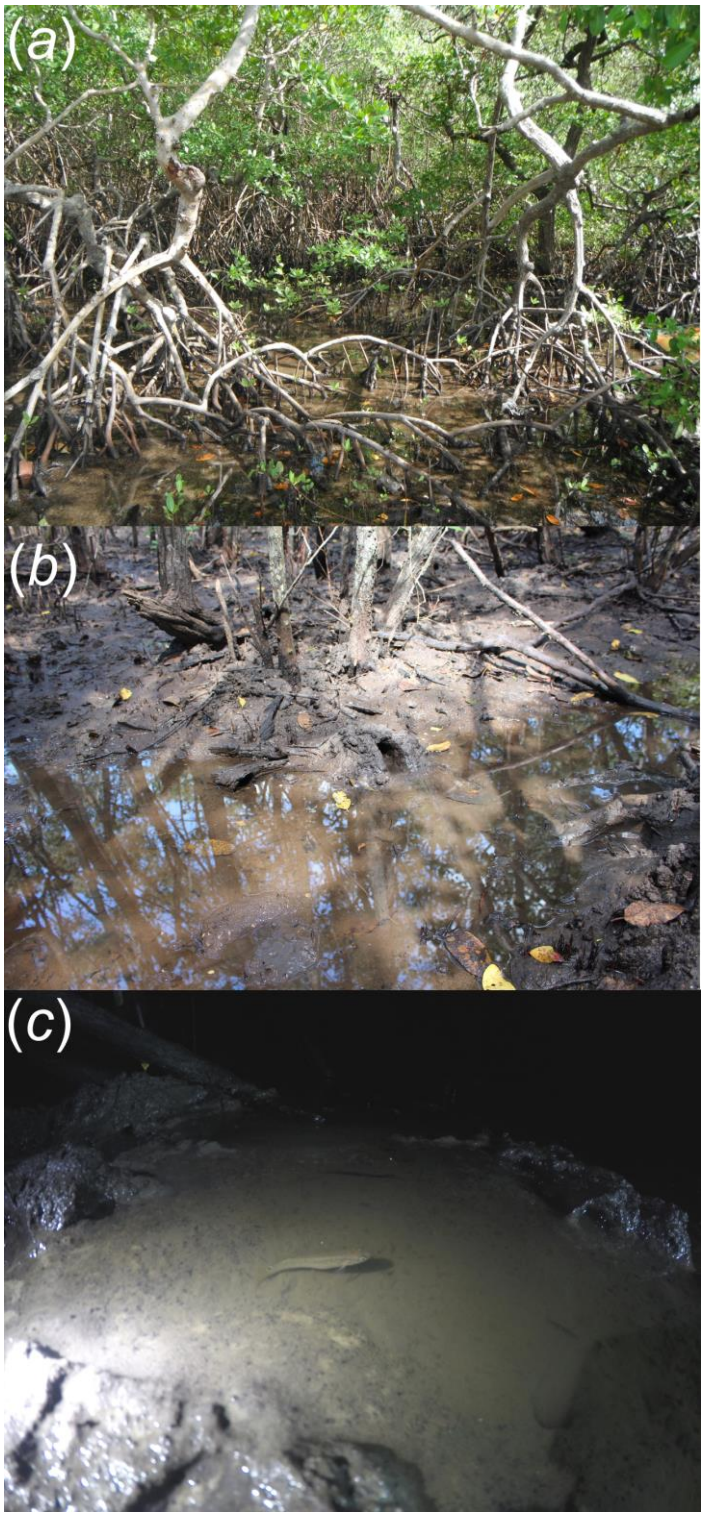


Figure I.5. (*a*) Temporary pools within mangrove forests, a typical habitat for mangrove killifishes. (*b*) Crab burrow within mangrove forests often used as microhabitats by mangrove killifishes. (*c*) *Kryptolebias hermaphroditus* individual sampled within a mangrove pool in Ceará-Mirim mangrove, northeast Brazil.

II. THESIS AIMS AND OBJECTIVES

The diversity of mating systems in the genus *Kryptolebias*, (predominantly selfing, mixed-mating, and obligately outcrossed, following mating systems definition by Barrett (2014)) especially within the mangrove killifishes clade (Costa 2016; Costa et al. 2010) represent unique opportunity among vertebrates to investigate the relative roles of genetic and epigenetic variation on ecology and evolutionary history of organisms. While there is evidence for the potential of epigenetic variation as an additional source of phenotypic diversity, several questions remain about the intricate relationship between genetic and epigenetic variation, especially how the former can interfere on the latter. The broad aim of this thesis was to investigate how genetic and environmental variation shape epigenetic variation in animals, using *Kryptolebias* species as model organisms. Using a combination of genetic, epigenetic, physiological and behavioural techniques in both laboratory-reared and wild *Kryptolebias*, this project aimed to understand several important questions on the relationships among genetic, epigenetic and environmental variation in the evolutionary history of organisms. Following are the main questions addressed on each chapter:

1. What is the structure and which factors shape epigenetic variation in natural populations of *Kryptolebias hermaphroditus*? **Chapter 1** used natural populations genotyped for microsatellite markers, amplified fragment length polymorphism (AFLP) and methylation sensitive amplified fragment length polymorphism (MS-AFLP) to investigate the genetic and epigenetic structure of the predominantly selfing species *K. hermaphroditus* in northeast Brazil and how ecological pressures (parasites) and/or genetic variability (genotypes and heterozygosity) influence DNA methylation

levels. Results are discussed in terms of the relationship between genetic and epigenetic variation and how it could contribute or even speed up local adaptation.

2. How are DNA methylation patterns transmitted to hybrids? **Chapter 2** investigates the genetic structure of *Kryptolebias ocellatus* across its known distribution in southeast Brazil using mitochondrial DNA (mtDNA) gene, microsatellites and single nucleotide polymorphisms (SNPs) and reports a remarkable case of hybridisation and introgression between the obligately outcrossed *K. ocellatus* with the predominantly selfing species *K. hermaphroditus*, representing the first case of hybridisation between species with different mating systems in vertebrates. This chapter further discusses the genetic structure patterns in *K. ocellatus* and the potential ecological causes and evolutionary consequences of the hybridisation between these two mangrove killifish species. **Chapter 3** further analyses the DNA methylation patterns of hybrids relatively to the parental species to investigate what are the major patterns of DNA methylation inheritance of hybrids in natural populations.

3. How autonomous are DNA methylation epialleles from the genetic background? **Chapter 4** uses reduced representation bisulphite sequencing (RRBS) from two selfing lines of *Kryptolebias marmoratus* reared under two contrasting environmental conditions (environmental enrichment) to investigate the distribution of DNA methylation epialleles according to its degree of autonomy from its underlying genetic background. Results are discussed in terms of the evolutionary relevance of epigenetic variation and how it can mediate the plasticity of downstream epigenetically-mediated phenotypes.

4. To what extent environmentally-induced epigenetic changes affect parental phenotypes and are inherited by the offspring? In **Chapter 5,** I analysed behavioural,

physiological (cortisol levels and metabolic rates) and epigenetic variation (RRBS data) of genetically-identical *Kryptolebias marmoratus* individuals reared in two contrasting environmental conditions. Using a fully-factorial experimental design, I evaluated the possibility of behavioural, physiological and epigenetic parental effects in the offspring and discussed its potential transmission mechanisms as well as its evolutionary implications.

CHAPTER 1: Genetic background and parasite load affect DNA methylation variation in a predominantly-self fertilising fish*

* A version of this work has been published as: **Berbel-Filho, W. M.**, C. G.de Leaniz, P. Móran, J. Cable, S. M. Q. Lima and S. Consuegra. 2019. Local parasite pressures and host genotype modulate epigenetic diversity in a mixed-mating fish. Ecology and Evolution 9:8736-8748.

1.1. Abstract

Parasite-mediated selection is one of the main drivers of genetic variation in natural populations. The persistence of long-term self-fertilization, however, challenges the notion that low genetic variation and inbreeding compromise the host's ability to respond to pathogens. DNA methylation represents a potential mechanism for generating additional adaptive variation under low genetic diversity. We compared genetic diversity (microsatellites and AFLPs), variation in DNA methylation (MS-AFLPs), and parasite loads in three populations of *Kryptolebias hermaphroditus*, a predominantly self-fertilizing fish, to analyse the potential adaptive value of DNA methylation in relation to genetic diversity and parasite loads. I found evidence that *K. hermaphroditus* occasionally outcross, being confirmed as the second example of mixed-mating species in vertebrates. We found strong genetic population structuring, as well as differences in parasite loads and methylation levels among sampling sites and selfing lineages. Globally, the interaction between parasites and inbreeding with selfing lineages influenced DNA methylation. At the local scale, when individuals were more genetically-similar, parasite load seemed more important than genotypic background in determining methylation levels.

1.2. Introduction

Mixed-mating organisms (alternating between self-fertilisation and outcrossing) benefit from the advantages of both biparental and uniparental reproduction: outcrossing generates genetic variability and adaptability potential, while selfing ensures reproduction without partners (Jarne and Charlesworth 1993), and reproductive assurance gives self-reproducing individuals an advantage when colonising new environments (Baker 1955). Offspring originated from selfing, however, can have reduced fitness compared to their outcrossed counterparts, known as inbreeding depression (Charlesworth and Willis 2009). Theory argues that occasional outcrossing should be beneficial when inbreeding is likely to reduce offspring fitness (Damgaard et al. 1992; Maynard-Smith 1978).

Among several theories for the origin and evolution of sex, the Red Queen hypothesis (RQH) (Bell 1982; van Valen 1973) is often invoked to explain the occurrence of sexual reproduction in face of the advantages of asexual reproduction (Blirt and Bell 1987; Lively 1987; Morran et al. 2011). According to the RQH, the more genetically diverse offspring of sexually reproducing individuals provide a "moving target" to coevolving antagonists (e. g. parasites, predators), making it more difficult for them to adapt compared to the "more static" offspring of asexual/uniparental individuals (Hamilton 1980; Maynard-Smith 1978). While the RQH is substantially supported by empirical data (Lively et al. 1990; Lively and Morran 2014) and sexual reproduction is observed in the vast majority of organisms (e. g. approximately 99% in animals Bell (1982)), asexual and self-fertilising lineages sometimes persist in a wide range of environments (Zhang et al. 2010). Given the short and long-term implications of inbreeding, the persistence of populations with limited capacity for genetically-based

adaptation (e.g. asexual, long-term self-fertilization) remain an evolutionary paradox (Verhoeven and Preite 2014).

Non-genetic mechanisms can represent an alternative source of phenotypic variation to genetically-depleted lineages deal with variable environments and coevolving antagonists (Angers et al. 2010; Castonguay and Angers 2012). Among non-genetic factors, epigenetic mechanisms (e.g. histone tail modifications, microRNAs, DNA methylation), can modulate changes in gene expression in response to environmental variation without involving changes in DNA sequence (Richards et al. 2017). DNA methylation is by far the best characterized epigenetic modification (Lea et al. 2017) 2017), and has important roles on pre-transcriptional control in several biological processes, such as cell differentiation and genomic imprinting (Jones 2012). DNA methylation variation is not completely independent from the genomes, with epialleles having different degrees of autonomy from the genotype (Richards 2006) and sometimes arising stochastically by errors in the enzymatic machinery in charge toreproduce and maintain cellular methylation patterns (Leung et al. 2016; Schmitz et al. 2011). Studies in plants and animals have been showing that indivdiuals with low levels of heterozygosity display higher genome-wide DNA methylation variation (Liebl et al. 2013; Richards et al. 2012; Schrey et al. 2012), which made some authors suggest that DNA methylation variation could be particularly important for organisms with limited genetic diversity (Castonguay and Angers 2012; Douhovnikoff and Dodd 2015; Verhoeven and Preite 2014). However, direct empirical evidence for the potential adaptive role of increased DNA methylation in inbred individuals has not been provided.

Mixed-mating organisms represent ideal models to test the associations between genetic and epigenetic variation with pathogen pressures because selfed and

outcrossed offspring can naturally coexist, usually displaying very different levels of genetic diversity. The genus *Kryptolebias* contains the only selfing known vertebrate species (*K. marmoratus* and *K. hermaphroditus*). *Kryptolebias marmoratus* populations are characterised by variable rates of selfing and outcrossing, however the possibility of outcrossing in *K. hermaphroditus* is yet to be confirmed. Populations of both species consist mainly of self-fertilizing hermaphrodites and varying levels of males at low frequencies (Tatarenkov et al. 2017; Berbel-Filho et al. 2016), and exhibit high levels of homozygosity (Tatarenkov et al. 2009, 2017), suggesting that self-fertilization is the most common mode of reproduction (Avise and Tatarenkov 2015). Negative associations between genetic diversity and parasite loads have been previously observed in *K. marmoratus* (Ellison et al. 2011). However, the relationship between epigenetic variation, parasites and mixed-mating species, has not been explored. Here, I evaluated the relationships between genetic background, heterozygosity, DNA methylation and parasite load in three natural populations of the mixed-mating mangrove killifish *K. hermaphroditus* distributed along the Brazilian coast (Tatarenkov et al. 2017). Based on the RQH, I predicted lower genetic diversity and higher parasite load in inbred compared to outbred individuals. Given the relationship between genetic background and DNA methylation reported in plants, I expected different DNA methylation profiles across selfing lines. I also predicted higher levels of DNA methylation in inbred individuals, which would be related to increased parasite loads, if DNA methylation played a role on host-pathogen interactions.

1.3. Material and Methods

1.3.1. Field sampling and parasite screening

Kryptolebias hermaphroditus is distributed along the Brazilian coast (Tatarenkov et al. 2017) and is typically found in shallow pools of high salinity levels (>30 ppt), clear waters and muddy substrates, where there are few other sympatric fish (Lira et al. 2015; Berbel-Filho et al. 2016). A total of 128 specimens of *K. hermaphroditus* were collected using hand-nets from three sampling sites on isolated mangroves on the North-eastern coast of Brazil between January and September 2015: Ceará-Mirim river – Site 1; Curimataú river – Site 2; Ipojuca river - Site 3 (Fig. 1.1). All specimens displayed the common hermaphrodite phenotype (dark colour with well-defined ocellus on the caudal fin; Costa 2011). Fish were euthanized using an overdose of tricaine methane-sulfonate (MS-222) following UK Home Office Schedule 1 (Hollands 1986), standard length was measured using a digital calliper (mm) and the whole fish were preserved in 95% ethanol at -20 ℃ for parasite screening and DNA extraction.

In the laboratory, fish were dissected and screened for both external and internal parasite infections using a dissecting microscope following the methods of Ellison et al. (2011). Macroscopic parasite analyses focused on the three most common types of parasites identified. To assess the reliability of parasite screening, a subsample of five fish was examined by a different observer and the agreement was 100%. I defined parasite load using a scaled measure of parasite abundance, where for each parasite morphotype (i), the number of parasites per individual (Ni) was divided by the maximum number found across all individuals (Nimax). The final value of the scaled parasite load represents the sum of scaled parasite loads across all parasite types. Given their uneven abundance (Table 1.1). This approach minimizes bias when parasite loads

are heavily influenced by a very abundant parasite type (in this case bacterial cysts) (Bolnick and Stutz 2017).

1.3.2. DNA extraction, microsatellite amplification and genetic analysis

Genomic DNA from all 128 fish was extracted from gill tissue using a Nexttec extraction kit for blood and tissue samples (Nexttec, Leverkusen, Germany). Gills are an important physical and immunological barrier to pathogens in fish (Press and Evensen 1999), and the organ where most parasites were found (Table 1.1). Twenty-seven microsatellite loci (Mackiewicz et al. 2006; Tatarenkov et al. 2017) were genotyped as in Ellison et al. (2011) and screened using GeneMapper v. 4.0 (Applied Biosystems, Foster City, USA). Loci were tested for linkage disequilibrium and Hardy-Weinberg equilibrium using GENEPOP v. 4.5.1 (Rousset 2008). Mean number of alleles per locus (N_{ma}), observed heterozygosity (H_o) and expected heterozygosity (H_e) were estimated using GenALEX v. 6.5 (Peakall and Smouse 2012). The inbreeding coefficient (F_{IS}) was calculated in GENEPOP. Global heterozygosity for individual fish was estimated using the homozygosity by locus index (HL) implemented by Cernicalin v. 1.3 using the default parameters (Aparicio et al. 2006).

I used the Bayesian clustering algorithm INSTRUCT (Gao et al. 2007) to estimate the optimal number of selfing lineages (K) given the expected deviation from Hardy-Weinberg equilibrium caused by self-fertilisation. INTRUSCT was run four simultaneous chains of 2,000,000 MCMC runs, 10 as thinning, and 100,000 of burn-in period, resulting in 100,000 interactions for each chain. The number of K tested ranged from 2 to 12. I used the individual *q*-values (the likelihood of membership to a particular genetic cluster or selfing lineage) from INSTRUCT to classify individuals as either selfed or outcrossed (Vähä and Primmer 2006). A threshold of *q*-value ≥0.9 was

used to classify selfed individuals, while <0.9 represented hybrids between two different selfing lineages, suggesting an outcrossing event (Ellison et al. 2011). Pairwise FST values among sampling sites and selfing lineages were estimated with Arlequin v. 3.5.2.2 (Excoffier and Lischer 2010) using 10,000 permutations. I used hierarchical analysis of molecular variance (AMOVA) to investigate population structuring among sampling sites and selfing lineages (according to individual *q*-values) using 10,000 randomizations. Differences between selfed and outcrossed groups in the total number of parasites and homozygosity by locus (microsatellites) were analysed using median Mann-Whitney rank tests implemented in R v. 3.3 (R Core Team, 2018).

1.3.3. Epigenetic analysis

Methylation-Sensitive Amplified Fragment Length Polymorphisms (MS-AFLPs) was used to assess genome-wide DNA methylation patterns. DNA extracted from gill filament tissue of 115 fish (33 classified as outcrossed and 82 as selfed according to the INSTRUCT *q*-values; 62, 36 and 17 from samplings sites 1, 2 and 3, respectively) was used for the MS-AFLPs analysis following Rodríguez-López et al. (2012). Briefly, a DNA aliquot of 100 ng per individual was split for digestion with two enzyme combinations: EcoRI/HpaII and EcoRI/MspI. The digested DNA was ligated to adaptors and a selective PCR was conducted using the primers ECORI-ACT: GACTGCGTACCAATTCACT and HPA-TAG: GATGAGTCTAGAACGGTAG following Ellison et al. (2015). The HpaII primer was end-labelled with 6-FAM. Fragments were run on an ABI PRISM 3100 (Applied Biosystems) and the resultant profiles were analysed using GENEMAPPER v. 4.0 (Applied Biosystems). To ensure reproducibility the following settings were applied: analysis range was 100-500 bp;

minimum peak height was 100 relative fluorescence units; pass range for sizing quality: 0.75-1.0; maximum peak width: 1.5 bp. To confirm MS-AFLP reproducibility, 24 individuals (~20% of the total; eight from each sampling site) were reanalysed and compared using the same protocols.

The R package msap v. 1. 1. 9 (Pérez-Figueroa 2013) was used to analyse MS-AFLP data. To increase reproducibility of the genotyping, I used an error threshold of 5% as suggested by Herrera and Bazaga (2010). According to the binary band patterns, each locus was classified as either methylation susceptible loci (MSL; i.e. displaying a proportion of HPA+/MSP- and/or HPA-/MSP+ sites which exceed the error threshold (5%) across all samples) or non-methylated loci (NML; if the same patterns did not exceed the error threshold) (Pérez-Figueroa 2013). MSL were used to assess epigenetic variation, while NML were used as a measure of AFLP genetic variation. Average group methylation percentages for inbreeding status were calculated using the different binary band patterns (hemimethylated pattern (HPA+/MSP-) + internal cytosine methylation pattern (HPA-/MSP+)/unmethylated pattern (HPA+/MSP+) + hypermethylation/absence of target (HPA/MSP-) *100) (Vergeer et al. 2012).

Epigenetic (MSL) and genetic at AFLPs (NML) differentiation among sampling sites, selfing lineages and between outcrossed and selfed groups, was assessed by AMOVA with 10,000 randomizations. Epigenetic (MSL) and genetic (AFLP and microsatellites) differentiation among sampling sites, selfing lineages and inbreeding status was visualized by principal coordinates analysis (PCoA). Mantel tests based on distance matrices (Mantel 1967) were used to test for potential correlations between epigenetic and genetic data for MSL, NML and microsatellites using GENALEX v. 6.5 with and 10,000 permutations. To identify disproportionately differentiated methylation states, I used a FST outlier approach implemented in Bayescan 2.1 (Foll

and Gaggiotti 2008), with $2x10^6$ iterations (thinning interval 20 after 20 pilot runs of 10^4 iterations each) and a burnin of $5x10^5$. I tested for outliers based on the MSL data generated on the comparisons among sampling sites, selfing lineages and between inbreeding status (inbred or outbred).

1.3.4. Statistical analyses

A Kruskal-Wallis test was used to examine the differences on scaled parasite load and bacterial cysts (the most prominent parasite) among selfing lineages. To test the relationship between genome-wide variation in global DNA methylation and parasite loads, the proportion of methylated loci per individual was calculated as the proportion of loci scored as methylated over the total number of loci observed per individual ("0" for unmethylated and "1" for methylated, excluding the missing data cells per individual). The proportion (or percentage) of methylated loci has been previously used to analyse differences in epigenetic profiles among groups (Groot et al. 2018; Vergeer et al. 2012). I then employed a generalized linear model with a binomial link to model proportion of methylated loci as a function of scaled parasite load, selfing lineage, sampling site and inbreeding status. To increase robustness, I repeated the analysis including only the most prominent parasite type (bacterial cysts). Model selection was conducted using the multi-model averaging approach implemented in the R package glmulti v 1.0.7 (Calcagno and de Mazancourt 2010). I chose the minimal adequate models based on the lowest AICc values (Akaike Information Criterion corrected for small sample size), Akaike weight (W_i) and evidence ratios (Burnham and Anderson 2004). Models (within 2 AIC units) were also reported.

To disentangle potential confounding effects arising from the unequal distribution of selfing lineages among sampling sites (i.e. five lineages are exclusive to a particular

sampling site, Table S1.1), I repeated the analyses (AMOVA, Mantel test, PCoA and GLMs) for both genetic (microsatellites and AFLPs) and epigenetic (MSL) data using only individuals from Site 1 (68 individuals for microsatellites and 62 for MS-AFLPs), as this was the only site with more than two selfing lineages (Table S1.1).

1.4. Results

1.4.1. Parasite screening

Bacterial cysts were present on the gills and consisted of white to yellow spherical cysts circumscribed by a capsule, which resulted in hypertrophied gill filaments. They were the most common type of pathogen appearing in 83.6% of the individuals screened, with a prevalence ranging from 1 to 19 (mean = 2.73, s.d. = ±2.99), and were most prevalent in Site 1 (mean = 3.16, s.d. = ±3.16), followed by Site 2 (mean = 2.66, s.d. = ±3.10) and Site 3 (mean = 1.27, s.d.= ±0.80). The second most common macroscopic parasites were protozoan cysts, which consisted of small dark oval cysts over the gills arch and filaments. In total, 19.53% of the total number of individuals were infected with these cysts, ranging from 1 to 6 (mean = 0.54, s.d. = ±1.26). Protozoan cysts were absent in Site 1, but present in Site 2 (mean = 1.52, s.d. = ±1.6) and Site 3 (mean = 0.33, s.d. = ±1.37). Finally, adult nematodes were found in the gut of only eight individuals (6.25%), ranging from 1 to 3 (mean = 0.09, s.d. = ±0.40). Nematodes were only detected in Sites 1 (mean = 0.3, s.d. = ±1.37) and 2 (mean = 0.02, s.d. = ±0.15) (Fig. S1.1; Tables 1.1; S1.1). Significant differences were found on scaled parasite loads (x^2= 32.14, p = <0.001, df = 5) and bacterial cysts (Chi square = 12.98, p = 0.01, df = 5) among selfing lineages.

1.4.2. Genetic diversity and population structuring based on microsatellites

No linkage disequilibrium was detected between any pair of microsatellite loci. As expected from the high levels of self-fertilisation of the species, no loci were found to be in Hardy-Weinberg equilibrium, and all 27 microsatellite loci showed an excess of homozygotes. The global homozygosity index (HL) was very high (mean = 0.95), as well the estimated selfing rates (Table 1.1). At the individual level, 93 individuals (72.6%) were homozygous across all 27 microsatellite loci. However, 17 individuals (13.28%), displayed intermediate to high levels of heterozygosity (ranging from 0.13 to 0.69).

The clustering Bayesian algorithm INSTRUCT indicated that six was the most likely number of selfing lineages (K). Selfing lineage 6 was shared between two different mangroves (Site 1 with seven individuals and Site 2 with one individual), separated by approximately 100 km. The other five lineages were solely represented in one of the mangroves (lineage 1 with 14 individuals, lineage 2 with 25 individuals and lineage 4 with 22 individuals in Site 1; lineage 3 with 41 individuals in Site 2; and lineage 5 with 18 individuals in Site 3) (Figs. 1.1; 1.2; Table S1.1). High F_{ST} values were found both among sampling sites (mean = 0.28, s.d. = ±0.02) and selfing lineages (mean = 0.32, s.d. = ±0.05). All pairwise comparisons were highly significant (Table S1.2).

Based on the resulting *q*-values from the INSTRUCT analysis, the fish were classified 92 fish (71% of the total; 46 from Site 1, 30 from Site 2, 16 from Site 3) were classified as selfed (with *q*-values ≥ 0.9), while 36 (29% of the total; 22 from Site 1, 12 from Site 2 and two in Site 3) as outcrossed (with *q*-values < 0.9) (Fig. 1.2; Table S1.1). The classification of individuals as selfed or outcrossed is based on the lineage composition, hence homozygote individuals can still be classified as originated from

outcrossing if they display alleles from different lineages, even if they appear in homozygosity after several generations of selfing. Overall, outcrossed individuals had significantly lower homozygosity by locus values (at microsatellites) and total parasite loads than selfed individuals (Table 1.2).

AMOVA analyses using microsatellites indicated strong and significant differentiation among sampling sites (F_{ST} = 0.28, P = 0.001) and selfing lineages (F_{ST} = 0.32, P = 0.001) (Table 1.3). Although significant, very low genetic differentiation was found between selfed and outcrossed individuals (F_{ST} = 0.01, P = 0.002) (Table 1.3; Fig. S1.2). These patterns were also seen on PCoA analysis, with individuals generally clustering by selfing lineages in the microsatellites data (25.84% of overall variation), with individuals from lineage 4 being the most differentiated from the other lineages on Site 1. In this site, substantial overlap was found among selfing lineages and between selfed and outcrossed, despite its significant differences (F_{ST} = 0.03, P = 0.001) (Table S1.4; Fig. S1.3).

1.4.3. Genetic and epigenetic variability and population structuring based on MS-AFLPs

The epigenetic analysis identified 381 MS-AFLP loci, of which 267 (70.07%) were methylation-susceptible loci (MSL) and 106 (27.82%) non-methylated loci (NML). Of the MSL loci, 236 (88.3%) were polymorphic and therefore used for the variability analysis. Reproducibility comparisons between original and replicated genotypes for 24 individuals revealed 238 loci with an average of 0.5% error rate (differences across individuals divided by the number of loci times number of replicates, as in Bonin et al. (2004)), which is within the normal reproducibility range for AFLPs genotyping (Bonin et al. 2004). AMOVA analysis for reproducibility also revealed no significant

differences between methylation and AFLP variation patterns between original and replicated set of individuals (Table S1.3). Averaged methylation ranged from 47.51% on lineage 2 to 38.17 % on lineage 5, and was 44.82% for inbred and 45.77% for outbred individuals.

AMOVA revealed very low but significant differentiation among sampling sites, for both genetic (AFLPs: $\phi_{ST} = 0.02$, P = 0.001) and epigenetic loci ($\phi_{ST} = 0.02$, P < 0.001). Significant differentiation among selfing lineages was also found on genetic (AFLPs: $\phi_{ST} = 0.02$, P = 0.004) and epigenetic loci ($\phi_{ST} = 0.02$, P = 0.001). Overall, higher genetic and epigenetic variance was found within than between groups (Table 1.3). As with microsatellites, no clear genetic at AFLPs or epigenetic differentiation was found between selfed and outcrossed individuals (Fig. S1.2). There was, however, a significant positive association between epigenetic (MSL) and genetic diversity, both using AFLPs (Mantel test, r = 0.11; P = 0.002) and microsatellites (r =0.09; P= 0.001). No MSL epiloci were identified as an F_{ST} outlier in any of the comparisons.

No significant differences among selfing lineages for individuals from Site 1 for AFLPs genetic data (selfing lineages: $\phi_{ST} = 0.008$, P = 0.12) or MSL epigenetic data (selfing lineages: $\phi_{ST} = 0.006$, P = 0.20) (Table S1.4). In the PCoA, substantial overlap was found among selfing lineages and between selfed and outcrossed individuals (Fig. S3). Mantel tests between genetic and epigenetic data indicated a significant positive association between AFLPs and MSL data (r = 0.21; P <0.001), but not between microsatellites and MSL (r = -0.005; P = 0.45).

1.4.4. Parasite loads, genetic and epigenetic variation

According to a multi-model testing approach, the most plausible model for the proportion of methylated DNA included selfing lineage, scaled parasite load,

inbreeding status and the interactions between selfing lineage and scaled parasite load and inbreeding. The proportion of methylated loci significantly varied among selfing lineages (estimate = 0.51, S. E.= 0.13, P < 0.001) and was affected by parasite loads and inbreeding status through its interactions with selfing lineage (parasite loads and selfing lineage: estimate = -0.55, S.E.=0.46, P = 0.005; inbreeding and selfing lineage interaction: estimate = -1.64, S.E.=0.14, P = 0.04) (Fig 1.3b-c; Tables 1.4 and S1.7). The second most likely model (ΔAICc=1.00) included only selfing lineage (estimate = -0.43, S. E. = 0.08, P < 0.001) and the interactions between inbreeding and selfing lineage (estimate = -1.10, S. E.=0.12, P = 0.04) as significant predictors. However, this model explained substantially less of the overall variation compared to the first model (weight: 0.17 vs. 0.28). and was 1.39 times less likely than the first one (Tables S1.5-S1.6).

Overall, the results of the single-taxa models (using number bacterial cysts) were very similar to those for scaled parasite loads. The best model to explain the proportion of methylated loci included selfing lineage, and the interactions between selfing lineage and bacterial cysts, and selfing lineage and inbreeding (Tables S1.7).

In the models with individuals from Site 1 (to remove any potential confounding effect between sampling site and selfing lineages) the model with the proportion of methylated loci with the lowest AIC indicated that selfing lineage, inbreeding and the interactions between inbreeding and selfing lineage and inbreeding and scaled parasite loads were all significant predictors (Table S1.8). However, the second best-fitting model (ΔAICc = 0.02) explained the same amount of variation (weight=0.39) and the evidence ratio (1.50) suggested that it was more likely than the first model. This second model indicated that overall, the proportion of methylated DNA significantly increased with scaled parasite loads (estimate = 0.43, S. E.= 0.11, P = 0.03) and that DNA

methylation levels were also affected by the interaction between scaled parasite loads and inbreeding (estimate = -1.29, S. E.=0.38, P <0.001), with inbred individuals having increased methylation levels with increased parasite loads, while outbred individuals had decreased methylation levels with increased parasite loads (Fig. 1.3d; Table 1.4).

1.5. Discussion

DNA methylation could play an important role in organisms with limited potential for genetically-based adaptation, including asexual and self-fertilising species, potentially increasing their plasticity capacity to cope with environmental change (Verhoeven and Pretie 2014; Douhovnikoff and Dodd 2015). However, my results did not indicate significant differences in genome-wide DNA methylation variation between selfed and outcrossed individuals, and my models only identified inbreeding status (defined as originating from selfing or outcrossing) significantly related to DNA methylation via its interaction with selfing lineage (all sampling sites) and parasites (at the local scale in Site 1). Higher variation in DNA methylation has been reported for clonal and inbred individuals (Liebl et al. 2013; Massicotte and Angers 2012; Richards et al. 2012), which has been suggested as a potential adaptive mechanism to surrogate genetic variation (Schrey et al. 2012), or alternatively, a consequence of inbreeding, representing a non-genetic mechanism contributing to inbreeding depression (Nakamura and Hosaka 2010; Vergeer et al. 2012). My results suggest that either inbreeding does not affect genome-wide DNA methylation variation or it does in a gene-specific manner (Venney et al. 2016).

Overall, inbred individuals harboured higher parasite loads than their outcrossed counterparts, supporting the prediction that low heterozygosity due to self-fertilisation may reduce fitness (considering parasite load as a proxy for pathogen pressure), as for other mixed-mating species (Ellison et al. 2011; King et al. 2011). Long periods of

self-fertilisation may reduce offspring fitness due to the accumulation of deleterious alleles and inbreeding depression (Charlesworth et al. 1993). Species with mixed-mating seem to overcome these problems through occasional outcrossing (Ellison et al. 2011; Morran et al. 2011), which can generate genetic diversity to face natural enemies, such as parasites (Lively 2014). Here, based on the high levels of individual heterozygosity observed in few individuals, I confirmed that outcrossing occasionally happens in *K. hermaphroditus*, making this species the second known vertebrate with mixed-mating system, together with its sister-species *K. marmoratus*. As shown in the latter, the relationship between parasites and inbreeding status (selfed or outcrossed) suggests that outcrossing might confer a fitness advantage, even when it occurs at very low frequencies (Ellison et al. 2011). However, despite the adaptive potential of outcrossing, the main reproductive strategy of *K. hermaphroditus* seems to be self-fertilisation. This suggests that other evolutionary mechanisms may be balancing the harmful effects of parasite infections, or that parasite selection is of low (Lively and Morran 2014). Theory predicts that low selection levels imposed by natural enemies are consistent with the maintenance of asexual reproduction (Judson 1997). For example, in the mixed-mating *Potamopyrgus* snails, the oldest asexual lineages are restricted to populations where parasites are rare (Neiman et al. 2005). Thus, the low number of parasites found in *K. hermaphroditus* (i.e. mean of 3.38 parasites per individual compared to 22.41 of *K. marmoratus* in Belize; Ellison et al. 2011), may explain the high prevalence of selfing in *K. hermaphroditus*.

I found that the different selfing lineages of *Kryptolebias hermaphroditus* distributed in three sampling sites of northeast Brazil differed significantly in parasite loads, genetic composition and DNA methylation patterns, which might indicate specific interactions between host genotypes, epigenotypes and parasites (Ebert 2008).

Previous studies on mangrove killifishes had identified extensive genetic structuring both between (Tatarenkov et al. 2015; Tatarenkov et al. 2017) and within mangrove systems even at close geographical proximity (Tatarenkov et al. 2012; Tatarenkov et al. 2007) (Tatarenkov et al. 2007; 2012; Ellison et al. 2012). I, here, also found strong evidence of genetic structuring between sampling sites and selfing lineages using microsatellites, but lower differentiation for AFLP genetic markers (likely due to the different mutation rate of the markers) and epigenetic markers (MS-AFLPs). As most of the selfing lineages were exclusive to specific sampling sites, I could not discard confounding effects between sampling sites and selfing lineages. Intriguing, however, are the differences revealed between selfing lineages within the same site. For example, lineage 2 is composed by nearly half outbred individuals, has the highest mean number of alleles, lowest inbreeding coefficient, homozygosity, selfing rate parasite loads among lineages from site 1. In contrast, lineage 6 is composed by only inbred individuals, has the lowest mean of alleles, highest inbreeding coefficient, heterozygosity, selfing rates among all sites, and it was also the mostly heavily parasitized. Site 1 was sampled in a relatively short period of time (January to March 2015), focusing on particular mangrove pools in an area no bigger than 200 m^2 in diameter (Lira et al. 2015 Berbel-Filho et al. 2016). Similarly, Ellison et al. (2012) found the highest genetic and ecological differences among *K. marmoratus* individuals from two sampling points only 150m distant to each other, suggesting that other factors than distance may be playing a role on the fine-scale genetic structure in mangrove killifishes.

Recent studies indicate that DNA methylation is likely to interact with genotypes in a genotype-by-environment manner to generate plastic responses (Herman and Sultan 2016). For example, Dubin and colleagues (2015) found strong influence of genetic

variants in DNA methylation levels in response to different temperature regimes in *Arabidopsis thaliana*. In humans, either the genotype alone or genotype-by-environment interactions in the uterus explained the variation of over a thousand differentially methylated regions on the methylome of neonates (Teh et al. 2014). Using data from all sampling sites, I found that genome-wide DNA methylation was strongly influenced by selfing lineage, and only at a smaller scale by inbreeding through its interaction with selfing lineage. Strong epigenetic differences between selfing lines had been identified previously in *K. marmoratus* (see Ellison et al. 2015), indicating an important role of the genetic background in the epigenetic variation of these species. In addition, I also found a significant correlation between DNA methylation and genetic variation (at both AFLP and microsatellites data), suggesting that autonomous variation in DNA methylation may be limited (Dubin et al. 2015).

Several abiotic and biotic factors, including parasites (Hu et al. 2018) as well as stochastic epimutations (Leung et al. 2016), are known to influence DNA methylation variation. My results showed that DNA methylation levels for all sampling sites were significantly influenced by parasite loads through the interaction with selfing lineage, suggesting a potential genotype-by-environment interaction on parasites responses. However, in Site 1, when selfing lineage were genetically more similar, genetic background did not have an effect on genome-wide DNA methylation, but parasites and their interaction with inbreeding status did. Overall, these findings suggest that genetic divergence have a major impact on DNA methylation profiles, however locally, when genetic differences are smaller, other environmental effects also effect genome-wide DNA methylation variation.

Increasing evidence has been showing that DNA methylation is involved in the modulation of host-pathogen interactions (Gómez-Díaz et al. 2012). The bacterial

parasite *Wolbachia*, for example, alters host genome-wide DNA methylation patterns resulting in the feminisation of infected leafhoppers (*Zyginidia pullulan*) to increase its transmission (Negri et al. 2009). Experiments in plants with both hyper- and hypomethylated mutants indicate that genome-wide DNA demethylation enhances immune responses to both bacterial (Dowen et al. 2012; Yu et al. 2013) and fungal infections (Le et al. 2014). Although I found evidence of parasites affecting DNA methylation variation, the anonymous nature of the genetic and epigenetic markers is a limiting factor to infer the potential adaptive/functional role of the DNA methylation variation in response to parasites. Further analyses, ideally under controlled experimental conditions and using higher resolution sequencing methods (i.e. whole-genome bisulphite sequencing, RNAseq), should help to clarify how DNA methylation may affect immune responses in mixed-mating *Kryptolebias* species.

The relationship between parasite loads and outcrossing seems to be common to several mixed-mating species (Steets et al. 2007; Ellison et al. 2011; King et al. 2011) in addition to *K. hermaphroditus*, suggesting that the influence of parasites in the regulation of mixed-mating could be generalised. The extent of this relationship, however, may depend on the severity of the selection imposed by coevolving parasites (Lively and Morran 2014). My results indicate that genotype composition (and its interaction with inbreeding) is important in DNA methylation responses to environmental variation in natural populations, and that, if DNA methylation responded in a genotypic-specific manner to parasites pressures, it could contribute to local adaptation (Smith et al. 2016). The mangrove killifish, with its naturally inbred populations and marked methylation differences between populations and genotypes, represents an ideal model to analyse the relative roles of genetic and epigenetic diversity in modulating local adaptation.

Table 1.1. Genetic diversity (at 27 microsatellite loci), mean parasite number (standard deviation in brackets) and parasite prevalence in *Kryptolebias hermaphroditus* at sampling sites in northeast Brazil. N = sampling size; N_a = mean number of alleles; H_e = expected heterozygosity; H_o = observed heterozygosity; F_{IS} = inbreeding coefficient; HL = homozygosity by locus; S = selfing rates.

	Site 1	**Site 2**	**Site 3**	**All sites**
Genetic diversity				
N	68	42	18	128
N_a	3.03	3.44	3.14	3.21
H_e	0.28	0.26	0.33	0.295
H_o	0.025	0.015	0.043	0.028
F_{IS}	0.91	0.94	0.87	0.93
HL	0.95	0.97	0.93	0.95
S	0.92	0.93	0.87	0.90
Parasite loads				
Bacterial gills cysts	3.16 (3.16)	2.66 (3.10)	1.27 (0.80)	2.73 (2.99)
Protozoan gills cysts	0	1.52 (1.60)	0.33 (1.37)	0.54 (1.26)
Nematodes	0.16 (0.53)	0.02 (0.15)	0	0.09 (0.40)
Total parasite load	3.33 (3.27)	4.21 (3.17)	1.61 (1.73)	3.38 (3.17)
Parasite prevalence (% of fish with infection)				
Bacterial gills cysts	91.17	71.42	83.33	83.59
Protozoan gills cysts	0	57.14	5.55	19.53
Nematodes	10.29	2.38	0	6.25

Table 1.2. Comparison of homozygosity by locus (HL) (at 27 microsatellite loci), mean parasites loads (standard error in brackets) and parasite prevalence between *Kryptolebias hermaphroditus* classed as either selfed or outcrossed based on q-values from selfing lineages structure estimated using INSTRUCT. P and z-values extracted from a two median Mann-Whitney test.

	Selfed	**Outcrossed**	**z**	***P* value**
Genetic diversity				
N	92	36		
HL	0.98	0.88	-4.76	**<0.001**
Parasite loads				
Bacterial gills cysts	3.25 (2.99)	1.69 (2.59)		
Protozoan gills cysts	0.57 (1.26)	0.47 (1.28)		
Nematodes	0.1 (0.4)	0.05 (0.42)		
Total parasite load	3.82 (3.47)	2.25 (1.94)	-2.84	**0.004**
Parasite prevalence (% of fish with infection)				
Bacterial gills cysts	89.13	69.44		
Protozoan gills cysts	18.47	22.22		
Nematodes	7.6	2.77		

Table 1.3. Hierarchical analysis of molecular variance (AMOVA) for microsatellites and MS-AFLPs data among (*a*) sampling sites, (*b*) selfing lineages and (*c*) selfed and outcrossed individuals in *Kryptolebias hermaphroditus*. df= degrees of freedom; SSD= sum of squared deviations; Mol. var. (%) = molecular variance percentages from variance components sigma 2; ϕ_{ST} = Phi statistics for population differentiation. *P* value derived from 10,000 permutations.

	Microsatellites				**NML**				**MSL**			
	df	Mol. var. (%)	F_{ST}	*P* value	df	Mol. var. (%)	ϕ_{ST}	*P* value	df	Mol. var. (%)	ϕ_{ST}	*P* value
(*a*) Sampling sites												
Among sites	2	28.46	0.28	**0.001**	2	2.20	0.02	**0.001**	2	2.96	0.02	**<0.001**
Within sites	227	71.54			112	97.80			112	97.05		
(*b*) Selfing lineages												

Among lineages	5	32.40	0.32	**0.001**	5	2.00	0.02	**0.004**	5	2.15	0.02	**0.001**
Within lineages	250	67.60			109	98.00			109	97.85		
(*c*) Inbreeding status												
Between selfed and outcrossed		1.28	0.01	**0.002**	1	0.15	0.02	0.32	1	0.82	0.02	0.06
Within selfed and outcrossed		98.72			113	99.85			113	99.18		

Table 1.4. Results of the best-fitting generalized linear models for proportion of methylated loci (binomial distribution) in *Kryptolebias hermaphroditus*, using the multi-model averaging approach (see appendix for the full model comparisons). df= degrees of freedom; Coeff = mean coefficient estimates.

Independent variable	**df**	**Coeff**	***z***	***P*-value**
Proportion of methylated loci				
Selfing lineage	5	-0.51	-4.50	**<0.001**
Scaled parasite load	1	-0.02	-0.02	0.83
Inbreeding	1	-0.50	1.73	0.15
Selfing lineage x parasite scaled	5	-0.55	-3.90	**0.005**
Selfing lineage x inbreeding	4	-1.64	-1.64	**0.04**
Proportion of methylated loci for site 1				
Scaled parasite load	1	-0.23	-11.49	**0.03**
Inbreeding	1	-0.31	-10.64	0.09
Inbreeding x scaled parasite load	1	-1.87	-17.93	**<0.001**

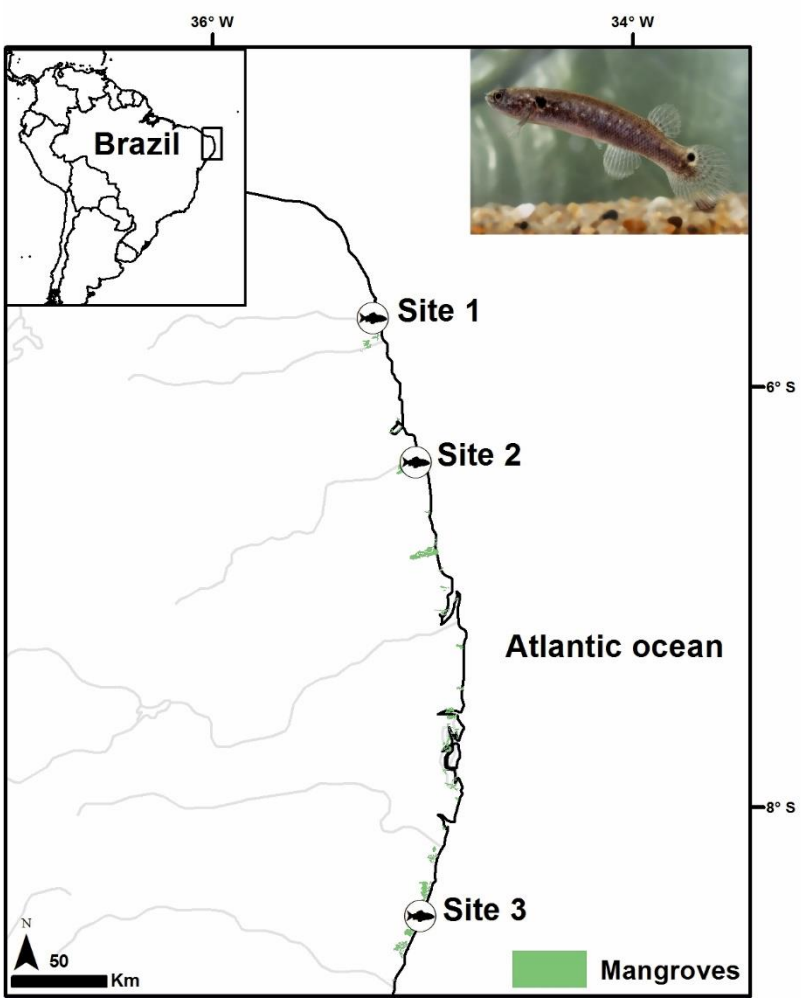


Figure 1.1. Sampling locations for *Kryptolebias hermaphroditus* (picture of a live individual on top-right corner) in northeast Brazil. Site 1 = Ceará-Mirim river; Site 2 = Curimataú river; Site 3 = Ipojuca river.

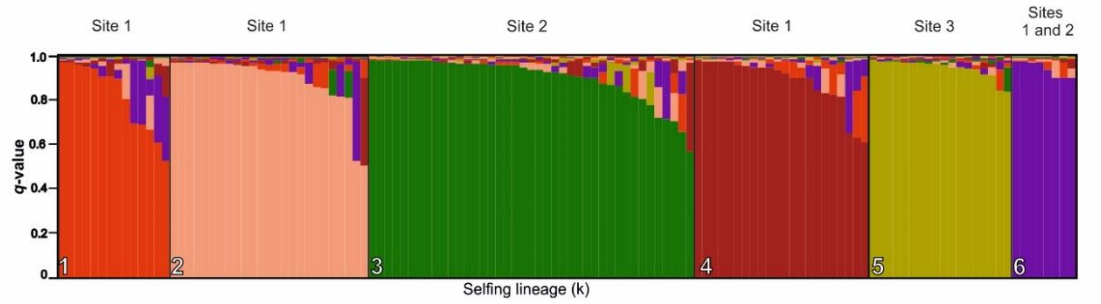


Figure 1.2. Genetic assignment of *Kryptolebias hermaphroditus* to six selfing lineages using INSTRUCT. Each individual is represented by a bar, which represents the likelihood of the individual to belong to a specific genetic cluster (colour)

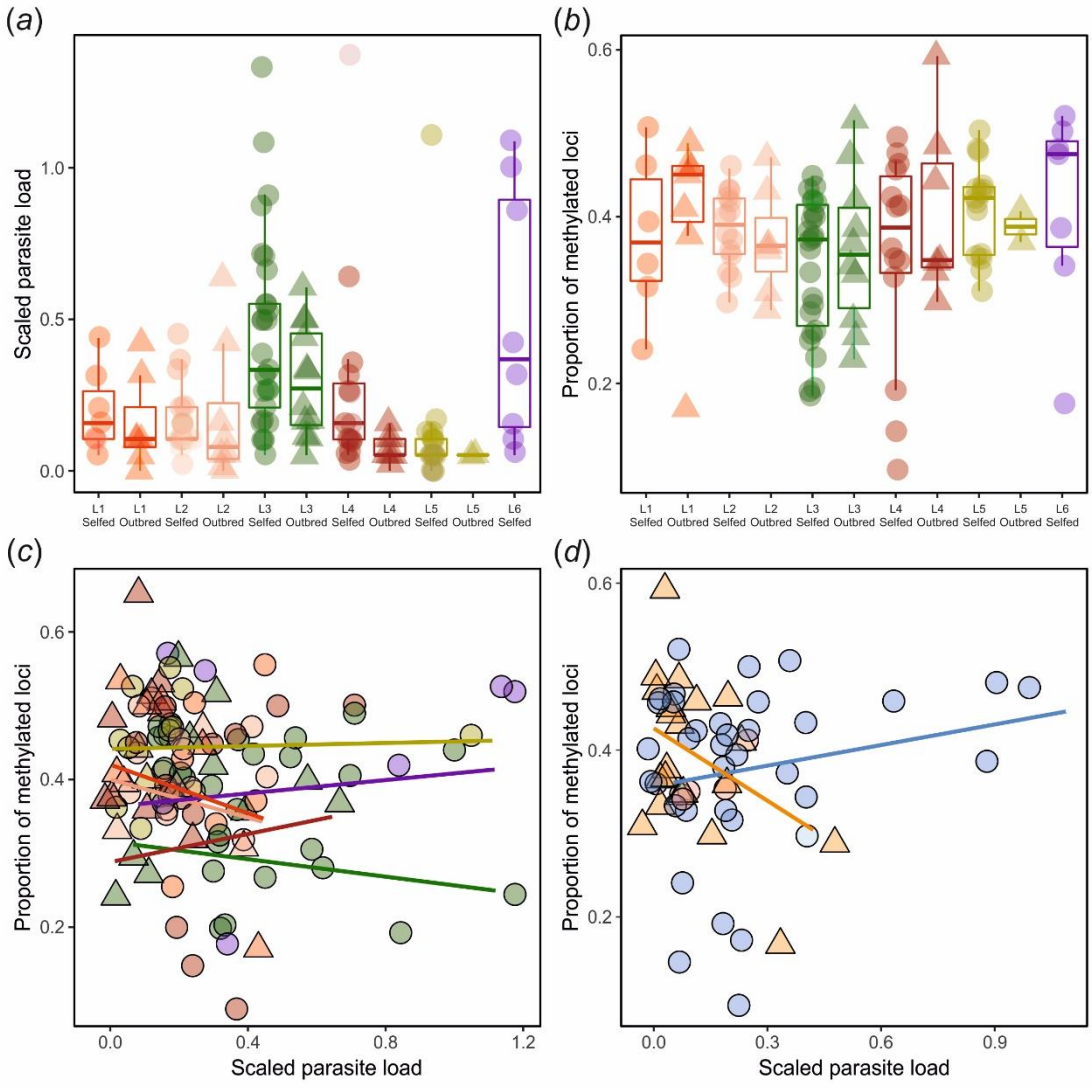
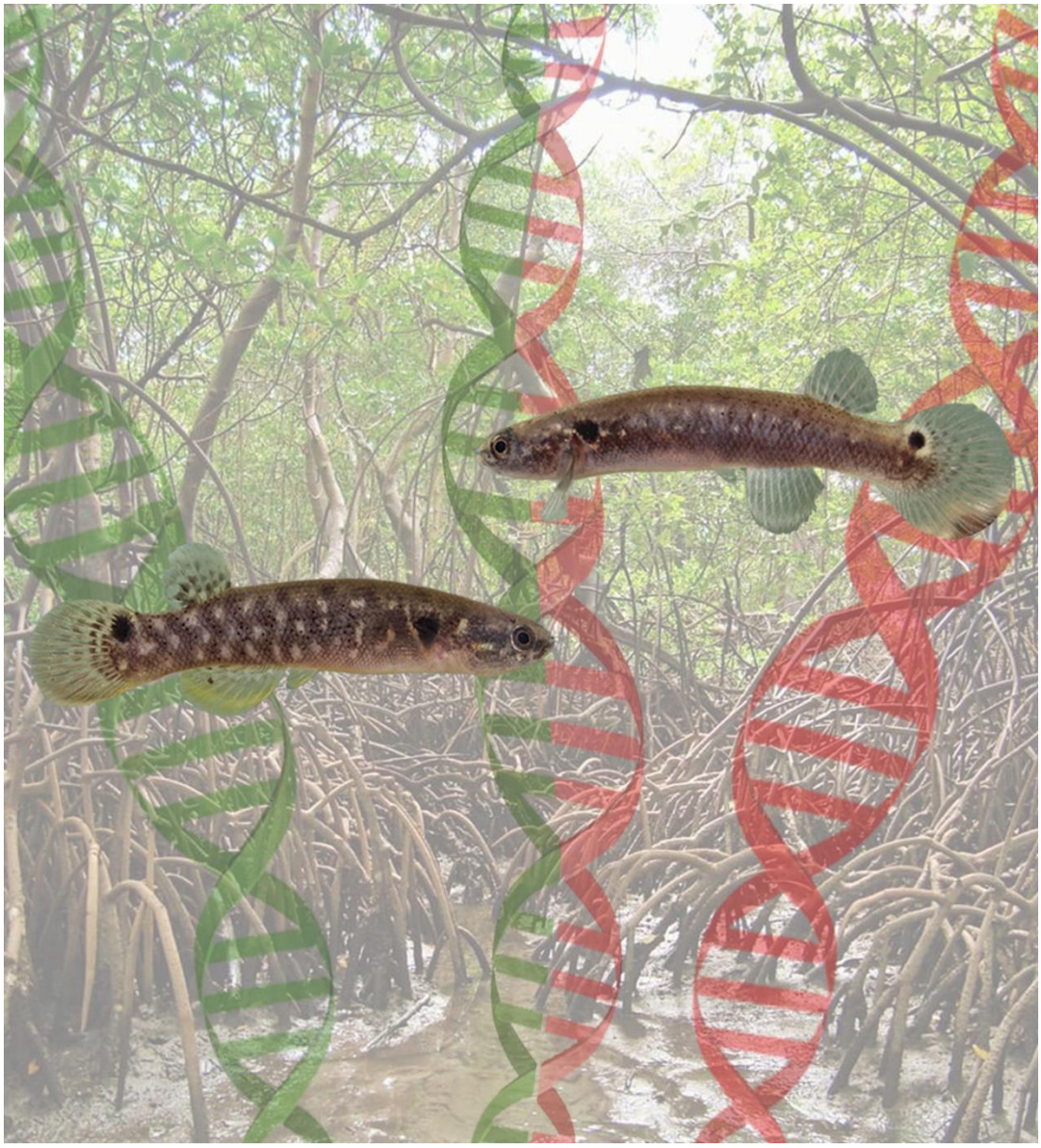


Figure 1.3. Relationships between (*a*) scaled parasite load across selfing lineages and inbreeding status, (*b*) proportion of methylated loci across selfing lineage and inbreeding status (selfed or outcrossed), (*c*) Proportion of methylated loci and selfing lineages and scaled parasite loads and (*d*) proportion of methylated loci across inbreeding status for sampling site 1 individuals. Circles for selfed, triangles for outcrossed individuals. Red = selfing lineage 1 (Site 1); salmon = selfing lineage 2 (Site 1); green = selfing lineage 3 (Site 2); brown = selfing lineage 4 (Site 1); yellow = selfing lineage 5 (Site 3); purple = selfing lineage 6 (Sites 1 and 2); orange = outcrossed individuals; blue = selfed individuals.

CHAPTER 2: Hybridisation between species with different mating systems revealed by the genetic structure of the mangrove killifish *Kryptolebias ocellatus*



2.1 Abstract

Mangrove forests are declining at an alarming rate. However, knowledge about the genetic structure of mangrove-associated organisms is still incipient. *Kryptolebias ocellatus* is an androdioecious (populations consisting only of males and hermaphrodites) mangrove killifish species inhabiting mangrove forests in southeastern Brazil. On its northernmost distribution, *K. ocellatus* is sympatric with the also androdioecious but predominantly selfing species *K. hermaphroditus* (composed of selfing hermaphrodites and very rare males). Here, I examined the genetic structure and phylogeographic patterns of *K. ocellatus* along its whole known distribution, using mitochondrial DNA, microsatellites and SNPs data. Specifically, I aimed to investigate whether *K. ocellatus* shows similar patterns of genetic structure along the southeastern Brazilian mangroves as its often-sympatric species *K. hermaphroditus*; and analyse any potential signal of hybridisation between *K. ocellatus* and *K. hermaphroditus* on areas where both species co-occur. Our results suggest absence of self-fertilisation in *K. ocellatus*, and a deeper genetic structuring within *K. ocellatus* in a much narrower range when compared to *K. hermaphroditus*, indicating older population establishment and/or lower dispersal capacity in *K. ocellatus*. I also found evidence of hybridization events in two of the three mangroves where both species are syntopic. Given the strong genetic divergence between the two species, the scarcity of males and the extremely low levels of outcrossing in *K. hermaphroditus*, the hybridisation between these two species is very unlikely and represents, the first example of hybridisation between species with different mating systems in vertebrates.

2.2. Introduction

Human activities, such as deforestation, urbanisation, harvesting and pollution modify the environmental conditions, ecological interactions and evolutionary pressures which organisms are subjected to (Allendorf and Hard 2009; Bull and Maron 2016; Heino et al. 2015). In particular, some of the most species-rich habitats are losing biodiversity and area at an alarming rate (Pimm and Raven 2000). Among those, mangrove forests are declining at a faster rate than tropical rainforests and coral reefs, mainly threatened by the growing aquaculture industry, urbanisation and pollution (Duke et al. 2007; Valiela et al. 2001). Despite this eminent conservation issue, the knowledge about the population dynamics and genetic structure of mangrove forests and their associated organisms is relatively scarce (Sandilyan and Kathiresan 2012).

Several fish species are known to inhabit or use mangroves during some period of their life cycle (e.g. foraging, breeding). However, very few are exclusively associated to mangroves (Taylor 2012). A rare example is the mangrove killifish clade from the genus *Kryptolebias.* The mangrove killifish clade is composed of the only known representatives among all rivulids (350+ species) living in brackish waters (Avise and Tatarenkov 2015; Costa 2011b; Costa et al. 2010), and perhaps more remarkably, the only two known examples of self-fertilising hermaphroditism among vertebrates (*K. marmoratus* and *K. hermaphroditus*, species that form the '*K. marmoratus* species complex', see Tatarenkov et al. (2017)). A third species compose the mangrove killifish clade: *K. ocellatus* (Costa 2011a; Costa et al. 2010). The tight association between western Atlantic mangrove environments and mangrove killifish clade distribution suggests that the current phylogeographic patterns observed in the killifish were influenced by the dispersal and distribution of the Atlantic mangrove forests. Therefore, information on the population structure of mangrove killifishes can be

informative in a broader context, possibly reflecting their genetic structure and colonisation patterns to other mangrove-dwelling organisms composing the Atlantic mangrove ecosystem (Tatarenkov et al. 2017).

Kryptolebias ocellatus (also known by its junior synonym *Kryptolebias caudomarginatus*, taxonomic nomenclature still under discussion (Costa 2011a; Huber 2017)) is a mangrove-dwelling species endemic from intermittent mangrove microhabitats in south and southeast Brazil (Costa 2016). Its populations are composed of males and hermaphrodites occurring approximately in an equal sex ratio (Costa et al. 2016; Costa et al. 2010). Based on the lack of departures from Hardy-Weinberg equilibrium, Tatarenkov et al. (2009) suggested that the species does not undergo selfing, however only two populations were analysed. On its northernmost known distribution (Guanabara and Sepetiba Bays, 22° S), it is often syntopic with *K. hermaphroditus* (Costa et al. 2016; Costa 2011a), a species mostly composed of self-fertilising hermaphrodites and extremely rare males (Berbel-Filho et al. 2016), which occasionally outcross but at very low frequencies (Berbel-Filho et al. 2019a). Extremely low levels of genetic diversity in *K. hermaphroditus*, especially in its southernmost distribution (where it is syntopic with *K. ocellatus*), suggest relatively recent dispersal and colonisation of this species in southeast Brazil (Tatarenkov et al. 2009; 2011; 2017). Recently, Tatarenkov et al. (2018) found evidence of hybridisation between *K. marmoratus* and *K. hermaphroditus* ('Central clade') suggesting that divergent *Kryptolebias* lineages are able to interbreed when in sympatry. Although some studies have reported that *K. ocellatus* is usually more abundant than *K. hermaphroditus* in the mangrove microhabitats where both species are syntopic (Costa 2016; Costa et al. 2010), further details about the ecological interactions between the two species remain unknown.

Considering that both species occupy shallow temporary pools in mangrove forests in southeast Brazil and show high similarities in general body shape and colour patterns (Costa 2016; Tatarenkov et al. 2017) (Fig.2.1), morphologically-based taxonomic classification of the two species has been historically confuse (Costa 2006; Costa 2011a; Costa 2016). However, their morphological similarity does not seem to reflect shared evolutionary history. Phylogenetic studies revealed that *K. ocellatus* is genetically very divergent from *K. hermaphroditus* and is likely to be the sister-species of the clade containing the selfing species from *K. marmoratus* species complex (*K. marmoratus and K. hermaphroditus*) (Kanamori et al. 2016; Tatarenkov et al. 2017; Tatarenkov et al. 2009; Vermeulen and Hrbek 2005). All the other known *Kryptolebias* species are freshwater and dioecious, suggesting that hermaphroditism arose within the mangrove killifish clade with reduction of males and selfing later evolving in the *K. marmoratus* species complex (Avise and Tatarenkov 2015; Costa et al. 2010).

The genus *Kryptolebias* has been of particular scientific interest due to its diversity of mating systems (dioecious, androdioecious, and mixed-mating species) (Avise and Tatarenkov 2015; Costa et al. 2010). Since the discovery of self-fertilisation in *K. marmoratus* (Harrington 1961), extensive information has been collected on the ecology, physiology and evolution of species within the *K. marmoratus* species complex (Kanamori et al. 2016; Tatarenkov et al. 2017; Tatarenkov et al. 2009; Taylor 2012; Wright 2012). However, the other South American *Kryptolebias*, including species within the mangrove killifish clade, remain understudied (Berbel-Filho et al. 2016; Costa et al. 2010; Lira et al. 2015). Therefore, I investigated the genetic structure and phylogeographic patterns of *Kryptolebias ocellatus* across its whole known distribution in southern and southeast Brazil using mitochondrial DNA, microsatellites and single-nucleotide polymorphisms (SNPs) data. More specifically, I aimed to: (1)

analyse the possibility of self-fertilisation in *K. ocellatus*; (2) investigate whether *K. ocellatus* shows similar patterns of genetic structure (low genetic structure and high gene flow) along the Brazilian mangroves as its often syntopic species *K. hermaphroditus*; (3) investigate any potential signal of hybridisation and introgression between *K. ocellatus* and *K. hermaphroditus* in populations where both species co-occur.

2.3. Material and Methods

2.3.1. Samples

I sampled *K. ocellatus* and *K. hermaphroditus* (on its syntopic populations) in south and southeast Brazil (Costa 2016), covering the entire known distribution of *K. ocellatus* between August and September 2017, using hand nets in mangrove microhabitats (e. g. mangrove temporary pools and crab burrows) (Fig. 2.1; Table 2.1). The sex of the fish (male or hermaphrodite) was inferred by body and fin coloration patterns as follows: in *K. ocellatus*, males were identified by a black spot on the dorsal part of the caudal fin, a distinguishable trait for males on this species (Costa 2016); in *K. hermaphroditus*, the presence of males was checked by the presence of a broad black margin along the whole caudal fin, bordered by a broad sub-marginal white zone as described in Costa (2016). Sampling work was carried out under license ICMBio/SISBIO 57145-1/2017.

2.3.2. Microsatellites and mtDNA markers

A subset of 16 microsatellites from (Mackiewicz et al. 2006a) were amplified and genotyped following protocols described in Tatarenkov et al. (2010). One mitochondrial gene, *cox1* (cytochrome oxidase subunit I gene) was also used to

investigate the genetic structure and major lineages distribution. Microsatellite genotypes were obtained for 190 individuals and *cox1* was sequenced in 129 individuals of *K. ocellatus*. Iriri (IRI) mtDNA and part of Guaratiba (GUA) microsatellite data was retrieved from individuals sampled in Tatarenkov et al. (2009). *Kryptolebias hermaphroditus* from three locations in southeast Brazil (Table 2.1), two of which were syntopic with *K. ocellatus*, were genotyped for the same markers.

A 618 bp region of the *cox1* was amplified with FishCOI-F (5′-TCAACYAATCAYAAAGACATYGGCAC-3′) and FishCOI-R (5′-ACTTCYGGGTGTCCRAARAAYCA-3′) primers as in Tatarenkov et al. (2017). PCR amplification conditions were as follow: initial denaturation of 95°C for 5 min; 30 cycles of denaturation at 95 °C for 40 s, annealing at 52 °C for 40 s and extension at 72 °C for 1 min; and a final extension step of 7 min at 72 °C. Both forward and reverse DNA strands were sequenced and assembled using Geneious v. 9.1.8 (www.geneious.com).

2.3.3. mtDNA phylogenetic and phylogeographic analysis

A Bayesian coalescent reconstruction was carried out using BEAST v. 2.5.1 (Bouckaert et al. 2019). The sequences dataset was reduced to the *cox1* haplotypes found for the 129 *K. ocellatus* individuals and using the only haplotype found for *K. hermaphroditus* in southeast Brazil as an outgroup. The best fitting model of nucleotide substitution was selected according to the Akaike and Bayesian Criteria on jModelTest2 (Darriba et al. 2012). The substitution model selected was 3-paratemer model with unequal base frequencies (TPM1uf). To time-calibrate the phylogenetic reconstruction and allow for rate variation among lineages, a lognormal relaxed molecular clock of 0.009 substitutions per site per million years was used, based on

the *cox1* Goodeidae fossil-calibrated molecular rate described in Webb et al. (2004). I performed three independent runs of 10,000,000 Markov Chain Monte Carlo (MCMC) runs, sampling every 1,000. I used Tracer v. 1.7.1 (Rambaut et al. 2018) to access convergence and effective sample sizes (≥ 200) among MCMC runs. I used the software TREEANNOTATOR v. 2.5.1 (Bouckaert et al. 2019) to discard the first 2000 trees (20%) as burn-in, and to generate a consensus tree with posterior probability value for each clade.

The number of haplotypes (*H*) and polymorphic sites (S), haplotype (*h*) and nucleotide diversities (π) for each sampling location and major clades were calculated using DNAsp v. 6.10.04 (Rozas et al. 2017). For performing neutrality tests and generating pairwise fixation indices (F_{ST}) among major clades and sampling locations, I used Arlequin v. 3.5.2.2 (Excoffier and Lischer 2010). I used Mega v. 7.0.26 (Kumar et al. 2016) to calculate Kimura-2-Parameter (K2P) genetic distances among major clades and sampled populations. To visualise haplotypes distribution and divergence, I reconstructed a *cox1* haplotype network using POPART (Leigh and Bryant 2015).

I used analyses of molecular variance (AMOVAs) with 1,000 permutations in Arlequin v. 3.5.2.2 to test different hypothesis of partition: (1) major clades recovered by Bayesian phylogenetic reconstruction of *cox1* (Northern/Parati/Southern), (2) major clades recovered by microsatellites analysis using Evanno's method ΔK (Northern/Southern, see below), (3) paleorivers basins according to Thomaz et al. (2015) (IRI+FUN/GUA+PRT/PAR/SFR/FLO).

2.3.4. Statistical analysis with microsatellite data

Micro-checker v. 2.2 (van Oosterhout et al. 2004) was used to check for errors in the data and presence of null alleles. To assess overall differentiation at the population

level, I used FSTAT v. 2.9.3.2 (Goudet 1995) to calculate F_{ST} and conduct exact G - tests based on 10,000 randomizations of alleles. FSTAT was also used to measure departures from Hardy–Weinberg equilibrium. *P* values for F_{IS} for each locus were based on 2240 randomizations, and *P* values over all loci were calculated from a weighted average of the statistic obtained for each locus. To ensure the neutrality of the markers, I tested for linkage disequilibrium and deviation from Hardy–Weinberg equilibrium using GENEPOP v. 4.2 (Rousset 2008). Unbiased expected (HE) and observed heterozygosity (HO) were calculated using MSA v. 4.05 (Dieringer and Schlötterer 2003).

