IMPLICATION OF ESTROGEN RESPONSE ELEMENTS IN EXPRESSION OF SECONDARY SEX-TRAITS IN THE SEX-ROLE REVERSED GULF PIPEFISH, *Syngnathus scovelli*

A Dissertation

by

ANDREW PATRICK ANDERSON

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Adam Jones
Co-Chair of Committee,	Heath Blackmon
Committee Members,	Jessica E. Light
	Gil Rosenthal
	Spencer Behmer
Head of Department,	Thomas McKnight

August 2019

Major Subject: Biology

ABSTRACT

Intralocus conflict, the differential selection optima for a genomic region, between males and females created by sexual selection can be resolved with the presence of sex-biased hormone response elements (HREs) in the genome which allow for gene regulation through sex-biased hormones. In general, *cis*-regulatory elements are found more frequently and in closer proximity to hormonally responsive genes. I predict that genes putatively under sexual selection are more likely to have a greater number of proximal sex-biased HREs than randomly selected genes. To investigate this I chose to use estrogen response elements (EREs) and the sex-role reversed Gulf pipefish, *Syngnathus scovelli*. I demonstrate the secondary sex traits, transverse bands, and body depth described in female *S. scovelli* do not have any confounding effects of age and confirm those traits do affect male mate choice. To scan the genome of *S. scovelli* for EREs, I developed and tested an algorithm for identifying putative estrogen binding regions. With confirmation of the secondary sex traits putatively under sexual selection, I used feminized males with female traits to elucidate genes that are involved in production of body depth and ornamentation (i.e., transverse bands). I was able to show that these genes have an excess of EREs compared to typically sex-differentiated genes thereby demonstrating the important role EREs can play in sexual selection.

DEDICATION

To my two children, Kaedence “KC” Carolina and Ignacio “Iggy” Wren Anderson.

ACKNOWLEDGEMENTS

I would like to thank my committee co-chair, Dr. Jones, for all of his guidance and investment in my scientific career, as well as my committee members, Dr. Light, Dr. Blackmon, Dr. Rosenthal, and Dr. Behmer for their input and support.

I would also like to give a huge thank you to my wife, Geysel Anderson, for her unconditional love and all that she has done to help me achieve my goals. In addition, I recognize my family and friends for just being there to listen or to provide distractions.

Lastly, I would like to thank my former lab mates, Dr. Rose and Dr. Flanagan, for their help in design and direction of my work here at Texas A&M.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a thesis committee consisting of Professors Jones, Blackmon, and Rosenthal of the Department of Biology, Professor Light of the Department of Wildlife and Fisheries Sciences, and Professor Behmer of the Department of Entomology.

The C++ code for EREFinder was developed by Professor Adam Jones. The analyses depicted in Chapter 3 were conducted in part by Dr. Sarah Flanagan of the School of Biological Sciences at the University of Canterbury.

All other work conducted for the thesis was completed by the student independently.

Funding Sources

Graduate study was supported by a fellowship from Texas A&M University.

This work was also made possible in part by the National Science Foundation under Grant Number DEB-1119261 awarded to Professor Adam Jones. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NSF.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

One possible consequence of sexual selection is different trait optima for males and females leading to intra- and inter-locus conflict within the genome (Parker 1979, Chapman et al. 2003, Connallon and Clark 2014). A potential resolution to intralocus conflict is the evolution of sex-biased hormonal control of the affected gene, leading to differential expression of that gene between the sexes (Connallon and Knowles 2005, Mank 2017, Wright et al. 2018). Sex-biased hormones (estrogens and androgens) occur in different circulating concentrations in males and females and have been shown to mediate sex-biased traits (Partridge et al. 2010, Bartos et al. 2012, Gonçalves et al. 2014, Lindsay et al. 2016). As such, genomic regions under sexual selection should be more sensitive to differing hormone levels and have a way to be activated in the presence of sex-biased hormones. Hormone response elements (HREs) represent the physical location on the genome where hormone-activated nuclear receptors bind to initiate or suppress transcription of the target gene. An expectation of sexual selection and intralocus conflict is the presence of greater number of HREs in genomic regions putatively affected by sexual selection.

While I could investigate systems that have conventional sex-roles to determine traits under selection and sex-biased hormone mediation, there are some disadvantages to using the male-biased hormone androgen. Androgen receptors bind to androgen response elements (AREs), but other receptors (progesterone, glucocorticoid, and mineralcorticoid) competitively bind to the response element as well (Nelson et al.

1999). Further there are many variations of the ARE that create additional uncertainty in identifying regions when scanning genomes. By choosing a sex-role reversed system I can instead look for estrogen response elements (EREs) as they do not have any other hormones that competitively bind to their response element. In addition to the clear canonical sequence of binding (Klein-Hitpass et al. 1988, Boyer et al. 2000), EREs are practical regions to investigate the effects of sexual selection on the sex-biased *cis*-regulatory elements.

To confirm that genomic regions under sexual selection in sex-role reversed species have an excess of EREs, I need a species with sex-role reversal that has a secondary sex-trait that is estrogen-mediated. Syngnathid fishes (e.g., seahorses, pipefishes, and seadragons) are an excellent group for investigations into the role of estrogens and EREs on secondary sex traits in sex-role reversed species. In particular, the Gulf pipefish *Syngnathus scovelli* is an ideal species because it has an annotated genome (Small et al. 2016), is known to have females under greater sexual selection (Jones and Avise 2001, Paczolt and Jones 2010, Flanagan and Jones 2017), and it has secondary sex traits that are estrogen-mediated (Partridge et al. 2010). Females of *S. scovelli* have distinctive transverse bands and a deeper body than males. These features are under sexual selection and are subject to male mate choice (Jones and Avise 1997, 2001, Flanagan et al. 2014). These features are also believed to be influenced by estrogens because males exposed to 17α-ethinylestradiol (EE2) develop a deeper abdomen and a banding pattern similar to females that is not present in unexposed males (Partridge et al. 2010). Determining the genes responsible for these traits would allow

me to investigate if the genomic coordinates of these genes have an over-representation of EREs.

Before I can be certain of the effect of sexual selection on bands and body depth in female *S. scovelli*, I need to investigate the role age might be playing in the expression of those traits. Secondary sex traits might change with age, thereby altering the selective forces on the trait (Lande and Arnold 1983). To ensure a proper description of selective pressures, verifying the effect of age is vital to understanding the selective pressures on an the traits of an organism (Arnold and Wade 1984). All of the studies on *S. scovelli* have used wild-caught specimens with unknown age. While males do prefer females with larger bands, this trait correlates with female size (Flanagan et al. 2014), a feature that also under sexual selection. Pipefish have indeterminate growth (Takahashi et al. 2003) and it is very likely that size correlates with age. It is necessary to confirm the effect of age on the secondary sex trait of banding and its effect on male mate choice.

To investigate the number and positions of EREs in a genome, I need a search algorithm to locate these regions. While many programs exist that can use a search algorithm based on position weight matrices (Frith et al. 2003, Kel et al. 2003, Tan and Lenhard 2016), these may not be appropriate for ERE searches (Man and Stormo 2001, Bulyk et al. 2002, Omidi et al. 2017). Since EREs are composed of two palindromic half-sites (Klein-Hitpass et al. 1988, Boyer et al. 2000), a loss of a perfect half-site has a greater effect on estrogen receptor binding than a point mutation (Tyulmenkov and Klinge 2001, Deegan et al. 2011). Using empirical evidence of sequence changes on binding affinity, Tyulmenkov and Klinge (2001) were able to develop an equation that

represented the binding affinity based on perfect half-site presence and base pair substitutions in EREs. Using this formula, I set out to make a search algorithm to scan genomes for regions of high estrogen-receptor binding. This algorithm uses a sliding window approach and reports all windows allowing the user to determine the best way to interpret and handle the data.

My thesis is composed of three parts, all leading to investigation of the role EREs play in sexually selected sex traits. The first chapter will determine if the secondary sex traits of bands and body depth in female *S. scovelli* have an age effect. The second chapter will establish the efficacy of the algorithm I developed for searching genomes for regions of high estrogen receptor binding. The third chapter will use a transcriptomic approach to look for genes associated with the estrogen-influenced sex traits and determine if those genes have an excess of EREs in those genomic regions.

Literature Cited

Arnold, S. J., and M. J. Wade. 1984. On the measurement of natural and sexual selection: applications. Evolution **38**:720-734.

Bartos, L., G. A. Bubenik, and E. Kuzmova. 2012. Endocrine relationships between rank-related behavior and antler growth in deer. Frontiers in Bioscience **4**:1111-1126.

Boyer, M., N. Poujol, E. Margeat, and C. A. Royer. 2000. Quantitative characterization of the interaction between purified human estrogen receptor α and DNA using fluorescence anisotropy. Nucleic acids research **28**:2494-2502.

Bulyk, M. L., P. L. Johnson, and G. M. Church. 2002. Nucleotides of transcription factor binding sites exert interdependent effects on the binding affinities of transcription factors. Nucleic acids research **30**:1255-1261.

Chapman, T., G. Arnqvist, J. Bangham, and L. Rowe. 2003. Sexual conflict. Trends in Ecology & Evolution **18**:41-47.

Connallon, T., and A. G. Clark. 2014. Evolutionary inevitability of sexual antagonism. Proceedings of the Royal Society of London B: Biological Sciences **281**:20132123.

Connallon, T., and L. L. Knowles. 2005. Intergenomic conflict revealed by patterns of sex-biased gene expression. Trends in Genetics **21**:495-499.

Deegan, B. J., A. M. Bona, V. Bhat, D. C. Mikles, C. B. McDonald, K. L. Seldeen, and A. Farooq. 2011. Structural and thermodynamic consequences of the replacement of zinc with environmental metals on estrogen receptor α–DNA interactions. Journal of Molecular Recognition **24**:1007-1017.

Flanagan, S. P., J. B. Johnson, E. Rose, and A. G. Jones. 2014. Sexual selection on female ornaments in the sex-role-reversed Gulf pipefish (Syngnathus scovelli). J Evol Biol **27**:2457-2467.

Flanagan, S. P., and A. G. Jones. 2017. Genome-wide selection components analysis in a fish with male pregnancy. Evolution **71**:1096-1105.

Frith, M. C., M. C. Li, and Z. Weng. 2003. Cluster-Buster: Finding dense clusters of motifs in DNA sequences. Nucleic acids research **31**:3666-3668.

Gonçalves, D., S. S. Costa, M. C. Teles, H. Silva, M. Inglês, and R. F. Oliveira. 2014. Oestradiol and prostaglandin F2α regulate sexual displays in females of a sex-role reversed fish. Proceedings of the Royal Society B: Biological Sciences **281**:20133070.

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

Jones, A. G., and J. C. Avise. 2001. Mating Systems and Sexual Selection in Male-Pregnant Pipefishes and Seahorses: Insights from Microsatellite-Based Studies of Maternity. Journal of Heredity **92**:150-158.

Kel, A. E., E. Gossling, I. Reuter, E. Cheremushkin, O. V. Kel-Margoulis, and E. Wingender. 2003. MATCHTM: a tool for searching transcription factor binding sites in DNA sequences. Nucleic acids research **31**:3576-3579.

Klein-Hitpass, L., G. U. Ryffel, E. Heitlinger, and A. C. Cato. 1988. A 13 bp palindrome is a functional estrogen responsive element and interacts specifically with estrogen receptor. Nucleic acids research **16**:647-663.

Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. Evolution **37**:1210-1226.

Lindsay, W. R., D. G. Barron, M. S. Webster, and H. Schwabl. 2016. Testosterone activates sexual dimorphism including male-typical carotenoid but not melanin plumage pigmentation in a female bird. The Journal of Experimental Biology **219**:3091-3099.

Man, T.-K., and G. D. Stormo. 2001. Non-independence of Mnt repressor–operator interaction determined by a new quantitative multiple fluorescence relative affinity (QuMFRA) assay. Nucleic acids research **29**:2471-2478.

Mank, J. E. 2017. The transcriptional architecture of phenotypic dimorphism. Nature Ecology & Evolution **1**:0006.

Nelson, C. C., S. C. Hendy, R. J. Shukin, H. Cheng, N. Bruchovsky, B. F. Koop, and P. S. Rennie. 1999. Determinants of DNA sequence specificity of the androgen, progesterone,

and glucocorticoid receptors: evidence for differential steroid receptor response elements. Molecular endocrinology **13**:2090-2107.

Omidi, S., M. Zavolan, M. Pachkov, J. Breda, S. Berger, and E. van Nimwegen. 2017. Automated incorporation of pairwise dependency in transcription factor binding site prediction using dinucleotide weight tensors. PLoS computational biology **13**:e1005176.

Paczolt, K. A., and A. G. Jones. 2010. Post-copulatory sexual selection and sexual conflict in the evolution of male pregnancy. Nature **464**:401-404.

Parker, G. 1979. Sexual selection and sexual conflict. Sexual selection and reproductive competition in insects **123**:166.

Partridge, C., A. Boettcher, and A. G. Jones. 2010. Short-term exposure to a synthetic estrogen disrupts mating dynamics in a pipefish. Horm Behav **58**:800-807.

Small, C., S. Bassham, J. Catchen, A. Amores, A. Fuiten, R. Brown, A. Jones, and W. Cresko. 2016. The genome of the Gulf pipefish enables understanding of evolutionary innovations. Genome biology **17**:258.

Takahashi, E., R. M. Connolly, and S. Y. Lee. 2003. Growth and Reproduction of Double-Ended Pipefish, Syngnathoides biaculeatus, in Moreton Bay, Queensland, Australia. Environmental Biology of Fishes **67**:23-33.

Tan, G., and B. Lenhard. 2016. TFBSTools: an R/bioconductor package for transcription factor binding site analysis. Bioinformatics **32**:1555-1556.

Tyulmenkov, V. V., and C. M. Klinge. 2001. A mathematical approach to predict the affinity of estrogen receptors α and β binding to DNA. Molecular and cellular endocrinology **182**:109-119.

Wright, A. E., M. Fumagalli, C. R. Cooney, N. I. Bloch, F. G. Vieira, S. D. Buechel, N. Kolm, and J. E. Mank. 2018. Male-biased gene expression resolves sexual conflict through the evolution of sex-specific genetic architecture. Evolution Letters **2**:52-61.

CHAPTER II

CHOOSY GULF PIPEFISH MALES IGNORE AGE BUT PREFER ACTIVE FEMALES WITH DEEPLY KEELED BODIES

Synopsis

Within Syngnathidae fishes (e.g., pipefish, seadragons and seahorses), male pregnancy often results in choosy males and competitive females. Females in these species often evolve secondary sexual traits and engage in courtship displays that make their ornaments more noticeable to males. Most syngnathids probably continue to grow larger throughout their lives, but we know little about the relationship between age and mating competition in these taxa. Here, I use the Gulf pipefish, *Syngnathus scovelli*, to investigate the roles of ornament size, courtship activity level, age, and fecundity in female mating competition. I conduct male choice trials that allow males to choose between similar sized females of different ages. I also measure age and size at maturity. My results show that females with larger ornaments are deeper bodied and engage in longer courtship displays, yet males choose females based on depth and display rather than ornamentation. This result suggests that ornamentation serves to help males assess female quality. Age plays no role in male choice or ornamentation. The finding that males care more about female phenotype than female age considerably simplifies the interpretation of mating patterns in natural populations of Gulf pipefish, which are characterized by considerable age structure.

Introduction

Sexual selection can act on multiple traits and these traits are often correlated with one another (Lande and Arnold 1983). One trait of special significance is age, as many phenotypic traits change over the course of an individual's lifespan. Because many phenotypes are age-related, disentangling the effects of secondary sexual traits from age can be challenging (Arnold and Wade 1984). Examples of traits that are strongly age dependent in some species include body size (McCann 1981, Madsen et al. 1993, Wikelski and Trillmich 1997) and weapon or ornament size (Kruuk et al. 2002, Vanhooydonck et al. 2005). Not only can traits change with age, but selection can also act on different traits at different ages (Freeman-Gallant et al. 2010). A further complication can arise when males engage in alternative strategies across different times in their lives to optimize lifetime fitness (Andersson 1994, Henson and Warner 1997, Engqvist and Taborsky 2016). When describing selection in a mixed-age population, cross-sectional studies cannot disentangle age effects while longitudinal studies can (Arnold and Wade 1984). However, longitudinal studies are logistically impossible for most taxa. For species in which it is unfeasible to follow lifetime success of particular individuals, an alternative is to accurately measure the age of each individual, thereby separating the effects of age and the secondary sex trait (McElligott et al. 2001). Regardless, the effects of age structure clearly should not be ignored in studies of sexual selection.

Mate choice is a multivariate problem, in the sense that choosers integrate information from multiple sources to decide whether or not to mate (Rosenthal 2017),

and the entire suite of traits involved in mate attraction may be affected by age. For instance, ornaments and weapons may be subject to allometric growth as body size increases over time (Kodric-Brown et al. 2006). A wide range of non-allometric correlations between ornaments and other phenotypic traits can also occur (Marchetti 1998, Loyau et al. 2005, McGlothlin et al. 2005). For almost every morphological trait associated with sexual selection, the bearer usually has a behavioral repertoire that goes along with it and is also subject to selection. Typical behaviors involved in sexual selection include aggression during contests with rivals (Carothers 1984, Dixon and Cade 1986, Lindenfors and S.Tullberg 2011), as well as song and display during mate choice (Searcy and Andersson 1986, Huber 2005, Byers et al. 2010). Several outcomes are possible from the interactions between morphological and behavioral traits. Behavior could reinforce the physical trait such that an increase in one results in an increase in the other (Møller 1990, Mateos 1998, Loyau et al. 2005). Alternatively, a compensatory relationship might exist as individuals may vary investment in traits that contribute to overall selection (Moller and Pomiankowski 1993, Taff et al. 2012). Finally, there may be no relationship if selection independently acts on behavioral and physical traits (Soma and Garamszegi 2015).

Fishes of the family Syngnathidae (pipefishes, seahorses, and seadragons) provide excellent opportunities to investigate interactions among behavior, ornamentation, and age in sexual selection. Many members of this clade are sex-role reversed in the sense that sexual selection operates more strongly on females than on males (Berglund et al. 1986b, Vincent et al. 1992, Jones and Avise 1997). Females of

sex-role-reversed species experience different selection pressures compared to those of conventional taxa. For instance, sex-role-reversed females have to balance resource allocation in secondary sex traits against the production of energetically costly eggs (Fitzpatrick et al. 1995). Much like most other fish species, larger female syngnathids generally have greater egg production than smaller females, and larger females are preferred by males (Berglund et al. 1986a, Paczolt and Jones 2010). In species with female ornamentation, ornament size and body size are often positively correlated, and multiple studies show evidence of selection acting on both of these traits (Berglund et al. 1986a, Berglund et al. 1997, Flanagan et al. 2014). Additionally, growing evidence suggests that female ornaments may be honest signals of female quality with respect to reproductive output (Berglund et al. 1997, Cunha et al. 2017, Mobley et al. 2018). Thus, the most preferred females manage to make substantial investments in both ornaments and reproductive output. In terms of behavior, larger females engage in courtship activities for longer periods of time and courtship duration influences male choice (Partridge et al. 2013). With so many energetic demands on females in terms ornaments, behavior, and egg production, I might predict that faster growing females would have to sacrifice investment along at least one of these phenotypic axes compared to slower growing females. Indeed, such a tradeoff has been demonstrated in *Syngnathus typhle*, where non-preferred females sacrifice reproduction for growth (Berglund 1991).

Given that syngnathid males produce many broods across a protracted breeding season (Berglund et al. 1989, Paczolt and Jones 2010, Mobley et al. 2011b), most syngnathid populations should have substantial age structure. Syngnathids show

indeterminate growth (Takahashi et al. 2003) with growth rate diminishing somewhat at maturity but not stopping (Whatley 1969, Parkinson and Booth 2016). While species can show a life span between just a few months (Parkinson and Booth 2016) to a few years (Svensson 1988), populations show a wide range of individual sizes within sampling periods (Takahashi et al. 2003, Bolland and Boettcher 2005, Ripley and Foran 2006). Many species have clear spawning periods (Vincent et al. 1995, Hiddink and Jager 2002, Bolland and Boettcher 2005), but with continuous mating even a species with a one year lifespan would have substantial age structure during the spawning season. With overall size as a key target of sexual selection, the interplay between body size, other sexually selected traits, and age *per se* is an important unresolved facet of mating patterns in syngnathid fishes.

Using the Gulf pipefish (*Syngnathus scovelli*), I test the following hypotheses: 1) Males distinguish among females on the basis of chronological age, 2) Female courtship behavior is correlated with ornamentation and fecundity, and 3) Males choose females on the basis of size, ornamentation, display behavior, and apparent fecundity, possibly in a way that interacts with female age. I use a lab-reared population of Gulf pipefish (*Syngnathus scovelli*), with males and females of known age, to test male preferences in binary choice trials. To test the effects of age *per se*, I allow males to choose between females of similar size that differ in age. My results show no evidence that males differentiate among females on the basis of age. However, males do prefer females with greater body depth, an indicator of female fecundity, and females that were more active in courtship.

Methods

Study species

The Gulf pipefish is a widely distributed coastal species from the Gulf of Mexico and Western Atlantic. The species is sex-role reversed with strong selection on females (Jones et al. 2001, Rose et al. 2013) and has been used frequently to investigate various aspects of sexual selection in role-reversed species (Jones and Avise 2001, Paczolt and Jones 2010, Flanagan and Jones 2017). Females have a distinctive ornamentation of transverse iridescent bands that have been shown to be a target of selection (Jones and Avise 1997, 2001, Flanagan et al. 2014). Larger females have larger band size and number with males preferring females that have larger band area (Flanagan et al. 2014). Additionally, males prefer females that are more active during courtship (Partridge et al. 2013), but no link has yet been established between ornamentation and courtship intensity.

Collection and rearing

Adult *S. scovelli* were collected by seine netting in Port Aransas, TX (27.8807, -97.1025) between May and July 2014 and were transported back to holding facilities at Texas A&M University. All fish were held at 26C in 26ppt saltwater (Instant Ocean™) on a 12:12 day:night cycle for the duration of their care. Vision between tanks was impeded using opaque barriers. Tanks were aerated using air stones and had 30% water changes every other day. Mature fish were held at a density of no more than 1 fish per 3.5L, while juveniles were held at no more than 1 fish per 1.7L. Fish were fed enriched (Algamac 3000™) live brine shrimp three times daily.

Wild caught fish were mated between July and September 2014 generating 16 successful broods from 10 different females. No males were used twice. When possible, a maximum of 20 offspring were taken from each brood and reared with their siblings. I used the presence of a complete brood pouch as an indication of maturity in male fish, and presence of a keeled abdomen and iridescent bands as indications of maturity in female fish. Once a fish reached maturity, the date and size of the individual was noted, and it was moved to an adult tank, where it lived with its same-sex, adult siblings. Two to four weeks prior to experimentation, females were presented with wild males captured from the same population in the summer of 2015. Upon successful mating the female was confirmed mature and housed individually in a 9.5L tank to prevent dominance interactions with other females. Gulf pipefish females continuously produce eggs, so a single mating event followed by a period of a week or more to replenish eggs should not affect fecundity (Begovac and Wallace 1987, 1988). Males were also housed individually in 9.5L tanks for two to four weeks prior to the experiment, but were not allowed to mate before the onset of the experiment. Visual inspection of the pouch reliably indicates whether or not a male is ready to mate.

Preference Trials

From the surviving offspring of the 16 broods, I selected 20 females and 20 males to use in the experiments. For each replicate, I matched females with respect to body length (difference of ~0.5 cm or less), and chose females that differed in age. I also ensured that no individuals within a replicate were related. Replicates were assigned three days prior to the beginning of the experiment. These criteria meant that female

pairs consisted of one older and one younger female (average difference in age: 32.6 ± 10.0 days) that were size-matched (average difference in standard length: 0.41± 0.15 cm).

