Missouri S&T Library and Learning Resources

Scholars' Mine

Masters Theses

Student Theses and Dissertations

Summer 2019

# Characterization of chromosomal translocations in a group of killifish species by using genome-wide high-density SNP mapping approach

Naznin Sultana Remex

Follow this and additional works at: https://scholarsmine.mst.edu/masters_theses

 Part of the Biology Commons, Environmental Sciences Commons, and the Genetics Commons

Department:

## Recommended Citation

Remex, Naznin Sultana, "Characterization of chromosomal translocations in a group of killifish species by using genome-wide high-density SNP mapping approach" (2019). *Masters Theses*. 7907.
https://scholarsmine.mst.edu/masters_theses/7907

This thesis is brought to you by Scholars' Mine, a service of the Missouri S&T Library and Learning Resources. This work is protected by U. S. Copyright Law. Unauthorized use including reproduction for redistribution requires the permission of the copyright holder. For more information, please contact scholarsmine@mst.edu.

# CHARACTERIZATION OF CHROMOSOMAL TRANSLOCATIONS IN A GROUP OF KILLIFISH SPECIES BY USING GENOME-WIDE HIGH-DENSITY SNP MAPPING APPROACH

by

NAZNIN SULTANA REMEX

A THESIS

Presented to the Faculty of the Graduate School of the

MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

In Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE IN APPLIED AND ENVIRONMENTAL BIOLOGY

2019

Approved by:

David D. Duvernell, Advisor
Ronald L. Frank
Gayla R. Olbricht
David J. Westenberg

© 2019

Naznin Sultana Remex

All Rights Reserved

## ABSTRACT

The role of chromosomal rearrangements in reproductive isolation and introgression between species is poorly understood. In heterozygous form, rearrangements may directly interrupt meiotic progression leading to partial sterility/subfertility (underdominance) or may suppress local meiotic segregation (recombination suppression). Such unbalanced meiotic segregation may also result in reproductive isolation and play roles as a driving force of speciation. The objective of this study was to gain insight into the pattern of chromosomal rearrangements in two closely related killifish species in the genus *Fundulus* (*F. notatus,* and *F. olivaceus*) by constructing genetic linkage maps using high-resolution single nucleotide polymorphism (SNP) markers. Markers associated with Robertsonian (Rb) translocations in *F. notatus* were generated by high-throughput genotyping-by sequencing (GBS) method and intra-specific SNPs were aligned to contigs in a reference *F. olivaceus* genome. This SNP-based mapping approach revealed 24 linkage groups (LGs) in *F. olivaceus* and 20 LGs in *F. notatus* including four Rb fusions (corresponding to chromosomes). We also found strong homology at the LG level between our maps and a previously constructed *F. heteroclitus* linkage map. Finally, using these maps and GBS-SNP data, we compared patterns of hybridization and introgression between populations of *F. olivaceus and F. notatus* from two natural hybrid zones. We observed weak prezygotic isolation, but stronger post-zygotic isolation between karyotypically different populations, which indicated multiple chromosomal fusions in *F. notatus* might have influenced reproductive viability of F1 hybrids, promoting reproductive isolation between these two species.

# ACKNOWLEDGMENTS

I would like to start with thanking my thesis advisor, Dr. David D. Duvernell, for being very supportive, helpful, and motivating mentor throughout my graduate years at Missouri S&T. His door was always open to me whenever I encountered any trouble regarding my research project or writing. I am very much grateful to our collaborator, Dr. Jake Schaefer at the University of Southern Mississippi, for collecting and taking care of all the fish samples used in this project, and helping us with some bioinformatics analysis. This thesis project would not be complete without the data (unpublished) provided by Dr. Andrew Whitehead lab at the University of California Davis. I would like to thank Dr. Rob Elshire for the GBS services to generate genetic markers.

I am very thankful to the members of my thesis advisory committee. I appreciate their valuable time and sound advice during our pleasant conversations. A special thanks to Dr. David J. Westenberg, who was not only a very thoughtful and insightful GTA advisor, but also very humble to agree to be on my committee at the last minute. All the members of Duvernell lab were also very helpful and cooperative while working together in the lab and deserve my gratitude. This work has been funded and supported by National Science Foundation (NSF) grant number 1556778.

Lastly, I would like to express my gratitude to my parents and my little brother for their unconditional love and support throughout my entire life. They were always by my side during my hard days. I am very much grateful to my husband, Imtiaz Khalil, for encouraging me to pursue a master's degree in the first place. He always supported me with his love and sacrifice during my time at Missouri S&T.

# TABLE OF CONTENTS

Page

# LIST OF ILLUSTRATIONS

Page

# LIST OF TABLES

Page

# 1. INTRODUCTION

Understanding the mechanisms that generate and maintain biodiversity is a major goal in the field of evolutionary biology. The coexistence of closely related species is common in nature and has been a significant interest to evolutionary biologists to understand the factors that determine species distributions, as well as outcomes of hybridization, introgression, and speciation (Barton and Hewitt, 1989; Hewitt, 1988; Swenson and Howard, 2005; Vamosi and Wilson, 2008). Hybridization may occur as a result of secondary contact between divergent populations when they lack physical barriers, and it provides a test of the mechanisms of reproductive isolation that accompany speciation (Mallet, 2007, 2005; Schumer et al., 2013). So, examining reproductive isolation of closely related species by studying hybrid zone dynamics is one of the ways to understand the process of evolution and speciation. Oftentimes members of genetically diverged groups of populations with incomplete reproductive isolation co-occur and mate in the wild to produce hybrid offspring. These mixed ancestry offspring may be sterile or inviable due to the deterioration of some existing chromosomal structure (e.g. underdominance and recombination suppression) (Abbott et al., 2013; Barton and Hewitt, 1985; Potter et al., 2015). Recent studies have demonstrated that hybridization can play a role directly in the process of speciation (Abbott et al., 2013) or in species diversification by producing novel phenotypes (Nolte et al., 2005). Even though it is very common in the wild, the ultimate consequence of hybridization in evolutionary process is not always clear and contact zones are valuable for studying reproductive isolation and introgression.

Chromosomal rearrangements, which alter the native structure of chromosomes, can drive genetic divergence and reproductive isolation (Noor et al., 2001). Such changes are often prevalent among closely related species and may also play a role in species adaptation and speciation. This may result in intergenomic incompatibilities and may reduce gene flow between populations strengthening reproductive isolation and promoting speciation (Barton and Hewitt, 1985; Schumer et al., 2013; Twyford and Ennos, 2012). When the rates of chromosomal change increase, the speciation rates also become higher (Navarro and Barton, 2003). A variety of types of chromosomal rearrangement (deletion, duplication, insertion, inversion, and translocation) are commonly found both within and among species. While chromosomal changes are often evident within and between species, the actual circumstances that lead to such rearrangements and their consequences are often not clear. Chromosomal reorganizations, including inversions and translocations, may impact genetic divergence in two different ways: by directly interrupting meiotic progression in heterozygotes, producing partially sterile or unfit hybrids (underdominance) (Potter et al., 2015; Rieseberg, 2001), or by reducing or suppressing local recombination or gene flow, possibly leading to inter-genomic incompatibilities (recombination suppression) (Navarro and Barton, 2003; Noor et al., 2001). Underdominance is more likely associated with Robertsonian translocation and recombination suppression is associated with chromosomal inversions. This reduced gene flow may act as a reproductive barrier, and with time, may lead to strong reproductive isolation and speciation. Studying chromosomal rearrangements is a key to see how genomic architecture may be shifted with divergence.

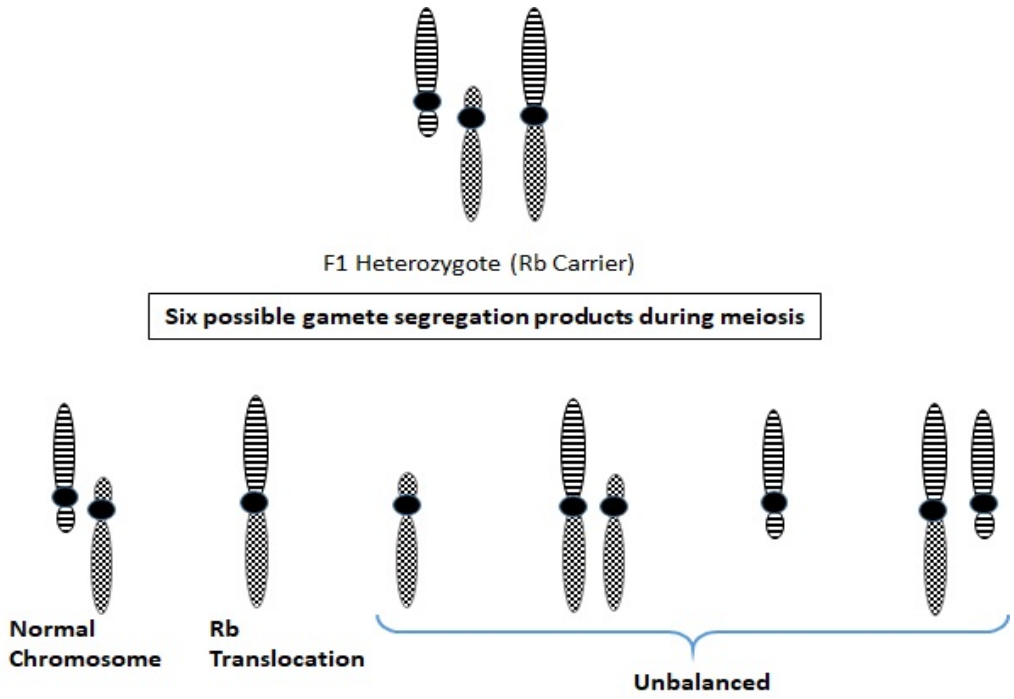


Figure 1.1. Pattern of gamete segregation of a Robertsonian translocation carrier during meiosis reduction division.

Robertsonian (Rb) translocation is a special category of translocation where two acrocentric or telocentric chromosomes fuse together to form one large metacentric chromosome with a single centromere. During this type of chromosomal rearrangement, breakage takes place across the centromere of small chromosomes and the long arms fuse together to form one chromosome with two long arms on either side of the centromere. The short arms of acrocentric chromosomes, with nonessential genes, may also fuse, but usually become lost within a few cell cycles. This type of rearrangement is very common, and has been widely documented in numerous mammal and fish species (Adega et al., 2009; Garagna et al., 2014; Piálek et al., 2005; Wójcik and Searle, 1988). An Rb fusion can segregate in a population for many generations and remain undetected.

Heterozygotes, who are carriers of an Rb translocation, have a balanced chromosomal complement with two copies of each gene. Problems may arise when these heterozygotes undergo meiotic reduction division, which can happen according to six different segregation patterns (Figure 1.1). All segregation patterns are equally possible, one sixth of the gametes will have balanced standard chromosomes, and one sixth will have balanced Rb chromosomes that will be transmitted to the offspring. The rest of the four possible outcomes will have either excess copies of genes (trisomy) or deficits of genes (monosomy) (Figure 1.1). Therefore, in contact zones, the rates of gene flow between the divergent populations with chromosomal translocations may not be uniform across loci (Turner et al., 2005; Wu, 2001). It is the next generation of F2 or backcross offspring where unbalanced segregation of alleles could reveal reduced fertility of Rb heterozygotes. Many human genetic diseases are associated with this condition, for example, familial Down syndrome, mental retardation, leukemia, and fertility problems have been associated with Rb fusions (Chapman and Hesketh, 2000; Fernhall et al., 1996; Niebuhr, 1974; Thirman et al., 1993). Thus, despite having a full genetic complement, the F1 heterozygotes could show reduced reproductive fitness (partial sterility) because of their chance of producing genetically unbalanced gametes. Mendel's second law of independent assortment dictates that segregation errors should occur independently at non-homologous Rb chromosomes. Therefore, unbalanced segregation probabilities combine multiplicatively due to several independent Rb translocations. So, Individuals who are heterozygous for multiple non-homologous Rb fusions could effectively be sterile with noticeably low reproductive fitness (Gropp, 1981). Thus, chromosomal

translocations may promote partial or complete sterility of hybrid offspring, leading to rapid and active reproductive isolation (Britton et al., 2000).

The *Fundulus notatus* species complex is an excellent model system to study the relationship between genetic divergence and chromosomal rearrangements. This complex contains three closely related species under the genus *Fundulus* (*F. olivaceus*, *F. notatus*, and *F. euryzonus*) which exhibit broadly overlapping geographic distributions throughout North America (Duvernell et al., 2013). Two members of this complex, the blackspotted topminnow (*F. olivaceus* Storer 1845) and the blackstipe topminnow (*F. notatus* Rafinesque 1820), are distributed throughout much of the Mississippi River drainages and the coastal drainages of the Gulf of Mexico (Howell and Black, 1981). The distribution of the third species (*F. euryzonus*) is restricted to two coastal drainages in Mississippi and Louisiana. Sympatric distributions are found where the ranges overlap (Figure 1.2). The rates of hybridization and introgression between *F. olivaceus* and *F. notatus* vary to a great extent among hybrid zones (Duvernell et al., 2013; Duvernell and Schaefer, 2014; Schaefer et al., 2016, 2011). Even though reproductive isolation among the species of this complex is quite strong, the barriers are incomplete, with hybridization observed in nature and in the lab (Duvernell et al., 2007; Vigueira et al., 2008). The co-occurrence of the topminnow species is very common in a broad range of river drainages (Figure 1.2) leading to secondary contact. The distribution of species usually follows an upstream-downstream pattern, and hybridization occurs where transitions of tributaries and large rivers take place. The offspring of mixed ancestry in many contact zones include individuals and backcrosses from multiple generations (Schaefer et al., 2016).

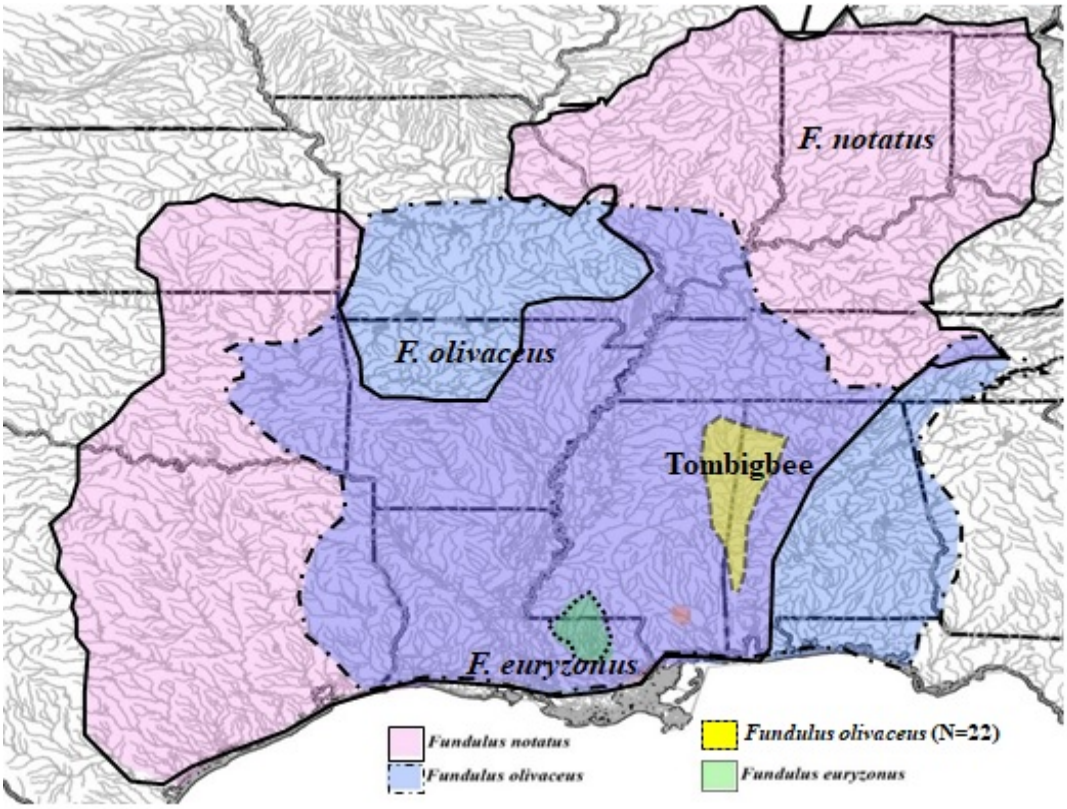


Figure 1.2. Broad overlapping distribution of the species of *Fundulus notatus* complex. Pink area represents the distribution of *F. notatus*; Blue area is the distribution of *F. olivaceus*; Green area is the habitat of *F. euryzonus*; Yellow is the distribution area of Tombigbee clade of *F. notatus* with more chromosomes (N= 22). Purple areas represent overlapping distributions of the species. (Image adapted from Duvernell and Schaefer 2013).

Our focal group of species for this study, *Fundulus notatus* complex, possess both intraspecific and interspecific chromosomal variations which makes this an excellent model system to study the role of Robertsonian fusion in reproductive isolation. Extensive chromosome studies have demonstrated that *F. olivaceus* and *F. euryzonus* share the same ancestral or standard karyotype of 24 chromosomes (N= 24), whereas, *F. notatus* populations have 20 chromosomes in haploid condition (N= 20) throughout most

of the species range (Figure 1.3) (Chen, 1971; Howell and Black, 1981). A population with a distinctive haploid karyotype of N=22 (2 large metacentric chromosomes) has been documented in Tombigbee clade of *F. notatus* (Figure 1.2) (Black and Howell, 1978). In hybrid zones, where individuals exhibiting both the standard and the rearranged chromosomes are found, there is an opportunity for generation of F1 hybrids. A diagnostic feature of these F1 hybrids is the generation of trivalent chromosome alignments at meiosis (Howell and Black, 1981) corresponding to the translocated chromosomes. Since chromosomal changes can directly alter or interrupt meiotic segregation, F1 individuals may become subfertile or sterile. So, the chances of F1 parents reproducing may be low. This raises the possibility that chromosomal rearrangements may play an important role in lineage divergence, promoting reproductive isolation, and limiting opportunities for genetic introgression.

To understand the role of chromosomal rearrangements in the process of reproductive isolation and introgression, characterization of such chromosomal changes is required. Construction of genetic/recombination maps, and use those maps to study chromosomal rearrangements, were main goals of this project. Linkage maps are useful tools to understand species-specific genomic architecture and how it differs between species (Berdan et al., 2014). A high-density linkage map study will facilitate fine-scale comparisons of single nucleotide polymorphisms (SNPs) across the loci between closely related species. SNPs provide appropriate genetic markers for high-resolution genome-wide association studies (GWAS) or fine gene-mapping studies because of their quantity throughout the genome and their stable inheritance over generations (Thomas et al., 2011). For this study, we used a high throughput SNP mapping approach in order to

identify genetic markers associated with Rb translocations. Using a large number of markers, we constructed linkage groups for both standard and translocated karyotypes of our study organisms to elucidate the role of such chromosomal rearrangements and genetic divergence on reproductive isolation and genomic patterns of introgression.

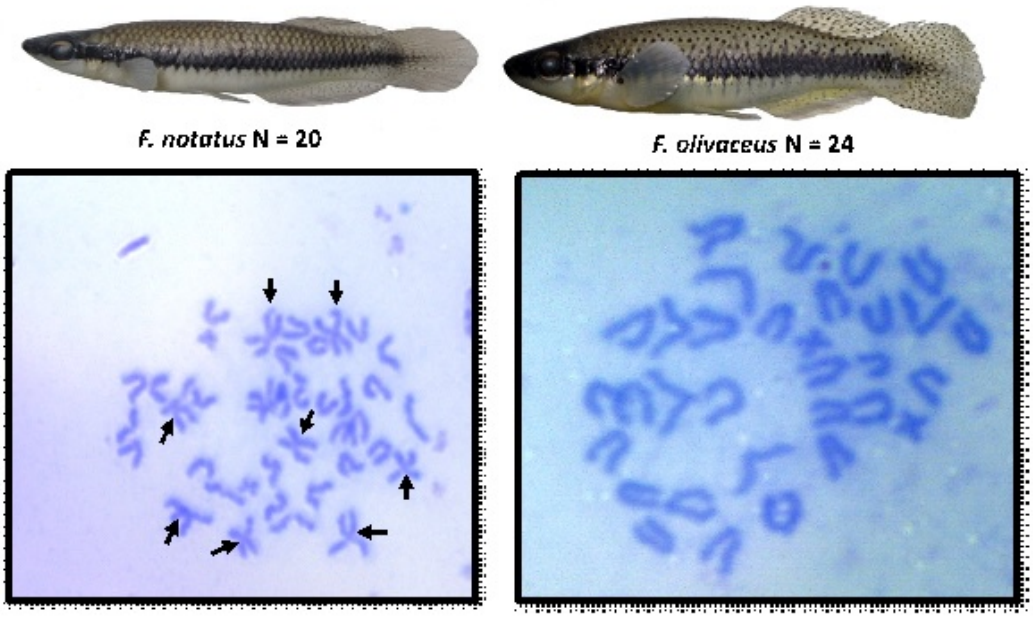


Figure 1.3. Karyotypes of F. notatus (left) and F. olivaceus (right). Arrows on the left picture indicate large metacentric chromosomes due to Robertsonian translocation. (Unpublished image prepared by Tyler McGowan, a former undergraduate student in the Duvernell lab).