2.3.5. Genetic structuring and clustering analysis with microsatellite data

I tested the same hypotheses of genetic partition as for the *cox1* haplotypes using microsatellites genotypes with AMOVA in Arlequin v. 3.5.2.2. The overall genotypic associations of individuals were visualized with a factorial correspondence analysis (FCA) using the procedure implemented in GENETIX v. 4.04 (Belkhir 2004).

I applied three different methods to estimate the most likely number of genetic clusters (K) across *K. ocellatus* distribution. First, using only microsatellite data I ran STRUCTURE 2.3.4 (Pritchard et al. 2000) with the following parameters: 10 iterations per K, a total of 1,000,000 MCMC and 100,000 burn-in, admixture model, independent allele frequencies, and testing K ranging 1–10. To identify the uppermost hierarchical level of genetic structure, I chose the most likely K value using second-order rate of change of likelihood ΔK method (Evanno et al. 2005), implemented in Structure Harvester (Earl 2012). Independent STRUCTURE runs were aligned and plotted using CLUMPAK (Kopelman et al. 2015).

Given the uneven number of individuals on my sampling, I input the STRUCTURE runs into STRUCTURESELECTOR (Li and Liu 2018), which used four metrics of cluster estimates (median of means (MedMeaK), maximum of means (MaxMeaK), median of medians (MedMedK) and maximum of medians (MaxMedK)) implemented by Puechmaille (2016) to investigate the most likely number of genetic clusters.

Finally, to integrate mtDNA, microsatellites data and spatial information, I used Geneland v. 4.0.8 (Guillot et al. 2008), which takes into account spatial information from each individual, also allowing for uncertainty in the positioning of sampled individuals. To first identify the spatial population distribution and then access individual assignment to the inferred most likely K value, I followed Guillot et al. (2005). First, I input individual mtDNA (129 individuals), geographical coordinates (geo-referenced according to the sampling points and allowing for uncertainty of ±0.05 in both latitude and longitude) and genotype data for 190 individuals with K range 1-10. Ten multiple runs were run with 10,000,000 MCMC iterations, sampled every 1,000 iterations. Once the K value was inferred from the modal value across the 10 multiple runs, I ran the MCMC again with other 10 multiple runs and K fixed to assigned value. These final 10 runs were postprocessed (with a burn-in of 20%) in order to obtain posterior probabilities of population membership for each individual. All Geneland analyses were performed using "geneland" R package (Guillot et al. 2008).

2.3.6. Hybridisation between *K. ocellatus* and *K. hermaphroditus*

In order to investigate the possibility of hybridisation between *K. ocellatus* and *K. hermaphroditus* in their syntopic populations (IRI, FUN and GUA) (Fig. 2.1), I genotyped 67 *K. hermaphroditus* (identified using *cox1*) individuals from three

populations in southeast Brazil (FUN, GUA, PIC) for the same 16 microsatellites studied for *K. ocellatus* (Table 2.1). I repeated the STRUCTURE analysis including both species with the same parameters described previously. NEWHYBRIDS v. 1.1 (Anderson and Thompson 2002) was used to estimate the posterior probability of each individual to belong to one of the six classes (parental species 1, parental species 2, F1, F2, backcross 1 (between parental species 1 and F1), and backcross 2 (between parental species 2 and F1) based on their allele frequencies. The analysis was run using the default genotype proportions, using the uniform prior option and a burn-in period of 50,000 iteration and 300,000 MCMC sweeps.

2.3.7. Genotype-by sequencing library

The genomes of a subset of 55 individuals (33 *K.* ocellatus and 22 *K. hermaphroditus* according to the *cox1* mtDNA haplotypes; Table S2.1) were sequenced using a reduced representation, which follows the original genotype-by-sequencing (GBS) protocol described in Poland et al. (2012). Briefly, genomic DNA was digested using a rare cutter restriction enzyme (RE) EcoRI (cutsite: GAATTC) and RE HpaII (cutsite: CCGG). Then, digested DNA was ligated to barcoded adapters with a HpaII cut site overhang and the EcoRI Y adapter. The ligation products were individually cleaned to remove excess of adapters using Agencourt AMPure XP purification system (#A63880, Beckman Coulter, Brea, CA, USA) at a ratio of 0.85 as described by the manufacturer's instructions. GBS libraries were produced by pooling 20ng of DNA from each sample. Each library was amplified in eight separate PCR reactions (25 µL each), containing 10 µL of library DNA, 5 µL of Q5 high fidelity buffer, 0.25 µL of Q5 high fidelity polymerase, and 1 µL of each Forward and Reverse common primer at 10 µM, 0.5 µL of 10 µM dNTP and 7.25 µL of pure sterile water. DNA fragments

were size-selected (range 200 – 350 bp) using Agencourt AMPure XP magnetic beads. Next, libraries were sequenced in an Illumina NextSeq500 sequencer platform (Cardiff University Genomics Research Hub, Cardiff, UK) to obtain 125 bp paired-end reads.

2.3.7.1 Data analysis

Paired-end reads were processed using a combination of packages in Linux bash shell environment. First, I used the software GBSX v 1.3 (Herten et al. 2015) to demultiplex the paired-end reads data allowing for one mismatch in the barcodes (-mb 1), no mismatch in the enzyme cut-site (-me 0), and ensuring that no common sequencing adapter was to be removed (-ca false). I then filtered (-qtrim r; -minlength 25) and merged the reads by individuals using BBmap tools (Bushnell 2014) (merging and filtering), mapped to *Kryptolebias marmoratus* reference genome (Rhee et al. (2017), assembly size = 680.3Mb; number of scaffolds = 3,073; N50 = 2,229,659; GC content = 37.76%) using Bowtie 2 v. 2.2.3 using default parameters (Langmead and Salzberg 2012) and used generated, filtered and indexed individual BAM files using samtools v. 1.9 (Li et al. 2009).

To call genotypes, I used the package ANGSD v 0.9.2.9 (Korneliussen et al. 2014) with the following parameters: minimum mapping quality (-minMapQ 30), minimum base quality (-minQ 20), missing data (-minInd 95%), Global Depth (-setMaxDepth 500X per individual), minimum genotype posterior probability (-postCutoff 0.95), single and double-tons were accordingly removed based on minimum minor allele frequencies (-MinMaf), anomalous reads (-remove_bads 1; SAM flag above 255), adjusted mapping quality for excessive mismatches (-C 50), performed BAQ computation (-baq 1), minimum coverage for genotype calling (-geno_minDepth 3), use of SAMtools genotype likelihood model (-GL 1), and estimated posterior genotype probabilities assuming an uniform prior (-doPost 2). In addition, I used the ANGSD

SNP calling method (-SNP_pval 1e-6), where a Likelihood Ratio Test is used to compare between the null (maf = 0) and alternative (estimated maf) hypotheses by using a X^2 distribution with one degree of freedom.

2.3.7.2 Population genetics estimates

For general population genetics estimates, I followed the instructions provided by ANGSD to compute the unfolded global estimate of the Site Frequency Spectrum (SFS) in order to calculate the observed proportion of heterozygous sites (PHt) per sample (Korneliussen et al. 2013). The observed fraction of heterozygous sites was calculated as the ratio between the number of heterozygotes and the total number of sites with information in percentage.

In order to visualise the overall relationships among individuals based on the SNPs data, a pairwise genetic distances matrix was computed directly from the genotype likelihoods outputted by ANGSD using the software ngsDist v.1.0.2 (Vieira et al. 2015). The matrix of genetic distances was used to conduct a Multidimensional Scaling (MDS) analysis using the R package *cmdscale*.

To estimate the individual ancestries, I used the software ngsAdmix v. 3.2 (Skotte et al. 2013) ranging K between 2-10 in 100 replicates using default parameters, except for tolerance for convergence (-tol 1e-6), log likelihood difference in 50 iterations (-tolLike50 1e-3), and a maximum number of EM iterations (-maxiter 10000).

2.3.7.3 Individual ancestry and hybridisation analyses

To investigate potential hybridisation between *K. ocellatus* and *K. hermaphroditus* with SNPs data, I subsampled the data containing only individuals from FUN and GUA for both species (39 individuals, 17 with *K. ocellatus cox1* haplotype, 22 with *K. hermaphroditus cox1* haplotype). I called SNPs with the same parameters described

above using the package ANGSD, with the only difference that I allowed for no missing data here. I selected SNPs with the highest pairwise F_{ST} values between species using ANGSD. I used the software ngsLD (Fox et al. 2019) based on genotype likelihoods using sites with 0.03 as minor allele frequency (--min_maf 0.03). I plotted the R^2 estimates using the fit_LDdecay.R script provided by ngsLD setting 500 kb as maximum distance between SNPs (--max_kb_dist 500), exhaustive fitting (--fit_level 100) and a fitting bin size of 200 bp (--fit_bin_size 200). Pair of SNPs with significant (LD) were removed and randomly replaced with other SNPs until a dataset of 200 SNPs (number limited by NEWHYBRIDS) with low-levels of LD and high FST values was reached. I then run NEWHYBRIDS with the same parameters described for the microsatellites to investigate the posterior probability of each individual to belong to one of the six hybrid classes.

2.4. Results

2.4.1. mtDNA phylogenetic and phylogeographic analysis

Twenty-two *cox1* haplotypes (618bp-long) were recovered from 129 *K. ocellatus* individuals sequenced. In contrast, only one cox1 haplotype was found for *K. hermaphroditus* across 61 samples (Table 2.2). A clear geographical pattern was found by the Bayesian reconstruction tree using *cox1* haplotypes in *K. ocellatus* (Fig. 2.2). Three major lineages were found: a clade composed of haplotypes from sampling locations within Guanabara and Sepetiba's Bays (IRI, FUN and GUA; hereafter called Northern clade), clustered with a clade containing haplotypes from the opposite side of Sepetiba Bay (PRT; hereafter called Parati clade), however the support for the grouping of Northern and Parati clades was low (PP: 0.66). The third clade was composed of haplotypes from sampling points in south Brazil (PAR, SFR, FLO;

hereafter called the Southern clade). Overall, the c*ox1* haplotype network and the distance-based tree using microsatellites data also supported the existence of these three major clades (Figs. 2.2; S2.1).

Overall haplotype diversity was 0.895, being the highest in the SFR (Southern clade) and the lowest in GUA (Northern clade) populations. Nucleotide diversity was generally low (π=0.007), but also followed a similar pattern to the haplotype diversity, being the highest at PAR (Southern clade) and the lowest at GUA (0.0004). The same pattern of haplotype and nucleotide diversity was also seen when sampling locations were grouped according to the major mtDNA clades, with the Southern clade being the most diverse, followed by the Northern and Parati clades, respectively (Table S2.3).

F_{ST} values were very high (mean= 0.72). All F_{ST} pairwise comparisons among sampling locations were significant, with the exception of the comparison between SFR and FLO (within the Southern clade). The highest F_{ST} value (0.92) was found in the comparison between FLO and PRT (Parati clade), while the lowest (0.14) was found between SFR and FLO (Table S2.4). All F_{ST} pairwise comparisons were significant when grouping sampling locations into the major mtDNA clades (mean = 0.80).K2P genetic distances followed a similar pattern to the F_{ST} comparisons, with the highest values (1.4%) being observed between samples from the Southern and Parati clades, while the lowest (0.2%) between samples of the same mtDNA clade (Table S2.4).

2.4.2. Microsatellite variation within populations

On average, 22.4 *K. ocellatus* and 16.8 *K. hermaphroditus* individuals were genotyped at 16 microsatellite loci per sampling location, but sample size varied considerably

(from 5 to 51) (Table 2.2). Overall, there was high level of variation at microsatellite loci in *K. ocellatus*, but not in *K. hermaphroditus*. Excluding the hybrid individuals (see results below), the number of alleles varied from two at locus R28 to 35 at locus R38, with an average of 17.6 alleles per locus considering all sampling locations combined. Average intrapopulation allelic richness (AR) based on 5 individuals was 4.21 for *K. ocellatus* and 1.38 for *K. hermaphroditus*. The mean expected heterozygosity (HE) was 0.56 (ranging from 0.47 to 0.60) for *K. ocellatus* and 0.07 (ranging from 0.03 to 0.10) for *K. hermaphroditus*. The *K. ocellatus* Northern clade populations (IRI, FUN and GUA) showed a higher average HE (0.59) than Parati (0.53) and Southern clades (0.53). Two samples (IRI and GUA) had significant heterozygote deficiency, whereas SFR had a significant heterozygote excess (Table 2.2). Examination of single-locus F_{IS} values indicated that significant values of mean F_{IS} in these populations were due to contribution of few loci (Table S2.5). Specifically, high F_{IS} values in IRI were caused by heterozygote deficiencies at loci R30 and R92, while high F_{IS} in GUA resulted from departures at locus R9, and significant heterozygote excess in SFR was caused by locus R34. Excluding these loci from analyses in respective populations rendered mean non-significant F_{IS} values. The great majority of single-locus F_{IS} were low and non-significant, and had approximately equal number of negative and positive values. Micro-Checker suggested that heterozygote deficiency at atypical loci mentioned above was probably a result of null alleles. Heterozygote excess at locus R34 in SFR may be a result of a locus duplication or an emergence of an additional priming site, that would make an appearance of additional 'spurious' alleles. Since mean F_{IS} were non-significant after the atypical loci were excluded, I consider that all studied populations of *K. ocellatus* are in HWE and did not exclude any locus from further analyses (Table S2.5).

2.4.3. Genetic differentiation at microsatellite loci

Genetic differentiation between populations in *K. ocellatus* was high and significant at many scales, either in global tests or in pairwise comparisons. For example, average F_{ST} among all comparisons was 0.25 ($P < 0.001$). In pairwise comparisons no difference was found between Guaratiba samples collected 10 years apart ($F_{ST} = 0.00$). F_{ST} in the remaining pairwise comparisons varied from 0.069 (between IRI and FUN; and SFR and FLO) to 0.386 (between IRI and PAR) (Table S2.6). The majority of pairwise F_{ST} were statistically significant after Bonferroni correction for multiple testing, with the exception of comparisons of FUN vs GUA-2017 ($F_{ST} = 0.072$), and FUN vs PAR ($F_{ST} = 0.382$), the latter most likely caused by small sample sizes in FUN and PAR (Table S2.6).

Two loci showed evident pattern of regional geographic differentiation. Sampling points from the Southern clade (PAR, SFR, and FLO) were fixed for allele $R18^{184}$, whereas Northern clade was nearly fixed for an alternative allele, $R18^{172}$ (allele frequencies $R18^{172}$ ranging from 95 to 98%). The Parati clade (geographically intermediate) had had both alleles, where 29 fish were homozygous at $R18^{172}$ but two fish were heterozygotes $R18^{172/184}$. Another locus with little overlap in allele distribution between Northern and Southern clades was R9. At this locus 91-100% of alleles in the Northern populations belonged to the class of 'short' alleles (alleles $R9^{94}$ – $R9^{226}$), whereas in the Southern populations 95-100% of alleles were in the class of 'long' alleles ($R9^{234}$ – $R9^{294}$). The two groups of populations shared alleles $R9^{230}$, $R9^{234}$ and $R9^{118}$, but combined frequencies of these alleles in any single populations were 9% at most.

2.4.4. Genetic partition and clustering analysis

AMOVA results using *cox1* haplotypes showed significant support for the division found among major mtDNA clades (78.5% of overall variation), followed by paleorivers configuration (62.74%), and major clades according to microsatellites genotypes (51.44%). Overall, variation within populations was low (<15%), while most of the variation (>51%) was found among groups (either major clades, paleorivers or geography), with most of the Φ_{CT} values being highly significant. Three-level AMOVA of microsatellite data partitioned according to the major mtDNA clades and microsatellite clusters indicated by Evanno's method (see below) had significant variation at all levels, with the latter explaining slightly more of the variation among groups than the former (19.26 vs 19.54%). Genetic partition according to the paleorivers configuration was not significant at highest level (Table 2.3).

Classification of individuals by the algorithm of STRUCTURE provided consistent results for each K across the 10 replicated runs. As expected in highly structured data, the most divergent groups separate into distinct clusters first. Evanno's ΔK method indicated the uppermost level of genetic structure, with K = 2. This analysis indicated one genetic cluster encompassing fish from the Southern clade (FLO, SFR, and PAR) and another composed by fish of fish from the northernmost sampling sites (IRI, FUN, GUA and PRT, mtDNA Northern and Parati clades) (Fig. 2.3). Inspection of outcomes at higher Ks added additional details on individual clustering. At K = 3, PRT population appeared distinct (corresponds to mtDNA Parati clade) (Fig. 2.3). Outcomes of K = 5 (indicated as the most likely number of genetic clusters by all metrics in STRUCTURESELECTOR) assigned most of the fish from GUA as a different genetic cluster. Interestingly, at this K, a subset of six individuals form FUN (FUN 08,11,13, 41,47,48) and five from GUA (GUA 09, 17, 20, 24, 62) was assigned

to its own genetic cluster. At K = 6, PAR was assigned to a separate genetic cluster (Fig S2.5). Runs at higher K did not change this pattern (Fig S2.3).

Geneland results incorporating mtDNA, microsatellites and spatial data generally agreed with those from STRUCTURE. Posterior distributions of the number of genetic clusters (K) showed a mode at K = 6 across all 10 replicated runs (Figs. 2.3; S2.2; S2.4). Spatially, cluster 1 was composed by individuals from IRI and a subset of individuals from FUN. Another subset of FUN individuals (composed of hybrids, see below) formed cluster 2. Individuals from GUA (including GUA hybrids), PRT and PAR composed each their own genetic cluster (clusters 3, 4 and 5 respectively). Cluster 6 was composed of the southernmost individuals from SFR and FLO (Fig. S2.4).

Factorial Correspondence Analysis (FCA) confirmed the uppermost subdivisions detected by Evanno's method (ΔK) from STRUCTURE (Fig 3). The plot along the two main axes showed that the major division was between the southern and the northern populations along axis 1, whereas six fish from FUN and five from GUA were separated from the rest along Axis 2. Another plot indicated that PRT and PAR fish were distinctive from other northern and southern populations along Axis 3 (Fig S2.5).

2.4.5. Hybridisation between *K. ocellatus* and *K. hermaphroditus*

Across all genetic clustering methods, strong evidence shows that FUN and GUA fish were subdivided into two groups: some had the same genetic background as other northern populations (IRI and GUA), while others showed an unique genetic composition, either being assigned to its own genetic cluster or not clustering with the other individuals from their respective population (Figs. 2.3; S2.3-S2.5). A closer

examination of individuals FUN 08, 11, 13, 41, 43, 47 and GUA 09, 17, 20, 24, 62) revealed that these fish had unique alleles at nine loci ($B86^{183}$, $R9^{194}$, $R11^{107}$, $R23^{300}$, $R28^{152}$, $R90^{242}$, $R92^{212}$, $R103^{127}$, and $R112^{191}$), which were not present in any other *K. ocellatus* populations across its known distribution. In all cases, these unique alleles were found in heterozygous state, with the second allele being commonly found in *K. ocellatus* from the Northern clade. FUN41 carried unique alleles only at five loci. Furthermore, allele $R28^{152}$ was found in a single fish from SFR population, so it is better characterized as 'nearly unique'. Several of the unique alleles found in the unusual FUN and GUA fish corresponded to alleles fixed in all populations of *K. hermaphroditus* (alleles $B86^{183}$, $R9^{194}$, $R11^{107}$, $R28^{152}$, $R92^{212}$, and $R103^{127}$). This strongly suggests that the divergent FUN and GUA fish are produced by hybridisation between *K.* ocellatus and *K. hermaphroditus*.

STRUCTURE analysis including *K. ocellatus* and *K. hermaphroditus* microsatellite data was consistent with the hybridisation hypothesis, with the eleven divergent fish appearing as admixture of two genetic pools (Fig 2.3). At K = 2 (most likely partition according to Evanno's ΔK) all *K. hermaphroditus* individuals were assigned with nearly 100% probability to one cluster, and almost all *K. ocellatus* were similarly assigned to another cluster, with exception of the subset of divergent FUN and GUA fish, which appeared to have genetic backgrounds of both species. At K = 7 (the best configuration according to metrics in STRUCTURESELECTOR) *K. hermaphroditus* individuals continued to be assigned to its own cluster, while *K. ocellatus* showed structure similar to that found in the separate analysis of this species (Figs. 2.3; S2.5), and the subset of divergent FUN and GUA fish continued to show strong signal of admixture between *K. ocellatus* and *K. hermaphroditus* (Fig 2.3).

According to STRUCTURE at K = 2, most of the divergent FUN fish (FUN 08, 11, 13, 47, 48) had about half of their genetic ancestry originated from *K. hermaphroditus* (average q-value for *K. hermaphroditus* cluster = 0.58, sd ± 0.01), with exception of FUN41, which had only 18% of *K. hermaphroditus* ancestry). In GUA, two individuals (GUA 09 and 62) also had approximately half of its ancestry coming from *K. hermaphroditus* (average q-value for *K. hermaphroditus* cluster = 0.51, sd ± 0.02), while the other three divergent individuals (GUA 17, 20, 24), showed approximately 35% of *K. hermaphroditus* ancestry (average q-value for *K. hermaphroditus* cluster = 0.35, sd ± 0.02). The NEWHYBRIDS classification using microsatellites data confirmed these patterns, with most of the individuals being classified as either "pure" *K. ocellatus* (179 individuals, average probability a posteriori = 0.99, sd ± 0.001) or "pure" *K. hermaphroditus* (67 individuals, average probability a posteriori = 0.99, sd ± 0.000001), while seven individuals (FUN 08, 11, 13, 47, 48; GUA 09 and 62) were classified as F1 (originating from outcrossing between *K. ocellatus* and *K. hermaphroditus*, average probability a posteriori = 0.97, sd ± 0.05), and four (FUN 41; GUA 17, 20, 24) being classified as backcrosses between *K. ocellatus* and a F1 hybrid (average probability a posteriori = 0.99, sd ± 0.003). No individuals were classified as F2 or as backcross between *K. hermaphroditus* and a F1 hybrid with the microsatellites (Fig. 2.3). Individuals classified as hybrids had substantially higher heterozygosity than the parental species, and showed strong signal of heterozygote excess (Tables 2.2; S2.7). With regards to the direction of the crosses, it appears that the female function was mostly performed by *K. ocellatus* hermaphrodites x males *K. hermaphroditus*, because all hybrids have a *cox1* haplotype of *K. ocellatus* (but see SNPs results below) (Fig 2.3).

Using genotype-by-sequencing, I produced two datasets using the same set of individuals using ANGSD-v.0.921 for specific analysis. Dataset 1 with 53 samples, 597,733 sites, coverage between 12.0X and 346.5X (mean 145.2X), missing data ranging from 0% to 7.2% (mean 0.50%); Dataset 2 with the same 53 samples, 5,477 SNPs, coverage between 12.4X and 382.6X (mean 152.9X), missing data ranging from 0% to 4.9% (mean 0.34%). To check whether the GBS library was able to produce SNPs evenly distributed across the reference *K. marmoratus* genome, I performed a regression between the size of each the 3,073 scaffolds and the number of sites found in each scaffold. I found a strong correlation ($R^2 = 0.93$, $p < 0.001$) indicating that the GBS protocol yielded genomic information uniformly-distributed throughout the reference genome.

Admixture results based on Dataset 2 SNPs (Fig 2.4) were consistent with the genetic structure and hybridisation evidence found with the microsatellite data, with the major genetic clusters corresponding to species, with some individuals from FUN and GUA showing strong evidence of admixture. At K = 2, four individuals from Fundão with *K. ocellatus cox1* haplotypes (FUN 08, 11, 43, 47) and two from Guaratiba (GUA 09, 17) showed a mixture of *K. ocellatus* and *K. hermaphroditus* genome. Two other individuals from Fundão (FUN 13, 41) indicated as hybrids by the microsatellites data failed to produce enough reads for the GBS library (cut-off ≥ 500k reads). Three other individuals identified as hybrids by microsatellites in Guaratiba (GUA 20, 24, 62) were not included in initial library (Table S2.1). One individual from Fundão (FUN 26), which displayed no evidence of admixture with the microsatellites data, showed substantial indication of genetic admixture with the SNPs. At K = 3, *K. ocellatus* from the Northern (FUN and GUA) and Southern (SFR and FLO) clades were split in separate clusters, similar to the uppermost genetic structure suggested by the

microsatellites data. The MDS analysis also supported these results (Fig 4), with individuals with shared ancestry of *K. ocellatus* and *K. hermaphroditus* positioned in between two opposite clusters composed by *K. ocellatus* and *K. hermaphroditus* individuals in the first dimension of variation (35.30% of the variation). The second dimension of variation (1.95% of overall variation) separates Northern and Southern clades of *K. ocellatus*. These patterns were reinforced by the F_{ST} comparisons (without hybrids), with the lowest values ($F_{ST} = 0.09$) found within clades of *K. ocellatus* (and between FUN and GUA for *K. hermaphroditus*, $F_{ST} = 0.06$), followed by the F_{ST} in comparisons of populations from different clades of *K. ocellatus* (average $F_{ST} = 0.45$), and finally F_{ST} between *K. ocellatus* and *K. hermaphroditus* (average $F_{ST} = 0.92$) (Table S2.8).

The NEWHYBRIDS analyses using a subset of 200 SNPs with high F_{ST} and low LD classified 11 individuals as "pure" *K. ocellatus* (probability a posteriori equals 1 in all individuals), 21 individuals as "pure" *K. hermaphroditus* (probability *a posteriori* equals 1 in all individuals), while five individuals (FUN 08, 11, 43, 47; GUA 09) were classified as F1 (probability *a posteriori* equals 1 in all individuals)and one individual (GUA 17) as a backcross between *K. ocellatus* and a F1 hybrid (probability *a posteriori* equals 1). All these classifications agree with the classifications found in the analyses of microsatellites data, with the exception of FUN 26, which was classified as "pure" *K. hermaphroditus* by microsatellites, but classified as a backcross between *K. hermaphroditus* and a F1 hybrid (probability *posteriori* equals 1) with the SNPs data. No individual was classified as F2. In terms of genetic diversity, individuals classified as hybrids (FUN 08, 11, 26, 43, 47; GUA 09, 17) had an average of proportion of heterozygous sites of 0.53, an approximately ten-fold increase from the values found in *K. ocellatus* (0.07) and *K. hermaphroditus* (0.03) (Fig 2.4).

2.5. Discussion

The genetic analysis using mtDNA, microsatellites and SNPs data revealed strong genetic structuring throughout the whole known distribution of the androdioecious mangrove killifish *K. ocellatus* in southeast Brazil. Correspondence of the observed and expected heterozygosity (and, respectively, low inbreeding coefficients), indicated absence of self-fertilisation in *K. ocellatus* across its distribution, as previously suggested by Tatarenkov et al. (2009). In addition, I found evidence of hybridisation and introgression (backcrosses) between the outcrossing androdioecious *K. ocellatus* and the predominantly selfing *K. hermaphroditus* mangrove killifish species in two out of three mangrove systems where the two species are known to co-occur, representing, to my knowledge, the first case of hybridisation between species with different mating systems in vertebrates.

2.5.1. Genetic structure in *Kryptolebias ocellatus*

Character mapping for the mangrove killifishes clade suggests that synchronous hermaphroditism had evolved in the common ancestor of all mangrove killifish species (*K. ocellatus*, *K. hermaphroditus* and *K. marmoratus*), while selfing had arisen later in the common ancestor of *K. hermaphroditus* and *K. marmoratus* (Avise and Tatarenkov 2015; Costa et al. 2010; Kanamori et al. 2016). Tatarenkov et al. (2009) genotyped two populations from Sepetiba Grande Bay (GUA) and Guanabara Bay (IRI) of *K. ocellatus* and based on low inbreeding coefficients, suggested that this species does not undergo selfing. The low F_{IS} values throughout the whole known distribution of the species found here reinforce the hypothesis that selfing is unlikely to occur in *K. ocellatus*, further confirming it is likely to have evolved later (in the common ancestor *K. hermaphroditus* and *K. marmoratus*) within the mangrove killifish clade.

Extensive genetic structure has been revealed at various geographic scales (Berbel-Filho et al. 2019a; Tatarenkov et al. 2015; Tatarenkov et al. 2007) and among selfing lines within the same mangrove system (Ellison et al. 2011; Ellison et al. 2012; Turko et al. 2018), in the selfing mangrove killifishes composing the *K. marmoratus* species complex. An exception to this general scenario of deep genetic structure seems to be the very low genetic differences found among *K. hermaphroditus* populations across the Brazilian coast (Tatarenkov et al. 2011). Tatarenkov et al. (2017) using mtDNA and microsatellite data found very little genetic differentiation in the populations *K. hermaphroditus* populations extending along 2500 km of the Brazilian coast, suggesting recent dispersal of the species into southeast Brazil, and challenging an earlier hypothesis that the selfing *Kryptolebias* species have originated in this region (based on the known distribution of the other *Kryptolebias* species) and have dispersed northwards (Costa et al. 2010; Murphy et al. 1999; Turner et al. 2006). I found here, that in much narrower geographic distribution (approximately 900 km along the coast from Magé/RJ to Florianópolis/SC), *K. ocellatus* showed a deeper genetic structure than *K. hermaphroditus*, being mostly divided in Northern and Southern genetic clades, but also showing genetic structure within these clades.

The mtDNA analysis revealed a clear genetic distinction between Northern (IRI, FUN and GUA) and Southern populations (PAR, SFR, FLO), with the first possibly subdivided in two minor clades (Northern and Parati), a pattern which was generally corroborated by uppermost genetic clustering detected by STRUCTURE, as well as in the FCA and MDS analyses with microsatellites and SNPs data, respectively. However, several within clades groups was also observed using microsatellites. This pattern of differentiation, with most of locales having their own genetic cluster agrees with the patterns of genetic structure found in *K. marmoratus* in the Caribbean (Avise

and Tatarenkov 2015), with exception of the absence of genetic differentiation in *K. hermaphroditus* in the same area in southeast Brazil previously reported by Tatarenkov et al. (2017) and confirmed here. On this species, populations from opposite sides of Sepetiba Bay, namely GUA and PIC (Fig 2.1), had only one *cox1* haplotype and were genetically homogeneous at microsatellites. In contrast, populations of *K. ocellatus* from opposite sides of Sepetiba Bay, GUA and PRT (only 80km distant from PIC through the coast line), belonged to different genetic lineages, with significant genetic differentiation between them. These findings, together with the extremely low levels of genetic diversity shown with mtDNA, microsatellites and SNPs in *K. hermaphroditus*, suggest that the morphologically similar and often syntopic mangrove killifishes species have different evolutionary history in southeast Brazil, with *K. hermaphroditus* being a recent coloniser of a mangrove area where *K. ocellatus* is likely to have settled much earlier and where it may have been originated.

Mangrove killifishes are amphibious (Turko and Wright 2015; Wright 2012), being able to spend prolonged periods of time out of water, moving overland by flipping across temporary water bodies within mangrove systems (Taylor 2012). As for most of the rivulids, however, long-distance dispersal seems to be limited by poor swimming capacity (Gibb et al. 2013) suggesting that other mechanisms (e. g. egg carriage by birds and storms, drifting on floating debris) may have helped mangrove killifishes to disperse across mangrove systems (Avise and Tatarenkov 2015). Recent studies have been found that the past connections brought by paleorivers had facilitated dispersal and colonisation of Neotropical coastal fishes into new river systems, strongly influencing their genetic structure (Lima et al. 2017; Thomaz et al. 2015; Thomaz et al. 2017). However, until now, this has been only tested in primarily freshwater fishes, on which dispersal capacity is intrinsically related to the availability

of freshwater environments. The AMOVA analysis partitioning the populations according to their potential paleorivers' connections (based on paleorivers in Thomaz et al. (2015)) showed a significant effect for *cox1*, but not with the microsatellites data. This incongruence is likely to be caused by the differences between mitochondrial and microsatellites markers, with the former mutating at a slower rate and being able to recover older historical connections, while the latter being likely to reveal more recent events of population dynamics given its faster mutation rate.

Mangrove killifishes are the only rivulid species living in brackish waters (Costa et al. 2010) and rarely share mangrove microhabitats with other fish species (Taylor 2012). Therefore, tracing parallels between the genetic structure of mangrove killifishes and other fish species is a challenging task. This scenario is even more challenging when comparing mangrove killifish genetic structure with that in other mangrove-dwelling organisms (e. g. crabs, shrimps), given their differences in life-cycle and dispersal capacities. Mangrove forests are composed of a polyphyletic group of plants which converged onto similar morphological and physiological adaptations to cope with tropical estuarine environments (Ball 1988). Mangrove plants release their propagules to be dispersed by the sea (Alleman and Hester 2011). For this reason, a multitude of studies using mangrove species from both Indo-West Pacific and Atlantic-East regions showed the direction of superficial ocean currents as the major factor influencing the genetic structure of mangrove plant species (Francisco et al. 2018; Yan et al. 2016). Studies using different mangrove tree species have shown weak genetic structure among estuaries in southeast Brazil, with a general north-south pattern of dispersal, guided by the Brazilian ocean current (Francisco et al. 2018; Mori et al. 2015; Pil et al. 2011). This panmixia scenario among different estuaries has also been observed in other mangrove-dwelling species in the same region which disperse through pelagic

larvae, such as crabs (Britto et al. 2018; Oliveira-Neto et al. 2007; Oliveira-Neto et al. 2008). While the strong genetic structure mostly among Northern and Southern estuaries found in *K. ocellatus* in southwestern Atlantic contrasts with the general patterns observed in these sea-dispersed mangrove species, it generally agrees with the one found for the cold-water philopatric penaeid shrimp *Farfantepenaeus paulensis* in southeast Brazil (Gusmão et al. 2005), which has two major genetic stocks, one in Rio de Janeiro and Santos (North), and other in Lagoa do Patos (South), with the latter representing post-larvae produced by individuals breeding in Santa Catarina (where SFR and FLO *K. ocellatus* are). Further research is needed to investigate which environmental and evolutionary factors have shaped the major genetic breaks in the absence of apparent geographical barrier between Northern and Southern populations of *K. ocellatus* and other coastal species in southeast Brazil.

2.5.2. Hybridisation and introgression between *K. ocellatus* and *K. hermaphroditus*

Hybridisation was once thought to be a rare and transient event in animals, usually happening while species are beginning to diverge (Mallet 2005; Taylor and Larson 2019). However, with the advance of new sequencing technologies, hybridisation reports have become increasingly common, being particularly high in groups such as plants (40%) (Mallet 2005; Whitney et al. 2010), insects and birds (10%) (Schwenk et al. 2008). The ubiquity of hybridisation reports revived the long-debated hypothesis that hybridisation is an adaptive mechanism, which can speed up adaptation by rapidly introducing locally-adapted alleles in large numbers at once (Lewontin and Birch 1966; Mitchell et al. 2019; Oziolor et al. 2019). Hybridisation has also been linked to increased invasiveness or success during range expansion, given enhanced fitness

often observed in hybrids when compared to their parental species (Hovick and Whitney 2014).

K. hermaphroditus was originally described in southeast Brazil (Guaratiba) as being exclusively composed of selfing hermaphrodites (Costa 2011a). The absence of males (and outcrossing) was substantiated by the extremely low levels of heterozygosity found in southern Brazil populations (Tatarenkov et al. 2011; Tatarenkov et al. 2009). However, recent sampling surveys have extensively expanded the known species distribution, as far as the mangroves surrounding the Amazon river estuary in northern Brazil (Guimarães-Costa et al. 2017; Lira et al. 2015), and the current genetic evidence suggests that the species has recently dispersed to southern Brazil from the north of South America or from the Caribbean (Tatarenkov et al. 2017). These recent surveys have also revealed the presence of rare males in some populations (Berbel-Filho et al. 2016; Costa 2016). Berbel-Filho et al. (2019a) have found low observed heterozygosity levels and high selfing rates in a population (Ceará-Mirim) where one male was reported (Berbel-Filho et al. 2016), suggesting that outcrossing (likely to happen between males and hermaphrodites, see Furness et al. (2015)) occurs in *K. hermaphroditus*, but at very low frequencies. Recently, Tatarenkov et al. (2018) reported unexpected hybridisation between two substantially divergent (3% K2P distance at *cox1*) *Kryptolebias* lineages in San Salvador, Bahamas, with hermaphrodites from the 'Central clade' (clade in the Caribbean closely related to *K. hermaphroditus*, see Tatarenkov et al. (2017) for details) hybridising with males of *K. marmoratus*. Here, I found evidence for hybridisation between species which are even more genetically distinct (11% K2P distance at *cox1*), namely *K. ocellatus* (androdioecious and outcrossing) and *K. hermaphroditus* (androdioecious and predominantly selfing) in two populations (FUN and GUA) in southeast Brazil. Eleven

out of twelve hybrids had a *K. ocellatus* mtDNA haplotype, while only one had a mtDNA haplotype of *K. hermaphroditus.* Assuming matrilineal inheritance of the mitochondrial DNA (as for most eukaryotes (Birky 1995)), these findings suggest that most of hybridisation events have happened between hermaphrodites of *K. ocellatus* and the very rare males of *K. hermaphroditus*. In addition, five of those individuals were classified as backcrosses between a F1 hybrid and a parental species, suggesting the hybrids can produce viable offspring. The absence of F2 individuals seems to suggest that the hybrids between *K. ocellatus* and *K. hermaphroditus* mostly reproduce via outcrossing with a parental species, however further research is needed to investigate the possibility that those hybrids also mate to each other or undergo self-fertilisation.

As for other mixed-mating species, in mangrove killifishes species (*K. marmoratus* and *K. hermaphroditus*) the increased genetic variation generated via male-mediated outcrossing events seems to allow the species to cope with rapidly changing selective pressures, such as parasites, while selfing assures reproduction and enhances colonisation of new microhabitats (Berbel-Filho et al. 2019a; Ellison et al. 2011; Mackiewicz et al. 2006c). Reinforcing this hypothesis is the fact that *K. marmoratus* males tend to associate with the most genetically dissimilar hermaphrodites (Ellison et al. 2013). Hybrids were found only in two of the sampling sites, in one of which (Guaratiba) introgression between the species had not been previously detected (Tatarenkov et al. 2011; Tatarenkov et al. 2009), suggesting that hybridisation events could be are relatively recent. The Fundão population, containing both *K. ocellatus* and *K. hermaphroditus* was only recently discovered, being one out of the three populations where males of *K. hermaphroditus* were described (Costa et al. 2016). However, no *K. hermaphroditus* males were ever reported in Guaratiba, despite

extensive collections efforts in the area (Costa 2011a; Costa 2016). Although the possibility of cryptic males cannot be discarded (Marson et al. 2019), the genetic data seems to indicate that outcrossing between males and hermaphrodites of *K. hermaphroditus* in Guaratiba is very rare (Tatarenkov et al. 2011; Tatarenkov et al. 2009).

Given its cryptic and complex environment, virtually no information is available about the mating behaviour of mangrove killifish species in the wild. However, observations on lab-reared *K. marmoratus* demonstrated that males and hermaphrodites display some of the typical killifish courtship/spawning behaviours (males closely positioning themselves sideways to the female/hermaphrodites with intense vibration) (Taylor 2012). Lomax et al. (2017), however, showed that *K. marmoratus* hermaphrodites rarely laid unfertilised eggs, regardless of whether they were accompanied by a male or another hermaphrodite. The asymmetrical direction of hybridisation found here may have been caused by the differences in the modes of reproduction of *K. ocellatus* and *K. hermaphroditus* hermaphrodites, as the former does not undergo selfing, suggesting that the hermaphrodites only lay unfertilised eggs, which may facilitate fertilisation by *K. hermaphroditus* males. The latter species, however, is predominantly selfing, suggesting that eggs laid by hermaphrodites are in its majority already fertilised, decreasing the likelihood of hybridisation between *K. ocellatus* males and *K. hermaphroditus* hermaphrodites. However, I cannot rule out that other mechanisms (e. g. spatial and temporal environmental fluctuations, different mating behaviours, differential survival of hybrids dependent on the direction of the crosses) may be driving the asymmetrical patterns of hybridisation found here. Further research is needed to investigate the behavioural and ecological interactions between *K. ocellatus* and *K. hermaphroditus* in syntopy.

Another question that remains unanswered is why the hybridisation seems to have happened in some populations (Fundão and Guaratiba), but not in others (Iriri). Fundão and Iriri (the latter in a less urbanised area) mangroves are situated in the Guanabara Bay, while Guaratiba is situated in the Sepetiba Bay. Both bays are known to be heavily impacted by human activities, such as intense deforestation of the mangrove forests (Farias et al. 2007) and in particular, pollution derived from industrial effluents, domestic sewage and agricultural run-offs (Carreira et al. 2004), resulting in high concentrations of heavy metals (Neto et al. 2006) and aromatic hydrocarbons (Figueiredo et al. 2008; Silva et al. 2007). Recently, Oziolor et al. (2019) found evidence that recent introgression between Gulf and Atlantic killifish (*Fundulus grandis* and *F. heteroclitus*, respectively) have conferred populations of the former species resistance to high levels of polycyclic aromatic hydrocarbons, suggesting that hybridisation may be a rapid mechanism to generate adaptive genetic variability to extreme environmental conditions. Further research should investigate if a similar scenario may be happening in the mangrove killifishes, as well as which are the environmental/evolutionary causes and consequences of the hybridisation between *K. ocellatus* and *K. hermaphroditus* in the polluted mangroves of southeast Brazil.

Table 2.1. *Kryptolebias ocellatus* and *Kryptolebias hermaphroditus* identified according *cox1* haplotypes across sampling locations and their respective sampling sizes. PR = Paraná State; RJ = Rio de Janeiro State; SC = Santa Catarina State; SP = São Paulo State. Asterisks represents individuals which were sequenced for *cox1* and/or genotyped in a different study (shown in Reference).

Sample ID	Location	Latitude	Longitude	*cox1*	Msats	Reference
(a) *Kryptolebias ocellatus*						
1. IRI	Iriri mangrove, Magé, RJ	22°39'48.80"S	43°05'12.20"W	22*	51*	Tatarenkov et al. (2009) *
2. FUN	Fundão mangrove, Rio de Janeiro, RJ	22°52'2.50"S	43°13'27.50"W	11	11	This study
3. GUA	Piracão mangrove, Guaratiba, RJ	23° 0'1.90"S	43°34'51.50"W	17	19/24*	This study/ Tatarenkov et al. (2009) *
4. PRT	Parati mangrove, Parati, RJ	23°11'58.80"S	44°43'26.20"W	39	31	This study
5. PAR	Paranaguá Bay, Paranaguá, PR	25°31'7.60"S	48°39'30.90"W	5	5	This study

6. SFR	Linguado channel, São Francisco do Sul, SC	26°22'0.02"S	48°39'58.40"W	19	19	This study
7. FLO	Rio Ratones estuary, Florianópolis, SC	27°28'3.84"S	48°29'33.76"W	16	30	This study
Total				**129**	**190**	
(b) *Kryptolebias hermaphroditus*						
2. FUN	Fundão mangrove, Rio de Janeiro, RJ	22°52'2.50"S	43°13'27.50"W	29	16	This study
3. GUA	Piracão mangrove, Guaratiba, RJ	23° 0'1.90"S	43°34'51.50"W	30	16/10*	This study/ Tatarenkov et al. (2011) *
8. PIC	Fazenda River, Picinguaba, SP	23°22′01.0″S	044°50′13.4″W	2	25*	Tatarenkov et al. (2011) *
Total				**61**	**67**	

Table 2.2. Descriptive statistics of genetic variation at microsatellite loci in (*a*) 179 *Kryptolebias ocellatus* (excluding hybrids) and (*b*) 11 hybrids and (*c*) 67 *K. hermaphroditus* 67 *K. hermaphroditus*. GUA samples from different years are separated. N = sample size; L = number of loci; P_{99} = proportion of polymorphic loci (99% criterion); A = average number of alleles; A_R = allelic richness based on 5 individuals; H_E = expected heterozygosity; H_O = observed heterozygosity; F_{IS} = coefficient of inbreeding. Asterisks represent significant departures from HWE.

Sampling location	***N***	***L***	P_{99}	***A***	A_R	H_E	H_O	F_{IS}
(*a*) *Kryptolebias ocellatus*								
1. IRI	51	16	0.93	11.19	4.49	0.57	0.54	0.05*
2. FUN	5	16	0.81	4.56	4.56	0.60	0.57	0.04
3. GUA 2017	14	16	0.93	8.00	4.65	0.60	0.56	0.06*
3. GUA 2007	24	16	0.93	8.19	4.42	0.60	0.53	0.11*
4. PRT	31	16	0.87	7.88	3.88	0.53	0.51	0.04
5. PAR	5	16	0.68	3.31	3.31	0.48	0.48	-0.02
6. SFR	19	16	0.87	7.25	4.17	0.58	0.63	-0.10*
7. FLO	30	16	0.81	9.87	4.27	0.54	0.52	0.03
Mean	**22.4**		**0.86**	**7.53**	**4.21**	**0.56**	**0.54**	**0.04**
(*b*) Hybrids								
2. FUN	6	16	0.88	4.25	3.89	0.62	0.81	-0.34*
3. GUA	5	16	1.00	4.69	4.69	0.69	0.86	-0.30*
Mean	**5.5**		**0.94**	**4.47**	**4.29**	**0.66**	**0.84**	**-0.32***
(*c*) *Kryptolebias hermaphroditus*								
2. FUN	16	15	0.13	1.33	1.16	0.03	0.00	0.73*

3. GUA 2017	16	14	0.13	1.73	1.48	0.09	0.01	0.87*
3. GUA 2007	10	14	0.21	1.57	1.48	0.10	0.00	0.93*
8. PIC	25	14	0.357	1.79	1.42	0.08	0.01	1.00*
Mean	**16.8**		**0.21**	**1.61**	**1.38**	**0.079**	**0.007**	**0.92***

Table 2.3. AMOVA results from different hypotheses for mtDNA *cox1* gene and microsatellites in *Kryptolebias ocellatus*. Hybrids were excluded from the microsatellite analysis.

Source of variation			***cox1* mDNA**					**Microsatellites**		
	df	**Variance**	**% of variation**	**P-value**	**F-statistics**	**df**	**Variance**	**% of variation**	**P-value**	**F-statistics**
Major clades according to mtDNA (3)										
Among groups	2	2.59	78.50	**<0.001**	0.78	2	1.21	19.26	**0.01**	0.19
Among populations within groups	11	0.26	7.88	**<0.001**	0.36	4	0.51	8.21	**<0.001**	0.10
Among individuals within populations	113	0.44	13.62	**<0.001**	0.86	172	0.24	3.93	**<0.001**	0.04
Genes within individulas						179	4.32	68.61	**<0.001**	0.31

Major clades according to Evanno method (2)										
Among groups	1	1.93	51.44	**<0.001**	0.75	1	1.28	19.54	**0.02**	0.19
Among populations within groups	5	1.37	36.49	**<0.001**	0.87	5	0.73	11.35	**<0.001**	0.14
Among individuals within populations	120	0.4	12.07	**0.02**	0.51	172	0.19	3.74	**<0.001**	0.05
Genes within individulas						179	4.45	65.37	**<0.001**	0.34
Paleorives										
Among groups	4	1.89	62.74	**<0.001**	0.62	4	0.61	10.68	0.15	0.10
Among populations within groups	9	0.67	22.34	**<0.001**	0.59	2	0.82	14.04	**<0.001**	0.15

Within populations	113	0.44	14.92	**<0.001**	0.85	172	0.19	4.08	**<0.001**	0.05
Genes within individulas						179	4.45	71.20	**<0.001**	0.28

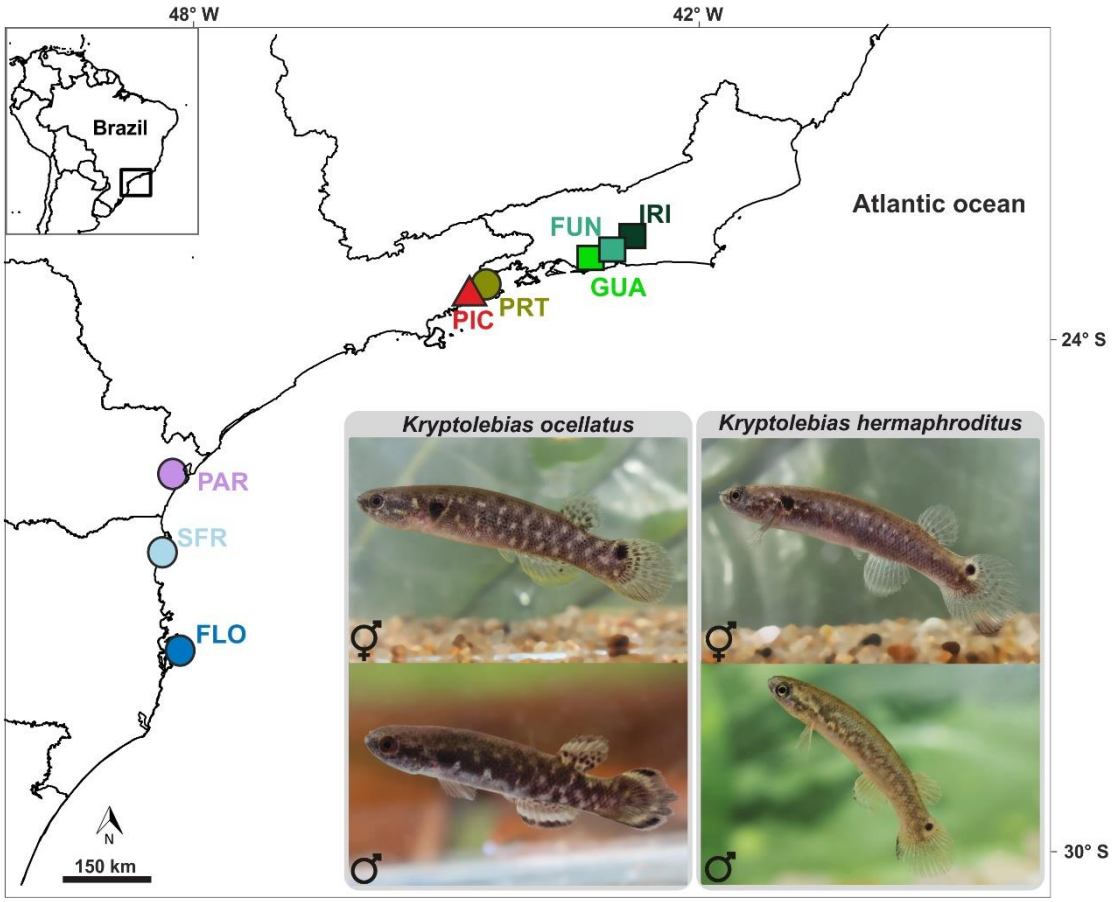


Figure 2.1. Sampling locations for *Kryptolebias ocellatus*. Squares represent locations where *K. ocellatus* and *K. hermaphroditus* are syntopic, circles are for locations where only *K. ocellatus* is found, while triangle designates site where only *K. hermaphroditus* is found. Labels for locations are described on Table 2.1.

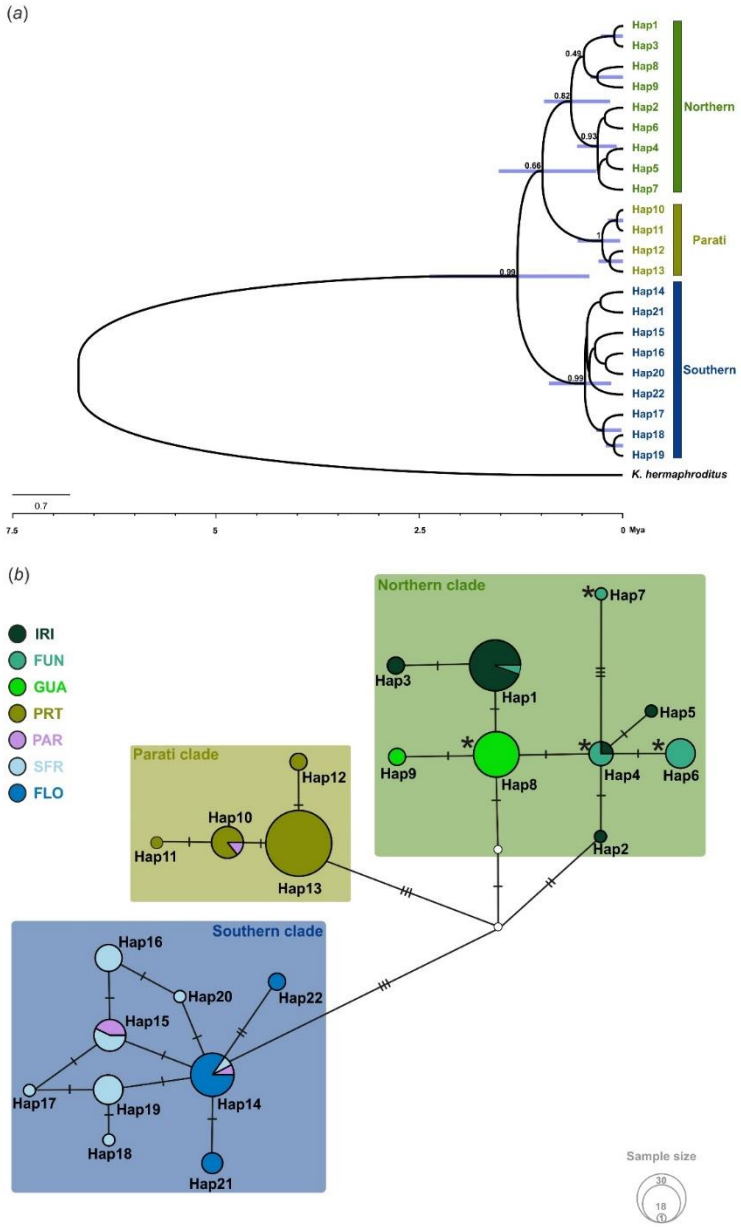


Figure 2.2. (*a*) Bayesian time-calibrated phylogenetic reconstruction for the 22 cox1 haplotypes from 129 specimens of *Kryptolebias ocellatus* and one *K. hermaphroditus* specimen used as an outgroup. (*b*) *cox1* haplotype network from 129 specimens of *K. ocellatus*. Each circle represents a haplotype and its size is proportional to the frequency of the haplotype. Ticks on branches connecting the haplotypes indicate nucleotide mutations. Asterisks represent haplotypes found in hybrid individuals (see results). Circles are coloured according to the different locations.

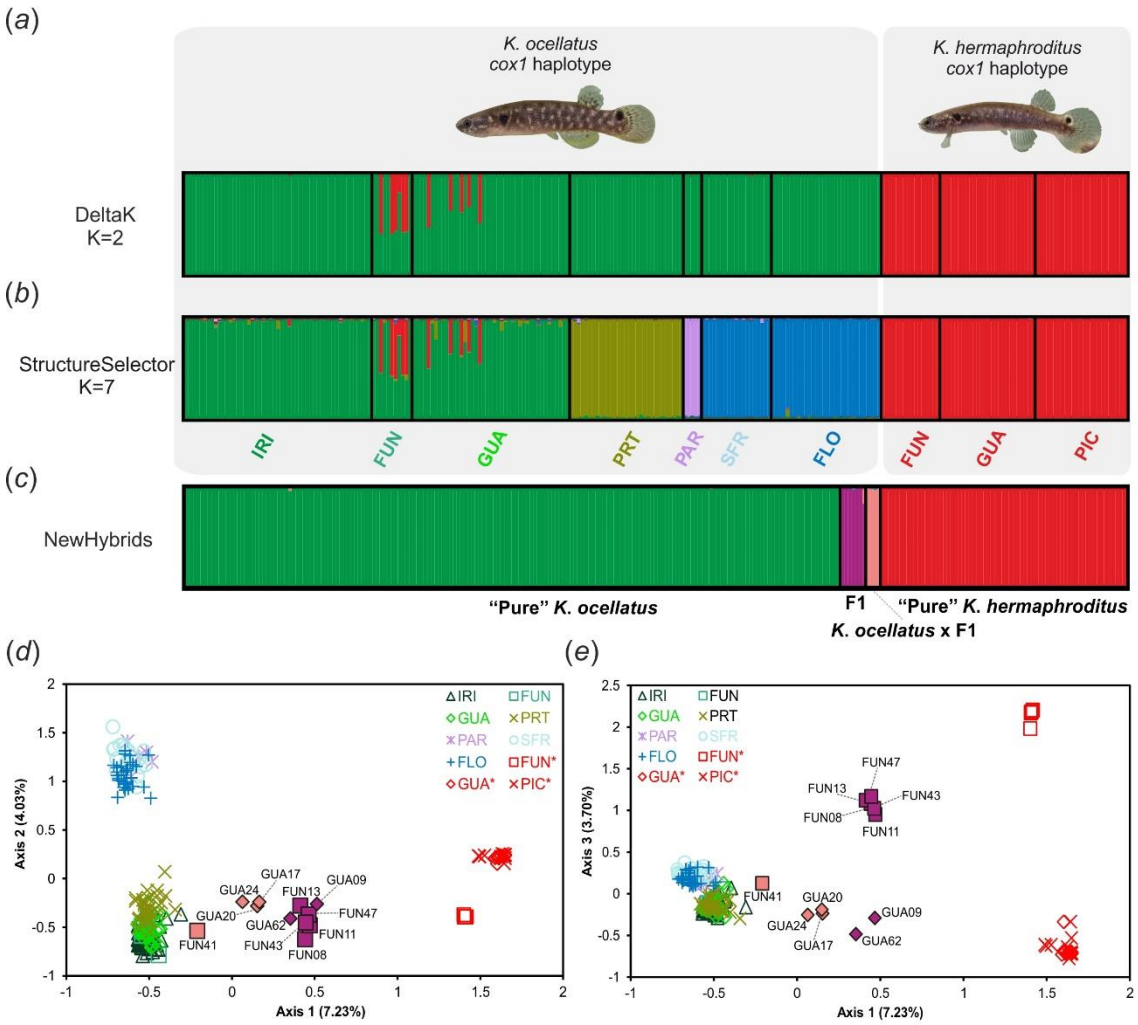


Figure 2.3. (*a*) Admixture plots showing the most likely genetic clusters (K) value for the microsatellites amplified in *Kryptolebias ocellatus* and *K. hermaphroditus* ran in Structure and determined by deltaK method in Evanno et al. (2005) and (*b*) STRUCTURESELECTOR results using the metrics defined in Puechmaille (2016) to account for unevenness on sampling sizes and hierarchical structure (*c*) NEWHYBRIDS individual classification using microsatellite data. (*d-e*) Factorial correspondence analysis for all *K. ocellatus* and *K. hermaphroditus* individuals coloured and shaped according to its sampling sites. Hybrid individuals (see results) are highlighted with their respective labels and coloured according to the hybrid class indicated by NEWHYBRIDS analysis. Asterisks represent sampling locations for *K. hermaphroditus*.

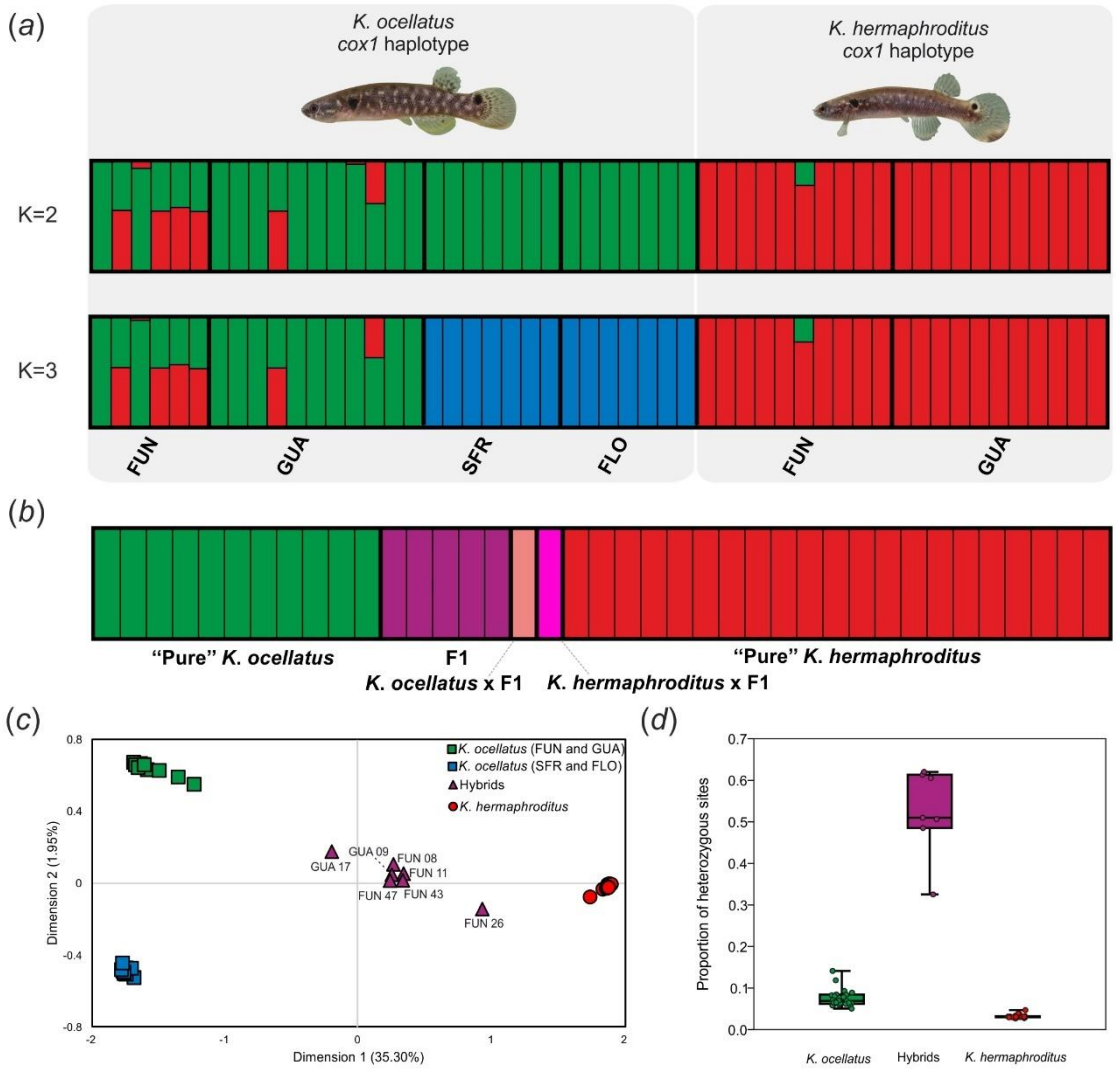
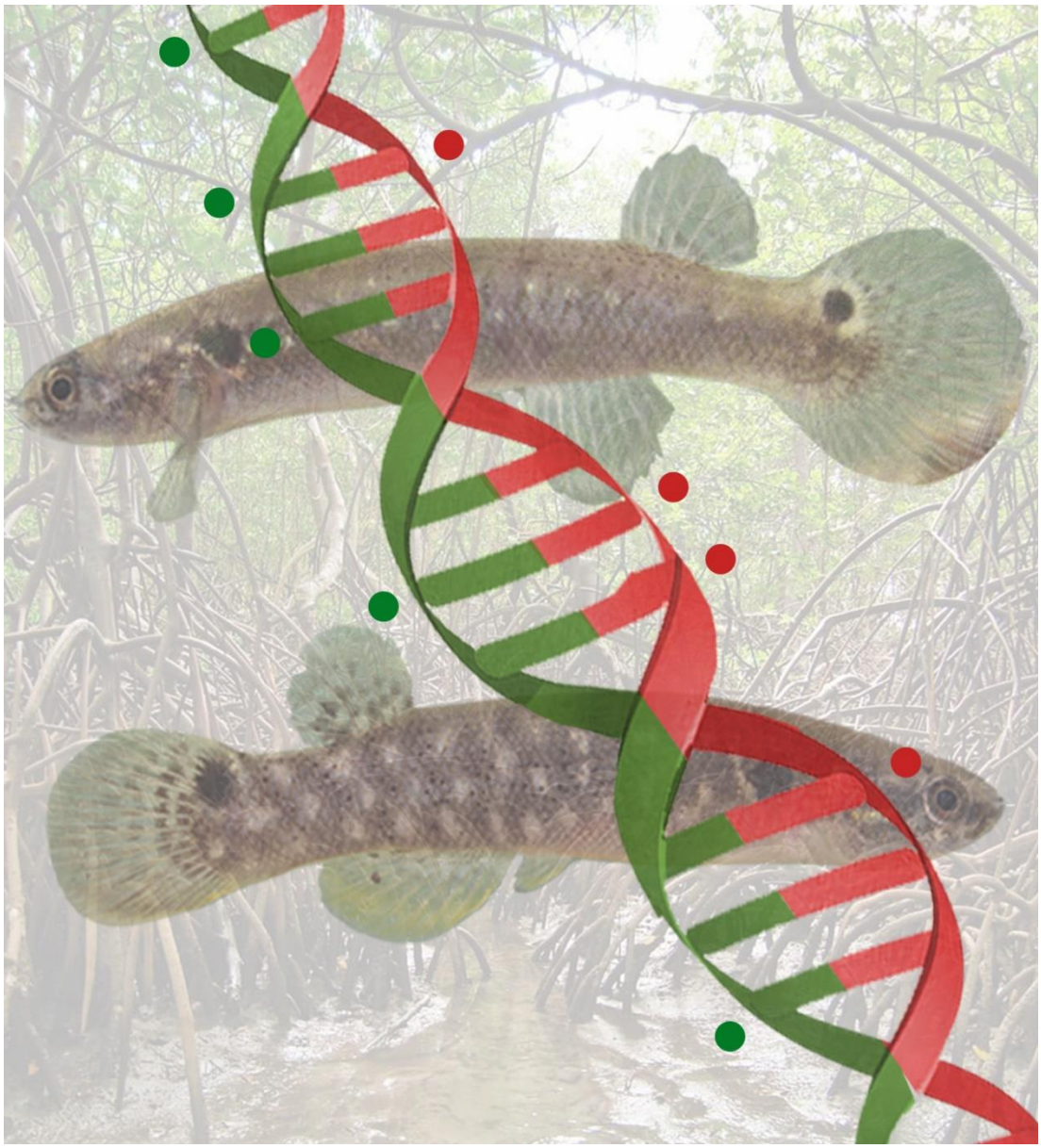


Figure 2.4. (*a*) Admixture plots showing the genetic clustering value for K= 2 and 3 for the 5,477 SNPs extracted from 53 individuals sequenced for the genotype-by-sequencing library and generated using ngsAdmix v. 3.2. Each individual is represented by a bar, and each colour represents a genetic cluster. (*b*) NEWHYBRIDS individual classification using a subset of 200 SNPs with low-levels of linkage-disequilibrium. (*c*) First two dimensions of the multidimensional Scaling (MDS) based on the genetic distances from 5,477 SNPs. Hybrid individuals (see results) are highlighted with their respective labels. (*d*) Proportion of heterozygous sites between *K. ocellatus*, *K. hermaphroditus* and the hybrid individuals.

CHAPTER 3: Intermediate patterns of epigenetic variation in hybrids of divergent mangrove killifish species in natural populations



3.1. Abstract

Hybridisation is one major source of evolutionary innovation. It's alternative outcomes (i.e. hybrid vigour, hybrid incompatibility) may depend not only of the of interactions between parental genetic alleles but also how the regulatory mechanisms from the two parental genomes are inherited and interact to each other. Besides the increasing reports of hybridisation in a wide range of taxa, its effects on cytosine methylation patterns are largely unexplored, particularly in animals. Here, I investigated the cytosine methylation patterns of wild hybrids between *Kryptolebias ocellatus* and *K. hermaphroditus* in southeast Brazil. My results revealed that the parental species, even living sintopically in the same mangrove microhabitats, diverge markedly on their cytosine methylation patterns, and their hybrids showed a predominant intermediate pattern of cytosine methylation patterns. In addition, the backcross individuals, showed cytosine methylation patterns which are more similar to the parental species than to the F1 hybrids. 217 gene ontologies, some of them, crucial to development, were shown to be overrepresented in hybrids in comparison to the parental species. The strong intermediate effects observed in the methylation patterns between of *K. ocellatus* and *K. hermaphroditus* hybrids suggest an important effect of genetic background on cytosine methylation inheritance in hybrids, which can contribute to hybrid divergence, and ultimately speciation, even under similar environmental conditions.

3.2. Introduction

Hybridisation is one of the major sources of evolutionary novelty, having important implications for phenotypic diversification, adaptation and speciation (Mitchell et al. 2019). Hybridisation can have alternative outcomes, in some cases resulting in heterosis, when hybrids are fitter than parental species (Hovick and Whitney 2014), or hybrid incapability, when hybrids have reduced fitness compared to the parental species, usually caused by the combination of incompatible alleles in hybrids, leading to increased postzygotic reproductive isolation in parental species (Blevins et al. 2017; Michalak 2009). New sequencing techniques suggest that hybridisation may be more common than previously thought (Taylor and Larson 2019), but the molecular mechanisms underlying its varied outcomes remain poorly understood (Bell et al. 2013; Lauss et al. 2018).