Experimental tanks were identical to those used by Partridge et al. (2013), with stimulus younger and older females and focal male placed as shown in Figure 1. Each tank was divided into three chambers by first dividing it in half lengthwise with a transparent barrier. One of the resulting long chambers was then divided in half by an opaque barrier. The focal male was placed in the largest chamber, which allowed him to view either of the smaller chambers through the transparent barrier. Females were placed in the smaller chambers, and they were not able to see each other (or sense each other in any way) due to the opaque barrier separating these chambers. This configuration has been used repeatedly in the literature for mate choice trials in syngnathids (Berglund and Rosenqvist 2001a, Partridge et al. 2010, Partridge et al. 2013). Immediately upon the start of the light cycle, the focal male and stimulus females were removed from their individual holding tanks, measured and photographed with a Canon EOS Rebel T5 (Canon USA, Inc. Melville, NY, USA), and placed in the experimental tank, where they were allowed to acclimate for ten minutes. After acclimation, behavior was recorded using either a Canon Vixia HFR10 (Canon USA, Inc. Melville, NY, USA) or Canon EOS Rebel T5 (Canon USA, Inc. Melville, NY, USA) for one hour. At the conclusion of the experiment, each female was paired with either the focal male or his sibling, which was individually housed and not allowed to participate in the experiment in any other way. Females were given two days to mate with the male with which they were

paired. Focal males were not reused. After the first ten replicates, each female pair was given one week of individual housing before being reused for the next ten replicates, ensuring that each female pair was used twice. Females within a pair swapped chambers between the first and second use to control for side bias. All fish were euthanized at the conclusion of the experiment using tricaine mesylate.

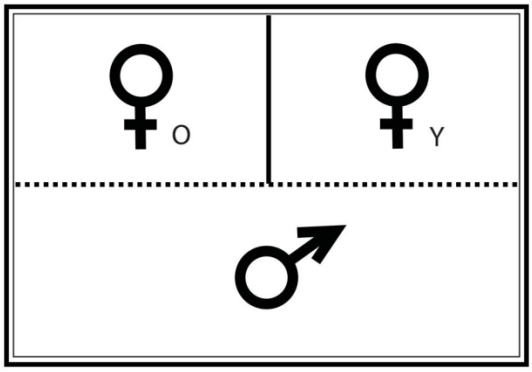

Figure 1: Cartoon of male choice between two females (O = Older , Y= Younger). Females had an opaque barrier between them and the male had a transparent barrier to females. Female age categories were alternated between left and right sides. Dimensions of the female sections were 25W x 28H x 12.5D cm each while male section was 50W x 28H x 12.5D cm. Design follows Partridge et al. (2013).

Measurement of phenotypes

Non-behavioral attributes of females used for this experiment consisted of age (days), standard length (SL; cm), depth (cm) and mean band area (mm^2). Ages of all fish were known since individuals were kept separate based on their family and by extension their birthdate. Standard length and depth were determined using the images

taken prior to the choice trial using ImageJ (Schneider et al. 2012). Depth was measured as the distance from the anterior point where the dorsal fin meets the body to the ventral keel of the abdomen making a line perpendicular to the body length (Figure 2). Mean band area followed that of Flanagan et al. (2014) where number of iridescent bands was counted and the total area of the bands was determined using tpsDig2.3.2 (Rohlf 2010). This value has been shown to play a role in premating sexual selection in *S. scovelli*.

Figure 2: Image of female *Syngnathus scovelli*. Physical measurements of the individual included standard length (tip of snout to base of tail), depth (dashed line), and area per band (one band is within box).

Behavioral responses of females and males were measured by hand-scoring videos of the choice experiment. Behavioral traits followed those designated in Partridge et al. (2013): male association time (time spent in front of a specific female's chamber), female color display (female with erect dorsal fin and darkened body to emphasize banding on body), female posing (vertical swimming in 'S'-shaped posture), and dancing (female posing in conjunction with male performing vertical swimming in close proximity). All behaviors were recorded as time spent engaged in activity divided by total time of recording to yield percentage of time spent in each behavior. Posing and

dancing behaviors were treated as exclusive to each other while color display was inclusive to all behaviors.

Statistical analysis

All data were analyzed in R. An exact binomial test was used to investigate sex bias in offspring survivorship. A Wilcoxon signed-rank test was used to determine if there was a difference between the sexes in size-at-maturity or age-at-maturity. Additionally a linear model for each sex was used to look for the effect of age-at-maturity on size-at-maturity.

I tested for correlations among female traits using the R package Hmisc (Harrel and Dupont 2014). To investigate what female traits were important to male mate choice, I used generalized linear models treating male association time as the response variable. I investigated which traits males used to distinguish between females and whether they preferred comparatively higher or lower values of those traits. To determine my explanatory variables for the models, the difference between age, SL, depth, mean band area, color display, posing, and dancing for the females on the right and left sides in a replicate were calculated. Both the response of male association time and the explanatory variables were calculated by arbitrarily using the female on the right side of the tank as the baseline female. For instance, for female depth, I subtracted body depth for the female on the left side of the tank from body depth of the female on the right side of the tank. This approach takes into account the paired design without introducing bias, as females were randomly assigned to a side of the tank. This approach also removes any bias introduced by males tending to move preferentially toward one

side of the tank. To confirm removal of side bias, I repeated my analysis by randomizing right and left side but still using each female once and found identical results. Data were transformed as needed and scaled. I used a stepwise optimization technique to find the optimal model starting with both conditions of no explanatory variables and all explanatory variables. If two starting points did not reach the same model, I utilized an ANOVA to find best model, choosing the model with the least variables when there was no difference between them.

Results

Maturation Data

Of the 292 offspring, 118 survived to maturity (68 males, 50 females). There was no sex bias in survivorship ($P = 0.1634$), assuming an equal sex ratio at birth. Fish took about seven months to reach maturity irrespective of sex (males: 214.5 ± 15.6 days, females: 210.3 ± 16.0 days, $P = 0.96$); however, females were longer at maturity than males (males: 9.59 ± 0.16 cm, females: 10.70 ± 1.92 cm, $P < 0.0001$). Age at maturation was not associated with length at maturation in females ($P = 0.4207$) but was weakly positively associated with length at maturation in males ($P = 0.0276$, $R^2 = 0.057$) (Figure 3).

Choice Experiment

Three of the 20 replicates were removed from the dataset due to the following occurrences: a female died one day following the experiment (cause unknown), a camera failed to record more than 15 minutes, and a male did not move for first 50 minutes of the experiment (all other males changed sides within first five minutes). Proportions of

time spent in various behaviors were transformed to reach normality as follows: color display: arcsine, posing: square root, dancing: cubic root. Because all three behaviors were strongly correlated, I performed a PCA on female behavior. The first principle component (henceforth "courtship") accounted for 65% of the variance and positively correlated with all three behaviors.

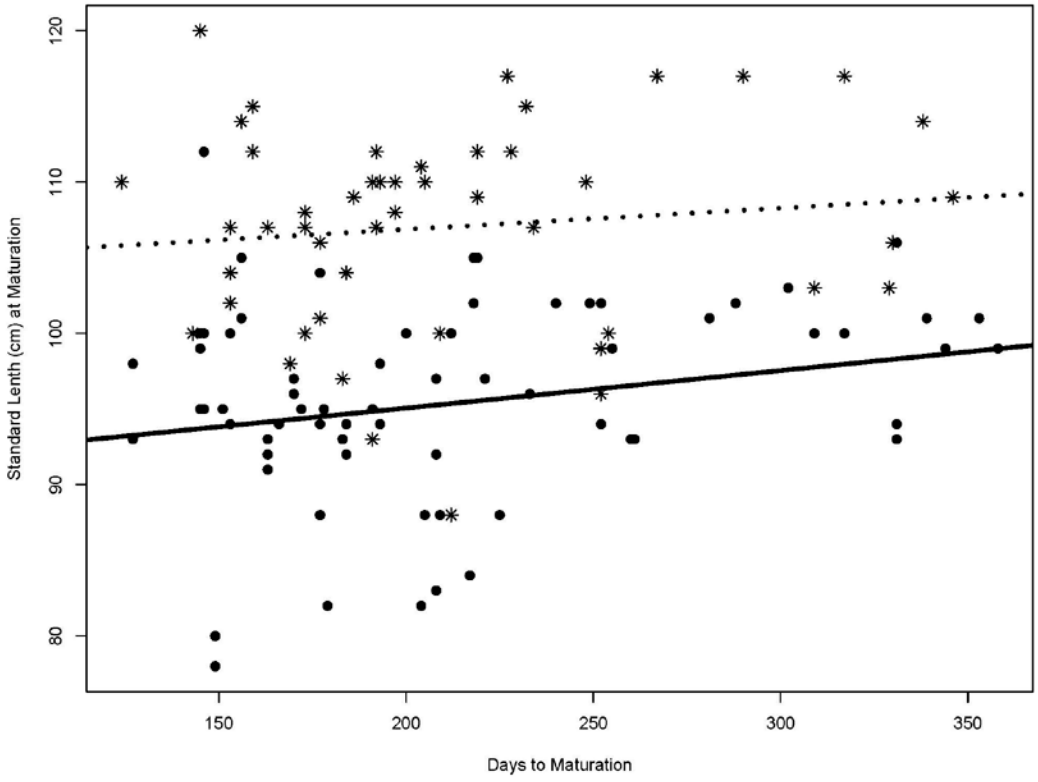

Figure 3: Scatter plot of age and length at maturation of *S. scovelli*. Dots and solid line represent male (R^2 = 0.057, $F_{1,66}$ = 5..076, *P* = 0.0276) individuals and asterisks and dashed line represent female (R^2 = 0.013, $F_{1,50}$ = 0.6595, *P* = 0.4207) individuals.

Male association time was significantly correlated with female courtship and female body depth (Figure 4). Mean band area in females was correlated with both her depth and courtship activity (Figure 5). The only individual behavior correlated to mean band area was color display and, using a stepwise optimization of a generalized linear model, both depth and color display were significant in explaining mean band area (Depth: coefficient—0.32 ± 0.16, $P = 0.05$, Color display: coefficient—0.34 ± 0.16, $P = 0.04$). As females aged, they expectedly increased in length, but unexpectedly did not get deeper or achieve a larger mean band area (Figure 6), despite these values correlating with length (Mean band area: Pearson's r = 0.37 $P = 0.0304$ and Depth: Pearson's r = 0.52 $P = 0.0017$). Age also had no effect on behaviors.

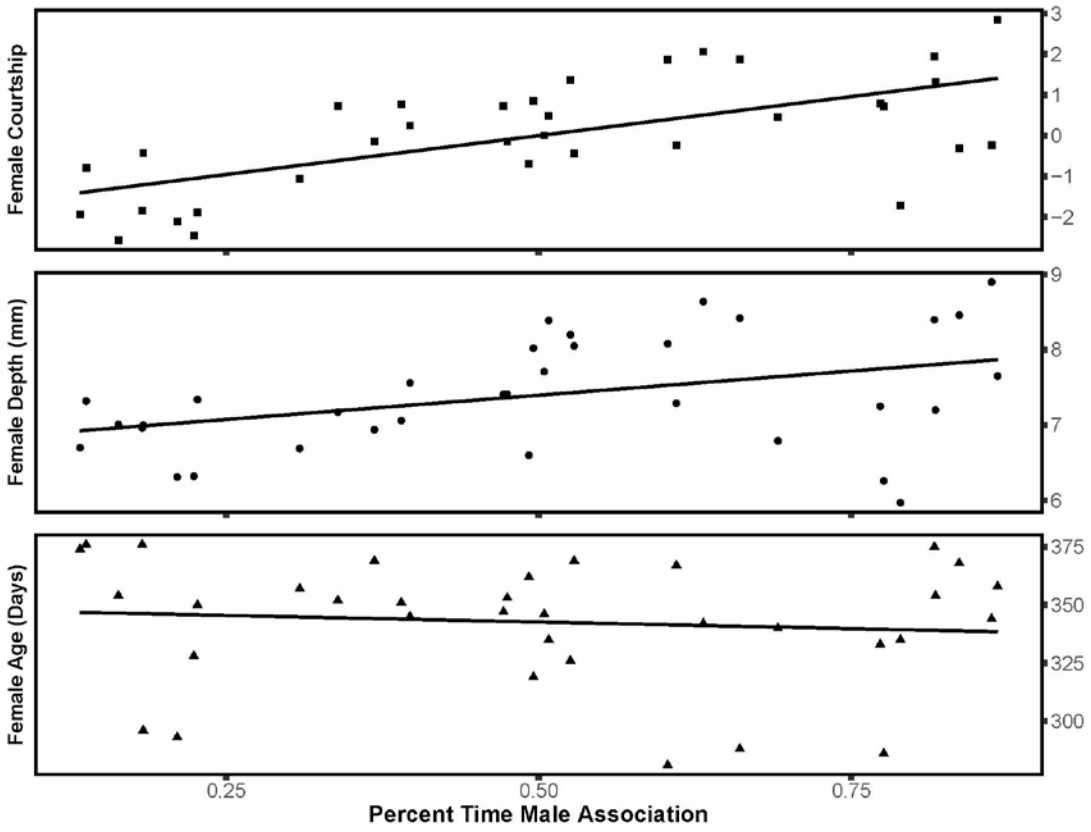

Figure 4: Scatter plots of percent of time a male associated with a female and her age, depth, or courtship effort. Courtship (first principle component of female behaviors) and association time are significantly correlated (Pearson's r = 0.65, *P* < 0.0001) as are depth and association time (Pearson's r = 0.40, *P* = 0.0181).

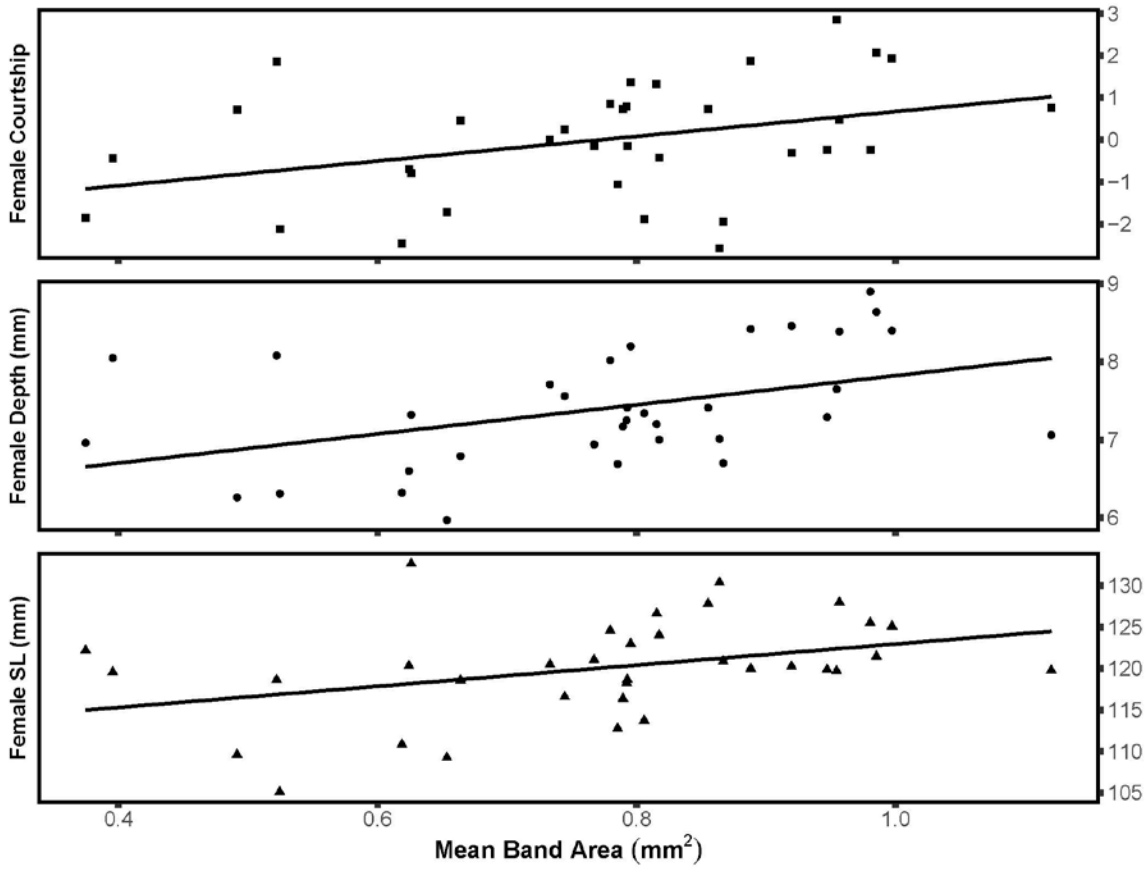

Figure 5: Scatter plots of female mean band area (i.e., size of her secondary sex ornament) and her standard length, depth, or courtship effort. All three traits are significantly correlated with mean band area (Courtship: Pearson's r = 0.65 *P* < 0.0001, Depth: Pearson's r = 0.44 *P* = 0.0098, Standard Length: Pearson's r = 0.37, *P* = 0.0304).

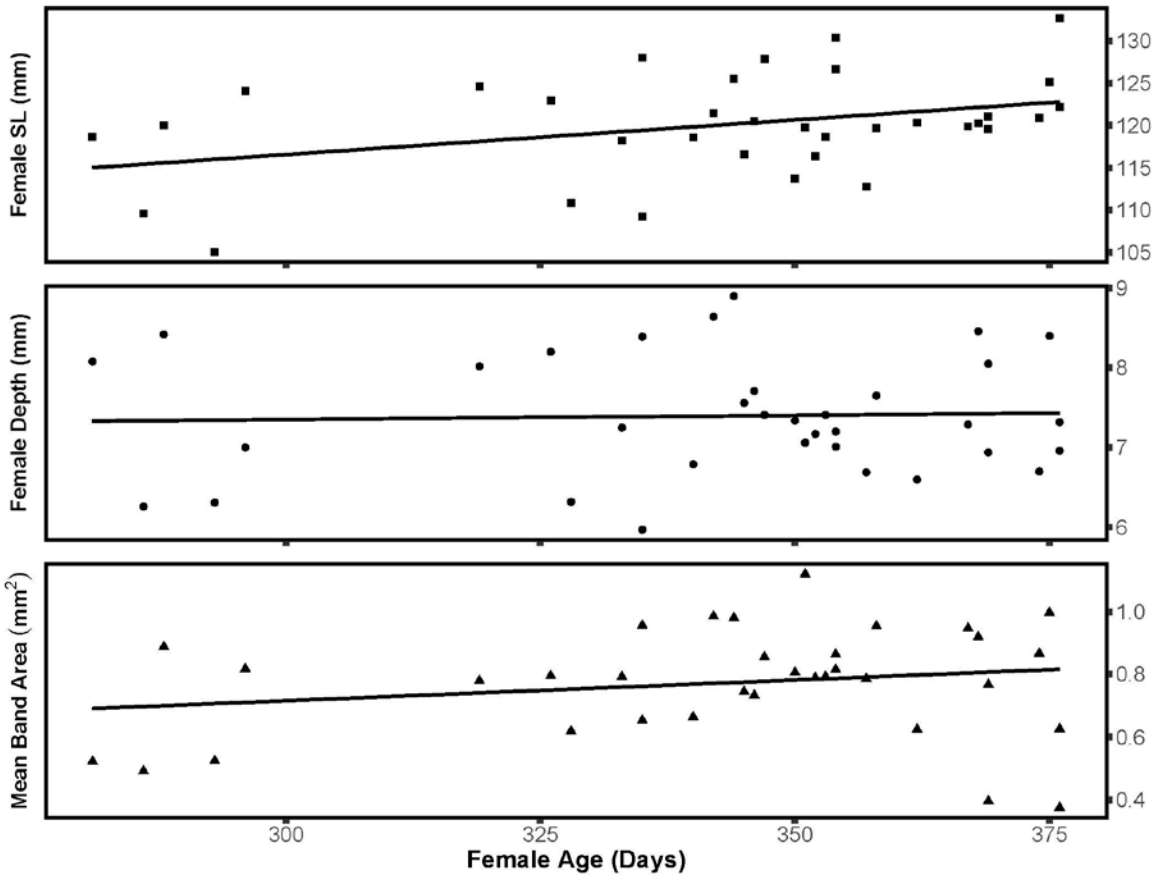

Figure 6: Scatter plots of age of female and her standard length, depth, or mean band area. Only female age and standard length are significantly correlated (Pearson's r = 0.37, *P* = 0.0331).

Model optimization found that courtship was the only explanatory variable retained (Table 1). When I treated each behavior separately, color display and dancing were the only explanatory variables retained in the model. Because behavior had such a large effect and is reinforced by male choice, I optimized without behavior. By doing so, model optimization found that depth was the only significant explanatory variable. Removing both depth and courtship made the model worse than the null.

Table 1 Generalized linear models for percent time the focal male associates with a female. Explanatory variables are the differences between the two experimental females in depth, standard length, mean band area, age, and overall courtship activity. The "Significant Variables" column identifies the variables that were significant in the model ($P < 0.05$). Models were tested against others with higher AIC using an ANOVA Chi Squared test ($P < 0.05$). The "Models Sig. Dif." column indicates models which were significantly different from the current model, whereas the "Models No Dif." column shows models that did not differ significantly from the current model. Note that some models are not nested within others and cannot be tested against each other.

Model	Explanatory Variables	AIC	Significant Variables	Models Sig. Dif.	Models No Dif.
1	~1	51.732	N/A	N/A	N/A
2	~Depth + Length + Band + Age	51.569	None	--	1
3	~Depth	46.083	Depth	1	2
4	~Courtship + Depth + Length + Band + Age	37.200	Courtship	1, 2, 3	--
5	~Courtship + Depth	35.056	Courtship	1, 3	4
6	~Courtship	34.397	Courtship	1	4, 5

Discussion

Here, I investigated the effects of female age on mate choice by male Gulf pipefish. Despite age being correlated with female body size, I found no effect of age on male mate choice in this species. Thus, in my experiment I observed no obvious trade-offs between ornamentation, courtship activity, and growth. Females that reached a large size quickly were no less ornamented, active, or fecund than their counterparts that took longer to reach that length. Furthermore, later maturing females were not any larger than early maturing females. Larger individuals in natural populations thus may represent individuals that matured more rapidly rather than necessarily being older (Halliday and Verrell 1988). Given a large variance in maturation time and a lack of male discrimination among females on the basis of age, the substantial age structure in natural *S. scovelli* populations may have little bearing on the interpretation of population-level effects of selection arising from male mate choice.

From my results, *S. scovelli* males are attracted to females that spend more time actively courting and these females have larger bands, a secondary sex ornament, confirming what has been previously shown in this species (Partridge et al. 2013, Flanagan et al. 2014) as well as other syngnathids (Rosenqvist 1990, Berglund et al. 1997, Berglund and Rosenqvist 2001b, Rosenqvist and Berglund 2011). I additionally find that body depth, a potential indicator of fecundity, predicts male preference and that this trait is also positively correlated with band area and courtship intensity.

Bands on female *S. scovelli* represent a striking sexually dimorphic trait and appear to be targets of sexual selection (Flanagan et al. 2014), so their smaller role in

male choice in my experiment is somewhat counter-intuitive. Given evidence of male preference for larger females across syngnathids (Berglund and Rosenqvist 1993, Berglund 1994, Paczolt and Jones 2010, Aronsen et al. 2013, Rose et al. 2013) and evidence that larger females are more fecund in pipefishes (Jones et al. 2000, Paczolt and Jones 2010, Mobley et al. 2011a) as well as in fishes in general (Barneche et al. 2018) the band ornamentation may serve as way to signal large body sizes and fecundity (Rosenqvist and Berglund 2011). I find that an increase in female body size corresponds to an increase in mean band size, as has been previously demonstrated (Flanagan et al. 2014), and I also find that body depth correlates with mean band size. By controlling for size, I show depth plays a role in male choice independent of size. Bands may therefore be an honest indicator of female fitness by directly signaling fecundity or by emphasizing the traits (depth and length) that are strong indicators of fecundity.