# 2. LITERATURE REVIEW

## 2.1. EXAMPLES OF CHROMOSOMAL REARRANGEMENTS

Different types of chromosomal rearrangements (inversions, translocations, and fusions) and their consequences have long been studied in a wide variety of species. Changes in the number or structure of chromosomes may lead to the formation of new species which are reproductively isolated. There are a number of examples of chromosomal mutations in nature that distinguish sister species. For example, chromosome inversions have extensively been studied in insects. Two desert fruit fly species, *Drosophila mojavensis* and *D. arizonae,* exhibit fixed inversions with associated increased divergence around the sites of inversion (Lohse et al., 2015) suggesting the role of inversions in suppressing genetic exchange. Moore and Taylor (1986) showed that two sympatric species of *Drosophila*, *D. pseudoobscura* and *D. persimilis*, are reproductively isolated due to large paracentric inversions between chromosome X and chromosome 2 resulting in sexual isolation, sterility of hybrid males, and inviability of hybrid backcrosses. Two sister species of malaria-transmitting mosquitoes, *Anopheles arabiensis* and *A. gambiae*, are also distinguished by chromosomal inversion (Coluzzi and Bradley, 1999; Wang et al., 2011). Chromosomal rearrangements have also been widely studied in vertebrate species. For example, chromosome fusions among salmonids (Kodama et al., 2014), reciprocal translocations and inversions in zebrafish (Talbot et al., 1998), and two fusions and one pericentric inversion between guinea fowl and chicken chromosomes (Shibusawa et al., 2002).

Robertsonian (Rb) translocation is also a frequently occurring chromosomal mutation that has been widely documented in mammals (Qumsiyeh et al., 1997). Masuda et al. (1980) revealed two Rb fusions in Japanese Black cattle (*Bos primigenius*) between chromosome 1 and 29 and chromosome 5 and 21. There are several other species and subspecies of cattle, such as Brown Swiss cattle, Swiss Simmental cattle, and British Friesian cattle, where researchers found evidences of chromosome fusion (Blazak and Eldridge, 1977; Gustavsson, 1979; Logue and Harvey, 1978). Five Rb translocations have been identified in the house shrew (*Suncus murinus* by Rogatcheva et al. (2000). In a study by Yang et al. (1995), the chromosome number in muntjac deer was shown to vary from 2N = 6 to 2N = 46. Other examples include- dramatic variations in the diploid number of chromosomes in Western European house mouse (*Mus musculus domesticus*), which may vary between 2N = 22 and 2N = 40 (Garagna et al., 2014a). Scientists reported over 100 geographically distinct chromosomal races under this subspecies of mouse (Hauffe et al., 2012). These remarkable variations in the number of chromosomes might be due to one or several chromosomal fusions. In great apes, multiple chromosomal rearrangements have separated human chromosomes (*Homo sapiens*: 2N= 46) from the chimpanzee (*Pan troglodytes*) and bonobo (*P. paniscus*) (2N= 48) (Nickerson and Nelson, 1998). One Robertsonian fusion and nine pericentric inversions between chimpanzee and human chromosomes were documented in a previous study by Szamalek et al. (2006), which may have limited gene flow early in the ancestry of these species.

## 2.2. REDUCED HYBRID FITNESS AND REPRODUCTIVE ISOLATION DUE TO ROBERTSONIAN TRANSLOCATIONS

Robertsonian rearrangement has long been considered as a driver of reproductive isolation by directly interrupting meiotic progression and becoming fixed in the population through meiotic drive or genetic drift (Gropp, 1982, 1981). Hybrids between populations with multiple Rb translocations typically show severe reduction in reproductive fitness because of mis-segregation of complex multivalent chains during meiosis (Baker and Bickham, 1986). This missegregation may lead to deletion or duplication of chromosomal segments in some gametes and the Rb heterozygotes exhibit partial sterility (Ayala and Coluzzi, 2005). This could diminish the likelihood of intercrossing leading to reinforcement, and complete reproductive isolation (White, 1978). Several previous studies have supported this theory, for example, Gustavsson (1979) reported reduced fertility of both male and female cattle who were heterozygous for a 1/29 Rb translocation. Schmutz et al. (1991) performed a study using cattle embryos where they found an impaired fertility rate of Rb carriers. So, underdominance at the Rb loci in heterozygotes could lead to selection against F1 hybrids, which would result in reinforcement of reproductive isolation between closely related species, distinguished by their chromosomal differences (White, 1974). Rock-wallabies are excellent examples of chromosomal rearrangements with a majority of them being Rb fusions, and few inversions and transpositions across centromeres (Potter et al., 2017). Potter et al. (2015) measured the gene flow between different chromosomal races with fusions of rock-wallabies and found relatively large amount of hybrid admixture, which was contradictory to their expectation of reduced gene flow due to underdominance and recombination suppression. This indicated that the actual consequences of Rb fusion on

gene flow is debated, and there might be other factors that influence the role of chromosomal rearrangements in driving reproductive isolation.

Within humans, Rb translocations can cause serious birth related problems. Wang et al. (1991) reported a case of a nine-year old girl suffering from abnormal congenital development or mental retardation due to chromosomal imbalance i.e. trisomy of chromosome 14 resulting from a 13/14 Rb fusion. Chromosomal anomalies may also play a role in the initiation as well as progression of tumorigenesis, and recently it has been a critical issue in cancer biology. Due to translocations, new combinations of DNA sequences can be created which can induce tumorigenesis sometimes by activating proto-oncogenes (cancer causing genes) or eliminating tumor-suppressor genes (Haigis and Dove, 2003). Other human genetic disorders due to Rb translocated chromosomes include familial Down syndrome, caused by the trisomy of chromosome 21 (the third copy of chromosome 21 is attached with chromosome 14 forming a Robertsonian fusion) (Niebuhr, 1974; Robinson et al., 1994). The children who suffer from translocation Down syndrome may experience intellectual disability, heart problems, delayed cognition, and behavioral abnormalities (Chapman and Hesketh, 2000; Fernhall et al., 1996). Spira et al. (1979) found evidence that T cell leukemia is another human genetic disorder which may be provoked by the trisomy of chromosome 15 which can be formed by the fusion with chromosomes 1, 5, or 6. Thus, Rb fusions create both inter- and intra-specific chromosomal polymorphisms and oftentimes promote fertility problems as well as birth/developmental defects in different species groups (Garagna et al., 2001).

## 2.3. GENETIC LINKAGE MAPPING TO REVEAL CHROMOSOMAL REARRANGEMENTS

Linkage maps are excellent tools for studying genome architecture, gene function, and chromosomal rearrangements between closely as well as distantly related taxa (Berdan et al., 2014). These maps can be used to associate genetic markers with chromosomes that are involved in Rb translocations and to compare linkage maps between species with different karyotypes. Genetic mapping is primarily based on accurate estimation of the rate of pair-wise recombination frequencies which have long been studied in the field of population genetics. According to Mendel's second law of independent assortment, during gamete segregation, alleles on one gene sort independently of alleles on another gene (unlinked markers). However, some genetic markers, which are in close proximity on the same chromosome, may become genetically linked and inherited together during gamete segregation. Linkage mapping takes advantage of this suppression of recombination resulting from physical linkage of genetic markers. If the frequency of recombination between two genetic markers is low, this means that these two markers are more likely to be linked and higher recombination frequency indicates markers on different chromosomes that are most likely to be unlinked (Ahn and Tanksley, 1993).

Recombination map construction and analysis of linkage groups can be performed in genetically divergent populations (e.g. F2, backcross, recombinant inbred lines, and double haploid populations) that are obtained from two parental lines (Meng et al., 2015). A wide variety of genetic markers can be used to create such linkage maps including traditional markers like microsatellites, or comparatively newer single nucleotide polymorphism (SNP) markers (Akkaya et al., 1995). New sequencing technologies have

been developed that facilitate the discovery of genetic markers for many species which has made it a lot easier to understand the genome architecture of species with large and complex genomes (Atwell et al., 2010; Cockram et al., 2010). The restriction-site associated DNA sequencing (RAD-seq) approach in combination with multiplex sequencing and next-generation sequencing (NGS) of the RAD tags was the first step towards reducing genome complexity and genetic mapping of mutations (Baird et al., 2008; Miller et al., 2007). Later, a more simple, robust, and affordable technique called genotyping-by sequencing (GBS) was developed to minimize the complexity of large genomes and discover SNPs (Elshire et al., 2011). The library construction through GBS is a more simplified approach than that of RAD-seq. This technique requires less quantity of DNA, a single restriction enzyme, eliminates random shearing, and requires fewer steps after PCR amplification of the library pool. We used this genome sequencing technology to generate thousands of SNP markers for linkage mapping.

Researchers use linkage mapping approaches to characterize chromosomal rearrangements in both plant and animal genomes, and track their patterns of inheritance in species. Doganlar et al. (2002) constructed 12 linkage groups for eggplant (*Solanum melongena*) in a mapping study and observed collinearity with one of the other solanaceae species, the tomato genome. They found evidence of 23 inversions and 5 translocations, which possibly have separated them from their last common ancestor. There are other plant species, including barley (Ramsay et al., 2000), bean (Pedrosa et al., 2003), and sunflower (Burke et al., 2004), whose genomes have been successfully sequenced and mapped. Geneticists have confirmed the presence of large numbers of inversions and translocations in these species.

Comparative genome studies using chromosome mapping are also common in animals to study evolution. A previous mapping study used microsatellite markers to reveal 20 Linkage Groups (LGs) corresponding to 20 autosomes in rhesus macaques, and helped geneticists to compare with that of humans (Rogers et al., 2006). Over the past two decades, linkage mapping studies have revealed chromosomal reorganizations among sheep (Maddox, 2001), cat (Pontius et al., 2007), deer (Huang et al., 2006), and mouse (Garagna et al., 2014b). In more recent years, scientists have been able to sequence the whole genome of human, and assemble informative genetic markers to linkage groups in order to learn more about the role of chromosomal mutations on different genetic diseases (Kong et al., 2002).

Genetic mapping has identified numerous examples of chromosomal rearrangements in fish genomes. Brenna-Hansen et al. (2012) mapped the chromosomes of Atlantic salmon from both European and North American origin. Their comprehensive comparisons between these two genomes uncovered three individual chromosomal fusions that separated karyotypes of these two species. A later study by Leitwein et al. (2017) constructed 40 LGs for brown trout (*Salmo trutta*) and compared with 29 LGs of Atlantic salmon (*S. salar*) to reveal multiple chromosomal fusion and fission events. Linkage mapping approaches thus have extensively been used by researchers to understand chromosome evolution and compare synteny between related species groups.

## 2.4. CHROMOSOME EVOLUTION IN KILLIFISH

Variation in the number of chromosomes within and among the group of killifish makes this group interesting to evolutionary geneticists for better understanding of the

process of speciation. An early chromosome cytogenetics study by Chen (1971) karyotyped 20 killifish (*Fundulus*) species and found the diploid number of chromosomes (2N) varied between 32 and 48. Most of the *Fundulus* species had a karyotype of 48 chromosomes (mostly acrocentric) with fewer species exhibiting lesser numbers of chromosomes (Black and Howell, 1978; Chen, 1971). The killifish with reduced chromosome number contained up to 16 large metacentrics which the author predicted were result of Rb fusions. This assumption was later supported by killifish mapping studies. Berdan et al. (2014) carried out a high-density genetic mapping study where they mapped the chromosomes of two closely related but karyotypically different killifish species- the Rainwater killifish (*Lucania parva*: N= 23) and the Bluefin killifish (*L. goodei*: N= 24). They used SNP markers and compared the synteny between these two species and with some other teleost fishes. They were able to uncover 23 and 24 LGs for these two sister species, respectively, which corresponded to their chromosome number from Chen's study. They also confirmed the presence of one Rb translocations in Rainwater killifish which resulted from the fusion of two acrocentric chromosomes in Bluefin killifish.

Another extensively studied killifish, *F. heteroclitus* (Atlantic killifish: N= 24), has been successfully sequenced and mapped into 24 linkage groups which is consistent with 24 previously documented chromosomes for this species (Waits et al., 2016). They used microsatellite markers combined with SNPs and observed a high degree of synteny between the genomes of Atlantic killifish, medaka, and zebrafish. A more recent unpublished study has constructed more refined and improved recombination maps for *F. heteroclitus* using high-quality and large number of RAD-seq markers (~5,600)

(Whitehead et al., unpublished data). They have been able to order about 84% of the genome scaffold assembly to 24 chromosomes. To get insight into finer scale resolution of genome structure, and understand more about the genome variations in killifish, we have constructed high-density maps for two *Fundulus* species (*F. olivaceus* and *F. notatus*) in this present study.

## 3. AIMS AND OBJECTIVES

Reproductive isolation is a driving factor in speciation (Coyne and Orr, 2004). Many closely related species differ by chromosomal rearrangements, and those chromosomal rearrangements can be responsible for reproductive isolation through underdominance (Noor et al., 2001). However, the role that chromosomal mutations may play in the speciation process itself is not well known (Rieseberg, 2001). Contact zones, where closely related species encounter one another, are places where the role of chromosomal rearrangements in driving reproductive isolation can be evaluated. The *Fundulus notatus* species complex possess both inter- and intraspecific chromosomal variations (Black and Howell, 1978; Chen, 1971). Studies of hybrid zones between karyotypically different *F. olivaceus* and *F. notatus* have demonstrated that hybridization and introgression occur, but vary to a great extent among geographic regions (Duvernell and Schaefer, 2014; Schaefer et al., 2011). An open question is the role of chromosomal rearrangements in contributing to reproductive isolation in topminnows. However, previous studies of chromosomal variation in these species have been conducted using cytogenetic techniques, but no one has previously mapped genetic markers onto the chromosomes to allow study of the fate of chromosome mutations in hybrid zones. In this present study, we used an advanced GBS technique to generate specific high-density SNP markers to localize and characterize such chromosomal fusions in *F. notatus* and *F. olivaceus*. In later part, we used these maps and GBS-SNP data to know more about the consequences of such chromosomal rearrangements on reproductive isolation by analyzing population genetics of two naturally replicated hybrid zones.

The main aims, objectives and specific hypotheses addressed in this study are stated below-

1. To prepare DNA samples from multiple families of F2-cross progeny for the construction of GBS libraries in order to generate high-density SNP markers suitable for recombination mapping of *F. notatus* and *F. olivaceus* chromosomes.
2. To assign SNP markers to linkage groups (i.e. chromosomes) and construct genetic recombination maps for each of the chromosomes in *F. notatus* and *F. olivaceus*.

   Hypothesis: *F. olivaceus* SNP markers will assign to 24 linkage groups, each with similar-length recombination maps. *F. notatus* SNP markers will assign to 20 linkage groups, comprised of sixteen short-length maps and four double-length maps.
3. To align *F. notatus* and *F. olivaceus* linkage groups to a reference genome to establish homology and infer the linkage groups in *F. olivaceus* that have been fused in *F. notatus*.

   Hypothesis: eight of the linkage groups identified in *F. olivaceus* will match the four largest linkage groups in *F. notatus*.
4. To generate GBS SNP data for population samples of *F. notatus* and *F. olivaceus* from two independent hybrid zones and conduct population genetic analyses to assign individuals to hybrid classes (i.e. parental, F1, F2, and backcross). We expected that hybridization rates would be similar between the two contact zones. However, due to karyotype differences in the *F. notatus*

populations. We predicted that backcross hybridization would be more limited in Spring River (N= 20, 24) than in the Tombigbee River (N= 22, 24).

Hypothesis 1: Individuals of hybrid origin will predominantly assign to F1 hybrid class.

Hypothesis 2: A higher proportion of individuals will assign to backcross classes in Tombigbee than in the Spring River hybrid zones.

# 4. MATERIALS AND METHODS

In order to accomplish our goals, this study has been performed in two stages. The first step was to use high-density genome-wide single nucleotide polymorphisms (SNPs) to construct genetic linkage maps for *F. olivaceus* and *F. notatus*. SNPs were generated using the genotyping-by-sequencing method (Elshire et al., 2011) and aligned against a draft *F. olivaceus* reference genome (Whitehead lab, unpublished data). The constructed recombination maps were then assembled and compared with the mapped reference genome of *F. heteroclitus* to align against individual chromosomes and identify the chromosomal translocations. The second step was to analyze the pattern of hybridization and introgression in two natural hybrid zones of two sister species with different karyotypes. We used GBS-generated SNPs as genetic markers and studied the genetic structure of these two drainages at both the SNP marker and chromosome level. The species-diagnostic SNPs in natural hybrid zones were then aligned to mapped contigs to study patterns of introgression at standard and fused chromosomes.

## 4.1. F2 GENETIC MAPPING

The construction of genetic linkage maps of *F. olicaveus* and *F. notatus* includes following steps-

**4.1.1. Creating F2 Mapping Populations.** Genetic crosses were constructed by Jake Schaefer (University of Southern Mississippi) and tissues were provided for this thesis project. We used an F2 cross design for this mapping study. For each species, crosses were created by using parents from geographically isolated and divergent

populations. The F2 populations were produced from multiple families for each species. Initial crosses were constructed from a minimum of three breeding pairs for each species (Figure 4.2). For *F. olivaceus*, the parents of F1s were drawn from the Bouie River (Pascagoula) (GPS coordinates: 31.425806, -89.414626) and the South Fork White Oak Creek, Arkansas River, Arkansas (GPS coordinates: 35.527143, -93.863363) (Figure 4.1). Similarly, parents of F1s for *F. notatus* were collected from Russet Creek, Texas (Ouachita River) (GPS coordinates: 33.428832, -94.548460) and Patterson Slough (Sabine River) (GPS coordinates: 30.307873, -93.720734) (Figure 4.1). The progeny of the grandparents (F1 progeny) were then raised to adulthood, and F1 progeny from different families (unrelated) were assigned as parents for the generation of mapping F2 progeny (Figure 4.2). Our crossing design and family sizes followed that of Berdan et al. (2014). The F2 progeny were genotyped for construction of genetic maps.

**4.1.2. DNA Preparation.** Fin clips from the grandparents and parents, and F2 larvae were preserved in a solution with high-salt concentration (Seutin et al., 1991) prior to DNA extraction. DNA extraction was performed using the Qiagen DNeasy blood and tissue extraction kit (Qiagen Inc.). Following extraction, the DNA samples were treated with *DN*ase-free *RN*ase A. The concentration of each DNA sample was quantified using a Qubit 3.0 fluorometer (broad range double-stranded DNA protocol). The samples were diluted or concentrated to a final concentration between 30 and 100 ng/µL. The quality of each DNA sample was confirmed by performing electrophoresis in a 1% agarose gel to test for DNA degradation. The digestibility of DNA samples was confirmed using a six-base-cutter restriction enzyme, *Eco*RI, in trial digestions conducted on a subset of samples (~ 10% of total samples).