The merging of two divergent genomes brought together by hybridisation often results in extensive changes in regulatory elements (Riddle and Birchler 2003), transposable elements activity (Ungerer et al. 2006), chromosomal rearrangements (Metcalfe et al. 2007), DNA methylation (Blevins et al. 2017; Lauss et al. 2018; Salmon et al. 2005) and gene expression patterns (McGirr and Martin 2019; Michalak and Noor 2003; Moehring et al. 2007). In particular, divergent genetic alleles in parental species can be incompatible in hybrids (Dobzhansky-Muller incompatibilities), causing reduced fitness (Johnson 2000). This can be caused by misregulated gene expression patterns (McGirr and Martin 2019), which are consequence of the complex interactions between parental genomes and their regulatory factors (e. g. epigenetic mechanisms) (Michalak 2009). Therefore, the adaptive success of hybrid individuals may depend on the interactions resulting from merging two divergent genomes together with their fine-scaled tuned levels of gene expression regulated by epigenetic mechanisms.

Hybridisation can lead to the generation of all modes of gene action (e. g. additivity and non-additivity) (Swanson-Wagner et al. 2006). Previous research on the effects of hybridisation on gene expression inheritance has thus far shown mixed results. For example, hybrids between farmed and wild Atlantic Salmon, and recently-diverged pupfish species have shown mostly non-additive patterns (e. g. either over or under-dominance) on the gene expression levels relatively to the parental species (McGirr and Martin 2019; Roberge et al. 2008), while predominately additive effects were found in hybrids of house mouse subspecies (Rottscheidt and Harr 2007) and *Drosophila* species (Hughes et al. 2006). Far less studied however, is the effects of hybridisation on epigenetic mechanisms, particularly in animals. Studies in plants have shown both additive (Radosavljević et al. 2019) and non-additive (Hegarty et al. 2011) effects of hybridisation on DNA methylation patterns in plants. As the effects of DNA methylation in gene expression are dependent of its location in the genome (Anastasiadi et al. 2018; Jones 2012), the use of anonymous markers on most of those studies made it difficult to investigate whether DNA methylation patterns in hybrids differed in a context-specific manner (e. g. promoters, gene bodies) relatively to its parental genomes.

Understanding epigenomic changes brought by hybridisation may be crucial to understand the factors regulating hybrids gene misexpression, as well as its viability and fitness. Using wild-caught individuals of *Kryptolebias ocellatus* and *K. hermaphroditus* sampled simultaneously with their F1 hybrids and backcrosses, I aimed to investigate the patterns of DNA methylation in the hybrids compared to the parental species. Given the indication of genotype-by-environment effects observed in DNA methylation (Berbel-Filho et al. 2019b; Dubin et al. 2015; Richards 2006), I expected contrasting DNA methylation patterns in *K. ocellatus* and *K.*

hermaphroditus, even in the same mangrove microhabitats. I hypothesise that hybrids will inherit epigenetic patterns from both parental species, predominantly showing intermediate patterns of DNA methylation relatively to their parental species across all genomic contexts.

3.3. Material and methods

3.3.1. Study populations

To avoid tissue and/or sex-specific effect on cytosine methylation, only pectoral-fin samples from hermaphrodite individuals were included in the library, which were initially assigned to species by the pattern of scales in the flanks (Costa 2006). Morphological species identification was later confirmed by barcoding through the amplification of the mitochondrial gene cytochrome oxidase subunit I (*cox1*).

To investigate the effect of hybridisation in DNA methylation patterns, I selected *K. ocellatus* and *K. hermaphroditus* individuals from two mangrove sites (GUA and FUN in Chapter 2) from the library amplified in Chapter 2, resulting in 39 individuals (11 *K. ocellatus*, 21 *K. hermaphroditus* and seven hybrids; Table S3.1). Individuals were sampled using the same methods described in Chapter 2. All individuals included were sampled at the same time (GUA mangrove in 30/08/2017; FUN mangrove in 31/08/201) minimising the effect temporal environmental variation on DNA methylation. Hybrids were classified into hybrid classes (F1, backcross 1: between *K. ocellatus* and F1, backcross 2: between *K. hermaphroditus* and F1) using NEWHYBRIDS v.1.1 (Anderson and Thompson 2002) based on SNP data (F1: FUN 08, 11, 43 and 47; GUA 09; backcross 1: GUA 17; backcross 2: FUN 26). The details for library preparation and bioinformatic prepossessing of reads (e. g. demultiplexing, filtering, alignment, sorting and indexing) are fully described in Chapter 2.

3.3.2. Differentially methylated cytosines and hybrid cytosine methylation patterns

Differentially methylated cytosines (DMC) were identified using a modification of the methylation-sensitive genotype-by-sequencing (ms-GBS) protocol (Kitimu et al. 2015) with methylation-sensitive restriction enzyme HpaII, which activity is blocked by the presence of methylation on any cytosine on its recognition site (i.e., CCGG). DMCs were identified using the R package msgbsR (Mayne et al. 2018). With this package, individual restriction-digested reads aligned to the reference genome are filtered out for correct cut sites and possible outliers (e. g. low number of cut sites). The function *diffMeth* uses a bridge with the gene expression R package edgeR (Robinson et al. 2010) to split data according to comparisons, normalise read counts according to library size and identify DMCs. I performed three comparisons: (1) *K. ocellatus* vs *K. hermaphroditus;* (2) Hybrids vs *K. ocellatus;* and (3) Hybrids vs *K. hermaphroditus*. Only loci with more than 1 count per million (CPM) reads in at least "n" individuals in each compared group, with "n" being determined by the group with the lowest number of samples in each comparison (11 in *K. ocellatus* vs *K. hermaphroditus;* seven in the comparisons including hybrids). DMCs were then filtered by correcting for multiple testing using Benjamini-Hochberg method with a false discovery rate (FDR) of 0.01. The logFC (logarithm 2 of fold-change in counts per million reads) values of those DMCs was retrieved to evaluate the intensity and direction of methylation changes. This approach uses fold change values as a relative inverse proxy for methylation state, with higher methylation of a specific DMC site reducing the number of reads for that locus, consequently reducing its CPM reads number (Konate et al. 2018). For example, a specific DMC is considered

hypomethylated in group two compared to group one if its logFC value is above 1 in the comparison between group one vs group two.

To identify the methylation patterns of hybrids relative to their parental species, I generated a list of common DMCs (FDR <0.01) present in the comparisons between hybrids vs *K. ocellatus* and hybrids vs *K. hermaphroditus.* The normalised counts of these DMCs across all individuals was used for the downstream analysis. To visualise overall variation in DMCs, I performed a multidimensional scale analysis (MDS) using Euclidean distance across all individuals. To compare DMCs profile across experimental groups using hierarchical clustering, normalised counts per DMC and individual were scaled as:

$$\%(i,j) = \frac{Nreads\ (i,j)}{Nmaxreads(j)} \text{ x } 100;$$

where *i* = individual; *j* = each DMC; *Nreads* = normalised counts for *i* in *j*; *Nmaxreads* = maximum number of normalised counts for *j*.

To classify the patterns of DNA methylation inheritance in the hybrids compared to the parental species, I quantified the differences in normalised counts for each site. Inheritance was considered potentially additive if normalised counts of DMCs with significant differences between parental species were intermediate in the hybrids. Inheritance was considered potentially either r dominant in hybrids if normalised counts were higher or lower in hybrids respectively, compared to the counts on the parental species (e. g. using DMCs common to the comparisons between hybrids vs parental species) (Fig. S3.1).

3.3.3. Genomic context and gene ontology enrichment analysis

Using the annotated *K. marmoratus* reference genome (Rhee et al. (2017), I identified the genomic context (within gene body, promoter region (≤2 kb 5'-3' upstream of the transcription start site (TSS)), or intergenic region (≥2kb upstream of TSS and/or downstream the gene bodies), for the 5,800 DMCs common to the two comparisons between hybrids and parental species. To identify any potential differences on DNA methylation patterns across hybrids and parental species across different genomic contexts, I rerun the MDS using DMCs present in gene promoters, gene bodies and intergenic regions, respectively.

The annotated regions affected by these DMCs were used for the gene ontology enrichment analysis using zebrafish (*Danio rerio*) gene orthologs in PANTHER v. 11 (Mi et al. 2016). I searched for enrichments across biological process ontologies curated for zebrafish. Only genes which matched with the genes names annotated for zebrafish were included in the gene ontology analysis.

3.4. Results

3.4.1. Methylation-sensitive genotype-by-sequencing library (ms-GBS)

The ms-GBS library only including individuals in the current analysis yielded in average 6,422,972 reads per individual, with 85.21% uniquely mapping reads to *K. marmoratus* reference genome (ASM164957v1) (Table S3.1). Uniquely mapping reads generated sequencing data for a total of 830,905 loci. The proportion of reads mapping to the reference genome did differ between hybrids, *K. ocellatus* and *K. hermaphroditus* (ANOVA, $p = 0.001$) with higher proportion of reads uniquely mapping in *K. hermaphroditus* (mean = 0.80; SD ± 0.08) when compared to hybrids

(mean = 0.71; SD ± 0.06) and *K. ocellatus* (mean = 0.69; SD ± 0.03) (Fig. S3.2; Table S2.1), reflecting phylogenetic relatedness.

3.4.2. Differentially-methylated cytosines among groups

In total, 56,705 DMCs were found in the comparisons between *K. ocellatus* and *K. hermaphroditus*, of which 37,664 (66.48%) had a false discovery rate < 0.01. Smaller numbers of significantly different DMCs were found in the comparison between hybrids and parental species (hybrids vs *K. ocellatus*: 50,143 DMCs of which 10,620 (21.17%) had FDR < 0.01; hybrids vs *K. hermaphroditus*: 56,390 DMCs of which 13,905 (24.65%) had FDR < 0.01) (Fig. S3.3). Relatively to the hybrids, *K. hermaphroditus* showed higher number of DMCs than *K. ocellatus*. However, the overall methylation patterns were more similar between hybrids and *K. hermaphroditus* than between hybrids and *K. ocellatus*, with 50.45% of the DMCs hypermethylated in relation to the hybrids (and vice-versa in terms of hypomethylated DMCs), vs. 64.23% in *K. ocellatus* (Fig. 3.1a). The MDS analysis using DMCs between parental species positioned the hybrids in between two opposite clusters representing the parental species. These results were also supported by the MDS using all reads normalised by library size (830,905 sites) (Fig. S3.4). In both cases, individuals identified as backcrosses occupied eigen spaces closer to the non-hybrid parental species than F1 hybrids.

5,800 DMCs were commonly differentially methylated in the comparisons between hybrids vs parental species, revealing a larger number of exclusively DMCs differences between hybrids and *K. hermaphroditus* than between hybrids and *K. ocellatus* (8,105 vs 4,820) (Fig. 3.1b). When comparing the common DMCs present in both comparisons among hybrids and parental species, *K. ocellatus* had 34.28% (vs.

65.19% in *K. hermaphroditus*) and 65.72% (vs. 34.81% in *K. hermaphroditus*) of hypermethylated and hypomethylated DMCs, respectively, relatively to the hybrids (Figs. 3.1c and S3.5). Again, the MDS analysis showed three distinct clusters, one composed by *K. ocellatus* individuals, another composed by *K. hermaphroditus* individuals, while an intermediate cluster containing the hybrids (both F1 and backcross individuals). As shown above, backcross individuals, GUA 17 and FUN 26 occupied eigen spaces closer to the backcross non-hybrid parental species than F1 hybrids. The contrasting methylation differences between parental species was reinforced and hierarchical clustering analysis, with the hybrids clearly showing intermediate levels of DNA methylation relatively to its parental species, with exception of the backcross individuals, GUA 17 and FUN 26, which clustered with their respective parental non-hybrid species, *K. ocellatus* and *K. hermaphroditus*, respectively (Fig. 3.2).

Out of the 37,664 DMCs between *K. ocellatus* and *K. hermaphroditus*, 33,329 (88.5%) had intermediated normalised read counts in the hybrids compared to the parental species. This intermediated pattern of DNA methylation in the hybrids compared to the parental species was also confirmed by the used of the significantly different in the analysis of the 5,800 DMCs common in the comparisons between hybrids vs parental species, on which all had intermediate normalised counts values compared to *K. ocellatus* and *K. hermaphroditus*, with no site being either over or under dominated in hybrids.

3.4.3. Genomic context

Of the 5,800 DMCs shared between the comparisons of hybrids versus parental species, 254 (4.38%) were within 2kb upstream gene bodies, representing putative

promoters, 3435 (59.23%) were overlapping gene bodies, while 831 (14.33%) represented potential intergenic regions. Of these, 1280 (22.06%) were within unannotated regions and 2786 unique genes were affected by DMCs in putative promoters and/or gene bodies.

The MDS analysis separating DMCs by its genomic context revealed a clear clustering between parental species across all genomic contexts (e. g. promoters, gene bodies, intergenic and unannotated regions), with the hybrids forming an intermediate cluster differentiated from the parental species clusters (Fig. S3.6).

1,322 genes affected by the 5,800 DMCs common to the comparisons between hybrids vs. parental species mapped to orthologs in the zebrafish genome. Using those for the gene ontology enrichment analysis, I found 217 significantly overrepresented ontologies, which influenced a wide range of biological processes, from cellular regulation, to neurogenesis and organ development (Table S3.2).

3.5. Discussion

Hybridisation events are likely to not only cause genetic, but major epigenetic rearrangements. Given the extensive genetic differences between *K. ocellatus* and *K. hermaphroditus* found in Chapter 2, the intricate relationship between genomic background and DNA methylation patterns found in Chapter 4 and the potential inheritance of DNA methylation patterns found in Chapter 5, I hypothesised that DNA methylation patterns of hybrids would be intermediate between parental species. My results reveal a potential additive effect on DNA methylation patterns of the hybrids compared to the parental species in an ongoing hybridisation occurring between *K. ocellatus* and *K. hermaphroditus* in syntopic natural populations.

Heritable epigenetic variation usually co-exists with genetic variation and are regulated by those either by *cis* and or *trans* effects (Dubin et al. 2015; Richards 2006). *K. ocellatus* and *K. hermaphroditus* are very divergent genetically (Chapter 2), and also showed many differentially methylated cytosines despite being found inhabiting the same mangrove microhabitats. Although I cannot rule out that ecological differences at a very small spatial/temporal scale (e. g. diet, habitat use) between species may have influenced the DNA methylation profiles, the strict relationship between genetic alleles and DNA methylation epialleles together with the evidence that genome and methylomes respond together to environmental variation (Asselman et al. 2015; Berbel-Filho et al. 2019b), suggest that major cause of the DNA methylation differences found between *K. ocellatus* and *K. hermaphroditus* are likely to be primarily-driven by genomic differences between the species.

Given the economic and agricultural interest on heterosis of crop hybrids, the study of the effects of hybridisation on DNA methylation inheritance has been largely explored in plants (Groszmann et al. 2013; Lauss et al. 2018; Michalak 2009), but it has lagged behind in animal systems. Most of the studies in plants have showed prominence of additive effects in DNA methylation patterns with typical Mendelian inheritance of DNA methylation (Hegarty et al. 2011; Moghaddam et al. 2011; Radosavljević et al. 2019; Yu et al. 2011). However, others also show a small number of non-additive epialleles that can arise independently in hybrids (Cara et al. 2013; Sakthivel et al. 2010). The inconsistent conclusions in these studies are likely due to the limited resolution of methylation-sensitive amplification fragment length (MS-AFLP) (He et al. 2013). Studies using high-throughput sequencing approaches have revealed many more non-additive DNA methylation states in crop hybrids (Greaves et al. 2012; He et al. 2010), with those differences generally altering gene expression levels and possibly

contributing to either hybrid heterosis or incompatibly (Chodavarapu et al. 2012). The strong intermediate effects observed in the methylation patterns between of *K. ocellatus* and *K. hermaphroditus* hybrids seem to reflect an important effect of genetic background on DNA methylation levels. This effect is reinforced by the patterns found in the backcross individuals, on which the patterns of DNA methylation are more similar to the parental species than to the F1 hybrids. To date, the few studies evaluating DNA methylation inheritance in hybrids in fishes found that in allopolyploid hybrids, the vast majority of DNA methylation levels is additive (Shao et al. 2018; Xiao et al. 2013). Given the evidence that DNA methylation levels affect gene expression levels and other RNA features (e.g. transcriptional noise) (Anastasiadi et al. 2018; Duncan et al. 2014; Huh et al. 2013), the intermediate effects found here contrast with the findings on gene expression inheritance in fish hybrids, which show predominant transgressive patterns of gene expression in F1 hybrids (McGirr and Martin 2019; Renaut et al. 2009; Roberge et al. 2008), and generally agree with the patterns of inheritance found in *Drosophila* and house mouse (Hughes et al. 2006; Rottscheidt and Harr 2007).

The direction of hybridisation effects in DNA methylation inheritance seem to be affected by the initial degree of divergence of the parental lineages (He et al. 2013). Greaves et al. (2012) using inbred lines of *Arabidopsis* found that non-additive changes in DNA methylation in hybrids are likely to be more frequent where parental methylation levels are different. My findings show extensive intermediate levels of DNA methylation in hybrids at both differentially and non-differentially methylated sites, which could reflect in additive patterns of gene expression. Rottscheidt and Harr (2007) argued that additivity instead of non-additivity is actually expected in more

divergent taxa as divergent traits are regulated by more and more genes, provided that most of them do not show any pattern of dominance among each other.

It is less clear how the differentially methylated cytosines between the parental species that show intermediate levels in hybrids could be related to hybrid fitness. I have shown that the 5,800 DMCs on which hybrids are different from their parental species could significantly affect a wide range of biological processes, many of them involved in important developmental process (i. e. central nervous system development, chordate embryonic development, eye development). Interspecific hybridisation can lead to increased fitness by phenotypic novelty in hybrids (Hovick and Whitney 2014). Alternatively, if the DMCs were causing misregulation, and therefore misexpression in hybrids, there is a possibility that those patterns of epigenetic inheritance may be causing hybrid incompatibility. However, the evidence that F1 hybrids are able to backcross (Chapter 2) indicates that the hybrids are reproductively viable, although I cannot rule out the they might already be showing non-lethal deleterious effects caused by gene misexpression. Another possibility that warrants further research is the effect of allele-specific compensatory effects in hybrids when stabilising selection favour an optimal level of gene expression by compensating the effect of one allele activity through cis and trans regulatory factors (Landry et al. 2007). Increasing research has been showing widespread compensatory effects in hybrids expression levels (Bell et al. 2013; McGirr and Martin 2019), however, the potential effect of compensatory effects in DNA methylation inheritance and the pattern of epiallele dominances is less understood (He et al. 2013).

The ongoing hybridisation between *K. ocellatus* and *K. hermaphroditus* is rather unlikely given the reproductive and genomic divergence between the parental species (Chapter 2). Using methylation-sensitive sequencing, the present study was one of the

first to explore the DNA methylation inheritance patterns in hybrids in animal natural populations, showing that hybrids have intermediate, and thus potentially additive, patterns of DNA methylation compared to the parental species, possibly caused by the strong influence of the genetic background on DNA methylation patterns. The new combination of not only genomic, but epigenomic content merged together by hybridisation may provide a mechanistic explanation for the rapid molecular and phenotypic variation often observed in hybrids.

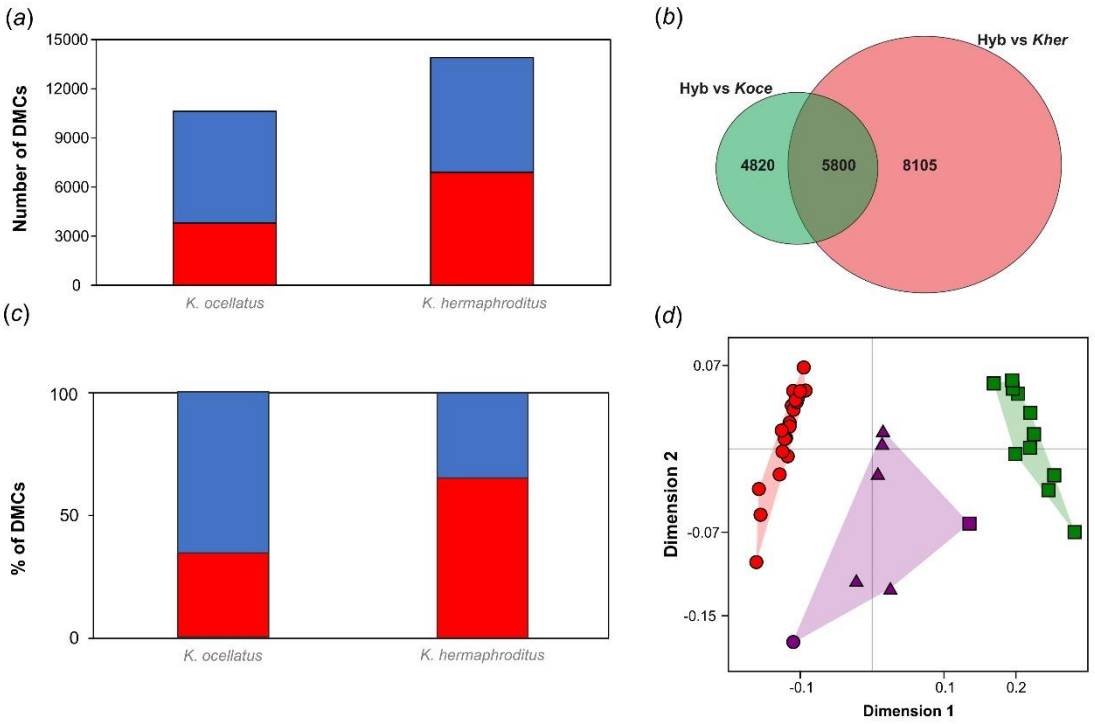


Figure 3.1. (*a*) Number of differentially methylated cytosines (DMCs) in *K. ocellatus* and *K. hermaphroditus* compared to their hybrids. Hypomethylated (logFC value > 1) and hypermethylated (logFC value < 1) DMCs in comparison to hybrids are shown in blue and red, respectively. (*b*) Number of DMCs of *K. ocellatus* and *K. hermaphroditus* in comparison with hybrids, and their overlap. (*c*) Percentage of DMCs either hypomethylated (blue) or hypermethylated (red) of the 5,800 DMCs common to the comparisons between hybrids vs parental species. (*d*) Multidimensional scaling analyses of the normalised counts for the 5,800 DMCs common to the comparisons between hybrids vs parental species. Squares represent *K. ocellatus*, circles represent *K. hermaphroditus*, and triangle represent hybrids. Backcrosses are represented by purple shapes according to their respective parental species.

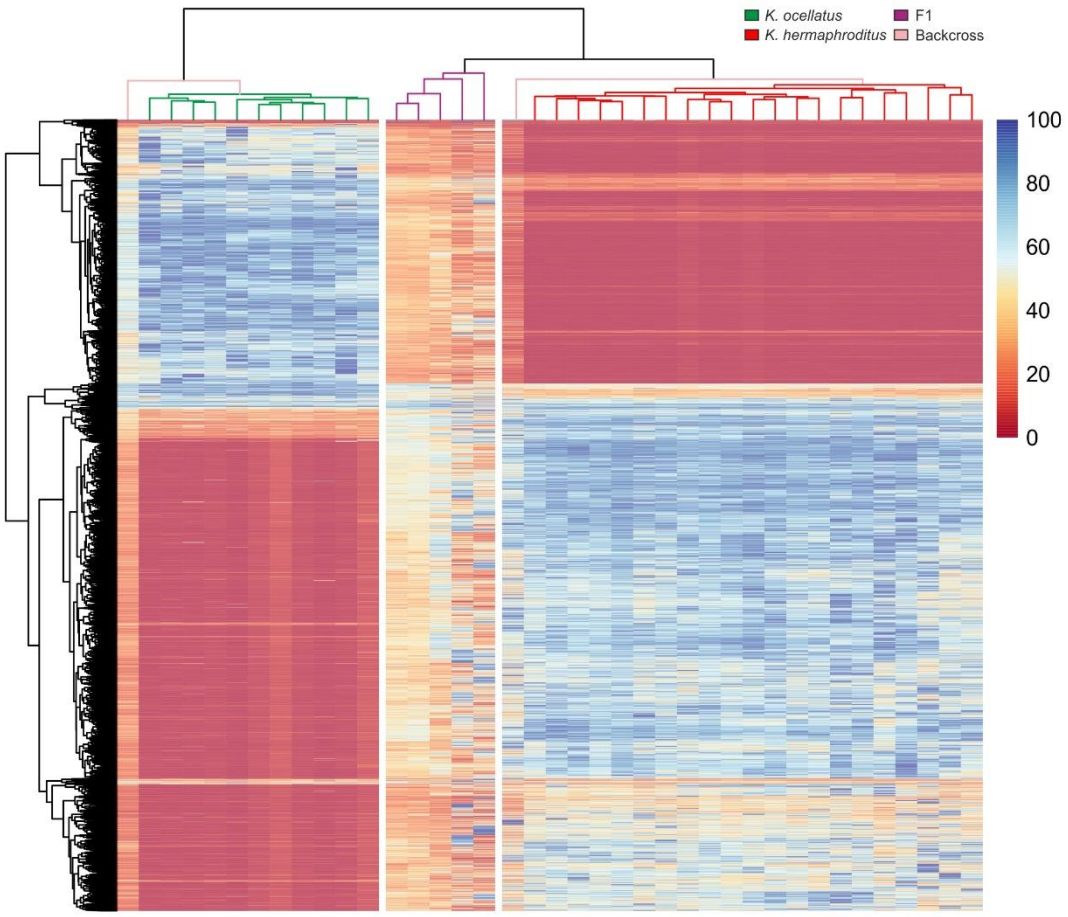


Figure 3.2. Hierarchical clustering of the scaled normalised counts for the 5,800 differentially-methylated cytosines (DMCs) common to both comparisons between hybrids and parental species. Each cell represents an individual DMC, and each column represent an individual fish. Hypomethylated DMCs in blue and hypermethylated DMCs in red.

CHAPTER 4: What does drive DNA methylation plasticity? Epigenetic responses of two inbred lines to different rearing environments*

*A version of this work has been published as: **Berbel-Filho, W. M.**, D. Rodríguez-Barreto, N. Berry, C. Garcia de Leaniz, and S. Consuegra. 2019. Contrasting DNA methylation responses of inbred fish lines to different rearing environments. Epigenetics 14:939-948.

4.1 Abstract

Epigenetic mechanisms generate plastic phenotypes that can become locally adapted across environments. Disentangling genomic from epigenomic variation is challenging in sexual species due to genetic variation among individuals, but it is easier in self-fertilizing species. I analysed cytosine methylation patterns of two highly inbred strains of a naturally self-fertilizing fish reared in two contrasting environments to investigate the obligatory (genotype-dependent), facilitated (partially depend on the genotype) or pure (genotype-independent) nature of the epigenetic variation. I found higher methylation differentiation between genotypes than between environments. Most methylation differences between environments common to both strains followed a pattern where the two genotypes (inbred lines) responded to the same environmental context with contrasting DNA methylation levels (facilitated epialleles). My findings suggest that, at least in part, DNA methylation could depend on the dynamic interaction between the genotype and the environment, which could explain the plasticity of epigenetically mediated phenotypes.

4.2. Introduction

Phenotypic plasticity is widespread phenomenon on nature, however its relative role on evolution is still a hotly debated topic in evolutionary biology (Des Marais et al. 2013; Fusco and Minelli 2010; Pigliucci et al. 2006), as, given its implications for individual fitness and organisms' differentiation, it has been considered by some authors as one of the major concepts currently missing into the framework of the modern evolutionary synthesis (Laland et al. 2014; Pigliucci 2007).

The variety of phenotypes produced by the interaction between genotypes and environments can be described by reaction norms (Des Marais et al. 2013; Pigliucci et al. 2006). Among those, the evidence for genotype-by-environment (G x E) interactions is pervasive (Sambandan et al. 2008; Sultan 2015) and its importance have been specially studied on human disease context (Hunter 2005), and plant/animals breeding systems (Hassen et al. 2018; Streit et al. 2012). Although several genetic mechanisms have been used to explain G x E interactions such as antagonistic pleiotropy, mutation contingency, epistasis, differential sensitivity and conditional neutrality (reviewed in Des Marais et al. 2013), little is known about its effects on non-genetic factors (i.e. epigenetics) (Herman and Sultan 2016).

Epigenetic modifications represent rapid intracellular mechanisms for changes on the transcriptional machinery, which might be influenced by environmental changes (Richards 2006), and ultimately, influence phenotypic variation (Richards et al. 2017). In phenotypic plasticity studies, the genome and epigenome are often experimentally confounded (Sultan 2015) and an implicit assumption is made that they react to environmental variation following similar norms of reaction. However, this may not necessarily be the case as epialleles may have different degree of sensibility/respond to environmental change, thereby generating additional phenotypic variation different

from the one generated by genomic responses (Richards et al. 2017; Verhoeven et al. 2016).

To what extent epigenetic modifications act independently from genomic variation is key, but still unexplored question to understand the potential role of epigenetics in evolution (Duncan et al. 2015; Richards et al. 2017; Verhoeven et al. 2016), as epigenetic variation completely under genetic control would not contribute any additional adaptive value (Verhoeven et al. 2016). Richards (2006) classified epigenetic variation in obligatory, facilitated or pure epialleles, based on their degree of autonomy from the underlying genotype. Obligatory epialleles would be fully dependent on genetic variation and should show no variation across environmental change (Schmitz et al. 2013), whereas facilitated and pure epigenetic variation would differ in their degree of autonomy from the genotype (from partially depend to independent) (Schmitz et al. 2013), acting as potential intermediaries between environmental conditions and genome responses.

Among the epigenetic modifications, DNA methylation is the best studied, and plays an important role in the pre-transcriptional control of several biological processes, such as cell differentiation and genomic imprinting (Koch et al. 2016; Moore et al. 2013). While the relationships between DNA methylation patterns, environmental conditions and phenotypic traits have been widely investigated (Baerwald et al. 2016; Keller et al. 2016) the plasticity of DNA methylation itself and its potential implications for downstream epigenetically-regulated phenotypes have rarely been considered (Herman and Sultan 2016; Teh et al. 2014). Here, I investigated the relative roles of the genotype and the rearing environment in the diversity of genome-wide cytosine methylation states in the brain of two genetically different and highly inbred self-fertilising lines of *Kryptolebias marmoratus* reared in two contrasting environments

(poor and enriched). I hypothesised that if DNA methylation was mostly autonomous and shaped by environmental change, a higher number of different epialleles would be found between environments, regardless of the genetic background, than if DNA methylation was mostly under genetic control, where most of the epigenetic differences would be between genotypes.

4.3. Material and methods

The self-fertilising hermaphroditic mangrove killifish (*Kryptolebias marmoratus*) (Tatarenkov et al. 2017), has naturally inbred lines (Ellison et al. 2015) which inhabit mangroves that differ markedly in habitat complexity, ranging from temporary pools to mangrove leaf litter and crab burrows, all influenced by tidal variation (Ellison et al. 2012; Taylor 2012). I used two highly inbred strains (R and DAN), originally from Belize (Lins et al. 2017; Tatarenkov et al. 2010) and kept in the laboratory for at least 20 selfing generations (Ellison et al. 2015).

I compared brain methylation of fish reared under enriched and impoverished conditions, as previous studies had shown environmental enrichment can affect behavioural flexibility (Roberts et al. 2011), brain size and cognition (Kihslinger et al. 2006), and induce epigenetic modifications during early development (Champagne 2008). I used two different habitats with different levels of environmental enrichment: 1) a physically enriched habitat, where fish were placed in contiguous rectangular tanks (9 cm depth x 12 cm width x 8 cm length) filled with 400ml of brackish with one perforated artificial log (3 cm depth x 4 cm width x 4 cm length) and three artificial plants to simulate a complex habitat (Fig. S4.1) and 2) a barren habitat (hereafter called poor) with the same tank conditions but without physical enrichment. For both strains, five initial hermaphrodite progenitors of similar size (mean=3.8 cm, sd= ±0.12) and age (mean = 417.3 days post hatchling, sd= ±13.4) were chosen. Eggs from these

progenitors were maintained individually in pots containing 100ml of brackish water and checked daily. Upon hatching, alevins were randomly distributed between the treatment tanks (enriched and poor). Given the difference in hatchlings per parent, the initial experimental set up consisted of 29 R fish (18 in enriched habitat, 11 in poor habitat) and 21 DAN fish (10 in enriched, 11 in poor). Fish were maintained under standard laboratory conditions (25-27 °C, 16-18‰ ppm salinity under a 12h light:12h dark photoperiod).

4.3.1. Genome-wide DNA methylation data

To minimise any potential age effect on methylation profiles, all fish sequenced were euthanized at 10 months old, using tricaine methane-sulfonate (MS-222) following Home Office Schedule 1. Individuals were dissected and their brains kept in molecular grade ethanol for DNA extraction. Given the effect of environmental enrichment on brain development and gene expression (Rampon et al. 2000; Salvanes et al. 2013), brain DNA was extracted from 22 individuals for epigenetic analysis (six DANs: three from poor, three from enriched environment; 16 Rs: six from poor, ten from enriched environment) using Qiagen DNeasy Blood and tissue kit (Qiagen®) following the manufacter protocol. This DNA was also used to amplify 23 microsatellites loci to confirm individuals' genotypes and degree of genetic variation, following Ellison et al. (2011).

Bisulphite converted genomic DNA libraries were prepared using Diagenode® Premium Reduced Representation Bisulphite Sequencing (RRBS) Kit according to manufacturer's indications and sequenced on an Illumina® NextSeq 500 platform (Cardiff University, Genomics Research Hub) using a 1x75pb single-end run. PCR

fully methylated and unmethylated spike controls from the RRBS kit were added to monitor bisulphite conversion efficiency.

4.3.2. Sequence quality check and alignment

Initial quality assessment of the raw sequences was performed using was performed using FastQC (Andrews 2010). TrimGalore! (Krueger 2016) was used to trim low-quality base calls, adapter contamination and positions filled in during end-repair. Trimmed reads were aligned to the *Kryptolebias marmoratus* reference genome (NCBI ASM164957v1; Rhee et al. (2017), assembly size = 680.3Mb; number of scaffolds = 3,073; N50 = 2,229,659; GC content = 37.76%; annotation date = 22/11/2018) prior bisulphite conversion using Bismark v0.17.0 (Krueger and Andrews 2011), which was also used for cytosine methylation calls. Non-CpG methylation levels were low in average (0.56% CHG and 0.49% CHH, Table S4.1), therefore I only considered methylation within CpG context for the analysis, as these are the most common methylation sites for vertebrates (Feng et al. 2010). To minimise potential biases due sequencing depth, I included only CpG sites with a minimum coverage of 10 reads in each sample across the 22 individuals sequenced (Shafi et al. 2017) for differential methylation analysis. To account for biological variation, samples were grouped into experimental groups representing genotypes and environments: “DAN enriched”, “DAN poor”, “R enriched”, “R poor”. Mapped reads were processed using SeqMonk (Andrews 2007).

4.3.3. Differentially methylated cytosines and regions

To identify differentially methylated cytosines (DMCs) across experimental groups, I used logistic regression with $p < 0.01$ after multiple testing correction (Benjamini-

Hochberg) and >20% minimal CpG methylation difference (|ΔM|), using R bridge in SeqMonk. I also performed t-tests across experimental group replicates using the same software, to generate a more conservative list of DMCs, only considering those shared by both statistical approaches. To identify differently methylated regions (DMRs), I performed a genome-wide unbiased DMR detection using tilling windows of 1000bp on windows with at least five CpGs with ≥10 reads across all individuals.

I used the scores of methylation for DMCs and DMRs between genotypes and environments for principal component analysis (PCA) in R v. 3.4.3. To test for the effect of the genotype, environment and their interaction on the methylation scores, I used linear models with the scores for the first two PCA axis (>70% of the total variation) as a function of genotype, environment and their interaction. I then individually compared DMCs and DMRs between genotypes, followed by a comparison between environments using the same methods described before for DMCs and DMRs identification. Subsequently, a comparison within each genotype between environments was carried out to identify potential environment-dependent DMCs and DMRs. From these comparisons, I identified annotated DMCs and DMRs shared between genotypes, which should represent commonly affected DMCs regardless of the genetic background.

4.3.4. Methylation patterns and epiallele classification

Following the epiallele classification in Richards (2006), I classified the DMCs and DMRs shared across genotypes between environments as facilitated, when displaying different directions of variation (non-parallel) on methylation scores across genotypes in the same environment (i.e. hypermethylated in an environment for one genotype and hypomethylated in the other), or pure when displaying the same direction of

variation (parallel) across genotypes and environments (i. e. hypermethylated or hypomethylated for both genotypes in the same environment).

4.3.5. Molecular network analysis and centrality metrics

To identify potential functional implications of variation in DNA methylation for the annotated DMCs identified across genotypes between environments, I built a functional gene network using GeneMANIA (Warde-Farley et al. 2010). To identify central genes and retrieve centrality metrics (del Rio et al. 2009) within the molecular network, I used NetworkAnalyzer plugin into Cytoscape v. 3.7.1 (Assenov et al. 2008). Panther GO terms (Mi et al. 2016) was used to identify biological process and pathways for the most connected genes (>10 connections) within the network.

4.4. Results

After quality filtering, approximately 273 million reads were retained, averaging 12 million reads per sample. Of those ~ 62.9 % were uniquely mapped reads to the reference genome. Overall bisulphite conversion was 99.6% (Table S4.1). Fish were genotyped for 23 microsatellites (Ellison et al. 2011). Genetic differences were identified between inbred lines (F_{ST} = 1.00, Table S4.2), but not within lines.

In total, I identified 5.5 million cytosine sites, of which 139,908 CpG sites fulfilled the minimum coverage requirements (10 reads across all individuals), representing 1.2% of the total number of cytosines om the mangrove killifish genome. This result is similar to recent RRBS studies in other fish (1% in rainbow trout (Baerwald et al. 2016); 1.5-2% in guppies (Hu et al. 2018)).

The majority of the cytosines surveyed mapped gene bodies (71.32%) or intergenic regions (19.10%), while only 2.54% were located on putative promoters. Linear

models using the PCA scores for 1,064 DMCs and 194 DMRs identified between genotypes and environments revealed that although genotype, environment, and genotype x environment interactions were significantly related DNA methylation scores, genotypes explained more of the variance for axis one (54% of overall variation for both DMCs and DMRs) than environment (Fig. S4.2; Table S4.3) When predictors were analysed individually, differences between genotypes also corresponded to a higher number of DMCs (817 vs 594, four DMCs shared) and DMRs (43 vs 17, no DMR shared) than differences between environments (Figs. 4.1; S4.3). Within genotypes, 357 and 3,632 DMCs (25 and 373 DMRs) were identified between environments, for DAN and R, respectively (Figs. 4.1; S4.3). An additional analysis on three different subsets of six R individuals (to match the number of DAN individuals) identified similar numerical differences in DMCs between lines (Table S4.4). Unsupervised hierarchical clustering revealed distinctive methylation profiles between groups, except for comparisons between environments, where one and two individuals from the poor environment clustered with individuals from the enriched environment for DMCs and DMRs respectively (Figs. 4.1; S4.3).

Twenty-five annotated DMCs and four DMRs between environments were shared across genotypes, potentially representing environmentally-affected DMCs, independently of genetic background. Of these, based on the direction of methylation across environments, 22 out 25 DMCs were classified as potentially facilitated, with methylation scores following a genotype-specific pattern under similar environments (Table 4.1). This pattern was supported by the PCA results based on the DMCs methylation scores, which indicated different methylation profiles between environments (PC1 explaining 55.83% of the variation), as well as the genotypes (PC2 explaining 22.39% of the variation) (Fig. 4.2a). Linear model analysis indicated that

PC1 values were significantly influenced by the environment (t-value= 1.63, df= 1, p=0.003) and the interaction between genotype and environment (t-value= -11.25, df= 1, p<0.001), while PC2 values were only significantly influenced by the genotype (t-value=-1.64, df= 1, p<0.001) (Table 4.2a). Methylation differences (with a lower threshold of 20%) for the facilitated DMCs ranged from 20.25% to 48.57% (Table 4.1). The facilitated DMCs were mostly hypermethylated on enriched environments with respect to poor environments for DAN fish while the opposite pattern was found for R fish (Table 4.1; Figs 4.2c; S4.4). The four DMRs between environments and shared by genotypes were also classified as facilitated, overall following the same methylation pattern found on the facilitated DMCs (Fig. S4.5; Table S4.5).

Only three of the annotated DMCs within or neighbouring gene bodies were considered pure (Table 4.1; Fig. S4.6). Average methylation differences for pure DMCs ranged from 25.4% to 34.37% (Table 4.1). The PCA only using pure DMCs showed a different pattern from the facilitated DMCs, with the PC1 separating environments explaining 72.38% of the variation, and the PC2 partially differentiating genotypes explaining 18.85% of the variation (Fig. 4.2b). PC1 loadings were significantly affected by the environment (t =-2.81, df= 1, p<0.001) and the genotype (t =-2.28, df= 1, p=0.008), while PC2 loadings were only significantly influenced by genotype (t =0.29, df= 1, p=0.003) (Table 4.2b, Fig. 4.2d).

Molecular network analysis revealed a highly connected network linked by genetic interactions and co-expression interactions, that was composed by 23 input annotated DMCs (the uncharacterised LOC108245430 and ubald1 with no identified connections were removed) and 20 neighbouring genes (Fig. S4.7). Centrality parameters, such as average degree (averaged number of connections per gene, mean=10.55; SD ± 5.89), closeness (average distance of the all the genes in the network to a certain gene;

mean=0.53 ± 0.06), and radiality (the easiness of reaching any gene from certain gene, 0 to 1 index with 1 being the highest; mean=0.77 ± 0.06) (Table S4.6), suggested that any alteration of the expression of the genes contained in the network was likely to have major effects on genetic interactions and gene expression levels.

Twelve of the 15 most connected genes within the network (>10 connections), were input genes (i.e. genes affected by DMCs between environments and shared by genotypes). Gene ontology analysis showed that some of these genes are involved on important cellular and metabolic processes in zebrafish, such as regulation of transcription by RNA polymerase and gene expression (myc), RNA modification (trit1), intracellular calcium content (ryr3), and lipid metabolism (sorcs2), as well as pathways related to angiogenesis and stress response (ryr3 and myc) (Table S4.7).

4.5. Discussion

A full comprehension of the evolutionary role of epigenetics in generating plastic phenotypes pass through the evaluation of the interplay between epigenetic variation, its genetic background and environmental changes (Richards et al. 2017; Sultan 2015). Given the important role of genetic background on DNA methylation (Dubin et al. 2015; Eichten et al. 2013; Schmitz et al. 2013), the diversity of epigenetic states in natural populations can be difficult to disentangle due to the high genetic variability among individuals. By using two highly inbred strains of the mangrove killifish *Kryptolebias marmoratus* reared under controlled environmental conditions, I have been able to estimate significant differences on methylation profiles among genotypes and environments, as well as detect epigenetic states with different levels of autonomy from its genetic background.

Environmental enrichment in fish affects brain structures (Kihslinger et al. 2006; Näslund et al. 2012; von Krogh et al. 2010), however few studies have investigated which molecular mechanisms underlie these changes (Lema et al. 2005; Salvanes et al. 2013) and whether it varies across different genetic backgrounds. *Kryptolebias marmoratus* populations are composed by naturally inbred lines living in highly variable habitats, both temporally and spatially (Ellison et al. 2012; Taylor 2012). Several studies using *K. marmoratus* isogenic lines have identified phenotypic differences between lines (e.g. mate choice (Ellison et al. 2013), reproduction (Lin and Dunson 1995), cannibalistic behaviour (Wells and Wright 2017), sexual differentiation (Turner et al. 2006), gene expression in response to infection (Pawluk et al. 2018)) even when exposed to similar environmental conditions (Earley et al. 2012). These findings suggest a potential role of non-genetic mechanisms, including DNA methylation (Fellous et al. 2018), in mediating these plastic phenotypes. My results indicate that genotypes have an overriding influence on brain DNA methylation patterns, and that their effect is greater than that caused by environmental enrichment. I only found a few DMCs that can be considered facilitated or pure epialleles, supporting the idea that environmentally-induced autonomous DNA methylation may be limited (Dubin et al., 2015). Yet, the DNA methylation patterns of these putative independent epialleles indicated that DNA methylation outcomes could depend on specific combinations of the genotype and environmental conditions.

Genotype-by-environment interactions are the most common reaction norms found on nature (Sultan 2015) being commonly observed from in behavioural or physiological traits to and intermediate phenotypes, such as gene expression (Herman and Sultan 2016). Although several genetic mechanisms have been described to explain G x E interactions (Des Marais et al. 2013), the role of DNA methylation remains largely

unexplored. The pattern found here could represent either a cause or a phenotypic consequence of G x E interactions, as changes in DNA methylation at single nucleotides tend to occur at a much higher rate than mutations in nucleotide sequence (van der Graaf et al. 2015).Functionally, most of the DMCs were found in gene bodies and were shown to be highly integrated within a gene network of genetic interactions and co-expression. Recent evidence in plants (Horvath et al. 2019) and animals (Huh et al. 2013) indicated that gene body methylation has a role on reducing erroneous transcription. Indeed, some of the genes affected by the facilitated DMCs are related to the regulation of RNA polymerase activity and gene expression patterns (myc and trit1). The G x E pattern found here may be affecting transcripts variability between genotypes even under similar condition, however further research is needed to address this question.

My results, suggest that, at least in part, cytosine methylation patterns are the result of a dynamic interaction between genotypes and the environment. Further research to investigate whether the patterns found here might influence transcription might provide a potential mechanistic explanation for the genotype by environment patterns often observed in phenotypically plastic responses (Des Marais et al. 2013; Sultan 2015)

Table 4.1. Methylation differences averaged (percentage) for differentially methylated cytosines (DMCs) between environments (poor, enriched), shared between genotypes (DAN, R) which overlap annotated genes (reference genome ASM164957v1). Epiallele classification (pure or facilitated) followed Richards (2006). Positive and negative values represent increased and decreased methylation in enriched and poor environments, respectively. Q-value is the p-value adjusted by False Discovery Rate (FDR=0.05).

Gene symbol	**Entrez gene name**	**Epiallele classification**	**Meth diff DAN**	***Q*-value**	**Meth diff R**	***Q*-value**
acvr2a	activin A receptor type 2A	P	34.37	0.007	29.17	<0.001
col25a1	collagen type XXV alpha 1 chain	F	43.61	0.005	-21.63	0.005
dmap1	DNA methyltransferase 1 associated protein 1	F	26.99	<0.001	-21.69	<0.001
foxp4	forkhead box P4	F	22.50	<0.001	-37.56	<0.001
gpc5	glypican 5	F	31.82	0.01	-32.06	<0.001
mipol1	mirror-image polydactyly 1	F	35.85	<0.001	-30.59	<0.001
necab2	N-terminal EF-hand calcium binding protein 2	F	20.25	0.01	-23.56	<0.001
neo1	neogenin 1	F	20.25	<0.001	-31.24	<0.001
nudcd1	NudC domain containing 1	F	39.76	<0.001	-21.05	<0.001
ramp3	receptor activity-modifying protein 3-like	P	-27.12	0.037	-29.15	<0.001
ryr3	ryanodine receptor 3	P	-30.48	0.003	-25.4	<0.001
sorcs2	sortilin-related VPS10 domain containing receptor 2	F	36.81	0.008	-33.20	<0.001
trit1	tRNA isopentenyltransferase 1	F	20.38	<0.001	-29.70	<0.001
trmt44	tRNA methyltransferase 44	F	23.98	<0.001	-28.21	<0.001
ubald1	UBA like domain containing 1	F	36.76	0.019	-31.02	0.004

zeb2	zinc finger E-box binding homeobox 2	F	31.83	<0.001	-28.35	<0.001
znf516	zinc finger protein 516	F	31.51	<0.001	-31.60	<0.001
zranb3	zinc finger RANBP2-type containing 3	F	41.41	<0.001	-33.79	<0.001
LOC108234847	adhesion G protein-coupled receptor L3-like	F	48.57	<0.001	-39.35	<0.001
LOC108240988	non-muscle caldesmon-like	F	32.49	0.005	-32.78	<0.001
LOC108243470	protein-methionine sulfoxide oxidase mical2b-like	F	33.35	<0.001	-25.70	<0.001
LOC108243852	receptor-type tyrosine-protein phosphatase N2-like	F	37.97	0.014	-30.28	<0.001
LOC108245430	uncharacterized protein	F	42.94	<0.001	-41.49	<0.001
LOC108247402	spectrin beta chain, non-erythrocytic 1-like	F	40.73	0.006	-27.58	<0.001
LOC108251479	transcriptional regulator Myc-B-like	F	30.49	0.010	-25.13	<0.001

Table 4.2. Linear model of principal component scores for mangrove killifish epialleles shared between genotypes (R, DAN) and environments (poor, enriched).

	t-values	Prop. of variance (%)	df	p-value
(*a*) Facilitated epialleles				
PC1 scores				
Genotype	1.63	0.06	1	0.80
Environment	1.58	12.41	1	0.003
Genotype x Environment	-11.25	68.08	1	<0.001
PC2 scores				
Genotype	-1.64	91.35	1	<0.001
Environment	8.28	0.21	1	0.49
Genotype x Environment	10.29	0.35	1	0.38
(*b*) Pure epialleles				
PC1 scores				
Genotype	2.28	13.18	1	0.008
Environment	-2.81	59.28	1	<0.001
Genotype x Environment	0.98	0.09	1	0.80
PC2 scores			1	
Genotype	0.29	37.96	1	0.003
Environment	-1.82	1.14	1	0.56
Genotype x Environment	-2.97	0.67	1	0.65

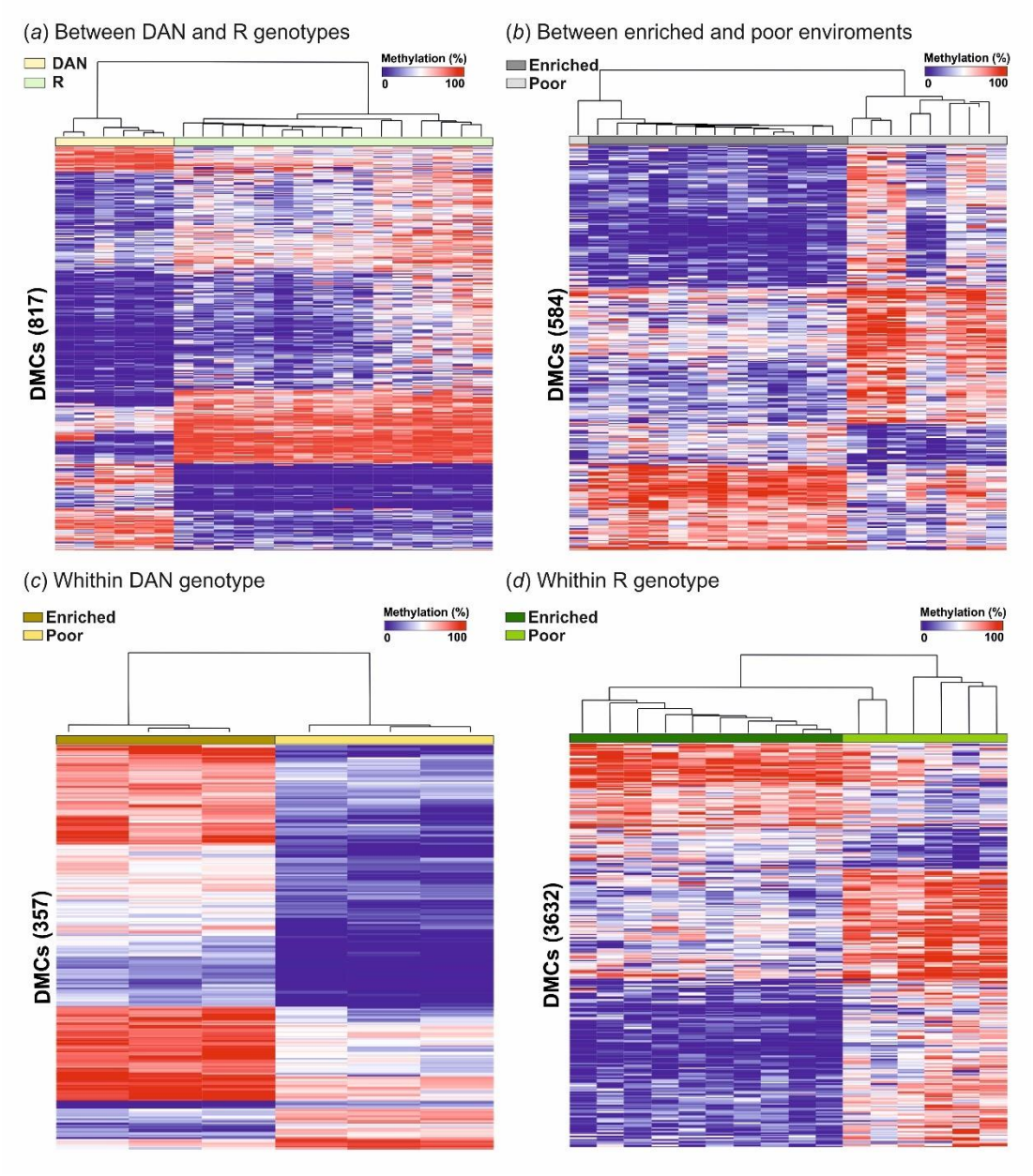


Figure 4.1. Heat map illustrating percentage of methylation for all differentially methylated cytosines (DMCs) identified (*a*) between genotypes, (*b*) between environments, (*c*) between environments for DAN strain, and (*d*) between environments for R strain (logistic regression q< 0.01 and |ΔM|>20%, and *t* test p <0.01) using unsupervised hierarchical clustering. Rows represent a unique CpG site and columns individual fish.

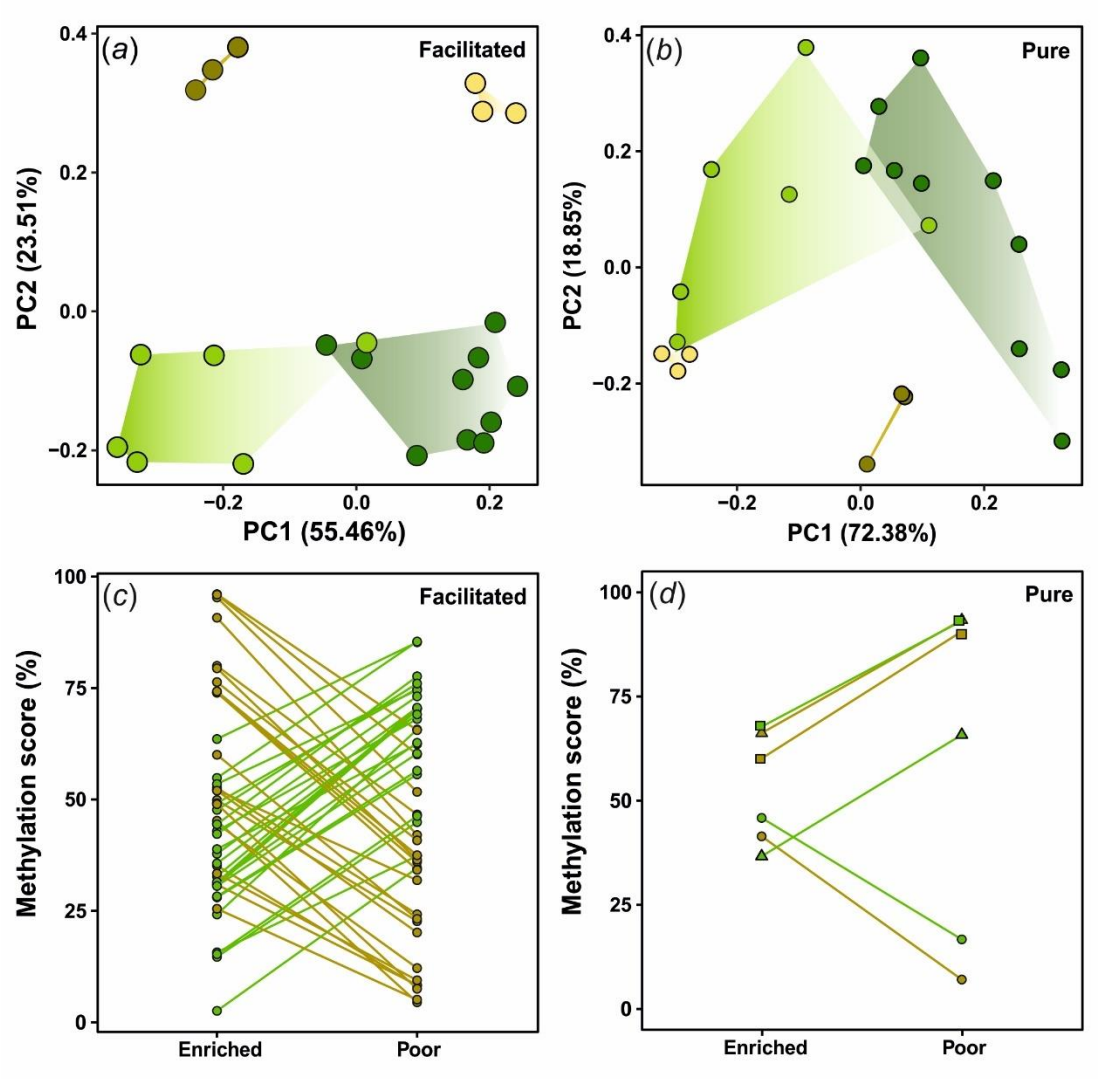
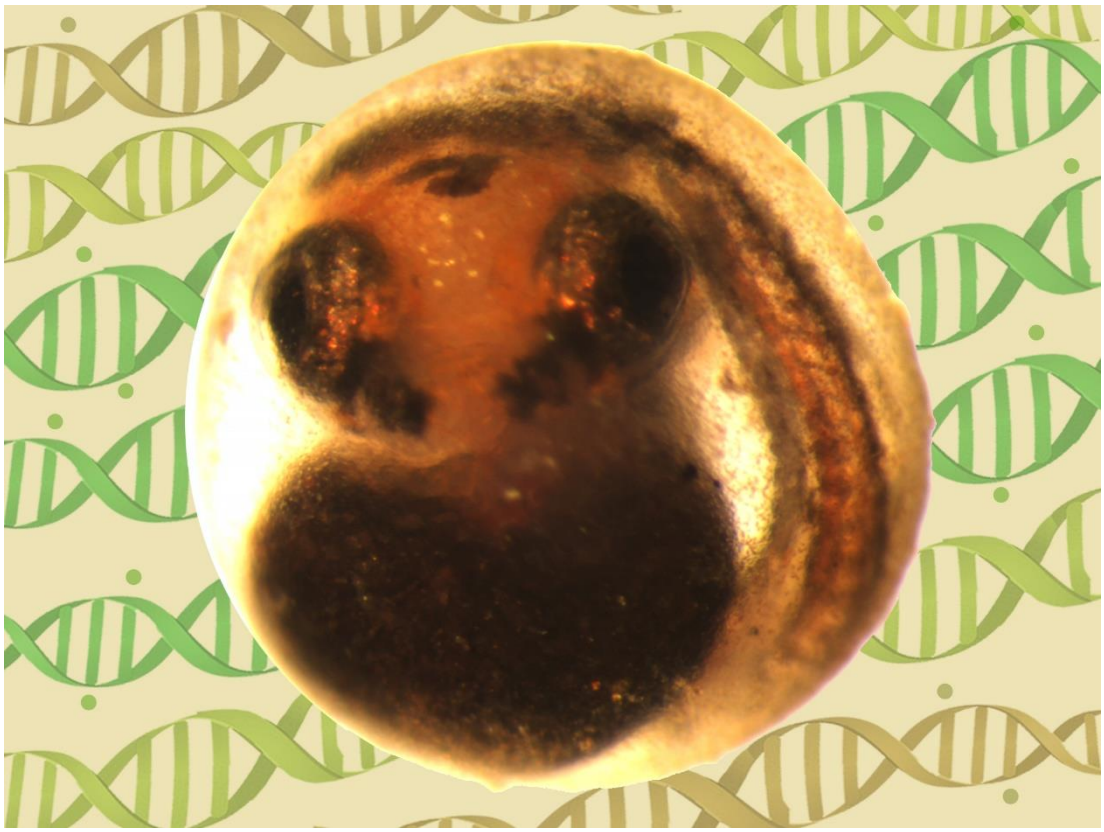


Figure 4.2. Principal component analysis (PCA) and reaction norms of epialleles across genotypes and environments. PCAs were based on individual scores of methylation across either (*a*) facilitated or (*b*) pure annotated differentially-methylated cytosines (DMCs). Dark yellow for DAN individuals on enriched environments; light yellow for DAN genotype on poor environments; dark green for R individuals on enriched environments; light green for R genotype on poor environments. Each reaction norm represents the change on averaged methylation scores (in percentage) for (*c*) facilitated and (*d*) pure epialleles annotated DMCs across environments. Different colours represent the genotypes (yellow for DAN; green for R). Different shapes (*d*) represent different annotated DMCs. Epialleles were classified according

to Richards (2006). Detailed information for each annotated DMCs methylation score across genotypes is available at Table 4.1.

CHAPTER 5: Environmental enrichment induces behavioural but limited epigenetic parental effects in a self-fertilising fish



5.1 Abstract

Parental effects influence offspring phenotypes through pre- and post-natal routes but little is known about their molecular basis. Epigenetic modifications are influenced by the environment and could be involved in inter- and transgenerational parental effects. Taking advantage of the self-fertilising nature of the mangrove killifish *Kryptolebias marmoratus*, I investigated the effects of rearing environment on parents and offspring by comparing neophobia, metabolic rate and brain DNA methylation patterns of genetically-identical fish reared in enriched or barren environments. Parental fish reared in enriched environments had lower cortisol levels, lower metabolic rates and were more active and neophobic than those reared in barren environments. They also differed in 1854 methylated cytosines (DMCs). I also found evidence of limited but significant intergenerational effects on the DNA methylation patterns of the offspring. Offspring activity was influenced by the parental activity. Among the DMCs of the parents, 98 followed the same methylation patterns in the offspring, three of which were significantly influenced by parental environments irrespective of their own rearing environment. My results suggest that the environment experienced by the parents influences behaviour and brain cytosine methylation patterns of the offspring in an environment-specific manner.

5.2. Introduction

Parental effects occur when maternal, paternal or both parental phenotypes affect offspring phenotypes (Bonduriansky and Day 2018; Uller et al. 2013). Such effects occur in a wide range of taxa (Uller 2008) via different pre- and post-natal routes (e. g. microhabitat selection for eggs, reproductive investment, intrauterine environment, parental care). Parental experiences can affect offspring fitness (Burton and Metcalfe 2014), however knowledge about the molecular mechanisms of the pre-natal parental effects is still limited (Gluckman et al. 2005; Illum et al. 2018; Jensen et al. 2014).

The transmission of some parental effects via germline has been related to genetic mechanisms, for example, the frequency of some deleterious mutations in sperm is associated with increasing male's age, affecting offspring fitness (Wyrobek et al. 2006). However, it is likely that pre-natal non-genetic mechanisms also play a major role in parental-offspring information transfer (Danchin et al. 2011; English et al. 2015; Jablonka and Raz 2009), as genetic-based inheritance solely cannot fully explain the variation of offspring phenotypes (Danchin et al. 2011). Epigenetic modifications, such as DNA methylation, histone modifications and microRNAs, mediate rapid changes in transcription influenced by environmental changes (Richards 2006) that can affect phenotypes (Richards et al. 2017; Verhoeven et al. 2016). Among the epigenetic mechanisms, cytosine methylation is the best characterized, being important on several biological processes, from genomic imprinting to cell differentiation (Jones 2012; Lea et al. 2017). In mammals, DNA methylation generally affects regulatory regions by supressing gene expression (Moore et al. 2013), whereas methylation in gene bodies can contribute to reduced transcriptional noise (Huh et al. 2013). Thus, differential methylation can affect gene expression and result in phenotypic plasticity (Herman and Sultan 2016). However, while the transmission of environmentally-induced epialleles via DNA methylation

from parents to offspring has been identified in plants, whether epigenetic mechanisms can provide a heritable memory of environmental influence in animals remains controversial (Heard and Martienssen 2014), as well as the potential adaptive value of this type of transmission (Perez and Lehner 2019).

The parental rearing environment can induce phenotypic modifications during early development which can be long-lasting and potentially intergenerational (Burton and Metcalfe 2014). A well-known example is the effect of structural environmental complexity (i.e. adding physical structures to a barren environment) on behaviour (Braithwaite and Salvanes 2005; Roberts et al. 2011), physiology (Näslund et al. 2013), cognitive capacity (Salvanes et al. 2013) and brain structure (Kihslinger et al. 2006) in fish. Physical structures are critical for most fish at different points of their life cycle (e. g. for spawning, sheltering, foraging), suggesting that structural complexity is an important ecological factor of their natural environment (Näslund and Johnsson 2016). Captive fish reared in enriched environments have increased survival in the wild compared to those reared in impoverished environments (D'Anna et al. 2012; Roberts et al. 2014), as well as enhanced cognitive capacity and behavioural flexibility (Salvanes et al. 2013; Spence et al. 2011; Strand et al. 2010). However, little is known about the molecular mechanisms underlying plastic responses to environmental enrichment, or whether these changes could be transmitted across generations (Näslund et al. 2012; Näslund and Johnsson 2016).

Kryptolebias marmoratus (Poey 1880) is a self-fertilising fish living in mangrove forests in North and Central America (Tatarenkov et al. 2017), occupying a wide range of mangrove fossorial microhabitats influenced by periodical tide variation (Ellison et al. 2012). *Kryptolebias marmoratus* displays aggression towards conspecifics (Taylor 2000) that vary depending on kinship relationship (Earley and Hsu 2008; Edenbrow and Croft 2012; Ellison et al. 2013). The species inhabits inherently heterogenous mangrove habitats, with multiple

genotypes being found in the same microhabitat (Ellison et al. 2012). Intraspecific aggression results in emersion (jumping out of water) and dispersal (Taylor 2000), suggesting that *K. marmoratus* environmental complexity plays an important role on their ecology and behaviour.

I reared of genetically identical *K. marmoratus* parents and their offspring in matched-mismatched environments with different levels of structural complexity to examine the influence of environmental enrichment on individual physiology and behaviour, and the potential role of epigenetic mechanisms (brain cytosine methylation) to mediate environmentally-induced parental effects.

5.3. Material and methods

5.3.1. Experimental design

I used a highly inbred strain of *K. marmoratus* (R, Chapter 4) kept under standard laboratory conditions (25-27 °C, 16-18‰, 12h light: 12h dark photoperiod) for at least 20 selfing generations (Ellison et al. 2015). Eggs from five fish of similar size and age were reared individually until hatching, when larvae were transferred to tanks with either enriched environment (shelter and plants, n = 14) or poor environment (identical except no enrichment, n=13), where they were kept for 10 months as the parents (P) in the tanks described in detail in Chapter 4.

To standardise potential age-related parental effects, eggs were only collected from P of similar age (7-10 months). The offspring (F1) of the first five parents which started laying eggs within 7-10 months (three from enriched and two from poor environments) were set up following a factorial design with matched or mismatched parent-offspring environments (Fig. S5.1; Table S5.1). F1 consisted of 15 mismatched individuals (seven poor to enriched and eight enriched

to poor), and 13 matched individuals (eight enriched to enriched and five poor to poor). All experiments mentioned below were approved by Swansea University Ethics Committee (permit STU_BIOL_30484_110717192024_3).

5.3.2. Metabolic rate and cortisol measurements

I measured basal metabolic rate (oxygen consumption) of 55 adults (14 P enriched; 13 P poor, eight F1 from enriched parents reared in enriched environment; ten F1 from enriched parents reared in poor environment; five F1 from poor parents reared in poor environment; five F1 from poor parents reared in enriched environment), 30 of which were also analysed for epigenetic variation using RRBS protocol (Table S5.1). Oxygen levels were calibrated in trials using saturated oxygenated water (100% dissolved oxygen) and anoxic water (2% dissolved oxygen). Fish were fastened for 48 hrs prior to acclimation for 20h. Oxygen consumption was measured once for each fish for 40 minutes after acclimation, with oxygen levels always above 60%. Chambers were drained and cleaned between runs. Basal metabolic rate was calculated taking into account rate of decrease of oxygen in the chamber, mass of the individual, volume of water and time of measurement (mg O2 g-1 min-1). Averaged background respiration levels across runs was 12.34% (sd=±9.71).

I used ultrasensitive graphene immunosensors (Barton et al. 2018) for measuring waterborne cortisol in a non-destructive way on parents reared in both enriched and poor environments. For this, 12 0ml of water were taken from the respirometer after each individual measurement of metabolic rate and kept at -80 °C until the analysis. A total of 10 ml was centrifuged at 1000 rpm for 5 minutes, and 10 ul of the supernatant were pipetted onto the modified sensor surface. Electrochemical measurements were conducted with a potentiostat/galvanostat (Autolab), controlled with NOVA software as in Barton et al. (2018).

5.3.3. Behavioural analyses

Neophobia was assessed for the same 55 individuals used in the metabolic rate analysis using a plastic test arena (7 cm depth x 7 cm width x 30 cm length) filled with 0.7 L of water. The arena was divided into six equally spaced zones: a covered acclimatisation section (zone 0) with a sliding opaque door, and five open test zones (5 cm each) without cover (zones 1, 2, 3, 4 and 5) delineated by marks at the top margins of the arena walls. A coloured toy block (0.5 cm depth x 0.5 cm width x 3.5 cm height) was glued at the middle of zone 3 to serve as a novel object (Fig. S5.2). Nine-month old fish were placed individually into zone 0 for 15 min acclimatisation, after which the removable gate was slowly lifted, and the fish behaviour recorded for 20 min with an overhead camera fixed 0.5m above the arena. After the test period, tanks were drained, rinsed with ethanol and distilled water and refilled. Videos were analysed by the same person using BORIS v. 7.1.4 (Friard and Gamba 2016) to ensure consistency. The following four behaviours were quantified for both P and F: (1) latency (s) to exit the acclimatisation zone, (2) number of inspections within 3 cm of the novel object (i. e. individual facing towards the novel for more than three seconds), (3) number of contacts with the novel object, and (4) number of movements between zones (activity).