Female depth is a significant contributor to male choice and correlates with female courtship intensity, more so than bands. I suggest depth is a direct indicator of fecundity (Mobley and Jones 2009) and it increases with length, which is another direct indicator of fecundity (Berglund et al. 1986a, Paczolt and Jones 2010). There is evidence that body depth is related to fecundity in other fish (Dadzie and Wangila 1980, Musa and Bhuiyan 2007) and an increase in depth would add to body volume, increasing the space females have to house eggs (Barneche et al. 2018). While *S. scovelli* has a prominent keel compared to other syngnathids (Jones and Avise 2001), females across syngnathids are deeper than males (Jones and Avise 2001, Mobley and Jones 2009)

suggesting depth is important for female fecundity. Males are likely choosing females with greater reproductive potential as mates.

The suite of female behaviors is independent of female size, but color display correlates with mean band area, whereas posing and dancing correlate with depth. It is important to consider that female courtship behaviors are sometimes initiated by males (Partridge et al. 2013) so the strong signals seen in courtship may be partially indicative of an interested male. A fecund female with greater depth is likely to generate male interest and would therefore engage in behaviors that encourage male interaction. Females with greater banding likely display for longer periods to further emphasize their bands. While color display is also stimulated by male interest, females spent greater amounts of time in display than males spent associating with them, suggesting display may represent a female's initial attempt to garner male interest. Display is involved in more than male choice. In other syngnathids, females use display and ornaments in intrasexual interactions (Berglund and Rosenqvist 2001a, 2009) and I have observed Gulf pipefish females displaying in same-sex groups as well as when housed individually. The opaque barrier should eliminate female-female competition in my experiment, suggesting female display in my experimental context is solely for male choice. I suggest that females display to attract males, the males then use bands to aid in assessing depth and fecundity, and finally males engage in courtship with the preferred female.

My data suggest that the female ornament in Gulf pipefish is a low-cost honest indicator, as has been suggested for ornaments in other syngnathids (Berglund et al.

1997, Cunha et al. 2017, Mobley et al. 2018), despite predictions of a costly trade-off (Trivers 1972, Fitzpatrick et al. 1995, Clutton-Brock 2009). It is important to keep in mind that my experiment looks only at one time point in each focal female and did not track life-history tradeoffs throughout female lifetimes (Arnold and Wade 1984, Grafen 1988). Despite some evidence and predictions that fast early growth might reduce life-time survival (Brooks 2000, Metcalfe and Monaghan 2003, Robinson et al. 2006, Lemaître et al. 2015), a sexually selected trait might serve to indicate health and positively correlate with survival (Møller Anders and Alatalo Rauno 1999, Jennions et al. 2001, Bergeron et al. 2008). While I are unable to demonstrate lifetime selection, given the apparent low cost of ornamentation in *S. scovelli*, it may be that the most rapidly growing individuals are in better health and that this good health is reflected in ornamentation and body depth.

In summary, I fail to detect an effect of female age on male mating preferences in the Gulf pipefish. This result partially alleviates concerns over whether previous results regarding mating patterns and sexual selection in natural populations of *S. scovelli* are confounded by age structure. Despite finding no preference based on age, I did find that males attend to female body depth, ornamentation, and courtship intensity. These results add further evidence to the idea that ornaments in sex-role-reversed pipefish are low-cost, honest signals. Females with larger bands and greater fecundity are more active in courtship, pointing to a reinforcing role of behavior in conveying female quality to the choosing males. I also find that the low cost of ornamentation allows fast-maturing individuals to avoid tradeoffs and successfully compete with similarly-sized individuals

that took longer to mature. Syngnathids as a group, and *S. scovelli* in particular, represent an interesting example of how sexual selection can act on a low-cost trait that signals female fecundity.

Acknowledgements

Fish were collected using a permit from Texas Parks and Wildlife (#SPR-0808-307) and the experiment was under Texas A&M University's animal use protocol (#2014-0042). I would like to thank A. Martinez, K. Herrera, P. Hale, E. Maynard, C. Gabeheart, and E. Dooley for their assistance with collection, care, and scoring of the animals and G. Anderson for her help with images.

Literature Cited

Andersson, M. B. 1994. Sexual selection. Princeton University Press.

Arnold, S. J., and M. J. Wade. 1984. On the measurement of natural and sexual selection: applications. Evolution **38**:720-734.

Aronsen, T., K. Mobley, A. Berglund, J. Sundin, A. Billing, and G. Rosenqvist. 2013. The operational sex ratio and density influence spatial relationships between breeding pipefish. Behavioral Ecology **24**:888-897.

Barneche, D. R., D. R. Robertson, C. R. White, and D. J. Marshall. 2018. Fish reproductive-energy output increases disproportionately with body size. Science **360**:642-645.

Begovac, P. C., and R. A. Wallace. 1987. Ovary of the pipefish, Syngnathus scovelli. Journal of Morphology **193**:117-133.

Begovac, P. C., and R. A. Wallace. 1988. Stages of oocyte development in the pipefish, Syngnathus scovelli. Journal of Morphology **197**:353-369.

Bergeron, P., M. Festa-Bianchet, A. Von Hardenberg, and B. Bassano. 2008. Heterogeneity in male horn growth and longevity in a highly sexually dimorphic ungulate. Oikos **117**:77-82.

Berglund, A. 1991. Egg competition in a sex-role reversed pipefish: subdominant females trade reproduction for growth. Evolution **45**:770-774.

Berglund, A. 1994. The operational sex ratio influences choosiness in a pipefish. Behavioral Ecology **5**:254-258.

Berglund, A., and G. Rosenqvist. 1993. Selective males and ardent females in pipefishes. Behavioral Ecology and Sociobiology **32**:331-336.

Berglund, A., and G. Rosenqvist. 2001a. Male pipefish prefer dominant over attractive females. Behavioral Ecology **12**:402-406.

Berglund, A., and G. Rosenqvist. 2001b. Male pipefish prefer ornamented females. Animal Behaviour **61**:345-350.

Berglund, A., and G. Rosenqvist. 2009. An intimidating ornament in a female pipefish. Behavioral Ecology **20**:54-59.

Berglund, A., G. Rosenqvist, and P. Bernet. 1997. Ornamentation predicts reproductive success in female pipefish. Behavioral Ecology and Sociobiology **40**:145-150.

Berglund, A., G. Rosenqvist, and I. Svensson. 1986a. Mate choice, fecundity and sexual dimorphism in two pipefish species (Syngnathidae). Behavioral Ecology and Sociobiology **19**:301-307.

Berglund, A., G. Rosenqvist, and I. Svensson. 1986b. Reversed sex-roles and parental energy investment in zygotes of two pipefish (Syngnathidae) species. Marine Ecology Progress Series **29**:209-215.

Berglund, A., G. Rosenqvist, and I. Svensson. 1989. Reproductive Success of Females Limited by Males in Two Pipefish Species. The American Naturalist **133**:506-516.

Bolland, J., and A. Boettcher. 2005. Population structure and reproductive characteristics of the gulf pipefish,Syngnathus scovelli, in Mobile Bay, Alabama. Estuaries **28**:957-965.

Brooks, R. 2000. Negative genetic correlation between male sexual attractiveness and survival. Nature **406**:67.

Byers, J., E. Hebets, and J. Podos. 2010. Female mate choice based upon male motor performance. Animal Behaviour **79**:771-778.

Carothers, J. H. 1984. Sexual Selection and Sexual Dimorphism in Some Herbivorous Lizards. The American Naturalist **124**:244-254.

Clutton-Brock, T. 2009. Sexual selection in females. Animal Behaviour **77**:3-11.

Cunha, M. A. B., A. Berglund, and N. M. Monteiro. 2017. Female ornaments signal own and offspring quality in a sex-role-reversed fish with extreme male parental care. Marine Ecology **38**:e12461.

Dadzie, S., and B. Wangila. 1980. Reproductive biology, length-weight relationship and relative condition of pond raised Tilapia zilli (Gervais). Journal of Fish Biology **17**:243-253.

Dixon, K. A., and W. H. Cade. 1986. Some factors influencing male-male aggression in the field cricket Gryllus integer (time of day, age, weight and sexual maturity). Animal Behaviour **34**:340-346.

Engqvist, L., and M. Taborsky. 2016. The evolution of genetic and conditional alternative reproductive tactics. Proceedings of the Royal Society B: Biological Sciences **283**:20152945.

Fitzpatrick, S., A. Berglund, and G. Rosenqvist. 1995. Ornaments or offspring: costs to reproductive success restrict sexual selection processes. Biological Journal of the Linnean Society **55**:251-260.

Flanagan, S. P., J. B. Johnson, E. Rose, and A. G. Jones. 2014. Sexual selection on female ornaments in the sex-role-reversed Gulf pipefish (Syngnathus scovelli). J Evol Biol **27**:2457-2467.

Flanagan, S. P., and A. G. Jones. 2017. Genome-wide selection components analysis in a fish with male pregnancy. Evolution **71**:1096-1105.

Freeman-Gallant, C. R., C. C. Taff, D. F. Morin, P. O. Dunn, L. A. Whittingham, and S. M. Tsang. 2010. Sexual selection, multiple male ornaments, and age- and condition-dependent signaling in the common yellowthroat. Evolution **64**:1007-1017.

Grafen, A. 1988. 28 On the Uses of Data on Lifetime Reproductive Success.

Halliday, T. R., and P. A. Verrell. 1988. Body Size and Age in Amphibians and Reptiles. Journal of Herpetology **22**:253-265.

Harrel, F., and C. Dupont. 2014. Hmisc: Harrell Miscellaneous. R package version 3.14-4.

Henson, S. A., and R. R. Warner. 1997. Male and Female Alternative Reproductive Behaviors in Fishes: A New Approach Using Intersexual Dynamics. Annual Review of Ecology and Systematics **28**:571-592.

Hiddink, J. G., and Z. Jager. 2002. Abundance and reproduction of Nilsson's pipefish on tidal flats. Journal of Fish Biology **61**:125-137.

Huber, B. A. 2005. Sexual selection research on spiders: progress and biases. Biological Reviews **80**:363-385.

Jennions, M. D., A. P. Moller, and M. Petrie. 2001. Sexually Selected Traits and Adult Survival: A Meta-Analysis. The Quarterly Review of Biology **76**:3-36.

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

Jones, A. G., and J. C. Avise. 2001. Mating Systems and Sexual Selection in Male-Pregnant Pipefishes and Seahorses: Insights from Microsatellite-Based Studies of Maternity. Journal of Heredity **92**:150-158.

Jones, A. G., G. Rosenqvist, A. Berglund, and J. C. Avise. 2000. Mate quality influences multiple maternity in the sex-role-reversed pipefish Syngnathus typhle. Oikos **90**:321-326.

Jones, A. G., D. Walker, and J. C. Avise. 2001. Genetic evidence for extreme polyandry and extraordinary sex-role reversal in a pipefish. Proceedings of the Royal Society of London. Series B: Biological Sciences **268**:2531-2535.

Kodric-Brown, A., R. M. Sibly, and J. H. Brown. 2006. The allometry of ornaments and weapons. Proceedings of the National Academy of Sciences **103**:8733-8738.

Kruuk, L. E. B., J. Slate, J. M. Pemberton, S. Brotherstone, F. Guinness, and T. Clutton-Brock. 2002. Antler size in red deer: hertiability and selection but no evolution. Evolution **56**:1683-1695.

Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. Evolution **37**:1210-1226.

Lemaître, J.-F., V. Berger, C. Bonenfant, M. Douhard, M. Gamelon, F. Plard, and J.-M. Gaillard. 2015. Early-late life trade-offs and the evolution of ageing in the wild. Proceedings of the Royal Society B: Biological Sciences **282**.

Lindenfors, P., and B. S.Tullberg. 2011. 2 - Evolutionary Aspects of Aggression: The Importance of Sexual Selection. Pages 7-22 *in* R. Huber, D. L. Bannasch, and P. Brennan, editors. Advances in Genetics. Academic Press.

Loyau, A., M. S. Jalme, and G. Sorci. 2005. Intra- and Intersexual Selection for Multiple Traits in the Peacock (Pavo cristatus). Ethology **111**:810-820.

Madsen, T., R. Shine, J. Loman, and T. Håkansson. 1993. Determinants of mating success in male adders, Vipera berus. Animal Behaviour **45**:491-499.

Marchetti, K. 1998. The evolution of multiple male traits in the yellow-browed leaf warbler. Animal Behaviour **55**:361-376.

Mateos, C. 1998. Sexual selection in the ring-necked pheasant: a review. Ethology Ecology & Evolution **10**:313-332.

McCann, T. S. 1981. Aggression and sexual activity of male Southern elephant seals, Mirounga leonina. Journal of Zoology **195**:295-310.

McElligott, A. G., M. P. Gammell, H. C. Harty, D. R. Paini, D. T. Murphy, J. T. Walsh, and T. J. Hayden. 2001. Sexual size dimorphism in fallow deer (Dama dama): do larger, heavier males gain greater mating success? Behavioral Ecology and Sociobiology **49**:266-272.

McGlothlin, J. W., P. G. Parker, V. Nolan, and E. D. Ketterson. 2005. Correlational selection leads to genetic integration of body size and an attractive plumage trait in dark-eyed juncos. Evolution **59**:658-671.

Metcalfe, N. B., and P. Monaghan. 2003. Growth versus lifespan: perspectives from evolutionary ecology. Experimental Gerontology **38**:935-940.

Mobley, K. B., and A. G. Jones. 2009. Environmental, demographic, and genetic mating system variation among five geographically distinct dusky pipefish (Syngnathus floridae) populations. Molecular Ecology **18**:1476-1490.

Mobley, K. B., C. Kvarnemo, I. Ahnesjö, C. Partridge, A. Berglund, and A. G. Jones. 2011a. The effect of maternal body size on embryo survivorship in the broods of pregnant male pipefish. Behavioral Ecology and Sociobiology **65**:1169-1177.

Mobley, K. B., J. R. Morrongiello, M. Warr, D. J. Bray, and B. B. M. Wong. 2018. Female ornamentation and the fecundity trade-off in a sex-role reversed pipefish. Ecol Evol **8**:9516-9525.

Mobley, K. B., C. M. Small, and A. G. Jones. 2011b. The genetics and genomics of Syngnathidae: pipefishes, seahorses and seadragons. J Fish Biol **78**:1624-1646.

Møller Anders, P., and V. Alatalo Rauno. 1999. Good-genes effects in sexual selection. Proceedings of the Royal Society of London. Series B: Biological Sciences **266**:85-91.

Møller, A. P. 1990. Sexual behavior is related to badge size in the house sparrow Passer domesticus. Behavioral Ecology and Sociobiology **27**:23-29.

Moller, A. P., and A. Pomiankowski. 1993. Why have birds got multiple sexual ornaments? Behavioral Ecology and Sociobiology **32**:167-176.

Musa, A., and A. S. Bhuiyan. 2007. Fecundity on Mystus bleekeri (Day, 1877) from the River Padma Near Rajshahi City. Turkish Journal of Fisheries and Aquatic Sciences **7**.

Paczolt, K. A., and A. G. Jones. 2010. Post-copulatory sexual selection and sexual conflict in the evolution of male pregnancy. Nature **464**:401-404.

Parkinson, K. L., and D. J. Booth. 2016. Rapid growth and short life spans characterize pipefish populations in vulnerable seagrass beds. Journal of Fish Biology **88**:1847-1855.

Partridge, C., A. Boettcher, and A. G. Jones. 2010. Short-term exposure to a synthetic estrogen disrupts mating dynamics in a pipefish. Horm Behav **58**:800-807.

Partridge, C., A. Boettcher, and A. G. Jones. 2013. The Role of Courtship Behavior and Size in Mate Preference in the Sex-Role-Reversed Gulf Pipefish, Syngnathus scovelli. Ethology **119**:692-701.

Ripley, J. L., and C. M. Foran. 2006. Population structure, growth rates, and seasonal abundance of twoSyngnathus pipefish species. Estuaries and Coasts **29**:1161-1171.

Robinson, M. R., J. G. Pilkington, T. H. Clutton-Brock, J. M. Pemberton, and L. E. B. Kruuk. 2006. Live fast, die young: trade-offs between fitness components and sexually antagonistic selection on weaponry in soay sheep. Evolution **60**:2168-2181.

Rohlf, F. 2010. TPS DIG, Suny Stony Brook.

Rose, E., K. A. Paczolt, and A. G. Jones. 2013. The contributions of premating and postmating selection episodes to total selection in sex-role-reversed Gulf pipefish. Am Nat **182**:410-420.

Rosenqvist, G. 1990. Male mate choice and female-female competition for mates in the pipefish Nerophis ophidion. Animal Behaviour **39**:1110-1115.

Rosenqvist, G., and A. Berglund. 2011. Sexual signals and mating patterns in Syngnathidae. J Fish Biol **78**:1647-1661.

Rosenthal, G. G. 2017. Mate choice: the evolution of sexual decision making from microbes to humans. Princeton University Press.

Searcy, W. A., and M. Andersson. 1986. Sexual selection and the evolution of song. Annual Review of Ecology and Systematics **17**:507-533.

Soma, M., and L. Z. Garamszegi. 2015. Evolution of courtship display in Estrildid finches: dance in relation to female song and plumage ornamentation. Frontiers in Ecology and Evolution **3**.

Svensson, I. 1988. Reproductive Costs in Two Sex-Role Reversed Pipefish Species (Syngnathidae). Journal of Animal Ecology **57**:929-942.

Taff, C. C., D. Steinberger, C. Clark, K. Belinsky, H. Sacks, C. R. Freeman-Gallant, P. O. Dunn, and L. A. Whittingham. 2012. Multimodal sexual selection in a warbler: plumage and song are related to different fitness components. Animal Behaviour **84**:813-821.

Takahashi, E., R. M. Connolly, and S. Y. Lee. 2003. Growth and Reproduction of Double-Ended Pipefish, Syngnathoides biaculeatus, in Moreton Bay, Queensland, Australia. Environmental Biology of Fishes **67**:23-33.

Trivers, R. 1972. Parental investment and sexual selection. Biological Laboratories, Harvard University Cambridge.

Vanhooydonck, B., A. Herrel, R. Van Damme, J. J. Meyers, and D. J. Irschick. 2005. The relationship between dewlap size and performance changes with age and sex in a Green Anole (Anolis carolinensis) lizard population. Behavioral Ecology and Sociobiology **59**:157-165.

Vincent, A., I. Ahnesjö, A. Berglund, and G. Rosenqvist. 1992. Pipefishes and seahorses: Are they all sex role reversed? Trends in Ecology & Evolution **7**:237-241.

Vincent, A. C. J., A. Berglund, and I. Ahnesjö. 1995. Reproductive ecology of five pipefish species in one eelgrass meadow. Environmental Biology of Fishes **44**:347-361.

Whatley, E. C. 1969. A study of Syngnathus scovelli in fresh waters of Louisiana and salt waters of Mississippi. Gulf and Caribbean Research **2**:437-474.

Wikelski, M., and F. Trillmich. 1997. Body size and sexual size dimorphism in marine iguanas fluctuates as a result of opposing natural and sexual selection: an island comparison. Evolution **51**:922-936.

CHAPTER III

EREFINDER: GENOME-WIDE DETECTION OF ESTROGEN RESPONSE ELEMENTS

Synopsis

Estrogen response elements (EREs) are specific DNA sequences to which ligand-bound estrogen receptors (ERs) physically bind, allowing them to act as transcription factors for target genes. Locating EREs and ER responsive regions is therefore a potentially important component of the study of estrogen-regulated pathways. Here, I report the development of a novel software tool, EREFinder, which conducts a genome-wide, sliding-window analysis of estrogen receptor binding affinity. I demonstrate the effects of adjusting window size and highlight the program's general agreement with ChIP studies. I further provide two examples of how EREFinder can be used for comparative approaches. EREFinder can handle large input files, has settings to allow for broad and narrow searches, and provides the full output to allow for greater data manipulation. These features facilitate a wide range of hypothesis testing for researchers and make EREFinder an excellent tool to aid in estrogen-related research.

Introduction

Estrogen receptors (ERs) are the most primitive steroid receptor in the chordate lineage (Thornton et al. 2003) and play a part in a wide range of biological processes within vertebrates. While many studies of ERs have focused primarily on pathologies and medical applications (Arnal et al. 2017, Jia et al. 2015), the pleiotropic effects of ER activity can be wide-ranging (McDonnel and Norris 2002). Consequently, the genome-

level details of estrogen signaling, especially with respect to ER-binding and its effects on transcription, are topics of considerable interest to a wide range of disciplines, from human health to ecology and evolution. Many factors contribute to the control of gene regulation by ERs, including the type of ligand (Anstead et al. 1997, Arnal et al. 2017), the presence of appropriate cofactors (Arnal et al. 2017, Glass and Rosenfeld 2000), and the genome-level distribution of estrogen response elements (EREs). Studies have highlighted the importance of both expression changes of estrogen-responsive genes (Young and Crews 1995) and changes in the structure of the ER (Tohyama et al. 2016), with focus shifting to the roles that EREs play at the level of the genome (Callard et al. 2001, Frankl-Viches et al. 2015).

The consensus ERE is a palindromic sequence of two half-sites with a 3-base-pair spacer (AGGTCAnnnTGACCT) that is preferentially bound by the ER (Klein-Hitpass et al. 1988, Boyer et al. 2000) making EREs a critical point of contact between ERs and the transcription of ER-regulated genes. Point mutations, resulting in deviations from the consensus sequence, negatively affect the strength of ER-binding (Tyulmenkov and Klinge 2001, Geserick et al. 2005, Deegan et al. 2011), and weaker ER-binding has been shown to negatively affect transcription of target genes (Klinge et al. 2000, Tyulmenkov and Klinge 2001). Of particular importance are the half-sites (AGGTCA and TGACCT), because strong binding of an ER requires at least one perfect half-site (Tyulmenkov and Klinge 2001, Deegan et al. 2011). Moreover, the number of available EREs in a given region positively correlates with expression of the target gene, whether those EREs are perfect or imperfect (Martinez and Wahli 1989, Kato et al. 1992,

Geserick et al. 2005). While some genes are regulated by distant EREs, the larger share of gene regulation occurs via EREs that lie in close proximity to the transcription start site (Carroll et al. 2006, Lin et al. 2007).

Empirical identification of transcription factor binding sites, such as EREs, has been accomplished by transfecting upstream regions to determine sequence responsiveness (Klein-Hitpass et al. 1988, Martinez and Wahli 1989, Kato et al. 1992) or by more current methods utilizing high-throughput techniques such as ChIP-seq (Johnson et al. 2007), HT-SELEX (Ogawa and Biggin 2011) and protein binding arrays (Mukherjee et al. 2004). Computational methods also exist that look for over-represented sequences in genomic regions affiliated with a response to a particular transcription factor (Bailey et al. 2006). While each technique has different assumptions with its own set of advantages and disadvantages, the end result is a dataset of motifs that have been preferentially bound by the target protein. This dataset is then converted into a position frequency matrix and a position weight matrix. Position weight matrices assume each substitution at a base pair position has an independent effect on the binding affinity of the protein to the motif and the magnitude of the effect is related to conservation of the base pair position in the frequency matrix (Stormo 2000). There is no shortage of programs that can search sequences based on the position weight matrix (Frith et al. 2003, Kel et al. 2003, Tan and Lenhard 2016, Wang et al. 2016); however, the generation of a position frequency matrix can involve bias (Teytelman et al. 2013) and the assumption of independence of base pairs may often be unwarranted (Man and Stormo 2001, Bulyk et al. 2002, Omidi et al. 2017). The lack of independence is

especially non-trivial in EREs, as loss of a perfect half site has a larger effect on the binding affinity than point mutations after the half site is lost (Tyulmenkov and Klinge 2001, Deegan et al. 2011). An online ERE detection program, Dragon ERE Finder, was available for about a decade beginning in 2003 (Bajic 2003). However, Dragon ERE Finder was severely limited for genome-wide analysis by the fact that it could scan a maximum of 10kb at a time, and it is no longer available for use (V. Bajic pers. comm.). Hence, I see a need for a search algorithm that uses the empirically estimated binding affinity of the ER to identify putative EREs on a genome-wide scale.