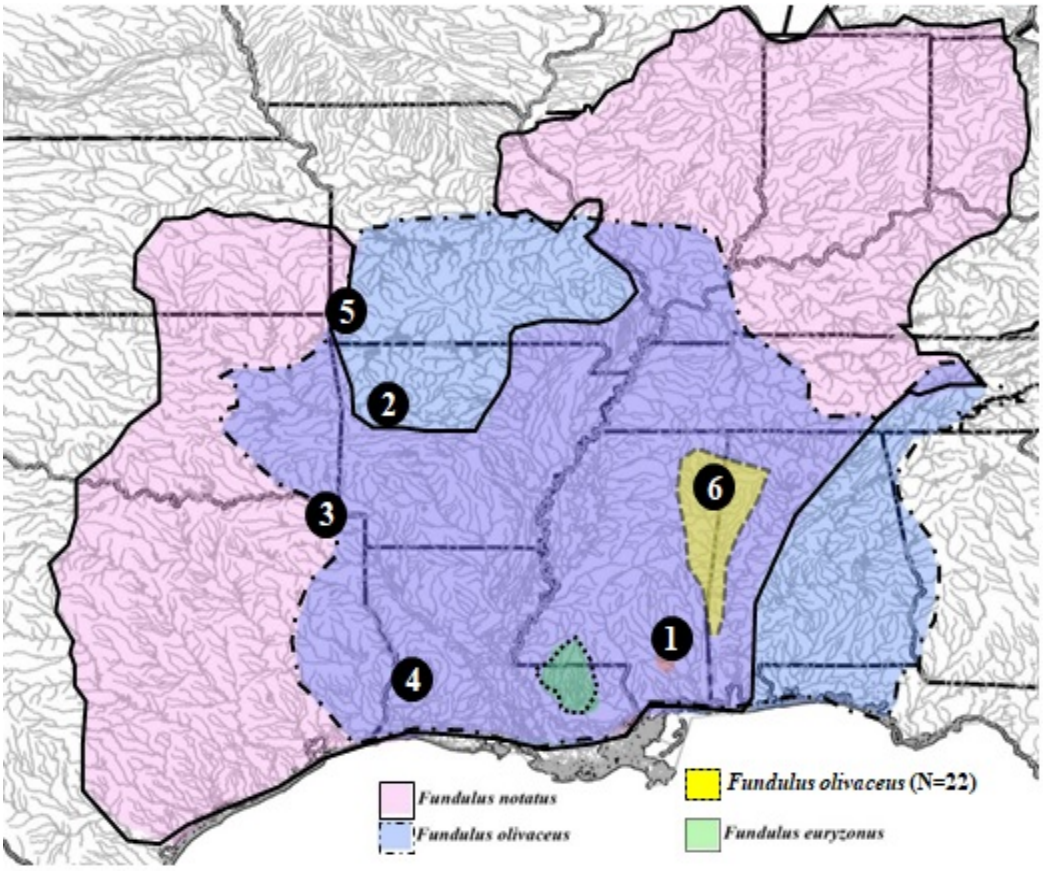


Figure 4.1. Sample collection sites of the grandparents for *F. olivaceus* and *F. notatus* mapping populations and locations of two natural hybrid zones. Geographically isolated drainages for *F. olivaceus* grandparents- 1. Bouie River (Pascagoula) and 2. South Fork White Oak Creek, Arkansas River. Geographically divergent populations for *F. notatus* grandparents- 3. Russet Creek, Ouachita River and 4. Patterson Slough, Sabine River. Two natural hybrid zones- 5. Spring River (contact zone of *F. olivaceus*: N= 24 and *F. notatus*: N= 20) and 6. Tombigbee River (contact zone of *F. olivaceus*: N= 24 and *F. notatus*: N= 22).

**4.1.3. GBS Library Construction and Sequencing.** Samples were genotyped by Elshire GBS Service at Palmerston North, New Zealand (The Elshire Group Ltd. https://www.elshiregroup.co.nz/). The Genotype-by-sequencing (GBS) method was used to discover large numbers of SNPs following the process described by Elshire et al.

(2011). Samples were digested with methylation sensitive restriction enzyme *Eco*T22I selected to eliminate repetitive fractions and reduce genome complexity. Pairs of enzyme-specific adapters along with unique barcodes were ligated with each resultant DNA digestion and individuals were pooled together for PCR amplification. After purifying PCR products, the multiplexed GBS libraries were sequenced using an Illumina HiSeq sequencer (Illumina Inc.) and millions of reads were produced in FASTQ (*fastq.gz). Protocols for *Eco*T221 digestion, adapter ligation, PCR amplification, and GBS library construction follow Elshire et al. (2011).

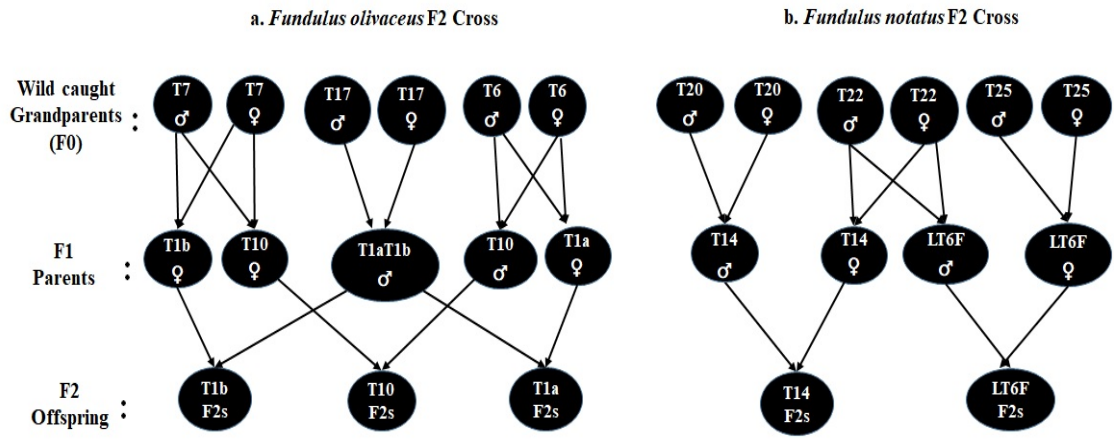


Figure 4.2. F2 mapping cross design of a) *Fundulus olivaceus* and b) *Fundulus notatus*. F2 offspring from families T1a and T1b in *F. olivaceus* shared one male parent.

**4.1.4. GBS Data Processing and SNPs Calling.** The raw sequence data generated through GBS were compressed FASTQ files containing multiplexed and barcoded sequence reads from the GBS library (Figure 4.3). They were first de-multiplexed, cleaned, and barcodes were removed by using STACKS version 1.48

(process_radtags module) (Catchen et al., 2013). The cleaned and trimmed master tags for each individual sample were aligned to the draft reference genome of *F. olivaceus* (https://osf.io/d54mx/) using the alignment tool, Bowtie2 version 2.3 (Langmead and Salzberg, 2012). Bowtie2 aligned the processed sequence data genotype-by-genotype against the reference genome using all default parameters and "end to end" option for "very sensitive" data. The Sequence Alignment/Map (SAM) files generated by Bowtie2 were converted to Binary Alignment/Map (BAM) format using SAMtools 1.9 (Li et al., 2009). BAM files are just the binary representation of SAM files and contain the same information but compressed to minimize space. The binary alignment (BAM) files were then used for SNP calling.

We used the "pstacks" program implemented in the STACKS pipeline v.1.48 (Catchen et al., 2013) for discovering polymorphic SNP loci for parents and offspring of both *F. olivaceus* and *F. notatus*. "Pstacks" created stacks of exactly matched short read sequences by using alignment files. We used all the default parameters for "pstacks" and specified a minimum depth of coverage value of 3 reads (-m) to report a stack. The model parameter alpha ($\alpha$) was set as 0.05 and a minimum mapping quality value was 10 for running "pstacks". For each species, we built a catalog of SNP loci using the F1 parents. These catalogs, created in "cstacks", contained all SNP loci to be mapped and alternative alleles present in the F1 parents. The next program in the pipeline, "sstacks", matched the stacks of putative loci in progeny to the catalog of parents and identified alleles in each offspring. The SNP allele data were converted to JoinMap format using the "genotype" module in STACKS. This module determined informative mappable SNP markers for constructing linkage groups.

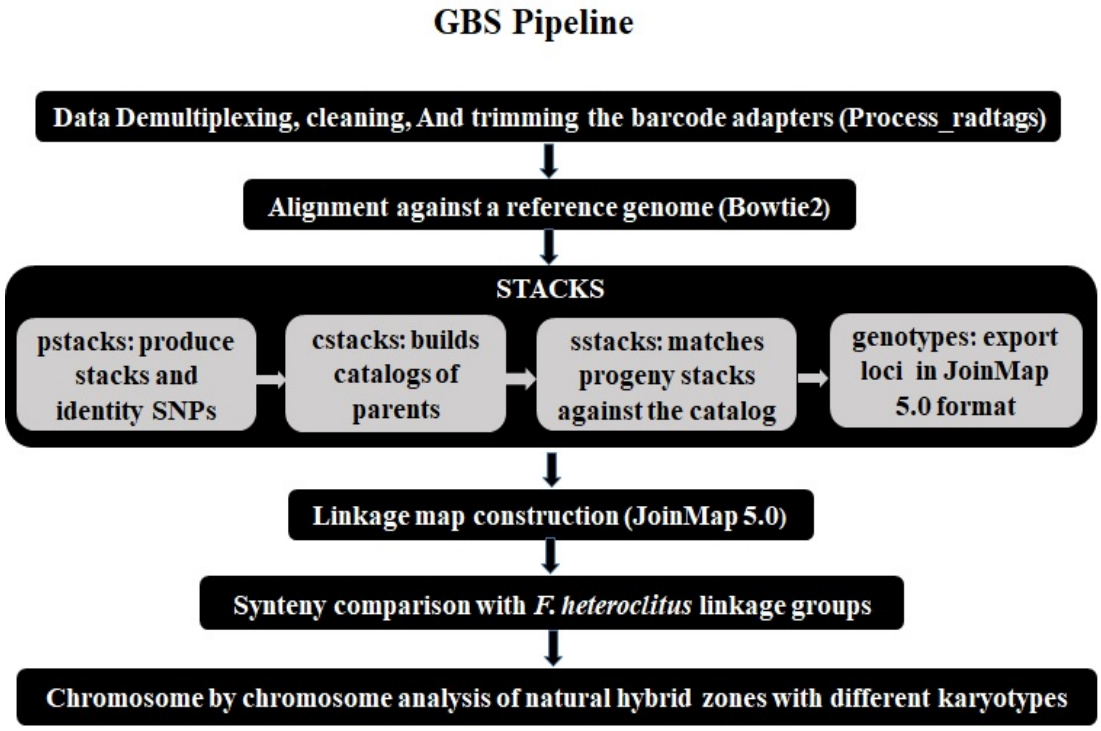


Figure 4.3. Workflow of data analysis after genotyping-by sequencing (GBS).

**4.1.5. Linkage Map Construction.** The SNP loci from STACKS were imported to JoinMap 5.0 (Ooijen, 2011) and the markers were mapped to linkage groups using the "cross-pollinated" (CP) mapping population design. We limited our analysis to biallelic loci. The segregation pattern were determined for each locus as either <nn × np>, <lm × ll>, or <hk × hk>, depending on if one or both parents were homozygous or heterozygous (Van Ooijen, 2006). SNPs with more than 50% missing data were filtered from the analysis. The pattern of allelic segregation for $\chi 2$ goodness-of-fit was calculated for each locus and the markers that significantly deviated from the Mendelian ratio (out of Hardy-Weinberg equilibrium, p value < 0.1) were excluded from the dataset.

Markers were grouped based on their logarithm of odds (LOD) scores and recombination frequency (RF) values. We used a minimum LOD score of 3.0 and a maximum recombination frequency of 0.40 in order to group the markers (Pootakham et al., 2015). Initially, the linkage groups were constructed for each family separately and later the maps from individual families within species were integrated to form consensus maps for each species. The map Integration tool was used for joining maps from each family if they shared two or more markers. For integration, we used the regression mapping algorithm. Map distances were calculated using the Kosambi mapping function to convert recombination fractions between markers to centiMorgan (cM) (Kosambi, 2016).

**4.1.6. Synteny Comparison and Characterization of RB Fusions.** *Fundulus heteroclitus* is a widely studied member of the genus *Fundulus* and consequently, a sequenced genome and mapping data are more fully developed for this species than for any other members in the genus (Adams et al., 2006; Waits et al., 2016). The *Fundulus heteroclitus* genome has been mapped and ordered into 24 linkage groups, which were used to construct 24 physical maps (in base-pair) corresponding to each chromosome (Miller and Whitehead, unpublished data). Contigs of the incomplete *F. olivaceus* draft reference genome were aligned to the contigs of a more complete *F. heteroclitus* reference genome using MUMmer version 4.0 (Marçais et al., 2018). This software package aligned these two genomes using default options. *F. olivaceus* contigs were assigned to *F. heteroclitus* linkage groups based on contig alignments. *F. olivaceus* and *F. notatus* consensus linkage groups were confirmed based on *F. olivaceus* reference genome contigs shared between LGs. If two groups of *F. olivaceus*/*F. heteroclitus* were

joined with one group of *F. notatus*, then that linkage group in *F. notatus* was considered as a Robertsonian fusion.

## 4.2. POPULATION STRUCTURE ANALYSIS OF NATURAL HYBRID ZONES

The following steps were performed to accomplish this part of the project-

**4.2.1. Sample Collection, DNA Preparation, and GBS.** The naturally replicated hybrid zones are excellent systems to study reproductive isolation and genetic introgression between closely related species. We selected two contact zones for this study, Spring River and Tombigbee River (Figure 4.1). These two drainages were selected because the incidence of hybridization were previously reported to be high in these two drainages (Duvernell and Schaefer, 2014). Another reason for choosing these two drainages is the difference in chromosome number of *F. notatus* between these two locations. Both of these zones are the contact zones of *F. olivaceus* and *F. notatus* with *F. olivaceus* having 24 chromosomes. However, an important difference is that in the Spring River, *F. notatus* exhibits 20 haploid chromosomes (4 Rb fusions), whereas, the Tombigbee population of *F. notatus* has 22 haploid chromosomes (2 Rb fusions). Therefore, these two contact zones allowed us to evaluate the impact of number of Rb fusions on reproductive isolation between species. Collection sites were selected to occur within contact zones based on previous studies (Duvernell and Schaefer, 2014). Fish were captured using dip nets, and fin clips were preserved in 100% ethanol. DNA extraction, quantification, and trial digestion for GBS were performed following the protocol described earlier. GBS libraries were constructed and sequenced by Elshire GBS service

following instructions of Elshire et al. (2011) (described in section 4.1.2. GBS library construction and sequencing).

**4.2.2. SNP Discovery.** GBS was used to generate SNP markers for the assessment of hybridization and introgression in two hybrid zones. For this part of the project, the raw GBS data were cleaned, processed, and SNPs were called using TASSEL version 5.0 (Bradbury et al., 2007; Glaubitz et al., 2014). Bowtie 2.0 (Langmead and Salzberg, 2012) was used to align the sequence reads against a previously sequenced *Fundulus heteroclitus* reference genome (https://my.mdibl.org/diplay/FGP/Home) (Reid et al., 2017) using "high-sensitive" and "end-to-end" options (Schaefer et al., 2018). SNPs were called in TASSEL following Schaefer et al. (2016). Data were filtered by locus (bialleleic SNPs with no gaps between alleles, minimum 10% coverage of locus, minimum minor allele frequency rate of 5%) and by individual (minimum read coverage of 10%). heterozygotes were called using a quantitative SNP calling function, "binomial likelihood" that exploits read counts allowing an expected sequencing error rate of 1% (Bradbury et al., 2007). Called SNPs were then exported to R as HapMap files for filtering out high-quality markers.

Additional filtering steps were executed in R using customized scripts before creating input datasets for further population structure analysis (Schaefer et al., 2016). Datasets contained only biallelic loci followed by filtering steps on missing data by locus (>10%) and missing data by individual (>20%). The SNP loci with excess observed heterozygosity ($H_o$) are likely due to miss-alignment of paralogs which is an artefact commonly occurring in GBS data (Nunez et al., 2015). To reduce technical error, the loci that deviated from Hardy-Weinberg Equilibrium (HWE) can be detected and excluded

from the analysis (Hosking et al., 2004). Likewise, we eliminated any markers that exhibited an observed heterozygosity greater than 70%. The loci that were less than 2000 bp apart from each other were further discarded from the analysis to reduce the effects of linkage-disequilibrium. It insures that the SNPs are independent and not on the same scaffold in the reference genome. The final SNP genotypes retained after all filtering steps were used for population analysis.

**4.2.3. Structure of Hybrid Zones at Population Level.** We used two methods to assign individuals to hybrid classes. The first was the STRUCTURE analysis, which determines the admixture proportion of each individual, with predicted admixture proportions of 0 and 1 for parents, 0.5 for F1 and F2 hybrids, and backcross individuals falling in between these values based on their admixture proportions. The second method, NewHybrids, assigned individuals to specific genotypic classes based on individual allelic compositions.

The pattern of hybridization and introgression of both naturally replicated hybrid zones were determined by using STRUCTURE v.2.3.4 (Pritchard et al., 2000) that is based on the parametric Bayesian model-based clustering method. It assigned admixture proportion scores (Q-score: membership coefficient value) to individuals based on their allele frequency into K clusters or population groups, where K was set to 2, corresponding to the two species. Model parameters included a burn-in cycle of 100,000 and a Markov Chain Monte Carlo (MCMC) repetitions of 2,000,000. Two independent runs were executed to confirm convergence of the results. Newhybrids version 1.1 (Anderson and Thompson, 2002) was used to assign individuals from two contact zones to discrete groups of parentals and different hybrid classes including F1s, F2s, and

multiple backcross generations (first, second, third, and fourth) with parentals. 300 SNPs were randomly selected because of limited marker handling capacity of NewHybrids. Analysis parameters for two independent runs were- 1,000,000 MCMC repetitive sweeps after 100,000 burn-in cycles using a Jeffreys-like prior.

**4.2.4. Analysis of Natural Hybrid Zones at Chromosome Level.** The results of *F. olivaceus* and *F. heteroclitus* genome alignment were exported from MUMmer4 and parsed into 24 files each representing one linkage group in *F. heteroclitus*. STRUCTURE analysis was run on each of the linkage groups/chromosomes. Contigs from reference *F. olivaceus* genome were aligned to 24 *F. heteroclitus* linkage groups based on contig alignments. Using the *F. olivaceus* reference, SNPs from Spring and Tombigbee Rivers were aligned to the *F. olivaceus* scaffolds of 24 chromosome groups. STRUCTURE analysis was performed on each chromosome independently using the same parameters described above. Runs were repeated for each drainage.

# 5. RESULTS

## 5.1. SNP DISCOVERY FROM GBS DATA FOR MAPPING

A total of 220 F2 offspring were produced from three and two families of *F. olivaceus* and *F. notatus*, respectively, for conducting genetic analysis (Table 5.1). Two families of *F. olivaceus* shared one male parent (Figure 4.2). After demultiplexing, cleaning, and trimming the raw GBS data, they were aligned to an incomplete *F. olivaceus* reference genome with an average alignment rate of 52.82% in *F. olivaceus* and 51.98% in *F. notatus*. Following analysis in the GBS pipeline software STACKS 1.48 (Catchen et al., 2013), a total of 93,919 SNPs were called from the tags for *F. olivaceus* (123 F2 individuals in three families) and a total of 67,371 SNP tags were called for *F. notatus* (97 F2 individuals in two families). The work flow of modules in STACKS GBS pipeline called, filtered, and genotyped SNP loci in order to produce high-quality markers for linkage groups (Table 5.2). Finally, there was a total of 2572 SNPs from three families in *F. olivaceus* and 1266 SNPs from two families in *F. notatus* that were exported from "genotypes" in STACKS to the mapping program for linkage mapping.

## 5.2. LINKAGE MAP DEVELOPMENT

Initially, the SNP data were processed in JoinMap by calculating genotype frequency and segregation distortion value for each locus. 396 SNP loci from *F. olivaceus* and 293 SNP markers from *F. notatus* datasets were excluded from the analysis because of significant distortion from Mendelian ratios during allele segregation (e.g.

1:2:1 for hk x hk cross) (p-value <0.1). We obtained 24 linkage groups for each family of *F. olivaceus* and 20 linkage groups for each family of *F. notatus* based on LOD value and recombination frequency (Figure 5.1 and 5.2). If the groups shared at least two or more markers between families within species, they were integrated. We found a total 24 consensus linkage groups in *F. olivaceus* containing 1051 SNP loci and 20 linkage groups in *F. notatus* containing 676 SNP markers (Table 5.3). The "suspect linkage" was also checked for each group in JoinMap 5.0 to see if any genetic marker that's present on one group is somehow related to other group and we found no evidence of suspect loci in our map development process. About 51.7% and 30.52% of the markers, imported into JoinMap for *F. notatus* and *F. olivaceus,* respectively, remained ungrouped during map construction. We did not include the ungrouped and excluded markers back in our maps later.