5.3.4. Statistical analyses

All statistical analysis was performed in R v. 3. 4. 3. Cortisol levels and basal metabolic rate were analysed for the P using a linear model with environment (poor vs enriched), and body weight as predictors. I used GLM with a quassipoisson link to account for overdispersion for the P behavioural count data (no. contacts, no. inspections and activity) and a Gaussian link for latency as a function of environment and body weight. One individual from poor environments (1st-PE-R03) had a much higher number of inspections (10) and contacts (14) than the average

(average for inspections: 1.23, s.d = 1.28; average for contacts: 1.38, s.d = 1.77), and was considered an outlier for those behaviours.

To test for parental effects on the F1 phenotypes, I only analysed those phenotypes shown to be significantly different between parental environments, using the same model structure as described above but including also the parental values and environment as predictors. I used the multi-model approach implemented in the R package glmulti v 1.0.7 (Calcagno and de Mazancourt 2010) for model selection, and considered models within 2 AIC units as being equivalent. i used the multi-model approach implemented in the R package glmulti v 1.0.7 (Calcagno and de Mazancourt 2010) for model selection, which tests all possible models and all interactions. To take into account the dependency of observations from individuals originated from the same parents, we selected the best-fit (highest Akaike weight) indicated by glmulti and ran generalized mixed-models including parents as a random factor using the R package mlmRev v.1.0-7 (Bates et al. 2019). Models were tested for overdispersion and individual observation (fish ID) was taken into account when models displayed overdispersion.

5.3.5. Genome-wide DNA methylation data

All individuals were analysed at the same age (10 months old). Fish were euthanized using an overdose of methane-sulfonate (MS-222) following UK Home Office Schedule 1, their brains were dissected and stored in molecular grade ethanol before DNA extraction using Qiagen DNeasy Blood and tissue kit using manufacter' s protocol.

Bisulphite converted genomic DNA libraries were prepared using Diagenode Premium Reduced Representation Bisulphite Sequencing (RRBS) Kit according manufacturer's indications. For the 17 P individuals (11 from enriched, six from poor) were multiplexed into a single library, pooled samples were bisulphite-converted, amplified by enrichment PCR and their quality assessed using Agilent D1000 ScreenTape System. The library was then

sequenced on an Illumina NextSeq 500 platform using a 1x75pb single-end run (Cardiff University, Genomics Research Hub). PCR fully methylated and unmethylated spike controls from the original kit were used to monitor bisulphite conversion efficiency.

A second library was created using fourteen individuals from the second generation (F1, five from enriched to enriched environments, three from enriched to poor, three from poor to poor, and three from poor to enriched). The library followed the same procedures and sequencing conditions as the first library.

5.3.6. Sequence quality and alignment

All analyses were conducted on a local server running NEBC Bio-Linux 8. I assessed the quality of the raw sequences using FastQC (Andrews 2010) and, after trimming of adaptors and low-quality reads was done using the RRBS default parameters (function: --rrbs) in TrimGalore! (Krueger 2016) reads were aligned to the reference genome (*Kryptolebias marmoratus*, NCBI ASM164957v1, Rhee et al. (2017)), trimmed and methylation calls were performed following the same parameters described in Chapter 4. Individuals were grouped into generations (P and F1), environments (own and parental, Table S5.1). Mapped reads were processed and compared using the R package methylKit v. 1. 10. (Akalin et al. 2012). Given its low number of reads into CpG context (16.9% vs mean 54.2%) and low bisulphite conversion efficacy of 70.2% (mean 99.5%), one P individual from enriched analysis was removed from downstream analysis.

5.3.7. Differentially methylated cytosines and methylation patterns

I first assessed differentially methylated cytosines (DMCs) between parental (P) environments (enriched vs poor), using logistic regression on quantitated normalised data with q-value < 0.01 after multiple testing correction (Benjamini-Hochberg) and 20% minimal CpG methylation

difference, using methylKit. To test whether the number of DMCs found in the original grouping is higher or lower than expected by chance, I generated a randomisation test with 4,000 random combinations of 16 parental individuals and tested for the number of DMCs for each combination following the same parameters as the ones described for the original grouping.

I then analysed whether the DNA methylation patterns (hypomethylated or hypermethylated) in the parents were maintained in the offspring. For this, I classified DMCs in two categories (i) environmentally-induced (differences in methylation patterns between the parents changed in the offspring depending on the environment) and (ii) intergenerational (differences in methylation patterns between parental environments were maintained/or changed in the offspring regardless of their rearing environment) (Fig. S5.1). I set up a threshold of ± 10% average methylation change offspring relatively to its parents to consider whether an individual epiallele methylation pattern has maintained the parental methylation state or not. For DMCs classified as intergenerational I identified the genomic location (within gene body, promoter region (≤2 kb upstream of the transcription start site (TSS)), or intergenic region (≥2kb upstream of TSS or downstream the gene bodies).

To test whether the methylation patterns of the offspring on the DMCs classified as potentially intergenerational were influenced by the parental environment, I analysed the methylation score of the offspring for each DMC (as a proportion index, ranging from 0 to 1) as a function of the parental environment (enriched or poor), the offspring environment and their interactions using a generalized linear model with a binomial link. A multiple correction testing was performed to correct for multiple comparisons.

5.4. Results

5.4.1. Parental physiology and exploratory behaviour

For the parental individuals (P), the multi-model selection approach indicated the multi-model selection approach identified two models within 2 ΔAIC, one of which cortisol levels were only affected by body weight (estimate: -3.39; t-value: 4.22, df = 1, p = 0.001) and a second model (ΔAICc = 1.87) included both body weight (estimate = -3.03; t-value: 2.86, df = 1, p < 0.001) and environments (enriched or poor) (t-value: 7.38, df = 1, p = 0.03) as significant factors affecting cortisol levels, with individuals from enriched environments having lower cortisol levels (Fig. 5.1a-b; Table S5.3). Basal metabolic rate decreased with body weight (estimate: -2.40; t-value: -2.35, df = 1, p = 0.01) but was no affected by parental environment (Table S5.3b). A linear regression analysis showed strong correlation between cortisol levels and basal metabolic rate (adjusted R^2 = 0.41; F-value = 16.73, df = 1, p < 0.001) (Fig. 5.1c).

Parental activity significantly decreased with body weight (estimate: -2.88; z-value: 31.58, df = 1, p = 0.01), was lower on individuals reared in poor environments (estimate: -0.25; z-value: 34.98, df = 1, p < 0.001) and varied with environment and body weight interaction (estimate: 2.45; z-value: 38.76, df = 1, p = 0.02) (Fig. 5.1d; Table S5.3). No other equivalent model was found for activity. Neither body weight (estimate: -2.82; z-value: -0.51, df = 1; p = 0.89) and/or the environment (estimate: 0.05; z-value: 0.12; df = 1; p = 0.60) affected number of inspections, nor number of contacts (body weight: estimate: 0.39; z-value: 0.08; df = 1; p = 0.93; environment: (estimate: 0.50; z-value: 1.33; df = 1; p = 0.18). Parental latency to leave acclimatisation zone variation was also not affected by body weight (estimate: 0.50; z-value: 1.33; df = 1; p = 0.18) or the environment (estimate: -1082; z-value: 0.81, df = 1, p = 0.42).

5.4.2. Parental-effect on offspring exploratory behaviour

For the offspring (F1), after correcting for overdispersion indicated that offspring activity was only significantly influenced by parental activity, increasing with increased parental activity (estimate = 5.28, z-value: 2.61, df = 1, p = 0.009) (Table S5.4). No other equivalent model was found.

5.4.3. Differentially methylated cytosines

After quality filtering, approximately 378 million reads were retained, averaging 12.5 million reads per sample. Of those, approximately 61.1% were uniquely mapped reads to the reference genome. Overall bisulphite conversion efficiency was 99.5% (Table S5.1).

In total, I identified 5.5 million cytosine sites, but only 39,205 CpG sites matched the minimum coverage requirements (minimum of 10 reads across all individuals). The majority of the cytosines surveyed mapped gene bodies (71.12%), followed by intergenic regions (12.51%), while only 2.61% were located on putative promoters.

Parental methylation profiles differed in 1854 methylated cytosines (DMCs) between environments. Most of these DMCs were overlapping gene bodies (67.31%), followed by intergenic regions (7.76%) and putative promoters (3.23%). The randomisation test revealed an average of 247.3 (s.d. ±158.5) DMCs across 4,000 combinations of 16 randomly assigned individuals between poor and enriched environments, revealing that the original grouping had approximately 10.3 times more DMCs than expected by chance.

5.4.4. Methylation patterns for parents and offspring

Of the 1854 DMCs identified between parental environments, 724 (39.05%) maintained the same methylation profile in the offspring reared in an environment matching their parent, but

changed in the offspring reared in a mismatching environment and were classified as environmentally-induced epialleles. Of the remaining 1130 DMCs (60.95% of the total), 98 (5.28% of total) maintained the parental methylation patterns in the offspring regardless of their rearing environment, although only five (scaffold: NW_016094248.1, position: 1049469; scaffold: NW_016094269.1., position: 1135514; scaffold: NW_016094316.1, position: 636543; scaffold: NW_016094324.1, position: 879262; scaffold: NW_016094376.1, position: 917192) within less than 10% change in methylation score across all experimental groups (classified as potentially intergenerational) (Table 5.1). When analysed separately by environment-specific context, 30 DMCs in the offspring originated from enriched environment, and 19 in the offspring originated from poor environment maintained their methylation score relatively to its parents regardless of the offspring environment within less than 10% change. The GLM models indicated that the three DMCs which maintained the parental methylation patterns on the offspring were significantly influenced by parental environment (Table 5.2; Fig. 5.3).

5.5. Discussion

The potential transmission of environmentally-induced epigenetic modifications to the offspring could have important implications for evolution (Richards et al. 2017; Verhoeven et al. 2016) but has proven challenging to study in natural populations, due to the confounding effects of genotype-by-environment interactions (Berbel-Filho et al. 2019b; Herman and Sultan 2016) and also to the unequal paternal and maternal contributions to epigenetic states (Soubry et al. 2014). By rearing the self-fertilising mangrove killifish *Kryptolebias marmoratus* under controlled environmental conditions, I identified significant physiological (basal metabolic rate and cortisol levels) behavioural (activity) and epigenetic differences among parents reared between enriched and barren environments, some of which were influenced the offspring phenotypes.

5.5.1. Environmental enrichment influence on physiology and behaviour

Extensive research has highlighted the consequences of the rearing environment on behaviour, physiology and performance in fishes (Jonsson and Jonsson 2014). Most fish exhibit association to physical structures during their life cycle, and consequently structural environmental enrichment has been deliberately used on captive fish as an attenuator of mal-adaptive or aberrant traits (Näslund and Johnsson 2016; Roberts et al. 2011; Roberts et al. 2014). Shelter-like structures (e. g. perforated logs, pipes) have generally been shown to have beneficial effects, such as decreased metabolic rates (Fischer 2000; Millidine et al. 2006) and reduced plasma cortisol levels (Barcellos et al. 2009; Näslund et al. 2013), particularly in aggressive species (Näslund et al. 2013). My results indicate that parental fish reared on enriched environments have lower basal metabolic rates and waterborne cortisol levels. They also show a negative relationship between cortisol levels (and potentially basal metabolic rate) and body mass, while metabolic rate was not related to the rearing environment of the parents, the tight correlation between waterborne cortisol levels and basal metabolic rates in parental fish, suggests that parental individuals reared in enriched environments were less stressed and spend less energy to maintain basal metabolic rate than individuals reared in barren environments.

In fish, structural environmental enrichment tends to decrease activity, mainly due to increased sheltering (Moberg et al. 2011; Roberts et al. 2011; von Krogh et al. 2010), and exploratory activity and boldness tend to be positively correlated (Champneys et al. 2018; Mazué et al. 2015). Here, parents reared on enriched environments were slightly more active than individuals reared in poor environments, but no significant differences were found between environments on neophobic behaviours (inspections and contacts), suggesting no clear boldness-exploratory relationship in response to environmental enrichment in *K. marmorat*us Edenbrow and Croft (2013) found that exploratory behaviour activity and boldness correlated

positively in response to conspecific presence and simulated predation risk in *K. marmoratus*, however these changes were not repeatable and were not necessarily correlated to aggression levels, suggesting plastic behavioural responses during the ontogeny of this species.

Although behavioural effects of environmental enrichment on fishes are well known (Jonsson and Jonsson 2014; Näslund and Johnsson 2016), the understanding of their potential inter- or transgenerational effects is scarce. My results indicate that, offspring activity was influenced by the parental activity. In general, offspring from parents reared in enriched environments had higher activity levels, regardless of their own environment, suggesting a sustained parental effect on activity levels. Previous studies in this species found that life-history traits (offspring size) but not behaviour (exploration and aggression) were affected by the parental environment (Edenbrow and Croft 2013). However, in mammals there is ample evidence of parental effects caused by environmental enrichment, where the offspring from enriched environments tend to be more exploratory (Dell and Rose 1987; Mychasiuk et al. 2012), have increased learning capacity and memory formation (Bygren 2013), than those reared in non-enriched environments. While in fish, increased cognitive capacity due to environmental enrichment has already been shown (Roberts et al. 2011; Salvanes et al. 2013), however this is the first evidence of behavioural intergenerational (parental) effects.

5.5.2. Environmental enrichment effect on DNA methylation

My results revealed strong effect of environmental enrichment on brain DNA methylation patterns, with 1584 differentially methylated cytosines (DMCs) (neighbouring or on gene bodies of 728 unique genes) differentiating genetically uniform parents reared on enriched and poor environments. Several studies have reported effects of environmental enrichment on brain growth (Näslund et al. 2012), cell proliferation (von Krogh et al. 2010), cognitive capacity (Salvanes et al. 2013; Spence et al. 2011), and gene expression levels (Evans et al. 2015) in

fishes. Therefore, the differential methylation patterns between rearing environments found here could be related to differential transcription levels (or noise reduction, depending of the genetic context) (Evans et al. 2015) in pathways of the brain development that could be the basis of the physiological and behavioural differences observed between enriched and poor reared individuals.

Due to epigenetic reprogramming during embryogenesis, only a small subset of epigenetic variants on the parents are likely to be transmitted to the offspring (Burggren 2016; Illum et al. 2018). DNA methylation changes during embryogenesis in *K. marmoratus* has a longer and later DNA methylation reprogramming period when compared to other fish and mammals (Fellous et al. 2018), which might represent an epigenetic window for environmental sensibility. In the offspring, most of my results revealed a stronger effect of their own rearing environment than that of their parents on cytosine methylation patterns. Thus, although there were clear effects of environmental enrichment on the brain cytosine methylation patterns of the parents, these changes may not have happened to the same extent in the germline, suggesting limited potential for cytosine methylation-mediated parental effects being transmitted trans-generationally and/or scape epigenetic reprogramming. I cannot discard that the low sample size for some offspring groups together with the somewhat conservative depth requirements (≥10 reads in all individuals) might have contributed to miss some potential epigenetic variants which contributed for parental effects. Yet, three DMCs maintained the same methylation patterns in both parents and offspring (within 10% methylation score) while additional DMCs maintained the parental methylation patterns in the offspring in a more environment-specific manner. To my knowledge, although limited, this is the first evidence of parental effects on the offspring cytosine methylation patterns caused by environmental enrichment in fish, extending previous results in mice, which demonstrated that parental enrichment can affect offspring brain weight, global methylation levels (Mychasiuk et al. 2012)

and learning capacity (Arai and Feig 2011). It is important to highlight here, that although changes in methylation status at particular genomic regions (i. e. promoters, gene bodies, first intron, transposons) are known to have functional effects (i.e. changes on gene expression, transcriptional noise, transcription factor binding) (Anastasiadi et al. 2018; Horvath et al. 2019; Huh et al. 2013; Morgan et al. 1999), the effects of changes in methylation at the single cytosine level, if any, are poorly known, particularly in animals (Gutierrez-Arcelus et al. 2013).

In summary, my results reveal behavioural parental effects in the offspring caused by parental environmental enrichment in the self-fertilising mangrove killifish *Kryptolebias marmoratus*. Although I found that some DMCs significantly maintained their methylation status in the offspring, their limited numbers indicated that cytosine methylation is perhaps not the best candidate for epigenetically-mediated parental effects. Still, while most studies on fish focus on behavioural and physiological consequences of environmental enrichment within one generation (Näslund and Johnsson 2016), my results provide evidence that environmental enrichment not only affect parents but is also transmitted to the offspring in a specific manner, however its epigenetic mechanisms still remains to be explored.

Table 5.1. Average methylation score differences (for parents) and for the offspring for methylated cytosines (DMCs) which maintained o methylation patterns (hyper or hypomethylated) relatively to the parental patterns (within ±10% change) Positive and negative values on methylation differences represent hyper or hypomethylation in relation to the enriched environment. "G", and "U" refer to gene bodies and unnanotated regions, respectively. Q-value is the p-value adjusted for the False Discovery Rate (FDR=0.05). Asterisks represent DMCs significantly affected by parental environment. Asterisks represent DMCs which were significantly influenced by parental environment.

Scaffold	**Position**	**Genomic context**	**Meth diff 1st (E-P)**	***Q*-value**	**Meth mean E>E**	**Meth mean P>P**	**Meth mean E>P**	**Meth mean P>E**
NW_016094248.1	1049469	G	20.79	<0.001	94.89	84.46	91.57	86.59
NW_016094269.1	1135514	G	-22.81	<0.001	78.56	82.11	72.52	84.14
NW_016094316.1*	636543	U	-22.17	<0.001	29.96	47.85	32.13	48.86
NW_016094324.1*	879262	G	26.35	<0.001	84.73	59.00	72.38	60.77
NW_016094376.1*	917192	G	-20.81	<0.001	22.36	34.30	29.18	38.27

Table 5.2. Results of most plausible binomial generalised models obtained by multi-model averaging approach for offspring methylation scores for the DMCs which maintained the parental methylation score (within 10% difference) across all offspring groups. Models are ranked according to the corrected Akaike Information Criterion (AICc), the difference with the most plausible fitting model (ΔAICc), and the Akaike weight (Wi), which represents the ratio between the weights of the best and competing models. Only models within two AICc units are shown.

	Df	**t-value**	***P*-value**	**AICc**	**Wi**
NW_016094316.1.636543					
Parental environment	1	2.70	0.002	-6.81	0.81
NW_016094324.1.879262					
Parental environment	1	-1.83	0.01	2.66	0.68
NW_016094376.1. 917192					
Parental environment	1	1.99	0.01	-10.78	0.72

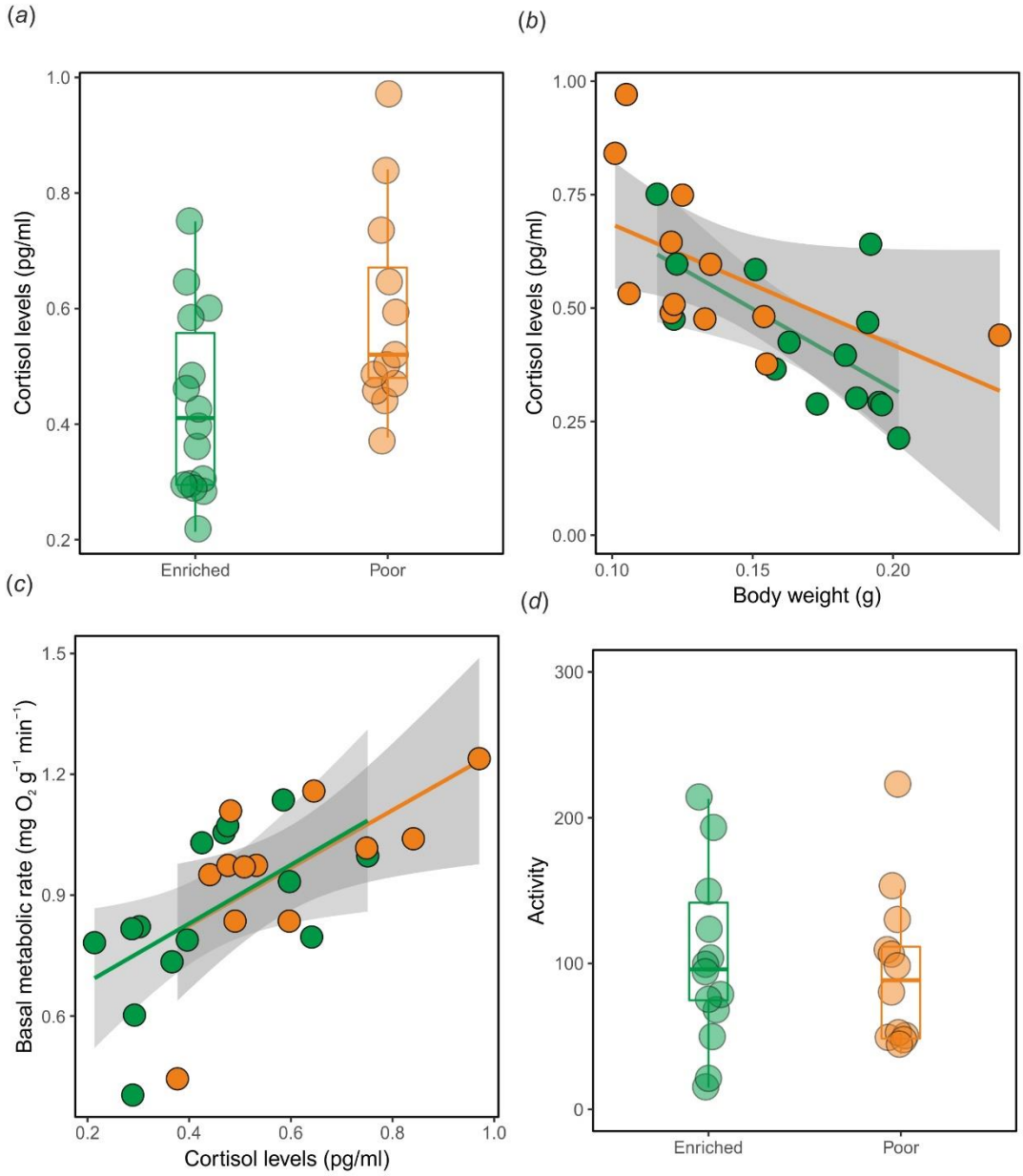


Figure 5.1. Physiological and behavioural relationships between (*a*) cortisol levels, and environments (enriched, poor) and (*b*) cortisol levels and body weight (g); (*c*) basal metabolic rate and cortisol levels; (*d*) activity (number of crosses among zones) and environments for parental individuals. Green and orange represent parents reared in enriched and poor environments, respectively.

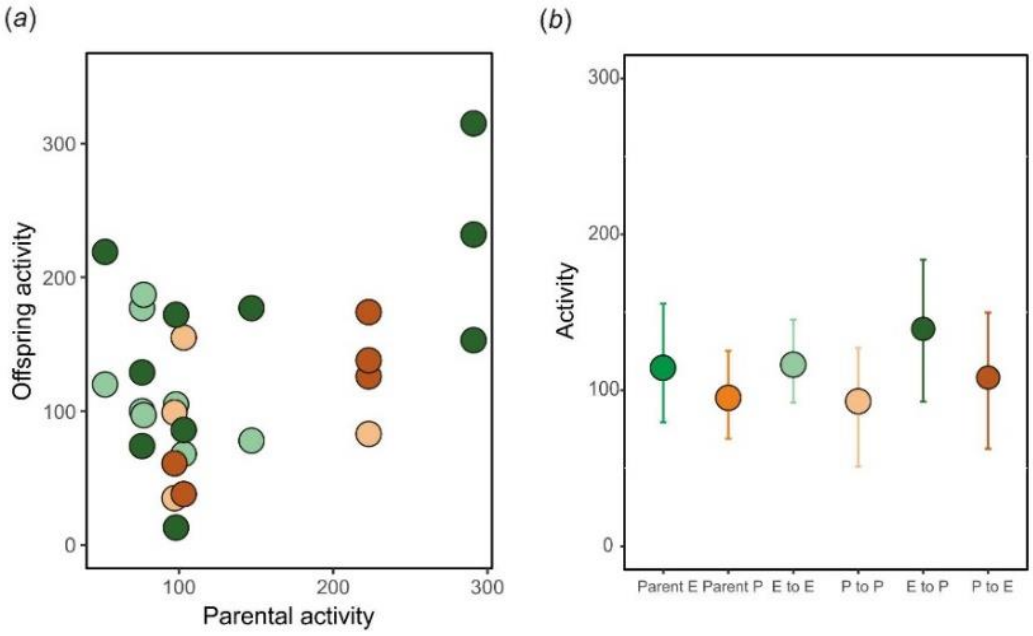


Figure 5.2. Relationships for behavioural metrics in the offspring. (*a*) Offspring activity (number of crosses among zones) and parental activity; (*b*) activity across experimental groups; Light green (parent reared on enriched) and orange (parent reared on poor) represent offspring which matched environment relative to their parents (green enriched, orange poor). Dark green (enriched) and dark orange (poor) represent individuals which mismatched environment relative to their parents. Dashed lines in (*a*) and solid coloured lines in (*c*) represent parental and offspring environments, respectively.

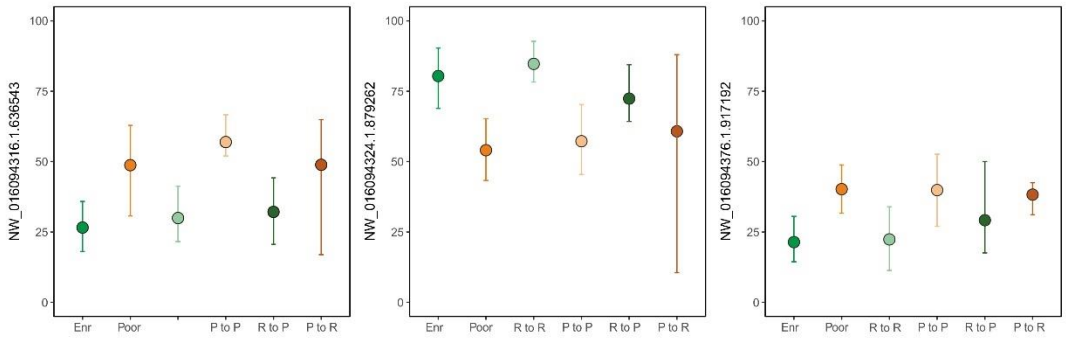


Figure 5.3. Average methylation scores for DMCs which maintained parental methylation patterns and score (within 10% change) across all offspring groups. Light-coloured circles indicate the mean of individuals which mismatched environment relative to their parental environment. And dark-coloured circles indicate means of individuals which matched the environment of their parents.

III. GENERAL DISCUSSION

Mixed-mating organisms, with their intermediate levels of selfing and outcrossing, represent a uniquely suitable opportunity to study the evolutionary implications of epigenetic variation. Its populations are composed by multiple divergent selfing isogenic lineages, allowing to test the effect of common environmental pressures on multiple epigenomes and the extent of its heritability, as well as the evolutionary implications of genetically-independent epialleles. The broad aim of this thesis was to evaluate the ecological and evolutionary implications of epigenetic variation in animals, using the unique diversity of mating systems of the killifish genus *Kryptolebias*. The results of the studies performed here have allowed to address several key questions about the relative contribution of genetic and epigenetic variation to the ecology and evolution of *Kryptolebias* species, and more broadly how those interact and respond to environmental pressures.

Chapter 1 aimed to assess the level of genetic and epigenetic structure in natural populations of *Kryptolebias hermaphroditus* as well as the potential factors (parasites, genotype background, heterozygosity) that shape its DNA methylation variation. It was confirmed that *K. hermaphroditus* is able to outcross, making it the second example of mixed-mating organism in vertebrates (together with its sister-species, *K. marmoratus*). Individual levels of genetic diversity (measured with microsatellites markers) correlated with parasite loads. Although outcrossing was confirmed, selfing seems to be the major mode of reproduction in *K. hermaphroditus*, which could be related to the low extent of parasite pressures. Inter and intra-mangroves genetic structure was found, even at small spatial scales. Taking into account all the selfing linages and sampling sites, DNA methylation levels majorly correlated with the genetic background, and to a lower extent with parasite loads. However, at the local scale, where selfing lineages are genetically more similar, DNA methylation was

affected by parasite loads, and their interactions with the inbreeding status (selfed or outcrossed). These findings suggest an intricate relationship between genotypic diversity and DNA methylation, with the genetic background having a stronger effect in differentiating DNA methylation profiles than individual heterozygosity. Therefore, the combined genomic and epigenomic responses to similar ecological pressures (Chapter 4), and its potential inheritance (Chapters 3 and 5) could speed up local adaptation in natural population even at small spatial scale, which can be particularly relevant to mixed-mating organisms, given the limited capacity of selfing lines for genetically-based adaptation. Future studies focusing on specific parasite effects together with high resolution epigenetic markers would help to clarify the functional effects of DNA methylation variation in natural populations, as well as to which extent different genotypic backgrounds converge and/or diverge in DNA methylation in response to parasite pressures.

To understand the effects of DNA methylation inheritance in natural populations, it is necessary first to assess the influence of the levels of genetic variation both within and between populations. *Kryptolebias ocellatus* is a mangrove killifish endemic to mangrove forests in southeast Brazil, for which mating-system, geographical distribution, population structure and interactions with the often sympatric selfing *K. hermaphroditus* were poorly known. In **Chapter 2**, I confirmed the outcrossing-only androdioecious nature of *K. ocellatus* and found deep genetic structure (using mtDNA, microsatellites and SNPs) across its whole known distribution. This finding contrasts with the extreme genetic homogeneity observed in *K. hermaphroditus* along the same area, corroborating the scenario of *K. hermaphroditus* as recent colonizer of a mangrove area (Tatarenkov et al. 2011; Tatarenkov et al. 2017) where *K. ocellatus* was already inhabiting. Despite the genetic and reproductive differences, I found evidence

of unlikely hybridisation and introgression events between the two species in two of the three mangrove areas where both species inhabit syntopically. This is the first case of hybridisation between species with different mating-systems in vertebrates. The ecological causes and evolutionary consequences of this unlikely hybridisation are still unknown, but the fact that the hybridisation seems to be ongoing in heavily polluted mangroves (Guaratiba and Fundão) and not in more pristine areas (Iriri) together with previous evidence of hybridisation-mediated pollution resistance in killifishes (Oziolor et al. 2019) suggest that this hybridisation may be mediating adaptation to environmental stress in those populations, although this warrants further research.

Taking advantage of the natural hybridisation between *K. ocellatus* and *K. hermaphroditus*, I evaluated the patterns of DNA methylation inheritance between parental species and hybrids in **Chapter 3.** DNA methylation patterns between parental species, even if under similar environmental pressures, was extensively different, reinforcing the significant effect of genetic background on DNA methylation profiles, and suggesting that environmentally-induced epialleles may be limited. Hybrids showed a predominant intermediate pattern of DNA methylation relatively to the parental species. The epialleles which were differentially methylated between hybrids and parental species affected genes important to several key developmental process, on which misregulation could have substantial phenotypic effects, or alternatively, create phenotypic novelty that ultimately can increase hybrid vigour and contribute to hybrid speciation. Backcrosses were more similar to the parental species than to F1 hybrids, reinforcing the evidence for an important effect of DNA background on DNA methylation patterns. Future studies covering a bigger area within this hybrid zone, investigating the environmental differences driving hybridisation as well as comparing the gene expression patterns together with fitness measures (e.g.

size, reproductive output, survival) will help to evaluate the long-term consequences of the ongoing hybridisation between *K. ocellatus* and *K. hermaphroditus*.

Given the tight association between genetic variants and DNA methylation epialleles, one of the most challenging tasks in ecological epigenetics is identifying the extent and potential evolutionary role genetically-independent epialleles. Using genetically-identical individuals of *K. marmoratus* from two different genetic backgrounds under constating environmental conditions (e. g. environmental enrichment), I evaluated the effect of environmental variation on the distribution of DNA methylation epialleles in **Chapter 4**. DNA methylation differences were more prominent between genotypes than environments (genetically-dependent). In the few epialleles which were commonly affected in both genotypes, most of them showed a direction of variation between environments which were dependent of the underlying genotype (genotype-by-environment interactions). These findings indicate that DNA methylation epialleles purely driven by environmental variation may be rare, and DNA methylation, as most of the other plastic traits, responds in a genotype-specific manner to environmental variation. Further studies, preferably using more genetic backgrounds and a more complete methylome coverage, are needed to analyse the potential downstream phenotypic consequences (i. e. gene expression, metabolic networks) of genotype-specific DNA methylation responses to common environmental pressures, and to provide a wider picture of the molecular mechanism behind the ubiquity of genotype-by-environment plasticity.

At the long-term, the evolutionary relevance of epigenetically-mediated phenotypic variation depends on the extent of its heritability. From the results of Chapter 4 it was evident that environmental enrichment affects brain DNA methylation patterns, so **Chapter 5** explored whether these epigenetic changes affected other phenotypes

(behaviour, physiology) and if those changes could be transmitted to the following generation though DNA methylation. I found evidence of environmental enrichment driving behavioural, physiological and DNA methylation differences in genetically-identical *K. marmoratus* individuals. Parental effects were found in activity levels of the offspring. Some putative transgenerational and intergenerational epialleles in the offspring were significantly influenced by parental DNA methylation patterns. These findings suggest that the offspring phenotypes can be influenced by the parental environment, although the role of cytosine methylation on its transmission may be limited, with other potential epigenetic mechanisms having higher chance to scape epigenetic reprogramming, such as microRNAs (Chen et al. 2016). being primary candidates for epigenetic inheritance. Further research investigating parental effects including other epigenetic mechanisms, their effects on gene expression as well as isolating inter from transgenerational effects would help to clarify the role of environmentally-induced epialleles on paternal effects.

A crucial reason for the increased interest on epigenetics in evolutionary biology resides in its potential for generating heritable phenotypes, providing an alternative mechanism of heritability residing outside the classic model of DNA-based inheritance (Bonduriansky and Day 2009). Although strong evidence for epigenetically-mediated phenotypes and heritable transmission have already been provided (Chen et al. 2016; Cubas et al. 1999; Morgan et al. 1999; Wolff et al. 1998), research in ecological epigenetics has only recently being expanded beyond the scope of cell lines or model-organisms in lab conditions (Richards et al. 2017; Verhoeven et al. 2016). This expansion has brought back important questions about the evolutionary implications of epigenetic variation, and the evidence that genetic features (e. g. genotype, heterozygosity) have a major influence on DNA methylation variability, particularly

in plants (Dubin et al. 2015), revived a crucial question regarding epigenetic variation: to which extent genetic background influence DNA methylation variation?

Genetic variants influence DNA methylation levels either directly, by changing a cytosine to a "non-methylable" nucleotide, or passively by altering transcription factors binding (Gutierrez-Arcelus et al. 2013; Stadler et al. 2011). DNA methylation levels can in turn, modify genetic features by regulating the movement of transposable elements (Richards et al. 2017). In this thesis, I found direct evidence of a prominent effect of genetic background in DNA methylation variation, both in natural (Chapters 1 and 3) and laboratory conditions (Chapter 4). These findings bring into reconsideration the role of DNA methylation as a primary epigenetic candidate to respond to environmental fluctuation independently of the genetic background. Given the evidence that differences between genotypes have overriding effect on DNA methylation levels driven by environmental pressures, ecological epigenetic studies should include the genetic background as an essential factor while evaluating DNA methylation variation, a conclusion that, while already highlighted in plants (Dubin et al. 2015; Herman and Sultan 2016), cell-lines and/or model organisms (Gutierrez-Arcelus et al. 2013; Shea et al. 2015; Teh et al. 2014), had been seldomly explored in non-model organisms in natural conditions before this study (Leung et al. 2016; Richards et al. 2017). Depending on the effect of DNA methylation transcriptional activity and its transgenerational capacity, these genotype-specific DNA methylation responses can further extend phenotypic differences in locally-divergent genotypes, which in the long term could represent preliminary steps for population differentiation, and ultimately, speciation (Smith et al. 2016).

In Chapter 3, I demonstrated intermediate, and potentially additive, patterns of DNA methylation when two divergent genomes are brought together by hybridisation

(Chapter 2), reinforcing an effect of genomic features driving methylation patterns under similar environmental conditions. The limited number of putative transgenerational epialleles identified in Chapter 5, corroborates the recent findings indicating that DNA methylation perhaps is not the primary epigenetic candidate for carrying parental environmentally-induced changes (Radford et al. 2014; Shea et al. 2015), possibly caused by limited scape of germline *de novo* DNA methylation patterns during epigenetic reprogramming (Iqbal et al. 2015). Although DNA methylation is a key regulatory factor for several important biological process (Jones 2012), the transmission of its genetically-independent epialleles may be limited, with other epigenetic markers, such as microRNAs, emerging as primary candidates on epigenetically-transmitted transgenerational effects (Chen et al. 2016; Posner et al. 2019).

IV. CONCLUSIONS

This thesis has provided direct empirical evidence for the determinants of cytosine methylation variation, taking advantage of the diversity of mating systems observed in the killifish genus *Kryptolebias*. Following are the main conclusions that can be derived from the findings of this thesis:

1. Confirmation of the mixed-mating (predominantly selfing with occasional outcrossing) status *K. hermaphroditus* and the outcrossing-only androdioecy in *K. ocellatus*.

2. Deep genetic structure at both small and large spatial scales in *K. hermaphroditus* and *K. ocellatus*, respectively, which may have important implications for conservation policies for those mangrove killifish species.

3. Hybridisation and introgression between *K. hermaphroditus* and *K. ocellatus* in natural populations in southeast Brazil, representing the first case of hybridisation between species with different mating systems in vertebrates. Additionally, this system can provide a natural model for the ecological causes and evolutionary consequences of hybridisation in human-altered environments.

4. Behavioural parental effects caused by environmental enrichment are likely to be transmitted to the offspring by other epigenetic mechanisms than cytosine methylation.

5. Cytosine methylation variation was shown to be largely influenced by genetic features, suggesting that its implications in evolutionary biology should evaluated into a genotypic-specific framework. Further studies should particularly focus on whether the effects genotype-specific cytosine methylation responses converge or diverge in downstream phenotypes, the extent of its heritability and its contribution to individual fitness and long-term evolution.

V. APPENDIX I: SUPPLEMENTARY FIGURES AND TABLES

IV.I. Figures

Supplementary figures for Chapter 1 - Genetic background and parasite load affect DNA methylation variation in a predominantly-self fertilising fish

Figure S1.1. Parasites of *Kryptolebias hermaphroditus* from northeast Brazil: (*a*) bacterial gill cyst; (*b*) protozoan gill cysts; (*c*) nematodes from the gut.

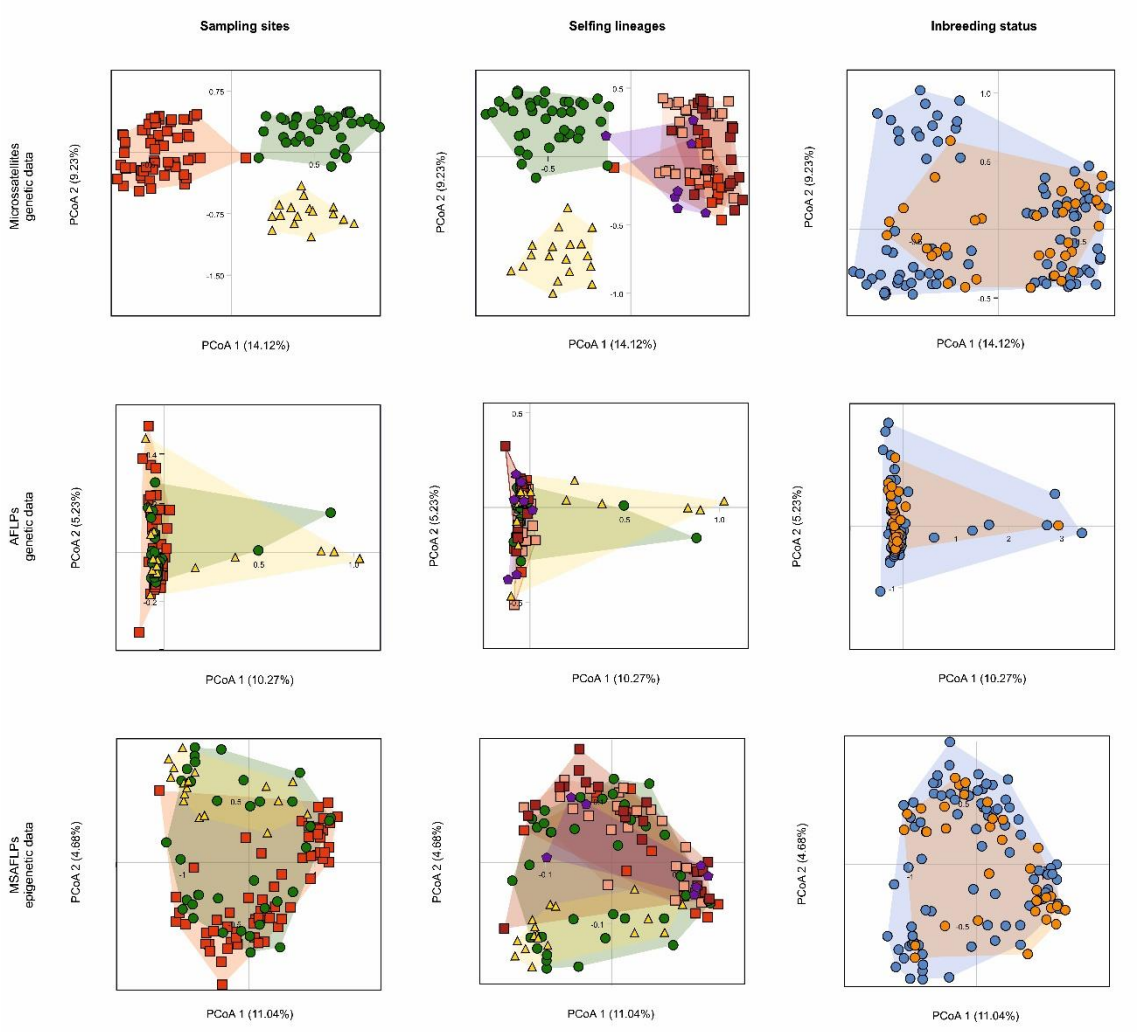


Figure. S1.2. Principal coordinate analysis (PCoA) of genetic (microsatellites and AFLPs) and epigenetic (MSL) variation for sampling sites, selfing lineages identified by INSTRUCT and inbreeding status (selfed or outcrossed). Squares, circles and triangles represent Site 1, Site 2 and Site 3, respectively. Lineages are coloured in red (1), salmon (2), green (3), brown (4) and yellow (5). Lineage 6 (shared between Sites 1 and 2) was represented by purple pentagons. Blue and orange circles represented inbred and outcrossed individuals, respectively.

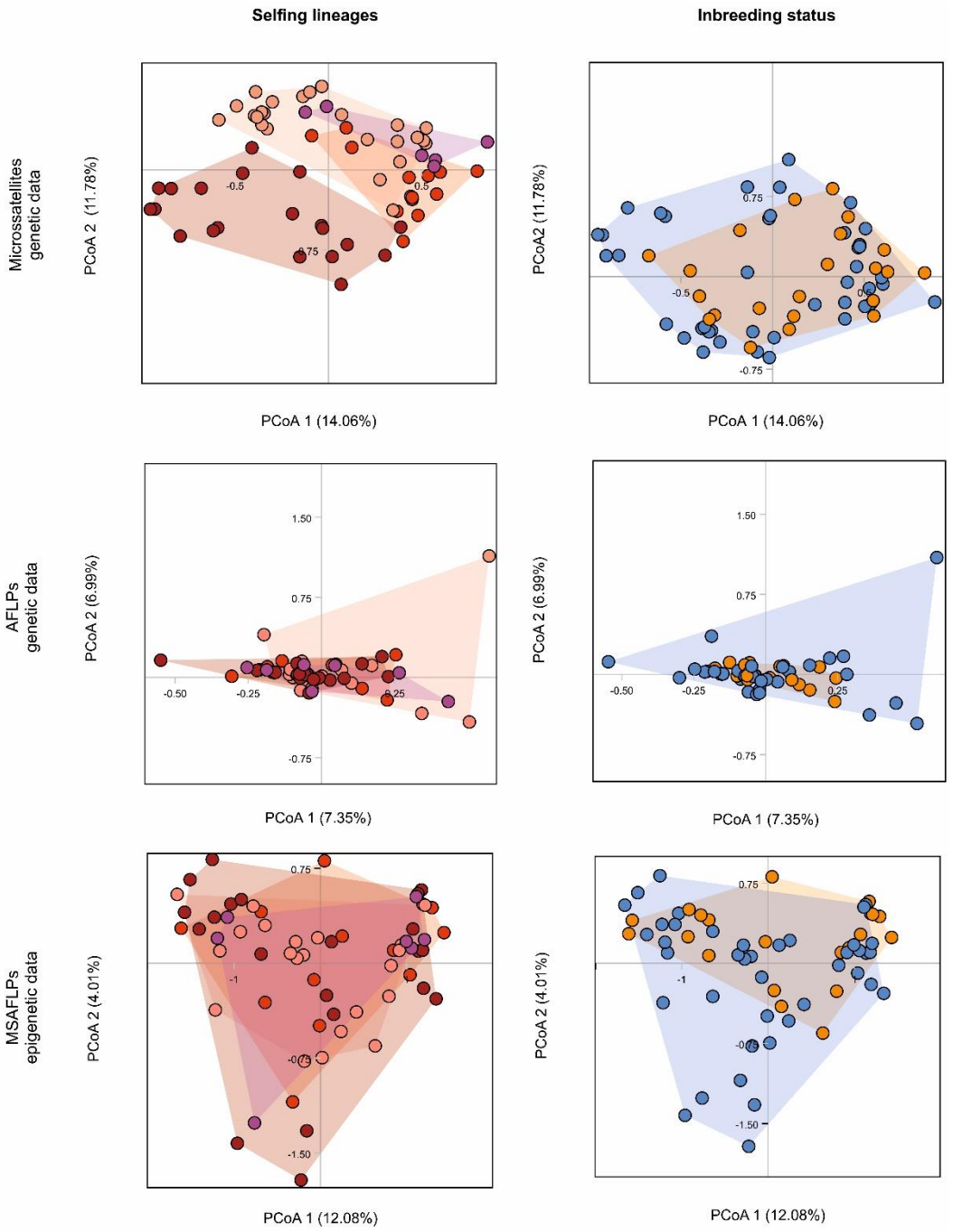


Figure S1.3. Principal coordinate analysis (PCoA) of genetic (microsatellites and AFLPs) and epigenetic (MSL) variation for selfing lineages identified by INSTRUCT and inbreeding status (selfed or outcrossed) for individuals in Site 1. Filled circles with different colours represent the different selfing lineages from Site 1 (red for lineage 1, salmon for lineage 2, brown for lineage 4, purple for lineage 6). Blue and orange circles represent inbred and outcrossed individuals, respectively.

Supplementary figures for Chapter 2 - Hybridisation between species with different mating systems revealed by the genetic structure of the mangrove killifish *Kryptolebias ocellatus*

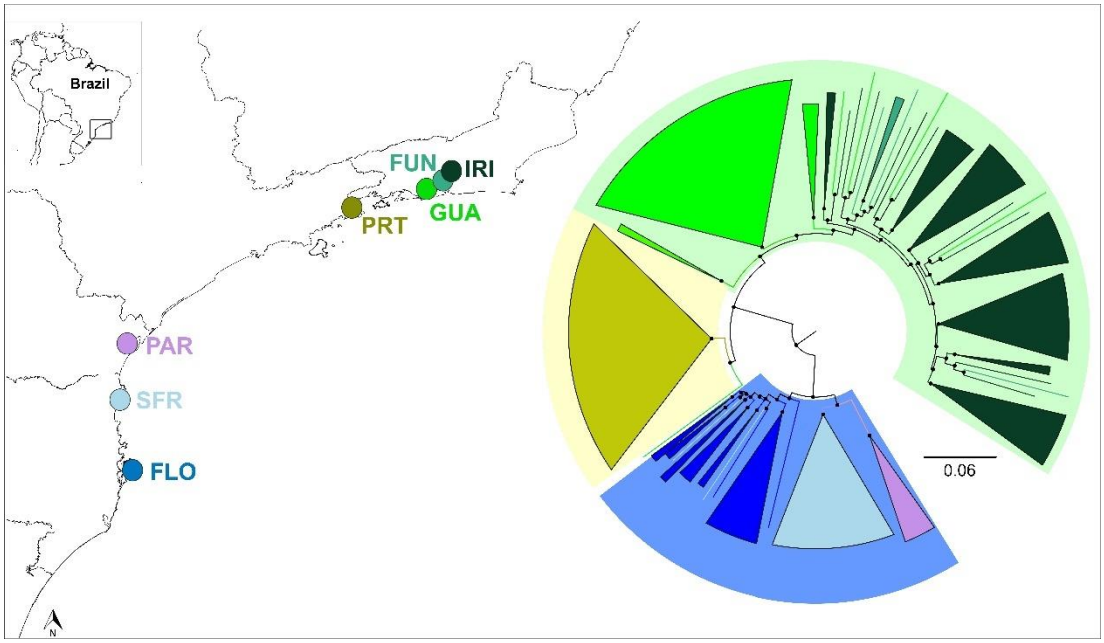


Figure S2.1. Neighbour-Joining tree based on the D_{PS} distances for the 179 *K. ocellatus* (excluding the hybrids) individuals amplified for 16 microsatellites loci and coloured according individuals sampling location and major clades.

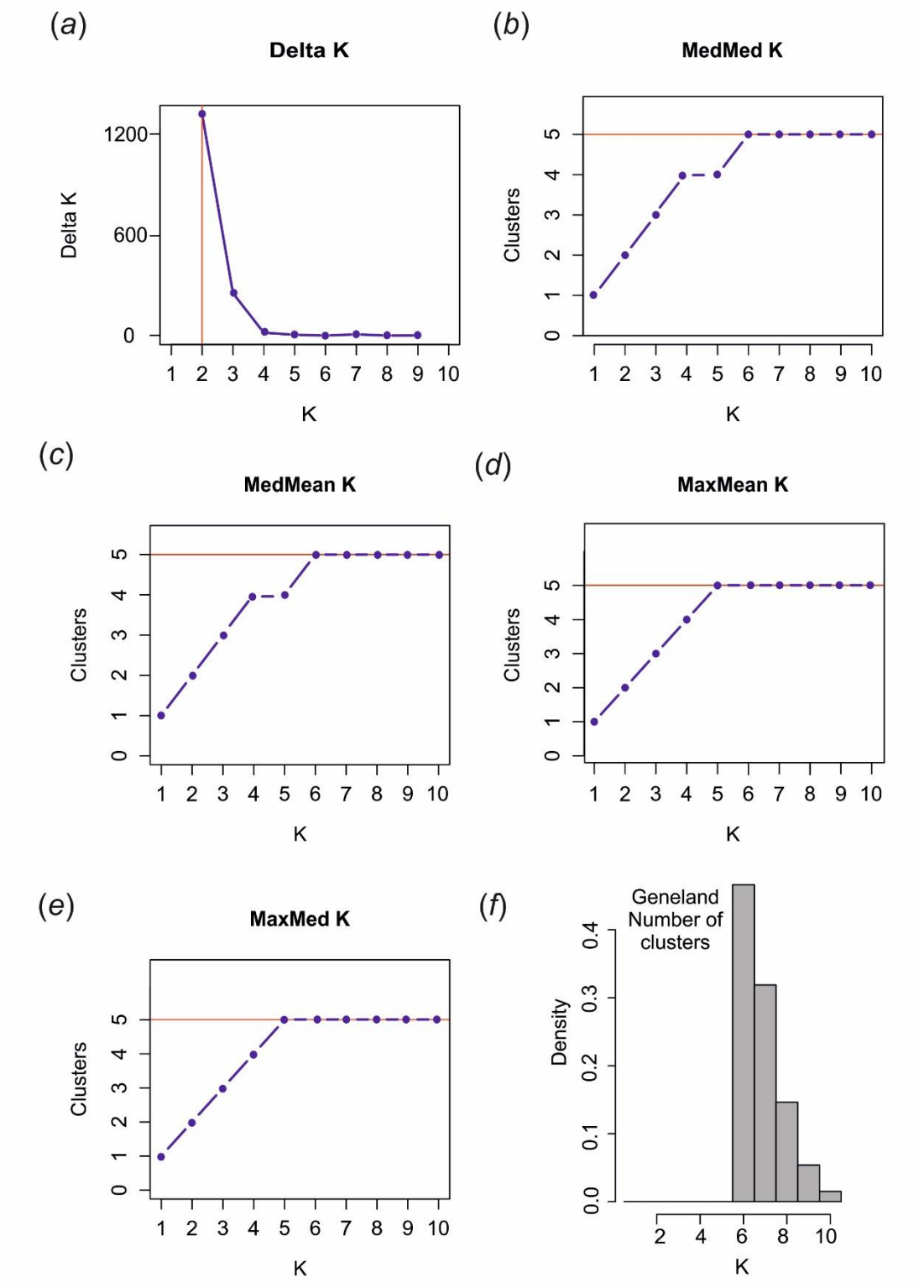


Figure S2.2. Estimated number of genetic clusters (K) from different methods. (*a*) deltaK method in Evanno et al. (2005); (*b*)median of medians (MedMedK), medians of (*c*) means (MedMeanK); (*d*)maximum of medians (MaxMedK) and (*e*) maximum of the means (MaxMeaK) as implemented by Puechmaille (2016) to account for unevenness on sampling sizes and hierarchical structure; (*f*) Posterior density distribution of the number of clusters estimated from Geneland.

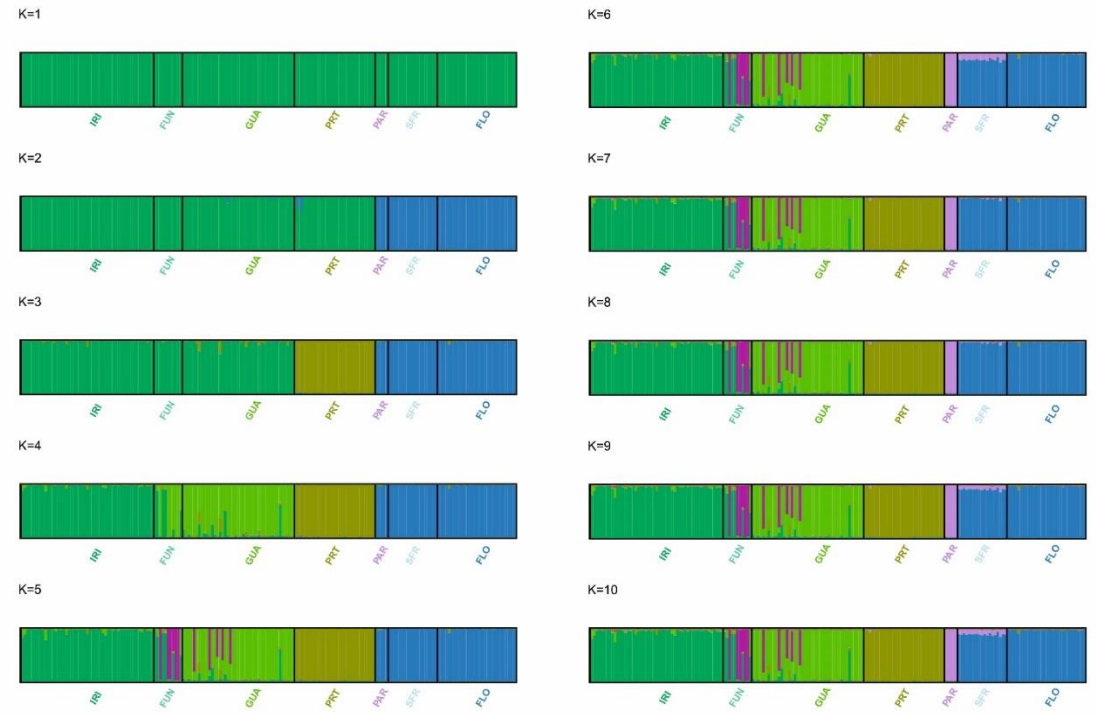


Figure S2.3. Panel for the lowest likelihood run for each K value ran in STRUCTURE with genotypes for 16 microsatellites for all 190 individuals with *K. ocellatus cox1* haplotypes.

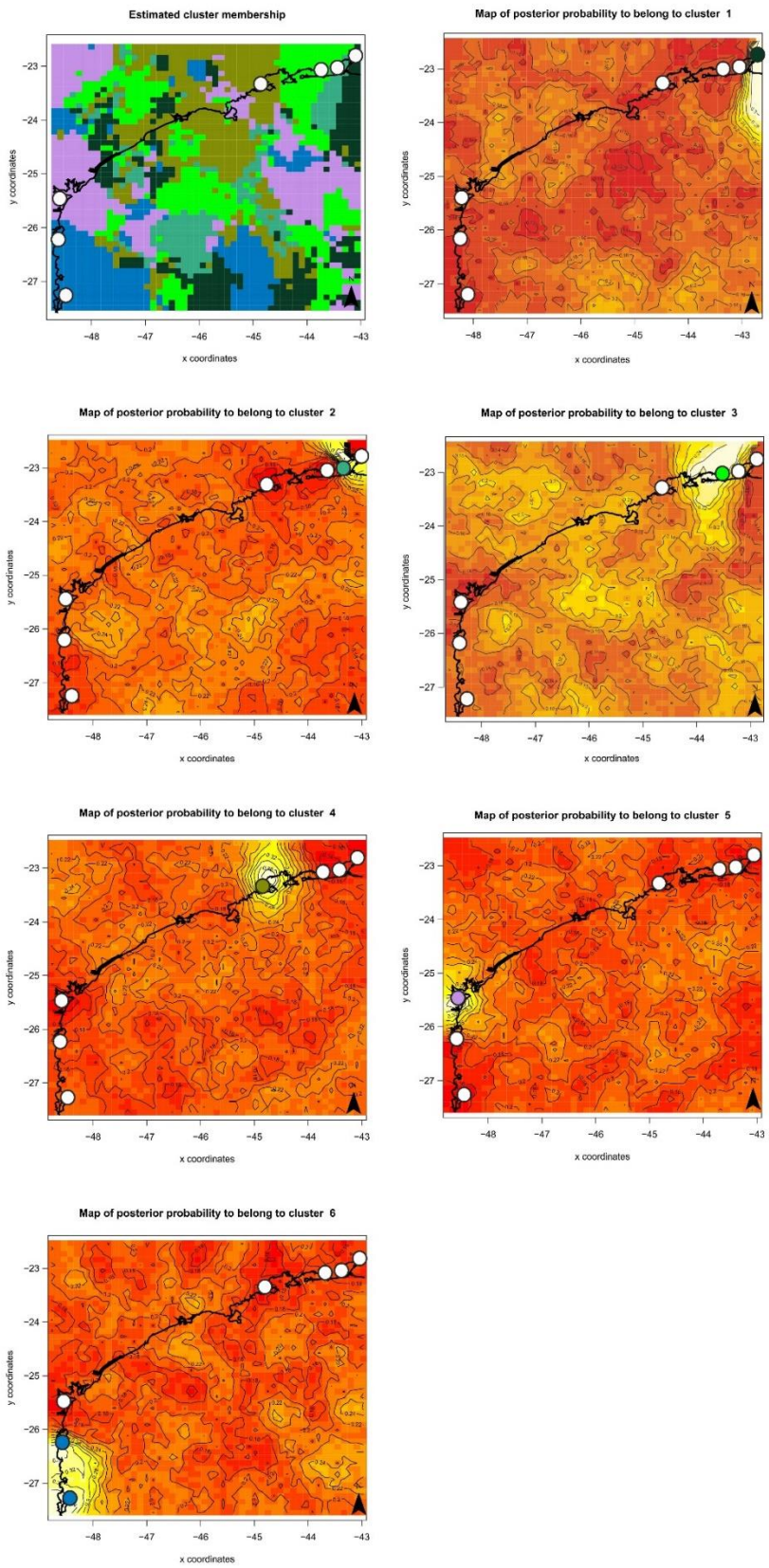


Figure S2.4. Maps of Geneland individual assignment to clusters for K = 6 with geographical coordinates. Different colours represent the different genetic clusters, while the highest membership values are in light yellow at the background.

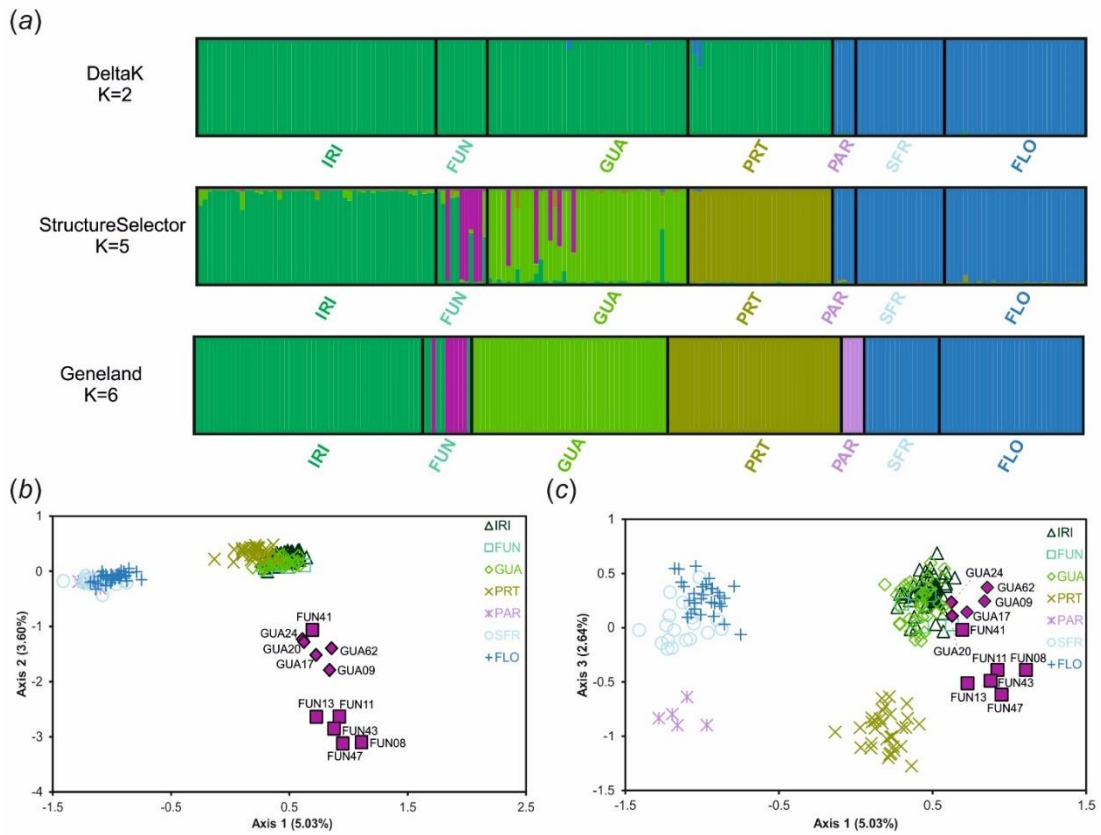


Figure S2.5. (*a*) Panels showing the most likely genetic clusters (K) value for the 16 microsatellites amplified in *Kryptolebias ocellatus* ran in Structure and determined by deltaK method in Evanno et al. (2005) and STRUCTURESELECTOR results using the metrics defined in Puechmaille (2016) to account for unevenness on sampling sizes and hierarchical structure. Geneland was also run to include spatial and mtDNA information. Each individual is represented by a bar, and each colour represents a genetic cluster. (*b*-*c*) Factorial correspondence analysis for all *K. ocellatus* individuals coloured and shaped according to its sampling sites. Hybrid individuals (see results) are highlighted with filled purple symbols and their respective labels.

Supplementary figures for Chapter 3 - Intermediate patterns of epigenetic variation in hybrids of divergent mangrove killifish species in natural populations

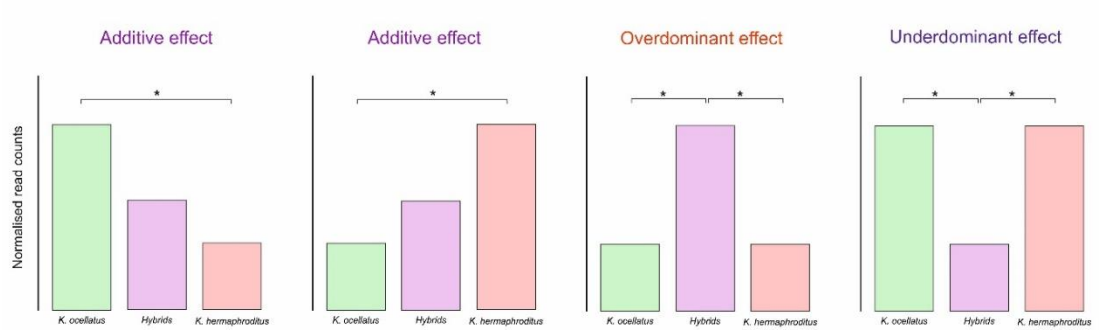


Figure S3.1. Schematic representation ot the classification of differentially methylated cytosines in either potentially additive, orverdominant or underdominant in the hybrids compared to the parental species. Asterisks represent signifcant differential methylation between groups.

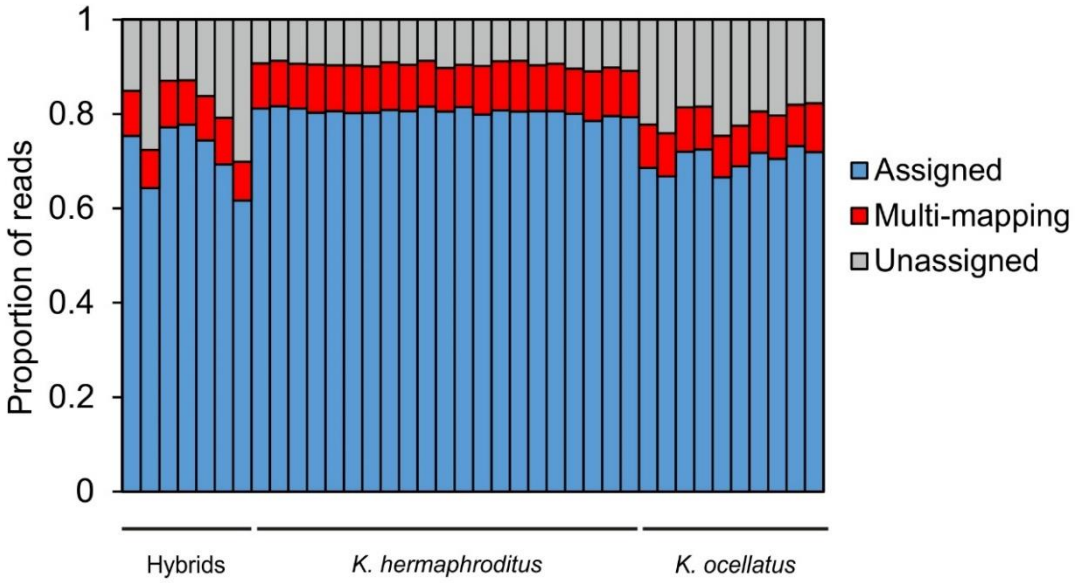


Figure S3.2. Proportion of reads uniquely mapped to features (blue), unassigned due to multi-mapping (red), and unassigned (grey) due to no match with reference genome feature using Bowtie 2 aligner.