Here I present EREfinder, a novel program for locating putative EREs. EREFinder utilizes the equations developed and empirically validated by Tyulmenkov and Klinge (2001), which predict the binding affinity (K_d) of ERα and ERβ to a given DNA sequence. The equations are based on the number of perfect half-sites in an ERE and also penalize mismatches that are not complementary substitutions (e.g. A→T or C→G) when a half-site is imperfect. The canonical binding sequences for estrogen receptor α or β (ERα and ERβ) are identical, as are the scoring criteria, but the weightings for perfect half-sites and mismatches are quantitatively different. By utilizing a sliding window defined by the user, EREFinder can function both as a detector of EREs and as an indicator of the average expected ER-binding affinity in a given subset of the sequence. Here I demonstrate EREFinder's ability to locate regions known to have *in vivo* ER binding and give two examples of how the output of EREFinder can be used in a comparative context.

EREFinder

Following the formulae presented by Tyulmenkov and Klinge (2001), EREFinder performs a scan across an entire fasta-format input file, evaluating binding affinities of either estrogen receptor α or β. Because the canonical ERE is palindromic, the program reads in only one direction along a given sequence. EREFinder uses a sliding window analysis, where the binding affinity (K_d) of every sequential set of fifteen base pairs is calculated. EREFinder takes the reciprocal of each K_d value (i.e., K_d^{-1}), because larger values of K_d^{-1} indicate stronger ER-binding. This approach also increases the influence of small K_d values on the local mean, increasing the likelihood of detecting canonical EREs. The mean of all K_d^{-1} values is then calculated across a user-specified width of the sliding window. In addition to designating the width, the user can also choose the slide interval for the sliding window. At one extreme, EREFinder can provide the ER-binding affinity for every overlapping 15-base pair subsequence present in a fasta file, if the user selects a width of one and slide interval of one. The output from EREFinder is given as a comma-delimited text file suitable for additional analysis in a spreadsheet program or a statistical package, such as R.

Methods/Results

Testing effects of window sizes

I sought to investigate how altering the window size affects the search algorithm's output. To do so I selected a 10Mb segment from the nuclear genome of *Homo sapiens* GRCh38.p9 Chromosome 17 (Coordinates: 40Mb-50Mb) that contained two perfect EREs. I used four different search criteria in EREFinder: individual base pair

score (that is, the K_d^{-1} of every overlapping 15-bp sequence), 100bp, 1kb, and 10kb sliding windows. For window sizes larger than one, slide intervals were half the window size, and I used the ERα K_d^{-1} values. Regardless of window size, EREFinder generated an output file in less than one minute on a standard desktop.

In the individual base pair score, 0.04% of all base pair scores had a K_d^{-1} value greater than 0.09, a value that indicates a perfect half-site with no more than two mismatches in the rest of the ERE. Windows containing the perfect EREs usually had the highest mean K_d^{-1} values, regardless of window size (Figure 7). As the windows get larger, the higher mean K_d^{-1} values begin to include regions enriched for half-sites, and my tests show that in some rare cases these regions have greater mean expected binding values than regions with perfect EREs (data not shown). Similarly, a nearly perfect ERE, which would have high binding affinity, can be in a low value window if there are no other high-binding locations nearby. Expectedly, increasing window size lowered the variance between mean K_d^{-1} values among windows (Figure 7).

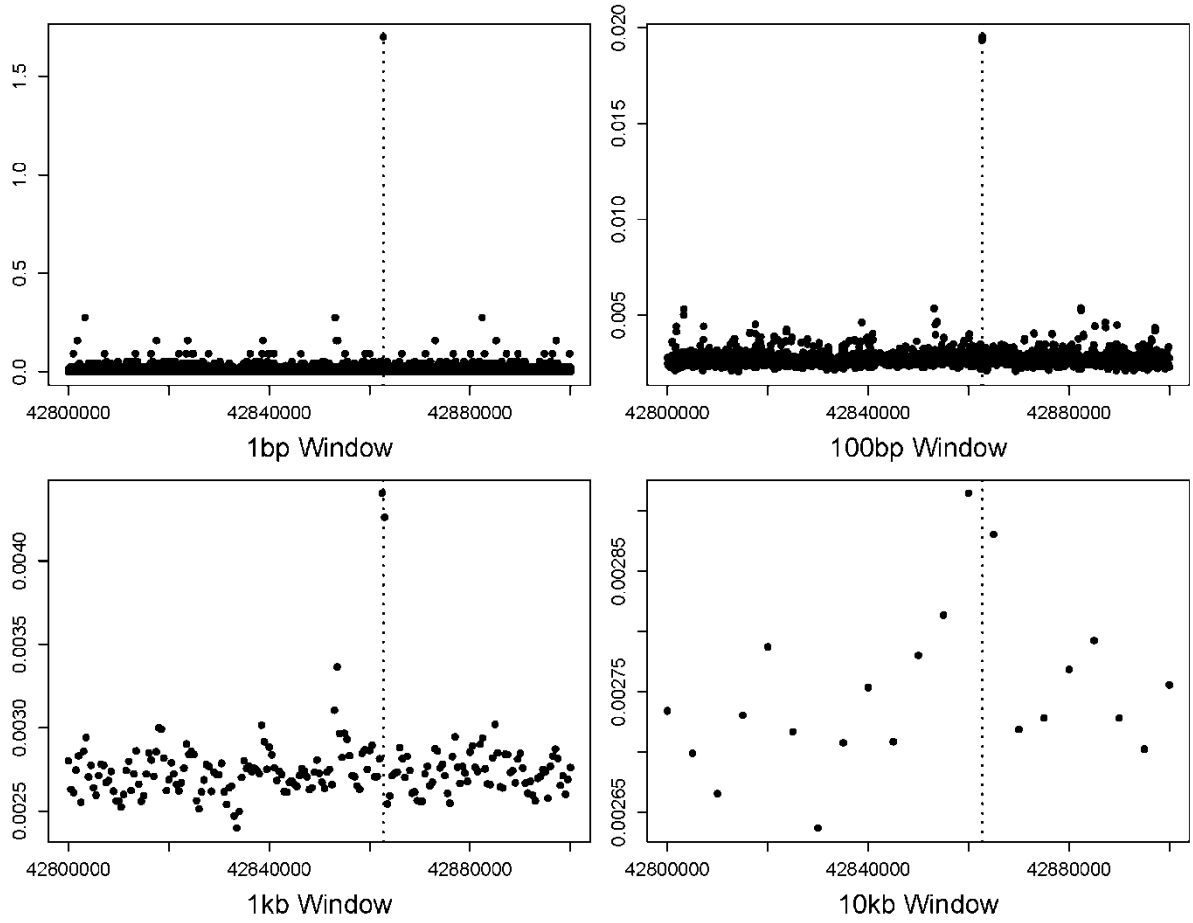

Figure 7: Effect of window size in EREFinder across a 100kb sequence from human chromosome 17. The x-axis shows base pair coordinates and the y-axis gives mean K_d^{-1} for the window, calculated by EREFinder. Each point represents a window (with the window size indicated below each panel) and the dashed vertical line represents location of a perfect ERE. Notice mean K_d^{-1} range (y-axis) drastically decreases as window sizes get larger.

Selecting a window size depends on the specific goals of the researcher. For searches of exact coordinates of singular putative EREs, small windows (100bp or individual scores) are appropriate to determine those locations. I suggest for large scans (1Mb or more) 100bp windows be used, because a file containing K_d^{-1} values for every overlapping 15-bp subsequence can be quite large and unwieldy. For searches of general

ER-binding, I suggest mid-sized windows (100bp or 1kb) are useful because both binding at one specific location and across multiple proximal locations are important for ER-binding and response. It is important to note that the output of EREFinder gives both mean binding values and counts of individual base pair scores at certain K_d^{-1} values, thereby allowing researchers to focus on whichever values they deem more appropriate for their investigation. Since I envision EREFinder to be used primarily in a comparative context, further demonstrations of the program will use 1kb windows, as I are interested in locating regions enriched for ER-binding.

Verification of ER-binding

Given that EREFinder is designed to locate putative regions of ER-binding (keeping in mind that not all regions with EREs will actually bind ER *in vivo*), I expect at least a weak positive relationship between EREFinder detections and empirical observations of ER-binding. I tested this expectation by hypothesizing that regions identified as ER-binding from ChIP-seq data should have higher mean K_d^{-1} values than background regions. Lin et al. (2007) identified 1,234 clusters of estrogen receptor binding in the MCF human cell line using a ChIP-PET approach and aligning to the human genome (Hg17). I selected 200 of these clusters from Chromosomes 4, 12, 17, and X for further investigation. Using EREFinder, I performed a 1kb sliding window analysis with 500bp slide intervals on the Hg17 human genome. I masked regions with missing data and used ERα values for binding affinity. Any windows containing more than 200 unknown bases (i.e., Ns) were removed from analysis. I identified the windows that fell within the coordinates provided by Lin et al. (2007) for high binding clusters. I

chose a 1kb window because binding can involve multiple sites and Lin et al. (2007) report both EREs and half-sites occurring in the same binding region. Since the ChIP-PET regions varied in their sequence size, I took the window with the highest binding in each cluster as that would represent the location of the active ER-binding. I predicted that the subset of windows found in the ChIP regions would have significantly higher ER-binding than expected for randomly chosen windows across the four chromosomes.

The mean K_d^{-1} for the ChIP regions fell in the 95th percentile of all window scores across the chromosomes, demonstrating that EREFinder assigns higher scores to regions with known ER binding. The density distribution of mean K_d^{-1} for the ChIP regions differs substantially from that for random windows across the chromosomes (Figure 8). The ChIP regions also include some chromosomal segments with extremely high scores due to perfect or nearly perfect EREs (Figure 8, outliers). However, these regions of exceptionally strong binding make up only 2.5% (in the case of perfect EREs) and 28% (in the case of nearly perfect EREs) of ChIP-identified regions. Rather, most ChIP regions contain multiple ERE-like sequences, each of which has a few mismatches (67% of regions) compared to the canonical ERE. This result is consistent with Lin et al.'s (2007) analysis, which shows that 69.45% of their ChIP regions contain at least one ERE motif with some mismatches and a perfect half-site. This comparison with ChIP-seq data illustrates that EREFinder is able to identify regions in the genome that are expected to have high ER-binding.

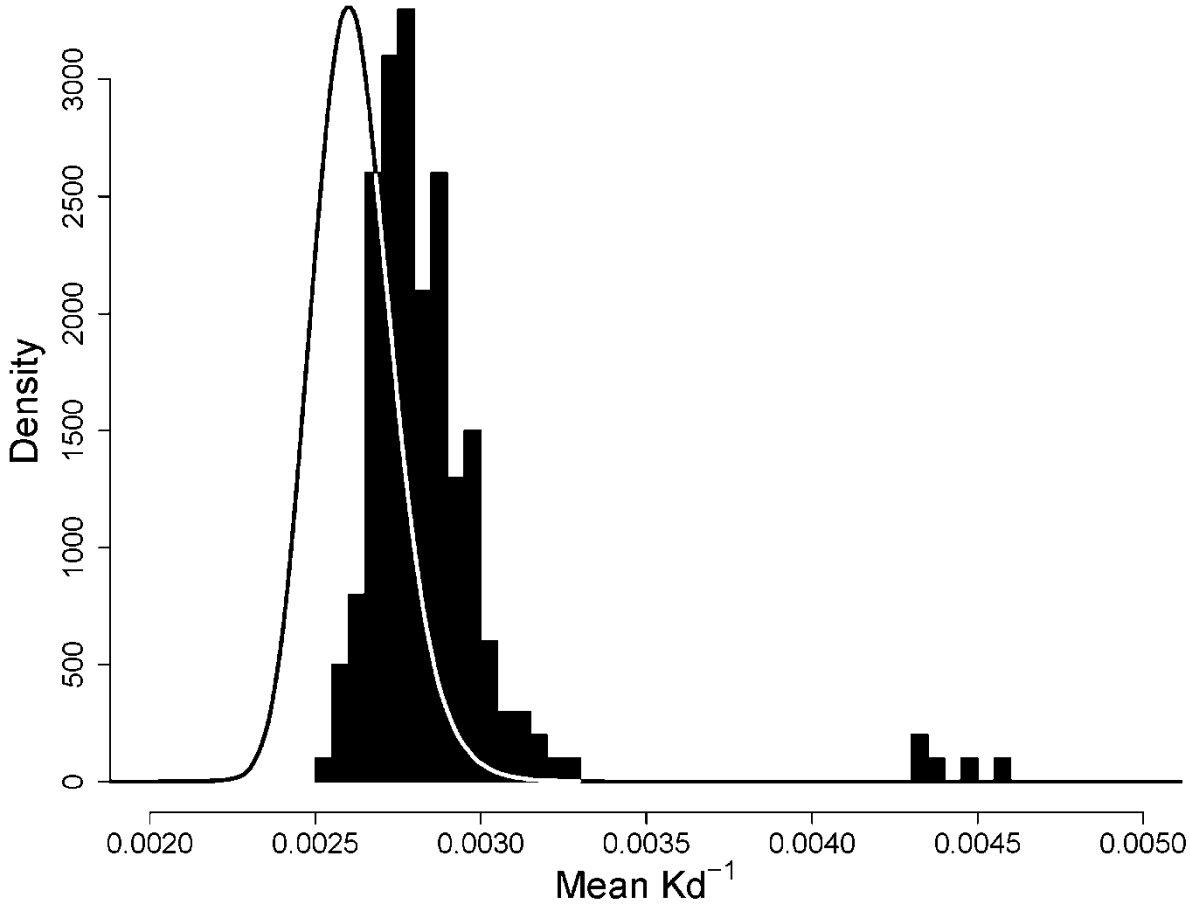

Figure 8: Density plot of mean K_d^{-1} values for 1kb windows generated from EREFinder on human chromosomes 4, 12, 17, and X. The smoothed line is the density plot of mean K_d^{-1} values for all windows (whether or not they contain known ER-binding). Dark bars show the density of mean K_d^{-1} values for the 200 windows associated with known ChIP-seq estrogen receptor binding regions (Lin et al. 2007).

Example Uses for EREFinder

Comparison of general ER-binding across human chromosomes

As an example of how EREFinder can be used to estimate ER-binding for comparative purposes, I used EREFinder to estimate mean K_d^{-1} for each entire chromosome (i.e., window size = chromosome size) for all chromosomes in the human genome. I hypothesized that I would find a correlation across chromosomes between average ER affinity and the number of estrogen responsive genes on that chromosome. Using a list of estrogen responsive genes (n = 356 genes) gathered from the published literature (Bourdeau et al. 2004, Lin et al. 2007), I calculated the frequency of estrogen responsive genes on each human chromosome (genes divided by size of chromosome in base pairs). I then tested for a correlation between mean K_d^{-1} for each chromosome and frequency of estrogen responsive genes on the chromosome. Both mean K_d^{-1} and gene per base pair were log transformed and genes per base pair were multiplied by the size of the Y chromosome (i.e., ~57Mb) for ease of visualization.

The correlation between chromosomal mean K_d^{-1} and density of estrogen responsive genes was significant but weak (P = 0.0258, R^2 = 0.172) (Figure 9). The correlation becomes stronger after removal of the Y chromosome, which had no estrogen responsive genes, and Chromosome 22, which had only two estrogen responsive genes ($P < 0.0001$, R^2 = 0.5335). While I have demonstrated a correlation across human chromosomes, these data could lead to more probing inquiries. For example, why does Chromosome X have an extremely low residual in the analysis despite being a sex chromosome? And, why do some chromosomes have higher binding than others despite

similar densities of estrogen responsive genes? Such patterns would not be apparent with narrower searches, and even these patterns could be investigated in a comparative context across species.

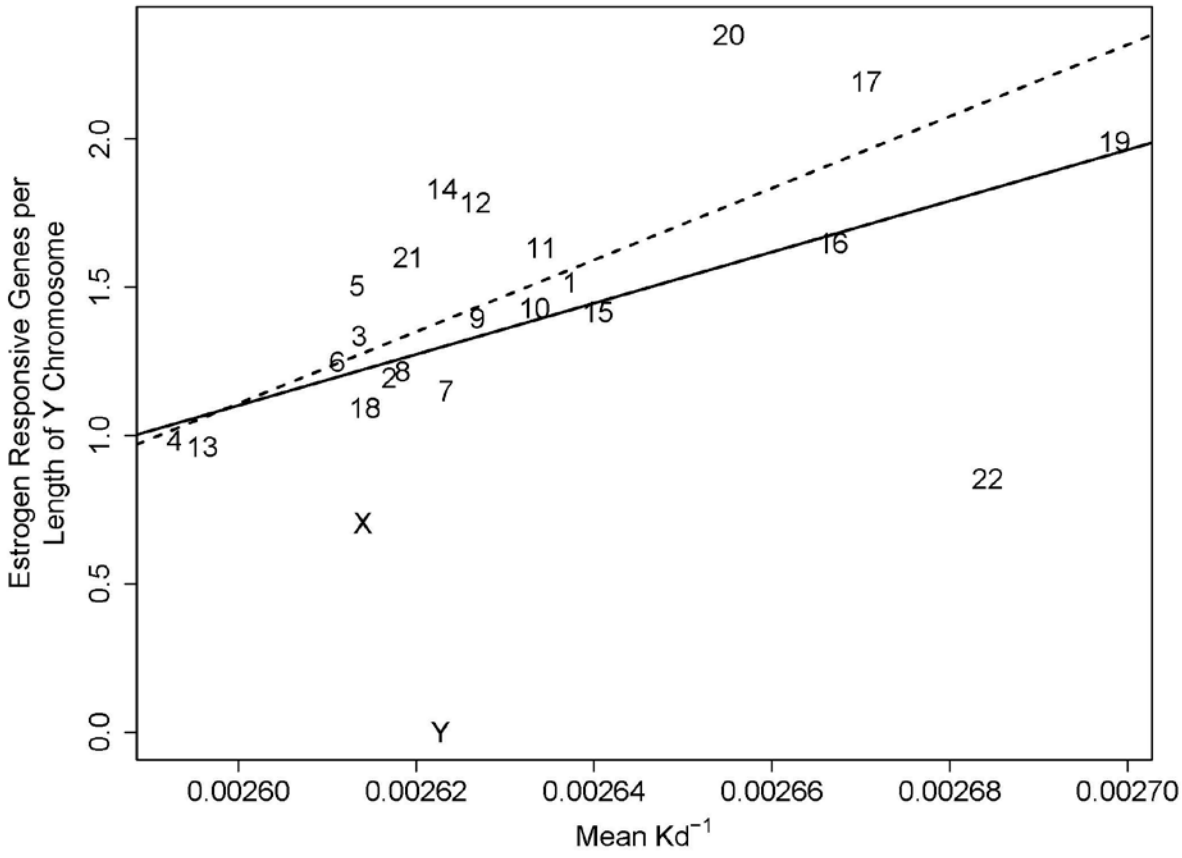

Figure 9: Using the human genome, scatter plot of chromosomal estrogen affinity and number of estrogen responsive genes found per chromosome (adjusted for size of chromosome). Both values are log transformed. Symbols (numeric, X and Y) indicate the identity of the relevant chromosome. Solid line includes all chromosomes ($P = 0.026$, $R^2 = 0.172$), dashed line excludes the extremely short, gene depauperate chromosomes 22 and Y ($P < 0.0001$, $R^2 = 0.534$).

Conservation of ER-binding across five primates

I hypothesized that ER-binding would be largely conserved across hominid genomes. To test this I download the genomes and gff files for the closest extant relatives of humans: orangutan (GenBank file: P_pygmaeus_2.0.2), gorilla (gorGor4), bonobo (panpan1.1), and chimpanzee (Pan_tro 3.0). I used the gff files to locate shared genes across these four non-human primate species corresponding to estrogen-responsive genes on the four human chromosomes I tested above (Bourdeau et al. 2004, Lin et al. 2007). These estrogen responsive genes were from breast cancer cell line MCF-7, were exposed to 17β-estradiol for one to two hours, and showed a transcription response between six and 48 hours. Fifty genes were selected and the corresponding chromosomes for each species were scanned with EREFinder using window size of 1kb with a 500bp interval for each chromosome. Using a custom R script (supplemental materials) I removed any windows with less than 80% of the base pairs scored, and I determined the mean K_d^{-1} for all remaining windows in the sample. To define high ER-binding peaks in my custom script, I used a hard cutoff of mean $K_d^{-1} > 0.003$ as this number is between the three standard deviation estimates for the four human chromosomes. My custom R script analyzed the trimmed input file, used the cut-off value, window size, and slide length size, and produced a file reporting the highest K_d^{-1} values within each detected peak of ER-binding potential. For each chromosome, I determined the location of each estrogen responsive gene and counted the number of high ER affinity regions 50kb up and downstream of the gene start and stop locations as well as the entire intragenic region (R script supplemental).

To determine conservation of peaks across species I considered only the number of peaks present at three different categories for each gene: upstream of the gene, downstream of the gene, and intragenic. If the number of peaks was unchanged in all three categories between two species then binding was said to be fully conserved. If, in only one of the three categories, the number of peaks differed by a value of one between two species, then binding was said to be mostly conserved. Exact position of peaks was not considered as insertions and deletions could cause a mismatch in location relative to the gene. Peak patterns were conserved more often between closely related species (Figure 10A) with chimpanzees and bonobos having strong conservation in over two-thirds of the genes investigated.

I sought to further demonstrate the comparative utility of the program by focusing on one gene with a strongly conserved pattern, *PDLIM3* (Figure 10B). Using seqinr (Charif and Lobry 2007) in the Biostrings package (Pages et al. 2018), I extracted the 50kb of sequence upstream of the gene for each species and then aligned them using the online MAFFT v7 program (Katoh and Standley 2013). Those alignments were read back through EREFinder, which was used to calculate K_d^{-1} for every overlapping 15-bp interval. Using these scores, I were able to locate regions of high binding, coordinates of perfect half sites, and the extent of conservation across species. For *PDLIM3*, the first region of high binding was driven by three perfect half-sites of ERE and two binding regions without perfect half-sites. Three of these, two perfect half-sites and one general binding region, occur within 200bp of each other (Figure 10C). The second high binding peak has no perfect half-sites but five binding regions close together. That particular

region has undergone a deletion in bonobos, which is why they do not show the second binding peak present in the other four species. Given the lack of perfect half-sites in this region and the loss in bonobos, this peak may not actually be important to estrogen binding, especially compared to the more proximal peak, but additional empirical work would be necessary to investigate this possibility. This exercise shows that EREFinder can be useful in the identification of ERE gains and losses over evolutionary time in a comparative context.