Table 5.1. Number of F2 progeny for each family of *F. olivaceus* and *F. notatus*, and the number of GBS-SNP markers that were genotyped and selected for linkage mapping.

| **Species** | **Family** | **Number of F2 Progeny** | **Number of SNP Markers that were Genotyped** |
|---|---|---|---|
| ***F. olivaceus*** | T1a | 30 | 954 |
| | T1b | 65 | 978 |
| | T10 | 28 | 640 |
| ***F. notatus*** | T14a | 44 | 640 |
| | LT6F | 53 | 626 |
| **Total** | | *F. olivaceus* = 123<br>*F. notatus* = 97<br>Total = 220 F2 offspring | *F. olivaceus* = 2572<br>*F. notatus* = 1266<br>Total = 3838 SNPs |

Table 5.2. Functions and summary statistics of each program of the GBS-pipeline.

| **GBS-pipeline program** | | **Function** | **Result** |
|---|---|---|---|
| **1) Bowtie2** | | Aligned GBS data to draft *F. olivaceus* reference genome | Alignment rate of *F. olivaceus* was 52.82% and *F. notatus* was 51.98% |
| **Stacks 1.48** | **2) pstacks** | Built stacks and identified SNPs for each individual | |
| | **3) cstacks** | Created catalogs of all consensus loci from parents | *F. olivaceus*= 93,919 SNPs<br>*F. notatus* = 43,052 SNPs |
| | **4) sstacks** | Progeny stacks were matched against the catalogs created by cstsacks | *F. olivaceus*: ~44,000 SNPs<br>*F. notatus*: ~30,000 SNPs |
| | **5) genotypes** | SNPs were genotyped and exported to mapping software | *F. olivaceus*= 2572 SNPs<br>*F. notatus* = 1266 SNPs |
| **6) JoinMap 5.0** | | Identified loci that deviated from Mendelian ratio and excluded those markers | *F. olivaceus*= 396 SNPs excluded<br>*F. notatus* = 293 SNPs excluded |
| | | Constructed LGs for each family and integrated maps to create consensus recombination maps | *F. olivaceus*= 24 LGs<br>*F. notatus* = 20 LGs |

The total length of linkage groups ranged from 15.4 cM to 63.1 cM in case of *F. olivaceus* and 15.5 cM to 58.7 cM for *F. notatus*. The average number of markers per map unit was 1.18 markers/cM in *F. olivaceus* and 1.04 markers/cM in *F. notatus* (Table 5.3). 24 integrated linkage groups in *F. olivaceus* covered a total size of 892 cM and the total map distance for *F. notatus* was 651.9 cM covered by 20 Linkage groups.

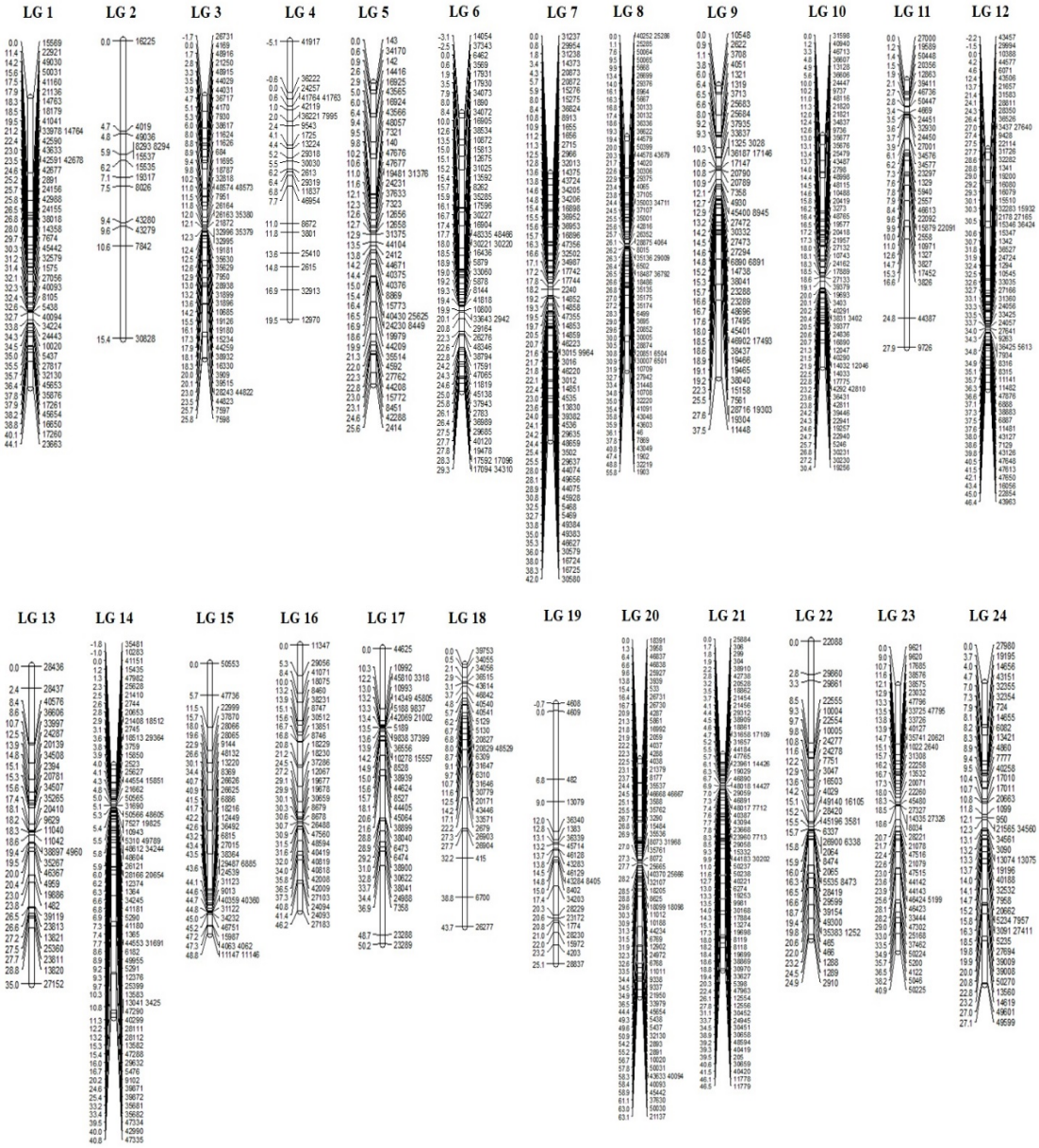


Figure 5.1. Linkage map of *F. olivaceus* constructed in JoinMap 5.0. Numbers on the left side of each linkage group are map positions in CentiMorgan and numbers on the right side represent SNP marker score.

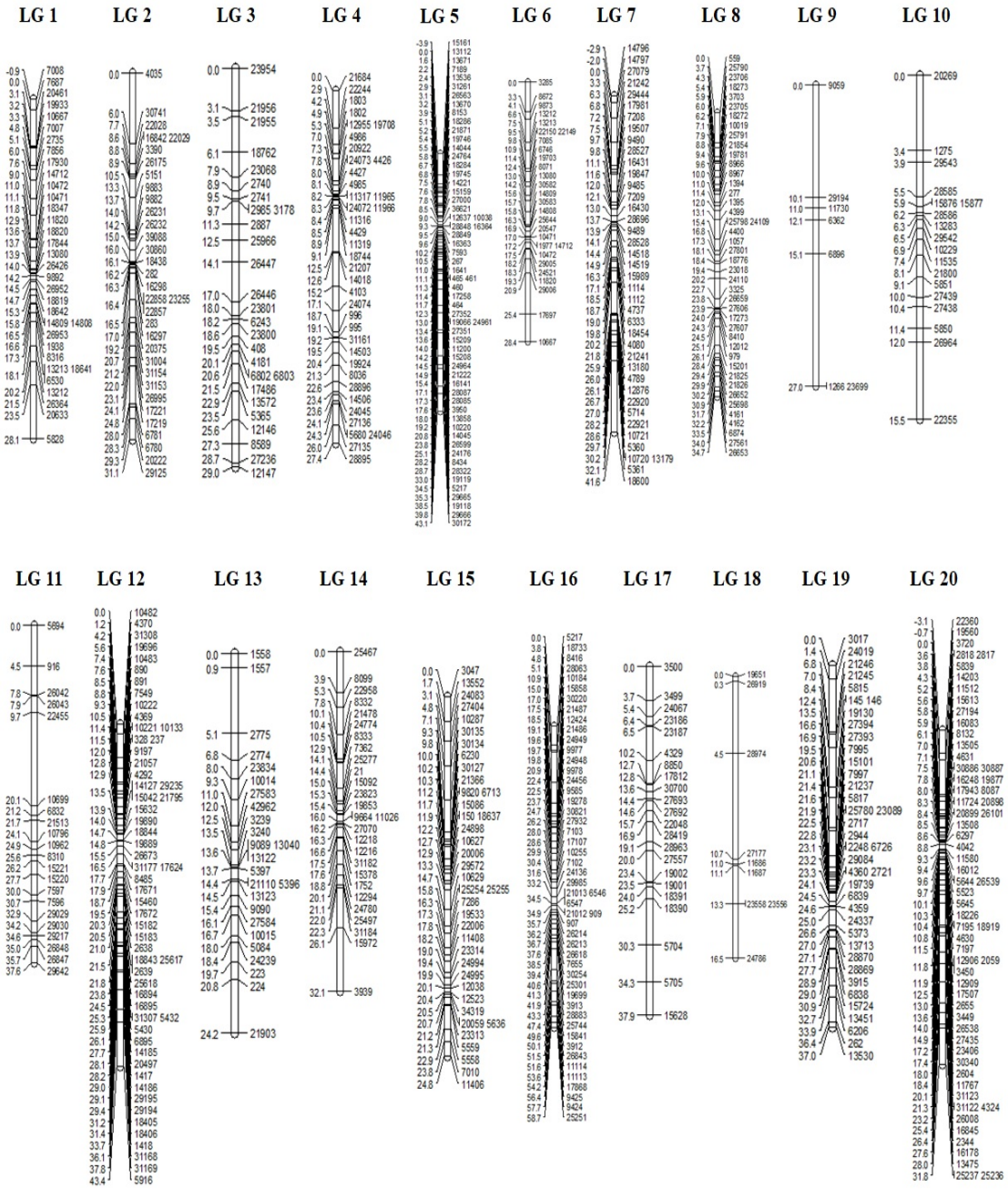


Figure 5.2. Linkage map of *F. notatus* constructed in JoinMap 5.0. Numbers on the left side of each linkage group are map positions in CentiMorgan and numbers on the right side represents SNP marker score.

Table 5.3. Summary of the main features of each linkage map in *F. olivaceus* and *F. notatus*.

| ***Fundulus olivaceus*** | | | | ***Fundulus notatus*** | | | |
|---|---|---|---|---|---|---|---|
| **LG** | **Number of SNPs mapped to each LG** | **Map size (cM)** | **Density of markers (SNP/cM)** | **LG** | **Number of SNPs mapped to each LG** | **Map size (cM)** | **Density of markers (SNP/cM)** |
| **1** | 44 | 44.1 | 1 | **1** | 34 | 28.1 | 1.21 |
| **2** | 13 | 15.4 | 0.84 | **2** | 33 | 31.1 | 1.06 |
| **3** | 48 | 25.8 | 1.86 | **3** | 27 | 29 | 0.93 |
| **4** | 23 | 19.5 | 1.18 | **4** | 38 | 27.4 | 1.39 |
| **5** | 42 | 25.6 | 1.64 | **5** | 60 | 43.1 | 1.39 |
| **6** | 42 | 29.3 | 1.77 | **6** | 28 | 28.4 | 0.99 |
| **7** | 62 | 42.0 | 1.48 | **7** | 40 | 41.6 | 0.96 |
| **8** | 64 | 55.8 | 1.15 | **8** | 43 | 34.7 | 1.24 |
| **9** | 47 | 37.5 | 1.25 | **9** | 7 | 27 | 0.26 |
| **10** | 59 | 30.4 | 1.94 | **10** | 18 | 15.5 | 1.16 |
| **11** | 31 | 27.9 | 1.11 | **11** | 21 | 37.6 | 0.56 |
| **12** | 63 | 46.4 | 1.36 | **12** | 57 | 43.4 | 1.31 |
| **13** | 30 | 35.0 | 0.86 | **13** | 25 | 24.2 | 1.03 |
| **14** | 69 | 40.8 | 1.69 | **14** | 27 | 32.1 | 0.84 |
| **15** | 34 | 48.8 | 0.70 | **15** | 39 | 24.8 | 1.57 |
| **16** | 31 | 46.2 | 0.67 | **16** | 50 | 58.7 | 0.85 |
| **17** | 34 | 50.2 | 0.68 | **17** | 22 | 37.9 | 0.58 |
| **18** | 27 | 43.7 | 0.62 | **18** | 9 | 16.5 | 0.55 |
| **19** | 22 | 25.1 | 0.88 | **19** | 37 | 39 | 0.95 |
| **20** | 65 | 63.1 | 1.03 | **20** | 61 | 31.8 | 1.92 |
| **21** | 68 | 46.5 | 1.46 | | | | |
| **22** | 37 | 24.9 | 1.49 | | | | |
| **23** | 44 | 40.9 | 1.08 | | | | |
| **24** | 42 | 27.1 | 1.55 | | | | |
| **Total** | **1051** | **892** | **1.18** | **Total** | **676** | **651.9** | **1.04** |

## 5.3. MAP COMPARISONS BETWEEN SPECIES USING REFERENCE LINKAGE MAPS

The *F. heteroclitus* (Atlantic killifish) reference genome has been mapped to 24 LGs (correspond to 24 chromosomes) in Dr. Whitehead's lab at the University of California Davis. Physical maps of *F. heteroclitus* were established using high-density RAD-Seq markers (Whitehead, personal contact). Those mapping data of 24 LGs were aligned with the *F. olivaceus* reference genome contigs. Using *F. heteroclitus* scaffolds for each LG as a reference, we identified a total of 280 anchored loci between the two mapped species who shared the common contigs of the *F. olivaceus* reference genome (Figure 5.3). SNPs from JoinMap-constructed *F. olivaceus* and *F. notatus* LGs were identified within those aligned contigs. We were able to establish homology between linkage groups, and found strong synteny between the maps of *F. olivaceus*, *F. notatus*, and *F. olivaceus* at LG level.

Furthermore, four LGs of *F. notatus* (LG15, LG16, LG19, and LG20) aligned against eight LGs of *F. heteroclitus* (LG4, LG16, LG9, LG15, LG10, LG19, LG14, and LG20) which we interpreted as confirmatory evidence of Robertsonian fusions (Figure 5.4). Apart from these translocations, the rest of the 16 LGs aligned with 16 LGs of *F. heteroclitus*. The map distances of all the anchor loci (that were in the same contigs) were exported to SigmaPlot and graphs were created to show the synteny between *F. olivaceus* and *F. notatus* along with *F. heteroclitus* (Figure 5.3). Four fused chromosomes in *F. notatus* that corresponded with eight linkage groups in *F. olivaceus* and that of *F. heteroclitus* are shown in Figure 5.4.

| | *F. olivaceus* LG | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 5 | 6 | 7 | 8 | 11 | 12 | 13 | 17 | 18 | 21 | 22 | 23 | 24 | 4 | 16 | 9 | 15 | 10 | 19 | 14 | 20 |
| ***F. notatus* LG** | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | **17** | | | | | | | | | | | | | | | | | | | | | | | |
| 2 | | **5** | | | | | | | | | | | | | | | | | | | | | | |
| 3 | | | **14** | | | | | | | | | | | | | | | | | | | | | |
| 5 | | | | **14** | | | | | | | | | | | | | | | | | | | | |
| 6 | | | | | **12** | | | | | | | | | | | | | | | | | | | |
| 7 | | | | | | **20** | | | | | | | | | | | | | | | | | | |
| 8 | | | | | | | **17** | | | | | | | | | | | | | | | | | |
| 11 | | | | | | | | **7** | | | | | | | | | | | | | | | | |
| 12 | | | | | | | | | **19** | | | | | | | | | | | | | | | |
| 13 | | | | | | | | | | **12** | | | | | | | | | | | | | | |
| 17 | | | | | | | | | | | **8** | | | | | | | | | | | | | |
| 18 | | | | | | | | | | | | **4** | | | | | | | | | | | | |
| 4 | | | | | | | | | | | | | **12** | | | | | | | | | | | |
| 9 | | | | | | | | | | | | | | **6** | | | | | | | | | | |
| 10 | | | | | | | | | | | | | | | **5** | | | | | | | | | |
| 14 | | | | | | | | | | | | | | | | **12** | | | | | | | | |
| 15 | | | | | | | | | | | | | | | | | **12** | **13** | | | | | | |
| 16 | | | | | | | | | | | | | | | | | | | **11** | **13** | | | | |
| 19 | | | | | | | | | | | | | | | | | | | | | **9** | **11** | | |
| 20 | | | | | | | | | | | | | | | | | | | | | | | **22** | **15** |
| | 1 | 2 | 3 | 5 | 6 | 7 | 8 | 11 | 12 | 13 | 17 | 18 | 21 | 22 | 23 | 24 | 4 | 16 | 9 | 15 | 10 | 19 | 14 | 20 |
| | ***F. heteroclitus* chromosomes** | | | | | | | | | | | | | | | | | | | | | | | |

Figure 5.3. Summary of homology comparisons between *F. olivaceus*, *F. notatus*, and *F. heteroclitus* linkage groups. Numbers on the top line are LGs in *F. olivaceus*; numbers on left are LGs in *F. notatus*; numbers listed on the bottom line are the identity of syntenic *F. heteroclitus* LGs. Bold numbers along the diagonal represent shared SNP loci between these three species that belong to same draft *F. olivaceus* reference genome scaffold.

## 5.4. POPULATION STRUCTURE ANALYSIS OF NATURAL HYBRID ZONES

SNP markers were used for the molecular dissection of two independent hybrid zones, in the Spring and Tombigbee Rivers, respectively. There were 135 and 157 fish fin clips used for DNA extraction from the Spring and Tombigbee drainages. So, a total of 292 DNA samples were genotyped through GBS to produce 75,411 SNP markers. For

Spring River, we started with 36,413 raw loci in 135 individuals, and after filtering, we had 326 loci in 123 individuals. Similarly, in the Tombigbee River the initial number of SNP loci was 38,998 in 157 individuals, and after filtering, 321 markers retained for 153 individuals for genetic structure analysis. Filtering parameters in R and the number of SNPs passed each step are reported in Table 5.4.

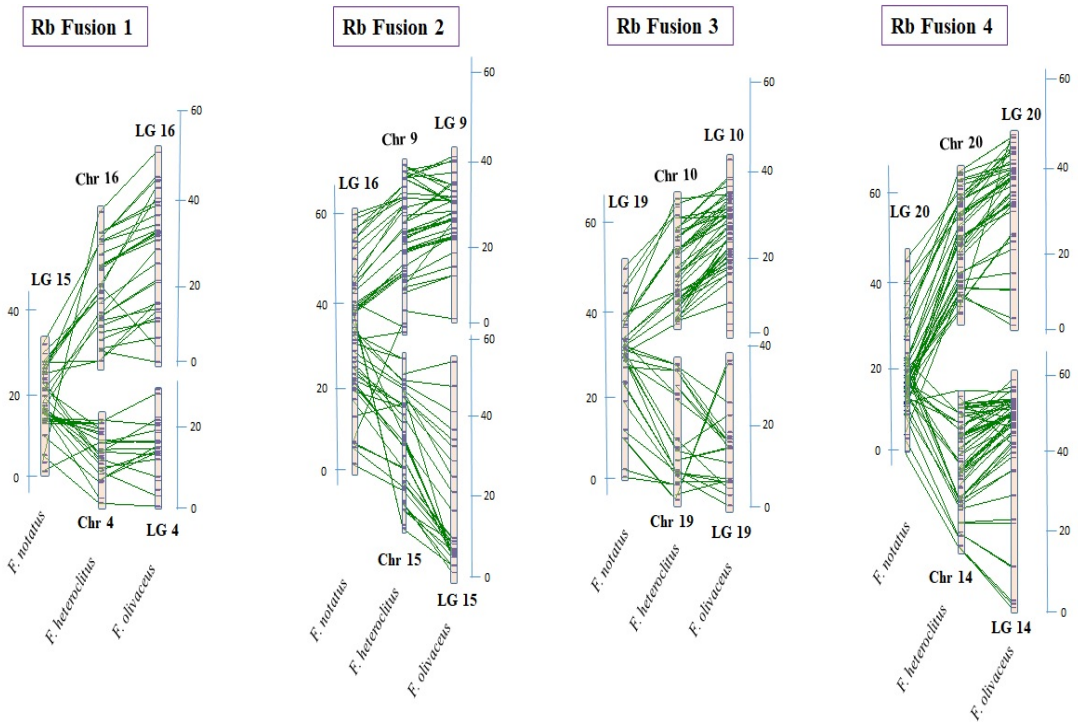


Figure 5.4. Homology between four Robertsonian LGs in *F. notatus* and non-fused single LGs in *F. olivaceus* and *F. heteroclitus*. Each LG of *F. notatus* (left side) aligned against two LG of *F. heteroclitus* (in middle) and *F. olivaceus* (right side). Green lines connected the position of same markers on different linkage groups. Scales are drawn in centiMorgan (cM).