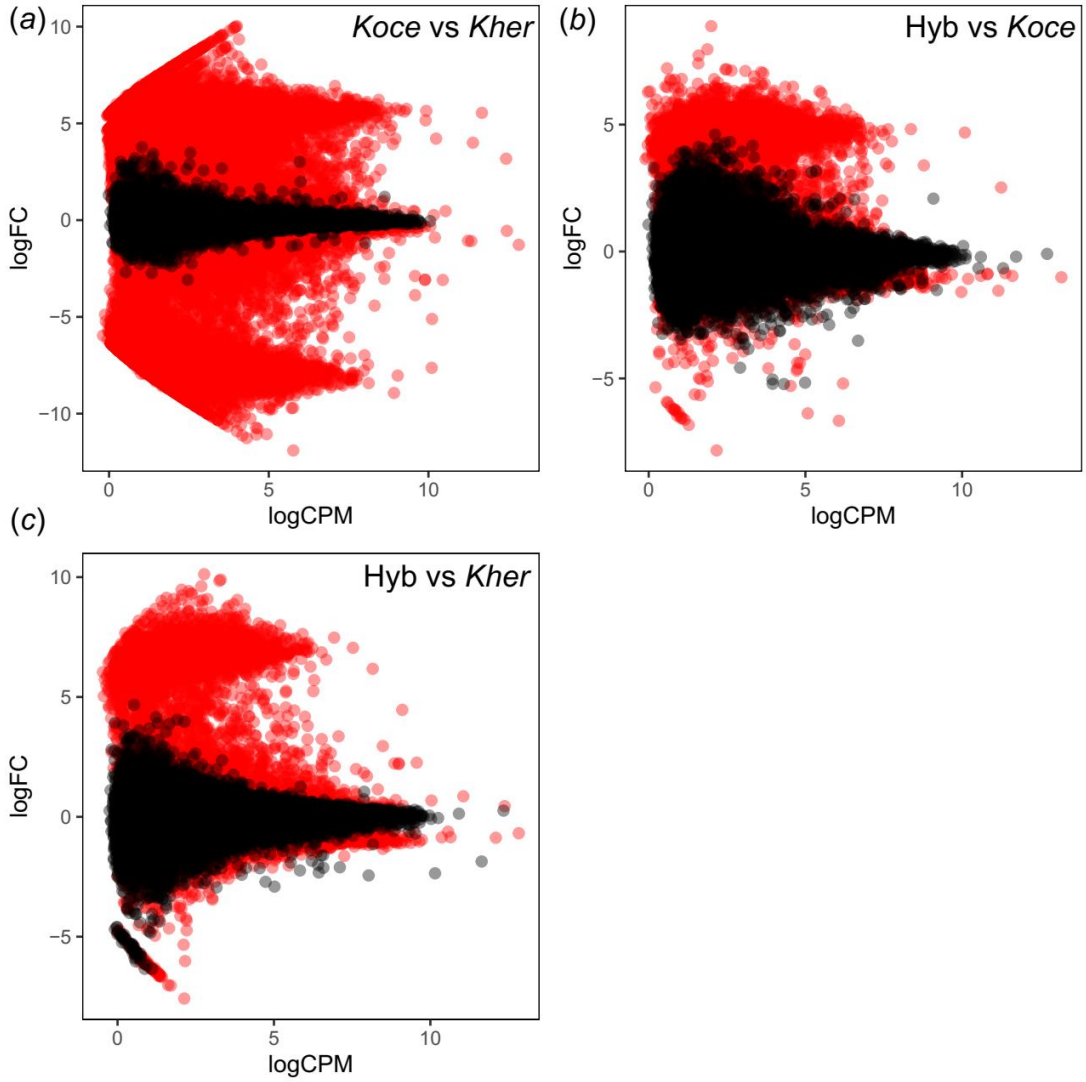


Figure S3.3. Differentially methylated cytosines (DMCs) identified in the comparisons between (*a*) *Kryptolebias ocellatus* and *Kryptolebias hermaphroditus*, (*b*) hybrids and *K. ocellatus*, and (*c*) hybrids and *K. hermaphroditus*. Black circles for FDR values ≥ 0.01; red circles for FDR values < 0.01.

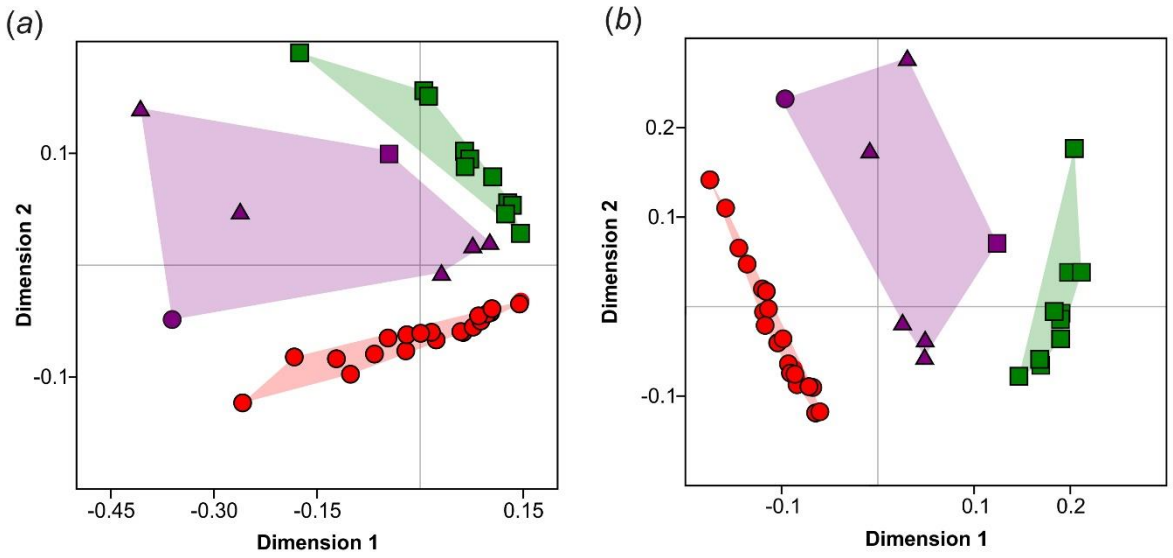


Figure S3.4. Multidimensional scaling analysis (MDS) of the normalised counts for (*a*) 830,950 sites with reads in the library and (*b*) the 37,774 DMCs between *K. ocellatus* and *K. hermaphroditus*. Squares represent *K. ocellatus*, circles represent *K. hermaphroditus*, and triangle represent hybrids. Backcrosses are represented by purple shapes according to their respective parental species.

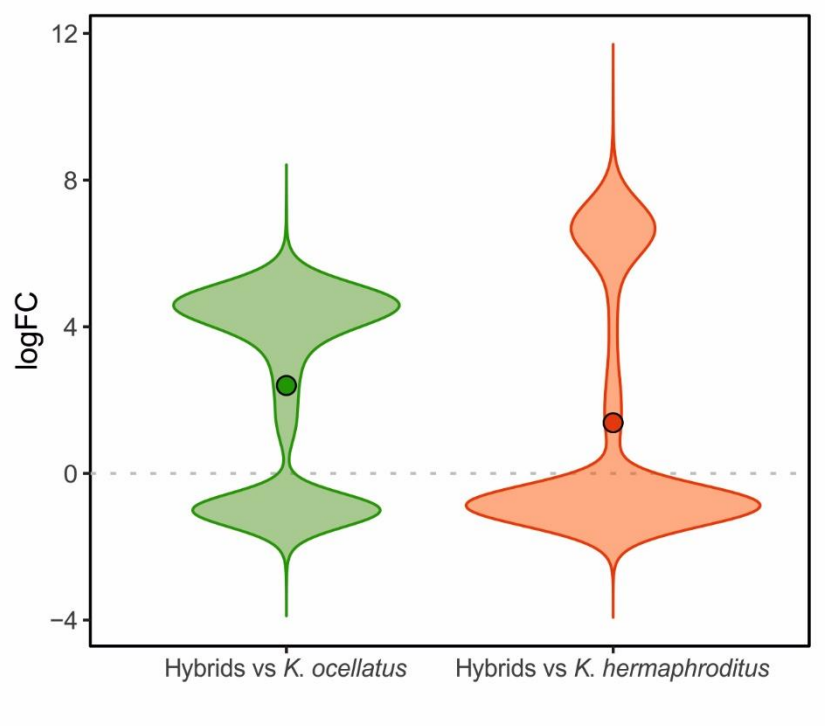


Figure S3.5. The direction of changes of the common 5,800 DMCs in the comparisons among hybrids vs parental species. logFC (logarithm of the fold change) values represent the intensity of changes compared to the hybrid values. logFC > 1 indicates higher CPM read counts (hypomethylation) in the hybrids compared to one of the parental species. logFC < 1 lower CPM read counts (hypermethylation) in the hybrids compared to one of the parental species. Central circles represent the mean for each comparison.

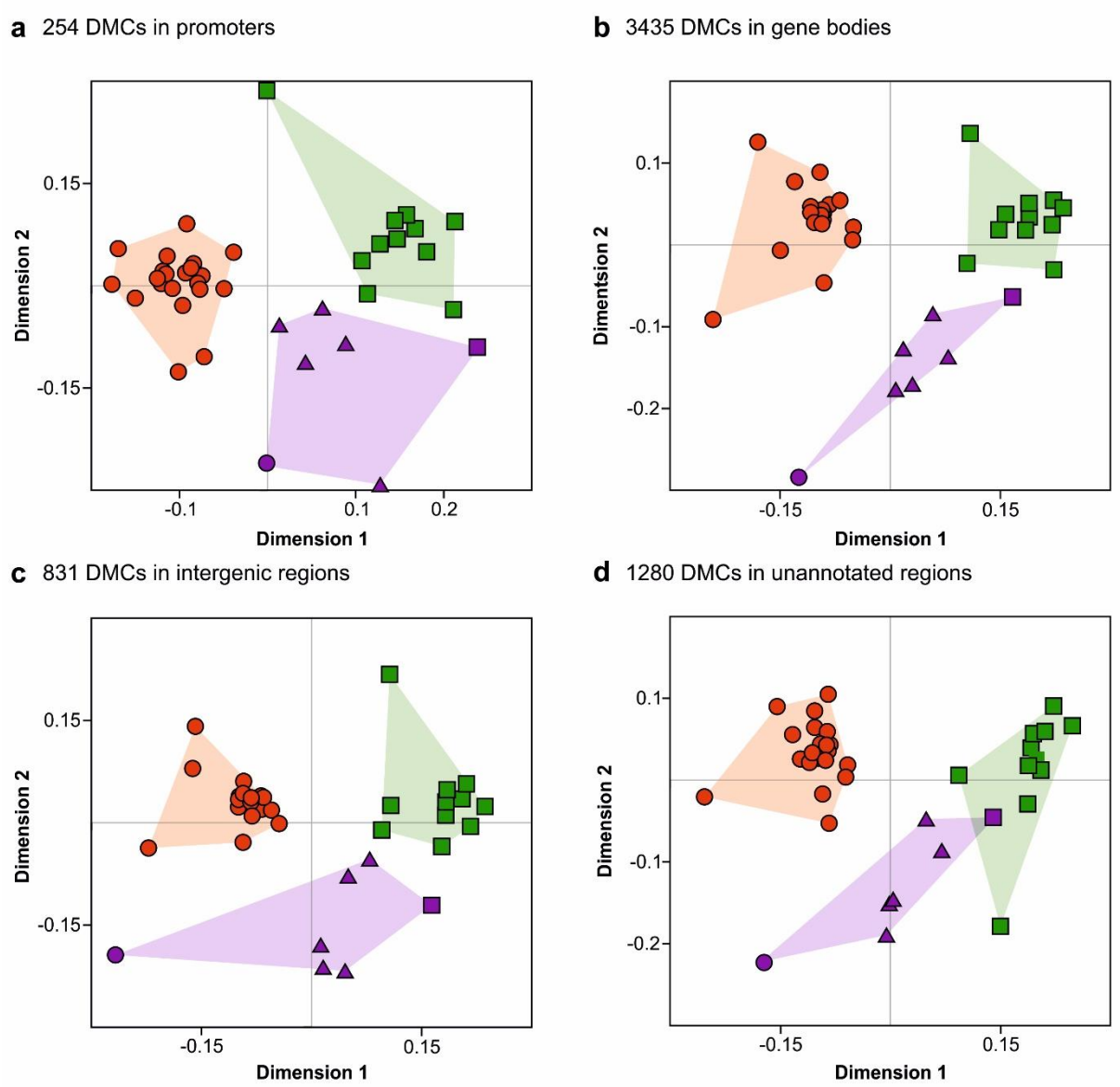


Figure S3.6. Multidimensional scaling analysis (MDS) of the normalised counts for the 5,800 DMCs common to the comparisons between hybrids and parental species in (*a*) putative promoters (within 2kb upstream the transcription start site), (*b*) gene bodies (*c*) intergenic regions (≥2kb upstream of TSS or downstream the gene bodies (*d*) unannotated regions. Squares for *K. ocellatus,* circles for *K. hermaphroditus,* and triangle for hybrids. Backcrosses are represented by purple shapes according to their respective parental species.

Supplementary figures for Chapter 4 - What does drive DNA methylation plasticity? Epigenetic responses of two inbred lines to different rearing environments

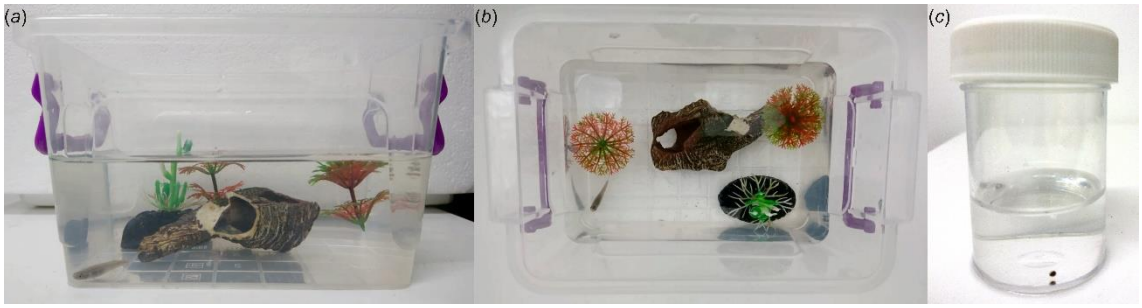


Figure S4.1. Pictures of the experimental tanks (9cm depth x 12cm width x 8cm length) filled with 400ml of brackish water with one perforated artificial log (3cm depth x 4 cm width x 4cm length) and three artificial plants to simulate a complex habitat in (*a*) lateral and (*b*) dorsal view. Poor environment tanks were identical except for the physical environmental enrichment (log and plants). (*c*) Circular plastic pots filled with 100ml of brackish water where eggs were maintained isolated until hatchling.

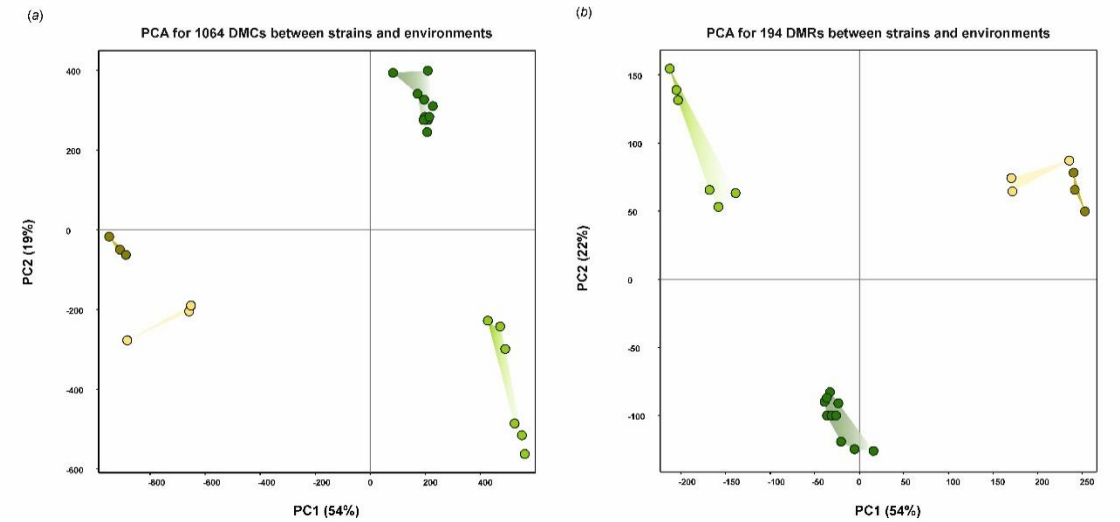


Figure S4.2. Principal component analysis (PCA) based on individual scores of methylation for (*a*) 1064 DMCs and (*b*) 194 DMRs found across genotypes and environments using logistic regression q< 0.01 and |ΔM|>20%, and ANOVA with alpha <0.01. Dark yellow for DAN individuals on enriched environments; light yellow for DAN genotype on poor environments; dark green for R individuals on enriched environments; light green for R genotype on poor environments.

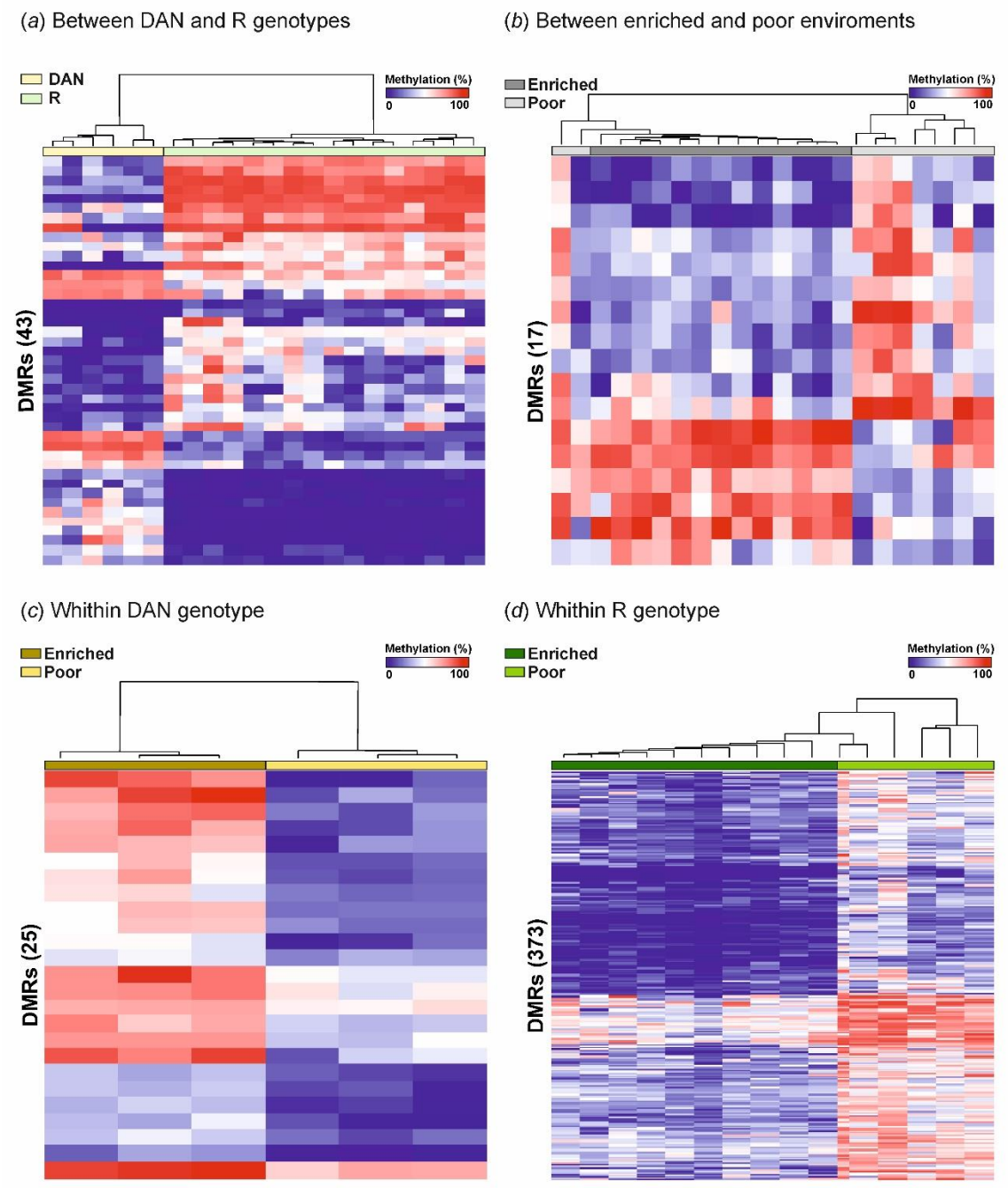


Figure S4.3. Heat map illustrating percentage of methylation for all differentially methylated regions (DMRs) identified (a) between genotypes, (b) between environments, (c) between environments for DAN strain, and (d) between environments for R strain (logistic regression $q < 0.01$ and $|\Delta M| > 20\%$, and t.test $p < 0.01$) using unsupervised hierarchical clustering. Rows represent a unique windowed region (5 CpG sites covering coverage requirements in a 1000bp window) and columns individual fish.

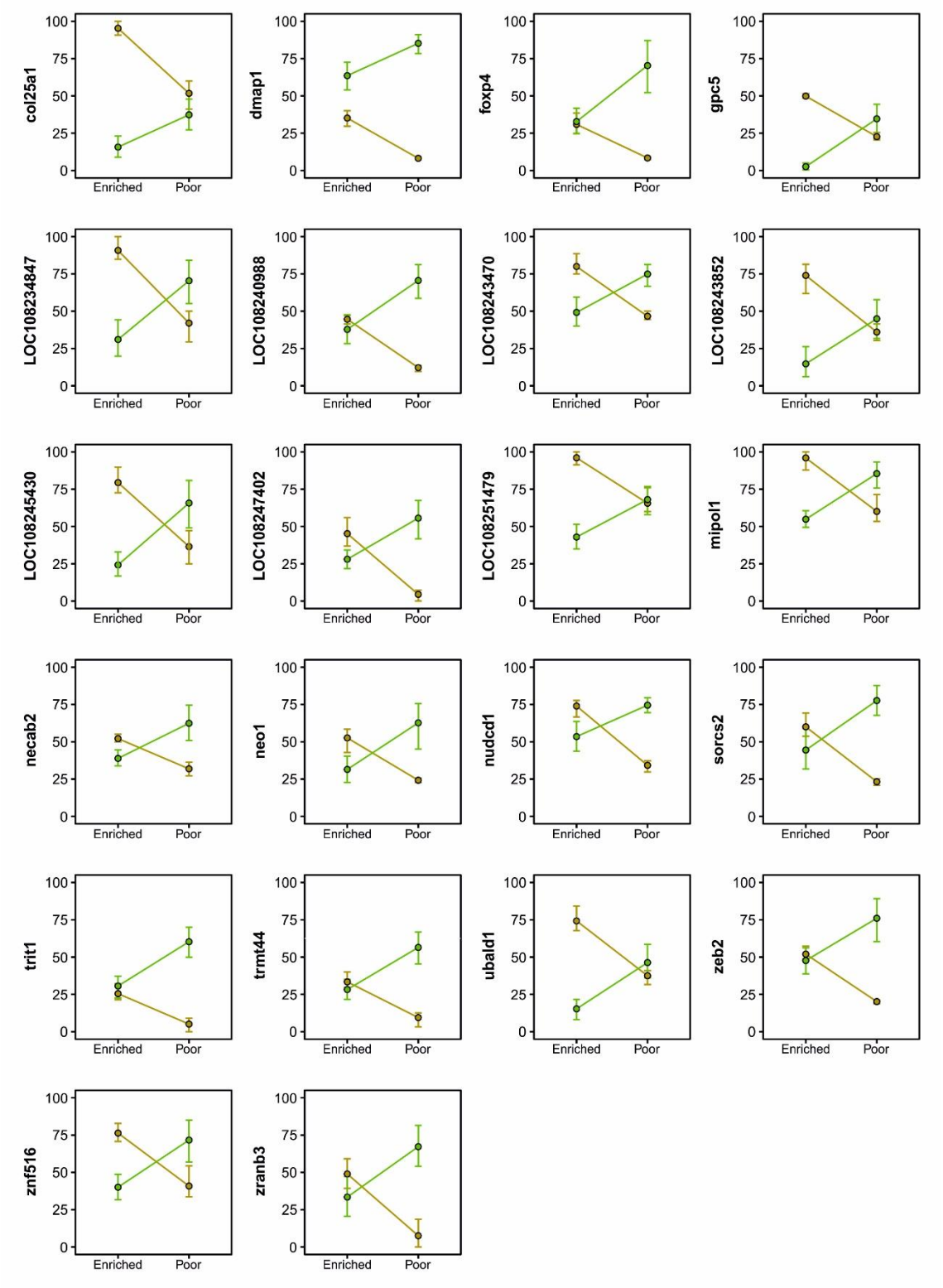


Figure S4.4. Reaction norms for methylation patterns of annotated DMCs neighbouring or overlapping gene bodies for facilitated epialleles. Methylation scores (in percentage) for each annotated DMC on the y-axis across genotypes and environments. Yellow for DAN individuals; green for R individuals.

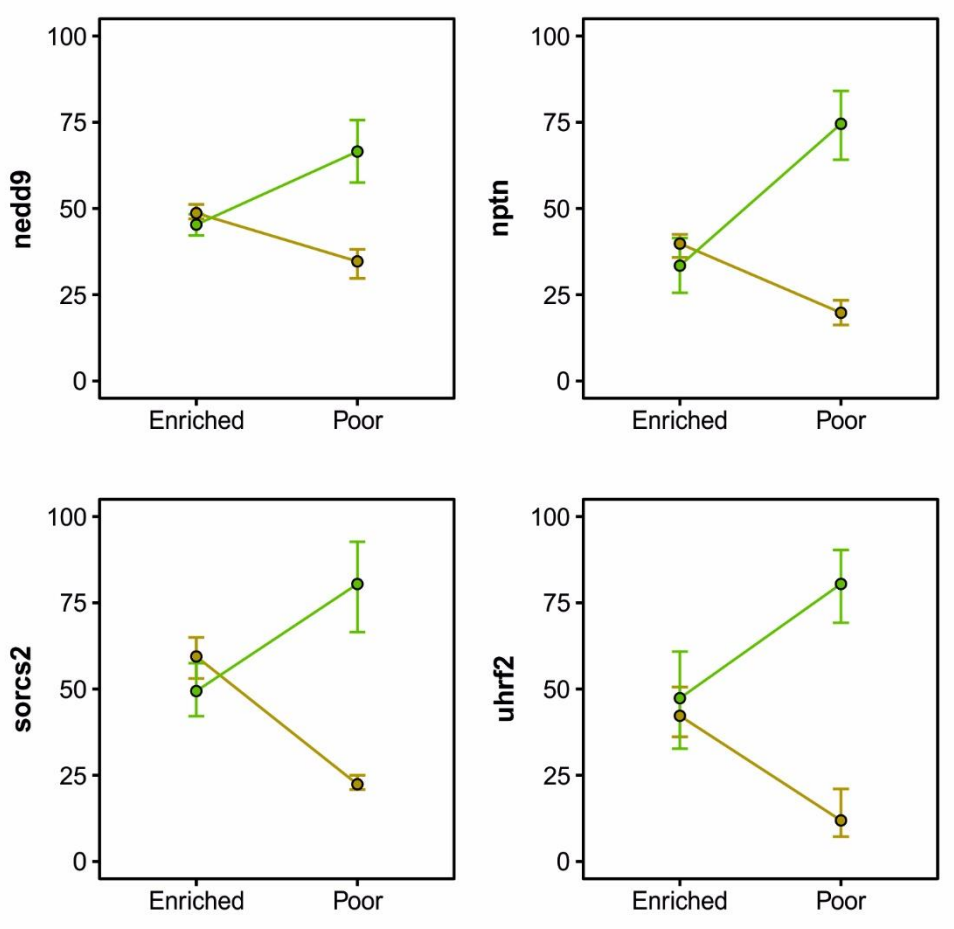


Figure S4.5. Reaction norms for methylation patterns of annotated DMRs neighbouring or overlapping gene bodies for DMRs commonly affected in both strains when compared between environments. Methylation scores (in percentage) for each annotated DMR on the y-axis across genotypes and environments. Yellow for DAN individuals; green for R individuals.

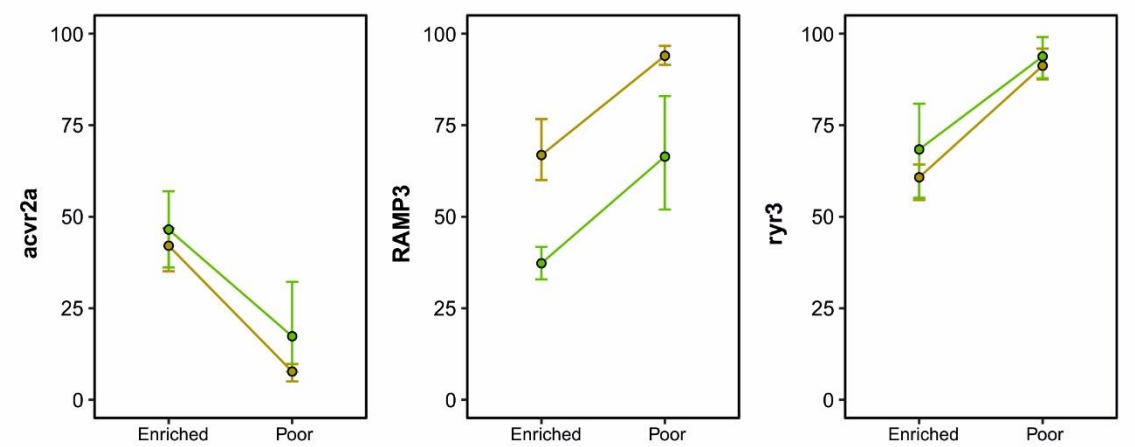


Figure S4.6. Reaction norms for methylation patterns of annotated DMCs neighbouring or overlapping gene bodies for pure epialleles. Methylation scores (in percentage) for each annotated DMC on the y-axis across genotypes and environments. Yellow for DAN individuals; green for R individuals.

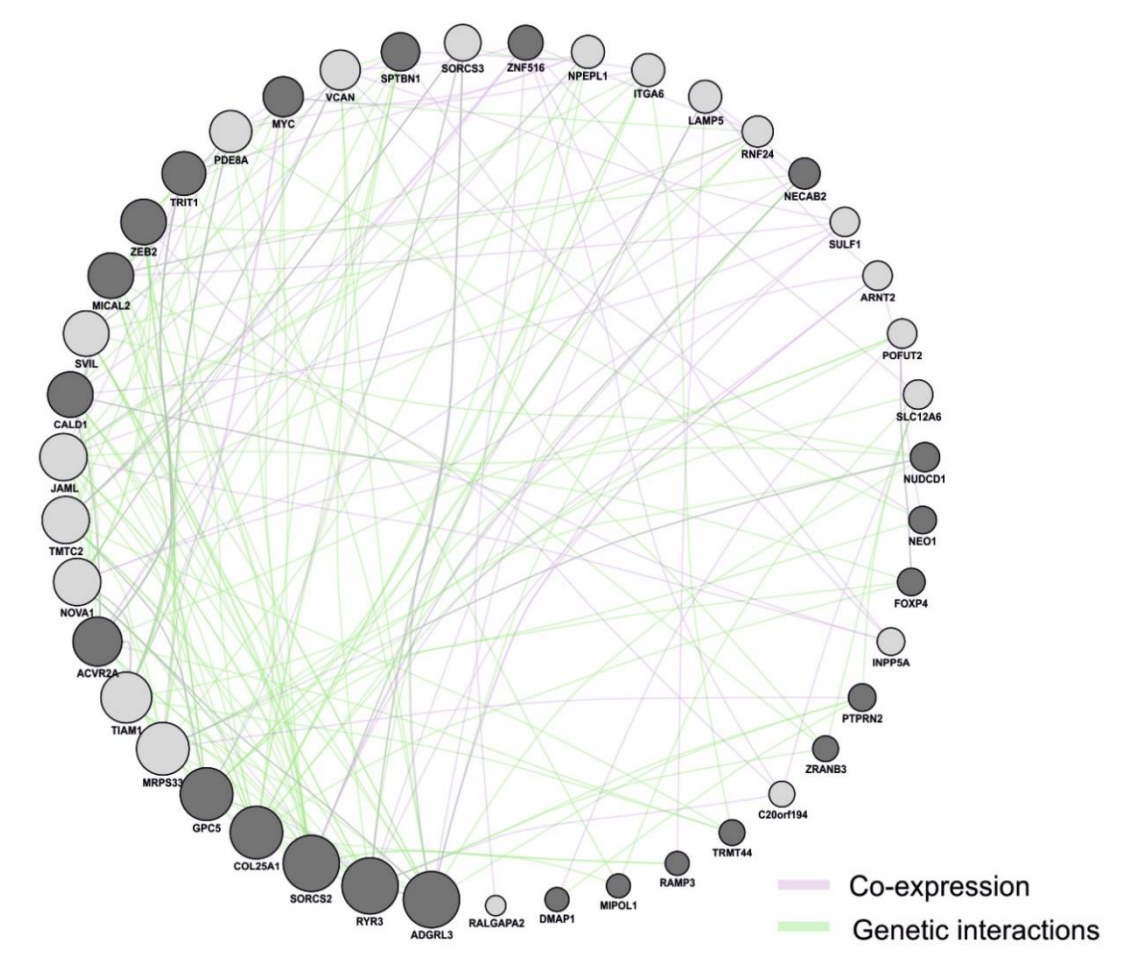


Figure S4.7. GeneMANIA network analysis of DMCs between environments (rich and poor) and shared between genotypes (DAN, R). Input genes in dark grey. Neighbouring genes indicated as part of the network but not part of the input genes indicated in light grey. Node size is proportional to the number of connections.

Supplementary figures for Chapter 5 - Environmental enrichment induces epigenetic and behavioural parental effects in a self-fertilising fish

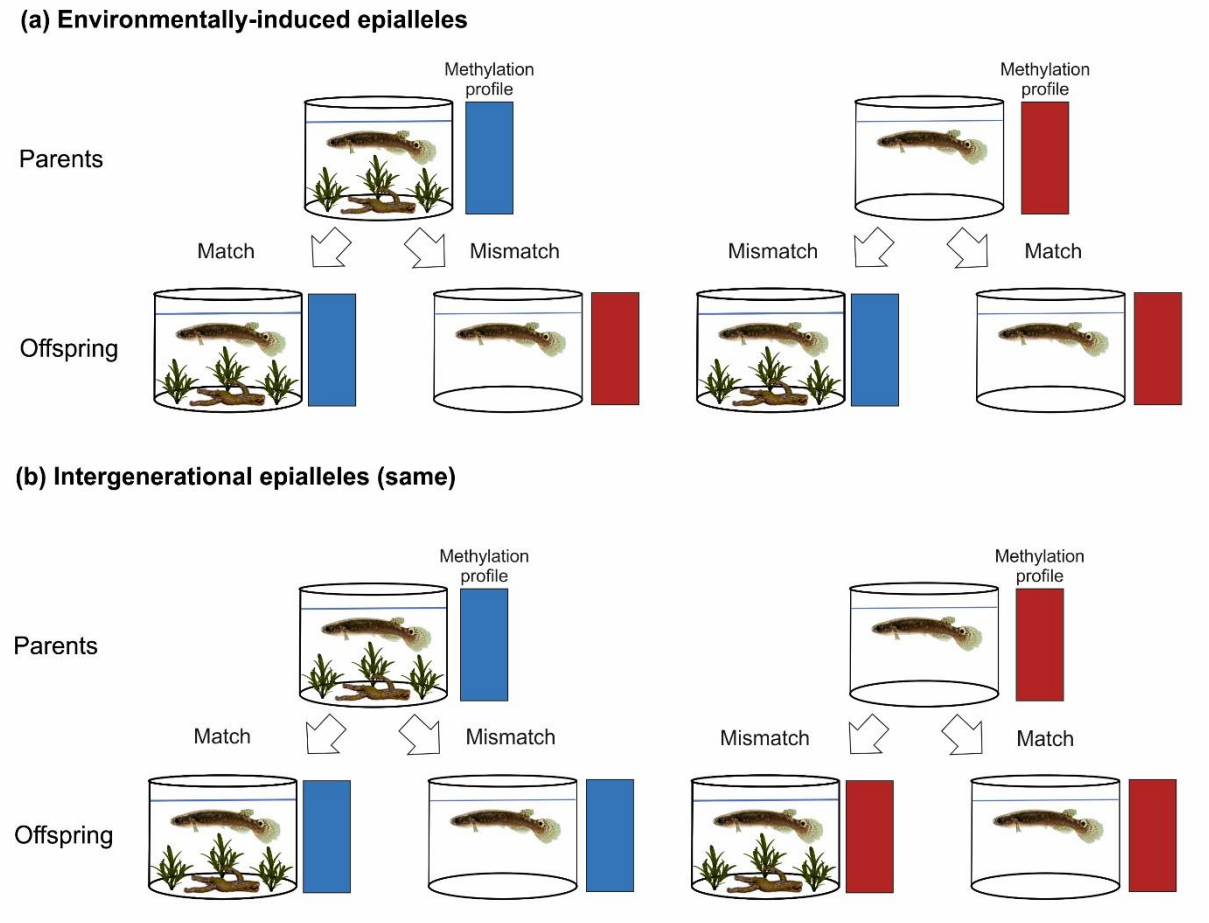


Figure S1.1. Schematic representation of epialleles classification according to its methylation patterns compared between parents and offspring. Epialleles were classified as (a) "environmentally-induced" if differences in methylation patterns found the parents changed in the offspring depending on their rearing environment; (b) intergenerational with parental patterns if methylation patterns in the parents were maintained in the offspring regardless of their rearing environment.

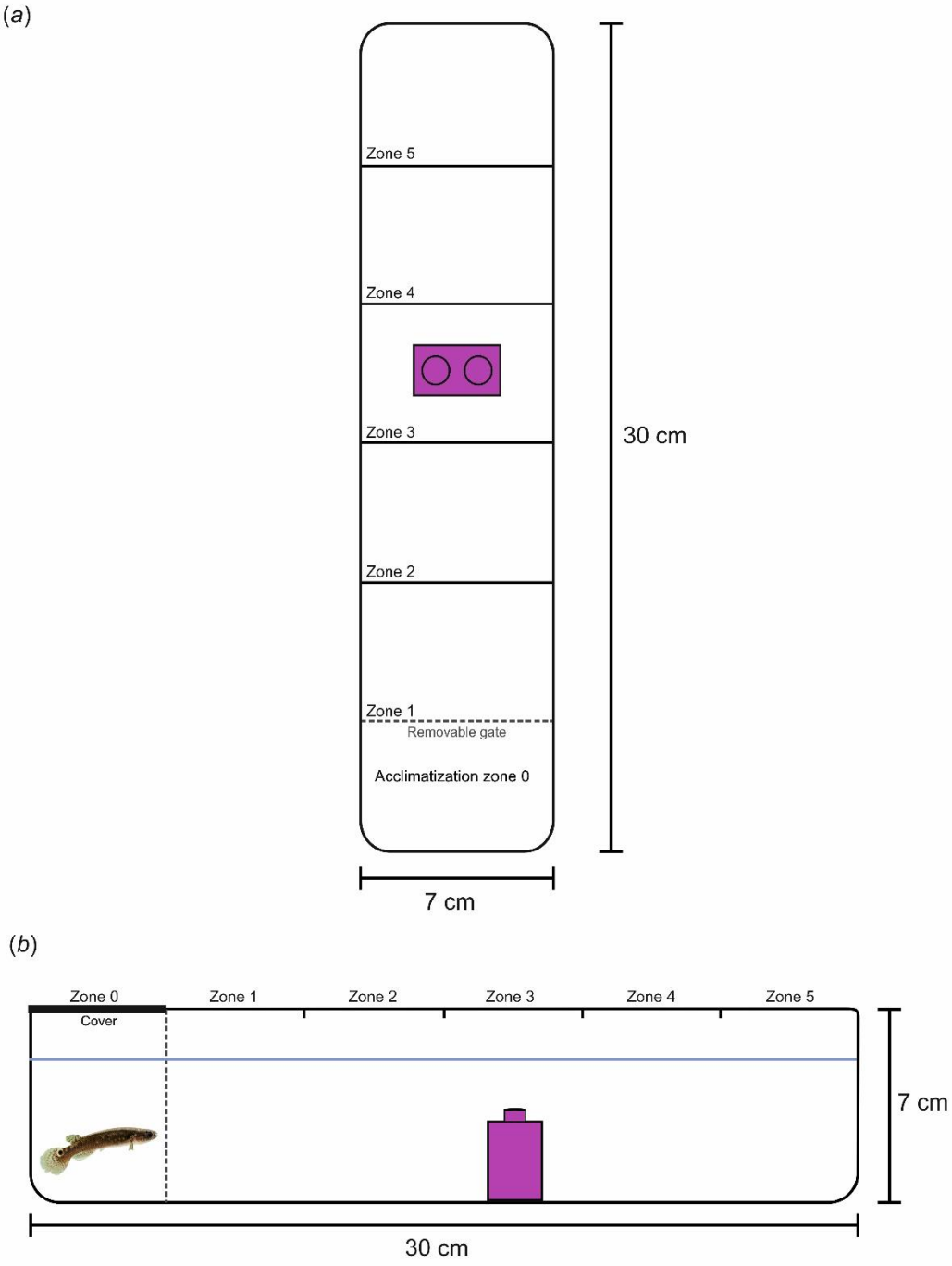


Figure S5.2. Schematic representation of the test arena used to assess neophobia in *Kryptolebias marmoratus* from (*a*) top and (*b*) side view. The acclimatization zone (zone 0) was separated by an opaque removable gate removed after acclimatization period. Five equally spaced zones of five centimetres were delimited by lines drawn at the top edge of the arena walls. The novel object, a conspicuous plastic brick toy was glued to the centre of zone.

IV.II. Tables

Supplementary figures for Chapter 1 - Genetic background and parasite load affect DNA methylation variation in a predominantly-self fertilising fish

Table S1.1. Genetic diversity (for 27 microsatellite loci), mean parasites number (standard deviation in brackets) and parasite prevalence in *Kryptolebias hermaphroditus* among selfing lineages identified by INTRUSCT in northeast Brazil. N= sampling size; N_a = mean number of alleles of alleles; H_e = expected heterozygosity; H_o = observed heterozygosity; F_{IS} = inbreeding coefficient; HL = homozygosity by locus; S = selfing rates.

	Lineage 1	Lineage 2	Lineage 3	Lineage 4	Lineage 5	Linage 6
N	14	25	41	22	18	8
Selfed/Outcrossed	7/7	17/8	29/12	15/7	16/2	8/0
Sampling site	1	1	2	1	3	1 (7)and 2 (1)
Genetic diversity						
N_{ma}	2.14	2.29	3.44	2.22	3.14	1.59
H_e	0.25	0.24	0.26	0.19	0.33	0.17
H_o	0.01	0.03	0.01	0.02	0.04	0.00
F_{IS}	0.97	0.85	0.89	0.88	0.87	0.97
HL	0.97	0.92	0.97	0.95	0.93	0.99
S	0.87	0.91	0.91	0.92	0.92	0.93
Parasite loads						
Bacterial gills cysts	2.92 (2.18)	2.24 (1.9)	2.53 (3.02)	3.13 (2.76)	1.27 (0.80)	7.12 (5.39)
Protozoan gills cysts	0	0	1.56 (1.60)	0	0.33 (1.37)	0
Nematodes	0.07 (0.25)	0.16 (0.46)	0.02 (0.15)	0.13 (0.62)	0	0.37 (0.69)
Total parasite load	3.14 (2.03)	2.36 (1.85)	4.12 (3.15)	3.27 (3.01)	1.61 (1.73)	7.5 (5.54)

Parasite prevalence (% of fish with infection)						
Bacterial gills cysts	92.85	84.00	70.73	95.45	83.33	100.00
Protozoan gills cysts	0	0	41.46	0	5.55	0
Nematodes	7.14	16.00	2.43	4.54	0	25.00

Table S1.2. Pairwise F_{ST} values among sampling sites and selfing lineages (k) in *Kryptolebias hermaphroditus*. Asterisks represent significance levels (***P ≤0.001).

Per sampling site

	1	**2**	**3**
1	-		
2	0.25***	-	
3	0.32***	0.28***	-

Per selfing lineage (K)

	1	**2**	**3**	**4**	**5**	**6**
1	-					
2	0.20***	-				
3	0.34***	0.28***	-			
4	0.27***	0.24***	0.35***	-		
5	0.34***	0.36***	0.29***	0.40***	-	
6	0.30***	0.28***	0.36***	0.40***	0.36***	-

Table S1.3. Hierarchical analysis of molecular variance partitioning (AMOVA) for MS-AFLPs data among original and replicated samples to test for reproducibility. df= degrees of freedom; SSD= sum of squared deviations; Mol. var. (%) = molecular variance percentages from variance components sigma 2; ϕ_{ST} = Phi statistics for population differentiation. *P* value derived from 10,000 permutations.

	df	**SSD**	**Mol. var. (%)**	**ϕ_{ST}**	***P* value**
Epigenetic data (MSL)					
Between original and replicate	1	30.48	0.19	0.002	0.53
Within original and replicate	46	1474	99.81		
AFLP genetic data (NML)					
Between original and replicate	1	1.35	1.00	0.01	0.94
Within original and replicate	46	91.46	99.00		

Table S1.4. Hierarchical analysis of molecular variance (AMOVA) for microsatellites and MS-AFLPs data **a)** among selfing lineages identified by INSTRUCT and **b)** between inbreeding status (selfed and outcrossed) in *Kryptolebias hermaphroditus* for Site 1 (Ceará-Mirim mangrove). df= degrees of freedom; Mol. var. (%) = molecular variance percentages from variance components sigma 2; *P* value derived from 10,000 permutations.

	Microsatellites				NML				MSL			
	df	Mol. var. (%)	F_{ST}	*P* value	df	Mol. var. (%)	ϕ_{ST}	*P* value	df	Mol. var. (%)	ϕ_{ST}	*P* value
a Selfing lineages												
Among lineages	3	27.91	0.27	**0.001**	3	0.83	0.008	0.12	3	0.15	0.006	0.52
Within lineages	132	72.09			58	99.17			58	99.85		
b Inbreeding status												
Between selfed and outcrossed	1	3.42	0.03	**0.001**	1	0.16	0.006	0.57	1	0.64	0.001	0.20
Within selfed and outcrossed	134	96.58			60	99.84			60	99.36		

Table S1.5. Results of the best generalized linear models indicated by the multi-model averaging approach for the proportion of methylated loci in *Kryptolebias hermaphroditus*. Models are ranked according to their corrected Akaike Information Criterion (AICc), the distance between a given model and the best fitting model (ΔAICc), the Akaike weight (Wi), and the evidence ratio (ER), which represents the ratio between the weights of the best and competing models. Only the best-fitting models within two AICc units are shown.

Independent variable	**df**	**z-value**	***P*-value**	**AICc**	**ΔAICc**	**W_i**	**ER**
Proportion of methylated loci							
Model 1				1065.5	0.00	0.28	1.00
Selfing lineage	5	-4.50	**<0.001**				
Scaled parasite load	1	-0.02	0.83				
Inbreeding	1	1.73	0.15				
Selfing lineage x parasite scaled	5	-3.90	**0.005**				
Selfing lineage x Inbreeding	4	-1.64	**0.04**				
Model 2				1066.5	1.00	0.17	1.39
Selfing lineage	5	-6.48	**<0.001**				
Inbreeding	1	-2.53	0.15				
Selfing lineage x Inbreeding	4	-2.22	0.04				

Table S1.6. Results of the best generalized linear models indicated by the multi-model averaging approach for the proportion of methylated loci in *Kryptolebias hermaphroditus* (including homozygosity index as independent variable). Models are ranked according to their corrected Akaike Information Criterion (AICc), the distance between a given model and the best fitting model (ΔAICc), the Akaike weight (Wi), and the evidence ratio (ER), which represents the ratio between the weights of the best and competing models. Only the best-fitting models within two AICc units are shown.

Independent variable	**df**	**z-value**	***P*-value**	**AICc**	**ΔAICc**	**W_i**	**ER**
Proportion of methylated loci							
Model 1				1021.9	0.00	0.04	1.00
Selfing lineage	5	-6.50	**<0.001**				
Scaled parasite load	1	-0.21	0.83				
Inbreeding	1	-1.19	**0.05**				
Homozygosity (HL)	1	-0.32	0.29				
Scaled parasite load x HL	1	3.17	0.28				
Inbreeding x HL	5	7.22	**<0.001**				
Selfing lineage xHL	1	12.13	**<0.001**				
Model 2				1023.2	1.30	0.02	3.19
Selfing lineage	5	-5.70	**<0.001**				
Scaled parasite load	1	-0.37	0.83				
Inbreeding	1	-1.19	0.05				
Homozygosity (HL)	1	-0.35	0.29				
Scaled parasite load x HL	1	3.14	0.10				
Inbreeding x HL	5	6.52	**<0.001**				
Selfing lineage xHL	1	11.22	**<0.001**				
Selfing lineage x Inbreeding	4	-3.10	**0.01**				

Table S1.7. Results of the best generalized linear model indicated by the multi-model averaging approach for the proportion of methylated loci including the number of bacterial cysts as a predictor in *Kryptolebias hermaphroditus*. Models are ranked according to their corrected Akaike Information Criterion (AICc), the distance between a given model and the best fitting model (ΔAICc), the Akaike weight (W_i), and the evidence ratio (ER), which represents the ratio between the weights of the best and competing models. Only the best-fitting models within two AICc units are shown.

Independent variable	**df**	**z-value**	***P*-value**	**AICc**	**ΔAICc**	**W_i**	**ER**
Proportion of methylated loci (with number of bacterial cysts as a predictor)							
Model 1				1060.8	0.00	0.51	1.00
Selfing lineage	5	-9.23	**<0.001**				
Bacterial cysts	1	-2.93	0.88				
Inbreeding	1	-1.17	0.15				
Selfing lineage x Bacterial cysts	5	9.09	**<0.001**				
Selfing lineage x inbreeding	4	-4.97	**0.04**				

Table S1.8. Results of the best generalized linear modes indicated by the multi-model averaging approach for the proportion of methylated loci in *Kryptolebias hermaphroditus* sampled on sampling site 1. Models are ranked according to their corrected Akaike Information Criterion (AICc), the distance between a given model and the best fitting model (ΔAICc), the Akaike weight (W_i), and the evidence ratio (ER), which represents the ratio between the weights of the best and competing models. Only the best-fitting models within two AICc units are shown.

Independent variable	**df**	**z/t-value**	***P*-value**	**AICc**	**ΔAICc**	**W_i**	**ER**
Proportion of methylated loci							
Model 1				621.86	0.00	0.29	1.00
Selfing lineage	3	-7.09	**0.04**				
Scaled parasite load	1	-5.02	0.31				
Inbreeding	1	-3.95	**0.04**				
Inbreeding x parasite scaled	1	-10.43	**0.01**				
Selfing lineage x inbreeding	2	-9.52	**0.001**				
Model 2				621.88	0.02	0.29	-0.66
Scaled parasite load	1	-11.49	**0.03**				
Inbreeding	1	-10.64	**0.09**				
Inbreeding x scaled parasite load	1	-17.93	**<0.001**				
Model 3				622.70	0.82	0.19	1.82
Selfing lineage	3	-3.61	**0.04**				
Scaled parasite load	1	-3.81	0.31				
Inbreeding	1	-2.15	**0.04**				
Inbreeding x parasite scaled	1	-7.04	**0.01**				
Selfing lineage x inbreeding	2	-7.10	**0.01**				
Selfing lineage x parasite scaled	3	-6.50	**0.005**				
Model 4				623.74	1.88	0.11	1.67

Selfing lineage	3	-10.72	**0.04**
Scaled parasite load	1	-5.88	0.31
Inbreeding	1	-4.08	**0.04**
Inbreeding x parasite scaled	1	-11.39	**<0.001**

Supplementary tables for Chapter 2 - Hybridisation between species with different mating systems revealed by the genetic structure of the mangrove killifish *Kryptolebias ocellatus*

Table S2.1. Summary of 55 samples sequenced on genotype-by-sequencing library. Individuals in red failed to pass the filtering cut-off (≥500kb reads). Hybrids marked with asterisks. Parameters 'proportion of heterozygous sites' (for Dataset 1), 'coverage and missing data' (for Dataset 2) are described in methods.

ID	Species (identified using cox1)	Sampling location	*cox1* haplotype	Number of reads	Uniquely mapped (%)	Coverage (X)	Missing data (%)	Proportion of Heterozygous sites
GUA_27	*K. hermaphroditus*	GUA	1 (Kher)	8,962,202	90.30	250.4	0.00	0.03
GUA_28	*K. hermaphroditus*	GUA	1 (Kher)	4,161,657	90.87	114.9	0.02	0.03
GUA_29	*K. hermaphroditus*	GUA	1 (Kher)	13,831,746	90.29	380.1	0.00	0.02
GUA_30	*K. hermaphroditus*	GUA	1 (Kher)	7,000,238	90.05	190.2	0.00	0.02
GUA_31	*K. hermaphroditus*	GUA	1 (Kher)	6,646,432	89.95	185.2	0.02	0.03
GUA_37	*K. hermaphroditus*	GUA	1 (Kher)	3,013,973	90.06	82.2	0.18	0.03
GUA_39	*K. hermaphroditus*	GUA	1 (Kher)	3,472,656	89.73	93.5	0.02	0.04
GUA_40	*K. hermaphroditus*	GUA	1 (Kher)	13,651,662	90.50	382.6	0.00	0.03
GUA_41	*K. hermaphroditus*	GUA	1 (Kher)	5,348,736	90.08	147.3	0.04	0.03
GUA_42	*K. hermaphroditus*	GUA	1 (Kher)	1,786,528	91.00	48.8	0.20	0.02

GUA_43	*K. hermaphroditus*	GUA	1 (Kher)	7,269,128	89.39	197.0	0.00	0.03
GUA_44	*K. hermaphroditus*	GUA	1 (Kher)	7,937,202	89.92	216.6	0.00	0.02
FUN_20	*K. hermaphroditus*	FUN	1 (Kher)	8,493,250	89.83	229.9	0.00	0.02
FUN_21	*K. hermaphroditus*	FUN	1 (Kher)	4,140,936	90.70	113.3	0.00	0.03
FUN_22	*K. hermaphroditus*	FUN	1 (Kher)	4,990,777	90.96	138.7	0.00	0.03
FUN_24	*K. hermaphroditus*	FUN	1 (Kher)	3,520,363	89.86	98.6	0.02	0.03
FUN_25	*K. hermaphroditus*	FUN	1 (Kher)	4,421,739	90.34	120.7	0.04	0.03
FUN_26*	*K. hermaphroditus*	FUN	1 (Kher)	1,361,947	69.56	24.9	0.84	0.32
FUN_29	*K. hermaphroditus*	FUN	1 (Kher)	7,611,704	89.30	208.2	0.00	0.02
FUN_30	*K. hermaphroditus*	FUN	1 (Kher)	2,892,977	88.63	77.2	0.00	0.03
FUN_31	*K. hermaphroditus*	FUN	1 (Kher)	9,160,363	89.53	251.5	0.00	0.03
FUN_32	*K. hermaphroditus*	FUN	1 (Kher)	2,322,888	88.84	62.4	0.02	0.02
GUA_04	*K. ocellatus*	GUA	8	5,033,372	77.40	102.4	0.15	0.08
GUA_05	*K. ocellatus*	GUA	9	11,979935	75.68	235.1	0.00	0.08
GUA_06	*K. ocellatus*	GUA	8	6,418,126	81.00	143.5	0.04	0.08
GUA_09*	*K. ocellatus*	GUA	8	5,784,849	84.54	138.1	0.05	0.61
GUA_10	*K. ocellatus*	GUA	8	11,097,202	81.18	243.2	0.04	0.08
GUA_11	*K. ocellatus*	GUA	8	2,208,760	75.05	44.5	0.46	0.06

GUA_13	*K. ocellatus*	GUA	8	11,446,137	77.19	230.1	0.00	0.09
GUA_16	*K. ocellatus*	GUA	8	14,282,577	80.18	313.2	0.00	0.11
GUA_17*	*K. ocellatus*	GUA	8	4,522,779	72.17	91.6	0.00	0.50
GUA_18	*K. ocellatus*	GUA	8	7,266,866	79.34	149.5	0.05	0.08
GUA_19	*K. ocellatus*	GUA	8	8,150,733	81.55	177.2	0.09	0.07
FUN_06	*K. ocellatus*	FUN	6	5,098,517	81.78	111.7	0.26	0.06
FUN_08*	*K. ocellatus*	FUN	4	9,433,950	86.75	234.2	0.00	0.60
FUN_10	*K. ocellatus*	FUN	1	6,856,002	81.10	145.5	0.02	0.14
FUN_11*	*K. ocellatus*	FUN	6	7,773,860	86.68	199.0	0.00	0.62
FUN_13	*K. ocellatus*	FUN	6	2,645	NA	NA	NA	NA
FUN_41	*K. ocellatus*	FUN	6	7,793	NA	NA	NA	NA
FUN_43*	*K. ocellatus*	FUN	6	578,575	83.31	13.8	4.14	0.48
FUN_47*	*K. ocellatus*	FUN	7	564,546	78.82	12.4	4.16	0.50
FLO_02	*K. ocellatus*	FLO	14	13,605,858	81.73	301.5	0.04	0.05
FLO_03	*K. ocellatus*	FLO	14	2,439,264	70.34	48.0	0.38	0.06
FLO_04	*K. ocellatus*	FLO	21	2,167,505	79.99	46.8	0.40	0.06
FLO_05	*K. ocellatus*	FLO	14	2,517,261	81.12	53.6	0.66	0.06
FLO_06	*K. ocellatus*	FLO	14	4,609,454	78.83	96.2	0.11	0.06

FLO_07	*K. ocellatus*	FLO	22	4,620,891	81.46	96.3	0.02	0.05
FLO_08	*K. ocellatus*	FLO	14	6,823,947	79.43	149.9	0.00	0.06
SFR_01	*K. ocellatus*	SFR	15	9,375,699	81.74	209.4	0.05	0.06
SFR_02	*K. ocellatus*	SFR	16	2,671,245	82.54	60.3	0.37	0.06
SFR_03	*K. ocellatus*	SFR	16	8,014,672	81.64	177.9	0.11	0.06
SFR_04	*K. ocellatus*	SFR	17	8,136,965	83.09	177.8	0.00	0.07
SFR_06	*K. ocellatus*	SFR	14	10,875,988	82.37	240.8	0.15	0.05
SFR_07	*K. ocellatus*	SFR	15	10,572775	82.07	235.7	0.05	0.06
SFR_08	*K. ocellatus*	SFR	19	684,671	77.07	14.5	4.93	0.05
Total				**337,622,523**	**83.90**	**152.98**	**0.34**	**0.11**

Table S2.2. cox*1* haplotype distribution across different locations found for *Kryptolebias ocellatus* and *K. hermaphroditus*.

	K. ocellatus																					*K. hermaphroditus*
Haplotype number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	1 (Kherm)
1. IRI	17	1	2	1	1																		
2. FUN	1			3		6	1																29
3. GUA								15	2														30
4. PRT										6	1	2	30										
5. PAR										1				1	3								
6. SFR														1	4	5	1	1	6	1			
7. FLO														11							3	2	
8.PIC																							2
Total	18	1	2	4	1	6	1	15	2	7	1	2	30	13	7	5	1	1	6	1	3	2	61

Table S2.3. Descriptive statistics of genetic variation at *cox1* mitochondrial DNA (mtDNA) gene for clades and sampling locations of *K. ocellatus*. H = number of haplotypes; S = number of polymorphic sites; *h* = haplotype diversity; π = nucleotide diversity.

	Population diversity indices				
	N	*H*	*S*	*h*	π
Overall	129	22	26	0.895	0.007
Major mtDNA clades					
Northern clade	50	9	10	0.775	0.002
Parati	39	4	3	0.391	0.0007
Southern clade	40	10	15	0.836	0.002
Sampling location					
1.IRI	22	5	5	0.407	0.001
2.FUN	11	4	6	0.673	0.002
3.GUA	17	2	1	0.248	0.0004
4.PRT	39	4	3	0.391	0.0007
5.PAR	5	3	9	0.700	0.006
6.SFR	19	7	4	0.819	0.002
7.FLO	16	3	3	0.508	0.002

Table S2.4. Pairwise F_{ST} values (below diagonal), Kimura-2-parameter (K2P) genetic distance (in percentage above diagonal) and within group K2P distance for mtDNA *cox1* gene among sampling locations for *Kryptolebias ocellatus*. Asterisks represent p-value ≤ 0.05.

	1	**2**	**3**	**4**	**5**	**6**	**7**
1. IRI	**0.1**	0.4	0.2	0.9	1.1	1.2	1.1
2. FUN	0.55*	**0.2**	0.3	1.0	1.2	1.3	1.2
3. GUA	0.58*	0.59*	**0**	0.7	0.9	1.1	0.9
4. PRT	0.88*	0.91*	0.88*	**0.1**	1.1	1.4	1.3
5. PAR	0.77*	0.79*	0.69*	0.85*	**0.6**	0.5	0.4
6. SFR	0.83*	0.84*	0.80*	0.90*	0.14	**0.3**	0.3
7. FLO	0.87*	0.90*	0.85*	0.92*	0.28*	0.30*	**0.1**

Table S2.5. F_{IS} values per locus and population for 179 individuals of *Kryptolebias ocellatus*. Asterisks indicate significance ($p<0.05$) after correction for multiple testing. Hybrids excluded from analysis. GUA samples from different years are separated.

	IRI	**FUN**	**GUA 2017**	**GUA 2007**	**PRT**	**PAR**	**SFR**	**FLO**
R9	-0.05	0.48	0.33*	0.59*	0.05	-0.17	-0.12	0.07
R11	0.09	-0.08	0.06	0.13	0.25*	-0.06	-0.10	-0.02
R18	-0.01	0.00	-0.04	0.00	-0.01	NA	NA	NA
R23	0.01	-0.05	-0.04	0.05	0.02	0.29	-0.04	0.17
R27	0.03	0.17	-0.06	-0.00	-0.02	0.05	-0.06	0.04
R28	NA	NA	NA	NA	NA	NA	0.000	NA
R30	0.26*	0.20	-0.15	-0.02	-0.13	0.03	-0.05	-0.01
R33	0.05	0.05	0.05	0.00	0.07	-0.20	-0.18	NA
R34	0.00	NA	0.00	-0.06	-0.01	NA	-0.71*	-0.09
R37	-0.01	-0.08	0.11	0.04	0.01	-0.14	0.00	0.05
R38	0.06	-0.08	0.26*	0.16	0.03	-0.11	-0.08	0.16
B86	-0.09	0.27	-0.13	0.07	-0.19	-0.33	-0.05	-0.18
R90	-0.01	NA	-0.02	-0.07	0.17	0.17	-0.32	-0.10
R92	0.42*	0.04	0.37	0.23	0.05	0.00	0.25	-0.05
R103	-0.11	-0.14	-0.06	0.21	-0.05	NA	-0.04	0.02
R112	-0.14	-0.28	-0.23	0.06	NA	NA	NA	0.00
All	**0.05***	**0.04**	**0.06***	**0.11***	**0.04**	**-0.02**	**-0.10***	**0.03**

Table S2.6. Pairwise F_{ST} values among *Kryptolebias ocellatus* populations based on 16 microsatellite loci. Hybrids are excluded from analysis. Asterisks indicate significance ($p<0.05$) after Bonferroni correction. GUA samples from different years are separated.

	IRI	**FUN**	**GUA 2017**	**GUA 2007**	**PRT**	**PAR**	**SFR**	**FLO**
1. IRI	-	*	*	*	*	*	*	*
2. FUN	0.07	-	NS	*	*	NS	*	*
3. GUA 2017	0.08	0.07	-	NS	*	*	*	*
3. GUA 2007	0.09	0.08	0	-	*	*	*	*
4. PRT	0.20	0.21	0.18	0.19	-	*	*	*
5. PAR	0.38	0.38	0.36	0.36	0.36	-	*	*
6.SFR	0.31	0.30	0.29	0.29	0.28	0.21	-	*
7. FLO	0.33	0.32	0.30	0.31	0.30	0.29	0.07	-

Table S2.7. Averaged descriptive statistics of genetic variation at microsatellite loci in 179 *Kryptolebias ocellatus* (excluding hybrids), 11 hybrids and 67 *K. hermaphroditus*. N = sample size; P_{99} = proportion of polymorphic loci (99% criterion); A = average number of alleles; A_R = allelic richness based on 11 individuals; H_E = expected heterozygosity; H_O = observed heterozygosity; F_{IS} = coefficient of inbreeding. Asterisks represent significant departures from HWE.

	N	***P*$_{99}$**	***A***	***A*$_R$**	***H*$_E$**	***H*$_O$**	***F*$_{IS}$**
Kryptolebias ocellatus	179	0.85	7.53	5.28	0.56	0.54	0.04
Hybrids	11	0.93	4.47	4.29	0.65	0.83	-0.31*
K. hermaphroditus	67	0.21	1.61	1.38	0.07	0.00	0.92*

Table S8.2. Pairwise F_{ST} values (below diagonal) and number of sites used in the comparison (upper diagonal) based on genotypes for *Kryptolebias ocellatus* and *K. hermaphroditus* among sampling points. Hybrids are excluded from analysis.

Species			*K. ocellatus*				*K. hermaphroditus*	
	Clade		Northern		Southern			
			FUN	GUA	SFR	FLO	FUN	GUA
K. ocellatus	Northern	FUN	-	1,088,202	1,019,797	921,267	1,172,844	1,158,256
		GUA	0.09	-	932,151	849,947	829,800	838,426
	Southern	SFR	0.47	0.42	-	861,114	806,433	809,800
		FLO	0.49	0.42	0.18	-	763,561	750,704
K. hermaphroditus		FUN	0.89	0.90	0.93	0.94	-	1,480,233
		GUA	0.91	0.91	0.94	0.94	0.06	-

Supplementary tables for Chapter 3 - Intermediate patterns of epigenetic variation in hybrids of divergent mangrove killifish species in natural populations

Table S3.1. Summary of the 39 samples included on methylation-sensitive genotype-by-sequencing analysis. Hybrids marked with asterisks.

ID	Species (identified using cox1)	Sampling location	Number of reads	Uniquely mapped (%)
GUA_27	*K. hermaphroditus*	GUA	8,962,202	90.30
GUA_28	*K. hermaphroditus*	GUA	4,161,657	90.87
GUA_29	*K. hermaphroditus*	GUA	13,831,746	90.29
GUA_30	*K. hermaphroditus*	GUA	7,000,238	90.05
GUA_31	*K. hermaphroditus*	GUA	6,646,432	89.95
GUA_37	*K. hermaphroditus*	GUA	3,013,973	90.06
GUA_39	*K. hermaphroditus*	GUA	3,472,656	89.73
GUA_40	*K. hermaphroditus*	GUA	13,651,662	90.50
GUA_41	*K. hermaphroditus*	GUA	5,348,736	90.08
GUA_42	*K. hermaphroditus*	GUA	1,786,528	91.00
GUA_43	*K. hermaphroditus*	GUA	7,269,128	89.39
GUA_44	*K. hermaphroditus*	GUA	7,937,202	89.92
FUN_20	*K. hermaphroditus*	FUN	8,493,250	89.83
FUN_21	*K. hermaphroditus*	FUN	4,140,936	90.70
FUN_22	*K. hermaphroditus*	FUN	4,990,777	90.96
FUN_24	*K. hermaphroditus*	FUN	3,520,363	89.86
FUN_25	*K. hermaphroditus*	FUN	4,421,739	90.34
FUN_26*	*K. hermaphroditus*	FUN	1,361,947	69.56
FUN_29	*K. hermaphroditus*	FUN	7,611,704	89.30
FUN_30	*K. hermaphroditus*	FUN	2,892,977	88.63
FUN_31	*K. hermaphroditus*	FUN	9,160,363	89.53
FUN_32	*K. hermaphroditus*	FUN	2,322,888	88.84
GUA_04	*K. ocellatus*	GUA	5,033,372	77.40
GUA_05	*K. ocellatus*	GUA	11,979935	75.68
GUA_06	*K. ocellatus*	GUA	6,418,126	81.00
GUA_09*	*K. ocellatus*	GUA	5,784,849	84.54
GUA_10	*K. ocellatus*	GUA	11,097,202	81.18
GUA_11	*K. ocellatus*	GUA	2,208,760	75.05
GUA_13	*K. ocellatus*	GUA	11,446,137	77.19
GUA_16	*K. ocellatus*	GUA	14,282,577	80.18
GUA_17*	*K. ocellatus*	GUA	4,522,779	72.17
GUA_18	*K. ocellatus*	GUA	7,266,866	79.34
GUA_19	*K. ocellatus*	GUA	8,150,733	81.55
FUN_06	*K. ocellatus*	FUN	5,098,517	81.78
FUN_08*	*K. ocellatus*	FUN	9,433,950	86.75
FUN_10	*K. ocellatus*	FUN	6,856,002	81.10
FUN_11*	*K. ocellatus*	FUN	7,773,860	86.68
FUN_43*	*K. ocellatus*	FUN	578,575	83.31
FUN_47*	*K. ocellatus*	FUN	564,546	78.82
Mean			**6,422,972**	**85.21**

Table S3.2. Summary of gene ontologies (GO) found for 1322 genes orthologous to zebrafish affected by 5,800 differentially methylated cytosines (DMCs) common to the comparisons between hybrids versus parental species. FDR = False discovery rate after Fisher test.