Discussion

Here, I present a novel software package, EREFinder, which quickly scans whole genomes and calculates estrogen-receptor binding potential. Using EREFinder I have demonstrated: 1) how altering the window size can affect the investigation and interpretation of the output, 2) that results from EREFinder are consistent with expectations from empirical studies of ER binding *in vivo*, and 3) that EREFinder is probably most useful for comparative investigations of ERE gains and losses. EREFinder has the advantages that it allows the user to access every binding score for every region of the genome and that it outputs the data in a format that can be manipulated easily in a statistical software package, such as R

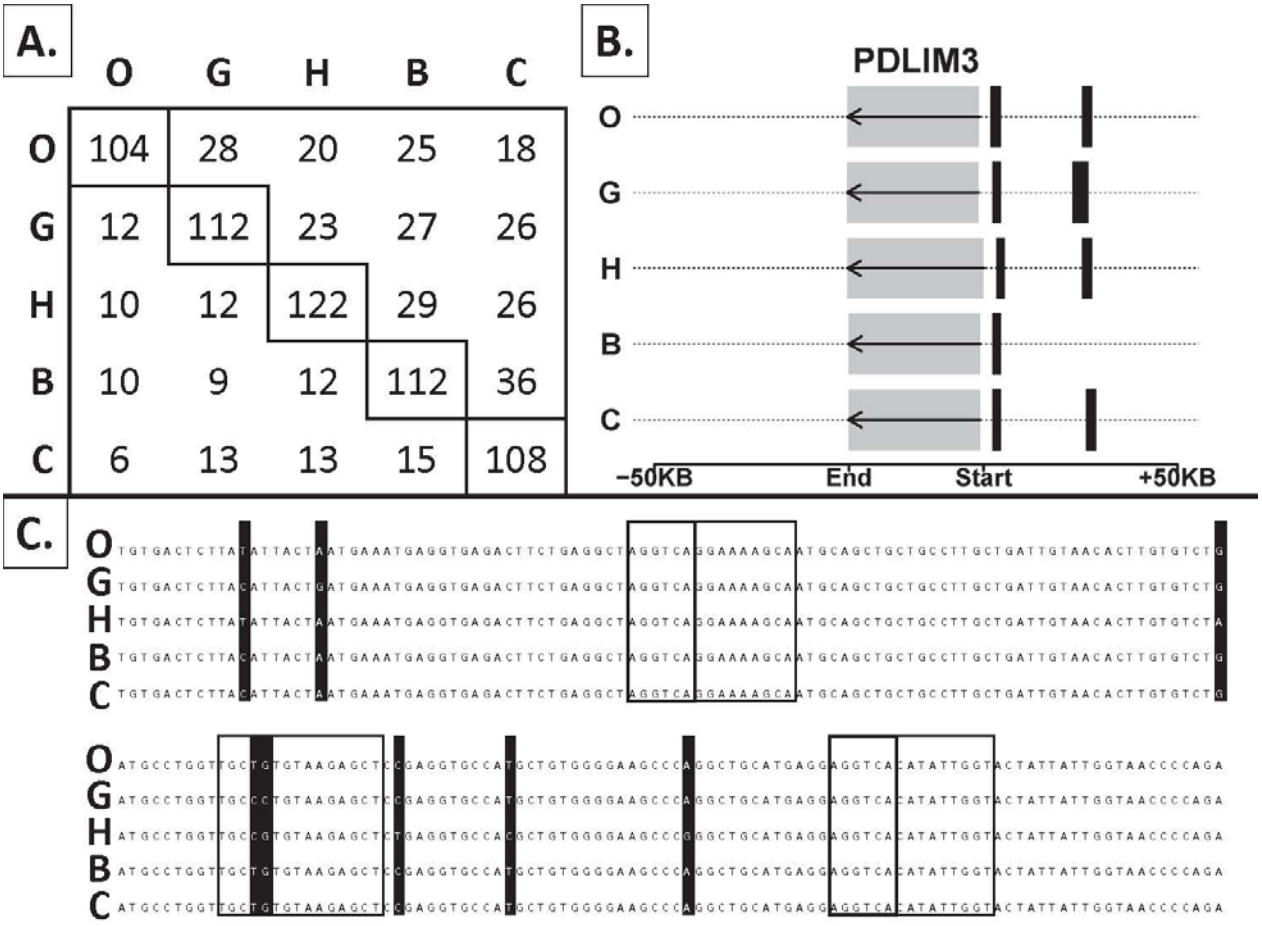

Figure 10: Comparative analysis of five hominid primates using known estrogen responsive genes in humans—*Pongo pygmaeus* (O-orangutan), *Gorilla gorilla* (G-gorilla), *Homo sapiens* (H-human), *Pan paniscus* (B-bonobo), and *Pan troglodytes* (C-chimpanzee). A. Below diagonal: Number of genes in species pair with shared number of ER-binding regions 50kb upstream and downstream, as well as in intragenic region (Total genes tested = 50). Above diagonal: Number of genes in species pair with difference of one ER-binding region either 50kb upstream, 50kb downstream, or in intragenic region (Total genes tested = 50). Diagonal: Total number of ER-binding regions for all 50 genes by species. B. Example gene, *PDLIM3*, with strong conservation of ER-binding regions across species. Gray block indicates gene, black boxes represent ER-binding regions, and arrow indicates read direction. C. 200bp piece found in the most proximal ER-binding region from alignment of 50kb upstream region of *PDLIM3*. Highlighted column indicates substitution, boxes represent predicted ER-binding with sub-boxes indicating a perfect ERE half-site.

While my methods for estimating regions of interest were simplified demonstrations of EREFinder's utility, users could choose their own methods for detection that are best suited for their needs (e.g., Affinity Density as described by Hazelett et al. 2009). For instance, a user could choose to focus on the number of perfect half-sites within a window or locate perfect EREs in the sequence. EREFinder can be used both for *a priori* searches for ER-binding in a given sequence using larger window sizes and for *post hoc* searches around a gene of interest using the base pair score; both of these methods were demonstrated in my comparative approach examples. These different types of searches facilitate the identification of regions of interest for hypothesis testing, comparative approaches, or evolutionary inquiries.

It is important to note that EREFinder is like other motif search algorithms in that it will identify many regions that could bind with the transcription factor despite the large share of them not binding *in vivo* (Wasserman and Sandelin 2004). As such, regions of high affinity should be interpreted only as a putative location for ER-binding. More information would be required if users desired prediction of ER-binding, as other cofactors, particularly FOXA1 (Carroll et al. 2005) in ER-binding (Hah et al. 2013), are necessary for completion of an active *cis*-regulatory module (Glass and Rosenfeld 2000, Shang et al. 2000, Suryamahon and Halfon 2014). As demonstrated with my comparison of genetic and ChIP-PET regions, matching coordinates is fairly straight-forward, making this kind of technique easy to implement for a variety of inquiries into other coupled factors. Epigenetic considerations can also refine predictive power by taking DNA shape (Yang and Ramsey 2015) and histone modification (Cuellar-Partida et al.

2011) into account. In addition, the interaction between ERα and ERβ can result in various expression patterns of up- and down-regulation, depending on the specific receptor that binds to the ERE (for review see Matthews and Gustafsson 2003).

Nonetheless, EREFinder did indicate high ER-binding in known ER-binding regions, demonstrating its ability to correctly identify putative binding capability. With the variety of different factors that can affect expression and ER-binding, it is important to consider that the primary function of EREFinder is to locate regions that have a sequence that could potentially bind an ER, without consideration of any other elements that may or may not be present on the given sequence. This potential is useful in a comparative context as it could represent regions that are active at unstudied ontogenetic stages or tissue types, or represent a possible gain or loss of ER binding that might be important across taxa. As I demonstrated with hominid primates, ER-binding regions are generally conserved and documenting changes in ER-binding would be a good starting point to identify genes of interest for a comparative study. A change in ER-binding could lead to transcriptional changes, thereby affecting how selection may act differently between species. Thus, the most useful application of EREFinder is probably in comparative studies of the evolution of ERE loss and gain among populations within a species or closely related species.

EREs represent the physical location on the genome for ER-binding, making EREs essential to the molecular process of ER gene activation and a useful point in the genome to search for putative ER-binding. While the presence of EREs is not necessarily a definitive indication of ER-binding, using EREFinder in conjunction with

other empirical and computational methods could prove useful for a variety of research questions. EREFinder demonstrates an ability to locate regions of high potential ER-binding in a given sequence quickly and with an output that allows users to manipulate the data in whatever manner they see fit.

Data Accessibility Statement

Source code and binaries freely available for download at https://github.com/AndersonDrew/EREfinder, implemented in C++ and supported on Linux and MS Windows. R scripts can be found at https://github.com/AndersonDrew/EREfinder

Dryad link to code at https://datadryad.org/handle/10255/dryad.196410

Literature Cited

Anstead, G. M., Carlson, K. E., and J. A. Katzenellenbogen (1997). The estradiol pharmacophore: Ligand structure-estrogen receptor bining affinity relationships and a model for the receptor binding site. *Steroids, 62,* 268-303.

Arnal, J.-F., Lenfant, F., Metivier, R., Fouriot, G.,Henrion, D.,Adlanmerini, M., Fontaine, C., Gourdy, P., Chambon, P., Katzenellenbogen, B., and J. Katzenellenbogen (2017). Membrane and nuclear estrogen receptor alpha actions: From tissue specficificit to medical implications. *Physiology Review, 97,* 1045-1087.

Bailey, T.L., Williams, N., Misleh, C., and W.W. Li (2006). MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Research, 34,* W369-W373.

Bourdeau, V., Deschenes, J., Metivier, R., Nagai, Y., Nguyen, D., Bretschneider, N., Gannon, F., White, J.H., and S. Mader (2004). Genome-wide identification of high-affinity estrogen response elements in human and mouse. *Molecular Endocrinology, 18*(6), 1411-1427.

Boyer, M., Poujol, N., Margeat, E., and C.A. Royer (2000). Quantitative characterization of the interaction between purified human estrogen receptor alpha and DNA using fluorescence anisotropy. *Nucleic Acids Research, 28*(13), 2494-2502.

Bulyk, M.L., Johnson, P.L.F., and G.M. Church (2002). Nucleotides of transcription factor binding sites exert interdependent effects on the binding affinities of transcription factors. *Nucleic Acids Research, 30*(5), 1255-1261.

Callard, G.V., Tchoudakova, A.V., Kishida, M., and E. Wood (2001). Differential tissue distribution, developmental programming, estrogen regulationand promoter characteristics of *cyp19* genes in teleost fish. *The Journal of Steroid Biochemistry and Molecular Biology, 79*, 305-314.

Carroll, J. S., Liu, X.S., Brodsky, A.S., Li, W., Meyer, C.A., Szary, A.J., Eeckhoute, J., Shao, W., Hestermann, E.V., Geistlinger, T.R., Fox, E.A., Silver, P.A., and M. Brown (2005). Chromosome-wise mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell, 122,* 33-43.

Carroll, J. S., Meyer, C.A., Song, J., Li, W., Geistlinger, T.R., Eeckhoute, J., Brodsky, A.S., Keeton, E.K., Fertuck, K.C., Hall, G.F., Wang, Q., Bekiranov, S., Sementchenko, V., Fox, E.A., Silver, P.A., Gingeras, T.R., Liu, X.S., and M. Brown (2006). Genome-wide analysis of estrogen receptor binding sites. *Nature Genetics, 38*(11), 1289-1297.

Charif, D. and J.R. Lobry (2007). Seqin{R} 1.0-2: a contributed package to the {R} project for statistical computing devoted to biological sequences retrieval and analysis. In *Structural Approaches to Sequence Evolution: Molecules, Networks, Populations* (pp. 207-232). New York, NY: Springer Verlag.

Cuellar-Parida, G., Buske, F.A., McLeay, R.C., Whitington, T., Nobel, W.S., and T.L. Bailey (2012). Epigenetic priors for identifying active transcription factor binding sites. *Bioinformatics, 28*(1), 56-62.

Deegan, B. J., Bhat, V., Seldeen, K.L., McDonald, C.B., and A. Farooq (2011). Genetic variation within the ERE motif modulate plasticity and energetics of binding of DNA to the ERalpha receptor. *Archives of Biochemistry and Biophysics, 507,* 262-270.

Frith, M. C., Li, M.C., and Z. Weng (2003). Cluster-Buster: finding dense clusters of motifs in DNA sequences. *Nucleic Acids Research, 31*(13), 3666-3668.

Geserick, C., Meyer, H.A., and B. Haendler (2005). The role of DNA response elements as allosteric modulators of steroid receptor function. *Molecular and Cellular Endocrinology, 236,* 1-7.

Glass, C. K., and M.G. Rosenfeld (2000). The coregulator exchange in transcriptional functions of nuclear receptors. *Genes and Development, 14,* 121-141.

Hah, N., Murakami, S., Nagari, A., Danko, C.G., and W.L. Kraus (2013). Enhancer transcripts mark active estrogen receptor binding sites. *Genome Research, 23,* 1210-1223.

Hazelett, D.J., Lakeland, D.L., and J.B. Weiss (2009). Affinity Density: a novel genomic approach to the identification of transcription factor regulatory targets. *Bioinformatics, 25*(13), 1617-1624.

Jia, M., Dahlman-Wright, K., and J-A. Gustafsson (2015). Estrogen receptor alpha and beta in health and disease. *Best Practice and Research Clinical Endocrinology and Metabolism, 29*, 557-568.

Johnson, D.S., Mortazavi, A., Myers, R.M., and B. Wold (2007). Genome-wide mapping of in vivo protein-DNA interactions. *Science 316,* 1497-1502.

Kato, K. and D.M. Standley (2013). MAFFT Multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution, 30*(4), 772–780.

Kato, S., Tora, L., Yamauchi, J., Masushige, S., Bellard, M., and P. Chambon (1992). A Far upstream estrogen response element of the ovalbumin gene contains several half-palindromic 5'-TGACC-3' motifs acting synergistically. *Cell, 68,* 731-742.

Kel, A. E., GoBling, E., Reuter, I., Cheremushkin, E., Kel-Margoulis, O.V., and E. Wingender (2003). MATCH: a tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Research, 31*(13), 3576-3579.

Klein-Hitpass, L., Ryffel, G.U., Heitlinger, E., and A.C.B. Cato (1988). A 13bp palindrome is a functional estrogen responsive element and interacts specifically with estrogen receptor. *Nucleic Acids Research, 16*(2), 647-663.

Klinge, C. M., Jernigan, S.C., Smith, S.L., Tyulmenkov, V.V., and P.C. Kulakosky (2000). Estrogen response element sequence impacts the conformation and transcriptional activity of estrogen receptor alpha. *Molecular and Cellular Endocrinology, 174*, 151-166.

Lin, C.-Y., Vega, V.B., Zhang, T., Kong, S.L., Xie, M., Chiu, K.P., Lipovich, L., Barnett, D.H., Stossi, F., Yeo, A., George, J., Kuznetsov, V.A., Lee, Y.K., Charn, T.H., Palanisamy, N., Miller, L.D., Cheung, E., Katzenellenbogen, B.S., Ruan, Y., Bourque, G., Wie, C-L., and E.T. Liu (2007). Whole-genome carography of estrogen receptor alpha binding sites. *PLoS Genetics, 3*(6), 867-885.

Man, T-K., and G.D. Stormo (2001). Non-independence of Mnt repressor-operator interaction determined by a new quantitative multiple fluorescence relative affinity (QuMFRA) assay. *Nucleic Acids Research, 29*(12), 2471-2478.

Martinez, E., and W. Wahli (1989). Cooperative binding of estroge nreceptor to imperfect estrogen-responsive DNA elements correlates with their synergistic hormone-dependent enhancer activity. *The EMBO Journal, 8*(12), 3781-3791.

Matthews, J., and J.-A. Gustafsson (2003). Estrogen Signaling: A subtle balance between ERalpha ERbeta. *Molecular Interventions, 3*(5), 281-292.

Matys, V., Fricke, E., Geffers, R., GoBling, E., Haubrock, M., Hehl, R., Hornischer, K., Karas, D., Kel, A.E., Kel-Margoulis, O.V., Kloos, D-U., Land, S., Lewicki-Potapov, B., Michael, H., Munch, R., Reuter, I., Rotert, S., Saxel, H., Scheeer, M., Thiele, S., and E. Wingender (2003). TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Research, 31*(1), 374-378.

McDonnel, D. P. a. J. D. Norris. (2002). Connections and regulation of the human estrogen receptor. *Science, 296*, 1642-1644.

Mukherjee, S., Berger, M.F., Jona, G., Wang, X.S., Muzzey, D., Snyder, M., Young, R.A., and M.L. Bulyk (2004). Rapid analysis of the DNA-binding specficities of transcription factors with DNA microarrays. *Nature Genetics, 36*(12), 1331-1339.

Ogawa, N. and M.D. Biggin (2012). High-throughput SELEX determination of DNa sequences bound by transcription factors *in vitro*. *Methods Molecular Biology, 786*, 51-63.

Omidi, S., Zavolan, M., Pachkov, M. Breda, J., Berger, S., and E. van Nimwegen (2017). Automated incorporation of pairwise dependency in transcription factor binding site prediction using dinucleotide weight tensors. *PLoS Computational Biology, 13*(7), e1005176.

Pages, H., Aboyoun, P., Gentleman, R., and S. DebRoy (2018). Biostrings: Efficient manipulation of biological strings. R package version 2.48.0

Shang, Y., Hu, X., DiRenzo, J., Lazar, M.A., and M. Brown (2000). Cofactor dynamcs and sufficiency in estrogen receptor-regulated transcription. *Cell, 103*, 843-852.

Stormo, G.D. (2000). DNA binding sites: representation and discovery. *Bioinformatics, 16*(1), 16-23.

Suryamohan, K. and M.S. Halfon (2015). Identifying transcriptional *cis*-regulatory modules in animal genomes. *WIREs Developmental Biology, 4*, 59-54.

Tan, G. and B. Lenhard (2016). TFBSTools: an R/Bioconductor package for transcription factor binding site analysis. *Bioinformatics, 32*(10), 1555-1556.

Teytelman, L., Thurtle, D.M., Rine, J., and A. van Oudenaarden (2013). Highly expressed loci are vulnerable to misleading ChIP localization of multiple unrelated proteins. *PNAS, 110*(46), 18602-18607.

Thornton, J. W., Need, E., and D. Crews (2003). Resurrecting the ancestral steroid receptor: ancient origin of estrogen signaling. *Science, 301*, 1714-1717.

Tohyama, S., Miyagawa, S., Lange, A., Ogino, Y., Mizutani, T., Ihara, M., Tanaka, H., Tatarazako, N., Kobayashi, T., Tyler, C.R., and T. Iguchi. Evolution of estrogen receptors in ray-finned fish and their comparative responses to estrogenic substances. *The Journal of Steroid Biochemistry and Molecular Biology, 158*, 189-197.

Tyulmenkov, V. V., and C. M. Klinge (2001). A mathematical approach to predict the affinity of estrogen receptors alpha and beta binding to DNA. *Molecular and Cellular Endocrinology, 182*, 109-119.

Young, L.J., and D. Crews (1995). Comparative neuroendocrinology of steroid receptor gene expression and regulation: relationship to physiology and behavior. *Trends in Endocrinology and Metabolism, 6*, 317-323.

Wang, Z., Martins, A.L., and C.G. Danko (2016). RTFBSDB: an integrated framework for transcription factor binding site analysis. *Bioinformatics, 32*(19), 3024-3026.

Wasserman, W.W. and A. Sandelin (2004). Applied bioinformatics for the identification of regulatory elements. *Nature Reviews Genetics, 5*, 276-287.

Yang, J. and S.A. Ramsey (2015). A DNA shape-based regulatory score improves position-weight matrix-based recognition of transcription factor binding sites. *Bioinformatics, 31*(21), 3445-3450.

CHAPTER IV

ROLE OF ESTROGEN RESPONSE ELEMENTS IN THE EXPRESSION OF GENES RELATED TO PATTERNING AND GROWTH FOUND IN SECONDARY SEX TRAITS IN SEX-ROLE-REVERSED GULF PIPEFISH

Synopsis

Sexual dimorphism often results from hormonally regulated trait differences between the sexes. In sex-role-reversed vertebrates, females often have ornaments used in mating competition and these traits are expected to be under hormonal control. Here, I take advantage of the fact that males of the sex-role-reversed Gulf pipefish (*Syngnathus scovelli*) develop female-typical traits when they are exposed to estrogens. By comparing control males, estrogen-exposed males, and females, I can potentially identify genes whose expression levels change during the development and maintenance of female-specific ornaments. I used RNA-sequencing to investigate the transcriptome of the sexually dimorphic ornament of the Gulf pipefish. I compared expression patterns of control males to those of estrogen-exposed males, estrogen-exposed females, and control females. My results identified a number of genes that differed among the sexes and confirmed that many of these were estrogen responsive. These estrogen-influenced loci included genes involved in cellular arrangement and muscular growth, as well as genes involved in anaerobic respiration and adipose tissue maintenance. These genes may be involved in the arrangement of chromatophores for color patterning, as well as in the growth of muscles to achieve the greater body depth typical of females in this species. In addition, anaerobic respiration and adipose tissue could be involved in the rigors of

female courtship and mating competition. A detailed analysis of estrogen-receptor binding sites also demonstrates that estrogen-regulated genes tend to have nearby *cis*-regulatory elements, a pattern which may explain their estrogen responsiveness. Overall, this study generates a number of interesting hypotheses regarding the genetic basis of a female ornament in a sex-role-reversed pipefish.

Introduction

Sexual dimorphism and its interplay with sexual selection is a well-known phenomenon that has been studied since the earliest days of evolutionary theory (Darwin 1871). An important cause of dimorphism is sexually antagonistic selection, in which different fitness optima for the sexes put the evolutionary interests of males and females at odds with each other (Parker 1979, Chapman et al. 2003, Connallon and Clark 2014). Males and females of the same species share anywhere from the majority to entirety of their genome, yet loci that benefit one sex may be detrimental to the other, a situation called intralocus sexual conflict (Lande 1980, Rice 1984, Bonduriansky and Chenoweth 2009). A common resolution to intralocus conflict is the differential expression of genes between the sexes, which is often facilitated by sex-biased hormones (Connallon and Knowles 2005, Mank 2017, Wright et al. 2018). In vertebrates, sex-biased hormones, which are produced following the development of the gonads, lead to numerous differences between males and females. Traits in males such as large body size (Cox et al. 2009), weaponry (Bartos et al. 2012), ornamentation (Lindsay et al. 2016), and

behavior (Ghosal and Sorensen 2016) have been shown to exhibit some form of androgen sensitivity.

While there is strong support for the importance of male sex hormones in the regulation of secondary sex traits in species with conventional sex roles, the function of sex hormones in sex-role reversed systems is a bit more ambiguous. Sex-role reversal is often defined as the situation in which sexual selection acts more strongly on females than on males, and it typically coincides with substantial male investment in offspring. In sex-role-reversed species, the females often evolve ornamentation and compete with one another for mating access to males (Trivers 1972, Williams 1975, Andersson 1994). Studies on the endocrinology of sex-role reversal have been mostly confined to birds but a handful of studies have investigated various species of fish. In reversed taxa, the classic sex-biased hormones, androgens and estrogens, tend to keep their traditional levels in their respective sexes, although in some cases females may reach similar levels of testosterone as males (Eens and Pinxten 2000). For example, in the sex-role-reversed black coucal bird, females express a higher level of androgen receptors in the brain, which could lead to greater testosterone sensitivity (Voigt and Goymann 2007). Females of the buttonquail, another sex-role-reversed bird, demonstrate increased ornament size with elevated circulating testosterone, despite lower levels of this hormone than males (Muck and Goymann 2011). Hormones other than estrogens and androgens can also show reversals, with progesterone modulating female aggression in the black coucal (Goymann et al. 2008) and prolactin apparently related to sex-role reversal in the spotted sandpiper (Oring et al. 1986) and Wilson's phalarope (Oring et al. 1988). In the

few studies of sex-role-reversed fishes, female-typical hormones sometimes have a clear role in the development secondary sex traits. For instance, estradiol has been found to be involved in female courtship and mating behavior in the peacock blenny (Gonçalves et al. 2014), as well as in the regulation of female ornamentation in Gulf pipefish (Partridge et al. 2010). In contrast, the two-spot goby, which shows a reversal of sex roles during part of the breeding season, possesses female nuptial coloration, and this coloration is controlled by prolactin rather than estrogens (Sköld et al. 2008).