Table 5.4. Summary of each filtering steps in R to select SNP markers for STRUCTURE analysis.

| **R-filters** | **Spring River** | | **Tombigbee River** | |
|---|---|---|---|---|
| **Filtering steps** | **Number of SNP loci retained** | **Number of Individuals retained** | **Number of SNP loci retained** | **Number of Individuals retained** |
| **Before filters** | 36413 | 135 | 38998 | 157 |
| **Missing data by locus (> 0.1)** | 2921 | 135 | 1977 | 157 |
| **Missing data by individual (> 0.2)** | 2921 | 123 | 1977 | 153 |
| **Observed heterozygosity ($H_o$) (>70%)** | 2708 | 123 | 1806 | 153 |
| **Minimum distance (<2000bp)** | 326 | 123 | 321 | 153 |

The high-quality polymorphic SNPs were then used to investigate population structures of Spring and Tombigbee Rivers. 300 SNPs were randomly selected for each drainage in order to analyze data in NewHybrids. This program assigned a probability value to each individual inferring which genotype class (parents, F1s, F2s, and backcross) that individual belongs to. Assignments were independently validated by the results from STRUCTURE, which assigned an admixture proportion to each individual. STRUCTURE was ran on whole data sets of two drainages independently for two genetic population clusters (K=2).

Individuals were sorted according to their Q-score (membership-coefficient value) and bar plots were constructed to visualize the distribution of admixture proportions

(Figure 5.5). The cutoffs of admixture proportion value or Q-score (ranges from 0 to 1) to determine F1 hybrids and backcross individuals were justified by using NewHybrids' assignment of individuals into different groups. On our admixture scale, a Q value of 0 indicated pure *F. olivaceus* and 1 denoted pure *F. notatus*. Q-values around 0.5 correspond to F1 and F2 hybrids, and first-generation backcross individuals have Q-value of 0.25 or 0.75. Graphs were then compared between two contact zones to understand the patterns of hybridization and introgression. From both analyses, we observed similar numbers of parents of each species and similar proportions of F1 hybrids in both contact zones. But, the extent of backcross individuals in both parental directions differed substantially across the two drainages (Figure 5.5).

Table 5.5. Delta-Q values for each pair of chromosomes that are fused in *F. notatus*.

| **Spring River** | | **Tombigbee River** | |
|---|---|---|---|
| **Chromosome pairs involved in one Rb-translocation** | **Delta-Q value between each pair** | **Chromosome pairs involved in one Rb-translocation** | **Delta-Q value between each pair** |
| Rb fusion-1 (Chromosome 4 vs. 16) | 0.173 | Rb fusion-1 (Chromosome 4 vs. 16) | 0.186 |
| Rb fusion-1 (Chromosome 9 vs. 15) | 0.254 | Rb fusion-1 (Chromosome 4 vs. 16) | 0.228 |
| Rb fusion-1 (Chromosome 10 vs. 19) | 0.222 | Rb fusion-1 (Chromosome 4 vs. 16) | 0.187 |
| Rb fusion-1 (Chromosome 4 vs. 16) | **0.03** | Rb fusion-1 (Chromosome 4 vs. 16) | 0.281 |

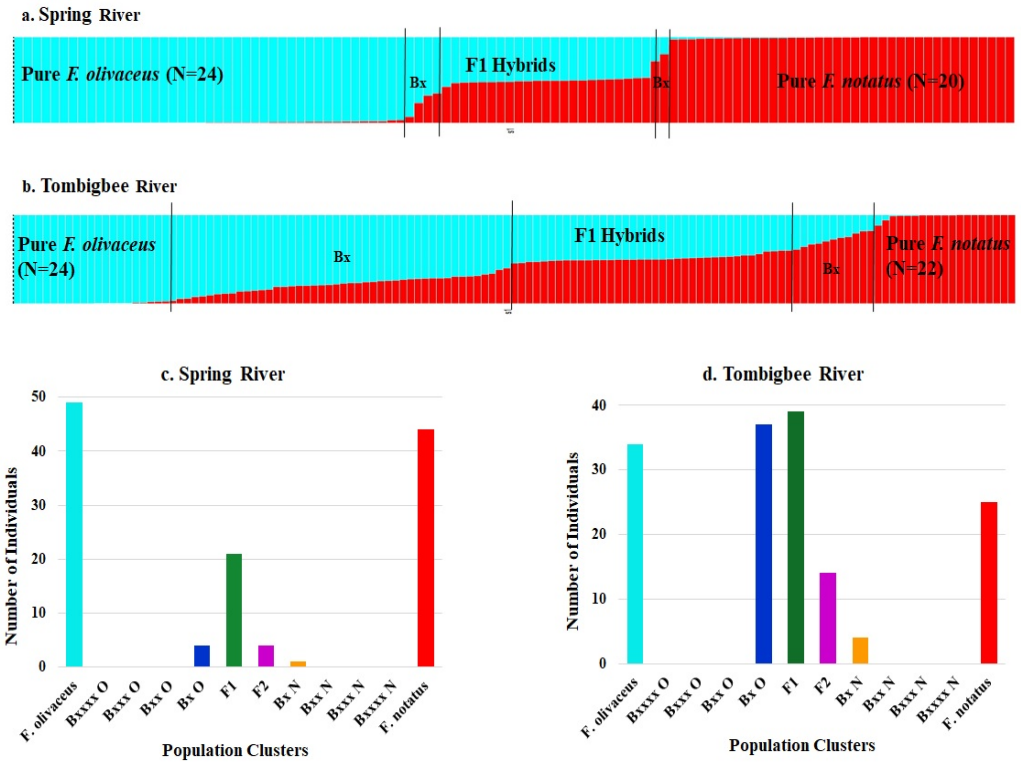


Figure 5.5. Population structure of two natural contact zones. a) Spring River and b) Tombigbee River drainages. STRUCTURE version 2.3.4 was used to assign admixture proportions to individuals using a quality-filtered dataset. The mean proportion of the membership of each cluster of two different populations is indicated by two colors- blue is pure *F. olivaceus* and red is pure *F. notatus*. Hybridization patterns of c) Spring River and d) Tombigbee River. NewHybrids version 1.1 was used to place individuals into discrete groups (parentals, F1, and backcrosses with parentals up to fourth generation).

## 5.5. CHROMOSOME-BY-CHROMOSOME ANALYSIS OF NATURAL HYBRID ZONES

After constructing 24 LGs in *F. olivaceus* and 20 LGs (with 4 Rb fusions) in *F. notatus,* and finding homology between species, we performed a chromosome-by-chromosome (one chromosome at a time) analysis of our hybrid zone data. Results (Q-values from STRUCTURE) were exported for 24 LGs of each contact zones and plotted on a dot plot to see the pattern of recombination for in each chromosome across these two

rivers (Figure 5.6). Data points at Q = 0 were homozygous for *F. olivaceus*, and Q = 1 were homozygous for *F. notatus*, and data points at Q = 0.5 were heterozygotes. Data points that fell between these values were generally in backcross individuals, and were putative F1 recombinant chromosomes.

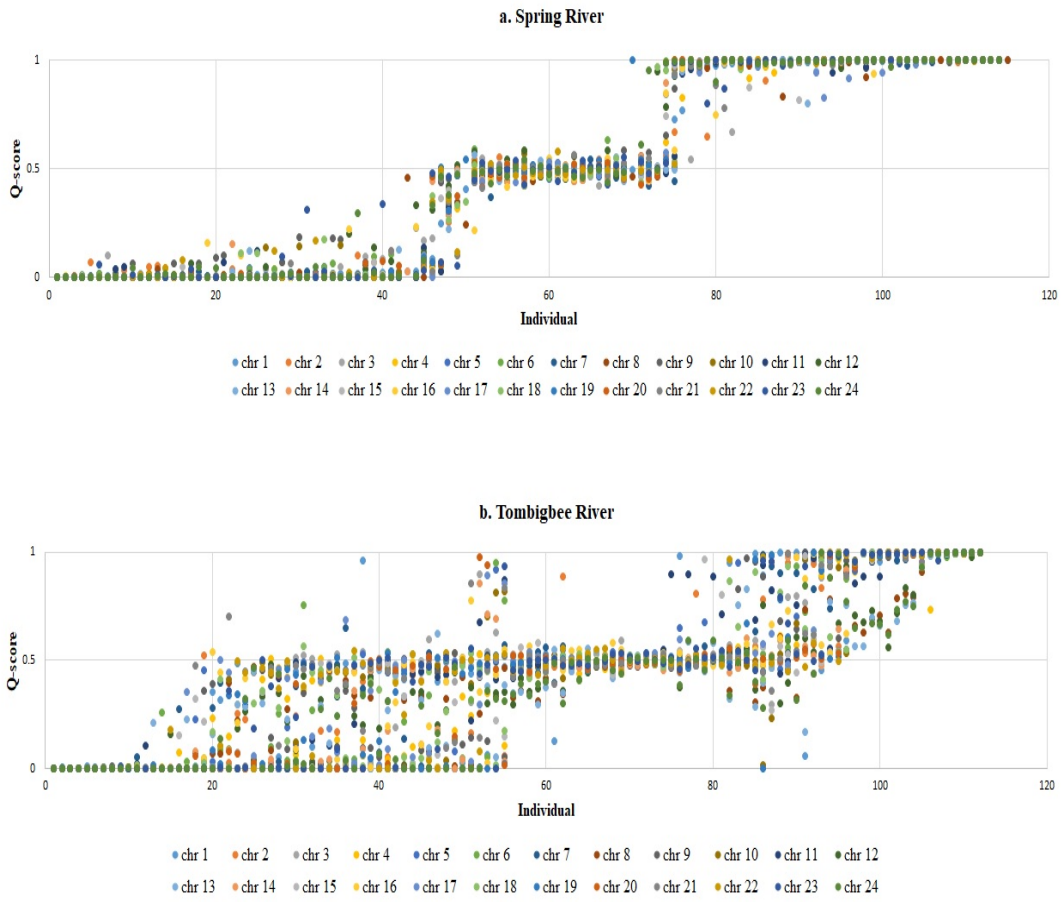


Figure 5.6. Chromosome-by-chromosome STRUCTURE dot plot of hybrid zones. a) Spring River and b) Tombigbee River. Model parameter K was set to 2 (two population clusters). Dots of different colors indicate 24 *Fundulus* chromosomes.

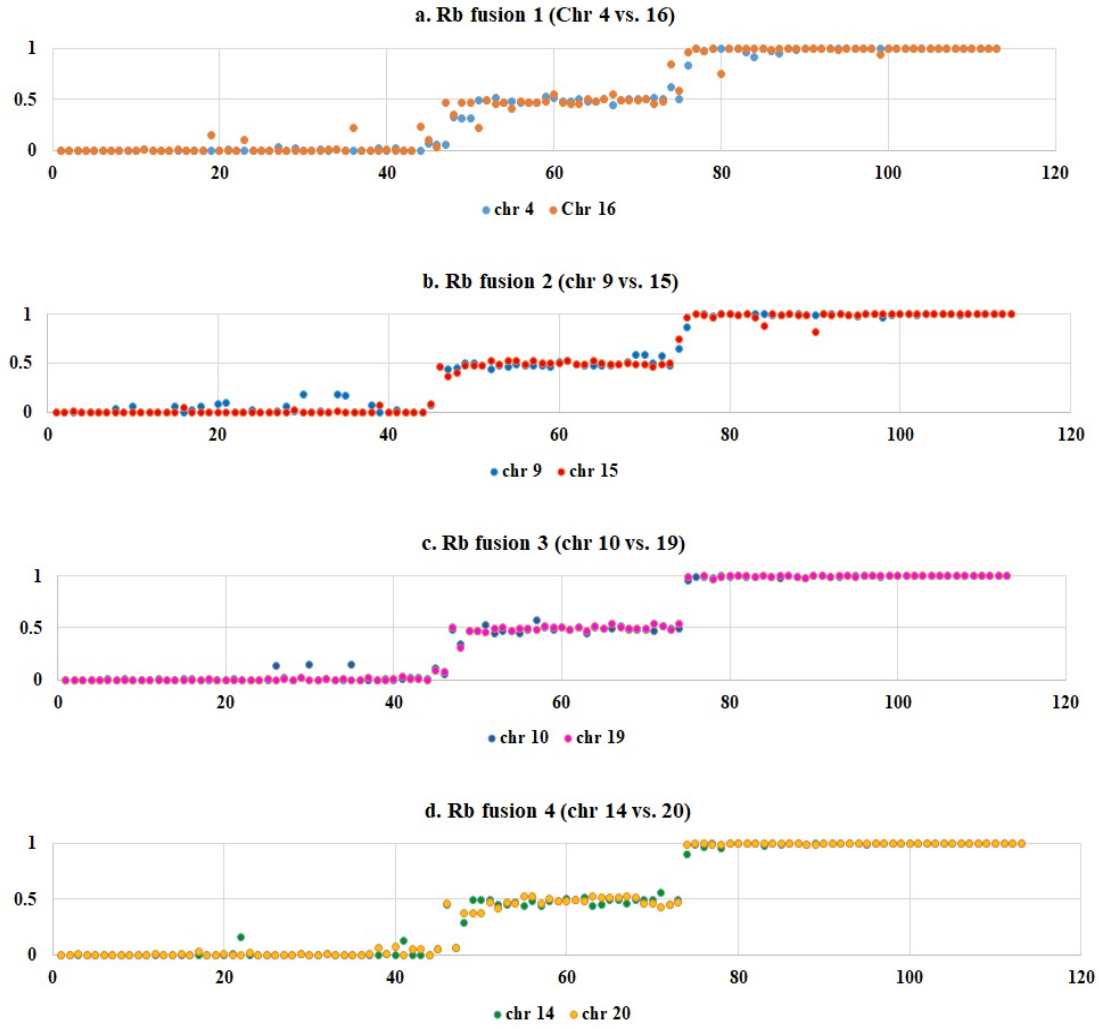


Figure 5.7. STRUCTURE dot plots for two chromosomes that were assumed to be involved in each Rb metacentric. a) Rb-fusion 1 b) Rb-fusion-2 c) Rb-fusion-3 and d) Rb-fusion-4 of Spring River. Dots of two different colors in each plot indicate two *Fundulus* chromosomes that were predicted to be fused to form one Rb translocation.

We compared the graphs between two contact zones and found contrasting patterns of hybridization and recombination. In Tombigbee River, we found a relatively large proportion of backcross individuals, with many exhibiting evidence of F1 recombination, whereas, in Spring River, only a very few backcross individuals were

inferred (Figure 5.6). Separate dotplots were also created (using Q-scores) for each drainages with chromosomes that were involved with Rb-fusions to see their pattern of segregation and recombination in each contact zone.

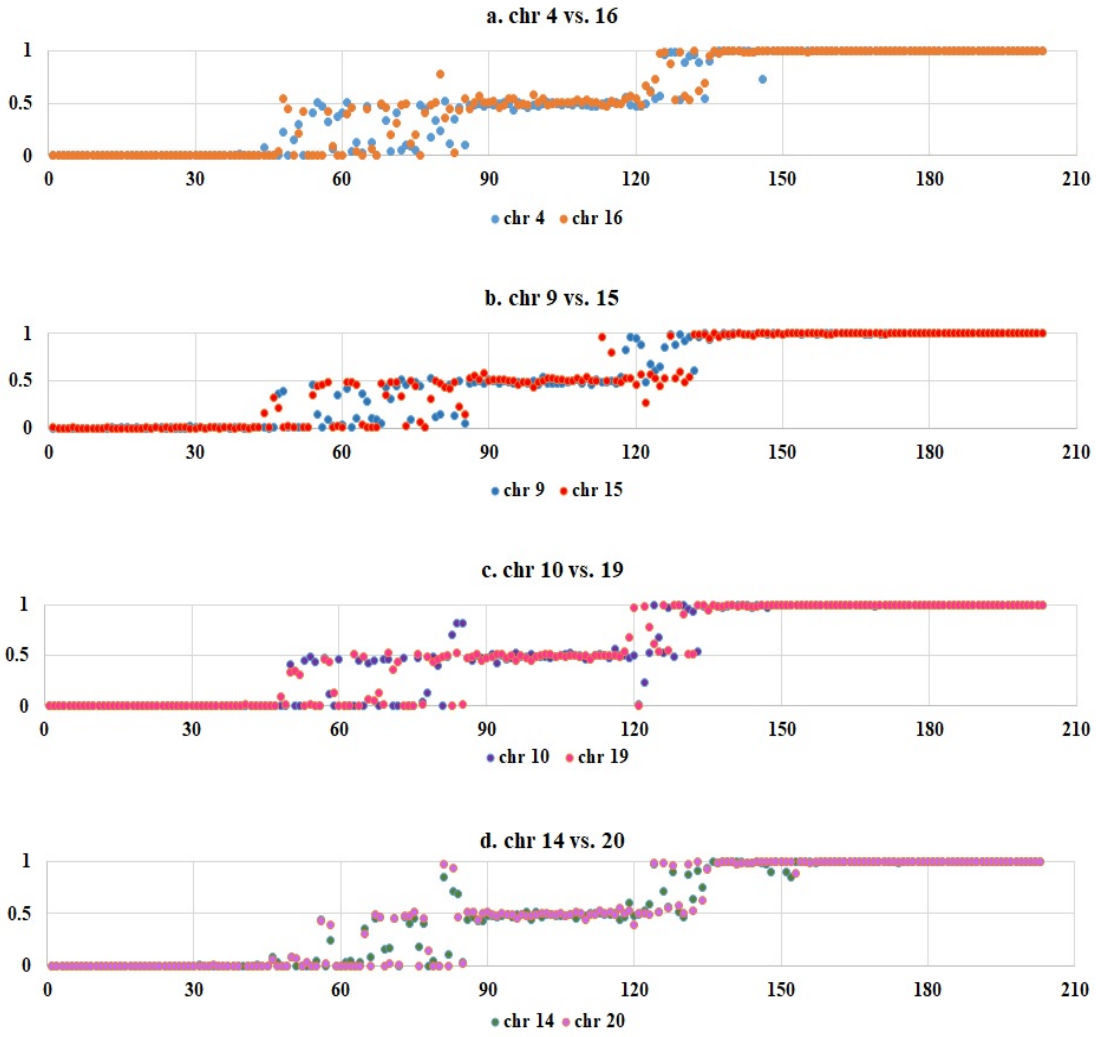


Figure 5.8. STRUCTURE dot plots for the two chromosomes that were assumed to be involved in each Rb metacentric. a) Rb-fusion 1 b) Rb-fusion-2 c) Rb-fusion-3 and d) Rb-fusion-4 of Tombigbee River (any of the four chromosomes should be fused to two Rb-chromosomes). Dots of two different colors in each plot indicate two *Fundulus* chromosomes that were predicted to be fused to form one Rb translocation.

The average delta-Q values were calculated between two chromosomes that we predicted to be fused together for the backcross individuals and compared between each pair (Figure 5.7 and 5.8). Since the viable gametes of F1 hybrids are assumed to get either one large metacentric or two small acrocentric chromosomes from the grandparents, the dots for two fused chromosomes should cluster. So, the delta-Q values were expected to be very small for each Rb fusion. We observed only one of these pairs (Rb-fusion 4) showed low delta-q value in Spring River and none of the translocated chromosomes showed low delta-Q value in Tombigbee River (Table 5.5). Two independent Rb translocations in Tombigbee clade of *F. notatus* (not mapped yet).