GO term	Number of genes	FDR
ameboidal-type cell migration (GO:0001667)	39	0.00
anatomical structure development (GO:0048856)	382	0.00
anatomical structure formation involved in morphogenesis (GO:0048646)	75	0.00
anatomical structure morphogenesis (GO:0009653)	209	0.00
animal organ development (GO:0048513)	227	0.00
animal organ morphogenesis (GO:0009887)	96	0.00
anterior/posterior axon guidance (GO:0033564)	5	0.04
axon development (GO:0061564)	51	0.00
axon guidance (GO:0007411)	42	0.00
axonogenesis (GO:0007409)	51	0.00
biological adhesion (GO:0022610)	53	0.00
biological regulation (GO:0065007)	609	0.00
brain development (GO:0007420)	46	0.00
camera-type eye morphogenesis (GO:0048593)	19	0.05
cardiovascular system development (GO:0072358)	49	0.00
cell adhesion (GO:0007155)	53	0.00
cell communication (GO:0007154)	266	0.00
cell development (GO:0048468)	133	0.00
cell differentiation (GO:0030154)	218	0.00
cell migration (GO:0016477)	69	0.00
cell morphogenesis (GO:0000902)	76	0.00
cell morphogenesis involved in differentiation (GO:0000904)	63	0.00
cell morphogenesis involved in neuron differentiation (GO:0048667)	55	0.00
cell motility (GO:0048870)	72	0.00
cell part morphogenesis (GO:0032990)	61	0.00
cell projection morphogenesis (GO:0048858)	60	0.00
cell projection organization (GO:0030030)	84	0.00
cell surface receptor signalling pathway (GO:0007166)	95	0.00
cell surface receptor signalling pathway involved in cell-cell signalling (GO:1905114)	20	0.05
cell-cell signaling (GO:0007267)	50	0.00
cellular component morphogenesis (GO:0032989)	84	0.00
cellular component organization (GO:0016043)	261	0.00
cellular component organization or biogenesis (GO:0071840)	275	0.00

cellular developmental process (GO:0048869)	225	0.00
cellular macromolecule biosynthetic process (GO:0034645)	77	0.04
cellular macromolecule metabolic process (GO:0044260)	248	0.00
cellular process (GO:0009987)	784	0.00
cellular response to stimulus (GO:0051716)	278	0.03
central nervous system development (GO:0007417)	63	0.00
chemotaxis (GO:0006935)	50	0.00
chordate embryonic development (GO:0043009)	55	0.00
circulatory system development (GO:0072359)	93	0.00
cranial skeletal system development (GO:1904888)	28	0.00
developmental growth (GO:0048589)	29	0.03
developmental growth involved in morphogenesis (GO:0060560)	14	0.04
developmental process (GO:0032502)	395	0.00
embryo development (GO:0009790)	102	0.00
embryo development ending in birth or egg hatching (GO:0009792)	55	0.00
embryonic cranial skeleton morphogenesis (GO:0048701)	26	0.00
embryonic epithelial tube formation (GO:0001838)	9	0.01
embryonic morphogenesis (GO:0048598)	80	0.00
embryonic organ development (GO:0048568)	56	0.00
embryonic organ morphogenesis (GO:0048562)	48	0.00
embryonic skeletal system development (GO:0048706)	28	0.00
embryonic skeletal system morphogenesis (GO:0048704)	27	0.00
endocardial progenitor cell migration to the midline involved in heart field formation (GO:0003262)	4	0.04
epithelial cell differentiation (GO:0030855)	28	0.00
epithelial tube morphogenesis (GO:0060562)	26	0.04
epithelium development (GO:0060429)	87	0.00
establishment of mitotic spindle localization (GO:0040001)	8	0.00
establishment of mitotic spindle orientation (GO:0000132)	8	0.00
establishment of spindle localization (GO:0051293)	9	0.00
establishment of spindle orientation (GO:0051294)	9	0.00
eye development (GO:0001654)	41	0.03
eye morphogenesis (GO:0048592)	28	0.00
gamma-aminobutyric acid signalling pathway (GO:0007214)	6	0.03
generation of neurons (GO:0048699)	112	0.00
growth (GO:0040007)	29	0.03
head development (GO:0060322)	46	0.01

heart development (GO:0007507)	51	0.00
hindbrain development (GO:0030902)	17	0.03
lateral line development (GO:0048882)	17	0.00
lateral line system development (GO:0048925)	18	0.01
localization (GO:0051179)	253	0.00
localization of cell (GO:0051674)	72	0.00
locomotion (GO:0040011)	103	0.00
macromolecule biosynthetic process (GO:0009059)	77	0.04
macromolecule metabolic process (GO:0043170)	295	0.03
macromolecule modification (GO:0043412)	166	0.03
mechanosensory lateral line system development (GO:0048881)	13	0.03
mesenchymal cell development (GO:0014031)	22	0.00
mesenchymal cell differentiation (GO:0048762)	25	0.00
mesenchyme development (GO:0060485)	31	0.00
microtubule cytoskeleton organization involved in mitosis (GO:1902850)	11	0.02
mitotic cell cycle process (GO:1903047)	26	0.04
modulation of chemical synaptic transmission (GO:0050804)	21	0.00
morphogenesis of an epithelium (GO:0002009)	42	0.02
morphogenesis of embryonic epithelium (GO:0016331)	11	0.02
motor neuron migration (GO:0097475)	5	0.04
movement of cell or subcellular component (GO:0006928)	117	0.00
multicellular organism development (GO:0007275)	354	0.00
multicellular organismal process (GO:0032501)	391	0.00
negative chemotaxis (GO:0050919)	9	0.05
negative regulation of axon extension (GO:0030517)	9	0.05
negative regulation of axonogenesis (GO:0050771)	9	0.05
negative regulation of biological process (GO:0048519)	147	0.02
negative regulation of canonical Wnt signalling pathway (GO:0090090)	11	0.02
negative regulation of cell communication (GO:0010648)	45	0.00
negative regulation of cell growth (GO:0030308)	10	0.03
negative regulation of cellular process (GO:0048523)	137	0.02
negative regulation of developmental process (GO:0051093)	24	0.03
negative regulation of growth (GO:0045926)	10	0.05
negative regulation of multicellular organismal process (GO:0051241)	25	0.03
negative regulation of response to stimulus (GO:0048585)	50	0.00

negative regulation of signal transduction (GO:0009968)	42	0.00
negative regulation of signalling (GO:0023057)	45	0.00
nervous system development (GO:0007399)	169	0.00
neural crest cell development (GO:0014032)	22	0.00
neural crest cell differentiation (GO:0014033)	23	0.00
neural crest cell migration (GO:0001755)	20	0.00
neural tube development (GO:0021915)	19	0.00
neural tube formation (GO:0001841)	9	0.01
neurogenesis (GO:0022008)	118	0.00
neuron development (GO:0048666)	79	0.00
neuron differentiation (GO:0030182)	92	0.00
neuron migration (GO:0001764)	8	0.05
neuron projection development (GO:0031175)	64	0.00
neuron projection guidance (GO:0097485)	43	0.00
neuron projection morphogenesis (GO:0048812)	60	0.00
plasma membrane bounded cell projection morphogenesis (GO:0120039)	60	0.00
plasma membrane bounded cell projection organization (GO:0120036)	81	0.00
positive regulation of biological process (GO:0048518)	171	0.00
positive regulation of cell migration (GO:0030335)	15	0.02
positive regulation of cell motility (GO:2000147)	16	0.02
positive regulation of cellular component movement (GO:0051272)	17	0.01
positive regulation of cellular process (GO:0048522)	155	0.00
positive regulation of developmental process (GO:0051094)	25	0.05
positive regulation of locomotion (GO:0040017)	16	0.03
positive regulation of nervous system development (GO:0051962)	16	0.02
posterior lateral line development (GO:0048916)	11	0.04
posterior lateral line system development (GO:0048915)	13	0.02
primary neural tube formation (GO:0014020)	6	0.03
regulation of anatomical structure morphogenesis (GO:0022603)	49	0.00
regulation of axon extension (GO:0030516)	15	0.00
regulation of axon extension involved in axon guidance (GO:0048841)	9	0.05
regulation of axon guidance (GO:1902667)	10	0.05
regulation of axonogenesis (GO:0050770)	20	0.00
regulation of biological process (GO:0050789)	557	0.00
regulation of biological quality (GO:0065008)	142	0.00
regulation of biosynthetic process (GO:0009889)	156	0.00
regulation of cell communication (GO:0010646)	131	0.00

regulation of cell development (GO:0060284)	35	0.01
regulation of cell differentiation (GO:0045595)	46	0.02
regulation of cell growth (GO:0001558)	18	0.00
regulation of cell migration (GO:0030334)	25	0.04
regulation of cell morphogenesis (GO:0022604)	26	0.00
regulation of cell morphogenesis involved in differentiation (GO:0010769)	21	0.00
regulation of cell motility (GO:2000145)	26	0.04
regulation of cell projection organization (GO:0031344)	29	0.00
regulation of cell size (GO:0008361)	18	0.00
regulation of cellular biosynthetic process (GO:0031326)	154	0.00
regulation of cellular component movement (GO:0051270)	30	0.01
regulation of cellular component organization (GO:0051128)	69	0.00
regulation of cellular component size (GO:0032535)	26	0.02
regulation of cellular macromolecule biosynthetic process (GO:2000112)	150	0.00
regulation of cellular metabolic process (GO:0031323)	219	0.00
regulation of cellular process (GO:0050794)	528	0.00
regulation of developmental growth (GO:0048638)	17	0.01
regulation of developmental process (GO:0050793)	86	0.00
regulation of extent of cell growth (GO:0061387)	16	0.00
regulation of gene expression (GO:0010468)	167	0.00
regulation of growth (GO:0040008)	19	0.02
regulation of intracellular signal transduction (GO:1902531)	63	0.00
regulation of localization (GO:0032879)	73	0.00
regulation of locomotion (GO:0040012)	29	0.03
regulation of macromolecule biosynthetic process (GO:0010556)	151	0.00
regulation of macromolecule metabolic process (GO:0060255)	218	0.00
regulation of membrane potential (GO:0042391)	26	0.00
regulation of metabolic process (GO:0019222)	234	0.00
regulation of molecular function (GO:0065009)	67	0.01
regulation of multicellular organismal development (GO:2000026)	76	0.00
regulation of multicellular organismal process (GO:0051239)	88	0.00
regulation of nervous system development (GO:0051960)	42	0.00
regulation of neurogenesis (GO:0050767)	34	0.00
regulation of neuron differentiation (GO:0045664)	30	0.00

regulation of neuron projection development (GO:0010975)	25	0.00
regulation of nitrogen compound metabolic process (GO:0051171)	208	0.00
regulation of non-canonical Wnt signalling pathway (GO:2000050)	5	0.05
regulation of nucleic acid-templated transcription (GO:1903506)	136	0.01
regulation of nucleobase-containing compound metabolic process (GO:0019219)	151	0.01
regulation of plasma membrane bounded cell projection organization (GO:0120035)	28	0.00
regulation of primary metabolic process (GO:0080090)	210	0.00
regulation of Ras protein signal transduction (GO:0046578)	26	0.00
regulation of response to stimulus (GO:0048583)	132	0.00
regulation of Rho protein signal transduction (GO:0035023)	21	0.01
regulation of RNA biosynthetic process (GO:2001141)	136	0.01
regulation of RNA metabolic process (GO:0051252)	147	0.01
regulation of signal transduction (GO:0009966)	109	0.00
regulation of signaling (GO:0023051)	132	0.00
regulation of small GTPase mediated signal transduction (GO:0051056)	28	0.00
regulation of synaptic plasticity (GO:0048167)	10	0.01
regulation of transcription, DNA-templated (GO:0006355)	136	0.01
regulation of trans-synaptic signalling (GO:0099177)	21	0.00
response to chemical (GO:0042221)	107	0.04
response to external stimulus (GO:0009605)	78	0.00
response to stimulus (GO:0050896)	347	0.00
semaphorin-plexin signalling pathway (GO:0071526)	10	0.04
sensory organ development (GO:0007423)	57	0.00
sensory organ morphogenesis (GO:0090596)	33	0.00
sensory system development (GO:0048880)	56	0.00
signal transduction (GO:0007165)	239	0.01
signaling (GO:0023052)	265	0.00
skeletal system development (GO:0001501)	39	0.00
skeletal system morphogenesis (GO:0048705)	30	0.00
spindle localization (GO:0051653)	9	0.00
stem cell development (GO:0048864)	22	0.00
stem cell differentiation (GO:0048863)	31	0.00
synapse assembly (GO:0007416)	10	0.01
synapse organization (GO:0050808)	14	0.03
system development (GO:0048731)	325	0.00
taxis (GO:0042330)	53	0.00

tissue development (GO:0009888)	147	0.00
tissue morphogenesis (GO:0048729)	52	0.01
tube development (GO:0035295)	75	0.00
tube morphogenesis (GO:0035239)	55	0.00
vasculature development (GO:0001944)	48	0.00
visual system development (GO:0150063)	41	0.03

Supplementary tables for Chapter 4 - What does drive DNA methylation plasticity? Epigenetic responses of two inbred lines to different rearing environments

Table S4.1. Individual genotype (line), rearing environment and read count information for *Kryptolebias marmoratus* individuals sequenced using Reduced Representation Bisulphite Sequencing (RRBS).

Sample ID	Strain	Environment	Total number of raw reads	Total number of raw reads per replicated group	Total number of clean reads	Total number of clean reads per replicated group	Bisulphite Conversion efficiency (%)	Mapping (%)	Multiple mapping (%)	Unique mapping (%)	CpG methylation (%)	CHG methylation (%)	CHH methylation (%)
PE-D03	DAN	Poor	13100686	35527426	12905073	34347143	99.8	76.8	8.7	68.1	23.2	0.5	0.4
PE-D08	DAN	Poor	12075916		11215314		99.7	73.9	7	66.9	16.8	0.5	0.4
PE-D11	DAN	Poor	10350824		10226756		99.6	56.9	8.4	48.5	28.1	0.5	0.5
RE-D07	DAN	Enriched	12750300	34452793	12625557	34064521	99.5	76.6	12.1	64.5	37.5	0.6	0.5
RE-D08	DAN	Enriched	10126731		9911130		99.7	75.6	12.4	63.2	39.3	0.6	0.5
RE-D13	DAN	Enriched	11575762		11527834		99.7	76.3	13	63.3	52.3	0.6	0.5
PE-R03	R	Poor	8895336	70667255	8787003	70001468	99.7	74.5	11.9	62.6	50.6	0.5	0.5
PE-R04	R	Poor	7979527		7783671		99.5	74.1	9.9	64.2	34.4	0.5	0.4
PE-R08	R	Poor	9302968		9218192		99.5	63.3	10.6	52.7	45.90	0.6	0.5
PE-R09	R	Poor	10897429		10870559		99.3	73.7	14.3	59.4	52	0.6	0.5
PE-R12	R	Poor	15205376		15164433		99.4	74.2	14	60.2	58.2	0.7	0.6
PE-R14	R	Poor	18386619		18177610		99.6	76.6	11.2	65.4	40	0.5	0.5
RE-R05	R	Enriched	19117048	136658204	18973122	135302833	99.7	73	12.5	60.5	39.9	0.5	0.5
RE-R06	R	Enriched	9625042		9602035		99.7	75.8	12.9	62.9	49	0.6	0.5
RE-R07	R	Enriched	17031887		16832764		99.6	74.3	12.7	61.6	43.3	0.6	0.5
RE-R08	R	Enriched	17731929		17373799		99.7	74.7	10.4	64.3	25.1	0.5	0.4
RE-R10	R	Enriched	18721051		18356361		99.7	75.2	11.2	64	32.9	0.5	0.5
RE-R11	R	Enriched	9478011		9464124		99.6	74.8	13.5	61.3	51.9	0.6	0.5
RE-R12	R	Enriched	6929800		6872621		99.6	75.4	12.3	63.1	48.1	0.6	0.5
RE-R13	R	Enriched	14934937		14894697		99.7	79.1	10.3	68.8	33.2	0.5	0.5
RE-R14	R	Enriched	11706829		11634216		99.5	78.9	11.5	67.4	43.7	0.6	0.5
RE-R15	R	Enriched	11381670		11299094		99.7	75.5	11.2	64.3	37.5	0.6	0.5
Total			277305678	277305678	273715965	273715965	99.6	74.05	11.45	62.6	40.13	0.56	0.49

Table S4.2. Microsatellite genotypes (23 microsatellites) amplified according Ellison et al. (2011) for all individuals *Kryptolebias marmoratus* analysed using Reduced Representation Bisulphite Sequencing (RRBS).

ID	Strain	Environment	R3	R10	R1	R34	R93	R92	R4	R11	R25	R5	R37	R17	R35	R9	R23	R18	R19	R38	R22	R86	R103	R30	R112
PE-D03	DAN	Poor	123	239	246	146	169	195	255	187	122	287	332	259	142	226	258	177	139	222	197	217	141	165	217
			123	239	246	146	169	195	255	187	122	287	332	259	142	226	258	177	139	222	197	217	141	165	217
PE-D08	DAN	Poor	123	239	246	146	169	195	255	187	122	287	332	259	142	226	258	177	139	222	197	217	141	165	217
			123	239	246	146	169	195	255	187	122	287	332	259	142	226	258	177	139	222	197	217	141	165	217
PE-D11	DAN	Poor	123	239	246	146	169	195	255	187	122	287	332	259	142	226	258	177	139	222	197	217	141	165	217
			123	239	246	146	169	195	255	187	122	287	332	259	142	226	258	177	139	222	197	217	141	165	217
RE-D13	DAN	Enriched	123	239	246	146	169	195	255	187	122	287	332	259	142	226	258	177	139	222	197	217	141	165	217
			123	239	246	146	169	195	255	187	122	287	332	259	142	226	258	177	139	222	197	217	141	165	217
RE-D07	DAN	Enriched	123	239	246	146	169	195	255	187	122	287	332	259	142	226	258	177	139	222	197	217	141	165	217
			123	239	246	146	169	195	255	187	122	287	332	259	142	226	258	177	139	222	197	217	141	165	217
RE-D08	DAN	Enriched	123	239	246	146	169	195	255	187	122	287	332	259	142	226	258	177	139	222	197	217	141	165	217
			123	239	246	146	169	195	255	187	122	287	332	259	142	226	258	177	139	222	197	217	141	165	217
PE-R03	R	Poor	123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
			123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
PE-R04	R	Poor	123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
			123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
PE-R08	R	Poor	123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
			123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
PE-R09	R	Poor	123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
			123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
PE-R12	R	Poor	123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
			123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203

PE-R14	R	Poor	123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
			123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
RE-R05	R	Enriched	123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
			123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
RE-R06	R	Enriched	123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
			123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
RE-R07	R	Enriched	123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
			123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
RE-R08	R	Enriched	123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
			123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
RE-R10	R	Enriched	123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
			123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
RE-R11	R	Enriched	123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
			123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
RE-R12	R	Enriched	123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
			123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
RE-R13	R	Enriched	123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
			123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
RE-R14	R	Enriched	123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
			123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
RE-R15	R	Enriched	123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203
			123	215	250	146	154	193	247	183	106	311	328	283	142	226	258	177	139	194	189	215	141	161	203

Table S4.3. Linear model of principal components scores for mangrove killifish (a) 1064 DMCs and (b) 194 DMRs found between environments (poor, enriched) and genotypes (R, DAN).

	t-values	**Prop. of variance (%)**	**df**	**p-value**
(*a*) 1064 DMCs				
PC1 scores				
Genotype	1.62	92.17	1	<0.001
Environment	-22.53	6.46	1	<0.001
Genotype x Environment	-24.09	0.28	1	0.04
PC2 scores				
Genotype	5.25	6.65	1	<0.001
Environment	-3.32	73.59	1	<0.001
Genotype x Environment	-6.87	13.24	1	<0.001
(*b*)194 DMRs				
PC1 scores				
Genotype	0.65	79.01	1	<0.001
Environment	14.68	16.60	1	<0.001
Genotype x Environment	-13.03	2.22	1	<0.001
PC2 scores				
Genotype	-5.07	9.83	1	<0.001
Environment	5.4	27.31	1	<0.001
Genotype x Environment	11.19	9.49	1	<0.001

Table S4.4. Number of DMCs found in the comparisons between enriched and poor environments using three different subsets of six individuals (three on each environment) from R genotype (original subset with 16 individuals). DMCs were found using logistic regression q< 0.01 and |ΔM|>20%, and t.test p <0.01, with the final list of composed by the DMCs shared by both approaches. This analysis identified similar numerical differences in DMCs between lines, suggesting that the differences in numbers of DMCs on the original dataset were not due to differences in sample size.

Individuals	**Linear regression**	**t-test**	**Both**
Subset 1			
PE-R03, PE-R08, PE-R12, RE-R05, RE-R06, RE-R07	6183	5777	2977
Subset 2			
PE-R04, PE-R09, PE-R14, RE-R08, RE-R10, RE-R11	4756	4930	2237
Subset 3			
PE-R03, PE-R09, PE-R12, RE-R11, RE-R12, RE-R13	4007	4664	1920

Table S4.5. Methylation differences averaged (in percentage) for differentially methylated cytosines (DMRs) between environments (poor, enriched) and shared between genotypes (DAN, R) which overlap annotated genes (reference genome *Kryptolebias marmoratus* NCBI ASM164957v1). Epiallele classification (pure or facilitated) was defined following Richards (2006). Positive and negative values represent increased and decreased methylation towards enriched and poor environments, respectively Q-value is the p-value adjusted for the False Discovery Rate (FDR=0.05).

Gene symbol	**Entrez gene name**	**Epiallele class**	**Meth diff DAN**	***Q*-value**	**Meth diff R**	***Q*-value**
nedd9	neural precursor cell expressed, developmentally down-regulated 9	F	23.99	<0.001	-21.19	<0.001
nptn	neuroplastin	F	20.03	<0.001	-41.11	<0.001
sorcs2	sortilin-related VPS10 domain containing receptor 2	F	37.04	<0.001	-31.03	<0.001
uhrf2	ubiquitin like with PHD and ring finger domains 2	F	42.24	<0.001	-33.13	<0.001

Table S4.6. Centrality measures to molecular network reconstructed using 23 DMCs between environments (enriched and poor) and shared between genotype (DAN, R) using NetworkAnalyzer. Origin represents whether the gene was on the input file or it was suggested by GeneMANIA as a neighbouring gene on the network. Degree represents the number of connections of a given gene. Closeness represent average distance of the all the genes in the network to a certain gene. Radiality represents the easiness of reaching any gene from certain gene.

Gene symbol	Origin	Degree	Closeness	Radiality
ACVR2A	Input	17	0.58	0.82
ADGRL3	Input	21	0.62	0.85
CALD1	Input	15	0.57	0.81
COL25A1	Input	19	0.62	0.85
DMAP1	Input	3	0.40	0.62
FOXP4	Input	5	0.46	0.71
GPC5	Input	19	0.62	0.85
MICAL2	Input	15	0.58	0.82
MIPOL1	Input	3	0.44	0.68
MYC	Input	12	0.56	0.80
NECAB2	Input	7	0.49	0.74
NEO1	Input	5	0.48	0.73
NUDCD1	Input	6	0.50	0.75
PTPRN2	Input	5	0.49	0.74
RAMP3	Input	3	0.44	0.68
RYR3	Input	21	0.64	0.86
SORCS2	Input	21	0.66	0.87
SPTBN1	Input	11	0.54	0.79
TRIT1	Input	14	0.56	0.80
TRMT44	Input	4	0.44	0.68
ZEB2	Input	15	0.60	0.83
ZNF516	Input	9	0.51	0.76
ZRANB3	Input	4	0.46	0.71
ARNT2	Neighbouring	6	0.48	0.73
C20orf194	Neighbouring	4	0.45	0.70
INPP5A	Neighbouring	5	0.48	0.73
ITGA6	Neighbouring	8	0.53	0.77
JAML	Neighbouring	16	0.61	0.84
LAMP5	Neighbouring	8	0.50	0.75
MRPS33	Neighbouring	19	0.64	0.86
NOVA1	Neighbouring	16	0.57	0.81
NPEPL1	Neighbouring	8	0.53	0.78
PDE8A	Neighbouring	13	0.54	0.79
POFUT2	Neighbouring	6	0.47	0.72
RALGAPA2	Neighbouring	1	0.34	0.52
RNF24	Neighbouring	7	0.50	0.75
SLC12A6	Neighbouring	6	0.46	0.71

SORCS3	Neighbouring	10	0.52	0.77
SULF1	Neighbouring	6	0.49	0.74
SVIL	Neighbouring	15	0.56	0.80
TIAM1	Neighbouring	18	0.60	0.83
TMTC2	Neighbouring	16	0.58	0.82
VCAN	Neighbouring	12	0.55	0.79
Mean (SD)		**10.55 ± 5.89**	**0.53 ± 0.06**	**0.77 ± 0.06**

Table S4.7. Gene ontology (GO) terms for biological processes and pathways for the most connected (> 10 connections) DMCs between environments shared by genotypes within GeneMANIA network. Gene ortholog was compared with zebrafish (*Danio rerio*) using PANTHER (Mi et al., 2016).

Major gene ontology terms level	**Specific functional level**	**Genes involved (symbols)**
(a) Biological process		
Metabolic process (GO:008152)	Transcription by RNA polymerase II (GO:006366)	myc
	Lipid metabolic process (GO:006629)	sorcs2
	tRNA modification (GO:0006400)	trit1
Localization (GO:0051179)	Intracellular protein transport (GO:0006886)	sorcs2
	Lipid transport (GO:0006869)	sorcs2
	Receptor-mediated endocytosis (GO:0006898)	sorcs2
	Sequestering of calcium ion (GO:0051208)	ryr3
Cellular process (GO:0009987)	Cell cycle (GO:0007049)	myc
Biological regulation (GO:0065007)	Regulation of transcription by RNA polymerase II (GO:006366)	myc
(b) Pathways		
	Beta1 adrenergic receptor signalling pathway (P04377)	ryr3
	Beta1 adrenergic receptor signalling pathway (P04378)	ryr3
	CCKR signalling map (P06959)	myc, ryr3
	Interleukin signalling pathway (P00036)	myc

	Oxidative stress response (P00046)	myc
	PDGF signalling pathway (P00047)	myc
	Wnt signalling pathway (P00057)	myc
	p53 pathway feedback loops 2 (P04398)	Myc

Supplementary tables for Chapter 5 - Environmental enrichment induces epigenetic and behavioural parental effects in a self-fertilising fish

Table S5.1. Individuals used in the experiment. Asterisks represent individuals sequenced for methylation patterns using reduced representation bisulphite sequencing.

ID	Generation	Parental Enviroment	Current Environment	Parent
1ST-RE-R01	F0		Enriched	
1ST-RE-R02	F0		Enriched	
1ST-RE-R05*	F0		Enriched	
1ST-RE-R06*	F0		Enriched	
1ST-RE-R07*	F0		Enriched	
1ST-RE-R08*	F0		Enriched	
1ST-RE-R10*	F0		Enriched	
1ST-RE-R11*	F0		Enriched	
1ST-RE-R12*	F0		Enriched	
1ST-RE-R13*	F0		Enriched	
1ST-RE-R14*	F0		Enriched	
1ST-RE-R15*	F0		Enriched	
1ST-RE-R16	F0		Enriched	
1ST-RE-R17	F0		Enriched	
1ST-PE-R01	F0		Poor	
1ST-PE-R03*	F0		Poor	

1ST-PE-R04*	F0		Poor	
1ST-PE-R05	F0		Poor	
1ST-PE-R07	F0		Poor	
1ST-PE-R08*	F0		Poor	
1ST-PE-R09*	F0		Poor	
1ST-PE-R10	F0		Poor	
1ST-PE-R11	F0		Poor	
1ST-PE-R12*	F0		Poor	
1ST-PE-R13	F0		Poor	
1ST-PE-R14*	F0		Poor	
1ST-PE-R15	F0		Poor	
2nd-RE-R02*	F1	Enriched	Enriched	1st-RE-R11
2nd-RE-R03*	F1	Enriched	Enriched	1st-RE-R07
2nd-RE-R04*	F1	Enriched	Enriched	1st-RE-R14
2nd-RE-R07	F1	Enriched	Enriched	1st-RE-R09
2nd-RE-R09	F1	Enriched	Enriched	1st-RE-R13
2nd-RE-R11*	F1	Enriched	Enriched	1st-RE-R11
2nd-RE-R13*	F1	Enriched	Enriched	1st-RE-R07
2nd-RE-R16	F1	Enriched	Enriched	1st-RE-R07
2nd-PE-R03*	F1	Poor	Poor	1st-PE-R03
2nd-PE-R05*	F1	Poor	Poor	1st-PE-R08
2nd-PE-R08	F1	Poor	Poor	1st-PE-R03
2nd-PE-R12	F1	Poor	Poor	1st-PE-R01
2nd-PE-R15*	F1	Poor	Poor	1st-PE-R03
2nd-PE-R01	F1	Enriched	Poor	1st-RE-R09
2nd-PE-R02*	F1	Enriched	Poor	1st-RE-R11
2nd-PE-R04*	F1	Enriched	Poor	1st-RE-R11

2nd-PE-R06	F1	Enriched	Poor	1st-RE-R13
2nd-PE-R07*	F1	Enriched	Poor	1st-RE-R11
2nd-PE-R09	F1	Enriched	Poor	1st-RE-R14
2nd-PE-R10	F1	Enriched	Poor	1st-RE-R12
2nd-PE-R11	F1	Enriched	Poor	1st-RE-R12
2nd-PE-R13	F1	Enriched	Poor	1st-RE-R07
2nd-PE-R16	F1	Enriched	Poor	1st-RE-R07
2nd-RE-R05	F1	Poor	Enriched	1st-PE-R01
2nd-RE-R08	F1	Poor	Enriched	1st-PE-R03
2nd-RE-R12*	F1	Poor	Enriched	1st-PE-R03
2nd-RE-R14*	F1	Poor	Enriched	1st-PE-R08
2nd-RE-R15*	F1	Poor	Enriched	1st-PE-R03

Table S5.2. Read counts, proportions of bisulphite conversion, mapping and methylation contexts for the different experimental groups (1st generation = F0, 2nd generation = F1) sequenced using reduced representation bisulphite sequencing.

Sample ID	Generation	Parent	Previous Environment	Current Environment	Total number of raw reads	Total number of clean reads	Bisulphite Conversion efficiency (%)	Mapping (%)	Multiple mapping (%)	Unique mapping (%)	CpG methylation (%)	CHG methylation (%)	CHH methylation (%)
1st-PE-R03	1st	NA	NA	Poor	8895336	8787003	99.7	74.5	11.9	62.6	50.6	0.5	0.5
1st-PE-R04	1st	NA	NA	Poor	7979527	7783671	99.5	74.1	9.9	64.2	34.4	0.5	0.4
1st-PE-R08	1st	NA	NA	Poor	9302968	9218192	99.5	63.3	10.6	52.7	45.90	0.6	0.5
1st-PE-R09	1st	NA	NA	Poor	10897429	10870559	99.3	73.7	14.3	59.4	52	0.6	0.5
1st-PE-R12	1st	NA	NA	Poor	15205376	15164433	99.4	74.2	14	60.2	58.2	0.7	0.6
1st-PE-R14	1st	NA	NA	Poor	18386619	18177610	99.6	76.6	11.2	65.4	40	0.5	0.5
1st-RE-R05	1st	NA	NA	Enriched	19117048	18973122	99.7	73	12.5	60.5	39.9	0.5	0.5
1st-RE-R06	1st	NA	NA	Enriched	9625042	9602035	99.7	75.8	12.9	62.9	49	0.6	0.5
1st-RE-R07	1st	NA	NA	Enriched	17031887	16832764	99.6	74.3	12.7	61.6	43.3	0.6	0.5
1st-RE-R08	1st	NA	NA	Enriched	17731929	17373799	99.7	74.7	10.4	64.3	25.1	0.5	0.4
1st-RE-R10	1st	NA	NA	Enriched	18721051	18356361	99.7	75.2	11.2	64	32.9	0.5	0.5
1st-RE-R11	1st	NA	NA	Enriched	9478011	9464124	99.6	74.8	13.5	61.3	51.9	0.6	0.5
1st-RE-R12	1st	NA	NA	Enriched	6929800	6872621	99.6	75.4	12.3	63.1	48.1	0.6	0.5
1st-RE-R13	1st	NA	NA	Enriched	14934937	14894697	99.7	79.1	10.3	68.8	33.2	0.5	0.5
1st-RE-R14	1st	NA	NA	Enriched	11706829	11634216	99.5	78.9	11.5	67.4	43.7	0.6	0.5
1st-RE-R15	1st	NA	NA	Enriched	11381670	11299094	99.7	75.5	11.2	64.3	37.50	0.6	0.5
2nd-PE-R03	2nd	1st-PE-R03	Poor	Poor	9436612	9434347	99.2	72.3	14	58.3	60.7	0.7	0.6
2nd-PE-R05	2nd	1st-PE-R08	Poor	Poor	14169788	14165283	99.1	73.9	15.8	58.1	65.6	0.7	0.6
2nd-PE-R15	2nd	1st-PE-R03	Poor	Poor	8264728	8262148	99.7	74.4	15.6	58.8	58.5	0.7	0.6
2nd-PE-R02	2nd	1st-RE-R11	Enriched	Poor	25246913	25236032	99.7	73.6	13.8	59.8	56.9	0.6	0.5
2nd-PE-R04	2nd	1st-RE-R11	Enriched	Poor	9598838	9595582	99.5	74	16.7	57.3	56.7	0.7	0.6
2nd-PE-R07	2nd	1st-RE-R11	Enriched	Poor	10307222	10300566	99.6	72.4	16.2	56.2	65.7	0.7	0.5
2nd-RE-R02	2nd	1st-RE-R11	Enriched	Enriched	9884878	9880493	99.4	74.2	13.1	61.1	67.2	0.7	0.5
2nd-RE-R03	2nd	1st-RE-R07	Enriched	Enriched	11806281	11798600	99.4	72.6	16.3	56.3	57.5	0.6	0.5
2nd-RE-R04	2nd	1st-RE-R14	Enriched	Enriched	10681971	10677008	99.2	72.7	12.9	59.8	63.9	0.7	0.6
2nd-RE-R11	2nd	1st-RE-R11	Enriched	Enriched	16291112	16267064	99.5	71.5	15.6	55.9	55	0.6	0.6
2nd-RE-R13	2nd	1st-RE-R07	Enriched	Enriched	7866499	8739487	99.8	72.2	16.6	55.6	58	0.7	0.6
2nd-RE-R12	2nd	1st-PE-R03	Poor	Enriched	15937636	15912236	99.7	75.9	12.1	63.8	43.1	0.5	0.5
2nd-RE-R14	2nd	1st-PE-R08	Poor	Enriched	6851691	6848684	99.3	72.8	15	57.8	59.9	0.7	0.7
2nd-RE-R15	2nd	1st-PE-R03	Poor	Enriched	14676322	14667955	99.1	79.1	7.7	71.4	37.6	0.4	0.4
Total					378345950	377089786	99.5	74.2	13.1	61.1	54.2	0.6	0.6

Table S5.3. Results of the most plausible linear models using a multi-model averaging approach for (*a*) cortisol levels (pg/ml) and (*b*) basal metabolic rates (mg O_2 g^{-1} min^{-1}) and (*c*-*e*) behavioural metrics for the parental individuals (F0). Models are ranked according to their corrected Akaike Information Criterion (AICc), the difference with the best fitting model (ΔAICc) and the Akaike weight (Wi), which represents the ratio between the weights of the best and competing models. Only models within two AICc units are shown.

	df	**z/t-value**	***P*-value**	**AICc**	**ΔAICc**	**Wi**
(a) Cortisol levels						
Model 1				-24.60	0.00	0.66
Body weight	1	4.22	0.001			
Model 2				-22.73	1.87	0.26
Body weight	1	2.86	<0.001			
Environment	1	7.38	0.03			
(b) Basal metabolic rate						
Model 1				-16.51	0.00	0.58
Cortisol levels	1	9.85	<0.001			
(c) Activity						
Model 1				2184.97	0.00	0.74
Body weight	1	31.58	0.01			
Environment	1	34.98	<0.001			
Body weight x Environment	1	38.76	0.02			

Table S5.4. Results of the most plausible generalized linear models (quasipoisson for count data and Gaussian for continuous data) indicated by multi-model averaging for activity and number of contacts in the offspring (F1). Models are ranked according to the corrected Akaike Information Criterion (AICc), the difference with the best fitting model (ΔAICc), and the Akaike weight (Wi) which represents the ratio between the weights of the best and competing models. Only models within two AICc units are shown.

	Df	**z-value**	***P*-value**	**AICc**	**ΔAICc**	**Wi**
Activity (number of crosses among zones)						
Model 1				871.59	0.00	0.50
Parental environment	1	-1.22	0.22			
Offsrping environment	1	0.703	0.48			
Parental activity	1	2.61	0.009			
Body weight	1	-1.32	0.18			
Offspring environment x Parental activity	1	0.38	0.70			
Parental activity x Body weight	1	1.18	0.18			

V.I. APPENDIX II: RELEVANT PUBLISHED MANUSCRIPTS

I either leaded or collaborated during the period of PhD candidature (October 2015 to October 2019) which are not part but directly relevant to the overall topics addressed in the current thesis. Briefly, Berbel-Filho et al. (2016) represents the first report of a male of *Kryptolebias hermaphroditus*, despite the species being described as composed only of selfing hermaphrodites. Tatarenkov et al. (2017) revealed deep genetic structure across the wide geographical distribution (29°N to 23°S) of the "*K. marmoratus* species complex", revealed a previously unknown genetic clade in the Caribbean, that despite its geographical proximity with *K. marmoratus*, it is phylogenetically more related to *K. hermaphroditus*.

1) Berbel-Filho, W. M., H. M. V. Espirito-Santo, and S. M. Q. Lima. 2016. First record of a male of *Kryptolebias hermaphroditus* Costa, 2011 (Cyprinodontiformes: Cynolebiidae). Neotropical Ichthyology 14:e160024.

2) Tatarenkov, A., S. M. Q. Lima, R. L. Earley, **W. M. Berbel-Filho**, F. B. M. Vermeulen, D. S. Taylor, K. Marson et al. 2017. Deep and concordant subdivisions in the self-fertilizing mangrove killifishes (*Kryptolebias*) revealed by nuclear and mtDNA markers. Biological Journal of the Linnean Society 122:558-578.

Berbel-Filho et al. (2016):

Neotropical Ichthyology, 14(3): e160024, 2016
DOI: 10.1590/1982-0224-20160024

Journal homepage: www.scielo.br/ni
Published online: 30 September 2016 (ISSN 1982-0224)

First record of a male of *Kryptolebias hermaphroditus* Costa, 2011 (Cyprinodontiformes: Cynolebiidae)

Waldir M. Berbel-Filho[1,2], Helder M. V. Espírito-Santo[3] and Sergio M. Q. Lima[1]

During an ichthyological survey in September 2015 at the Ceará-Mirim River estuary, Rio Grande do Norte State, northeastern Brazil, we collected a male of *Kryptolebias hermaphroditus*, *a* cynolebiid species that had been previously described as containing exclusively self-fertilizing hermaphrodites. This is the first record of a male in this species, over 140 years after the discovery of the mangrove rivulid species from Brazil. Our discovery reinforces the need for more studies in *K. hermaphroditus*, as well as the potential of this species as a model for evolutionary studies due to its unique mating system.

Durante uma amostragem ictiológica em setembro de 2015 no estuário do rio Ceará-Mirim, Estado do Rio Grande do Norte, Nordeste do Brasil, nós coletamos um macho de *Kryptolebias hermaphroditus*, uma espécie de rivulídeo que foi descrita como contendo apenas hermafroditas auto-fertilizantes. Este é o primeiro registro de um macho dessa espécie, mais de 140 anos depois da descoberta das espécies de rivulídeos de manguezais do Brasil. Nossa descoberta reforça a necessidade de mais estudos em *K. hermaphroditus*, assim como, o potencial desta espécie como um modelo para estudos evolutivos devido ao seu sistema reprodutivo único.

Keywords: Hermaphroditism, Mangrove killifish, Mid-Northeastern Caatinga ecoregion, Mixed-mating system, Sexual dimorphism.

Introduction

Although the overall taxonomy of *Kryptolebias* Costa, 2004 is confusing (Costa, 2006), the *K. marmoratus* species group is well supported by both molecular (Tartarenkov *et al*., 2009) and morphological data (Costa *et al*., 2010). The species comprising this group, *K. marmoratus* (Poey, 1880), *K. ocellatus* (Hensel, 1868) and *K. hermaphroditus* Costa, 2011, inhabit estuarine areas associated with mangroves, ranging from southeastern United States of America to southern Brazil (Costa, 2011). Besides this unusual brackish habitat, these species are also unique among aplocheiloid killifishes due to the absence of females, which are replaced by hermaphrodites (Costa *et al*., 2010).

Indeed, self-fertilizing hermaphroditism is the dominant mode of reproduction in *Kryptolebias marmoratus* distributed from Florida to Venezuela (Tatarenkov *et al*., 2011) and *K. hermaphroditus* occurring in the Atlantic Forest mangroves of Brazil (Lira *et al*., 2015). This clade represents the only occurrence of self-fertilizing hermaphroditism among vertebrates (Tartarenkov *et al*., 2009; Costa *et al*., 2010). Because of its reproductive mixed-mated system, *K. marmoratus* has been a popular model species for embryological, physiological, behavioral and genetic studies (Avise & Tatarenkov, 2015). While males are rare in *K. marmoratus*, they are considered to be absent in *K. hermaphroditus* (Costa, 2011; Tatarenkov *et al*., 2011), despite this species having been known at least since 1868 (previously known as *K. ocellatus*) and numerous museum and unpreserved specimens have been examined (Costa, 2011). Due to the absence of males in fish collections, and following a re-examination of the holotype of *K. ocellatus*, the monomorphic hermaphrodite species from Brazil was described as *K. hermaphroditus* (Costa, 2011).

Hermaphrodites of both species (*K. marmoratus* and *K. hermaphroditus*) typically have a black ocellus on the caudal peduncle and a dark-grey body side (Soto & Noakes, 1994; Costa, 2011). In *K. marmoratus*, males exhibit a bright orange coloration on the body flank and fins, an absent or poorly-defined ocellus on the caudal peduncle, and black margins on the caudal fins (Davis *et al*., 1990; Soto & Noakes, 1994; Costa, 2011).

[1]Laboratório de Ictiologia Sistemática e Evolutiva, Departamento de Botânica e Zoologia, Universidade Federal do Rio Grande do Norte, Av. Senador Salgado Filho 3000, 56078-970 Natal, RN, Brazil. (SMQL) smaialima@gmail.com, (WMBF) waldirmbf@hotmail.com (corresponding author)
[2]Department of BioSciences, College of Science, Swansea University, SA2 8PP, Swansea, Wales, United Kingdom
[3]Post-Doctoral Program, Instituto Nacional de Pesquisas da Amazônia, Av. André Araújo, 2936, Petrópolis, 69067-375 Manaus, AM, Brazil. espiritosantohm@gmail.com

In a recent collection in the Ceará-Mirim River estuary, we found a *Kryptolebias hermaphroditus* individual with a male phenotype. Considering that this species has been described as being completely composed of hermaphrodites, the main objective of the present study is to record and describe the first occurrence of a male *K. hermaphroditus*.

Material and Methods

In September 2015, during a nocturnal exploratory fieldwork being conducted with the objective of designing an ecological study on *Kryptolebias hermaphroditus* at the Ceará-Mirim River estuary (05°40'25.9"S 35°14'14.5"W, see Lira *et al*., 2015 for a map and detailed description of the area), we observed a bright orange individual of *Kryptolebias* (Fig. 1). This specimen together with specimens exhibiting a more typical hermaphroditic appearance, were captured by hand and with the help of a hand sieve (50 x 60 cm, 1 mm mesh), under permit #30532-1/2011 issued by ICMBio/SISBIO. The fishes were euthanized using clove oil, fixed in formalin solution, transferred to an 70% alcohol solution and then deposited in the ichthyological collection of Universidade Federal do Rio Grande do Norte (UFRN 3774, 8 ex.). Some other individuals, as well as a small fin-clip (TIUFRN 3481; tissue collection of UFRN), were directly stored in 96% ethanol for molecular analysis.

Measurements and counts for the orange individual follow Costa (1995). *Kryptolebias* species identification was carried out using the key proposed by Costa (2009), and later confirmed using Costa (2011). To identify male traits, we used the information described for the closely related species *K. marmoratus* (Davis *et al*., 1990; Soto & Noakes, 1994; Costa, 2011).

Results

All specimens were identified as *Kryptolebias hermaphroditus* on the basis of having 1-4 teeth in the vomer and a color pattern with silvery or pale golden spots on the dorsal part of the flank. The male individual had bright orange coloration on the body flank and fins (mainly dorsal, anal and caudal), and black margins on the anal and caudal fins (Fig. 1), as in males of *K. marmoratus*.

Comparing male and hermaphroditic phenotypes of the captured specimens, both have a dark ocellus on the caudal peduncle; however, the ocellus is faint on the male specimen. The hermaphrodites consistently exhibit a brown or grey flank coloration, as well as hyaline unpaired fins with pale grey dots on the basal portion, whereas the male individual exhibits unpaired fins with orange coloration (Fig. 2).

Description. Morphometric data of male (30.1 mm SL) in Table 1. Urogenital papilla cylindrical in male, pocket-like shape in hermaphrodites (Fig. 3). Dorsal fin rounded. Anal fin sub-trapezoidal, distal margin slightly damaged; anal-fin rays 4 and 5 longer than other anal-fin rays. Caudal fin oval, deeper than long. Pectoral fin short and rounded. Pelvic fin elliptical. Pelvic-fin bases medially separated by short interspace. Dorsal-fin origin at vertical between 9th and 10th anal-fin rays. Dorsal-fin rays 8; anal-fin rays 11; caudal-fin rays 30; pectoral-fin rays 13; pelvic-fin rays 6.

Scales small, cycloid. Frontal squamation E-patterned. Longitudinal series of scales 47; transverse series of scales 13; scale rows around caudal peduncle 25. Contact organs absent.

Fig. 1. *Kryptolebias hermaphroditus*, male, UFRN 3774, 30.1 mm SL (three days after sampling); Ceará-Mirim River estuary, Extremoz, Rio Grande do Norte, Brazil.

Fig. 2. *Kryptolebias hermaphroditus*, male (below) 30.1 mm SL, hermaphrodite (above) 35.5 mm SL, UFRN 3774, Ceará-Mirim River estuary, Extremoz, Rio Grande do Norte, Brazil.

Table 1. Morphometric data of the *Kryptolebias hermaphroditus*, male (UFRN 3774).

Standard length (mm)	30.16
Percents of standard length	
Body depth	17.11
Caudal-peduncle depth	13.23
Predorsal length	73.18
Prepelvic length	62.10
Length of dorsal-fin base	12.33
Length of anal-fin base	15.78
Caudal-fin length	25.86
Pectoral-fin length	17.90
Pelvic-fin length	6.63
Head length	25.10
Percents of head length	
Head depth	59.31
Head width	71.07
Snout length	29.33
Lower-jaw length	23.51
Eye diameter	26.55

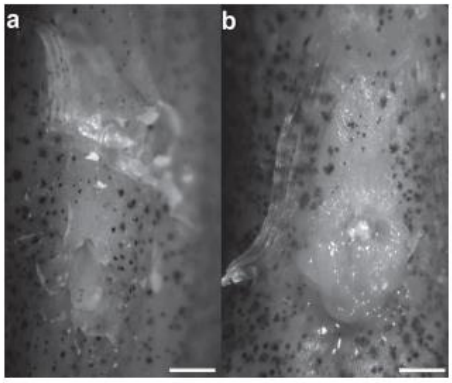


Fig. 3. Urogenital papillae of *Kryptolebias hermaphroditus* in ventral view; UFRN 3374: a, cylindrical shape in male, 30.1 mm SL; b, pocket shape in hermaphrodite, 35.5 mm SL. Scale bars = 1 mm.

Coloration. Flank dark grey on anterior and dorsal portion of head, from greyish to orange on posterior and ventral portions. Side of the body intensely pigmented with faint dark grey stripe between postorbital region and caudal-fin base, and small irregular somewhat zigzag pattern composed of small orange spots more concentrated on anterior half of body side. Inconspicuous dark grey humeral blotch; rounded dark spot with faint yellowish grey margin on dorsal end of caudal peduncle. Side of head light brown on dorsal portion, pale golden with dark grey dots ventrally. Jaws grey. Iris brown, with narrow pale yellow margin around pupil. Dorsal, anal and caudal fins hyaline with dark orangish-brown on basal and medial portions of fins; black stripe on distal anterior margin of anal and lower portion of caudal fin. Pectoral and pelvic fins hyaline. Bright orange coloration faded a little after sampling, probably due to stress and non-natural conditions (Figs. 1-2).

Ecological notes. The sampling locality had clear water and the fishes could be observed using flashlights. The *Kryptolebias hermaphroditus* male was captured in an ephemeral pool at higher elevations among the estuary's mangroves. This pool had at least four fiddler crabs (*Uca maracoani* Latreille, 1802) burrow openings and, prior to our approach, the male fish was resting on the water's surface near one of these burrows. During our sampling attempts, the fish repeatedly hid and re-appeared among the crab burrows, appearing periodically in different openings until it was caught. Time between reappearances was about one minute.

Large (23.2-35.5 mm SL) hermaphrodites were also collected in the same area as the male, in shallow pools formed in the edge of the mangroves; in the inner and deeper pools, *Poecilia vivipara* Bloch & Schneider, 1801 was the dominant species. *Guavina guavina* (Valenciennes, 1837), potential predator of *Kryptolebias hermaphroditus*, were also active during the night and observed entering the same large crab burrows. Salinity in these pools, previously recorded in January 2015 at the same site, was 34-36 ppm during the low tide.

Discussion

Over the 140 years, since the description of *Kryptolebias ocellatus* (as was long called hermaphroditic species who was later described as *K. hermaphroditus*) from a mangrove in Rio the Janeiro, no males of this hermaphroditic species from the Brazilian coast have ever been found (Costa, 2011). However, potential miswritten on previous literature (Costa, 2006; Lira *et al*., 2015) needs to be clarified for a better understanding of the issue. Both studies used the subjective term "rare" while discussing the existence of males in *K. hermaphroditus.* Costa (2006) had already mentioned two reports of potential *K. hermaphroditus* males based on pictures and brief descriptions of live specimens with light grey and dark orange spots on the flanks, grey humeral and caudal peduncle spots and dark grey to black distal zones of unpaired fins. However, this author recognized that this material was not preserved for study in any scientific collection and needed posterior confirmation. Posteriorly, Costa (2011) described *K. hermaphroditus* as only composed by hermaphrodites, discarding the relevance of the two previous unconfirmed reports. In both cases male phenotypes were reported for Rio de Janeiro, where *K. ocellatus* (males and hermaphrodites, previously known as K. caudomarginatus Seegers) occurs in syntopy with *K. hermaphroditus* making these records more doubtful. The potential records described above, the presence of males on populations of *K. marmoratus* (Avise & Tatarenkov, 2015), as well as the cryptic and complex mangrove microhabitats inhabited by *K. hermaphroditus*, may have led Costa (2006) and Lira *et al*. (2015) to not fully discard the presence of males of *K. hermaphroditus*, despite extensive sampling and no reliable scientific evidence of males have been made available up to that moment.

The recently collected specimen of *Kryptolebias hermaphroditus* with male appearance, described herein, matches the description of the putative *K. hermaphroditus* males mentioned by Costa (2006) and *K. marmoratus* males (Davis *et al*., 1990; Soto & Noakes, 1994; Costa, 2011). Although gonadal histological evidence was not surveyed in order to maintain the unique specimen integrity, its morphological characters support this first documentation of a male *K. hermaphroditus*. This evidence includes orange body and fins, faded ocellus, black margins of anal and caudal fins and cylindrical urogenital papillae (Costa, 2011). Once the color pattern of males is the most conspicuous character for differentiating cynolebiid species (Costa, 2003), the discovery of a *K. hermaphroditus* male may reveal important morphological differences between this species and *K. marmoratus*.

Regardless the presence of an ocellus on the caudal peduncle, it appeared faded in comparison to that of hermaphrodite individuals. According to Soto & Noakes (1994), the most distinctive characters attributed to *Kryptolebias marmoratus* males are: a bright orange overall coloration and the absence of a distinct caudal ocellus. However, they also considered that individuals with intermediate phenotypes (with orange color and a faded ocellus) could also be males. Davis *et al*. (1990) had previously mentioned that it is not possible yet to distinguish primary and secondary males, the latter of which result from sex change of functional hermaphrodites into males due to environmental conditions (Harrington, 1971). Huber (1992) cited the presence of ocellus in secondary males of *K. ocellatus*, which agrees with our specimen description. However, as discussed above, this report was not afterwards confirmed.

Primary males of *Kryptolebias marmoratus* can be readily produced in laboratories by inducing eggs to temperatures lower than those usually found in mangrove

areas (<20°C), while secondary males can be induced by maintaining individuals at higher temperatures (>30°C), which are similar to those found on populations with higher rate of males (Turner *et al.*, 2006; Earley *et al.*, 2012). Their induction in the lab at naturally occurring temperatures suggests that, in the wild, secondary males may be the most ecologically relevant type of male (Turner *et al.*, 2006). This may also be the case for *K. hermaphroditus* in northeastern Brazil, where higher temperatures are predominant due to the Equator's line proximity.

Davis *et al.* (1990) demonstrated distinct male phenotypes within *Kryptolebias marmoratus*. One, from Twin Cays, Belize, exhibited a dark orange color pattern, an inconspicuous caudal ocellus, and caudal black margins along the dorsal, anal and caudal fins. The other one, from Florida, USA, exhibited a bright orange coloration almost devoid of dark chromatophores, except the head. Ellison *et al.* (2012) also found a male of *K. marmoratus* with orange coloration, faded ocellus and black borders on the anal and caudal fins in Belize; a pattern similar to that of the *K. hermaphroditus* male herein described.

Between 2011 and 2015, during extensive surveys at the Ceará-Mirim River estuary (Lira *et al.*, 2015), we collected 107 hermaphrodites before finding this single male. This apparent rarity of males aligns with the extremely low degree of heterozygosis reported for populations of *Kryptolebias hermaphroditus* from southeastern Brazil (Tartarenkov *et al.*, 2011). Together with absence of male specimens in all known populations (Costa, 2011; Lira *et al.*, 2015), these findings suggest that self-fertilizing hermaphroditism is the major mode of reproduction in *K. hermaphroditus*. However, the persistent occurrence of male individuals in *K. marmoratus*, but at low densities (varying from 1% in Florida to 10-25% in Belize, Davis *et al.*, 1990), suggests that sexual outcrossing serves an adaptive role within natural populations, possibly contributing to genetic diversity that helps the species to deal with new environmental pressures, such as parasites loads (Ellison *et al.*, 2011, 2013).

The occurrence of self-fertilization in *Kryptolebias marmoratus* has made that species a popular model for population genetics, development, evolutionary biology, behavior, among others (for reviews see Earley *et al.*, 2012; Taylor, 2012; Avise & Tatarenkov, 2015; Turko & Wright, 2015). Nonetheless, there have been relatively few studies on the South American lineage of *K. hermaphroditus*, possibly due to its misidentification as *K. marmoratus* (Seegers, 1984). Additionally, the difficulties of working in mangrove environments, which represent a complex, multi-dimensional, and cryptic habitat with variable environmental and physical features could also contributed to the scarcity of studies on the South American lineage (Taylor, 2012).

The discovery of a male specimen encourages more studies on *Kryptolebias hermaphroditus*. While researchers have addressed some aspects of systematics (Costa, 2006, 2011), evolutionary history (Costa *et al.*, 2010; Tatarenkov *et al.*, 2009, 2011), and geographic distribution (Sarmento-Soares *et al.*, 2014; Lira *et al.*, 2015), there is little available information regarding natural history, ecology and population dynamics. Despite the large gap in our knowledge of the biology of *K. hermaphroditus*, the recent studies above mentioned, including the present one, suggest some interesting perspectives for studies on evolution, morphology, behavior and physiology across the species' wide geographic distribution in the mangroves of the Brazilian coast.

Acknowledgements

We thank Lucas Medeiros, Mateus Lira and Roney Paiva for helping with fieldwork, and to Luciano Barros-Neto for providing the pictures of the live specimens. Carlos Eduardo Alencar provided information about mangrove crabs in the Ceará-Mirim estuary. Graco Viana and Centro de Treinamento em Aquicultura da Universidade Federal do Rio Grande do Norte provided logistical support. Waldir Berbel-Filho receives a PhD scholarship from Science without Borders Program/CNPq (process #233161/2014-7) and Helder Espírito-Santo received a Postdoctoral fellowship from PDJ Program/CNPq (process #168228/2014-9).

References

Avise, J. C. & A. Tatarenkov. 2015. Population genetics and evolution of the mangrove rivulus *Kryptolebias marmoratus*, the world's only self-fertilizing hermaphroditic vertebrate. Journal of Fish Biology, 87: 519-538.

Costa, W. J. E. M. 1995. Pearl killifishes: the Cynolebiatinae: systematics and biogeography of the Neotropical annual fish subfamily. Neptune city, TFH Publications, 128p.

Costa, W. J. E. M. 2003. Family Rivulidae. Pp. 526-548. In: Reis, R. E., S. O. Kullander & C. J. Ferraris Jr. (Eds.). Check list of freshwater fishes of South and Central America. Porto Alegre, Edipucrs.

Costa, W. J. E. M. 2006. Redescription of *Kryptolebias ocellatus* (Hensel) and *K. caudomarginatus* (Seegers) (Teleostei: Cyprinodontiformes: Rivulidae), two killifishes from mangroves of south-eastern Brazil. Aqua Journal of Ichthyology and Aquatic Biology, 11: 5-12.

Costa, W. J. E. M. 2009. Peixes aploqueilóideos da Mata Atlântica brasileira: história, diversidade e conservação. Rio de Janeiro, Museu Nacional. 171p.

Costa, W. J. E. M. 2011. Identity of *Rivulus ocellatus* and a new name for a hermaphroditic species of *Kryptolebias* from southeastern Brazil (Cyprinodontiformes: Rivulidae). Ichthyological Exploration of Freshwaters, 22: 185-192.

Costa, W. J. E. M., S. M. Q. Lima & R. Bartolette. 2010. Androdioecy in *Kryptolebias* and the evolution of self-fertilizing hermaphroditism. Biological Journal of the Linnean Society, 99: 344-349.

Davis, W. P., D. S. Taylor & B. J. Turner. 1990. Field observations of the ecology and habits of mangrove rivulus (*Rivulus marmoratus*) in Belize and Florida (Teleostei: Cyprinodontiformes: Rivulidae). Ichthyological Exploration of Freshwaters, 1: 123-134.

Earley, R. L., A. F. Hannineb, A. Fuller, M. J. Garcia & E. A. Lee. 2012. Phenotypic plasticity and integration in the mangrove rivulus (*Kryptolebias marmoratus*): a prospectus. Integrative and Comparative Biology, 52: 814-827.

Ellison, A., J. Cable & S. Consuegra. 2011. Best of both worlds? Association between outcrossing and parasite loads in a selfing fish. Evolution, 65: 3021-3026.

Ellison, A., P. Wright, D. S. Taylor, C. Cooper, K. Regan, S. Currie & S. Consuegra. 2012. Environmental diel variation, parasite loads, and local population structuring of a mixed-mating mangrove fish. Ecology and Evolution, 2: 1682-1695.

Ellison, A., J. Jones, C. Inchley & S. Consuegra. 2013. Choosy males could help explain androdioecy in a selfing fish. The American Naturalist, 181: 855-862.

Harrington, R. W. 1971. How ecological and genetic factors interact to determine when self-fertilizing hermaphrodites of *Rivulus marmoratus* change into functional secondary males, with a reappraisal of the mode of intersexuality among fishes. Copeia, 1971: 389-432.

Huber, J. H. 1992. Review of Rivulus: ecobiogeography-relationships. Paris, Cybium, Société Française d'Ichtyologie, 572p.

Lira, M. G. S., R. E. C. Paiva, T. P. A. Ramos & S. M. Q. Lima. 2015. First record of *Kryptolebias hermaphroditus* Costa, 2011 (Cyprinodontiformes: Rivulidae) in the extreme north Atlantic Forest mangroves, Rio Grande do Norte state, Brazil. Check List, 11: 1656.

Sarmento-Soares, L. M., L. F. S. Ingenito, L. F. Duboc, R. F. Martins-Pinheiro, R. Borçato & J. P. Silva. 2014. Primeiros registros de *Kryptolebias ocellatus* (Hensel) (Cyprinodontiformes, Rivulidae) para riachos de Mata Atlântica no Espírito Santo. Boletim da Sociedade Brasileira de Ictiologia, 111: 15-19.

Soto, C. G. & D. L. G. Noakes. 1994. Coloration and gender in the hermaphroditic fish *Rivulus marmoratus* Poey (Teleostei: Rivulidae). Ichthyological Exploration of Freshwaters, 5: 79-90.

Tatarenkov, A., S. M. Q. Lima, D. S. Taylor & J. C. Avise. 2009. Long-term retention of self-fertilization in a fish clade. Proceedings of the National Academy of Sciences, 106: 14456-14459.

Tatarenkov, A., S. M. Q. Lima & J. C. Avise. 2011. Extreme homogeneity and low genetic diversity in *Kryptolebias ocellatus* from south-eastern Brazil suggest a recent foundation for this androdioecious fish population. Journal of Fish Biology, 79: 2095-2105.

Taylor, D. S. 2012. Twenty-four years in the mud: What have we learned about the natural history and ecology of the mangrove rivulus, *Kryptolebias marmoratus*? Integrative and Comparative Biology, 52: 724-736.

Turko, A. J. & P. A. Wright. 2015. Evolution, ecology and physiology of amphibious killifishes (Cyprinodontiformes). Journal of Fish Biology, 87: 815-835.

Turner, B. J., M. T. Fisher, D. S. Taylor, W. P. Davis & B. L. Jarrett. 2006. Evolution of 'maleness' and outcrossing in a population of the self-fertilizing killifish, *Kryptolebias marmoratus*. Evolutionary Ecology Research, 8: 1475-1486.

Submitted March 6, 2016
Accepted August 18, 2016 by Paulo Lucinda

Tatarenkov et al. (2017):

Biological Journal of the Linnean Society, 2017, **122**, 558–578. With 6 figures.