Syngnathid fishes (seahorses, pipefishes, and seadragons) are an excellent group for investigation of the effects of estrogens on secondary sex traits in sex-role reversed species. Many species of syngnathids are sex-role reversed, with females experiencing stronger sexual selection than males (Berglund et al. 1986b, Vincent et al. 1992, Jones and Avise 1997). In species with female ornamentation, ornaments are subject to male mate choice and are correlated with size (Berglund et al. 1986a, Flanagan et al. 2014) thereby serving as possible honest signals fertility? (Berglund et al. 1997, Cunha et al. 2017, Mobley et al. 2018). Ornamented female syngnathids often engage in courtship rituals, with males preferring females with greater courtship effort (Partridge et al. 2013). Additionally, strongly sexually selected females are polyandrous (Jones and Avise 2001, Rose et al. 2013a, Flanagan et al. 2014) indicating long-term mating effort is likely higher in females than in males. Studies on the Gulf pipefish, *Syngnathus scovelli*, have revealed that ornamentation is estrogen-mediated (Partridge et al. 2010, Rose et al. 2013b). Female *S. scovelli* have deeply keeled abdomens and transverse iridescent bands (Jones and Avise 1997, 2001). When males are exposed to 17α-

ethinylestradiol (EE2) they develop a deeper abdomen and a banding pattern similar to females that is not present in unexposed males (Partridge et al. 2010). Interestingly, EE2 does not induce behavioral changes in male pipefish (Sária et al. 2013).

The connection from hormone to trait entails a genetic underpinning, as hormonally-mediated expression is the proximate cause of the observed traits. Given the evidence of estrogen-induced banding and body depth in male *S. scovelli*, it would be expected that genes involved in the development of those traits would have sexually dimorphic expression profiles. It is unknown what genes or cell types take part in the banding pattern in *S. scovelli*, but it is likely that iridophores are involved. Iridophores are a type of pigmentation cell (chromatophore) that creates shiny silver coloration, among other colors, by layering guanine plates within the cell to reflect incoming light (Lythgoe et al. 1984). Other types of chromatophores include melanophores and xanthophores, which produce dark browns to yellows and reds (Ligon and McCartney 2016). In terms of coloration, male and female *S. scovelli* have similar patterns with the exception of the iridescent bands, making it unlikely that genes involved in coloration outside of the banding pattern would be differentially expressed between the sexes. The previously mentioned chromatophores have well-documented gene pathways, interactions, and development (Kelsh et al. 2009, Budi et al. 2011, Patterson and Parichy 2013) that are most often mediated by melanin secreting hormone (MSH) (Sköld et al. 2016). It is also unclear what genes are involved in body depth, but the two expected tissue types involved would be muscle and adipose tissue. Muscle tissue of exercised fishes shows an increase in muscle mass maintenance genes and myogenic genes (Garcia

de la serrana et al. 2012, Palstra et al. 2013, Palstra et al. 2014). Adipose tissue deposits are often used for energy storage and can be found in various locations on fishes that show unique expression profiles depending on various intrinsic and extrinsic factors (Imrie and Sadler 2010, Weil et al. 2013).

An additional bridge from hormone to transcript involves the physical binding of the hormone nuclear receptor to the genome. The estrogen response element (ERE) is a *cis*-regulatory factor composed of two palindromic half-sites (i.e., AGGTCAnnnTGACCT) that are preferentially bound by the estrogen receptor (Klein-Hitpass et al. 1988, Boyer et al. 2000). Substitutions to the canonical ERE affect binding and transcriptional responses to estrogens (Tyulmenkov and Klinge 2001, Geserick et al. 2005, Deegan et al. 2011). Genes that respond to estrogen typically have EREs in greater numbers (Martinez and Wahli 1989, Kato et al. 1995, Geserick et al. 2005) and in closer proximity (Carroll et al. 2006, Lin et al. 2007) compared to non-responsive genes. While the presence of an ERE is not diagnostic of an estrogen-responsive gene (Wasserman and Sandelin 2004), a population of estrogen-responsive genes is nevertheless expected to show an increase in proximal EREs when compared to genes that are not estrogen responsive. Therefore, at the genomic level I should see an elevation in number of EREs near genes that show transcriptional responses to the presence of estrogen.

With an annotated genome available (Small et al. 2016), *S. scovelli* is one the few sex-role-reversed species that lends itself to transcriptomic work to connect hormonal mediation to gene expression to morphological traits. Male and female transcriptomes have already been compared for livers (Rose et al. 2015), brains (Beal et al. 2018), and

brood pouches (Small et al. 2013). Here, I investigate transcriptional patterns in sexually dimorphic skin and muscle tissue of *S. scovelli*. I seek to elucidate genes with sexually dimorphic, estrogen-mediated expression in the skin and muscle tissue where the female ornament forms. These genes are putative candidates for regulation of the differences in ornamentation and morphology that distinguish male and female *S. scovelli* from one another. By comparing groups of estrogen exposed and non-exposed male and females, I aim to link estrogen-induced genes and estrogen-induced sexually dimorphic traits in the muscle and skin. I hypothesize that genes involved in iridophore production and muscle mass should show an increase in expression in both estrogen-exposed males and non-exposed females compared to non-exposed males. Given the tendency of EREs to be proximal to estrogen-mediated genes, I also hypothesize that estrogen-induced genes should have higher densities of proximal EREs compared to genes that are not estrogen regulated.

Methods

Experimental Design

This study used the same Gulf pipefish specimens as described in Rose et al. (2015). All fish were preserved in RNA*later*™ (ThermoFisher Scientific) at -80°C, and they were separated into four groups of five individuals each based on EE2 exposure treatments: Control non-brooding adult males (control males), estrogen-exposed non-brooding adult males (EE2 males), control adult females (control females), and estrogen-exposed adult females (EE2 females). Collection of specimens and estrogen exposure techniques are detailed in Rose et al. (2015). I used the same sequenced individuals that

had their livers dissected out and were exposed to a concentration of 5ng/L of EE2. To obtain enough tissue for RNA sequencing I took a cross-section of the eviscerated individual that was wide enough to include two ornamental bands on each side. This cross-section included muscle, skin, and bone tissue together, but no other organs. Total RNA was extracted and sent to Michigan State University RTSF Genomics Core for library prep and Illumina HiSeq 2500 sequencing following Rose et al. (2015). A separate, bar-coded library was prepared for each individual used in this study (that is, RNA was not pooled across specimens).

Differential Expression Analysis

A total of 660 million 125bp paired-end reads were sequenced and delivered from the MSU Genomics core. Reads were trimmed using Trimmomatic (Bolger et al. 2014) and visualized using FastQC (Andrews 2010) to verify quality of trimmed reads. HISAT2 (Kim et al. 2015) was used to align forward and reverse reads to the genome provided by Small et al. (2016). Mapped reads were processed with Cufflinks (Trapnell et al. 2012) with all individual assemblies and the reference gtf file merged using cuffmerge. The merged gtf file and mapped reads were read into R using featureCounts in Rsubread (Smyth et al. 2013). Using edgeR (Robinson et al. 2010) reads were filtered to ensure all five individuals in any one group showed expression of that transcript. Reads were then normalized and grouped by condition and sex. Using the exactTest function all six pairwise comparisons were made and merged into a final table for analysis. Since edgeR provides normalized counts per million (CPM) I utilized the raw count data and a customized R script that found the longest open reading frame to

estimate the transcript per million (TPM) for visualization of data. For transcripts that did not have a match to the original genome I used Blast2GO (Conesa et al. 2005) to determine the blastp hit when possible; otherwise I report blastp hit from Small et al. (2016).

Since I am interested in genes showing an estrogen response in males, I investigated all three pairwise comparisons to control males. I consider a false discovery rate (FDR) of less than 0.05 as a statistically significant value for a pairwise expression difference. I further filtered results based on biological significance by requiring at least one group in the pairwise comparison to have a mean TPM greater than 1. I consider transcripts that are estrogen-influenced in males to be those that show significant expression differences between control and EE2 males. Transcripts that are mostly sex-influenced are those that have significant expression differences between control males and both control and EE2 females. I consider estrogen-influenced with sex-bias to be those transcripts with significant expression differences between control males and the three other groups. The remaining three pairwise comparisons were only used to investigate deviations from expected patterns (e.g., differential expression between control female and EE2 female for sex-biased genes).

Investigation of Differential Expression Patterns

I further identified the differential expression patterns of known pigmentation genes. From the literature I identified 53 genes with a demonstrable causal role in coloration in fishes. Using a customized R script, I blasted the nucleotide sequences of these genes from multiple fish species against the *S. scovelli* and *Hippocampus comes*

(tiger-tail seahorse) genomes (Lin et al. 2016) to confirm the presence of each gene in these Syngnathidae taxa. Matches to *S. scovelli* were visually inspected to ascertain the best match. The locus under consideration was then identified in my transcriptome data to produce a list of transcripts putatively associated with a pigmentation pathway.

Lastly, I investigated the roles EREs may play in differential expression patterns in *S. scovelli* and in other fishes. Using EREFinder I performed a whole genome, sliding window scan of mean estrogen receptor binding affinity. I used a window size of 100bp with 50bp intervals. With a customized R script I found peaks of high estrogen binding (mean K_d^{-1} of window ≥ 0.0045) across the genome from the EREFinder output. The selected cutoff value was 5.7 standard deviations above the mean K_d^{-1} of the entire genome, thereby selecting windows of extremely high binding compared to the background. I then took the coordinates of differentially expressed genes and matched them to any high estrogen binding coordinates within 25kb upstream and 10kb downstream of the transcription start site of the gene. For the estrogen-influenced genes I determined the number of estrogen binding regions within my selected range and the number of genes with any estrogen binding within my selected range. To test if either the number of peaks or genes exceeded a random sampling, I sampled an equal number of randomly chosen annotated genes 1,000 times and determined where my gene set fell on the resulting distribution. I also did this for sex-biased only genes.

Results

Sequencing of reads across all four experimental groups generated 17,360 transcripts of which 2,982 did not have an identifier from the published genome. The most highly

expressed genes were those associated with muscle tissue, and a few are associated with skin and skeletal tissue (Table 2). Seventeen transcripts showed a significant fold change between control males and EE2 males, of which two were down-regulated (Figure 11). None of the transcripts showed a $\log_2$ fold change of greater than 10 between control and EE2 males. Of these 17 transcripts, 11 also had significant fold changes between control males and both female groups, all of which were up-regulated compared to control males. The blast results obtained from the annotation file or my blasted hits allowed us to place transcripts into three categories: those that blasted to genes typical to hepatic/adipose tissues (seven), those that blasted to genes typical to muscle tissue (three), and those that blasted to a cellular organization gene which can be found in skin tissue (one; Table 3). The six transcripts that were uniquely estrogen-influenced in males blasted to muscle development and structural genes. EE2 males had similar expression levels to females for eight of the 11 genes with shared differential expression. The exceptions to this pattern were transcripts identified as vitellogenin or choriogenin.

Table 2: Most highly expressed genes for control males and females, as determined by mean transcript per million (TPM) for the five samples of each treatment. Gene ID refers to either identifier in annotated genome (Small et al. 2016) with prefix "SSCG" or transcript ID from this study that did not match an annotated region with prefix "XLOC". Bolded genes had significant differential expression (FDR < 0.05) between control males and control females. Genes with asterix (*) are found in top 25 expressed genes in both sexes.

Top 25 Expressed Female Genes	TPM	Gene ID	Top 25 Expressed Male Genes	TPM	Gene ID
PREDICTED: C-type mannose receptor 2-like*	1118	SSCG00000009477	lectin protein type III*	3189	SSCG00000009491
lectin protein type III*	1105	SSCG00000009491	glyceraldehyde-3-phosphate dehydrogenase*	1566	SSCG00000018902
glyceraldehyde-3-phosphate dehydrogenase*	911	SSCG00000018902	PREDICTED: actin, alpha skeletal muscle A*	1220	SSCG00000002073
PREDICTED: RNA-binding protein 5-like*	907	SSCG00000000622	creatine kinase M-type-like*	1216	XLOC_010522
PREDICTED: actin, alpha skeletal muscle A*	884	SSCG00000002073	PREDICTED: RNA-binding protein 5-like*	976	SSCG00000000622
creatine kinase M-type-like*	812	XLOC_010522	PREDICTED: C-type mannose receptor 2-like*	933	SSCG00000009477
actin, alpha skeletal muscle*	713	SSCG00000002064	NA*	872	XLOC_008770
beta-enolase*	472	SSCG00000016242	cholesterol side chain cleavage cytochrome P450*	794	SSCG00000015888
NA*	445	XLOC_024549	beta-enolase*	682	SSCG00000016242
cholesterol side chain cleavage cytochrome P450*	410	SSCG00000015888	skeletal muscle fast troponin T isoform 2*	667	SSCG00000019593
NA*	399	XLOC_008770	**creatine kinase M-type***	**643**	**XLOC_003319**
creatine kinase M-type*	**353**	**XLOC_003319**	PREDICTED: lanC-like protein 1 isoform X2*	537	SSCG00000009383
skeletal muscle fast troponin T isoform 2*	314	SSCG00000019593	**Alanyl-tRNA editing protein Aarsd1**	**497**	**SSCG00000013141**
PREDICTED: lanC-like protein 1 isoform X2*	299	SSCG00000009383	actin, alpha skeletal muscle*	471	SSCG00000002064
NA*	277	XLOC_001421	NA*	460	XLOC_024549
CD2 antigen cytoplasmic tail-binding protein 2*	253	SSCG00000018621	CD2 antigen cytoplasmic tail-binding protein 2*	425	SSCG00000018621
tropomyosin alpha-1 chain isoform X9	246	SSCG00000019708	**collagen alpha-1(I) chain-like**	**408**	**XLOC_024075**
PREDICTED: cytochrome c oxidase subunit 6B1*	244	SSCG000000179591	NA*	401	XLOC_016642
Osteocalcin	242	XLOC_008578	PREDICTED: myomesin-3	389	SSCG00000018900
keratin, type I cytoskeletal 19-like	235	SSCG00000008212	PREDICTED: cytochrome c oxidase subunit 6B1*	379	SSCG000000179591
PREDICTED: protein FAM72A*	230	SSCG00000013900	PREDICTED: protein FAM72A*	275	SSCG00000013900
NA*	221	XLOC_016642	myosin heavy chain, fast skeletal muscle-like	242	SSCG00000000630
PREDICTED: syntaxin-16 isoform X4	202	SSCG00000000898	NA*	242	XLOC_001421
keratin, type I cytoskeletal 13	198	SSCG00000008449	PREDICTED: solute carrier family 35 member C2	240	SSCG00000002943
PREDICTED: transcription factor SOX-6-like*	196	SSCG00000003665	PREDICTED: transcription factor SOX-6-like*	224	SSCG00000003665

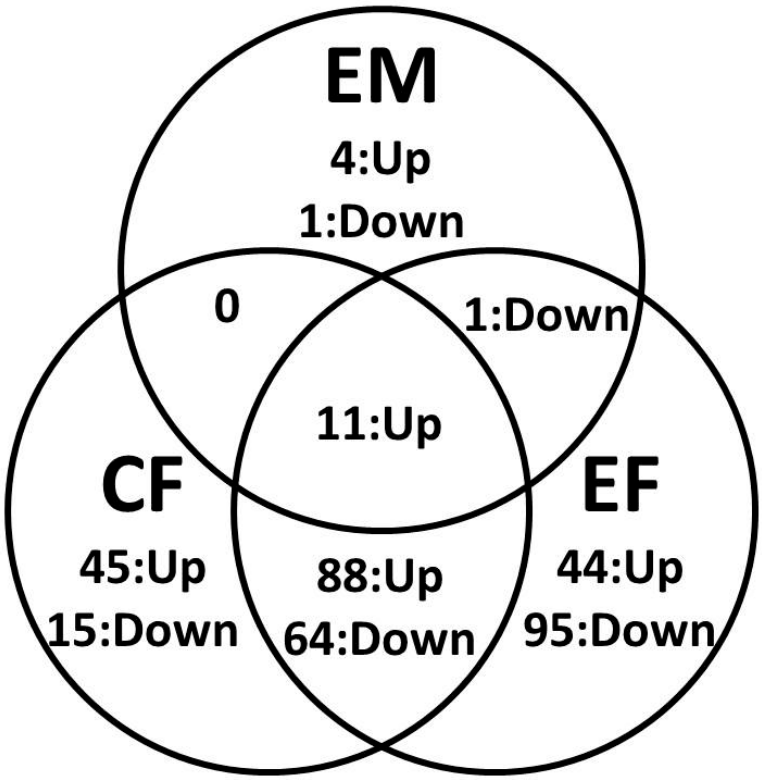

Figure 11: Number of transcripts differentially expressed between three pairwise comparisons. Each circle represents a pairwise comparison of control males to control females (CF), control males to EE2 exposed males (EM), or control males to EE2 exposed females (EF). Overlapping circles indicate shared differential expression in pairwise comparisons. Direction of fold change is relative to control males and all shared expression change directions were in the same direction relative to control males. The 11 genes in the center are putatively estrogen-influenced with sex-bias and may play a role in observed feminization of male phenotype upon exposure to EE2.

Table 3: Transcripts that showed significant expression differences between control males and EE2 exposed males (EE2 Male). Gene ID refers to either identifier in annotated genome (Small et al. 2016) with prefix "SSCG" or transcript ID from this study that did not match an annotated region with prefix "XLOC". Mean transcript per million (TPM) of control male group is shown. The last three columns show the log_2 fold change of EE2 males, control females, and EE2 exposed females relative to control males. A positive fold change indicates lower expression levels in control males; a negative fold change indicates higher expression levels in control males. EE2 male, control female, and EE2 female values shown in bold indicate a significant difference compared to control males (FDR < 0.05).

Gene Description	Gene ID	Control Male TPM	EE2 Male	Control Female	EE2 Female
5β-cholestane-3α,7α-diol 12α hydroxylase	SSCG00000010315	0.10	***9.01***	***8.57***	***8.59***
myosin heavy chain, fast skeletal muscle	XLOC_018346	0.02	***8.14***	2.80	0.58
vitellogenin C [Thunnus thynnus]	SSCG00000005247	0.00	***7.96***	***14.86***	***13.89***
Coagulation factor XIII A chain [Larimichthys crocea]	SSCG00000015268	0.01	***6.78***	***7.31***	***6.62***
PREDICTED: alcohol dehydrogenase 1-like isoform X1	SSCG00000005272	1.44	***5.68***	***4.85***	***4.88***
tropomyosin alpha-1 chain-like isoform X4	SSCG00000019892	0.12	***5.22***	-0.41	-0.61
PREDICTED: bile salt export pump isoform X1	SSCG00000009270	1.07	***4.48***	***3.26***	***3.85***
choriogenin L	SSCG00000003459	0.15	***3.78***	***7.60***	***6.20***
PREDICTED: RNA-binding protein 14 isoform X1	SSCG00000002232	2.54	***3.20***	***2.94***	***2.50***
Vitellogenin	SSCG00000001288	0.22	***3.17***	***10.96***	***9.36***
PREDICTED: 7-dehydrocholesterol reductase	SSCG00000004357	1.84	***2.71***	***2.40***	***2.33***
PREDICTED: lysosome membrane protein 2-like isoform X1	SSCG00000001035	3.13	***2.40***	***1.93***	***2.21***
PREDICTED: SPARC-related modular calcium-binding protein 1 isoform X2	SSCG00000003774	0.54	***2.17***	***2.10***	***2.44***
Rho guanine nucleotide exchange factor 18	SSCG00000015802	0.33	***1.86***	-0.10	-0.59
PREDICTED: FH1/FH2 domain-containing protein 3 isoform X2	SSCG00000016309	0.54	***1.16***	-0.05	0.49
unnamed protein product	SSCG00000010971	28.47	***-1.78***	-0.50	-1.00
PREDICTED: protein S100-G-like	SSCG00000010975	27.13	***-2.12***	-1.05	***-1.38***

The majority of transcripts differentially expressed between control males and females were differentially expressed whether or not females were exposed to EE2 (163 genes fell into this category). Despite pairwise comparisons of control male to control female or EE2 female generating some different transcripts, none of these transcripts (which numbered 363) were significantly different between control and EE2 females. None of the transcripts showed a $\log_2$ fold change between control males and either female greater than 10, with the exception of the two vitellogenin transcripts which were expressed at higher levels in females but still did not exceed a $\log_2$ fold change of 15 (Table 4). Blast2Go was unable to determine functional group for nearly 65% of transcripts, but the top differentially expressed genes suggest an abundance of hepatic/adipose tissue related genes. To confirm I did not accidently include liver tissue in my sample I compared my lists of top expressed and differentially expressed genes to those identified in a study of the *S. scovelli* liver transcriptome (Rose et al. 2015). With the exception of vitellogenin, the most highly expressed and differentially expressed genes in my set are different than those found by Rose et al. (2015) and at lower values of expression differences.

Table 4: Transcripts with the largest expression change between control males and control females. All expression changes shown are significant for both pairwise comparisons of control males to control females and control males to EE2 exposed females. Significant expression changes (FDR < 0.05) between control males and EE2 exposed males are in bold. Gene ID refers to either identifier in annotated genome (Small et al. 2016) with prefix "SSCG" or transcript ID from this study that did not match an annotated region with prefix "XLOC". Mean transcript per million (TPM) of control male group is shown as well as the log_2 fold change relative to EE2 exposed males (EE2 Male), control females, and EE2 exposed females (EE2 Female); a negative fold change indicates higher expression levels in control males**.**

Gene Description	Gene ID	Control Male TPM	EE2 Male	Control Female	EE2 Female
vitellogenin C	SSCG00000005247	0.00	***7.96***	***14.86***	***13.89***
Vitellogenin	SSCG00000001288	0.03	***3.17***	***10.96***	***9.36***
NA	XLOC_023457	0.01	2.18	**10.70**	**9.35**
galactose-specific lectin nattectin-like	XLOC_012482	0.05	8.61	**8.92**	**8.23**
5β-cholestane-3α,7α-diol 12α hydroxylase	SSCG00000010315	0.10	***9.01***	***8.57***	***8.59***
choriogenin L	SSCG00000003459	0.01	***3.78***	***7.60***	***6.20***
Coagulation factor XIII A chain	SSCG00000015268	0.01	***6.78***	***7.31***	***6.62***
vitellogenin Aa	SSCG00000015789	0.16	1.16	**6.33**	**4.87**
PREDICTED: alcohol dehydrogenase 1-like isoform X1	SSCG00000005272	1.44	***5.68***	***4.85***	***4.88***
elastase-1-like	SSCG00000017895	0.32	0.24	**4.65**	**2.51**
PREDICTED: uncharacterized protein	SSCG00000008823	0.31	-0.77	**4.41**	**3.69**
histidine-rich glycoprotein-like	SSCG00000005882	0.19	-0.69	**3.72**	**2.81**
PREDICTED: plasma membrane calcium-transporting ATPase 1-like isoform X5	SSCG00000014800	0.34	3.96	**3.39**	**3.25**
PREDICTED: complement factor H-like	SSCG00000006026	0.21	-0.70	**3.30**	**3.24**
PREDICTED: GRAM domain-containing protein 3-like isoform X3	SSCG00000000578	0.68	3.18	**3.29**	**3.88**
PREDICTED: bile salt export pump isoform X1	SSCG00000009270	0.06	***4.48***	***3.26***	***3.85***
PREDICTED: hemopexin	SSCG00000017720	0.88	-1.05	**3.15**	**2.32**
PREDICTED: RNA-binding protein 14 isoform X1	SSCG00000002232	0.31	***3.20***	***2.94***	***2.50***
hypothetical protein AAES_135341	SSCG00000001172	0.14	0.72	**2.88**	**2.57**
syntaxin-binding protein 6-like	XLOC_007852	2.15	0.87	**-2.99**	**-3.06**
PREDICTED: insulin gene enhancer protein ISL-2	SSCG00000009125	12.93	-0.01	**-3.37**	**-3.08**
PREDICTED: G0/G1 switch protein 2-like	SSCG00000006896	6.61	-2.37	**-3.43**	**-3.15**
PREDICTED: collagenase 3-like	SSCG00000002143	19.25	1.36	**-3.71**	**-4.94**
Immunoglobulin-like and fibronectin type III domain-containing protein 1	SSCG00000012567	3.15	-0.11	**-4.05**	**-3.54**
NA	XLOC_008235	72.07	-0.12	**-4.06**	**-4.25**

Of the 53 pigmentation genes investigated, 51 were successfully annotated in the *S. scovelli* genome, and 36 of those were matched in my transcriptome data set. Locating these genes in my data set found only two had a significant FDR value in any of my pairwise comparisons (Table 5). Both of these genes were up-regulated in control females relative to control males and in EE2 females relative to control males. These transcripts are the genes *tfec* and *atic*, both of which have a putative role in the development and maintenance of iridophores. The transcript matching *tfec* had expression levels below my biological threshold and showed significant up-regulation between EE2 males and both female groups. The other transcript, *atic*, was biologically significant and had no significant expression differences between EE2 males and the other female groups.