# 6. DISCUSSION

## 6.1. GENETIC LINKAGE MAPS OF TWO *FUNDULUS* SPECIES

We constructed genetic recombination maps for both *F. olivaceus* and *F. notatus* based on high density SNP markers (1.18 SNP loci per CentiMorgan in *F. olivaceus* and 1.04 SNP loci per Centimorgan in *F. notatus*) which were generated using the Genotype by-Sequencing method. These linkage maps provide valuable genomic resources for these two species in order to address evolution and speciation related questions. These can also provide useful groundwork for future mapping studies, synteny comparison studies with other closely related teleost fish species, as well as for population genetic studies at the molecular level. There are only a couple other *Fundulus* species that have been mapped to linkage groups to reveal chromosomal rearrangements or study molecular genetics (Berdan et al., 2014; Waits et al., 2016). However, multiple chromosomal translocations in *F. olivaceus* and *F. notatus*, evident from karyotypic studies (Black and Howell, 1978; Chen, 1971; Setzer, 1970), have never been characterized before using high-resolution molecular markers. In this study, we established 24 linkage groups for Blackspotted topminnow and 20 linkage groups for Blackstripped topminnow with molecular markers (SNPs) at an average interval of 0.89cM. Our results are similar to other species in the family Fundulidae (Berdan et al., 2014; Waits et al., 2016).

The number of linkage groups for each species corresponded to the number of chromosomes observed from karyotypic studies (Chen, 1971) (Figure 1.3). The genetic recombination maps were more saturated in markers when the maps from individual

families were integrated to create consensus maps. In most of the linkage groups, the SNP markers were distributed uniformly along the central regions, except for a few distal regions. We noticed only three gaps in *F. olivaceus* (two on LG 17 one on LG 18) and three gaps in *F. notatus* (two on LG 9 and one on LG 11) that were larger than 10 cM on each consensus linkage group. These gaps were possibly due to the failure of GBS to detect polymorphic markers (SNPs) in that specific regions of the genome. Other possibilities include- these regions are representing recombination hotspots across the genome, or polymorphisms were absent in those particular sections because of being identical-by-descent among the parent species (Pootakham et al., 2015).

## 6.2. MAP RELATIONSHIPS BETWEEN SPECIES, HOMOLOGY, AND ROBERTSONIAN TRANSLOCATIONS

We compared the linkage maps of both Blackspotted topminnow (*F. olivaceus*) and Blackstripped topminnow (*F. notatus*) with a recently sequenced and mapped Atlantic Killifish (*F. heteroclitus*) reference genome (Waits et al., 2016) and found strong synteny between the linkage groups of each species. While combining and comparing our recombination maps with the map of *F. heteroclitus* (Miller et al., unpublished data) and draft *F. olivaceus* reference contigs (Whitehead, unpublished data), we found evidence of strong synteny. The 24 linkage groups constructed for *F. olivaceus* in this study aligned to 24 chromosomes in *F. heteroclitus* confirming that these two species shared the same ancestral karyotypes of N= 24. The 16 linkage groups of *F. notatus* showed one-to-one correspond to *F. heteroclitus* chromosomes and the remaining four linkage groups aligned against eight chromosomes in the Atlantic killifish. This confirmed that karyotypic differences between species are best explained by Robertsonian fusion of

acrocentric chromosomes in *F. notatus*. Chen (1971), karyotyped 20 *Fundulus* species and found that the majority of them exhibited 24 chromosomes with only a few (8 species) having a reduced number of chromosomes. He showed four large metacentrics in *F. notatus* and our results supported the hypothesis of possessing four Rb translocations in this species. A similar mapping study has been performed by Berdan et al. (2014), where they established 24 LGs in Bluefin killifish and 23 LGs in Rainwater killifish along with the presence of one Rb fusion. Several lines of evidence support that these large metacentic chromosomes were the result of chromosomal fusions in an F. *notatus* ancestor, and not because of fission (one large metacentric chromosome breaking across the centromere to form two small acrocentric chromosomes). First, Chen's study (1971) supported that the ancestral karyotype included N = 24, and that N = 20 in *F. notatus* and some other killifish species is derived from this typical ancestral number. Second, *F. heteroclitus* is another closely related species of the genus *Fundulus* (Rodgers et al., 2018) and Waits et al. (2016) established 24 linkage groups mapped to 24 chromosomes in that species. Our study connected these studies and provided support to the hypothesis that the derived karyotype of *F. notatus* is the result of four individual Robertsonian translocation events. For each fused group, the SNP markers on the top half and the markers on the bottom half aligned with two individual linkage groups in *F. olivaceus*. These four fused groups were expected to be the largest linkage groups, but our analysis could not support that hypothesis. This might be due to not having enough GBS-SNP markers for *F. notatus* to cover the entire chromosome if markers on only a portion of the large metacentric chromosomes were detected.

We found large-scale conservation of synteny at the linkage group level between *F. olivaceus*, *F. notatus*, and *F. heteroclitus* while comparing genetic maps. Including *F. heteroclitus*, some other fish also showed conservation of synteny between species. Waits et al. found *F. heteroclitus* chromosomes syntenous with Medaka (*Oryzias latipes*-Beloniformes N= 24) which is a distantly related teleost fish. The fused and non-fused chromosomes/linkage groups in *Lucania goodei* and *L. parva* also showed preservation of synteny (Berdan et al., 2014). Other evidences, such as, one fused chromosome in guppies (*Poecilia reticulate*) (Tripathi et al., 2009), two fused chromosomes in tilapia (*Oreochromis niloticus*) (Liu et al., 2013), also showed synteny with medaka chromosomes.

During comparison of synteny, the SNP marker orders on each linkage group occasionally showed local inconsistencies between the maps of *F. olivaceus*, *F. notatus*, and *F. heteroclitus*. These discrepancies could be indicators of the presence of local inversions, duplications, deletions, or other types of rearrangements in the parents' genotypes (Han et al., 2011). However, the relatively small number of markers included in our final maps make these conclusions tenuous, and additional mapping efforts would be required to rule out technical errors within linkage groups.

## 6.3. CONTRASTING PATTERN OF HYBRIDIZATION BETWEEN TWO NATURAL HYBRID ZONES

Naturally replicated hybrid zones are the result of secondary contacts between closely related species and occurred where tributaries transit large rivers (Schaefer et al., 2016). Both Spring and Tombigbee River drainages contain *F. olivaceus* with 24 chromosomes but the number of chromosomes of *F. notatus* is different across these two

hybrid zones (N = 20 and 22 in Spring and Tombigbee respectively) (Black and Howell, 1978). Using 647 SNP markers generated through genotyping-by-sequencing in 292 fish samples from these two hybrid zones, we measured the level of genetic divergence and patterns of hybridization across these two rivers. We observed a fairly large proportion of F1 hybrids in both drainages (21.14% of the total population in Spring River and 23.53% in Tombigbee River) suggesting that the propensity of the two species to hybridize is similar in both drainages. So, it appears that prezyogotic barriers to hybridization are limited in these drainages. However, the proportions of backcross individuals (reproduction between F1 hybrids and either one of the parents) differed strikingly between the two rivers. This discrepancy in the proportion of backcross individuals could result from differences in reproductive viability of F1 hybrids, and their ability to mate and produce offspring, which could be influenced by differences in *F. notatus* karyotypes between these two drainages (since there are twice as many Rb fusions in Spring *F. notatus* population than in Tombigbee *F. notatus* population). If Rb heterozygotes experienced partial sterility due to aneuploidy, then we predicted that Spring River F1 individuals would have lower fertility than Tombigbee River F1 hybrids. Our data were consistent with this prediction.

Our data do not provide direct evidence of reproductive viability of F1 males and females. We do not know if males were sterile, or if females exhibited reduced fertility. The degree of subfertility in male versus female F1 hybrids could be determined experimentally. A previous study demonstrated that both pre- and post-zygotic reproductive barriers may have a role on the reproductive isolation in *F. notatus* complex. Vigueira et al. (2008) observed a six-fold reduction in hatching success of eggs produced

by either male or female hybrids between *F. olivaceus* and *F. notatus*. Therefore, subfertility of F1 hybrids is likely aid on the order of about 17% relative to members of either species. Theoretically, due to aneuploidy, fitness reduction should be multiplicative, though empirical studies indicate that this is in fact not the case. For example, Baker and Bickham (1986) reported that individuals that are Rb heterozygous for three centric fusions in house mice (*Mus musculus*) suffered only about 25% fertility reduction or less. Rock-wallabies showed to have no suppression of gene flow due to either the simple fusion (populations differ by one or more non-overlapping Rb fusions) or the complex fusions (different populations have different fusions involving the same chromosomes) (Potter et al., 2015). They predicted that complex fusions would result in low interspecific gene flow (reproductive barrier) while simple fusions would exhibit high gene flow (no reproductive barrier). They even have documented that male F1 are sterile, while female F1 exhibit subfertility.

We also observed the pattern of segregation of each LG and compared recombination rate between Spring and Tombigbee River. The pattern of introgression at LG level was same as we found while analyzing population genetic structures using the whole genome. We noticed a remarkably large number of recombinants among the backcrosses in Tombigbee River than that of Spring River. In an F1 Rb heterozygote, the only viable gametes are the ones that either get one metacentric chromosome from one grandparent, or both acrocentric chromosomes from the other grandparent. Under those circumstances, the two acrocentric chromosomes that we predicted to fuse together should be linked and exhibit similar admixture proportions. However, there are lots of dots that were separated among the backcross individuals in both drainages (Figure 5.7

and 5.8) which would imply that recombination took place at Rb chromosomes in those individuals. We were not able to use correlations in admixture proportions between LG pairs to identify fusions in the Tombigbee race of *F. notatus*. Discernment of chromosomal fusions in those *F. notatus* populations must await further mapping efforts.

# APPENDIX A.

# MAPPING TOOLS AND PARAMETERS

**BOWTIE2 2.3**

# This shell script aligned the raw sequence data (FASTQ.gz) to a reference genome (*F. olivaceus.*

```
#!/bin/bash
FILES=*fastq.gz
n=1
for f in $FILES
      do
   t=${f##/*/}
   ar=(${t//[_.]/ })
   s=${ar[1]}_${ar[2]}_${ar[3]}_${ar[0]}
   echo Processing $n. $s
echo $s
   ((n=n+1))
   done > fasta_name.readgroup.txt
```

**SAMTOOLS 1.9**

# This shell script converted SAM files to BAM format, indexed, and sorted the files to a directory.

```
#!/bin/bash
# Convert each .sam file to .bam file, sort and index each file
FILES=aligned/*.sam
P=
```

```
for f in $FILES
    do
    echo Convert .sam alignments to .bam, sort and index $f
    t=${f##/*/}
    #echo $t
    n=${t/.*/}
    #echo $n.bam
    samtools view -bS $P"$t" > $P"$n".bam
    samtools sort $P"$n".bam -o $P"$n"_sorted.bam
    samtools index $P"$n"_sorted.bam $P"$n"_sorted.bai
#     rm $f
    done
```

**PSTACKS**

#This command line was used to extract stackts, which had been aligned to a reference genome, and then identified SNPs.

#pstacks -t bam -f ./aligned/sample_map.bam -o ./stacks -i 1 -m 3 -p 18

| Model Parameter | Description | Value |
|---|---|---|
| t | Input file type | BAM |
| f | Path to where the input files are | |
| o | Path to output directory | |

| Model Parameter | Description | Value |
|---|---|---|
| i | An integer ID which should be unique for each sample | 1,2,3,…….. |
| m | Minimum depth of coverage to consider as a stack | 3 |
| p | Number of threads for parallel execution | 18 |

# Shell script for pstacks to process all samples at a time.

```
#!/bin/bash
set -e
set -u
set -o pipefail
ID=1
for i in ./aligned/*.bam ;
        do
        pstacks -t bam -f "$i" -o ./stacks -i $ID -m 3 -p 18
        ID=$((ID+1))
        Done
```

## CSTACKS

# This command line was used to build a catalog containing a set of all possible consensus loci expected from parents.

# Construct catalog of *F. olivaceus* parents

#cstacks -b 1 -s ./stacks/F1_T6_T1aF_map -s ./stacks/F1_T17_T1aT1bM_map -s ./stacks/F1_T7_T1bF_map -s ./stacks/F1_T6_T10M_map -s ./stacks/F1_T7_T10F_map -o ./stacks -p 18 --aligned

# Construct catalog of *F. notatus* parents

#cstacks -b 1 -s ./stacks/F1_T20_T14M_map -s ./stacks/F1_T22_T14F_map -s ./stacks/NF1_T22_LT5BF_map -s ./stacks/NF1_T25_LT5BM_map -s ./stacks/NF1_T22_LT6FM_map -s ./stacks/NF1_T25_LT6FF_map -s ./stacks/NF1_T25_T17M_map -s ./stacks/NF1_T22_T17F_map -o ./stacks -p 18 --aligned

| Model Parameter | Description | Value |
|---|---|---|
| b | Batch ID | 1 |
| s | Path to where the input files resides | Parents only |
| o | Path to output directory | |
| p | Number of threads for parallel execution | 18 |

## SSTACKS

# SNP loci determined by pstacks will be searched against the catalog of parents.

# Three families separately for *F. olivaceus*

**T1a:**

#sstacks -b 1 -c ./stacks/batch_1 -o ./stacks -s ./stacks/F2_T1a_1_map -s ./stacks/F2_T1a_2_map -s ./stacks/F2_T1a_3_map -s ./stacks/F2_T1a_4_map -s

./stacks/F2_T1a_5_map -s ./stacks/F2_T1a_6_map -s ./stacks/F2_T1a_7_map -s ./stacks/F2_T1a_8_map -s ./stacks/F2_T1a_9_map -s ./stacks/F2_T1a_10_map -s ./stacks/F2_T1a_11_map -s ./stacks/F2_T1a_12_map -s ./stacks/F2_T1a_13_map -s ./stacks/F2_T1a_14_map -s ./stacks/F2_T1a_15_map -s ./stacks/F2_T1a_16_map -s ./stacks/F2_T1a_17_map -s ./stacks/F2_T1a_18_map -s ./stacks/F2_T1a_19_map -s ./stacks/F2_T1a_20_map -s ./stacks/F2_T1a_21_map -s ./stacks/F2_T1a_22_map -s ./stacks/F2_T1a_24_map -s ./stacks/F2_T1a_25_map -s ./stacks/F2_T1a_26_map -s ./stacks/F2_T1a_27_map -s ./stacks/F2_T1a_28_map -s ./stacks/F2_T1a_29_map -s ./stacks/F2_T1a_30_map -s ./stacks/F2_T1a_31_map -s ./stacks/F1_T6_T1aF_map -s ./stacks/F1_T17_T1aT1bM_map -o ./stacks -c ./stacks/batch_1 -p 18 --aligned

**T1b:**

#sstacks -b 1 -c ./stacks/batch_1 -o ./stacks -s ./stacks/F2_T1b_32_map -s ./stacks/F2_T1b_33_map -s ./stacks/F2_T1b_34_map -s ./stacks/F2_T1b_35_map -s ./stacks/F2_T1b_36_map -s ./stacks/F2_T1b_37_map -s ./stacks/F2_T1b_38_map -s ./stacks/F2_T1b_39_map -s ./stacks/F2_T1b_40_map -s ./stacks/F2_T1b_41_map -s ./stacks/F2_T1b_42_map -s ./stacks/F2_T1b_43_map -s ./stacks/F2_T1b_44_map -s ./stacks/F2_T1b_45_map -s ./stacks/F2_T1b_46_map -s ./stacks/F2_T1b_49_map -s ./stacks/F2_T1b_50_map -s ./stacks/F2_T1b_51_map -s ./stacks/F2_T1b_52_map -s ./stacks/F2_T1b_53_map -s ./stacks/F2_T1b_54_map -s ./stacks/F2_T1b_55_map -s ./stacks/F2_T1b_56_map -s ./stacks/F2_T1b_57_map -s ./stacks/F2_T1b_58_map -s ./stacks/F2_T1b_59_map -s ./stacks/F2_T1b_60_map -s ./stacks/F2_T1b_61_map -s ./stacks/F2_T1b_62_map -s ./stacks/F2_T1b_63_map -s ./stacks/F2_T1b_64_map -s ./stacks/F2_T1b_65_map -s ./stacks/F2_T1b_66_map -s ./stacks/F2_T1b_67_map -s

./stacks/F2_T1b_68_map -s ./stacks/F2_T1b_69_map -s ./stacks/F2_T1b_70_map -s ./stacks/F2_T1b_71_map -s ./stacks/F2_T1b_72_map -s ./stacks/F2_T1b_73_map -s ./stacks/F2_T1b_74_map -s ./stacks/F2_T1b_75_map -s ./stacks/F2_T1b_76_map -s ./stacks/F2_T1b_77_map -s ./stacks/F2_T1b_78_map -s ./stacks/F2_T1b_79_map -s ./stacks/F2_T1b_80_map -s ./stacks/F2_T1b_81_map -s ./stacks/F2_T1b_82_map -s ./stacks/F2_T1b_83_map -s ./stacks/F2_T1b_84_map -s ./stacks/F2_T1b_85_map -s ./stacks/F2_T1b_86_map -s ./stacks/F2_T1b_87_map -s ./stacks/F2_T1b_88_map -s ./stacks/F2_T1b_89_map -s ./stacks/F2_T1b_90_map -s ./stacks/F2_T1b_91_map -s ./stacks/F2_T1b_92_map -s ./stacks/F2_T1b_93_map -s ./stacks/F2_T1b_94_map -s ./stacks/F2_T1b_95_map -s ./stacks/F2_T1b_96_map -s ./stacks/F2_T1b_97_map -s ./stacks/F2_T1b_98_map -s ./stacks/F1_T17_T1aT1bM_map -s ./stacks/F1_T7_T1bF_map -p 18 --aligned

**T10:**

#sstacks -b 1 -c ./stacks/batch_1 -o ./stacks -s ./stacks/F2_T10_99_map -s ./stacks/F2_T10_100_map -s ./stacks/F2_T10_101_map -s ./stacks/F2_T10_102_map -s ./stacks/F2_T10_103_map -s ./stacks/F2_T10_104_map -s ./stacks/F2_T10_105_map -s ./stacks/F2_T10_106_map -s ./stacks/F2_T10_107_map -s ./stacks/F2_T10_108_map -s ./stacks/F2_T10_109_map -s ./stacks/F2_T10_110_map -s ./stacks/F2_T10_111_map -s ./stacks/F2_T10_112_map -s ./stacks/F2_T10_113_map -s ./stacks/F2_T10_114_map -s ./stacks/F2_T10_115_map -s ./stacks/F2_T10_116_map -s ./stacks/F2_T10_117_map -s ./stacks/F2_T10_118_map -s ./stacks/F2_T10_119_map -s ./stacks/F2_T10_120_map -s ./stacks/F2_T10_121_map -s ./stacks/F2_T10_122_map -s ./stacks/F2_T10_123_map -s

./stacks/F2_T10_124_map -s ./stacks/F2_T10_125_map -s ./stacks/F2_T10_126_map -s ./stacks/F1_T6_T10M_map -s ./stacks/F1_T7_T10F_map -p 18 --aligned

# Two families separately for *F. notatus*

**T14:**

#sstacks -b 1 -c ./stacks/batch_1 -o ./stacks -s ./T14a/F1_T20_T14M_map -s ./T14a/F1_T22_T14F_map -s ./T14a/F2_T14a_2_map -s ./T14a/F2_T14a_3_map -s ./T14a/F2_T14a_4_map -s ./T14a/F2_T14a_8_map -s ./T14a/F2_T14a_9_map -s ./T14a/F2_T14a_10_map -s ./T14a/F2_T14a_11_map -s ./T14a/F2_T14a_13_map -s ./T14a/F2_T14a_15_map -s ./T14a/F2_T14a_17_map -s ./T14a/F2_T14a_18_map -s ./T14a/F2_T14a_19_map -s ./T14a/F2_T14a_20_map -s ./T14a/F2_T14a_21_map -s ./T14a/F2_T14a_22_map -s ./T14a/F2_T14a_55_map -s ./T14a/F2_T14a_56_map -s ./T14a/F2_T14a_57_map -s ./T14a/F2_T14a_58_map -s ./T14a/F2_T14a_59_map -s ./T14a/F2_T14a_60_map -s ./T14a/F2_T14a_61_map -s ./T14a/F2_T14a_62_map -s ./T14a/F2_T14a_63_map -s ./T14a/F2_T14a_64_map -s ./T14a/F2_T14a_65_map -s ./T14a/F2_T14a_66_map -s ./T14a/F2_T14a_68_map -s ./T14a/F2_T14a_69_map -s ./T14a/F2_T14a_70_map -s ./T14a/F2_T14a_71_map -s ./T14a/F2_T14a_72_map -s ./T14a/F2_T14a_73_map -s ./T14a/F2_T14a_74_map -s ./T14a/F2_T14a_75_map -s ./T14a/F2_T14a_76_map -s ./T14a/F2_T14a_77_map -s ./T14a/F2_T14a_78_map -s ./T14a/F2_T14a_79_map -s ./T14a/F2_T14a_80_map -s ./T14a/F2_T14a_81_map -s ./T14a/F2_T14a_82_map -s ./T14a/NF2_T14_110_map -s ./T14a/NF2_T14_111_map -p 18 –aligned