Deep and concordant subdivisions in the self-fertilizing mangrove killifishes (*Kryptolebias*) revealed by nuclear and mtDNA markers

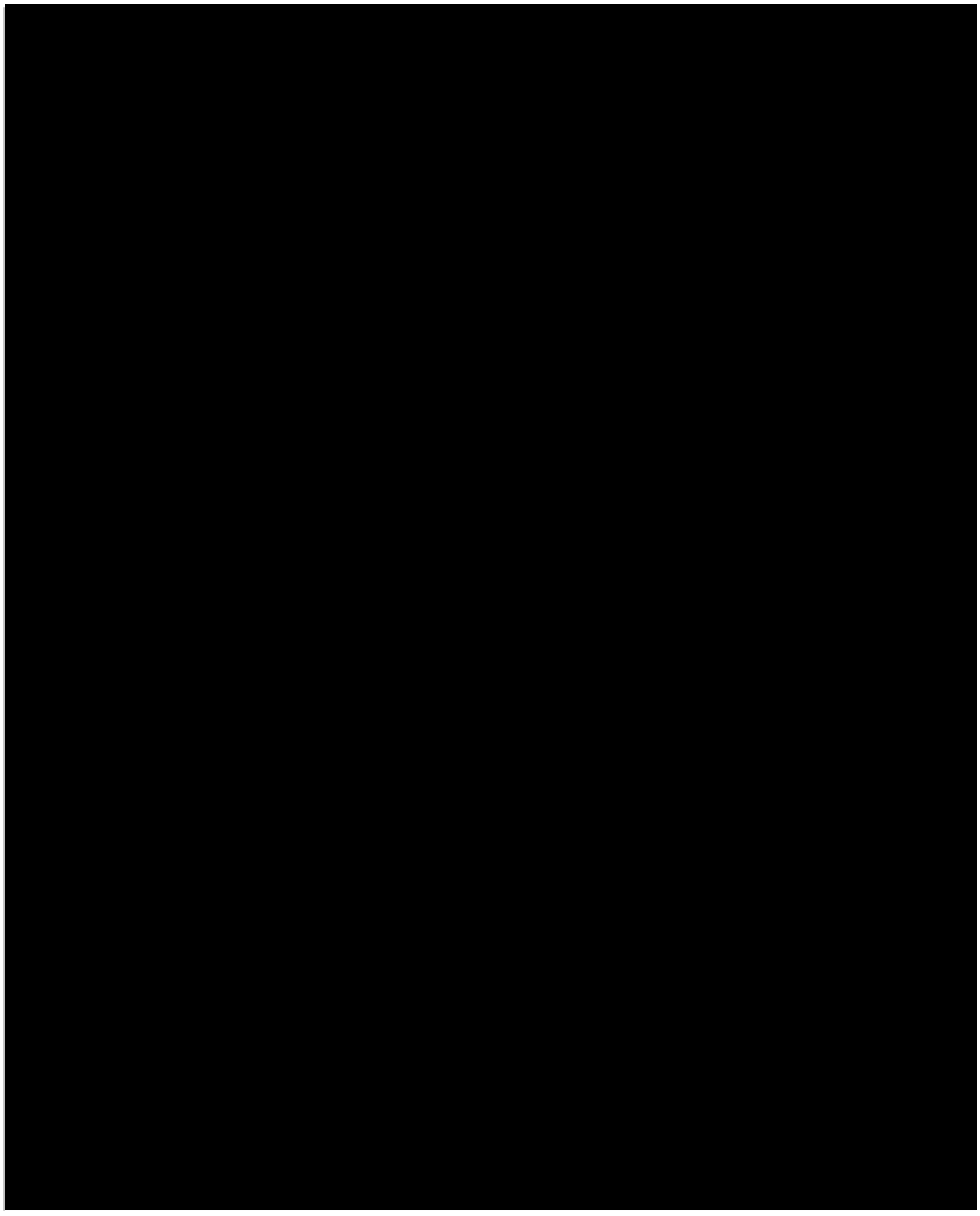
ANDREY TATARENKOV[1*], SERGIO M. Q. LIMA[2], RYAN L. EARLEY[3], WALDIR M. BERBEL-FILHO[2,4], FRANS B. M. VERMEULEN[5], D. SCOTT TAYLOR[6], KRISTINE MARSON[3], BRUCE J. TURNER[7] and JOHN C. AVISE[1]

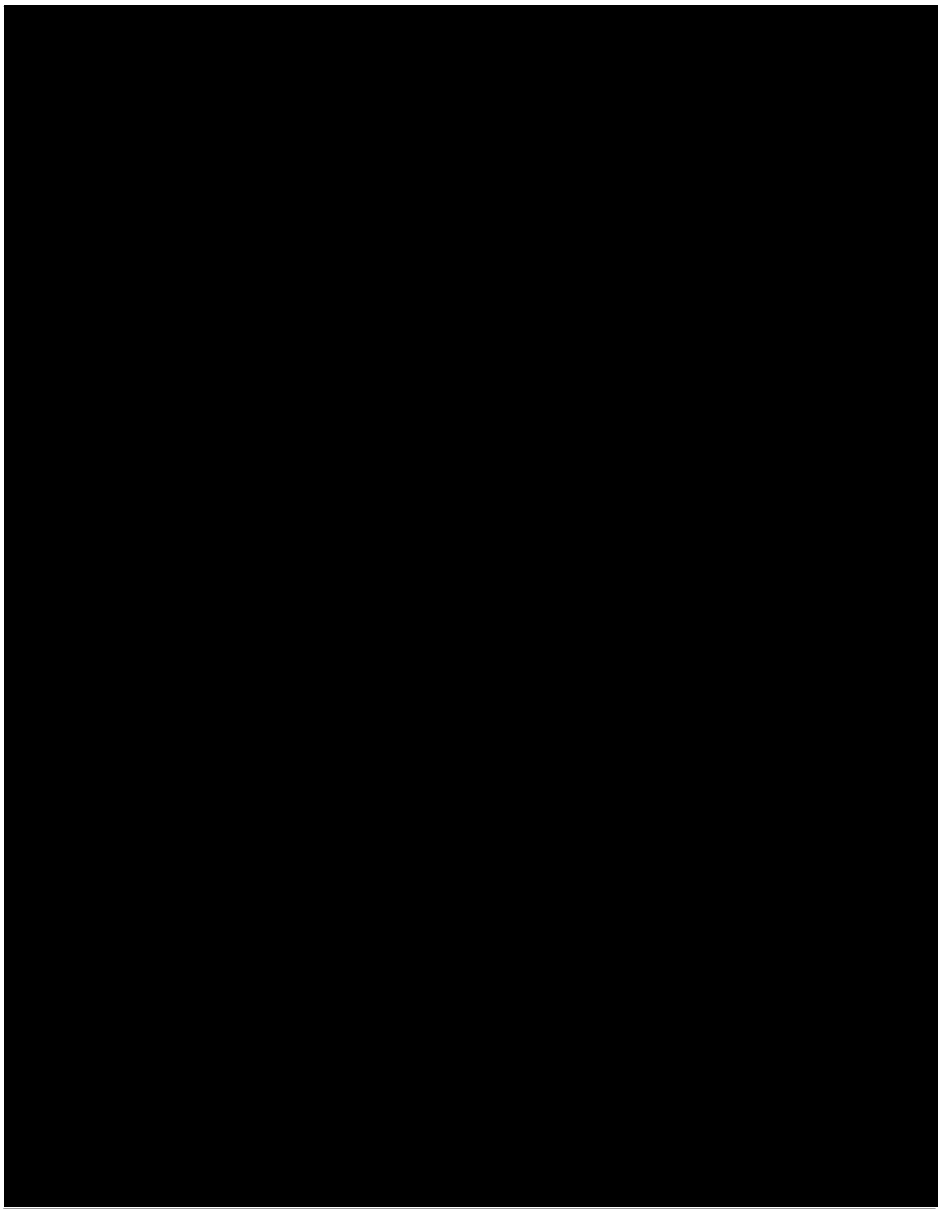
[1]*Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697, USA*
[2]*Laboratório de Ictiologia Sistemática e Evolutiva, Departamento de Botânica e Zoologia, Universidade Federal do Rio Grande do Norte, Av. Senador Salgado Filho 3000, 56078-970 Natal, RN, Brazil*
[3]*Department of Biological Sciences, University of Alabama, Tuscaloosa, AL 35487, USA*
[4]*Department of BioSciences, College of Science, Swansea University, Swansea, Wales SA2 8PP, UK*
[5]*Tanki Leendert 194 c, Noord, Aruba*
[6]*Brevard County Environmentally Endangered Lands Program, Melbourne, FL 32904, USA*
[7]*Department of Biological Sciences, Virginia Tech, Blacksburg, VA 24061, USA*

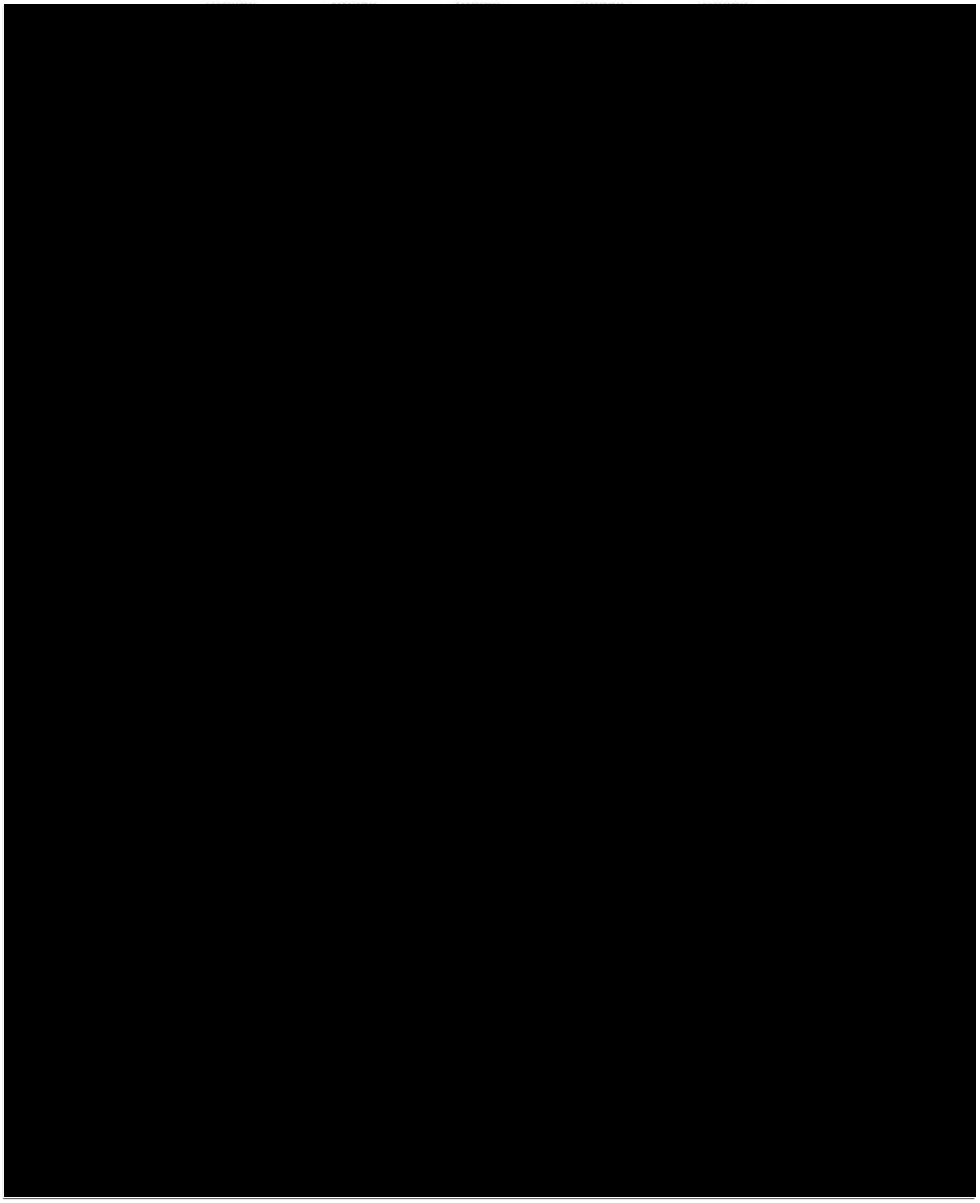
Received 6 May 2017; revised 27 July 2017; accepted for publication 28 July 2017

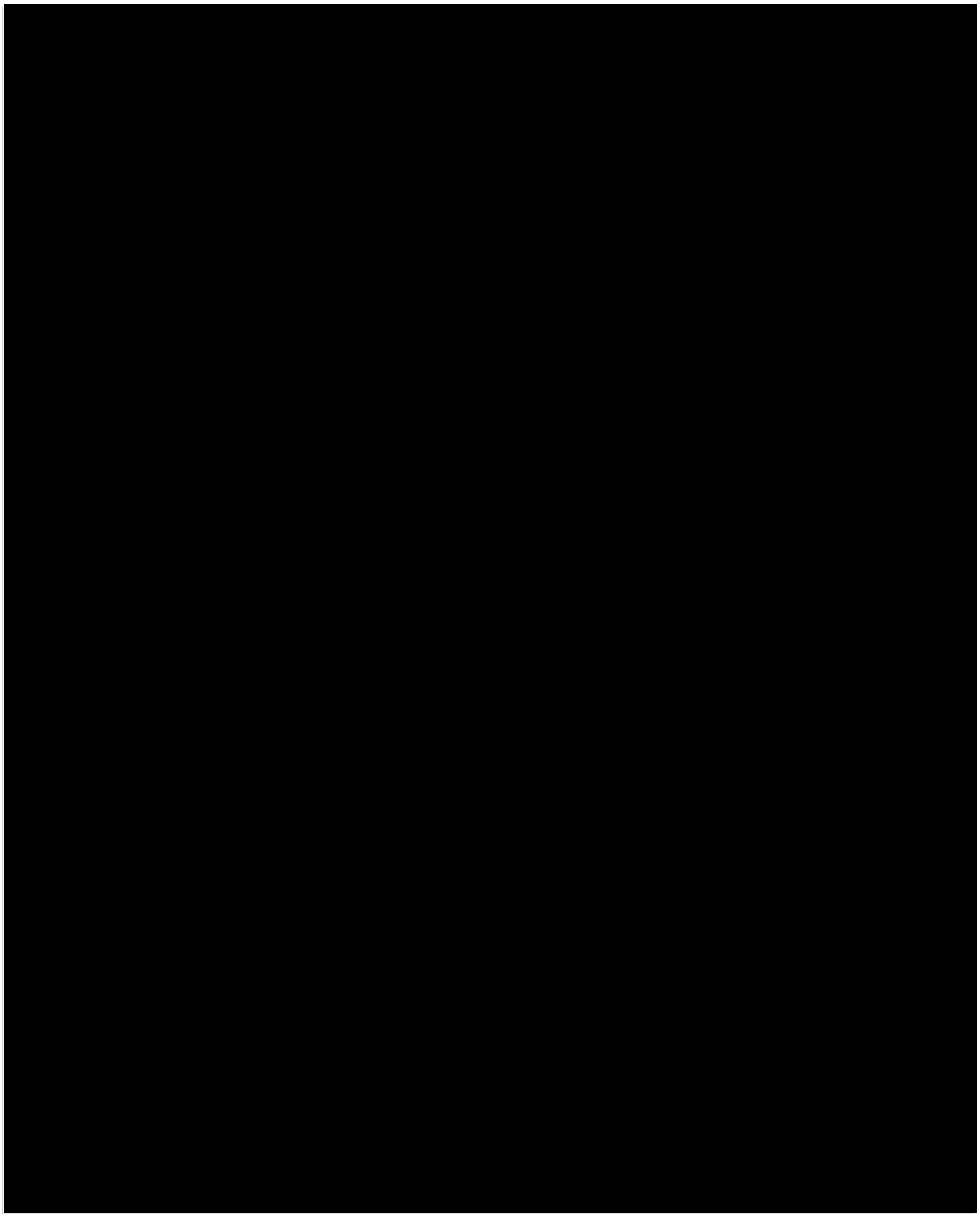
We use extensive geographical sampling and surveys of nuclear microsatellite and mitochondrial DNA loci to investigate the phylogeographic structure of the only recognized self-fertilizing vertebrates, the mangrove killifishes, currently thought to comprise two cryptic species, *Kryptolebias marmoratus* and *Kryptolebias hermaphroditus*. All genetic markers revealed three concordant main clades. The Northern clade includes populations from Florida, northern Cuba, Bahamas, Belize and Honduras and corresponds to *K. marmoratus*. The Southern clade encompasses populations from Brazil and corresponds to *K. hermaphroditus*. This species was considered endemic to southeastern Brazil, but molecular data corroborate its occurrence in northeastern Brazil. The Central clade, not previously resolved with genetic data, includes populations from Panama and Antilles. Despite the geographic proximity of the Northern and Central clades, the latter is genetically closer to the Southern clade. The discovery of the Central clade raises some taxonomic issues – it can either be considered a distinct species or united with the Southern clade into a single species with two subspecies. Another possible taxonomic solution is a single selfing species, *K. marmoratus*, with three subspecies. We show that the Central and Southern clades are highly selfing (97–100%), whereas selfing rates of the Northern clade populations vary geographically (39–99%). Genetic patterns indicate that populations in SE Brazil are recent, contrary to expectations based on the known distributions of related species.

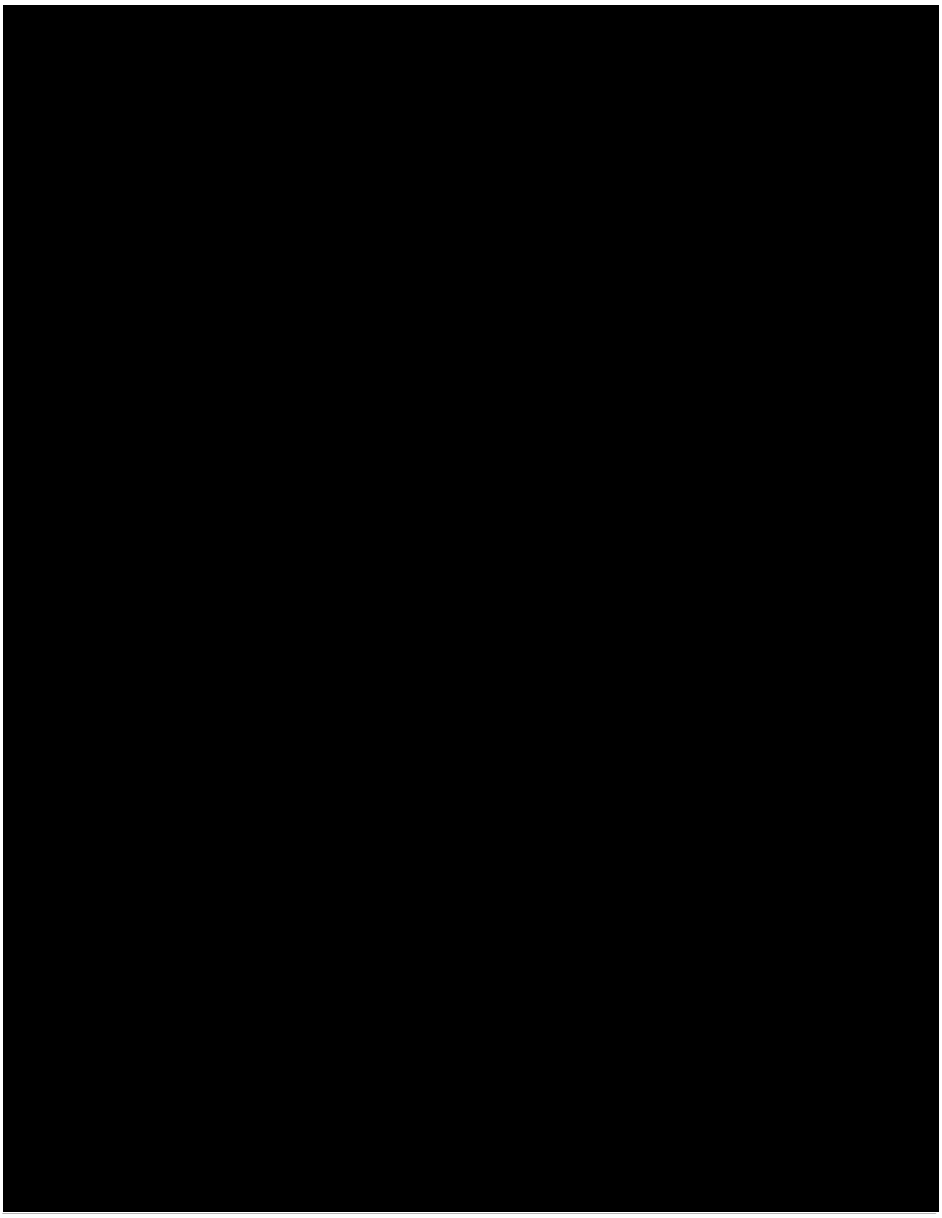
ADDITIONAL KEYWORDS: Caribbean – cryptic species – Cynolebiidae – mangrove rivulus – microsatellites – selfing – species concept – species delimitation.

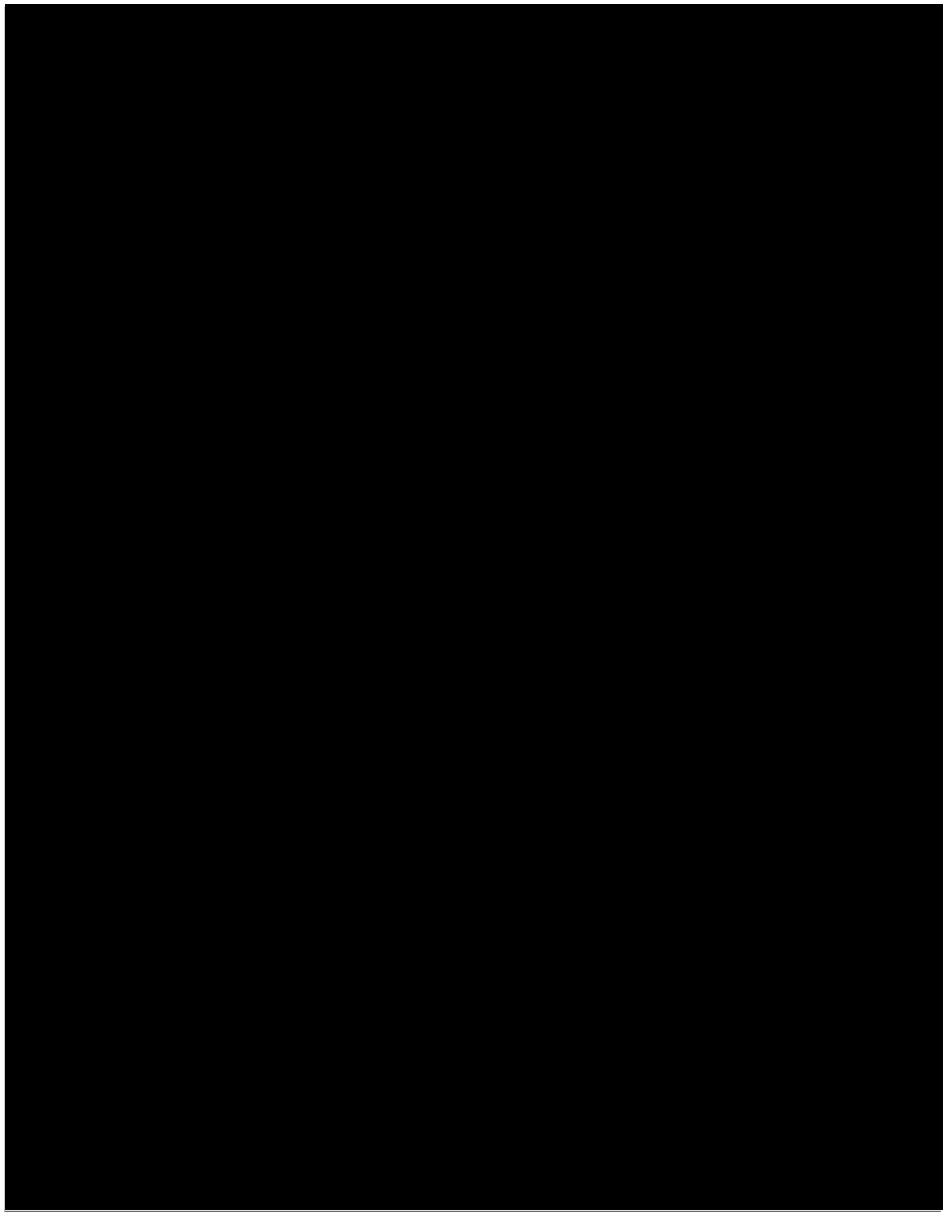


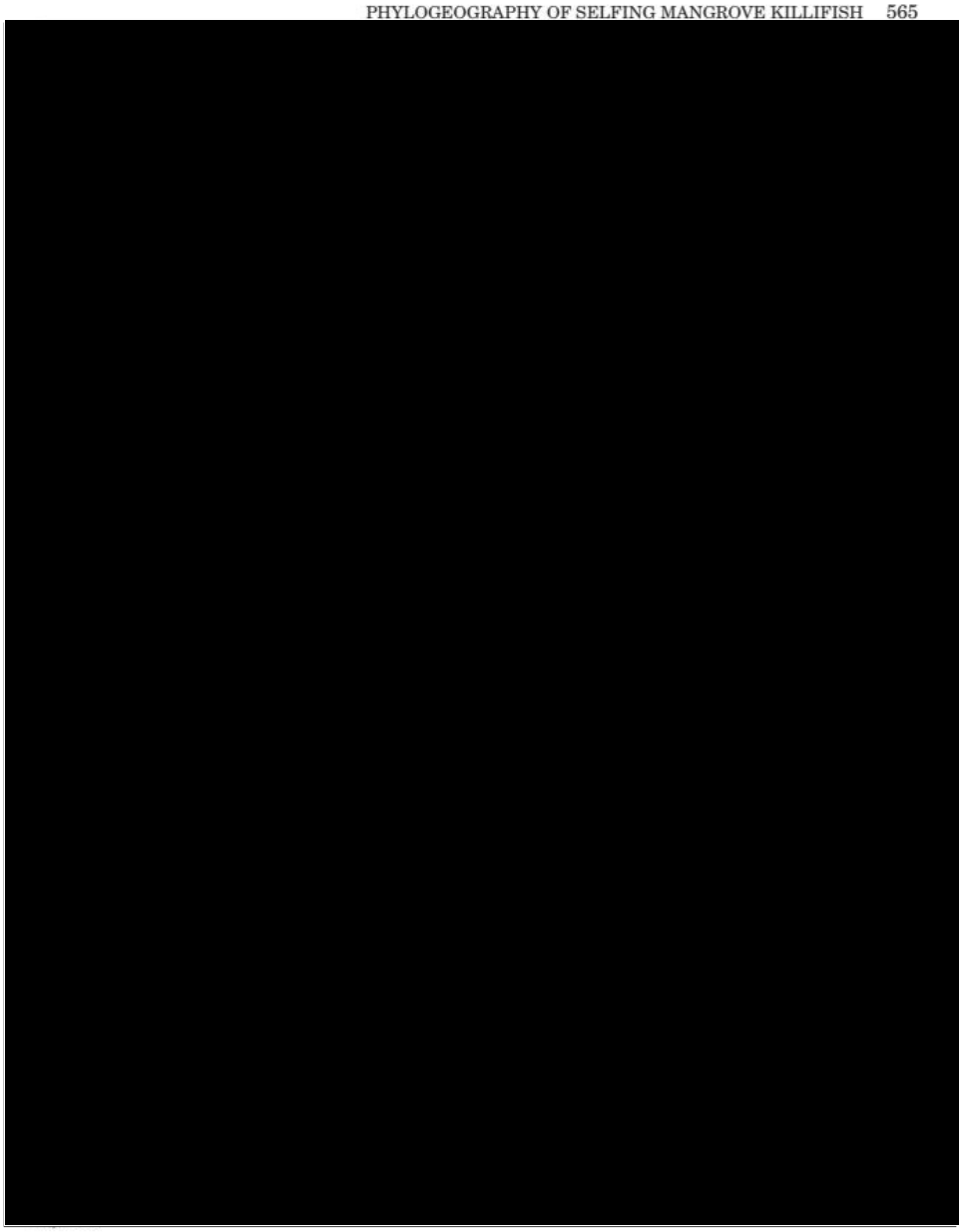


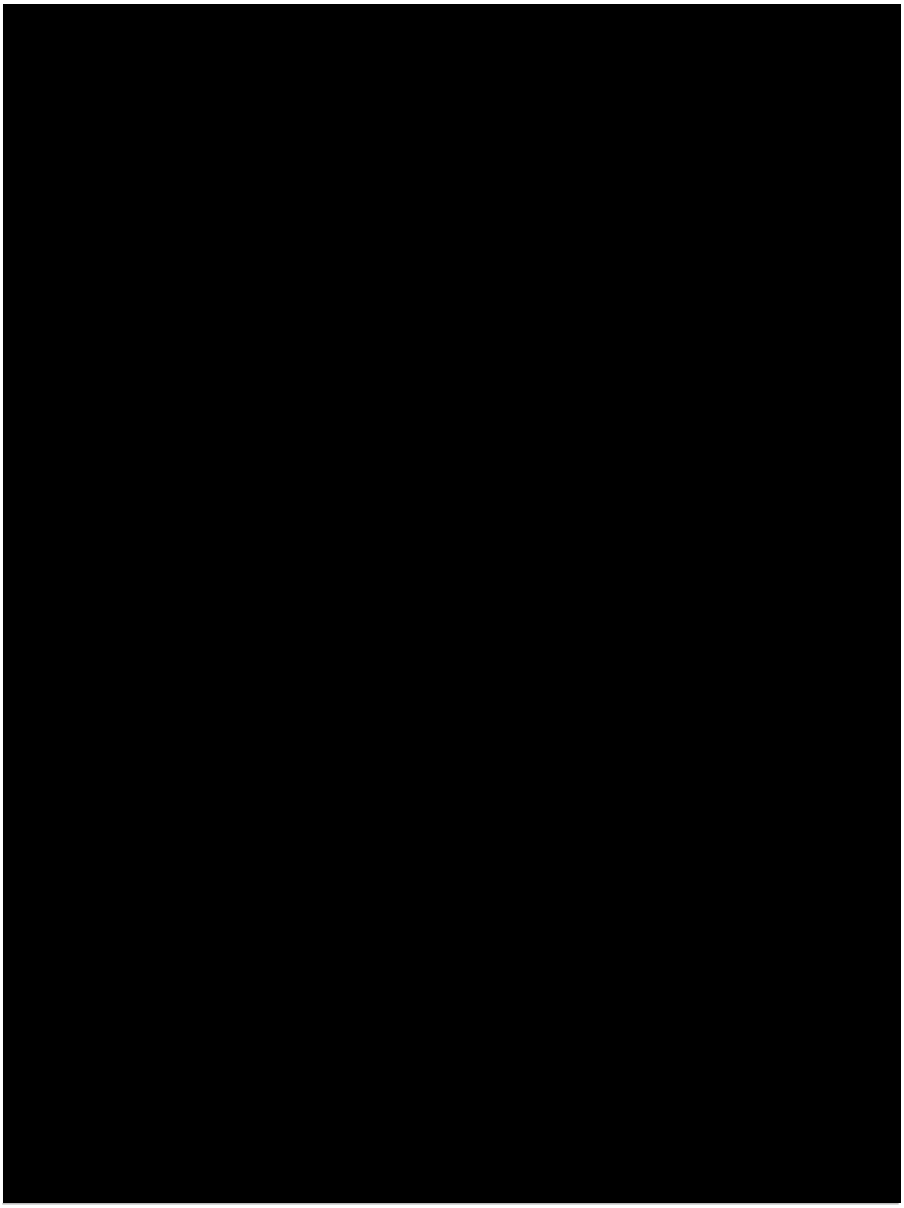


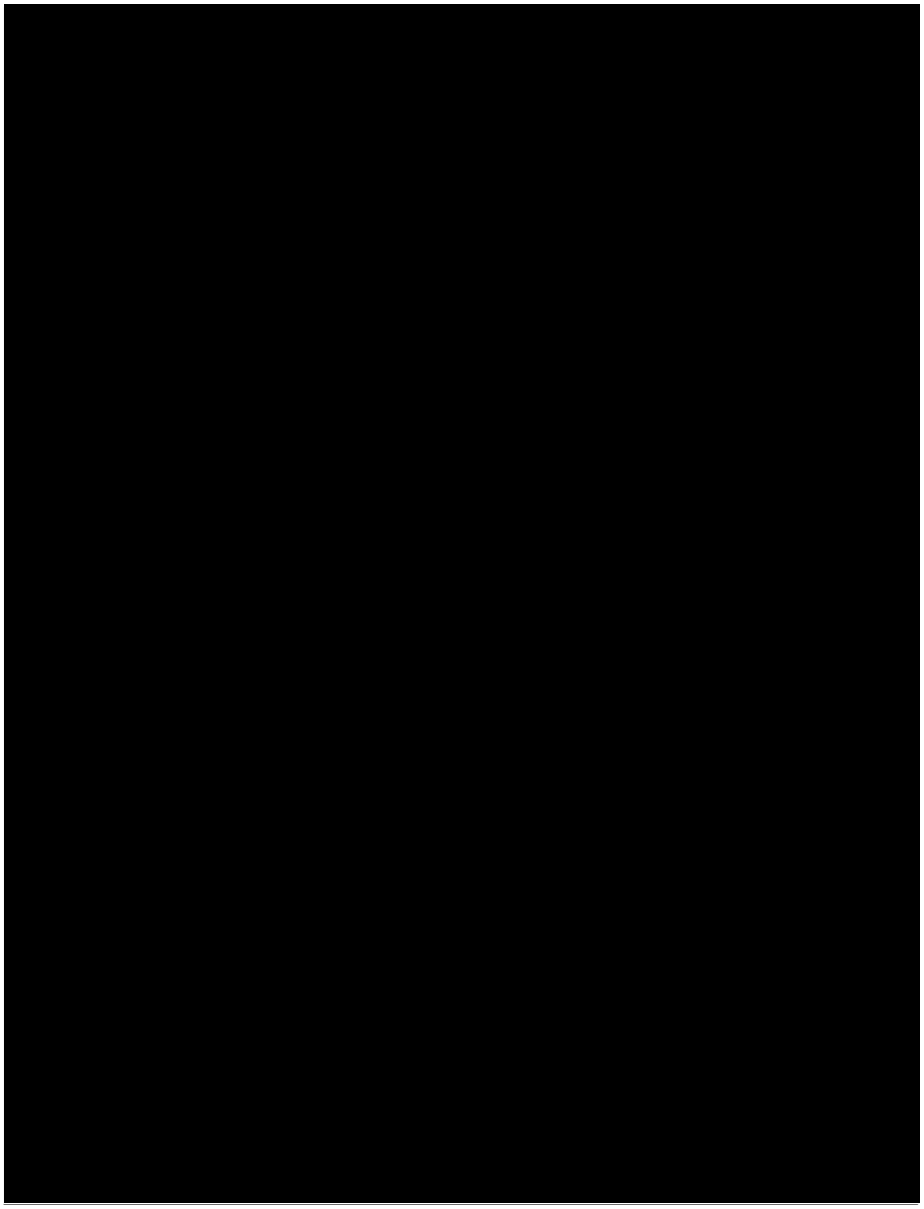


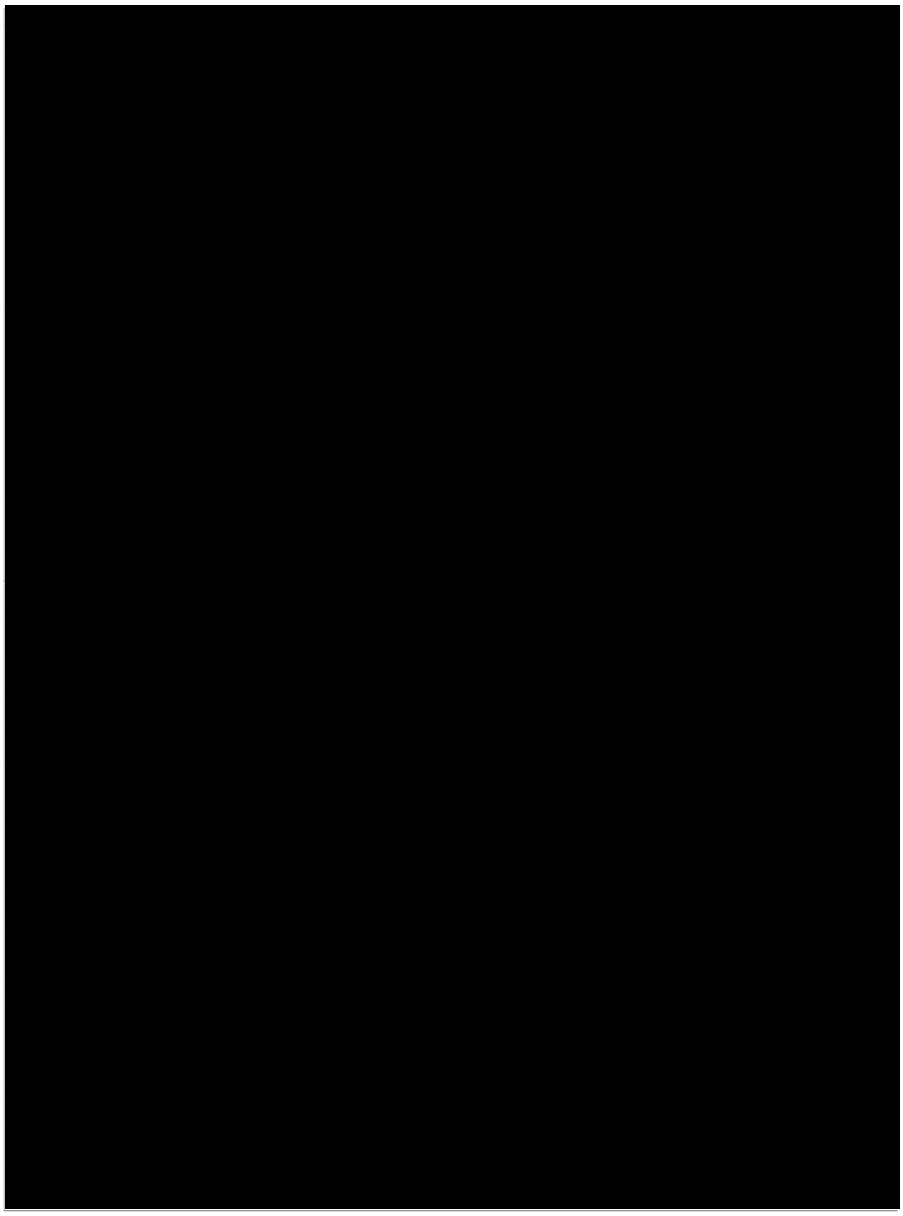


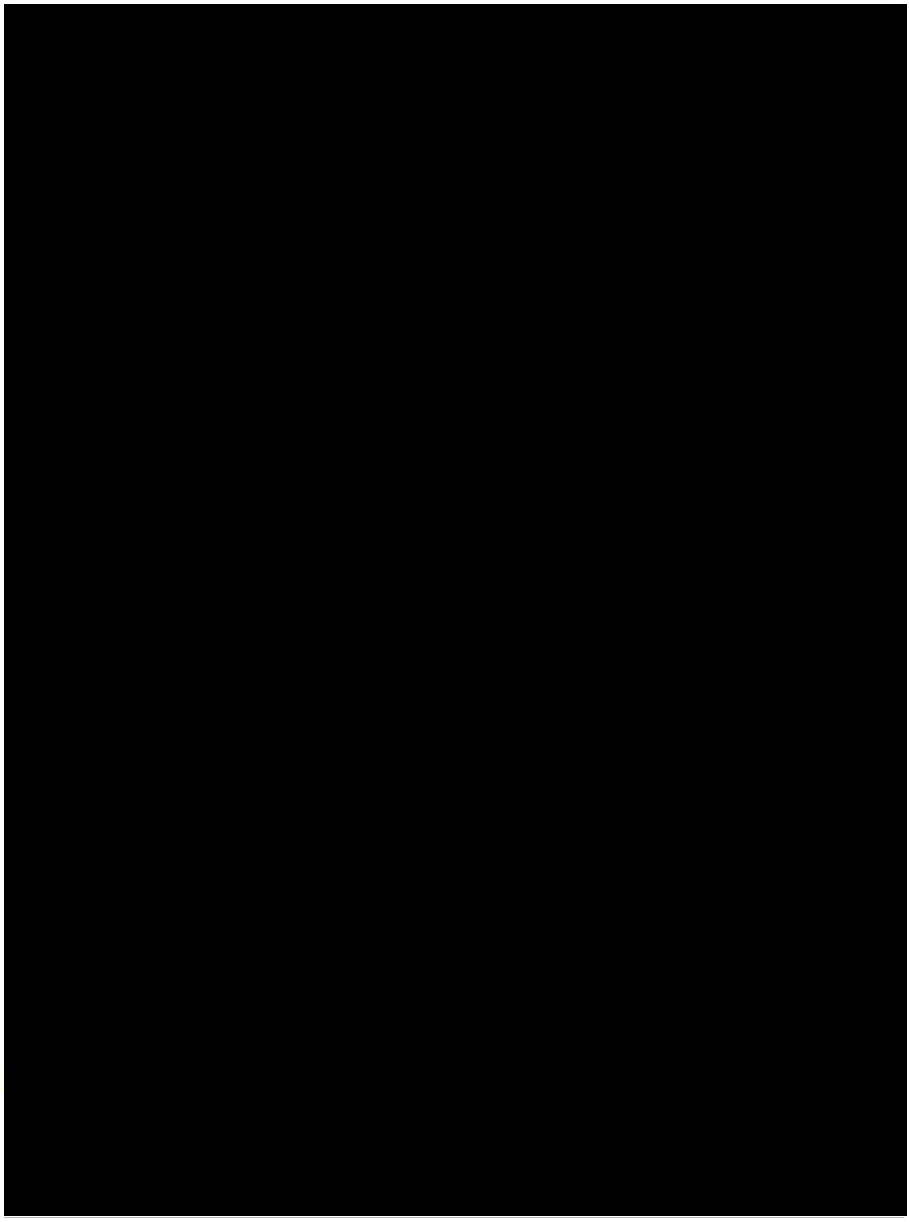


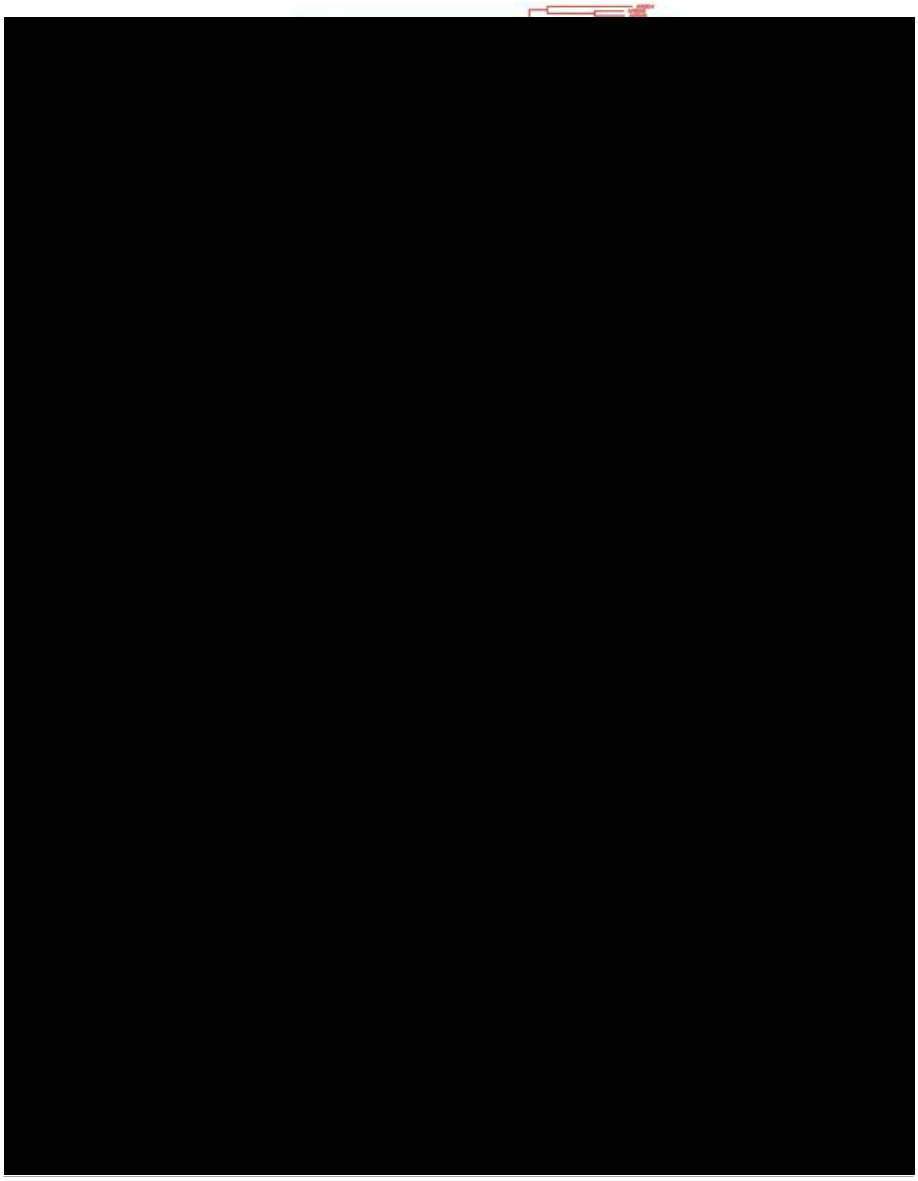


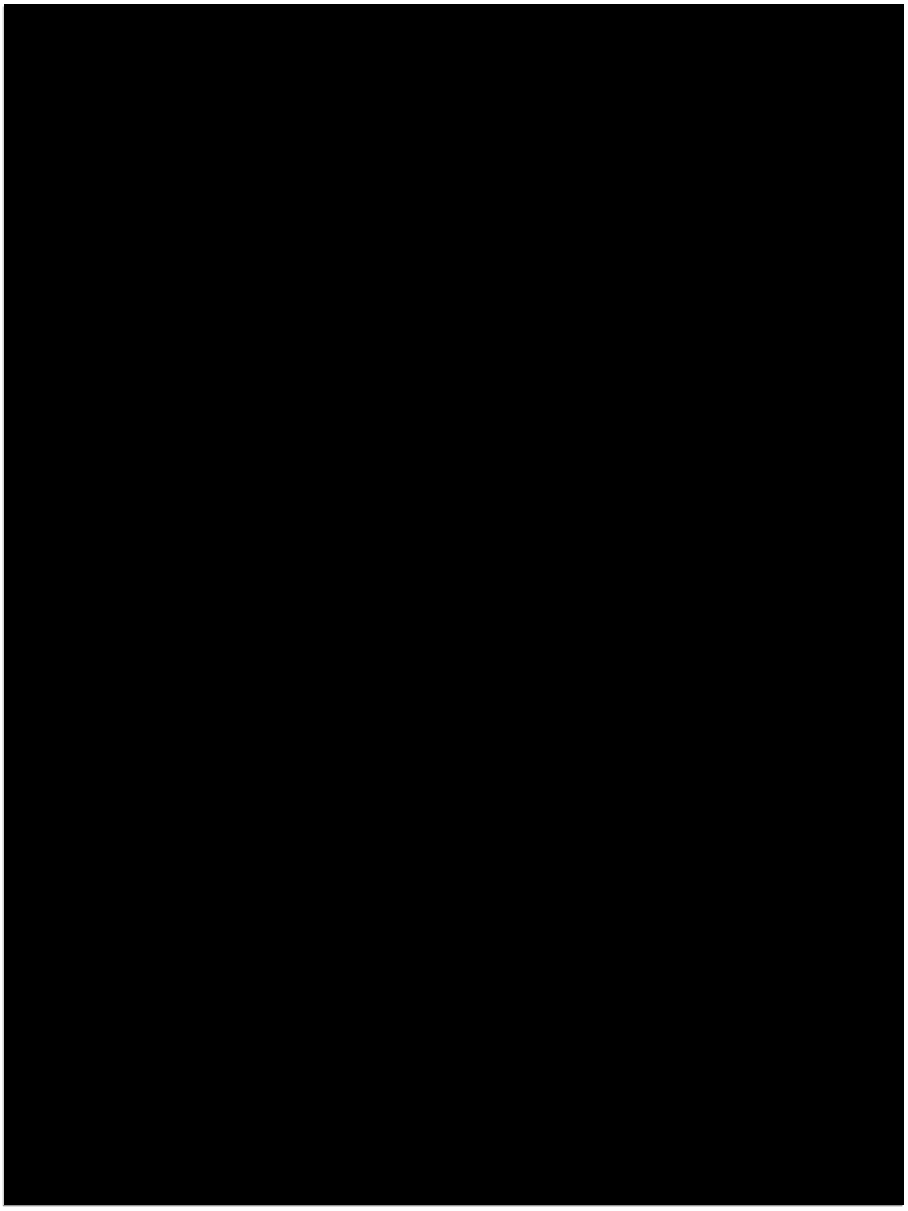


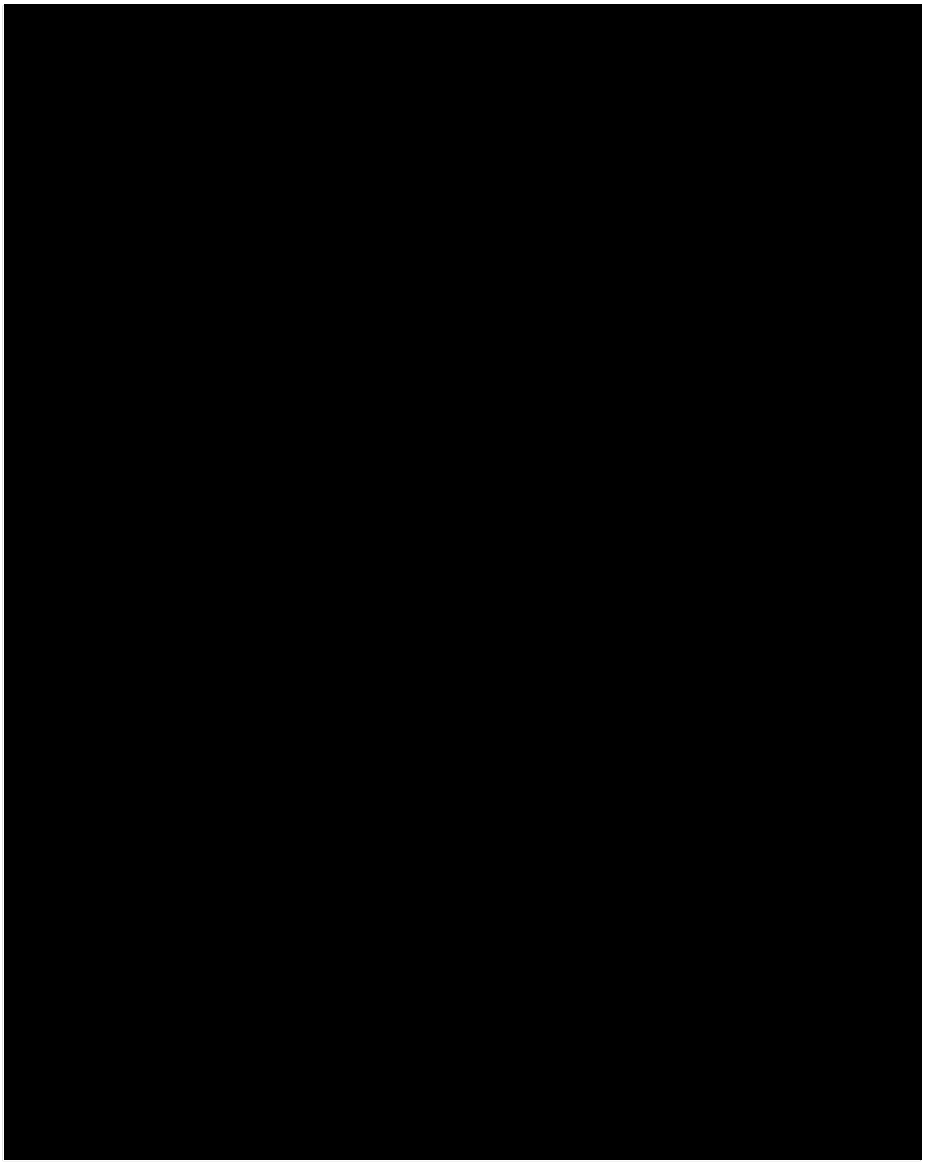


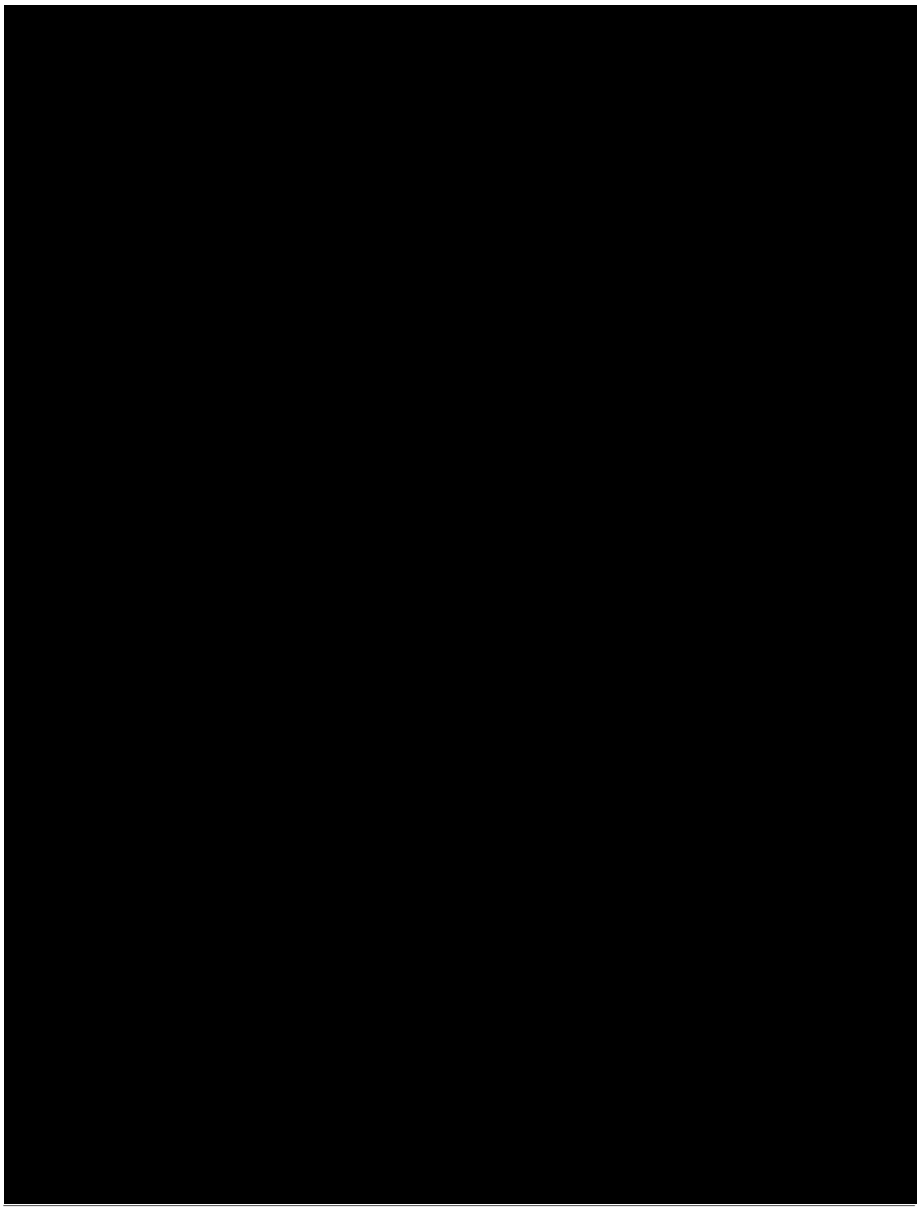


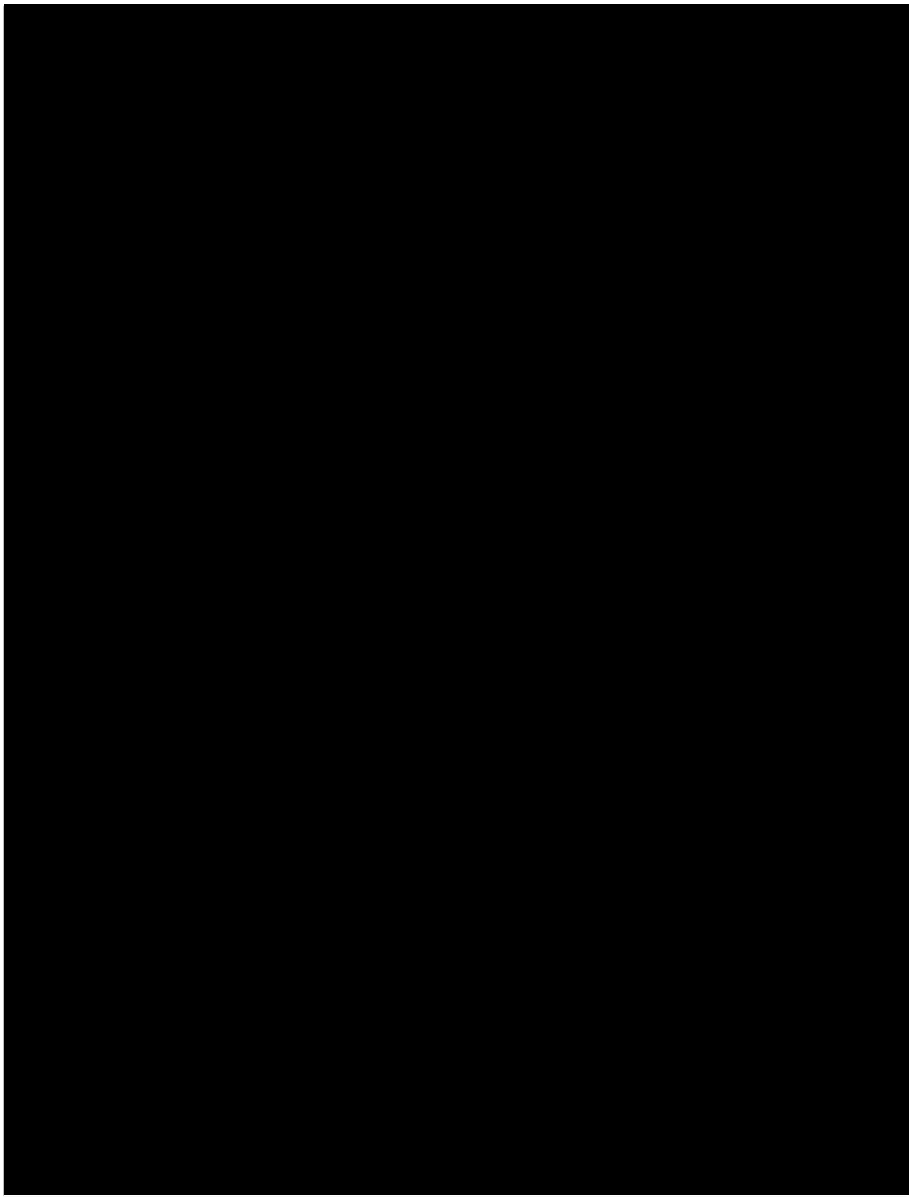


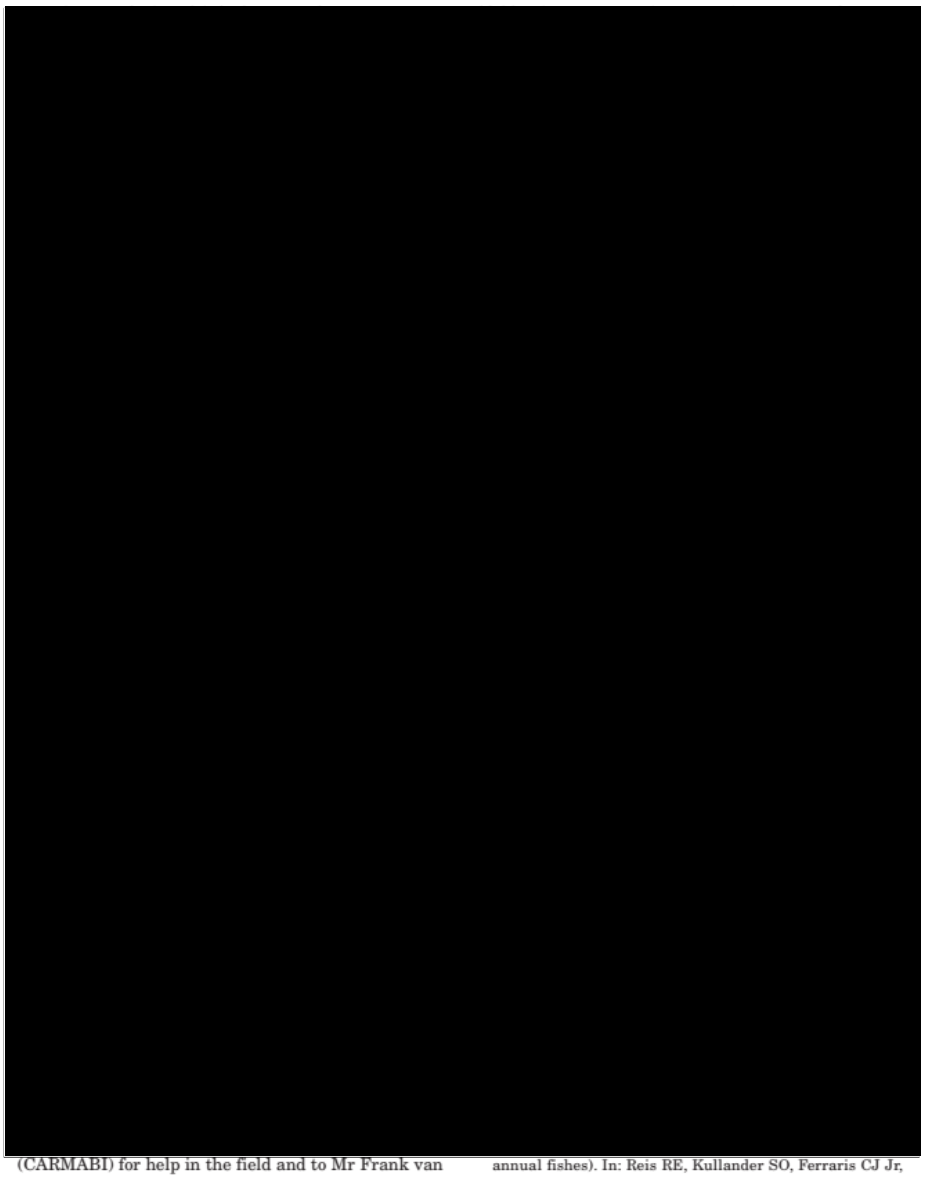






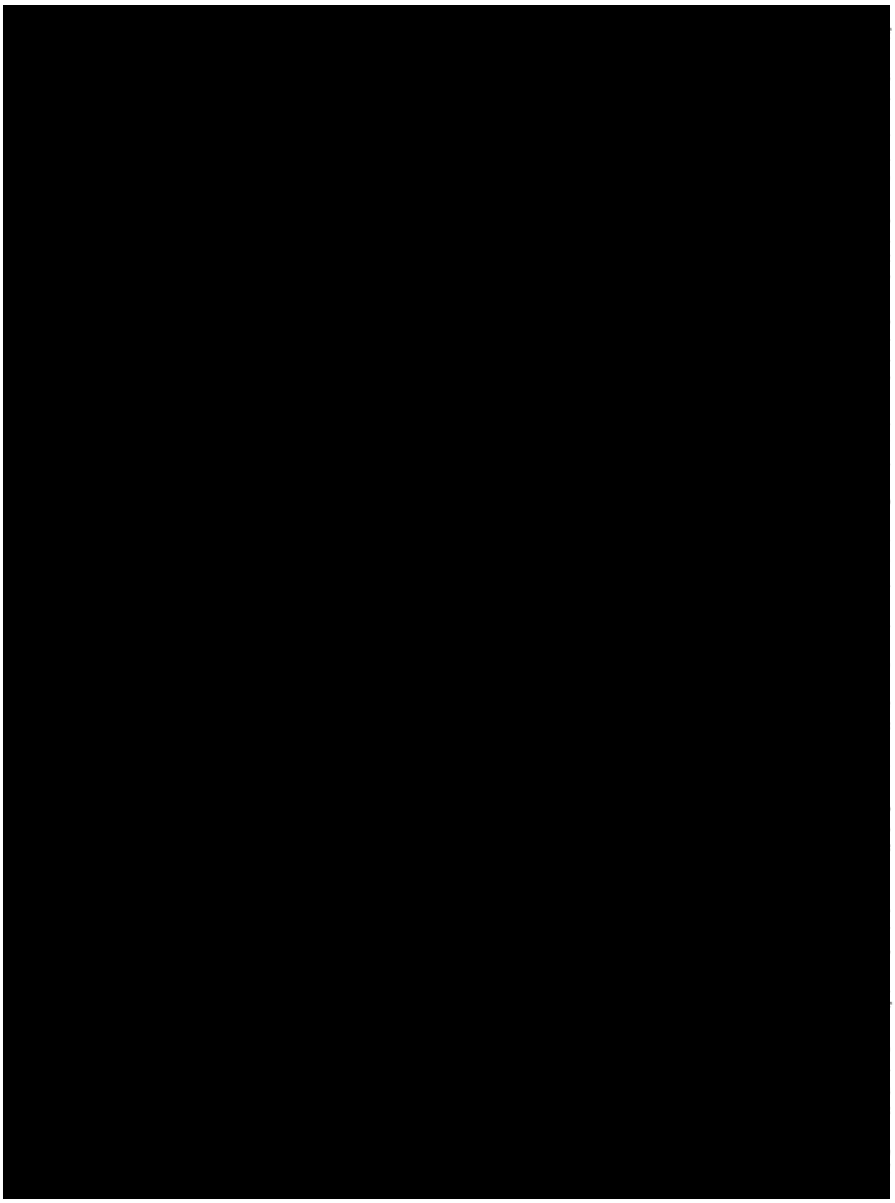


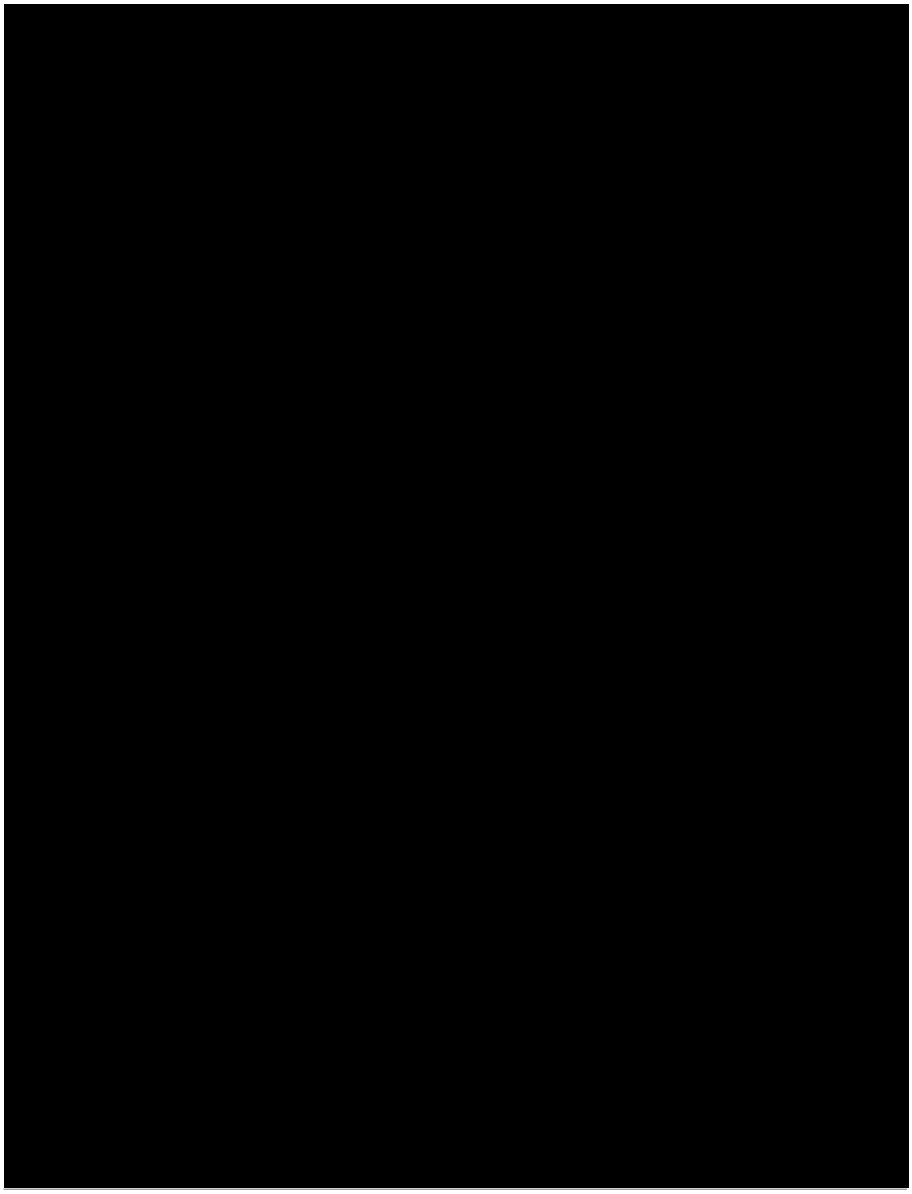


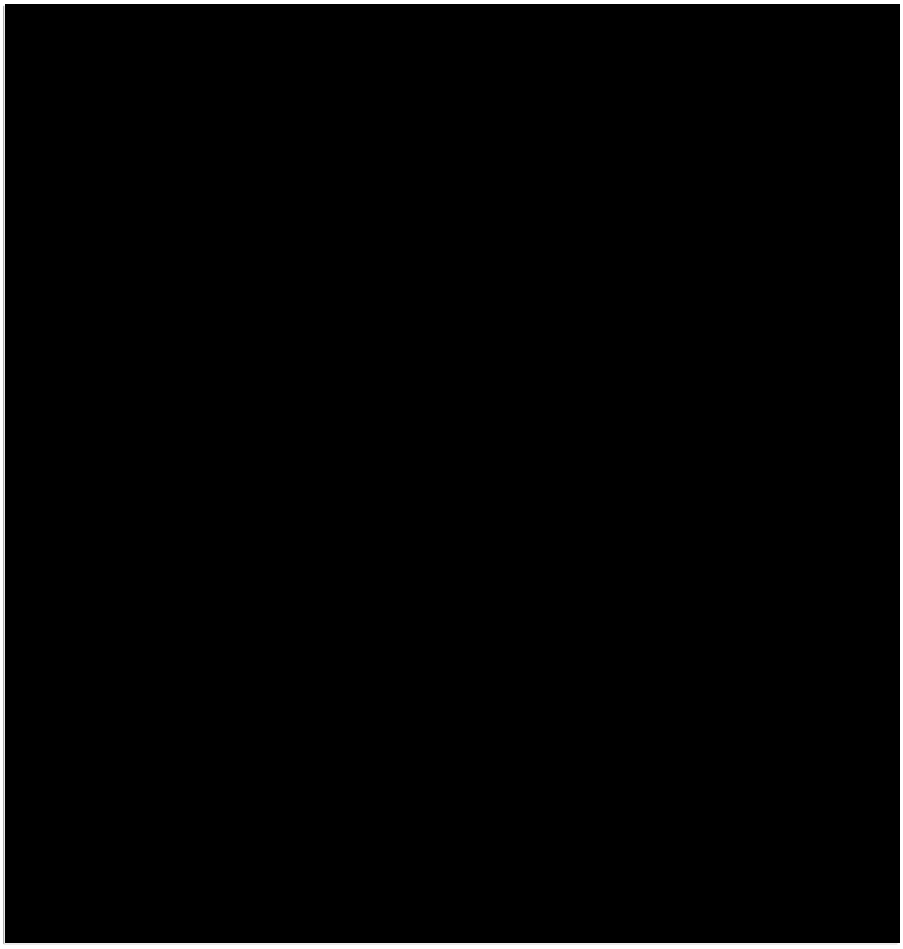


(CARMABI) for help in the field and to Mr Frank van

annual fishes). In: Reis RE, Kullander SO, Ferraris CJ Jr,







REFERENCES

Agrawal, A. A., C. Laforsch, and R. Tollrian. 1999. Transgenerational induction of defences in animals and plants. Nature 401:60-63.

Akalin, A., M. Kormaksson, S. Li, F. E. Garrett-Bakelman, M. E. Figueroa, A. Melnick, and C. E. Mason. 2012. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. Genome biology 13:R87.

Alleman, L. K., and M. W. Hester. 2011. Reproductive ecology of black mangrove (*Avicennia germinans*) along the Louisiana coast: propagule production cycles, dispersal limitations, and establishment elevations. Estuaries and Coasts 34:1068-1077.

Allendorf, F. W., and J. J. Hard. 2009. Human-induced evolution caused by unnatural selection through harvest of wild animals. Proceedings of the National Academy of Sciences 106:9987-9994.

Allis, C. D., and T. Jenuwein. 2016. The molecular hallmarks of epigenetic control. Nature Reviews in Genetics 17:487-500.

Alonso, C., R. Perez, P. Bazaga, M. Medrano, and C. M. Herrera. 2016. MSAP markers and global cytosine methylation in plants: a literature survey and comparative analysis for a wild-growing species. Molecular Ecology Resources 16:80-90.

Anastasiadi, D., A. Esteve-Codina, and F. Piferrer. 2018. Consistent inverse correlation between DNA methylation of the first intron and gene expression across tissues and species. Epigenetics & Chromatin 11:37.

Anderson, E., and E. Thompson. 2002. A model-based method for identifying species hybrids using multilocus genetic data. Genetics 160:1217-1229.

Andrews, S. 2007. SeqMonk: A tool to visualise and analyse high throughput mapped sequence data. https://www.bioinformatics.babraham.ac.uk/projects/seqmonk/.

—. 2010. FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Angers, B., E. Castonguay, and R. Massicotte. 2010. Environmentally induced phenotypes and DNA methylation: how to deal with unpredictable conditions until the next generation and after. Molecular Ecology 19:1283-1295.

Aparicio, J. M., J. Ortego, and P. J. Cordero. 2006. What should we weigh to estimate heterozygosity, alleles or loci? Molecular Ecology 15:4659-4665.

Arai, J. A., and L. A. Feig. 2011. Long-lasting and transgenerational effects of an environmental enrichment on memory formation. Brain Research Bulletin 85:30-35.

Asselman, J., D. I. De Coninck, M. B. Vandegehuchte, M. Jansen, E. Decaestecker, L. De Meester, J. Vanden Bussche et al. 2015. Global cytosine methylation in *Daphnia magna* depends on genotype, environment, and their interaction. Environmental Toxicology and Chemistry 34:1056-1061.

Assenov, Y., F. Ramirez, S. E. Schelhorn, T. Lengauer, and M. Albrecht. 2008. Computing topological parameters of biological networks. Bioinformatics 24:282-284.

Avise, J. C., and A. Tatarenkov. 2015. Population genetics and evolution of the mangrove rivulus *Kryptolebias marmoratus*, the world's only self-fertilizing hermaphroditic vertebrate. Journal of Fish Biology 87:519-538.

Badyaev, A. V. 2005. Developmental plasticity and evolution. American Journal of Physical Anthropology 128:703-706.

Baerwald, M. R., M. H. Meek, M. R. Stephens, R. P. Nagarajan, A. M. Goodbla, K. M. Tomalty, G. H. Thorgaard et al. 2016. Migration-related phenotypic divergence is associated with epigenetic modifications in rainbow trout. Molecular Ecology 25:1785-1800.

Baker, H. G. 1955. Self-compatibility and establishment after '"long-distance" dispersal. Evolution 9:347-349.

Ball, M. C. 1988. Ecophysiology of mangroves. Trees 2:129-142.

Banta, J. A., and C. L. Richards. 2018. Quantitative epigenetics and evolution. Heredity 121:210-224.

Barcellos, L. J. G., L. C. Kreutz, R. M. Quevedo, J. G. S. da Rosa, G. Koakoski, L. Centenaro, and E. Pottker. 2009. Influence of color background and shelter availability on jundiá (Rhamdia quelen) stress response. Aquaculture 288:51-56.

Barrett, S. C. 2014. Evolution of mating systems: outcrossing versus selfing. The Princeton guide to evolution:356-362.

Barton, H., W. M. Berbel-Filho, S. Consuegra, L. Francis, C. Tizaoui, R. S. Conlan, and S. R. Teixeira. 2018. Ultrasensitive environmental assessment of xeno-estrogens in water samples using label-free graphene immunosensors. Analytical biochemistry 548:102-108.