Table 5: Genes identified as involved in chromatophore pathways with evidence and citation provided. Gene ID refers to the identifier in annotated genome (Small et al. 2016). Mean transcript per million (TPM) of control male group is shown as well as the log_2 fold change of the EE2 males, control females, and EE2 females relative to control males. Significant expression differences (FDR < 0.05) are shown in bold.

Symbol	Name	Evidence	Citation	Gene ID	Control Male TPM	EE2 Male	Control Female	EE2 Female
atic	Atic	increased expression in iridophore compared to other chromatophores	(Higdon et al. 2013)	SSCG00000013410	1.83	0.45	***1.23***	***1.08***
tfec	Transcription E factor	plays role in regulating iridophore specification	(Petratou et al. 2018)	SSCG00000016281	0.17	0.18	***1.30***	***1.18***

Using the set of 17 genes identified as differentially expressed between control and EE2 males, I matched genomic coordinates of those genes to regions of high estrogen binding. Compared to a random sample of genes, these 17 genes showed an increase in number of proximal estrogen receptor binding regions and 13 of the 17 genes had at least one binding region nearby, which is more than expected (Figure 12). For a comparison of genes not suspected to have estrogen binding, I chose the 152 genes identified as differentially expressed between control males and both female groups but not between control males and EE2 males (Figure 11). Both the number of proximal estrogen receptor binding regions and genes with at least one binding region fell well within the expected distribution from random sampling (Figure 12).

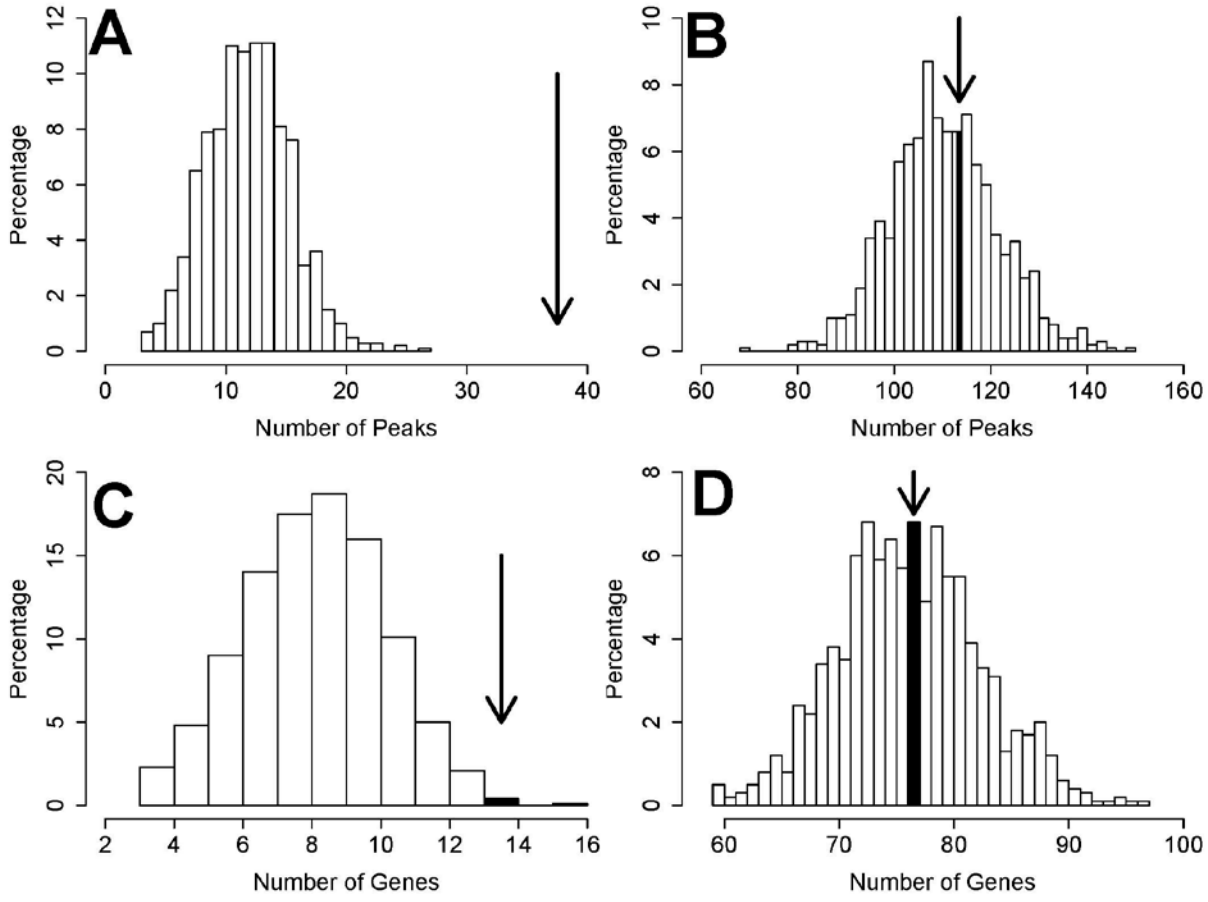

Figure 12: Histogram of number of estrogen responsive peaks identified by EREFinder proximal to 1,000 random gene sets (A and C) and number of genes from 1,000 random sets with at least one estrogen responsive peak (B and D). A and C use a random set size of 17 genes while B and D use a set size of 152 genes. Dark bars and arrows show where on the distribution the set of genes obtained from differential expression fall on the distribution. A and C use the set of 17 genes with differential expression between control and EE2 exposed males. B and C use the set of 152 genes with differential expression between control males and both control and EE2 females but exclude the 11 genes with differential expression between control males and EE2 exposed males (see Figure 11).

Discussion

Pairwise comparisons of the male and female, both control and EE2 exposed, transcriptomes in the phenotypically dimorphic muscle and skin tissue of *S. scovelli* elucidated several genes of interest. I was able to use the annotated genome to delve into the expected functions of these differentially expressed transcripts, allowing us to posit some hypotheses regarding the mechanisms by which these genes contribute to the sex-biased phenotypes. With the identification of iridophore genes from my search of pigmentation genes and the finding of genes involved with cellular arrangement in the dermal layer, I are able to offer some possible genetic hypotheses regarding how banding might occur (see below). Moreover, the differential expression of muscle growth genes between EE2 males and control males suggests some pathways for the development of the female's deeply keeled body. I find the typical markers of estrogen exposure in vitellogenins and coagulation factor 13 commonly found in adipose, skin, and muscle tissue (Wang et al. 2005, Hao et al. 2013, Zhong et al. 2014). The remaining genes from my study might be involved in adipose tissue production or other energetic pathways. The majority of transcriptional sex differences appear to be independent of the female sex-biased hormone estrogen or could possibly require additional cofactors not present in males. Additionally, these expression differences are less pronounced compared to those seen in the more typically dimorphic liver (Zheng et al. 2013, Rose et al. 2015, Qiao et al. 2016). Lastly, I show that the estrogen-induced genes are influenced by local genomic regions of estrogen-receptor binding, which differentiate them from other sex-differentiated genes (i.e., those not directly affected by estrogen exposure).

The most obvious effect of EE2 on *S. scovelli* is the induction in males of the iridescent bands unique to females (Partridge et al. 2010). Surprisingly, no genes associated with pigmentation show a significant transcriptional response to EE2 exposure. Instead, there is a sex difference in expression of *atic* and *tfec*, both of which are implicated in the formation of iridophores (Higdon et al. 2013, Petratou et al. 2018, Salis et al. 2019), although the hormonal control of these genes has not yet been established in fishes. A gene putatively expressed in the skin tissue, which does show estrogen-mediated differences between males and females is *sparc*, a matricellular protein found during development and in tissues with high turnover (Lane and Sage 1994, Brekken and Sage 2000, Bradshaw and Sage 2001). In fishes, *sparc* shows expression in bone (Renn et al. 2006, Weigele et al. 2015) and skin tissue (Karsi et al. 2002, Iimura et al. 2012). The primary function of *sparc* is arranging and distributing cells in the extra-cellular matrix (Tremble et al. 1993, Bradshaw and Sage 2001), and it is up-regulated in cases of melanoma where it increases the motility of melanocytes (Ledda et al. 1997, Robert et al. 2006). Another potential player in coloration is lysosome membrane protein 2 (*scarb2*). Although usually implicated in muscle function (Eskelinen et al. 2003, Zeigler et al. 2014), there is evidence it can dilute melanosomes within melonacytes thereby giving lighter colors to the skin (Navarro et al. 2008, Strömberg et al. 2008). It appears *sparc* (Lehane et al. 1999, Witkiewicz et al. 2010) and *scarb2* (Williams et al. 2007) are both down-regulated by estrogen in other organisms, a pattern in opposition to my results for *S. scovelli*.

The size of the bands generated in estrogen-exposed male *S. scovelli* do not reach the size of female bands, even when males are exposed to EE2 levels 20-fold greater than I used in the present experiment (Partridge et al. 2010). The observation of both *atic* and *tfec* expression in control males suggests that iridophores comprise the dermal layer in normal male, albeit in lower quantities compared to females. Presumably, male iridiphores are also not normally arranged in males in a way that creates the distinct banding pattern typical of females. Indeed, banding patterns would normally involve an interaction of multiple chromatophores. The processes regulating interactions among chromatophores are not well defined beyond some observations related to direct interactions among specific cell types (Singh and Nüsslein-Volhard 2015). Arrangement in the extra-cellular matrix, distribution of pigments within the cell, and the shapes of the chromatophores create most of the observed coloration patterns attributable to chromatophores (Ligon and McCartney 2016, Sköld et al. 2016, Nüsslein-Volhard and Singh 2017). Such arrangements could be mediated by a matricellular protein like *sparc*, but no studies have yet addressed the role of *sparc* in color patterning.

Given the small percentage of skin that comprised my sample tissue and my small sample sizes of five individuals per treatment, I may have had insufficient statistical power to detect a slight induction of genes involved in coloration that respond to estrogen. This possibility seems especially likely for *atic* as the EE2 males had expression levels intermediate between control males and females, and EE2 male expression was significantly different from either group. The link from estrogen to gene to banding pattern is still unclear, but I have now identified two genes involved in

iridophore production, one gene in cellular arrangement, and one gene in melanosome distribution, any of which could play a role in the distinctive sexually differentiated banding pattern in *S. scovelli*.

The four genes uniquely up-regulated by estrogen in males all have a role in muscle growth and tissue organization. Both myosin and tropomyosin are involved in muscle growth in short-term recovery (Garcia de la serrana et al. 2012, Palstra et al. 2014) and long-term growth in fishes (Heeley and Hong 1994, Gauvry and Fauconneau 1996, Mommsen 2001). Rho guanine nucleotide exchange factor 18 (*arhgef18*) plays a role in forming tight junctions between cells as well as cell growth and motility (Blomquist et al. 2000). Consequently, *arhgef18* is found in epithelial cells and plays a role in tissue development and remodeling (Terry et al. 2011, Herder et al. 2013, Balda and Matter 2016). FH1/FH2 domain-containing protein 3 (*FHOD3*) is a type of formin protein that plays a role in actin binding and cellular organization (Westendorf et al. 1999, Zigmond 2004, Higgs 2005). Specifically *FHOD3* is found in the muscles and heart of mammals (Taniguchi et al. 2009, Iskratsch et al. 2010) and has also been found in fishes (Gurgul et al. 2018).

It is likely I do not see an up-regulation of these muscular and structural genes in females because they have already developed their musculature and depth by the end of their maturation. Fish in my experiment were exposed for seven days, so males were likely still undergoing feminization and had not yet reached the phenotypic endpoint when they were sacrificed for analysis. Longer exposure to EE2 causes further deepening of the body in males (pers. comm. E. Rose). What I are possibly observing is

part of the expression profile of a maturing female pipefish. A way to confirm this suspicion would be to produce a time series of the female transcriptome during maturation in *S. scovelli*. The estrogen control and up-regulation of these genes is at odds with expectations for a species without sex-role reversal. While estrogens have a diverse array of effects on muscles (Kadi et al. 2002, Enns and Tiidus 2010, Kitajima and Ono 2016), there is some evidence of the down-regulation of myosin in the presence of estrogen in fishes (Olin and von der Decken 1987).

While there is no apparent up-regulation of muscular and structural genes in females, my data do produce an interesting case of putative regulation of muscle performance in females. Alcohol dehydrogenase (*adh1*), which is up-regulated in EE2 males and both female *S. scovelli* groups, is classically considered a hepatic gene, functioning in the breakdown of alcohols (Pikkarainen and Räihä 1967). In addition, long-term exposure to alcohol can cause up-regulation of *adh1* in muscles as the tissue breaks down (Khayrullin et al. 2016, Kimball and Lang 2018). In fishes there is evidence of anaerobic respiration leading to an end product of ethanol, which is easily excreted, rather than lactate (Shoubridge and Hochachka 1980). In a comparative study, fishes that naturally experience oxygen-poor environments show higher levels of *adh* expression in muscle than close relatives in more oxygen-rich environments (Torres et al. 2012). There is evidence of estrogen up-regulation as well (Qulali et al. 1991). Female *S. scovelli*, which may court and mate with multiple males in a single day, may face greater energetic demands than males. Given that longer and more vigorous courtships lead to more successful matings (Partridge et al. 2013), a female able to recover from anaerobic

activity more quickly would have an advantage over her competition. These considerations lead to the hypothesis that, in addition to species of fish in oxygen-depleted zones having more *adh* expression, fishes that expend a great deal of effort on courtship effort may also have high levels of *adh* expression.

The remaining genes worth highlighting are implicated in adipose tissue, which could serve as energetic reserves for female performance. Both RNA binding protein 14 (*RBM14*) and 7-dehydrocholesterol reductase (*dhcr7*) occur in skin and adipose tissue, but they are engaged in different functions. *RBM14* plays a role in adipose tissue differentiation (Firmin et al. 2017) and is expressed in swordtail, *Xiphophorus maculatus*, skin tissue, where it is up-regulated by exposure to UVB light (Downs 2013). *dhcr7* is the final step in cholesterol catabolism (Mitsche et al. 2015) and is also involved in vitamin D synthesis (Prabhu et al. 2016, Prabhu et al. 2017). While vitamin D production in fishes is still uncertain (Lock et al. 2009) there is some evidence of endogenous production and even production through exposure to blue light (Pierens and Fraser 2015). While it is interesting both of the genes are known to have a response to light exposure, I have no firm explanations for why I see an estrogen-induction or sex difference in these genes.

A bit farther removed from known links to adipose tissues are 5β-cholestane-3α,7α-diol 12α hydroxylase and the bile salt export pump (*abcb11*), which are involved in the production of bile acids, particularly cholic acid, from the catabolism of cholesterol (Hansson and Wikvall 1982, Chiang 2017). While this process takes place largely in the liver, these genes were not among the top expressed genes in the liver

regardless of estrogen exposure (Rose et al. 2015) as they are in my sample. Recent work has found that bile acids serve as signaling molecules, particularly for *TGR5* (Watanabe et al. 2006, Lefebvre et al. 2009, Iguchi et al. 2010) which is involved in adipose tissue formation (Chiang 2017, Velazquez-Villegas et al. 2018). While *abcb11* has been identified in fish, unlike other members of the *abc* gene family it is confined mostly to expression in the liver (Lončar et al. 2010, Ferreira et al. 2014). Along with *RMB14*, it could be that these genes play a role in adipose deposits for female *S. scovelli* as they build energetic reserves for egg production and courtship.

Lastly, I demonstrate that estrogen-responsive genes are more likely to have high estrogen receptor binding than expected by chance. While it is known that EREs occur in greater numbers (Martinez and Wahli 1989, Kato et al. 1995, Geserick et al. 2005) and in closer proximity (Carroll et al. 2006, Lin et al. 2007) to estrogen responsive genes, my findings confirmed this pattern without the need for a separate molecular technique such as ChIP-Seq. An interesting aspect of my search for estrogen receptor binding is that I did not find any canonical EREs near the estrogen responsive genes. Instead I found high binding values due to the presence of perfect half-sites, and I found an excess of these half-sites near the estrogen-responsive genes. While I cannot be certain of any *in vivo* binding locations (Wasserman and Sandelin 2004), the pattern suggests that many of my estrogen responsive genes have proximal estrogen-receptor control. It is also worth noting that most sex-differentiated genes do not have a higher than expected amount of estrogen receptor binding, suggesting other sex-mediated differences beyond estrogen levels *per se* contribute to much of the sexual dimorphism observed in *S. scovelli*.

Estrogen plays a clear role in the expression of sexually dimorphic traits in *S. scovelli*, and my study detected a suite of genes affected by estrogen. Primarily I find evidence that structural and organizational regulatory changes likely lead to the observed differences in banding pattern and body depth between males and females. I also show that these differences are caused by up-regulation in response to estrogen, even though some of these structural genes are down-regulated by estrogen in other, non-sex-role-reversed systems. Further, I suggest that my data are consistent with the higher energetic demands on females during mating. Given that some of the genes were involved in muscle development and adipose tissue, female pipefish may have recruited genes involved in muscle recovery and energy storage. Finally, I show that estrogen-biased genes have proximal estrogen receptor binding, setting the stage for future comparative work on the evolution of *cis*-acting regulatory factors during the evolution of female ornaments in syngnathid fishes. Thus, my results will contribute to future reductionist studies focused on the roles of particular genes in the development and maintenance of sexually dimorphic traits and to broader evolutionary studies addressing how changes in hormonal regulation shape patterns of sexual dimorphism.

Literature Cited

Andersson, M. B. 1994. Sexual selection. Princeton University Press.

Andrews, S. 2010. FastQC: a quality control tool for high throughput sequence data.

Balda, M. S., and K. Matter. 2016. Tight junctions as regulators of tissue remodelling. Current Opinion in Cell Biology **42**:94-101.

Bartos, L., G. A. Bubenik, and E. Kuzmova. 2012. Endocrine relationships between rank-related behavior and antler growth in deer. Frontiers in Bioscience **4**:1111-1126.

Beal, A. P., F. D. Martin, and M. C. Hale. 2018. Using RNA-seq to determine patterns of sex-bias in gene expression in the brain of the sex-role reversed Gulf Pipefish (Syngnathus scovelli). Marine Genomics **37**:120-127.

Berglund, A., G. Rosenqvist, and P. Bernet. 1997. Ornamentation predicts reproductive success in female pipefish. Behavioral Ecology and Sociobiology **40**:145-150.

Berglund, A., G. Rosenqvist, and I. Svensson. 1986a. Mate choice, fecundity and sexual dimorphism in two pipefish species (Syngnathidae). Behavioral Ecology and Sociobiology **19**:301-307.

Berglund, A., G. Rosenqvist, and I. Svensson. 1986b. Reversed sex-roles and parental energy investment in zygotes of two pipefish (Syngnathidae) species. Marine Ecology Progress Series **29**:209-215.

Blomquist, A., G. Schworer, H. Schablowski, A. Psoma, M. Lehnen, K. H. Jakobs, and U. Rumenapp. 2000. Identification and characterization of a novel Rho-specific guanine nucleotide exchange factor. Biochemical Journal **352**:319-325.

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

Bonduriansky, R., and S. F. Chenoweth. 2009. Intralocus sexual conflict. Trends in Ecology & Evolution **24**:280-288.

Boyer, M., N. Poujol, E. Margeat, and C. A. Royer. 2000. Quantitative characterization of the interaction between purified human estrogen receptor α and DNA using fluorescence anisotropy. Nucleic acids research **28**:2494-2502.

Bradshaw, A. D., and E. H. Sage. 2001. SPARC, a matricellular protein that functions in cellular differentiation and tissue response to injury. The Journal of clinical investigation **107**:1049-1054.

Brekken, R. A., and E. H. Sage. 2000. SPARC, a matricellular protein: at the crossroads of cell–matrix. Matrix Biology **19**:569-580.

Budi, E. H., L. B. Patterson, and D. M. Parichy. 2011. Post-Embryonic Nerve-Associated Precursors to Adult Pigment Cells: Genetic Requirements and Dynamics of Morphogenesis and Differentiation. PLoS Genetics **7**:e1002044.

Carroll, J. S., C. A. Meyer, J. Song, W. Li, T. R. Geistlinger, J. Eeckhoute, A. S. Brodsky, E. K. Keeton, K. C. Fertuck, and G. F. Hall. 2006. Genome-wide analysis of estrogen receptor binding sites. Nature genetics **38**:1289.

Chapman, T., G. Arnqvist, J. Bangham, and L. Rowe. 2003. Sexual conflict. Trends in Ecology & Evolution **18**:41-47.

Chiang, J. Y. 2017. Recent advances in understanding bile acid homeostasis. F1000Research **6**:2029-2029.

Conesa, A., S. Götz, J. M. García-Gómez, J. Terol, M. Talón, and M. Robles. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics **21**:3674-3676.

Connallon, T., and A. G. Clark. 2014. Evolutionary inevitability of sexual antagonism. Proceedings of the Royal Society of London B: Biological Sciences **281**:20132123.

Connallon, T., and L. L. Knowles. 2005. Intergenomic conflict revealed by patterns of sex-biased gene expression. Trends in Genetics **21**:495-499.

Cox, R., D. Stenquist, and R. Calsbeek. 2009. Testosterone, growth and the evolution of sexual size dimorphism. Journal of Evolutionary Biology **22**:1586-1598.

Cunha, M. A. B., A. Berglund, and N. M. Monteiro. 2017. Female ornaments signal own and offspring quality in a sex-role-reversed fish with extreme male parental care. Marine Ecology **38**:e12461.

Darwin, C. 1871. The descent of man, and selection in relation to sex. London: Murray **415**.

Deegan, B. J., A. M. Bona, V. Bhat, D. C. Mikles, C. B. McDonald, K. L. Seldeen, and A. Farooq. 2011. Structural and thermodynamic consequences of the replacement of zinc with environmental metals on estrogen receptor α–DNA interactions. Journal of Molecular Recognition **24**:1007-1017.

Downs, K. P. 2013. Characterization of UVB Inducible Gene Expression in Xiphophorus Skin.

Eens, M., and R. Pinxten. 2000. Sex-role reversal in vertebrates: behavioural and endocrinological accounts. Behavioural Processes **51**:135-147.

Enns, D. L., and P. M. Tiidus. 2010. The influence of estrogen on skeletal muscle. Sports medicine **40**:41-58.