**LT6F:**

#sstacks -b 1 -c ./LT6F/batch_1 -o ./LT6F -s ./LT6F/NF1_T22_LT6FM_map -s ./LT6F/NF1_T25_LT6FF_map -s ./LT6F/F2_LT6F_23_map -s ./LT6F/F2_LT6F_24_map -s ./LT6F/F2_LT6F_25_map -s ./LT6F/NF2_LT6F_26_map -s ./LT6F/NF2_LT6F_27_map -s ./LT6F/NF2_LT6F_28_map -s ./LT6F/NF2_LT6F_29_map -s ./LT6F/NF2_LT6F_30_map -s ./LT6F/NF2_LT6F_31_map -s ./LT6F/NF2_LT6F_32_map -s ./LT6F/NF2_LT6F_33_map -s ./LT6F/NF2_LT6F_34_map -s ./LT6F/NF2_LT6F_35_map -s ./LT6F/NF2_LT6F_36_map -s ./LT6F/NF2_LT6F_37_map -s ./LT6F/NF2_LT6F_38_map -s ./LT6F/NF2_LT6F_39_map -s ./LT6F/NF2_LT6F_40_map -s ./LT6F/NF2_LT6F_41_map -s ./LT6F/NF2_LT6F_42_map -s ./LT6F/NF2_LT6F_43_map -s ./LT6F/NF2_LT6F_44_map -s ./LT6F/NF2_LT6F_45_map -s ./LT6F/NF2_LT6F_46_map -s ./LT6F/NF2_LT6F_47_map -s ./LT6F/NF2_LT6F_93_map -s ./LT6F/NF2_LT6F_94_map -s ./LT6F/NF2_LT6F_95_map -s ./LT6F/NF2_LT6F_96_map -s ./LT6F/NF2_LT6F_97_map -s ./LT6F/NF2_LT6F_98_map -s ./LT6F/NF2_LT6F_99_map -s ./LT6F/NF2_LT6F_100_map -s ./LT6F/NF2_LT6F_101_map -s ./LT6F/NF2_LT6F_102_map -s ./LT6F/NF2_LT6F_103_map -s ./LT6F/NF2_LT6F_104_map -s ./LT6F/NF2_LT6F_105_map -s ./LT6F/NF2_LT6F_106_map -s ./LT6F/NF2_LT6F_117_map -s ./LT6F/NF2_LT6F_118_map -s ./LT6F/NF2_LT6F_119_map -s

./LT6F/NF2_LT6F_120_map -s ./LT6F/NF2_LT6F_121_map -s

./LT6F/NF2_LT6F_122_map -s ./LT6F/NF2_LT6F_123_map -s

./LT6F/NF2_LT6F_124_map -s ./LT6F/NF2_LT6F_125_map -s

./LT6F/NF2_LT6F_126_map -s ./LT6F/NF2_LT6F_127_map -s

./LT6F/NF2_LT6F_128_map -s ./LT6F/NF2_LT6F_129_map -s

./LT6F/NF2_LT6F_130_map -p 18 –aligned

| Model Parameter | Description | Value |
|---|---|---|
| b | Batch ID | 1 |
| c | Path to the catalog directory | |
| s | Path to where the input files resides | Parents and progeny |
| o | Path to output directory to store results | |
| p | Number of threads for parallel execution | 18 |

**GENOTYPES**

# This command line was executed to generate input files for mapping program.

# -m and -r options were modified for each family of each species

#genotypes -b 1 -t CP -o joinmap -P ./T1a -r 21 -c -m 40

#genotypes -b 1 -t CP -o joinmap -P ./T1b -r 45 -c -m 40

#genotypes -b 1 -t CP -o joinmap -P ./T10 -r 20 -c -m 40

#genotypes -b 1 -t CP -o joinmap -P ./T14a -r 30 -c -m 25

#genotypes -b 1 -t CP -o joinmap -P ./T1a -r 35 -c -m 25

| Model Parameter | Description | Value |
|---|---|---|
| b | Batch ID | 1 |
| t | Population map type | CP |
| o | Type of output file | joinmap |
| P | Path to output directory to store files | |
| c | Allow automated corrections to the data | |
| r | Minimum number of F2 progeny to call a SNP marker | modified |
| m | Minimum stack depth to export a SNP locus for one individual | modified |

**JOINMAP 5.0**

| Parameter | Value |
|---|---|
| Maximum recombination frequency between markers to consider linkages | 0.40 |
| Minimum LOD to group markers | 3.00 |
| Goodness-of-fit jump threshold for removal of loci | 5.00 |
| Mapping algorithm | Regression model |
| Mapping function | Kosambi's |
| Maximum number of neighboring markers used to construct maps | 5 |
| Burn-in-chain length | 10,000 |
| Minimum number of markers shared between families to integrate maps | 2 |

# APPENDIX B.

# POPULATION GENETICS ANALYSIS OF TWO HYBRID ZONES

**R-SCRIPT**

```
# This customized R-script was used to filter out quality loci for STRUCTURE and
NewHybrids analysis.

library(hapmap)
unzip("hapmap.zip",list=T)
#load tombigbee data
temp<-read_hapmap("tombigbee.hmp.txt")
#or
#load spring data
temp<-read_hapmap("spring.hmp.txt")

# changed function - supply the name of the file with meta data
# this file has to be in the hamap.zip file with the hapmap file
hmp<-parse_hapmap(temp,"tombigbee_meta.csv")

# delete the raw hapmap data to free memory
rm(temp)

# use this to drop unwanted samples prior to any filtering
# case sensitive, complete sample names in quotes, commas between all except the last
one
drop_samples<-c(              )
```

```
hmp<-remove_sample_name(hmp,drop_samples)
# use this to drop all samples based on similar names
# example here will drop all samples that contain "BLANK"
drop_samples<-row.names(hmp[[1]])[grep("BLANK",row.names(hmp[[1]]))]
hmp<-remove_sample_name(hmp,drop_samples)

# filtering
hmp<-biallelic(hmp)
hmp<-dropindels(hmp)
hmp<-filtermissingloci(hmp,0.1,show_hist = T)
hmp<-filtermissingind(hmp,0.2,show_hist=T)
hmp<-filterhetero(hmp,0.7,show_hist = T)
hmp<-filterdistance(hmp,min_bp = 2000,show_hist = F)
# rarefy - pick random subset of SNPs for NewHybrids
hmp<-hmp_rarefy(hmp,300)

# export functions
# for all of these, populations set based on first three letters of sample name
hmp_structure(hmp,"tombigbee.str")
#or
hmp_structure(hmp,"spring.str”)
```

# BIBLIOGRAPHY

Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J.E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C.A., Buggs, R., Butlin, R.K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S.H., Hermansen, J.S., Hewitt, G., Hudson, A.G., Jiggins, C., Jones, J., Keller, B., Marczewski, T., Mallet, J., Martinez-Rodriguez, P., Möst, M., Mullen, S., Nichols, R., Nolte, A.W., Parisod, C., Pfennig, K., Rice, A.M., Ritchie, M.G., Seifert, B., Smadja, C.M., Stelkens, R., Szymura, J.M., Väinölä, R., Wolf, J.B.W., Zinner, D., 2013. Hybridization and speciation. J. Evol. Biol. 26, 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x

Adams, S.M., Lindmeier, J.B., Duvernell, D.D., 2006. Microsatellite analysis of the phylogeography, Pleistocene history and secondary contact hypotheses for the killifish, Fundulus heteroclitus: MICROSATELLITE ANALYSIS OF F. HETEROCLITUS. Mol. Ecol. 15, 1109–1123. https://doi.org/10.1111/j.1365-294X.2006.02859.x

Adega, F., Guedes-Pinto, H., Chaves, R., 2009. Satellite DNA in the Karyotype Evolution of Domestic Animals – Clinical Considerations. Cytogenet. Genome Res. 126, 12–20. https://doi.org/10.1159/000245903

Ahn, S., Tanksley, S.D., 1993. Comparative linkage maps of the rice and maize genomes. Proc. Natl. Acad. Sci. 90, 7980–7984. https://doi.org/10.1073/pnas.90.17.7980

Akkaya, M.S., Shoemaker, R.C., Specht, J.E., Bhagwat, A.A., Cregan, P.B., 1995. Integration of Simple Sequence Repeat DNA Markers into a Soybean Linkage Map. Crop Sci. 35, 1439–1445. https://doi.org/10.2135/cropsci1995.0011183X003500050030x

Anderson, E.C., Thompson, E.A., n.d. A Model-Based Method for Identifying Species Hybrids Using Multilocus Genetic Data 14.

Atwell, S., Huang, Y.S., Vilhjálmsson, B.J., Willems, G., Horton, M., Li, Y., Meng, D., Platt, A., Tarone, A.M., Hu, T.T., Jiang, R., Muliyati, N.W., Zhang, X., Amer, M.A., Baxter, I., Brachi, B., Chory, J., Dean, C., Debieu, M., de Meaux, J., Ecker, J.R., Faure, N., Kniskern, J.M., Jones, J.D.G., Michael, T., Nemri, A., Roux, F., Salt, D.E., Tang, C., Todesco, M., Traw, M.B., Weigel, D., Marjoram, P., Borevitz, J.O., Bergelson, J., Nordborg, M., 2010. Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. Nature 465, 627–631. https://doi.org/10.1038/nature08800

Ayala, F.J., Coluzzi, M., 2005. Chromosome speciation: Humans, Drosophila, and mosquitoes. Proc. Natl. Acad. Sci. 102, 6535–6542. https://doi.org/10.1073/pnas.0501847102

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A., Johnson, E.A., 2008. Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. PLOS ONE 3, e3376. https://doi.org/10.1371/journal.pone.0003376

Baker, R.J., Bickham, J.W., 1986. Speciation by monobrachial centric fusions. Proc. Natl. Acad. Sci. 83, 8245–8248. https://doi.org/10.1073/pnas.83.21.8245

Barton, N.H., Hewitt, G.M., 1989. Adaptation, speciation and hybrid zones. Nature 341, 497. https://doi.org/10.1038/341497a0

Barton, N.H., Hewitt, G.M., 1985. Analysis of Hybrid Zones. Annu. Rev. Ecol. Syst. 16, 113–148. https://doi.org/10.1146/annurev.es.16.110185.000553

Berdan, E.L., Kozak, G.M., Ming, R., Rayburn, A.L., Kiehart, R., Fuller, R.C., 2014. Insight Into Genomic Changes Accompanying Divergence: Genetic Linkage Maps and Synteny of Lucania goodei and L. parva Reveal a Robertsonian Fusion. G3 Genes Genomes Genet. 4, 1363–1372. https://doi.org/10.1534/g3.114.012096

Black, A., Howell, W.M., 1978. A Distinctive Chromosomal Race of the Cyprinodontid Fish, Fundulus notatus, from the Upper Tombigbee River System of Alabama and Mississippi. Copeia 1978, 280. https://doi.org/10.2307/1443564

Blazak, W.F., Eldridge, F.E., 1977. A Robertsonian Translocation and its Effect upon Fertility in Brown Swiss Cattle. J. Dairy Sci. 60, 1133–1142. https://doi.org/10.3168/jds.S0022-0302(77)83999-4

Bradbury, P.J., Zhang, Z., Kroon, D.E., Casstevens, T.M., Ramdoss, Y., Buckler, E.S., 2007. TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23, 2633–2635. https://doi.org/10.1093/bioinformatics/btm308

Brenna-Hansen, S., Li, J., Kent, M.P., Boulding, E.G., Dominik, S., Davidson, W.S., Lien, S., 2012. Chromosomal differences between European and North American Atlantic salmon discovered by linkage mapping and supported by fluorescence in situ hybridization analysis. BMC Genomics 13, 432. https://doi.org/10.1186/1471-2164-13-432

Britton-Davidian, J., Catalan, J., Ramalhinho, M. da G., Ganem, G., Auffray, J.-C., Capela, R., Biscoito, M., Searle, J.B., Mathias, M. da L., 2000. Rapid chromosomal evolution in island mice. Nature 403, 158. https://doi.org/10.1038/35003116

Burke, J.M., Lai, Z., Salmaso, M., Nakazato, T., Tang, S., Heesacker, A., Knapp, S.J., Rieseberg, L.H., 2004. Comparative Mapping and Rapid Karyotypic Evolution in the Genus Helianthus. Genetics 167, 449–457. https://doi.org/10.1534/genetics.167.1.449

Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., Cresko, W.A., 2013. Stacks: an analysis tool set for population genomics. Mol. Ecol. 22, 3124–3140. https://doi.org/10.1111/mec.12354

Chapman, R.S., Hesketh, L.J., 2000. Behavioral phenotype of individuals with Down syndrome. Ment. Retard. Dev. Disabil. Res. Rev. 6, 84–95. https://doi.org/10.1002/1098-2779(2000)6:2<84::AID-MRDD2>3.0.CO;2-P

Chen, T.R., 1971. A comparative chromosome study of twenty killifish species of the genus Fundulus (Teleostei: Cyprinodontidae). Chromosoma 32. https://doi.org/10.1007/BF00285254

Cockram, J., White, J., Zuluaga, D.L., Smith, D., Comadran, J., Macaulay, M., Luo, Z., Kearsey, M.J., Werner, P., Harrap, D., Tapsell, C., Liu, H., Hedley, P.E., Stein, N., Schulte, D., Steuernagel, B., Marshall, D.F., Thomas, W.T.B., Ramsay, L., Mackay, I., Balding, D.J., Consortium, T.A., Waugh, R., O'Sullivan, D.M., 2010. Genome-wide association mapping to candidate polymorphism resolution in the unsequenced barley genome. Proc. Natl. Acad. Sci. 107, 21611–21616. https://doi.org/10.1073/pnas.1010179107

Coluzzi, M., Bradley, D.J., 1999. The malaria challenge : after one hundred years of malariology, in: The Malaria Challenge : After One Hundred Years of Malariology. Presented at the The Malariology Centenary Conference, Lombardo Editore.

Coyne, J., Orr, H.A., 2004. Speciation. Sinauer Associates, Inc., Sunderland, Massachusetts U.S.A.

Doganlar, S., Frary, A., Daunay, M.-C., Lester, R.N., Tanksley, S.D., 2002. A Comparative Genetic Linkage Map of Eggplant (Solanum melongena) and Its Implications for Genome Evolution in the Solanaceae. Genetics 161, 1697–1711.

Duvernell, D.D., Meier, S.L., Schaefer, J.F., Kreiser, B.R., 2013. Contrasting phylogeographic histories between broadly sympatric topminnows in the Fundulus notatus species complex. Mol. Phylogenet. Evol. 69, 653–663. https://doi.org/10.1016/j.ympev.2013.07.013

Duvernell, D.D., Schaefer, J.F., 2014. Variation in contact zone dynamics between two species of topminnows, Fundulus notatus and F. olivaceus, across isolated drainage systems. Evol. Ecol. 28, 37–53. https://doi.org/10.1007/s10682-013-9653-z

Duvernell, D.D., Schaefer, J.F., Hancks, D.C., Fonoti, J.A., Ravanelli, A.M., 2007. Hybridization and reproductive isolation among syntopic populations of the topminnows Fundulus notatus and F. olivaceus. J. Evol. Biol. 20, 152–164. https://doi.org/10.1111/j.1420-9101.2006.01213.x

Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., Mitchell, S.E., 2011. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. PLOS ONE 6, e19379. https://doi.org/10.1371/journal.pone.0019379

Fernhall, B., Pitetti, K.H., Rimmer, J.H., McCubbin, J.A., Rintala, P., Millar, A.L., Kittredge, J., Burkett, L.N., 1996. Cardiorespiratory capacity of individuals with mental retardation including Down syndrome. Med. Sci. Sports Exerc. 28, 366–371. https://doi.org/10.1097/00005768-199603000-00012

Garagna, S., Page, J., Fernandez-Donoso, R., Zuccotti, M., Searle, J.B., 2014a. The Robertsonian phenomenon in the house mouse: mutation, meiosis and speciation. Chromosoma 123, 529–544. https://doi.org/10.1007/s00412-014-0477-6

Garagna, S., Page, J., Fernandez-Donoso, R., Zuccotti, M., Searle, J.B., 2014b. The Robertsonian phenomenon in the house mouse: mutation, meiosis and speciation. Chromosoma 123, 529–544. https://doi.org/10.1007/s00412-014-0477-6

Glaubitz, J.C., Casstevens, T.M., Lu, F., Harriman, J., Elshire, R.J., Sun, Q., Buckler, E.S., 2014. TASSEL-GBS: A High Capacity Genotyping by Sequencing Analysis Pipeline. PLOS ONE 9, e90346. https://doi.org/10.1371/journal.pone.0090346

Gropp, A., 1982. Consequences of Robertsonian heterozygosity: segregational impairment of fertility versus male-limited sterility. Genet. Control Gamete Prod. Funct. 115–151.

Gropp, A., 1981. Robertsonian translocations : cytology, meiosis, segregation pattern and biological consequences of heterozygosity. Symp Zool Soc Lond 47, 141–181.

Gustavsson, I., 1979. Distribution and Effects of the 1/29 Robertsonian Translocation in Cattle. J. Dairy Sci. 62, 825–835. https://doi.org/10.3168/jds.S0022-0302(79)83334-2

Haigis, K.M., Dove, W.F., 2003. A Robertsonian translocation suppresses a somatic recombination pathway to loss of heterozygosity. Nat. Genet. 33, 33–39. https://doi.org/10.1038/ng1055

Han, Y., Zheng, D., Vimolmangkang, S., Khan, M.A., Beever, J.E., Korban, S.S., 2011. Integration of physical and genetic maps in apple confirms whole-genome and segmental duplications in the apple genome. J. Exp. Bot. 62, 5117–5130. https://doi.org/10.1093/jxb/err215

Hauffe, H.C., Giménez, M.D., Searle, J.B., 2012. Chromosomal hybrid zones in the house mouse. Cambridge university press, pp. 407–430. https://doi.org/10.1017/CBO9781139044547.018

Hewitt, G.M., 1988. Hybrid zones-natural laboratories for evolutionary studies. Trends Ecol. Evol. 3, 158–167. https://doi.org/10.1016/0169-5347(88)90033-X

Hosking, L., Lumsden, S., Lewis, K., Yeo, A., McCarthy, L., Bansal, A., Riley, J., Purvis, I., Xu, C.-F., 2004. Detection of genotyping errors by Hardy–Weinberg equilibrium testing. Eur. J. Hum. Genet. 12, 395. https://doi.org/10.1038/sj.ejhg.5201164

Howell, W.M., Black, A., 1981. Karyotypes in Populations of the Cyprinodontid Fishes of the Fundulus notatus Species-complex: A Geographic Analysis. Bull. Ala. Mus. Nat. Hist. 19–30.