Bates, D., M. Maechler, B. Bolker, and M. S. Walker. 2019. Package 'mlmRev'.

Belkhir, K. 2004. GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. http://www.genetix.univ-montp2.fr/genetix/genetix.htm.

Bell, G. 1982, The Masterpiece of Nature: The Evolution and Genetics of Sexuality. Berkeley, University of California Press.

Bell, G. D., N. C. Kane, L. H. Rieseberg, and K. L. Adams. 2013. RNA-seq analysis of allele-specific expression, hybrid effects, and regulatory divergence in hybrids compared with their parents from natural populations. Genome Biology and Evolution 5:1309-1323.

Bell, J. T., A. A. Pai, J. K. Pickrell, D. J. Gaffney, R. Pique-Regi, J. F. Degner, Y. Gilad et al. 2011. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. Genome Biology 12:R10.

Berbel-Filho, W. M., C. G. de Leaniz, P. Morán, J. Cable, S. M. Lima, and S. Consuegra. 2019a. Local parasite pressures and host genotype modulate epigenetic diversity in a mixed-mating fish. Ecology and Evolution 9:8736-8748.

Berbel-Filho, W. M., H. M. V. Espirito-Santo, and S. M. Q. Lima. 2016. First record of a male of *Kryptolebias hermaphroditus* Costa, 2011 (Cyprinodontiformes: Cynolebiidae). Neotropical Ichthyology 14:e160024.

Berbel-Filho, W. M., D. Rodríguez-Barreto, N. Berry, C. Garcia de Leaniz, and S. Consuegra. 2019b. Contrasting DNA methylation responses of inbred fish lines to different rearing environments. Epigenetics 14:939-948.

Best, C., H. Ikert, D. J. Kostyniuk, P. M. Craig, L. Navarro-Martin, L. Marandel, and J. A. Mennigen. 2018. Epigenetics in teleost fish: From molecular mechanisms to physiological phenotypes. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology 224:210-244.

Birky, C. W., Jr. 1995. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. Proceedings of the National Academy of Sciences 92:11331-11338.

Blevins, T., J. Wang, D. Pflieger, F. Pontvianne, and C. S. Pikaard. 2017. Hybrid incompatibility caused by an epiallele. Proceedings of the National Academy of Sciences 114:3702-3707.

Blirt, A., and G. Bell. 1987. Mammalian chiasma frequencies as a test of two theories of recombination. Nature 326:803.

Bolnick, D. I., and W. E. Stutz. 2017. Frequency dependence limits divergent evolution by favouring rare immigrants over residents. Nature 546:285.

Bonduriansky, R., and T. Day. 2009. Nongenetic inheritance and its evolutionary implications. Annual Review of Ecology Evolution and Systematics 40:103-125.

Bonduriansky, R., and T. Day. 2018, Extended Heredity: a new understanding of inheritance and evolution, Princeton University Press.

Bonin, A., E. Bellemain, P. Bronken Eidesen, F. Pompanon, C. Brochmann, and P. Taberlet. 2004. How to track and assess genotyping errors in population genetics studies. Molecular Ecology 13:3261-3273.

Bossdorf, O., D. Arcuri, C. L. Richards, and M. Pigliucci. 2010. Experimental alteration of DNA methylation affects the phenotypic plasticity of ecologically relevant traits in *Arabidopsis thaliana*. Evolutionary Ecology 24:541-553.

Bossdorf, O., C. L. Richards, and M. Pigliucci. 2008. Epigenetics for ecologists. Ecology Letters 11:106-115.

Bouckaert, R., T. G. Vaughan, J. Barido-Sottani, S. Duchene, M. Fourment, A. Gavryushkina, J. Heled et al. 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLoS Computational Biology 15:e1006650.

Braithwaite, V. A., and A. G. Salvanes. 2005. Environmental variability in the early rearing environment generates behaviourally flexible cod: implications for rehabilitating wild populations. Proceedings of the Royal Society of London B: Biological Sciences 272:1107-1113.

Bressman, N. B., M. Simms, and M. A. Ashley-Ross. 2018. Where do fish go when stranded on land? Terrestrial orientation and navigation of the mangrove rivulus, *Kryptolebias marmoratus*. Integrative and Comparative Biology 58:E22-E22.

Britto, F. B., A. J. Schmidt, A. M. Carvalho, C. C. Vasconcelos, A. M. Farias, P. Bentzen, and F. M. Diniz. 2018. Population connectivity and larval dispersal of the exploited mangrove crab *Ucides cordatus* along the Brazilian coast. PeerJ 6:e4702.

Bull, J. W., and M. Maron. 2016. How humans drive speciation as well as extinction. Proceedings of the Royal Society B: Biological Sciences 283:20160600.

Burgess, S. C., and D. J. Marshall. 2014. Adaptive parental effects: the importance of estimating environmental predictability and offspring fitness appropriately. Oikos 123:769-776.

Burggren, W. 2016. Epigenetic inheritance and its role in evolutionary biology: re-evaluation and new perspectives. Biology 5:1-22.

Burnham, K. P., and D. R. Anderson. 2004. Multimodel inference - understanding AIC and BIC in model selection. Sociological Methods & Research 33:261-304.

Burton, T., and N. B. Metcalfe. 2014. Can environmental conditions experienced in early life influence future generations? Proceedings of the Royal Society B: Biological Sciences 281:20140311.

Bushnell, B. 2014. BBMap: a fast, accurate, splice-aware aligner, Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States).

Bygren, L. O. 2013. Intergenerational health responses to adverse and enriched environments. Annual Review of Public Health 34:49-60.

Calcagno, V., and C. de Mazancourt. 2010. glmulti: An R package for easy automated model selection with (generalized) linear models. Journal of Statistical Software 34:1-29.

Callahan, H. S., M. Pigliucci, and C. D. Schlichting. 1997. Developmental phenotypic plasticity: Where ecology and evolution meet molecular biology. Bioessays 19:519-525.

Cara, N., C. F. Marfil, and R. W. Masuelli. 2013. Epigenetic patterns newly established after interspecific hybridization in natural populations of *Solanum*. Ecology and Evolution 3:3764-3779.

Carreira, R. S., A. L. Wagener, and J. W. Readman. 2004. Sterols as markers of sewage contamination in a tropical urban estuary (Guanabara Bay, Brazil): space–time variations. Estuarine, Coastal and Shelf Science 60:587-598.

Castonguay, E., and B. Angers. 2012. The key role of epigenetics in the persistence of asexual lineages. Genetics Research International 2012:534289.

Champagne, F. A. 2008. Epigenetic mechanisms and the transgenerational effects of maternal care. Frontiers in Neuroendocrinology 29:386-397.

Champneys, T., G. Castaldo, S. Consuegra, and C. Garcia de Leaniz. 2018. Density-dependent changes in neophobia and stress-coping styles in the world's oldest farmed fish. Royal Society Open Science 5:181473.

Charlesworth, B., M. T. Morgan, and D. Charlesworth. 1993. The effect of deleterious mutations on neutral molecular variation. Genetics 134:1289-1303.

Charlesworth, D., N. H. Barton, and B. Charlesworth. 2017. The sources of adaptive variation. Proc Biol Sci 284.

Charlesworth, D., and J. H. Willis. 2009. The genetics of inbreeding depression. Nature Reviews Genetics 10:783-796.

Chen, Q., W. Yan, and E. Duan. 2016. Epigenetic inheritance of acquired traits through sperm RNAs and sperm RNA modifications. Nature Reviews Genetics 17:733.

Chodavarapu, R. K., S. Feng, B. Ding, S. A. Simon, D. Lopez, Y. Jia, G.-L. Wang et al. 2012. Transcriptome and methylome interactions in rice hybrids. Proceedings of the National Academy of Sciences 109:12040-12045.

Cochrane, P. V., G. S. Rossi, L. Tunnah, M. G. Jonz, and P. A. Wright. 2019. Hydrogen sulphide toxicity and the importance of amphibious behaviour in a mangrove fish inhabiting sulphide-rich habitats. Journal of Comparative Physiology B:1-13.

Costa, W. J., P. F. Amorim, and J. L. Mattos. 2016. Molecular phylogeny and evolution of internal fertilization in South American seasonal cynopoeciline killifishes. Molecular Phylogenetics and Evolution 95:94-99.

Costa, W. J. E. M. 2006. Redescription of *Kryptolebias ocellatus* (Hensel) and *K. caudomarginatus* (Seegers)(Teleostei: Cyprinodontiformes: Rivulidae), two killifishes from mangroves of south-eastern Brazil. Aqua: Journal of Ichthyology & Aquatic Biology 11:5-13.

—. 2011a. Identity of *Rivulus ocellatus* and a new name for a hermaphroditic species of *Kryptolebias* from south-eastern Brazil (Cyprinodontiformes: Rivulidae). Ichthyological Exploration of Freshwaters 22:185-192.

—. 2011b. Phylogenetic position and taxonomic status of *Anablepsoides*, *Atlantirivulus*, *Cynodonichthys*, *Laimosemion* and *Melanorivulus*

(Cyprinodontiformes: Rivulidae). Ichthyological Exploration of Freshwaters 22:233-249.

—. 2016. Colouration, taxonomy and geographical distribution of mangrove killifishes, the *Kryptolebias marmoratus* species group, in southern Atlantic coastal plains of Brazil (Cyprinodontiformes: Rivulidae). Ichthyological Exploration of Freshwaters 27:183-192.

Costa, W. J. E. M., S. M. Q. Lima, and R. Bartolette. 2010. Androdioecy in *Kryptolebias* killifish and the evolution of self-fertilizing hermaphroditism. Biological Journal of the Linnean Society 99:344-349.

Cubas, P., C. Vincent, and E. Coen. 1999. An epigenetic mutation responsible for natural variation in floral symmetry. Nature 401:157.

D'Anna, G., V. M. Giacalone, T. V. Fernández, A. M. Vaccaro, C. Pipitone, S. Mirto, S. Mazzola et al. 2012. Effects of predator and shelter conditioning on hatchery-reared white seabream *Diplodus sargus* (L., 1758) released at sea. Aquaculture 356:91-97.

Damgaard, C., D. Couvet, and V. Loeschcke. 1992. Partial selfing as an optimal mating strategy. Heredity 69:289-295.

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. Nature Reviews Genetics 12:475.

Danchin, E., A. Pocheville, O. Rey, B. Pujol, and S. Blanchet. 2019. Epigenetically facilitated mutational assimilation: epigenetics as a hub within the inclusive evolutionary synthesis. Biological Reviews 94:259-282.

Darriba, D., G. L. Taboada, R. Doallo, and D. Posada. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nature Methods 9:772.

Davis, W. P., D. S. Taylor, and B. Turner. 1990. Field observations of the ecology and habits of mangrove rivulus (*Rivulus marmoratus*) in Belize and Florida (Teleostei: Cyprinodontiformes: Rivulidae). Ichthyological Exploration of Freshwaters 1:123-134.

Daxinger, L., and E. Whitelaw. 2012. Understanding transgenerational epigenetic inheritance via the gametes in mammals. Nature Reviews Genetics 13:153-162.

Dell, P. A., and F. D. Rose. 1987. Transfer of Effects from Environmentally Enriched and Impoverished Female Rats to Future Offspring. Physiology & Behavior 39:187-190.

Des Marais, D. L., K. M. Hernandez, and T. E. Juenger. 2013. Genotype-by-environment interaction and plasticity: exploring genomic responses of plants to the abiotic environment. Annual Review of Ecology, Evolution, and Systematics 44:5-29.

Dickins, T. E., and Q. Rahman. 2012. The extended evolutionary synthesis and the role of soft inheritance in evolution. Proceedings of the Royal Society B 279:2913-2921.

Dieringer, D., and C. Schlötterer. 2003. Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. Molecular Ecology Notes 3:167-169.

Douhovnikoff, V., and R. S. Dodd. 2015. Epigenetics: a potential mechanism for clonal plant success. Plant Ecology 216:227-233.

Dowen, R. H., M. Pelizzola, R. J. Schmitz, R. Lister, J. M. Dowen, J. R. Nery, J. E. Dixon et al. 2012. Widespread dynamic DNA methylation in response to biotic stress. Proceedings of the National Academy of Sciences 109:E2183-E2191.

Dubin, M. J., P. Zhang, D. Meng, M.-S. Remigereau, E. J. Osborne, F. Paolo Casale, P. Drewe et al. 2015. DNA methylation in Arabidopsis has a genetic basis and shows evidence of local adaptation. eLife 4:e05255.

Duke, N. C., J. O. Meynecke, S. Dittmann, A. M. Ellison, K. Anger, U. Berger, S. Cannicci et al. 2007. A world without mangroves? Science 317:41-42.

Duncan, E. J., P. D. Gluckman, and P. K. Dearden. 2014. Epigenetics, plasticity, and evolution: How do we link epigenetic change to phenotype? Journal of Experimental Zoology Part B: Molecular and Developmental Evolution 322:208-220.

Earl, D. A. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4:359-361.

Earley, R. L., A. F. Hanninen, A. Fuller, M. J. Garcia, and E. A. Lee. 2012. Phenotypic plasticity and integration in the mangrove rivulus (*Kryptolebias marmoratus*): a prospectus. Integrative and Comparative Biology 52:814-827.

Earley, R. L., and Y. Hsu. 2008. Reciprocity between endocrine state and contest behavior in the killifish, Kryptolebias marmoratus. Hormones and behavior 53:442-451.

Ebert, D. 2008. Host-parasite coevolution: Insights from the *Daphnia*-parasite model system. Current Opinion in Microbiology 11:290-301.

Edenbrow, M., and D. Croft. 2012. Kin and familiarity influence association preferences and aggression in the mangrove killifish *Kryptolebias marmoratus*. Journal of Fish Biology 80:503-518.

Edenbrow, M., and D. P. Croft. 2013. Environmental and genetic effects shape the development of personality traits in the mangrove killifish *Kryptolebias marmoratus*. Oikos 122:667-681.

Eichten, S. R., R. Briskine, J. Song, Q. Li, R. Swanson-Wagner, P. J. Hermanson, A. J. Waters et al. 2013. Epigenetic and genetic influences on DNA methylation variation in maize populations. The Plant Cell 25:2783-2797.

Ellison, A., J. Cable, and S. Consuegra. 2011. Best of both worlds? Association between outcrossing and parasite loads in a selfing fish. Evolution 65:3021-3026.

Ellison, A., J. Jones, C. Inchley, and S. Consuegra. 2013. Choosy males could help explain androdioecy in a selfing fish. The American Naturalist 181:855-862.

Ellison, A., C. M. R. López, P. Moran, J. Breen, M. Swain, M. Megias, M. Hegarty et al. 2015. Epigenetic regulation of sex ratios may explain natural variation in self-fertilization rates. Proceedings of Royal Society B:20151900.

Ellison, A., P. Wright, D. S. Taylor, C. Cooper, K. Regan, S. Currie, and S. Consuegra. 2012. Environmental diel variation, parasite loads, and local population structuring of a mixed-mating mangrove fish. Ecology and Evolution 2:1682-1695.

English, S., I. Pen, N. Shea, and T. Uller. 2015. The information value of non-genetic inheritance in plants and animals. PLoS One 10:e0116996.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14:2611-2620.

Evans, M. L., T. S. Hori, M. L. Rise, and I. A. Fleming. 2015. Transcriptomic responses of Atlantic Salmon (*Salmo salar*) to environmental enrichment during juvenile rearing. PloS one 10:e0118378.

Excoffier, L., and H. E. Lischer. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10:564-567.

Farias, C. O., C. Hamacher, A. d. L. R. Wagener, R. C. d. Campos, and J. M. Godoy. 2007. Trace metal contamination in mangrove sediments, Guanabara Bay, Rio de Janeiro, Brazil. Journal of the Brazilian Chemical Society 18:1194-1206.

Fellous, A., T. Labed-Veydert, M. Locrel, A. S. Voisin, R. L. Earley, and F. Silvestre. 2018. DNA methylation in adults and during development of the self-fertilizing mangrove rivulus, *Kryptolebias marmoratus*. Ecology and Evolution 8:6016-6033.

Feng, S., S. J. Cokus, X. Zhang, P.-Y. Chen, M. Bostick, M. G. Goll, J. Hetzel et al. 2010. Conservation and divergence of methylation patterning in plants and animals. Proceedings of the National Academy of Sciences 107:8689-8694.

Figueiredo, L. H., A. d. L. Wagener, J. Dagaut, and A. Saliot. 2008. Non-aromatic hydrocarbons in recent sediments of Sepetiba and Ilha Grande Bays, Brazil. Journal of the Brazilian Chemical Society 19:516-527.

Fischer, P. 2000. An experimental test of metabolic and behavioural responses of benthic fish species to different types of substrate. Canadian Journal of Fisheries and Aquatic Sciences 57:2336-2344.

Foll, M., and O. Gaggiotti. 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics 180:977-993.

Francisco, P. M., G. M. Mori, F. M. Alves, E. V. Tambarussi, and A. P. de Souza. 2018. Population genetic structure, introgression, and hybridization in the genus *Rhizophora* along the Brazilian coast. Ecology and Evolution 8:3491-3504.

Friard, O., and M. Gamba. 2016. BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. Methods in Ecology and Evolution 7:1325-1330.

Frick, N., and P. Wright. 2002. Nitrogen metabolism and excretion in the mangrove killifish *Rivulus marmoratus* II. Significant ammonia volatilization in a teleost during air-exposure. Journal of Experimental Biology 205:91-100.

Fricke, R. 2019. Eschmeyer's catalog of fishes: genera, species, references. http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp, California Academy of Natural Sciences San Francisco.

Furness, A. I., A. Tatarenkov, and J. C. Avise. 2015. A genetic test for whether pairs of hermaphrodites can cross-fertilize in a selfing killifish. Journal of Heredity 106:749-752.

Fusco, G., and A. Minelli. 2010. Phenotypic plasticity in development and evolution: facts and concepts. Philosophical Transactions of the Royal Society B: Biological Sciences 365:547-556.

Gao, H., S. Williamson, and C. D. Bustamante. 2007. A Markov chain Monte Carlo approach for joint inference of population structure and inbreeding rates from multilocus genotype data. Genetics 176:1635-1651.

Gao, L., Y. Geng, B. Li, J. Chen, and J. Yang. 2010. Genome-wide DNA methylation alterations of *Alternanthera philoxeroides* in natural and manipulated habitats: implications for epigenetic regulation of rapid responses to environmental

fluctuation and phenotypic variation. Plant, Cell & Environment 33:1820-1827.

Garver, K. L., and B. Garver. 1991. Eugenics: past, present, and the future. American Journal of Human Genetics 49:1109.

Gertz, J., K. E. Varley, T. E. Reddy, K. M. Bowling, F. Pauli, S. L. Parker, K. S. Kucera et al. 2011. Analysis of DNA methylation in a three-generation family reveals widespread genetic influence on epigenetic regulation. PLoS Genetics 7:e1002228.

Gibb, A. C., M. A. Ashley-Ross, and S. T. Hsieh. 2013. Thrash, flip, or jump: the behavioral and functional continuum of terrestrial locomotion in teleost fishes. Integrative Comparative Biology 53:295-306.

Gibson, D. J., E. V. Sylvester, A. J. Turko, G. J. Tattersall, and P. A. Wright. 2015. Out of the frying pan into the air—emersion behaviour and evaporative heat loss in an amphibious mangrove fish (*Kryptolebias marmoratus*). Biology Letters 11:20150689.

Gluckman, P. D., M. A. Hanson, H. G. Spencer, and P. Bateson. 2005. Environmental influences during development and their later consequences for health and disease: implications for the interpretation of empirical studies. Proceedings of the Royal Society B 272:671-677.

Goudet, J. 1995. FSTAT (version 1.2): a computer program to calculate F-statistics. Journal of Heredity 86:485-486.

Greaves, I. K., M. Groszmann, H. Ying, J. M. Taylor, W. J. Peacock, and E. S. Dennis. 2012. Trans chromosomal methylation in *Arabidopsis* hybrids. Proceedings of the National Academy of Sciences 109:3570-3575.

Groot, M. P., N. Wagemaker, N. J. Ouborg, K. J. F. Verhoeven, and P. Vergeer. 2018. Epigenetic population differentiation in field- and common garden-grown *Scabiosa columbaria* plants. Ecology and Evolution 8:3505-3517.

Groszmann, M., I. K. Greaves, R. Fujimoto, W. J. Peacock, and E. S. Dennis. 2013. The role of epigenetics in hybrid vigour. Trends in Genetics 29:684-690.

Guillot, G., F. Mortier, and A. Estoup. 2005. GENELAND: a computer package for landscape genetics. Molecular Ecology Notes 5:712-715.

Guillot, G., F. Santos, and A. Estoup. 2008. Analysing georeferenced population genetics data with Geneland: a new algorithm to deal with null alleles and a friendly graphical user interface. Bioinformatics 24:1406-1407.

Guimarães-Costa, A., H. Schneider, and I. Sampaio. 2017. New record of the mangrove rivulid *Kryptolebias hermaphroditu*s Costa, 2011 (Cyprinodontiformes: Cynolebiidae) in the Pará state, northern Brazil. Check List 13:2093.

Gusmão, J., C. Lazoski, and A. M. Solé-Cava. 2005. Population genetic structure of Brazilian shrimp species (*Farfantepenaeus* sp., *F. brasiliensis*, *F. paulensis* and *Litopenaeus schmitti*: Decapoda: Penaeidae). Genetics and Molecular Biology 28:165-171.

Gutierrez-Arcelus, M., T. Lappalainen, S. B. Montgomery, A. Buil, H. Ongen, A. Yurovsky, J. Bryois et al. 2013. Passive and active DNA methylation and the interplay with genetic variation in gene regulation. eLife 2:e00523.

Hamilton, W. D. 1980. Sex versus non-sex versus parasite. Oikos 35:282-290.

Harrington, R. W. 1961. Oviparous hermaphroditic fish with internal self-fertilization. Science 134:1749-1750.

—. 1967. Environmentally controlled induction of primary male gonochorists from eggs of the self-fertilizing hermaphroditic fish, *Rivulus marmoratus* Poey. The Biological Bulletin 132:174-199.

Harrington, R. W., and K. D. Kallman. 1968. The homozygosity of clones of the self-fertilizing hermaphroditic fish *Rivulus marmoratus* Poey (Cyprinodontidae, Atheriniformes). The American Naturalist 102:337-343.

Hassen, M. B., J. Bartholomé, G. Valè, T.-V. Cao, and N. Ahmadi. 2018. Genomic prediction accounting for genotype by environment interaction offers an effective framework for breeding simultaneously for adaptation to an abiotic stress and performance under normal cropping conditions in rice. G3: Genes, Genomes, Genetics 8:2319-2332.

He, G., H. He, and X. W. Deng. 2013. Epigenetic variations in plant hybrids and their potential roles in heterosis. Journal of Genetics and Genomics 40:205-210.

He, G., X. Zhu, A. A. Elling, L. Chen, X. Wang, L. Guo, M. Liang et al. 2010. Global epigenetic and transcriptional trends among two rice subspecies and their reciprocal hybrids. The Plant Cell 22:17-33.

Heard, E., and R. A. Martienssen. 2014. Transgenerational epigenetic inheritance: myths and mechanisms. Cell 157:95-109.

Hegarty, M. J., T. Batstone, G. L. Barker, K. J. Edwards, R. J. Abbott, and S. J. Hiscock. 2011. Nonadditive changes to cytosine methylation as a consequence of hybridization and genome duplication in *Senecio* (Asteraceae). Molecular Ecology 20:105-113.

Heino, M., B. D. Pauli, and U. Dieckmann. 2015. Fisheries-induced evolution. Annual Review of Ecology, Evolution, and Systematics 46.

Hensel, R. F. 1868. Beiträge zur kenntniss der wirbelthiere Südbrasiliens. Archiv für Naturgeschichte 34:323-375.

Herman, J. J., and S. E. Sultan. 2016. DNA methylation mediates genetic variation for adaptive transgenerational plasticity. Proceeding of the Royal Society B 283:20160988.

Herrera, C. M., and P. Bazaga. 2010. Epigenetic differentiation and relationship to adaptive genetic divergence in discrete populations of the violet *Viola cazorlensis*. New Phytologist 187:867-876.

Herrera, C. M., M. Medrano, and P. Bazaga. 2016. Comparative spatial genetics and epigenetics of plant populations: heuristic value and a proof of concept. Molecular Ecology 25:1653-1664.

Herten, K., M. S. Hestand, J. R. Vermeesch, and J. K. Van Houdt. 2015. GBSX: a toolkit for experimental design and demultiplexing genotyping by sequencing experiments. BMC Bioinformatics 16:73.

Horvath, R., B. Laenen, S. Takuno, and T. Slotte. 2019. Single-cell expression noise and gene-body methylation in *Arabidopsis thaliana*. Heredity 123:81-91.

Hovick, S. M., and K. D. Whitney. 2014. Hybridisation is associated with increased fecundity and size in invasive taxa: meta-analytic support for the hybridisation-invasion hypothesis. Ecology Letters 17:1464-1477.

Huber, J. 2017. Reanalysis of single type of *Rivulus ocellatus* Hensel, 1880 in Berlin Museum pending its putative molecular analysis, with the proposal of revalidation of *Rivulus caudomarginatus*. Killi-Data Series 2016-2017:4-12.

Hughes, K. A., J. F. Ayroles, M. M. Reedy, J. M. Drnevich, K. C. Rowe, E. A. Ruedi, C. E. Caceres et al. 2006. Segregating variation in the transcriptome: cis regulation and additivity of effects. Genetics 173:1347-1355.

Huh, I., J. Zeng, T. Park, and S. V. Yi. 2013. DNA methylation and transcriptional noise. Epigenetics & Chromatin 6:1-10.

Hunter, D. J. 2005. Gene–environment interactions in human diseases. Nature Reviews Genetics 6:287.

Huxley, J. 1942. Evolution. The modern synthesis. Evolution. The Modern Synthesis.

Illum, L. R. H., S. T. Bak, S. Lund, and A. L. Nielsen. 2018. DNA methylation in epigenetic inheritance of metabolic diseases through the male germ line. Journal of Molecular Endocrinology 60:R39-R56.

Iqbal, K., D. A. Tran, A. X. Li, C. Warden, A. Y. Bai, P. Singh, X. Wu et al. 2015. Deleterious effects of endocrine disruptors are corrected in the mammalian germline by epigenome reprogramming. Genome Biology 16:59.

Jablonka, E. 2006. Developmental plasticity and evolution. Biology & Philosophy 21:143-154.

Jablonka, E., and G. Raz. 2009. Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. The Quarterly Review of Biology 84:131-176.

Jarne, P., and J. R. Auld. 2006. Animals mix it up too: the distribution of self-fertilization among hermaphroditic animals. Evolution 60:1816-1824.

Jarne, P., and D. Charlesworth. 1993. The evolution of the selfing rate in functionally hermaphrodite plants and animals. Annual Review of Ecology and Systematics 24:441-466.

Jensen, N., R. M. Allen, and D. J. Marshall. 2014. Adaptive maternal and paternal effects: gamete plasticity in response to parental stress. Functional Ecology 28:724-733.

Jeremias, G., J. Barbosa, S. M. Marques, J. Asselman, F. J. Gonçalves, and J. L. Pereira. 2018. Synthesizing the role of epigenetics in the response and adaptation of species to climate change in freshwater ecosystems. Molecular Ecology 27:2790-2806.

Jiang, L., J. Zhang, J.-J. Wang, L. Wang, L. Zhang, G. Li, X. Yang et al. 2013. Sperm, but not oocyte, DNA methylome is inherited by zebrafish early embryos. Cell 153:773-784.

Johnson, N. A. 2000. Speciation: Dobzhansky–Muller incompatibilities, dominance and gene interactions. Trends in Ecology & Evolution 15:480-482.

Jones, P. A. 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nature Reviews Genetics 13:484-492.

Jonsson, B., and N. Jonsson. 2014. Early environment influences later performance in fishes. Journal of Fish Biology 85:151-188.

Judson, O. P. 1997. A model of asexuality and clonal diversity: cloning the red queen. Journal of Theoretical Biology 186:33-40.

Kanamori, A., Y. Sugita, Y. Yuasa, T. Suzuki, K. Kawamura, Y. Uno, K. Kamimura et al. 2016. A genetic map for the only self-fertilizing vertebrate. G3: Genes, Genomes, Genetics 6:1095-1106.

Keller, T. E., J. R. Lasky, and S. V. Yi. 2016. The multivariate association between genomewide DNA methylation and climate across the range of *Arabidopsis thaliana*. Molecular Ecology 25:1823-1837.

Kevles, D. J., and L. E. Hood. 1992, The code of codes: Scientific and social issues in the human genome project, Harvard University Press.

Kihslinger, R. L., S. C. Lema, and G. A. Nevitt. 2006. Environmental rearing conditions produce forebrain differences in wild Chinook salmon

Oncorhynchus tshawytscha. Comparative Biochemistry and Physiology A 145:145-151.

King, K. C., J. Jokela, and C. M. Lively. 2011. Parasites, sex, and clonal diversity in natural snail populations. Evolution 65:1474-1481.

Kitimu, S. R., J. Taylor, T. J. March, F. Tairo, M. J. Wilkinson, and C. M. Rodríguez López. 2015. Meristem micropropagation of cassava (*Manihot esculenta)* evokes genome-wide changes in DNA methylation. Frontiers in Plant Science 6:590.

Konate, M., M. Wilkinson, B. Mayne, S. Pederson, E. Scott, B. Berger, and C. Rodriguez Lopez. 2018. Salt stress induces non-CG methylation in coding regions of barley seedlings (*Hordeum vulgare*). Epigenomes 2:1-12.

Kopelman, N. M., J. Mayzel, M. Jakobsson, N. A. Rosenberg, and I. Mayrose. 2015. CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. Molecular Ecology Resources 15:1179-1191.

Korneliussen, T. S., A. Albrechtsen, and R. Nielsen. 2014. ANGSD: analysis of next generation sequencing data. BMC Bioinformatics 15:356.

Korneliussen, T. S., I. Moltke, A. Albrechtsen, and R. Nielsen. 2013. Calculation of Tajima's D and other neutrality test statistics from low depth next-generation sequencing data. BMC Bioinformatics 14:289.

Krueger, F. 2016. TrimGalore! A wrapper around cutadapt and FastQC to consistently apply adapter and quality trimming to FastQ files, with extra functionality for RRBS data. . https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.

Krueger, F., and S. R. Andrews. 2011. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics 27:1571-1572.

Kumar, S., G. Stecher, and K. Tamura. 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Molecular Biology and Evolution 33:1870-1874.

Laland, K., T. Uller, M. Feldman, K. Sterelny, G. B. Müller, A. Moczek, E. Jablonka et al. 2014. Does evolutionary theory need a rethink? Nature News 514:161.

Laland, K. N., T. Uller, M. W. Fellman, K. Sterelny, G. B. Muller, A. Moczek, E. Jablonka et al. 2015. The extended evolutionary synthesis: its structure, assumptions and predictions. Proceedings of the Royal Society B 282:20151019.

Landry, C., D. Hartl, and J. Ranz. 2007. Genome clashes in hybrids: insights from gene expression. Heredity 99:483-493.

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

Lauss, K., R. Wardenaar, R. Oka, M. H. van Hulten, V. Guryev, J. J. Keurentjes, M. Stam et al. 2018. Parental DNA methylation states are associated with heterosis in epigenetic hybrids. Plant Physiology 176:1627-1645.

Le, T. N., U. Schumann, N. A. Smith, S. Tiwari, P. C. K. Au, Q. H. Zhu, J. M. Taylor et al. 2014. DNA demethylases target promoter transposable elements to positively regulate stress responsive genes in *Arabidopsis*. Genome Biology 15:458.

Lea, A. J., T. P. Vilgalys, P. A. P. Durst, and J. Tung. 2017. Maximizing ecological and evolutionary insight in bisulfite sequencing data sets. Nature Ecology & Evolution 1:1074-1083.

Lee, J. S., S. Raisuddin, and D. Schlenk. 2008. *Kryptolebias marmoratus* (Poey, 1880): a potential model species for molecular carcinogenesis and ecotoxicogenomics. Journal of Fish Biology 72:1871-1889.

Leigh, J. W., and D. Bryant. 2015. popart: full-feature software for haplotype network construction. Methods in Ecology and Evolution 6:1110-1116.

Lema, S. C., M. J. Hodges, M. P. Marchetti, and G. A. Nevitt. 2005. Proliferation zones in the salmon telencephalon and evidence for environmental influence on proliferation rate. Comparative Biochemistry and Physiology A 141:327-335.

Leung, C., S. Breton, and B. Angers. 2016. Facing environmental predictability with different sources of epigenetic variation. Ecology and Evolution 6:5234-5245.

Levis, N. A., A. J. Isdaner, and D. W. Pfennig. 2018. Morphological novelty emerges from pre-existing phenotypic plasticity. Nature Ecology and Evolution 2:1289-1297.

Lewontin, R. C., and L. C. Birch. 1966. Hybridization as a source of variation for adaptation to new environments. Evolution 20:315-336.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth et al. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078-2079.

Li, Y. L., and J. X. Liu. 2018. StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. Molecular Ecology Resources 18:176-177.

Liebl, A. L., A. W. Schrey, C. L. Richards, and L. B. Martin. 2013. Patterns of DNA methylation throughout a range expansion of an introduced songbird. Integrative and Comparative Biology 53:351-358.

Lima, S. M. Q., W. M. Berbel-Filho, T. F. P. Araújo, H. Lazzarotto, A. Tatarenkov, and J. C. Avise. 2017. Headwater capture evidenced by paleo-rivers reconstruction and population genetic structure of the armored catfish (*Pareiorhaphis garbei*) in the Serra do Mar mountains of southeastern brazil. Frontiers in Genetics 8.

Lin, H.-C., and W. A. Dunson. 1995. An explanation of the high strain diversity of a self-fertilizing hermaphroditic fish. Ecology 76:593-605.

Lira, M. G. S. L., R. E. C. Paiva, T. P. A. Ramos, and S. M. Q. Lima. 2015. First record of *Kryptolebias hermaphroditus* Costa, 2011 (Cyprinodontiformes: Rivulidae) in the extreme north Atlantic Forest mangroves, Rio Grande do Norte state, Brazil. Checklist 11:1-5.

Lively, C. M. 1987. Evidence from a New-Zealand snail for the maintenance of sex by parasitism. Nature 328:519-521.

Lively, C. M., C. Craddock, and R. C. Vrijenhoek. 1990. Red Queen hypothesis supported by parasitism in sexual and clonal fish. Nature 344:864.

Lively, C. M., and L. T. Morran. 2014. The ecology of sexual reproduction. Journal of Evolutionary Biology 27:1292-1303.

Lomax, J. L., R. E. Carlson, J. W. Wells, P. M. Crawford, and R. L. Earley. 2017. Factors affecting egg production in the selfing mangrove rivulus (*Kryptolebias marmoratus*). Zoology 122:38-45.

Loureiro, M., R. d. Sá, S. W. Serra, F. Alonso, L. E. K. Lanés, M. V. Volcan, P. Calviño et al. 2018. Review of the family Rivulidae (Cyprinodontiformes, Aplocheiloidei) and a molecular and morphological phylogeny of the annual fish genus *Austrolebias* Costa 1998. Neotropical Ichthyology 16:e180007.

Mackiewicz, M., A. Tatarenkov, A. Perry, J. R. Martin, J. F. Elder, D. L. Bechler, and J. C. Avise. 2006a. Microsatellite documentation of male-mediated outcrossing between inbred laboratory strains of the self-fertilizing mangrove killifish (*Kryptolebias marmoratus*). Journal of Heredity 97:508-513.

Mackiewicz, M., A. Tatarenkov, D. S. Taylor, B. J. Turner, and J. C. Avise. 2006b. Extensive outcrossing and androdioecy in a vertebrate species that otherwise reproduces as a self-fertilizing hermaphrodite. Proceedings of the National Academy of Sciences 103:9924-9928.

Mackiewicz, M., A. Tatarenkov, B. J. Turner, and J. C. Avise. 2006c. A mixed-mating strategy in a hermaphroditic vertebrate. Proceeding of the Royal Society B 273:2449-2452.

Mallet, J. 2005. Hybridization as an invasion of the genome. Trends in Ecology & Evolution 20:229-237.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. Cancer Research 27:209-220.

Marson, K. M., D. S. Taylor, and R. L. Earley. 2019. Cryptic male phenotypes in the mangrove rivulus fish, *Kryptolebias marmoratus*. The Biological Bulletin 236:000-000.

Massicotte, R., and B. Angers. 2012. General-purpose genotype or how epigenetics extend the flexibility of a genotype. Genetics Research International 2012:1-7.

Massicotte, R., E. Whitelaw, and B. Angers. 2011. DNA methylation: a source of random variation in natural populations. Epigenetics 6:421-427.

Maynard-Smith, J. 1978, The evolution of sex. Cambridge, United Kingdom, Cambridge University Press Cambridge.

Mayne, B. T., S. Y. Leemaqz, S. Buckberry, C. M. Rodriguez Lopez, C. T. Roberts, T. Bianco-Miotto, and J. Breen. 2018. msgbsR: An R package for analysing methylation-sensitive restriction enzyme sequencing data. Scientific Reports 8:2190.

Mazué, G. P., F.-X. Dechaume-Moncharmont, and J.-G. J. Godin. 2015. Boldness–exploration behavioral syndrome: interfamily variability and repeatability of personality traits in the young of the convict cichlid (*Amatitlania siquia*). Behavioral Ecology 26:900-908.

McGirr, J. A., and C. H. Martin. 2019. Hybrid gene misregulation in multiple developing tissues within a recent adaptive radiation of Cyprinodon pupfishes. PloS One 14:e0218899.

Metcalfe, C. J., K. V. Bulazel, G. C. Ferreri, E. Schroeder-Reiter, G. Wanner, W. Rens, C. Obergfell et al. 2007. Genomic instability within centromeres of interspecific marsupial hybrids. Genetics 177:2507-2517.

Mi, H., X. Huang, A. Muruganujan, H. Tang, C. Mills, D. Kang, and P. D. Thomas. 2016. PANTHER version 11: expanded annotation data from gene ontology and reactome pathways, and data analysis tool enhancements. Nucleic Acids Research 45:D183-D189.

Michalak, P. 2009. Epigenetic, transposon and small RNA determinants of hybrid dysfunctions. Heredity 102:45-50.

Michalak, P., and M. A. Noor. 2003. Genome-wide patterns of expression in *Drosophila* pure species and hybrid males. Molecular Biology and Evolution 20:1070-1076.

Millidine, K., J. Armstrong, and N. Metcalfe. 2006. Presence of shelter reduces maintenance metabolism of juvenile salmon. Functional Ecology 20:839-845.

Mitchell, N., G. L. Owens, S. M. Hovick, L. H. Rieseberg, and K. D. Whitney. 2019. Hybridization speeds adaptive evolution in an eight-year field experiment. Scientific Reports 9:6746.

Moberg, O., V. A. Braithwaite, K. H. Jensen, and A. G. V. Salvanes. 2011. Effects of habitat enrichment and food availability on the foraging behaviour of juvenile Atlantic Cod (*Gadus morhua* L). Environmental Biology of Fishes 91:449-457.

Moehring, A. J., K. C. Teeter, and M. A. Noor. 2007. Genome-wide patterns of expression in *Drosophila* pure species and hybrid males. II. Examination of multiple-species hybridizations, platforms, and life cycle stages. Molecular Biology and Evolution 24:137-145.

Moghaddam, A. M. B., F. Roudier, M. Seifert, C. Bérard, M. L. M. Magniette, R. K. Ashtiyani, A. Houben et al. 2011. Additive inheritance of histone modifications in *Arabidopsis thaliana* intra-specific hybrids. The Plant Journal 67:691-700.

Moore, L. D., T. Le, and G. Fan. 2013. DNA methylation and its basic function. Neuropsychopharmacology 38:23-38.

Morgan, H. D., H. G. Sutherland, D. I. Martin, and E. Whitelaw. 1999. Epigenetic inheritance at the agouti locus in the mouse. Nature Genetics 23:314.

Mori, G. M., M. I. Zucchi, and A. P. Souza. 2015. Multiple-geographic-scale genetic structure of two mangrove tree species: the roles of mating system, hybridization, limited dispersal and extrinsic factors. PLoS One 10:e0118710.

Morran, L. T., O. G. Schmidt, I. A. Gelarden, R. C. Parrish, 2nd, and C. M. Lively. 2011. Running with the Red Queen: host-parasite coevolution selects for biparental sex. Science 333:216-218.

Murphy, W. J., J. E. Thomerson, and G. E. Collier. 1999. Phylogeny of the Neotropical killifish family Rivulidae (Cyprinodontiformes, Aplocheiloidei) inferred from mitochondrial DNA sequences. Molecular Phylogenetics and Evolution 13:289-301.

Mychasiuk, R., S. Zahir, N. Schmold, S. Ilnytskyy, O. Kovalchuk, and R. Gibb. 2012. Parental enrichment and offspring development: Modifications to brain, behavior and the epigenome. Behavioural Brain Research 228:294-298.

Nakamura, S., and K. Hosaka. 2010. DNA methylation in diploid inbred lines of potatoes and its possible role in the regulation of heterosis. Theoretical and Applied Genetics 120:205-214.

Näslund, J., K. Aarestrup, S. T. Thomassen, and J. I. Johnsson. 2012. Early enrichment effects on brain development in hatchery-reared Atlantic salmon (Salmo salar): no evidence for a critical period. Canadian Journal of Fisheries and Aquatic Sciences 69:1481-1490.

Näslund, J., and J. I. Johnsson. 2016. Environmental enrichment for fish in captive environments: effects of physical structures and substrates. Fish and Fisheries 17:1-30.

Näslund, J., M. Rosengren, D. Del Villar, L. Gansel, J. R. Norrgård, L. Persson, J. J. Winkowski et al. 2013. Hatchery tank enrichment affects cortisol levels and shelter-seeking in Atlantic salmon (*Salmo salar*). Canadian Journal of Fisheries and Aquatic Sciences 70:585-590.

Negri, I., A. Franchini, E. Gonella, D. Daffonchio, P. J. Mazzoglio, M. Mandrioli, and A. Alma. 2009. Unravelling the *Wolbachia* evolutionary role: the reprogramming of the host genomic imprinting. Proceedings of the Royal Society B 276:2485-2491.

Neiman, M., J. Jokela, and C. M. Lively. 2005. Variation in asexual lineage age in *Potamopyrgus antipodarum*, a New Zealand snail. Evolution 59:1945-1952.

Neto, J. A. B., F. X. Gingele, T. Leipe, and I. Brehme. 2006. Spatial distribution of heavy metals in surficial sediments from Guanabara Bay: Rio de Janeiro, Brazil. Environmental Geology 49:1051-1063.

Oliveira-Neto, J. F., W. A. Boeger, M. R. Pie, A. Ostrensky, and D. B. Hungria. 2007. Genetic structure of populations of the mangrove crab *Ucides cordatus* (Decapoda: Ocypodidae) at local and regional scales. Hydrobiologia 583:69-76.

Oliveira-Neto, J. F., M. R. Pie, M. A. Chammas, A. Ostrensky, and W. A. Boeger. 2008. Phylogeography of the blue land crab, *Cardisoma guanhumi* (Decapoda: Gecarcinidae) along the Brazilian coast. Journal of the Marine Biological Association of the United Kingdom 88:1417-1423.

Oziolor, E. M., N. M. Reid, S. Yair, K. M. Lee, S. G. VerPloeg, P. C. Bruns, J. R. Shaw et al. 2019. Adaptive introgression enables evolutionary rescue from extreme environmental pollution. Science 364:455-457.

Pawluk, R. J., T. M. Uren Webster, J. Cable, C. Garcia de Leaniz, and S. Consuegra. 2018. Immune-related transcriptional responses to parasitic infection in a naturally inbred fish: roles of genotype and individual variation. Genome Biology and Evolution 1:319-327.

Peakall, R., and P. E. Smouse. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. Bioinformatics 28:2537-2539.

Pérez-Figueroa, A. 2013. msap: a tool for the statistical analysis of methylation-sensitive amplified polymorphism data. Molecular Ecology Resources 13:522-527.

Perez, M. F., and B. Lehner. 2019. Intergenerational and transgenerational epigenetic inheritance in animals. Nature cell biology 21:143-151.

Pigliucci, M. 2007. Do we need an extended evolutionary synthesis? Evolution 61:2743-2749.

Pigliucci, M., and G. Muller. 2010, Evolution–the extended synthesis. Cambridge, MA, MIT press.

Pigliucci, M., C. J. Murren, and C. D. Schlichting. 2006. Phenotypic plasticity and evolution by genetic assimilation. Journal of Experimental Biology 209:2362-2367.

Pil, M. W., M. R. Boeger, V. C. Muschner, M. R. Pie, A. Ostrensky, and W. A. Boeger. 2011. Postglacial north–south expansion of populations of *Rhizophora mangle* (Rhizophoraceae) along the Brazilian coast revealed by microsatellite analysis. American Journal of Botany 98:1031-1039.

Pimm, S. L., and P. Raven. 2000. Biodiversity: extinction by numbers. Nature 403:843.

Poey, F. 1880, Revisio piscium cubensium, Sociedad Española de Historia Natural.

Poland, J. A., P. J. Brown, M. E. Sorrells, and J.-L. Jannink. 2012. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PloS One 7:e32253.

Posner, R., I. A. Toker, O. Antonova, E. Star, S. Anava, E. Azmon, M. Hendricks et al. 2019. Neuronal small rnas control behavior transgenerationally. Cell 177:1814-1826.

Press, C. M., and O. Evensen. 1999. The morphology of the immune system in teleost fishes. Fish & Shellfish Immunology 9:309-318.

Price, T. D., A. Qvarnstrom, and D. E. Irwin. 2003. The role of phenotypic plasticity in driving genetic evolution. Proceedings of the Royal Society B 270:1433-1440.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959.

Puechmaille, S. J. 2016. The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. Molecular Ecology Resources 16:608-627.

Radford, E. J., M. Ito, H. Shi, J. A. Corish, K. Yamazawa, E. Isganaitis, S. Seisenberger et al. 2014. In utero effects. In utero undernourishment perturbs the adult sperm methylome and intergenerational metabolism. Science 345:1255903.

Radosavljević, I., S. Bogdanović, F. Celep, M. Filipović, Z. Satovic, B. Surina, and Z. Liber. 2019. Morphological, genetic and epigenetic aspects of homoploid hybridization between *Salvia officinalis* L. and *Salvia fruticos*a Mill. Scientific Reports 9:3276.

Raj, S., K. Brautigam, E. T. Hamanishi, O. Wilkins, B. R. Thomas, W. Schroeder, S. D. Mansfield et al. 2011. Clone history shapes *Populus* drought responses. Proceedings of the National Academy of Sciences 108:12521-12526.

Rambaut, A., A. J. Drummond, D. Xie, G. Baele, and M. A. Suchard. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Systematic Biology 67:901-904.

Rampon, C., C. H. Jiang, H. Dong, Y. P. Tang, D. J. Lockhart, P. G. Schultz, J. Z. Tsien et al. 2000. Effects of environmental enrichment on gene expression in the brain. Proceedings of the National Academy of Sciences 97:12880-12884.

Regan, K. S., M. G. Jonz, and P. A. Wright. 2011. Neuroepithelial cells and the hypoxia emersion response in the amphibious fish *Kryptolebias marmoratus*. The Journal of Experimental Biology 214:2560-2568.

Reik, W. 2007. Stability and flexibility of epigenetic gene regulation in mammalian development. Nature 447:425-432.

Renaut, S., A. Nolte, and L. Bernatchez. 2009. Gene expression divergence and hybrid misexpression between lake whitefish species pairs (*Coregonus* spp. Salmonidae). Molecular Biology and Evolution 26:925-936.

Rhee, J.-S., B.-S. Choi, J. Kim, B.-M. Kim, Y.-M. Lee, I.-C. Kim, A. Kanamori et al. 2017. Diversity, distribution, and significance of transposable elements in the genome of the only selfing hermaphroditic vertebrate *Kryptolebias marmoratus*. Scientific Reports 7:40121.

Richards, C. L., C. Alonso, C. Becker, O. Bossdorf, E. Bucher, M. Colomé-Tatché, W. Durka et al. 2017. Ecological plant epigenetics: Evidence from model and non-model species, and the way forward. Ecology Letters 20:1576-1590.

Richards, C. L., O. Bossdorf, and M. Pigliucci. 2010. What role does heritable epigenetic variation play in phenotypic evolution? BioScience 60:232-237.

Richards, C. L., A. W. Schrey, and M. Pigliucci. 2012. Invasion of diverse habitats by few Japanese knotweed genotypes is correlated with epigenetic differentiation. Ecology Letters 15:1016-1025.

Richards, E. J. 2006. Inherited epigenetic variation—revisiting soft inheritance. Nature Reviews Genetics 7:395.

Riddle, N. C., and J. A. Birchler. 2003. Effects of reunited diverged regulatory hierarchies in allopolyploids and species hybrids. Trends in Genetics 19:597-600.

Roberge, C., E. Normandeau, S. Einum, H. Guderley, and L. Bernatchez. 2008. Genetic consequences of interbreeding between farmed and wild Atlantic salmon: insights from the transcriptome. Molecular Ecology 17:314-324.

Roberts, L., J. Taylor, and C. G. de Leaniz. 2011. Environmental enrichment reduces maladaptive risk-taking behavior in salmon reared for conservation. Biological Conservation 144:1972-1979.

Roberts, L., J. Taylor, P. Gough, D. Forman, and C. Garcia de Leaniz. 2014. Silver spoons in the rough: can environmental enrichment improve survival of hatchery Atlantic salmon *Salmo salar* in the wild? Journal of Fish Biology 85:1972-1991.

Robinson, M. D., D. J. McCarthy, and G. K. Smyth. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139-140.

Rodgers, A. B., C. P. Morgan, N. A. Leu, and T. L. Bale. 2015. Transgenerational epigenetic programming via sperm microRNA recapitulates effects of paternal stress. Proceedings of the National Academy of Sciences 112:13699-13704.

Rodriguez-Barreto, D., C. G. de Leaniz, E. Verspoor, H. Sobolewska, M. Coulson, and S. Consuegra. 2019. DNA methylation changes in the sperm of captive-reared fish: a route to epigenetic introgression in wild populations. Molecular Biology and Evolution.

Rodríguez-López, C. M., P. Moran, F. Lago, M. Espineira, M. Beckmann, and S. Consuegra. 2012. Detection and quantification of tissue of origin in salmon

and veal products using methylation sensitive AFLPs. Food Chemistry 131:1493-1498.

Rottscheidt, R., and B. Harr. 2007. Extensive additivity of gene expression differentiates subspecies of the house mouse. Genetics 177:1553-1567.

Rousset, F. 2008. GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources 8:103-106.

Rozas, J., A. Ferrer-Mata, J. C. Sánchez-DelBarrio, S. Guirao-Rico, P. Librado, S. E. Ramos-Onsins, and A. Sánchez-Gracia. 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Molecular Biology and Evolution 34:3299-3302.

Sakthivel, K., K. Girishkumar, G. Ramkumar, V. Shenoy, S. Kajjidoni, and P. Salimath. 2010. Alterations in inheritance pattern and level of cytosine DNA methylation, and their relationship with heterosis in rice. Euphytica 175:303-314.

Salmon, A., M. L. Ainouche, and J. F. Wendel. 2005. Genetic and epigenetic consequences of recent hybridization and polyploidy in *Spartina* (Poaceae). Molecular Ecology 14:1163-1175.

Salvanes, A. G., O. Moberg, L. O. Ebbesson, T. O. Nilsen, K. H. Jensen, and V. A. Braithwaite. 2013. Environmental enrichment promotes neural plasticity and cognitive ability in fish. Proceedings of Royal Society B 280:20131331.

Sambandan, D., M. A. Carbone, R. R. Anholt, and T. F. Mackay. 2008. Phenotypic plasticity and genotype by environment interaction for olfactory behavior in *Drosophila melanogaster*. Genetics 179:1079-1088.

Sandilyan, S., and K. Kathiresan. 2012. Mangrove conservation: a global perspective. Biodiversity and Conservation 21:3523-3542.

Sarmento-Soares, L. M., L. F. Ingenito, L. Duboc, R. Martins-Pinheiro, R. Borçato, and J. Silva. 2014. Primeiros registros de *Kryptolebias ocellatu*s (Hensel)(Cyprinodontiformes, Rivulidae) para riachos de Mata Atlântica no Espírito Santo. Boletim Sociedade Brasileira de Ictiologia 111:15-19.

Sati, S., V. S. Tanwar, K. A. Kumar, A. Patowary, V. Jain, S. Ghosh, S. Ahmad et al. 2012. High resolution methylome map of rat indicates role of intragenic DNA methylation in identification of coding region. PLoS One 7:e31621.

Scheiner, S. M., M. Barfield, and R. D. Holt. 2012. The genetics of phenotypic plasticity. XI. Joint evolution of plasticity and dispersal rate. Ecology and Evolution 2:2027-2039.

Schmitz, R. J., Y. He, O. Valdés-López, S. M. Khan, T. Joshi, M. A. Urich, J. R. Nery et al. 2013. Epigenome-wide inheritance of cytosine methylation variants in a recombinant inbred population. Genome Research 23:1663-1674.

Schmitz, R. J., M. D. Schultz, M. G. Lewsey, R. C. O'Malley, M. A. Urich, O. Libiger, N. J. Schork et al. 2011. Transgenerational epigenetic instability is a source of novel methylation variants. Science 334:369-373.

Schrey, A. W., M. Alvarez, C. M. Foust, H. J. Kilvitis, J. D. Lee, A. L. Liebl, L. B. Martin et al. 2013. Ecological epigenetics: beyond MS-AFLP. Integrative and Comparative Biology 53:340-350.

Schrey, A. W., C. A. Coon, M. T. Grispo, M. Awad, T. Imboma, E. D. McCoy, H. R. Mushinsky et al. 2012. Epigenetic variation may compensate for decreased genetic variation with introductions: a case study using house sparrows (*Passer domesticus*) on two continents. Genetics Research International 2012:1-7.

Schwenk, K., N. Brede, and B. Streit. 2008. Introduction. Extent, processes and evolutionary impact of interspecific hybridization in animals. Philosophical Transactions of the Royal Society B: Biological Sciences 363:2805-2811.

Seegers, L. 1984. Zur revision der *Rivulus*-arten sudost-Brasiliens, mit einer Neubeschreibung von Rivulus luelingi n. spp. und. *Rivulus caudomarginatus* n. spp.(Pisces: Cyprinodontidae: Rivulinae). Zoologische Beiträge 28:271-320.

Sentis, A., R. Bertram, N. Dardenne, F. Ramon-Portugal, G. Espinasse, I. Louit, L. Negri et al. 2018. Evolution without standing genetic variation: change in transgenerational plastic response under persistent predation pressure. Heredity 121:266-281.

Shafi, A., C. Mitrea, T. Nguyen, and S. Draghici. 2017. A survey of the approaches for identifying differential methylation using bisulfite sequencing data. Briefings in Bioinformartics:737-753.

Shao, G.-M., X.-Y. Li, Y. Wang, Z.-W. Wang, Z. Li, X.-J. Zhang, L. Zhou et al. 2018. Whole genome incorporation and epigenetic stability in a newly synthetic allopolyploid of gynogenetic gibel carp. Genome biology and evolution 10:2394-2407.

Shea, J. M., R. W. Serra, B. R. Carone, H. P. Shulha, A. Kucukural, M. J. Ziller, M. P. Vallaster et al. 2015. Genetic and epigenetic variation, but not diet, shape the sperm methylome. Developmental Cell 35:750-758.

Sheldon, E., A. Schrey, S. Andrew, A. Ragsdale, and S. Griffith. 2018. Epigenetic and genetic variation among three separate introductions of the house sparrow (*Passer domesticus)* into Australia. Royal Society Open Science 5:172185.

Silva, T. F. d., D. d. A. Azevedo, and F. R. d. Aquino Neto. 2007. Distribution of polycyclic aromatic hydrocarbons in surface sediments and waters from Guanabara Bay, Rio de Janeiro, Brazil. Journal of the Brazilian Chemical Society 18:628-637.

Skotte, L., T. S. Korneliussen, and A. Albrechtsen. 2013. Estimating individual admixture proportions from next generation sequencing data. Genetics 195:693-702.

Smith, T. A., M. D. Martin, M. Nguyen, and T. C. Mendelson. 2016. Epigenetic divergence as a potential first step in darter speciation. Molecular Ecology 25:1883-1894.

Soubry, A., C. Hoyo, R. L. Jirtle, and S. K. Murphy. 2014. A paternal environmental legacy: evidence for epigenetic inheritance through the male germ line. Bioessays 36:359-371.

Spence, R., A. E. Magurran, and C. Smith. 2011. Spatial cognition in zebrafish: the role of strain and rearing environment. Animal Cognition 14:607-612.

Stadler, M. B., R. Murr, L. Burger, R. Ivanek, F. Lienert, A. Schöler, E. van Nimwegen et al. 2011. DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature 480:490.

Steets, J. A., D. E. Wolf, J. R. Auld, and T. L. Ashman. 2007. The role of natural enemies in the expression and evolution of mixed mating in hermaphroditic plants and animals. Evolution 61:2043-2055.

Strand, D. A., A. C. Utne-Palm, P. J. Jakobsen, V. A. Braithwaite, K. H. Jensen, and A. G. V. Salvanes. 2010. Enrichment promotes learning in fish. Marine Ecology Progress Series 412:273-282.

Streit, M., F. Reinhardt, G. Thaller, and J. Bennewitz. 2012. Reaction norms and genotype-by-environment interaction in the German Holstein dairy cattle. Journal of Animal Breeding and Genetics 129:380-389.

Sultan, S. E. 2015, Organism and environment: ecological development, niche construction, and adaptation, Oxford University Press, USA.

Svensson, E. I. 2018. On reciprocal causation in the evolutionary process. Evolutionary Biology 45:1-14.

Swanson-Wagner, R. A., Y. Jia, R. DeCook, L. A. Borsuk, D. Nettleton, and P. S. Schnable. 2006. All possible modes of gene action are observed in a global comparison of gene expression in a maize F1 hybrid and its inbred parents. Proceedings of the National Academy of Sciences 103:6805-6810.

Tatarenkov, A., R. L. Earley, B. M. Perlman, D. S. Taylor, B. J. Turner, and J. C. Avise. 2015. Genetic subdivision and variation in selfing rates among central american populations of the mangrove rivulus, *Kryptolebias marmoratus*. Journal of Heredity 106:276-284.

Tatarenkov, A., R. L. Earley, D. S. Taylor, and J. C. Avise. 2012. Microevolutionary distribution of isogenicity in a self-fertilizing fish (*Kryptolebias marmoratus*) in the Florida Keys. Integrative Comparative Biology 52:743-752.

Tatarenkov, A., R. L. Earley, D. S. Taylor, W. P. Davis, and J. C. Avise. 2018. Natural hybridization between divergent lineages in a selfing hermaphroditic fish. Biology Letters 14:20180118.

Tatarenkov, A., H. Gao, M. Mackiewicz, D. S. Taylor, B. J. Turner, and J. C. Avise. 2007. Strong population structure despite evidence of recent migration in a selfing hermaphroditic vertebrate, the mangrove killifish (*Kryptolebias marmoratus*). Molecular Ecology 16:2701-2711.

Tatarenkov, A., S. M. Q. Lima, and J. C. Avise. 2011. Extreme homogeneity and low genetic diversity in *Kryptolebias ocellatus* from south-eastern Brazil suggest a recent foundation for this androdioecious fish population. Journal of Fish Biology 79:2095-2105.

Tatarenkov, A., S. M. Q. Lima, R. L. Earley, W. M. Berbel, F. B. M. Vermeulen, D. S. Taylor, K. Marson et al. 2017. Deep and concordant subdivisions in the self-fertilizing mangrove killifishes (*Kryptolebias*) revealed by nuclear and mtDNA markers. Biological Journal of the Linnean Society 122:558-578.

Tatarenkov, A., S. M. Q. Lima, D. S. Taylor, and J. C. Avise. 2009. Long-term retention of self-fertilization in a fish clade. Proceedings of the National Academy of Sciences 106:14456-14459.

Tatarenkov, A., B. C. Ring, J. F. Elder, D. L. Bechler, and J. C. Avise. 2010. Genetic composition of laboratory stocks of the self-fertilizing fish *Kryptolebias marmoratus*: a valuable resource for experimental research. PLoS One 5:e12863.

Taylor, D. S. 1990. Adaptive specializations of the cyprinodont fish *Rivulus marmoratus*. Florida Scientist:239-248.

—. 2000. Biology and ecology of *Rivulus marmoratus*: new insights and a review. Florida Scientist:242-255.

Taylor, D. S. 2012. Twenty-Four Years in the mud: What have we learned about the natural history and ecology of the mangrove rivulus, *Kryptolebias marmoratus*? Integrative and Comparative Biology 52:724-736.

Taylor, D. S., M. T. Fisher, and B. J. Turner. 2001. Homozygosity and heterozygosity in three populations of *Rivulus marmoratus*. Environmental Biology of Fishes 61:455-459.

Taylor, S. A., and E. L. Larson. 2019. Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. Nature Ecology and Evolution 3:170-177.

Teh, A. L., H. Pan, L. Chen, M. L. Ong, S. Dogra, J. Wong, J. L. MacIsaac et al. 2014. The effect of genotype and in utero environment on interindividual variation in neonate DNA methylomes. Genome Res 24:1064-1074.

Thomaz, A. T., L. R. Malabarba, S. L. Bonatto, and L. L. Knowles. 2015. Testing the effect of palaeodrainages versus habitat stability on genetic divergence in riverine systems: study of a Neotropical fish of the Brazilian coastal Atlantic Forest. Journal of Biogeography 42:2389-2401.

Thomaz, A. T., L. R. Malabarba, and L. L. Knowles. 2017. Genomic signatures of paleodrainages in a freshwater fish along the southeastern coast of Brazil: genetic structure reflects past riverine properties. Heredity 119:287-294.

Turko, A. J., A. Tatarenkov, S. Currie, R. L. Earley, A. Platek, D. S. Taylor, and P. A. Wright. 2018. Emersion behaviour underlies variation in gill morphology and aquatic respiratory function in the amphibious fish *Kryptolebias marmoratus*. Journal of Experimental Biology 221.

Turko, A. J., and P. A. Wright. 2015. Evolution, ecology and physiology of amphibious killifishes (Cyprinodontiformes). Journal of Fish Biology 87:815-835.

Turner, B. J., M. T. Fisher, D. S. Taylor, W. P. Davis, and B. L. Jarrett. 2006. Evolution of 'maleness' and outcrossing in a population of the self-fertilizing killifish, Kryptolebias marmoratus. Evolutionary Ecology Research 8:1475-1486.

Uller, T. 2008. Developmental plasticity and the evolution of parental effects. Trends in Ecology & Evolution 23:432-438.

Uller, T., S. Nakagawa, and S. English. 2013. Weak evidence for anticipatory parental effects in plants and animals. Journal of Evolutionary Biology 26:2161-2170.

Ungerer, M. C., S. C. Strakosh, and Y. Zhen. 2006. Genome expansion in three hybrid sunflower species is associated with retrotransposon proliferation. Current Biology 16:R872-R873.

Vähä, J. P., and C. R. Primmer. 2006. Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. Molecular Ecology 15:63-72.

Valiela, I., J. L. Bowen, and J. K. York. 2001. Mangrove forests: one of the world's threatened major tropical environments. BioScience 51:807-815.

van der Graaf, A., R. Wardenaar, D. A. Neumann, A. Taudt, R. G. Shaw, R. C. Jansen, R. J. Schmitz et al. 2015. Rate, spectrum, and evolutionary dynamics of spontaneous epimutations. Proceedings of the National Academy of Sciences 112:6676-6681.

van Oosterhout, C., W. F. Hutchinson, D. P. Wills, and P. Shipley. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4:535-538.

van Valen, L. 1973. A new evolutionary law. Evolutionary Theory 1:1-30.

Venney, C. J., M. L. Johansson, and D. D. Heath. 2016. Inbreeding effects on gene-specific DNA methylation among tissues of Chinook salmon. Molecular Ecology 25:4521-4533.

Vergeer, P., N. Wagemaker, and N. J. Ouborg. 2012. Evidence for an epigenetic role in inbreeding depression. Biology Letters 8:798-801.

Verhoeven, K. J., and T. P. van Gurp. 2012. Transgenerational effects of stress exposure on offspring phenotypes in apomictic dandelion. PLoS One 7:e38605.

Verhoeven, K. J. F., J. J. Jansen, P. J. van Dijk, and A. Biere. 2010. Stress-induced DNA methylation changes and their heritability in asexual dandelions. New Phytologist 185:1108-1118.

Verhoeven, K. J. F., and V. Preite. 2014. Epigenetic variation in asexually reproducing organisms. Evolution 68:644-655.

Verhoeven, K. J. F., B. M. Vonholdt, and V. L. Sork. 2016. Epigenetics in ecology and evolution: what we know and what we need to know. Molecular Ecology 25:1631-1638.

Vermeulen, F. B., and T. Hrbek. 2005. *Kryptolebias sepia* n. sp. (Actinopterygii: Cyprinodontiformes: Rivulidae), a new killifish from the Tapanahony River drainage in southeast Surinam. Zootaxa 928:1-20.

Vieira, F. G., F. Lassalle, T. S. Korneliussen, and M. Fumagalli. 2015. Improving the estimation of genetic distances from Next-Generation sequencing data. Biological Journal of the Linnean Society 117:139-149.

von Krogh, K., C. Sorensen, G. E. Nilsson, and O. Overli. 2010. Forebrain cell proliferation, behavior, and physiology of zebrafish, Danio rerio, kept in enriched or barren environments. Physiology & Behavior 101:32-39.

Waddington, C. H. 1942. The epigenotype. Endeavour 1:18-20.

Warde-Farley, D., S. L. Donaldson, O. Comes, K. Zuberi, R. Badrawi, P. Chao, M. Franz et al. 2010. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Research 38:214-220.

Webb, S. A., J. A. Graves, C. Macias-Garcia, A. E. Magurran, D. O. Foighil, and M. G. Ritchie. 2004. Molecular phylogeny of the livebearing Goodeidae (Cyprinodontiformes). Molecular Phylogenetics and Evolution 30:527-544.

Wells, M. W., A. J. Turko, and P. A. Wright. 2015. Fish embryos on land: terrestrial embryo deposition lowers oxygen uptake without altering growth or survival in the amphibious fish *Kryptolebias marmoratus*. Journal of Experimental Biology 218:3249-3256.

Weeks, S. C. 2013. The role of androdioecy and gynodioecy in mediating evolutionary transitions between dioecy and hermaphroditism in the Animalia. Evolution 66:3670-3686.

Wells, M. W., and P. A. Wright. 2017. Do not eat your kids: embryonic kin recognition in an amphibious fish. Behavioral Ecology and Sociobiology 71:140.

West-Eberhard, M. J. 2003, Developmental plasticity and evolution, Oxford University Press.

Whitney, K. D., J. R. Ahern, L. G. Campbell, L. P. Albert, and M. S. King. 2010. Patterns of hybridization in plants. Perspectives in Plant Ecology, Evolution and Systematics 12:175-182.

Wolff, G. L., R. L. Kodell, S. R. Moore, and C. A. Cooney. 1998. Maternal epigenetics and methyl supplements affect agouti gene expression in A vy/a mice. The FASEB Journal 12:949-957.

Wright, P. A. 2012. Environmental physiology of the mangrove rivulus, *Kryptolebias marmoratus,* a cutaneously breathing fish that survives for weeks out of water. Integrative and Compative Biology 52:792-800.

Wyrobek, A. J., B. Eskenazi, S. Young, N. Arnheim, I. Tiemann-Boege, E. W. Jabs, R. L. Glaser et al. 2006. Advancing age has differential effects on DNA

damage, chromatin integrity, gene mutations, and aneuploidies in sperm. Proceedings of the National Academy of Sciences 103:9601-9606.

Xiao, J., C. Song, S. Liu, M. Tao, J. Hu, J. Wang, W. Liu et al. 2013. DNA methylation analysis of allotetraploid hybrids of red crucian carp (*Carassius auratus* red var.) and common carp (*Cyprinus carpio* L.). PLoS One 8:e56409.

Yan, Y.-B., N. C. Duke, and M. Sun. 2016. Comparative analysis of the pattern of population genetic diversity in three Indo-West Pacific *Rhizophora* mangrove species. Frontiers in Plant science 7:1434.

Yeh, P. J., and T. D. Price. 2004. Adaptive phenotypic plasticity and the successful colonization of a novel environment. American Naturalist 164:531-542.

Yu, A., G. Lepere, F. Jay, J. Y. Wang, L. Bapaume, Y. Wang, A. L. Abraham et al. 2013. Dynamics and biological relevance of DNA demethylation in *Arabidopsis* antibacterial defense. Proceedings of the National Academy of Sciences of the United States of America 110:2389-2394.

Yu, X., X. Li, X. Zhao, L. Jiang, G. Miao, J. Pang, X. Qi et al. 2011. Tissue culture-induced genomic alteration in maize (*Zea mays*) inbred lines and F1 hybrids. Annals of Applied Biology 158:237-247.

Zhang, Y. Y., D. Y. Zhang, and S. C. Barrett. 2010. Genetic uniformity characterizes the invasive spread of water hyacinth (*Eichhornia crassipes*), a clonal aquatic plant. Molecular Ecology 19:1774-1786.