Huang, L., Chi, J., Nie, W., Wang, J., Yang, F., 2006. Phylogenomics of several deer species revealed by comparative chromosome painting with Chinese muntjac paints. Genetica 127, 25–33. https://doi.org/10.1007/s10709-005-2449-5

Kodama, M., Brieuc, M.S.O., Devlin, R.H., Hard, J.J., Naish, K.A., 2014. Comparative Mapping Between Coho Salmon (Oncorhynchus kisutch) and Three Other Salmonids Suggests a Role for Chromosomal Rearrangements in the Retention of Duplicated Regions Following a Whole Genome Duplication Event. G3amp58 GenesGenomesGenetics 4, 1717–1730. https://doi.org/10.1534/g3.114.012294

Kong, A., Gudbjartsson, D.F., Sainz, J., Jonsdottir, G.M., Gudjonsson, S.A., Richardsson, B., Sigurdardottir, S., Barnard, J., Hallbeck, B., Masson, G., Shlien, A., Palsson, S.T., Frigge, M.L., Thorgeirsson, T.E., Gulcher, J.R., Stefansson, K., 2002. A high-resolution recombination map of the human genome. Nat. Genet. 31, 241–247. https://doi.org/10.1038/ng917

Kosambi, D.D., 2016. The Estimation of Map Distances from Recombination Values, in: Ramaswamy, R. (Ed.), D.D. Kosambi: Selected Works in Mathematics and Statistics. Springer India, New Delhi, pp. 125–130. https://doi.org/10.1007/978-81-322-3676-4_16

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359. https://doi.org/10.1038/nmeth.1923

Leitwein, M., Guinand, B., Pouzadoux, J., Desmarais, E., Berrebi, P., Gagnaire, P.-A., 2017. A Dense Brown Trout (Salmo trutta) Linkage Map Reveals Recent Chromosomal Rearrangements in the Salmo Genus and the Impact of Selection on Linked Neutral Diversity. G3 Genes Genomes Genet. 7, 1365–1376. https://doi.org/10.1534/g3.116.038497

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Liu, F., Sun, F., Li, J., Xia, J.H., Lin, G., Tu, R.J., Yue, G.H., 2013. A microsatellite-based linkage map of salt tolerant tilapia (Oreochromis mossambicus x Oreochromis spp.) and mapping of sex-determining loci. BMC Genomics 14, 58. https://doi.org/10.1186/1471-2164-14-58

Logue, D.N., Harvey, M.J., 1978. A 14/20 Robertsonian translocation in Swiss Simmental cattle. Res. Vet. Sci. 25, 7–12.

Lohse, K., Clarke, M., Ritchie, M.G., Etges, W.J., 2015. Genome-wide tests for introgression between cactophilic *Drosophila* implicate a role of inversions during speciation: GENOME-WIDE TESTS FOR INTROGRESSION. Evolution 69, 1178–1190. https://doi.org/10.1111/evo.12650

Maddox, J.F., 2001. An Enhanced Linkage Map of the Sheep Genome Comprising More Than 1000 Loci. Genome Res. 11, 1275–1289. https://doi.org/10.1101/gr.GR-1350R

Mallet, J., 2007. Hybrid speciation. Nature 446, 279–283. https://doi.org/10.1038/nature05706

Mallet, J., 2005. Hybridization as an invasion of the genome. Trends Ecol. Evol., Special issue: Invasions, guest edited by Michael E. Hochberg and Nicholas J. Gotelli 20, 229–237. https://doi.org/10.1016/j.tree.2005.02.010

Marçais, G., Delcher, A.L., Phillippy, A.M., Coston, R., Salzberg, S.L., Zimin, A., 2018. MUMmer4: A fast and versatile genome alignment system. PLOS Comput. Biol. 14, e1005944. https://doi.org/10.1371/journal.pcbi.1005944

Masuda, H., Shioya, Y., Fukuhara, R., 1980. Robertsonian translocation in Japanese Black cattle. Jpn. J. Zootech. Sci. 51, 26–32.

Meng, L., Li, H., Zhang, L., Wang, J., 2015. QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. Crop J., Special Issue: Breeding to Optimize Agriculture in a Changing World 3, 269–283. https://doi.org/10.1016/j.cj.2015.01.001

Miller, M.R., Dunham, J.P., Amores, A., Cresko, W.A., Johnson, E.A., 2007. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. Genome Res. 17, 240–248. https://doi.org/10.1101/gr.5681207

Moore, B.C., Taylor, C.E., 1986. Drosophila of Southern California: III. Gene arrangements of Drosophila persimilis. J. Hered. 77, 313–323. https://doi.org/10.1093/oxfordjournals.jhered.a110248

Navarro, A., Barton, N.H., 2003. Chromosomal Speciation and Molecular Divergence--Accelerated Evolution in Rearranged Chromosomes. Science 300, 321–324. https://doi.org/10.1126/science.1080600

Nickerson, E., Nelson, D.L., 1998. Molecular Definition of Pericentric Inversion Breakpoints Occurring during the Evolution of Humans and Chimpanzees. Genomics 50, 368–372. https://doi.org/10.1006/geno.1998.5332

Niebuhr, E., 1974. Down's syndrome. Humangenetik 21, 99–101. https://doi.org/10.1007/BF00278575

Nolte, A. W, Freyhof, J., Stemshorn, K.C, Tautz, D., 2005. An invasive lineage of sculpins, Cottus sp. (Pisces, Teleostei) in the Rhine with new habitat adaptations has originated from hybridization between old phylogeographic groups. Proc. R. Soc. B Biol. Sci. 272, 2379–2387. https://doi.org/10.1098/rspb.2005.3231

Noor, M. A. F., Grams, K.L., Bertucci, L.A., Almendarez, Y., Reiland, J., Smith, K.R., 2001. The Genetics of Reproductive Isolation and the Potential for Gene Exchange Between Drosophila Pseudoobscura and D. Persimilis Via Backcross Hybrid Males. Evolution 55, 512–521. https://doi.org/10.1111/j.0014-3820.2001.tb00785.x

Noor, M. A. F., Grams, K.L., Bertucci, L.A., Reiland, J., 2001. Chromosomal inversions and the reproductive isolation of species. Proc. Natl. Acad. Sci. 98, 12084–12088. https://doi.org/10.1073/pnas.221274498

Nunez, J.C.B., Seale, T.P., Fraser, M.A., Burton, T.L., Fortson, T.N., Hoover, D., Travis, J., Oleksiak, M.F., Crawford, D.L., 2015. Population Genomics of the Euryhaline Teleost Poecilia latipinna. PLOS ONE 10, e0137077. https://doi.org/10.1371/journal.pone.0137077

Ooijen, J.W.V., 2011. Multipoint maximum likelihood mapping in a full-sib family of an outbreeding species. Genet. Res. 93, 343–349. https://doi.org/10.1017/S0016672311000279

Pedrosa, A., Vallejos, C., Bachmair, A., Schweizer, D., 2003. Integration of common bean (Phaseolus vulgaris L.) linkage and chromosomal maps. Theor. Appl. Genet. 106, 205–212. https://doi.org/10.1007/s00122-002-1138-3

Pialek, J., Hauffe, H.C., Searle, J.B., 2005. Chromosomal variation in the house mouse. Biol. J. Linn. Soc. 84, 535–563. https://doi.org/10.1111/j.1095-8312.2005.00454.x

Pontius, J.U., Mullikin, J.C., Smith, D.R., Agencourt Sequencing Team, Lindblad-Toh, K., Gnerre, S., Clamp, M., Chang, J., Stephens, R., Neelam, B., Volfovsky, N., Schaffer, A.A., Agarwala, R., Narfstrom, K., Murphy, W.J., Giger, U., Roca, A.L., Antunes, A., Menotti-Raymond, M., Yuhki, N., Pecon-Slattery, J., Johnson, W.E., Bourque, G., Tesler, G., NISC Comparative Sequencing Program, O'Brien, S.J., 2007. Initial sequence and comparative analysis of the cat genome. Genome Res. 17, 1675–1689. https://doi.org/10.1101/gr.6380007

Pootakham, W., Ruang-Areerate, P., Jomchai, N., Sonthirod, C., Sangsrakru, D., Yoocha, T., Theerawattanasuk, K., Nirapathpongporn, K., Romruensukharom, P., Tragoonrung, S., Tangphatsornruang, S., 2015. Construction of a high-density integrated genetic linkage map of rubber tree (Hevea brasiliensis) using genotyping-by-sequencing (GBS). Front. Plant Sci. 6. https://doi.org/10.3389/fpls.2015.00367

Potter, S., Bragg, J.G., Blom, M.P.K., Deakin, J.E., Kirkpatrick, M., Eldridge, M.D.B., Moritz, C., 2017. Chromosomal Speciation in the Genomics Era: Disentangling Phylogenetic Evolution of Rock-wallabies. Front. Genet. 8. https://doi.org/10.3389/fgene.2017.00010

Potter, S., Moritz, C., Eldridge, M.D.B., 2015. Gene flow despite complex Robertsonian fusions among rock-wallaby ( *Petrogale* ) species. Biol. Lett. 11, 20150731. https://doi.org/10.1098/rsbl.2015.0731

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of Population Structure Using Multilocus Genotype Data. Genetics 155, 945–959.

Qumsiyeh, M.B., Coate, J.L., Peppers, J.A., Kennedy, P.K., Kennedy, M.L., 1997. Robertsonian chromosomal rearrangements in the short-tailed shrew, Blarina carolinensis, in western Tennessee. Cytogenet. Genome Res. 76, 153–158. https://doi.org/10.1159/000134534

Ramsay, L., Macaulay, M., Ivanissevich, S. degli, MacLean, K., Cardle, L., Fuller, J., Edwards, K.J., Tuvesson, S., Morgante, M., Massari, A., Maestri, E., Marmiroli, N., Sjakste, T., Ganal, M., Powell, W., Waugh, R., 2000. A Simple Sequence Repeat-Based Linkage Map of Barley. Genetics 156, 1997–2005.

Reid, N.M., Jackson, C.E., Gilbert, D., Minx, P., Montague, M.J., Hampton, T.H., Helfrich, L.W., King, B.L., Nacci, D.E., Aluru, N., Karchner, S.I., Colbourne, J.K., Hahn, M.E., Shaw, J.R., Oleksiak, M.F., Crawford, D.L., Warren, W.C., Whitehead, A., 2017. The Landscape of Extreme Genomic Variation in the Highly Adaptable Atlantic Killifish. Genome Biol. Evol. 9, 659–676. https://doi.org/10.1093/gbe/evx023

Rieseberg, L.H., 2001. Chromosomal rearrangements and speciation. Trends Ecol. Evol. 16, 351–358. https://doi.org/10.1016/S0169-5347(01)02187-5

Robinson, W.P., Bernasconi, F., Basaran, S., Haluza, R., Farah, L.M.S., LUWeci, G., Schinzel, A.A., n.d. A Somatic Origin of Homologous Robertsonian Translocations and Isochromosomes 13.

Rodgers, R., Roach, J.L., Reid, N.M., Whitehead, A., Duvernell, D.D., 2018. Phylogenomic analysis of Fundulidae (Teleostei: Cyprinodotiformes) using RNA-sequencing data. Mol. Phylogenet. Evol. 121, 150–157. https://doi.org/10.1016/j.ympev.2017.12.030

Rogatcheva, M.B., Ono, T., Sonta, S., Oda, S., Borodin, P.M., 2000. Robertsonian metacentrics of the house musk shrew (Suncus murinus, Insectivora, Soricidae) lose the telomeric sequences in the centromeric area. Genes Genet. Syst. 75, 155–158. https://doi.org/10.1266/ggs.75.155

Rogers, J., Garcia, R., Shelledy, W., Kaplan, J., Arya, A., Johnson, Z., Bergstrom, M., Novakowski, L., Nair, P., Vinson, A., Newman, D., Heckman, G., Cameron, J., 2006. An initial genetic linkage map of the rhesus macaque (Macaca mulatta) genome using human microsatellite loci. Genomics 87, 30–38. https://doi.org/10.1016/j.ygeno.2005.10.004

Schaefer, J., Duvernell, D., Campbell, D.C., 2016. Hybridization and introgression in two ecologically dissimilar *Fundulus* hybrid zones: HYBRIDIZATION IN DISSIMILAR HYBRID ZONES. Evolution 70, 1051–1063. https://doi.org/10.1111/evo.12920

Schaefer, J., Kreiser, B., Flanagan, S., 2018. Population genomics of Fundulus grandis exposed to oil from Deepwater Horizon. J. Exp. Mar. Biol. Ecol. 509, 82–90. https://doi.org/10.1016/j.jembe.2018.09.003

Schaefer, J.F., Duvernell, D.D., Kreiser, B.R., 2011. Ecological and genetic assessment of spatial structure among replicate contact zones between two topminnow species. Evol. Ecol. 25, 1145–1161. https://doi.org/10.1007/s10682-011-9461-2

Schmutz, S.M., Moker, J.S., Barth, A.D., Mapletoft, R.J., 1991. Embryonic loss in superovulated cattle caused by the 1;29 Robertsonian translocation. Theriogenology 35, 705–714. https://doi.org/10.1016/0093-691X(91)90411-6

Schumer, M., Cui, R., Boussau, B., Walter, R., Rosenthal, G., Andolfatto, P., 2013. An Evaluation of the Hybrid Speciation Hypothesis for Xiphophorus Clemenciae Based on Whole Genome Sequences. Evolution 67, 1155–1168. https://doi.org/10.1111/evo.12009

Setzer, P.Y., 1970. An Analysis of a Natural Hybrid Swarm by Means Chromosome Morphology. Trans. Am. Fish. Soc. 99, 139–146. https://doi.org/10.1577/1548-8659(1970)99<139:AAOANH>2.0.CO;2

Seutin, G., White, B.N., Boag, P.T., 1991. Preservation of avian blood and tissue samples for DNA analyses. Can. J. Zool. 69, 82–90. https://doi.org/10.1139/z91-013

Shibusawa, M., Nishida-Umehara, C., Masabanda, J., Griffin, D.K., Isobe, T., Matsuda, Y., 2002. Chromosome rearrangements between chicken and guinea fowl defined by comparative chromosome painting and FISH mapping of DNA clones. Cytogenet. Genome Res. 98, 225–230. https://doi.org/10.1159/000069813

Spira, J., Wiener, F., Ohno, S., Klein, G., 1979. Is trisomy cause or consequence of murine T cell leukemia development? Studies on Robertsonian translocation mice. Proc. Natl. Acad. Sci. 76, 6619–6621. https://doi.org/10.1073/pnas.76.12.6619

Swenson, N.G., Howard, D.J., 2005. Clustering of Contact Zones, Hybrid Zones, and Phylogeographic Breaks in North America. Am. Nat. 166, 581–591. https://doi.org/10.1086/491688

Szamalek, J.M., Cooper, D.N., Schempp, W., Minich, P., Kohn, M., Hoegel, J., Goidts, V., Hameister, H., Kehrer-Sawatzki, H., 2006. Polymorphic micro-inversions contribute to the genomic variability of humans and chimpanzees. Hum. Genet. 119, 103–112. https://doi.org/10.1007/s00439-005-0117-6

Talbot, W.S., Egan, E.S., Gates, M.A., Walker, C., Ullmann, B., Neuhauss, S.C.F., Kimmel, C.B., Postlethwait, J.H., 1998. Genetic Analysis of Chromosomal Rearrangements in the cyclops Region of the Zebrafish Genome. Genetics 148, 373–380.

Thirman, M.J., Gill, H.J., Burnett, R.C., Mbangkollo, D., McCabe, N.R., Kobayashi, H., Ziemin-van der Poel, S., Kaneko, Y., Morgan, R., Sandberg, A.A., Chaganti, R.S.K., Larson, R.A., Le Beau, M.M., Diaz, M.O., Rowley, J.D., 1993. Rearrangement of the MLL Gene in Acute Lymphoblastic and Acute Myeloid Leukemias with 11q23 Chromosomal Translocations. N. Engl. J. Med. 329, 909–914. https://doi.org/10.1056/NEJM199309233291302

Thomas, P.E., Klinger, R., Furlong, L.I., Hofmann-Apitius, M., Friedrich, C.M., 2011. Challenges in the association of human single nucleotide polymorphism mentions with unique database identifiers. BMC Bioinformatics 12, S4. https://doi.org/10.1186/1471-2105-12-S4-S4

Tripathi, N., Hoffmann, M., Willing, E.M., Lanz, C., Weigel, D., Dreyer, C., 2009. Genetic linkage map of the guppy, *Poecilia reticulata*, and quantitative trait loci analysis of male size and colour variation. Proc. R. Soc. B Biol. Sci. 276, 2195–2208. https://doi.org/10.1098/rspb.2008.1930

Turner, T.L., Hahn, M.W., Nuzhdin, S.V., 2005. Genomic Islands of Speciation in Anopheles gambiae. PLoS Biol. 3, e285. https://doi.org/10.1371/journal.pbio.0030285

Twyford, A.D., Ennos, R.A., 2012. Next-generation hybridization and introgression. Heredity 108, 179–189. https://doi.org/10.1038/hdy.2011.68

Vamosi, J.C., Wilson, J.R.U., 2008. Nonrandom extinction leads to elevated loss of angiosperm evolutionary history. Ecol. Lett. 11, 1047–1053. https://doi.org/10.1111/j.1461-0248.2008.01215.x

Van-Ooijen, J.W., 2006. JoinMap 4: Software for the calculation of genetic linkage maps in experimental populations of diploid species. Kyazma BV Wagening. Neth.

Vigueira, P.A., Schaefer, J.F., Duvernell, D.D., Kreiser, B.R., 2008. Tests of reproductive isolation among species in the Fundulus notatus (Cyprinodontiformes: Fundulidae) species complex. Evol. Ecol. 22, 55–70. https://doi.org/10.1007/s10682-007-9158-8

Waits, E.R., Martinson, J., Rinner, B., Morris, S., Proestou, D., Champlin, D., Nacci, D., 2016. Genetic Linkage Map and Comparative Genome Analysis for the Atlantic Killifish (<i>Fundulus heteroclitus</i>). Open J. Genet. 06, 28–38. https://doi.org/10.4236/ojgen.2016.61004

Wang, J.C., Passage, M.B., Yen, P.H., Shapiro, L.J., Mohandas, T.K., 1991. Uniparental heterodisomy for chromosome 14 in a phenotypically abnormal familial balanced 13/14 Robertsonian translocation carrier. Am. J. Hum. Genet. 48, 1069–1074.

Wang, M.-H., Marinotti, O., Vardo-Zalik, A., Boparai, R., Yan, G., 2011. Genome-Wide Transcriptional Analysis of Genes Associated with Acute Desiccation Stress in Anopheles gambiae. PLOS ONE 6, e26011. https://doi.org/10.1371/journal.pone.0026011

White, M.J., 1978. Modes of speciation Freeman, San Francisco. Rev Suisse Zool. 61, 173–176.

White, M.J.D., 1974. Speciation in the Australian Morabine Grasshoppers — the cytogenetic evidence, in: White, M.J.D. (Ed.), Genetic Mechanisms of Speciation in Insects: Symposia Held at the XIVth International Congress of Entomology, Canberra, Australia August 22–30, 1972, Sponsored by the Australian Academy of Science and the Australian Entomological Society. Springer Netherlands, Dordrecht, pp. 57–68. https://doi.org/10.1007/978-94-010-2248-4_4

Wojcik, J.M., Searle, J.B., 1988. The chromosome complement of Sorex granarius –the ancestral karyotype of the common shrew ( Sorex araneus )? Heredity 61, 225. https://doi.org/10.1038/hdy.1988.109

Wu, C.-I., 2001. The genic view of the process of speciation. J. Evol. Biol. 14, 851–865. https://doi.org/10.1046/j.1420-9101.2001.00335.x

Yang, F., Carter, N.P., Shi, L., Ferguson-Smith, M.A., 1995. A comparative study of karyotypes of muntjacs by chromosome painting. Chromosoma 103, 642–652. https://doi.org/10.1007/BF00357691

# VITA

Naznin Sultana Remex was born and grew up in Naogaon, Bangladesh. She started her bachelor's in Zoology at the University of Dhaka, Bangladesh in 2012 and graduated in April 2016. She first came to the United States in September 2016 to spend time with her husband at Western Illinois University (WIU). While staying in Macomb, Illinois, she began to help her husband in his master's thesis project in a genetics lab. Her interest to pursue a master's degree in genetics and genomics started from Dr. Musser's lab at WIU and she joined Duvernell lab at Missouri University of Science and Technology (Missouri S&T) as a graduate student in fall 2017. During her time at Missouri S&T, Naznin has had the opportunity to work as both Graduate Teaching Assistant as well as Graduate Research Assistant. Besides her master's thesis project, she also helped the lab in some other projects and became a co-author in one of the papers published from Duvernell lab. Naznin received her master's degree in Applied and Environmental Biology from Missouri S&T in July 2019. After graduating, she started her Ph.D. in Molecular and Cellular Physiology at Louisiana State University Health Science Center (LSUHSC) in Shreveport, Louisiana.