Eskelinen, E.-L., Y. Tanaka, and P. Saftig. 2003. At the acidic edge: emerging functions for lysosomal membrane proteins. Trends in Cell Biology **13**:137-145.

Ferreira, M., J. Costa, and M. A. Reis-Henriques. 2014. ABC transporters in fish species: a review. Frontiers in Physiology **5**.

Firmin, F. F., F. Oger, C. Gheeraert, J. Dubois-Chevalier, A.-S. Vercoutter-Edouart, F. Alzaid, C. Mazuy, H. Dehondt, J. Alexandre, B. Derudas, Q. Dhalluin, M. Ploton, A. Berthier, E. Woitrain, T. Lefebvre, N. Venteclef, F. Pattou, B. Staels, J. Eeckhoute, and P. Lefebvre. 2017. The RBM14/CoAA-interacting, long intergenic non-coding RNA Paral1 regulates adipogenesis and coactivates the nuclear receptor PPARγ. Scientific reports **7**:14087.

Flanagan, S. P., J. B. Johnson, E. Rose, and A. G. Jones. 2014. Sexual selection on female ornaments in the sex-role-reversed Gulf pipefish (Syngnathus scovelli). J Evol Biol **27**:2457-2467.

Garcia de la serrana, D., A. Estévez, K. Andree, and I. A. Johnston. 2012. Fast skeletal muscle transcriptome of the Gilthead sea bream (Sparus aurata) determined by next generation sequencing. BMC Genomics **13**:181.

Gauvry, L., and B. Fauconneau. 1996. Cloning of a trout fast skeletal myosin heavy chain expressed both in embryo and adult muscles and in myotubes neoformed in vitro. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology **115**:183-190.

Geserick, C., H.-A. Meyer, and B. Haendler. 2005. The role of DNA response elements as allosteric modulators of steroid receptor function. Molecular and cellular endocrinology **236**:1-7.

Ghosal, R., and P. W. Sorensen. 2016. Male-typical courtship, spawning behavior, and olfactory sensitivity are induced to different extents by androgens in the goldfish suggesting they are controlled by different neuroendocrine mechanisms. General and comparative endocrinology **232**:160-173.

Gonçalves, D., S. S. Costa, M. C. Teles, H. Silva, M. Inglês, and R. F. Oliveira. 2014. Oestradiol and prostaglandin F2α regulate sexual displays in females of a sex-role reversed fish. Proceedings of the Royal Society B: Biological Sciences **281**:20133070.

Goymann, W., A. Wittenzellner, I. Schwabl, and M. Makomba. 2008. Progesterone modulates aggression in sex-role reversed female African black coucals. Proceedings of the Royal Society B: Biological Sciences **275**:1053-1060.

Gurgul, A., K. Pawlina-Tyszko, M. Bugno-Poniewierska, T. Szmatoła, I. Jasielczuk, S. Dobosz, and K. Ocalewicz. 2018. Transcriptome analysis of rainbow trout (Oncorhynchus mykiss) eggs subjected to the high hydrostatic pressure treatment. International journal of genomics **2018**.

Hansson, R., and K. Wikvall. 1982. Hydroxylations in Biosynthesis of Bile Acids. European Journal of Biochemistry **125**:423-429.

Hao, R., M. Bondesson, A. V. Singh, A. Riu, C. W. McCollum, T. B. Knudsen, D. A. Gorelick, and J.-Å. Gustafsson. 2013. Identification of Estrogen Target Genes

during Zebrafish Embryonic Development through Transcriptomic Analysis. PLoS One **8**:e79020.

Heeley, D., and C. Hong. 1994. Isolation and characterization of tropomyosin from fish muscle. Comparative Biochemistry and Physiology Part B: Comparative Biochemistry **108**:95-106.

Herder, C., J. M. Swiercz, C. Müller, R. Peravali, R. Quiring, S. Offermanns, J. Wittbrodt, and F. Loosli. 2013. ArhGEF18 regulates RhoA-Rock2 signaling to maintain neuro-epithelial apico-basal polarity and proliferation. Development **140**:2787-2797.

Higdon, C. W., R. D. Mitra, and S. L. Johnson. 2013. Gene Expression Analysis of Zebrafish Melanocytes, Iridophores, and Retinal Pigmented Epithelium Reveals Indicators of Biological Function and Developmental Origin. PLoS One **8**:e67801.

Higgs, H. N. 2005. Formin proteins: a domain-based approach. Trends in Biochemical Sciences **30**:342-353.

Iguchi, Y., M. Yamaguchi, H. Sato, K. Kihira, T. Nishimaki-Mogami, and M. Une. 2010. Bile alcohols function as the ligands of membrane-type bile acid-activated G protein-coupled receptor. Journal of lipid research **51**:1432-1441.

Iimura, K., H. Tohse, K. Ura, and Y. Takagi. 2012. Expression Patterns of runx2, sparc, and bgp During Scale Regeneration in the Goldfish Carassius auratus. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution **318**:190-198.

Imrie, D., and K. C. Sadler. 2010. White adipose tissue development in zebrafish is regulated by both developmental time and fish size. Developmental Dynamics **239**:3013-3023.

Iskratsch, T., S. Lange, J. Dwyer, A. L. Kho, C. d. Remedios, and E. Ehler. 2010. Formin follows function: a muscle-specific isoform of FHOD3 is regulated by CK2 phosphorylation and promotes myofibril maintenance. The Journal of cell biology **191**:1159-1172.

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

Jones, A. G., and J. C. Avise. 2001. Mating Systems and Sexual Selection in Male-Pregnant Pipefishes and Seahorses: Insights from Microsatellite-Based Studies of Maternity. Journal of Heredity **92**:150-158.

Kadi, F., C. Karlsson, B. Larsson, J. Eriksson, M. Larval, H. Billig, and I. H. Jonsdottir. 2002. The effects of physical activity and estrogen treatment on rat fast and slow skeletal muscles following ovariectomy. Journal of Muscle Research & Cell Motility **23**:335.

Karsi, A., D. Cao, P. Li, A. Patterson, A. Kocabas, J. Feng, Z. Ju, K. D. Mickett, and Z. Liu. 2002. Transcriptome analysis of channel catfish (Ictalurus punctatus): initial analysis of gene expression and microsatellite-containing cDNAs in the skin. Gene **285**:157-168.

Kato, S., H. Endoh, Y. Masuhiro, T. Kitamoto, S. Uchiyama, H. Sasaki, S. Masushige, Y. Gotoh, E. Nishida, and H. Kawashima. 1995. Activation of the estrogen receptor through phosphorylation by mitogen-activated protein kinase. Science **270**:1491-1494.

Kelsh, R. N., M. L. Harris, S. Colanesi, and C. A. Erickson. 2009. Stripes and belly-spots—A review of pigment cell morphogenesis in vertebrates. Seminars in Cell & Developmental Biology **20**:90-104.

Khayrullin, A., L. Smith, D. Mistry, A. Dukes, Y. A. Pan, and M. W. Hamrick. 2016. Chronic alcohol exposure induces muscle atrophy (myopathy) in zebrafish and alters the expression of microRNAs targeting the Notch pathway in skeletal muscle. Biochemical and Biophysical Research Communications **479**:590-595.

Kim, D., B. Langmead, and S. L. Salzberg. 2015. HISAT: a fast spliced aligner with low memory requirements. Nature methods **12**:357.

Kimball, S. R., and C. H. Lang. 2018. Mechanisms Underlying Muscle Protein Imbalance Induced by Alcohol. Annual Review of Nutrition **38**:197-217.

Kitajima, Y., and Y. Ono. 2016. Estrogens maintain skeletal muscle and satellite cell functions. The Journal of endocrinology **229**:267-275.

Klein-Hitpass, L., G. U. Ryffel, E. Heitlinger, and A. C. Cato. 1988. A 13 bp palindrome is a functional estrogen responsive element and interacts specifically with estrogen receptor. Nucleic acids research **16**:647-663.

Lande, R. 1980. Sexual dimorphism, sexual selection, and adaptation in polygenic characters. Evolution **34**:292-305.

Lane, T. F., and E. H. Sage. 1994. The biology of SPARC, a protein that modulates cell-matrix interactions. The FASEB Journal **8**:163-173.

Ledda, F., A. I. Bravo, S. Adris, L. Bover, J. Mordoh, and O. L. Podhajcer. 1997. The expression of the secreted protein acidic and rich in cysteine (SPARC) is associated with the neoplastic progression of human melanoma. Journal of Investigative Dermatology **108**:210-214.

Lefebvre, P., B. Cariou, F. Lien, F. Kuipers, and B. Staels. 2009. Role of Bile Acids and Bile Acid Receptors in Metabolic Regulation. Physiological reviews **89**:147-191.

Lehane, D. B., N. McKie, R. G. G. Russell, and I. W. Henderson. 1999. Cloning of a Fragment of the Osteonectin Gene from Goldfish,Carassius auratus:Its Expression and Potential Regulation by Estrogen. General and comparative endocrinology **114**:80-87.

Ligon, R. A., and K. L. McCartney. 2016. Biochemical regulation of pigment motility in vertebrate chromatophores: a review of physiological color change mechanisms. Current Zoology **62**:237-252.

Lin, C.-Y., V. B. Vega, J. S. Thomsen, T. Zhang, S. L. Kong, M. Xie, K. P. Chiu, L. Lipovich, D. H. Barnett, and F. Stossi. 2007. Whole-genome cartography of estrogen receptor α binding sites. PLoS Genetics **3**:e87.

Lin, Q., S. Fan, Y. Zhang, M. Xu, H. Zhang, Y. Yang, A. P. Lee, J. M. Woltering, V. Ravi, and H. M. Gunter. 2016. The seahorse genome and the evolution of its specialized morphology. Nature **540**:395.

Lindsay, W. R., D. G. Barron, M. S. Webster, and H. Schwabl. 2016. Testosterone activates sexual dimorphism including male-typical carotenoid but not melanin plumage pigmentation in a female bird. The Journal of Experimental Biology **219**:3091-3099.

Lock, E.-J., R. Waagbø, S. Wendelaar Bonga, and G. Flik. 2009. The significance of vitamin D for fish: A review.

Lončar, J., M. Popović, R. Zaja, and T. Smital. 2010. Gene expression analysis of the ABC efflux transporters in rainbow trout (Oncorhynchus mykiss). Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology **151**:209-215.

Lythgoe, J. N., J. Shand, and R. G. Foster. 1984. Visual pigment in fish iridocytes. Nature **308**:83-84.

Mank, J. E. 2017. The transcriptional architecture of phenotypic dimorphism. Nature Ecology & Evolution **1**:0006.

Martinez, E., and W. Wahli. 1989. Cooperative binding of estrogen receptor to imperfect estrogen-responsive DNA elements correlates with their synergistic hormone-dependent enhancer activity. The EMBO journal **8**:3781-3791.

Mitsche, M. A., J. G. McDonald, H. H. Hobbs, and J. C. Cohen. 2015. Flux analysis of cholesterol biosynthesis in vivo reveals multiple tissue and cell-type specific pathways. Elife **4**:e07999.

Mobley, K. B., J. R. Morrongiello, M. Warr, D. J. Bray, and B. B. M. Wong. 2018. Female ornamentation and the fecundity trade-off in a sex-role reversed pipefish. Ecol Evol **8**:9516-9525.

Mommsen, T. P. 2001. Paradigms of growth in fish. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology **129**:207-219.

Muck, C., and W. Goymann. 2011. Throat patch size and darkness covaries with testosterone in females of a sex-role reversed species. Behavioral Ecology **22**:1312-1319.

Navarro, R. E., J. L. Ramos-Balderas, I. Guerrero, V. Pelcastre, and E. Maldonado. 2008. Pigment Dilution Mutants from Fish Models with Connection to Lysosome-Related Organelles and Vesicular Traffic Genes. Zebrafish **5**:309-318.

Nüsslein-Volhard, C., and A. P. Singh. 2017. How fish color their skin: A paradigm for development and evolution of adult patterns. BioEssays **39**:1600231.

Olin, T., and A. von der Decken. 1987. Estrogen treatment and its implication on vitellogenin and myosin synthesis in salmon (Salmo salar). Physiological zoology **60**:346-351.

Oring, L. W., A. J. Fivizzani, M. A. Colwell, and M. E. El Halawani. 1988. Hormonal changes associated with natural and manipulated incubation in the sex-role reversed Wilson's phalarope. General and comparative endocrinology **72**:247-256.

Oring, L. W., A. J. Fivizzani, M. E. El Halawani, and A. Goldsmith. 1986. Seasonal changes in prolactin and luteinizing hormone in the polyandrous spotted sandpiper, Actitis macularia. General and comparative endocrinology **62**:394-403.

Palstra, A. P., S. Beltran, E. Burgerhout, S. A. Brittijn, L. J. Magnoni, C. V. Henkel, H. J. Jansen, G. E. E. J. M. van den Thillart, H. P. Spaink, and J. V. Planas. 2013. Deep RNA Sequencing of the Skeletal Muscle Transcriptome in Swimming Fish. PLoS One **8**:e53171.

Palstra, A. P., M. Rovira, D. Rizo-Roca, J. R. Torrella, H. P. Spaink, and J. V. Planas. 2014. Swimming-induced exercise promotes hypertrophy and vascularization of fast skeletal muscle fibres and activation of myogenic and angiogenic transcriptional programs in adult zebrafish. BMC Genomics **15**:1136.

Parker, G. 1979. Sexual selection and sexual conflict. Sexual selection and reproductive competition in insects **123**:166.

Partridge, C., A. Boettcher, and A. G. Jones. 2010. Short-term exposure to a synthetic estrogen disrupts mating dynamics in a pipefish. Horm Behav **58**:800-807.

Partridge, C., A. Boettcher, and A. G. Jones. 2013. The Role of Courtship Behavior and Size in Mate Preference in the Sex-Role-Reversed Gulf Pipefish, Syngnathus scovelli. Ethology **119**:692-701.

Patterson, L. B., and D. M. Parichy. 2013. Interactions with Iridophores and the Tissue Environment Required for Patterning Melanophores and Xanthophores during Zebrafish Adult Pigment Stripe Formation. PLoS Genetics **9**:e1003561.

Petratou, K., T. Subkhankulova, J. A. Lister, A. Rocco, H. Schwetlick, and R. N. Kelsh. 2018. A systems biology approach uncovers the core gene regulatory network governing iridophore fate choice from the neural crest. PLoS Genetics **14**:e1007402.

Pierens, S. L., and D. R. Fraser. 2015. The origin and metabolism of vitamin D in rainbow trout. The Journal of Steroid Biochemistry and Molecular Biology **145**:58-64.

Pikkarainen, P., and N. Räihä. 1967. Development of alcohol dehydrogenase activity in the human liver. Pediatric Research **1**:165.

Prabhu, A. V., W. Luu, D. Li, L. J. Sharpe, and A. J. Brown. 2016. DHCR7: A vital enzyme switch between cholesterol and vitamin D production. Progress in Lipid Research **64**:138-151.

Prabhu, A. V., W. Luu, L. J. Sharpe, and A. J. Brown. 2017. Phosphorylation regulates activity of 7-dehydrocholesterol reductase (DHCR7), a terminal enzyme of cholesterol synthesis. The Journal of Steroid Biochemistry and Molecular Biology **165**:363-368.

Qiao, Q., S. Le Manach, B. Sotton, H. Huet, E. Duvernois-Berthet, A. Paris, C. Duval, L. Ponger, A. Marie, and A. Blond. 2016. Deep sexual dimorphism in adult medaka fish liver highlighted by multi-omic approach. Scientific reports **6**:32459.

Qulali, M., R. A. Ross, and D. W. Crabb. 1991. Estradiol induces class I alcohol dehydrogenase activity and mRNA in kidney of female rats. Archives of Biochemistry and Biophysics **288**:406-413.

Renn, J., M. Schaedel, J.-N. Volff, R. Goerlich, M. Schartl, and C. Winkler. 2006. Dynamic expression of sparc precedes formation of skeletal elements in the Medaka (Oryzias latipes). Gene **372**:208-218.

Rice, W. R. 1984. SEX CHROMOSOMES AND THE EVOLUTION OF SEXUAL DIMORPHISM. Evolution **38**:735-742.

Robert, G., C. Gaggioli, O. Bailet, C. Chavey, P. Abbe, E. Aberdam, E. Sabatié, A. Cano, A. G. de Herreros, and R. Ballotti. 2006. SPARC represses E-cadherin and induces mesenchymal transition during melanoma development. Cancer research **66**:7516-7523.

Robinson, M. D., D. J. McCarthy, and G. K. Smyth. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics **26**:139-140.

Rose, E., S. P. Flanagan, and A. G. Jones. 2015. The Effects of Synthetic Estrogen Exposure on the Sexually Dimorphic Liver Transcriptome of the Sex-Role-Reversed Gulf Pipefish. PLoS One **10**:e0139401.

Rose, E., K. A. Paczolt, and A. G. Jones. 2013a. The contributions of premating and postmating selection episodes to total selection in sex-role-reversed Gulf pipefish. Am Nat **182**:410-420.

Rose, E., K. A. Paczolt, and A. G. Jones. 2013b. The effects of synthetic estrogen exposure on premating and postmating episodes of selection in sex-role-reversed Gulf pipefish. Evol Appl **6**:1160-1170.

Salis, P., T. Lorin, V. Lewis, C. Rey, A. Marcionetti, M.-L. Escande, N. Roux, L. Besseau, N. Salamin, M. Sémon, D. Parichy, J.-N. Volff, and V. Laudet. 2019. Developmental and comparative transcriptomic identification of iridophore

contribution to white barring in clownfish. Pigment Cell & Melanoma Research **32**:391-402.

Sárria, M. P., M. M. Santos, L. F. C. Castro, N. M. Vieira, and N. M. Monteiro. 2013. Estrogenic chemical effects are independent from the degree of sex role reversal in pipefish. Journal of Hazardous Materials **263**:746-753.

Shoubridge, E., and P. Hochachka. 1980. Ethanol: novel end product of vertebrate anaerobic metabolism. Science **209**:308-309.

Singh, Ajeet P., and C. Nüsslein-Volhard. 2015. Zebrafish Stripes as a Model for Vertebrate Colour Pattern Formation. Current Biology **25**:R81-R92.

Sköld, H. N., T. Amundsen, P. A. Svensson, I. Mayer, J. Bjelvenmark, and E. Forsgren. 2008. Hormonal regulation of female nuptial coloration in a fish. Hormones and Behavior **54**:549-556.

Sköld, H. N., S. Aspengren, K. L. Cheney, and M. Wallin. 2016. Fish chromatophores—from molecular motors to animal behavior. Pages 171-219 International Review of Cell and Molecular Biology. Elsevier.

Small, C., S. Bassham, J. Catchen, A. Amores, A. Fuiten, R. Brown, A. Jones, and W. Cresko. 2016. The genome of the Gulf pipefish enables understanding of evolutionary innovations. Genome biology **17**:258.

Small, C. M., A. D. Harlin-Cognato, and A. G. Jones. 2013. Functional similarity and molecular divergence of a novel reproductive transcriptome in two male-pregnant Syngnathus pipefish species. Ecol Evol **3**:4092-4108.

Smyth, G. K., W. Shi, and Y. Liao. 2013. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics **30**:923-930.

Strömberg, S., M. G. Björklund, A. Asplund, R. Rimini, J. Lundeberg, P. Nilsson, F. Pontén, and M. J. Olsson. 2008. Transcriptional profiling of melanocytes from patients with vitiligo vulgaris. Pigment Cell & Melanoma Research **21**:162-171.

Taniguchi, K., R. Takeya, S. Suetsugu, M. Kan-o, M. Narusawa, A. Shiose, R. Tominaga, and H. Sumimoto. 2009. Mammalian formin fhod3 regulates actin assembly and sarcomere organization in striated muscles. Journal of Biological Chemistry **284**:29873-29881.

Terry, S. J., C. Zihni, A. Elbediwy, E. Vitiello, I. V. Leefa Chong San, M. S. Balda, and K. Matter. 2011. Spatially restricted activation of RhoA signalling at epithelial junctions by p114RhoGEF drives junction formation and morphogenesis. Nature Cell Biology **13**:159.

Torres, J. J., M. D. Grigsby, and M. E. Clarke. 2012. Aerobic and anaerobic metabolism in oxygen minimum layer fishes: the role of alcohol dehydrogenase. The Journal of Experimental Biology **215**:1905-1914.

Trapnell, C., A. Roberts, L. Goff, G. Pertea, D. Kim, D. R. Kelley, H. Pimentel, S. L. Salzberg, J. L. Rinn, and L. Pachter. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature protocols **7**:562.

Tremble, P. M., T. F. Lane, E. H. Sage, and Z. Werb. 1993. SPARC, a secreted protein associated with morphogenesis and tissue remodeling, induces expression of metalloproteinases in fibroblasts through a novel extracellular matrix-dependent pathway. The Journal of cell biology **121**:1433-1444.

Trivers, R. 1972. Parental investment and sexual selection. Biological Laboratories, Harvard University Cambridge, MA.

Tyulmenkov, V. V., and C. M. Klinge. 2001. A mathematical approach to predict the affinity of estrogen receptors α and β binding to DNA. Molecular and cellular endocrinology **182**:109-119.

Velazquez-Villegas, L. A., A. Perino, V. Lemos, M. Zietak, M. Nomura, T. W. H. Pols, and K. Schoonjans. 2018. TGR5 signalling promotes mitochondrial fission and beige remodelling of white adipose tissue. Nature communications **9**:245.

Vincent, A., I. Ahnesjö, A. Berglund, and G. Rosenqvist. 1992. Pipefishes and seahorses: Are they all sex role reversed? Trends in Ecology & Evolution **7**:237-241.

Voigt, C., and W. Goymann. 2007. Sex-role reversal is reflected in the brain of African black coucals (Centropus grillii). Developmental Neurobiology **67**:1560-1573.

Wang, H., J. T. T. Tan, A. Emelyanov, V. Korzh, and Z. Gong. 2005. Hepatic and extrahepatic expression of vitellogenin genes in the zebrafish, Danio rerio. Gene **356**:91-100.

Wasserman, W. W., and A. Sandelin. 2004. Applied bioinformatics for the identification of regulatory elements. Nature Reviews Genetics **5**:276.

Watanabe, M., S. M. Houten, C. Mataki, M. A. Christoffolete, B. W. Kim, H. Sato, N. Messaddeq, J. W. Harney, O. Ezaki, T. Kodama, K. Schoonjans, A. C. Bianco, and J. Auwerx. 2006. Bile acids induce energy expenditure by promoting intracellular thyroid hormone activation. Nature **439**:484-489.

Weigele, J., T. A. Franz-Odendaal, and R. Hilbig. 2015. Expression of SPARC and the osteopontin-like protein during skeletal development in the cichlid fish Oreochromis mossambicus. Developmental Dynamics **244**:955-972.

Weil, C., F. Lefèvre, and J. Bugeon. 2013. Characteristics and metabolism of different adipose tissues in fish. Reviews in Fish Biology and Fisheries **23**:157-173.

Westendorf, J. J., R. Mernaugh, and S. W. Hiebert. 1999. Identification and characterization of a protein containing formin homology (FH1/FH2) domains. Gene **232**:173-182.

Williams, C., K. Edvardsson, S. A. Lewandowski, A. Ström, and J. Å. Gustafsson. 2007. A genome-wide study of the repressive effects of estrogen receptor beta on estrogen receptor alpha signaling in breast cancer cells. Oncogene **27**:1019.

Williams, G. C. 1975. Sex and evolution. Princeton University Press.

Witkiewicz, A. K., B. Freydin, I. Chervoneva, M. Potoczek, W. Rizzo, H. Rui, J. R. Brody, G. F. Schwartz, and M. P. Lisanti. 2010. Stromal CD10 and SPARC expression in ductal carcinoma in situ (DCIS) patients predicts disease recurrence. Cancer Biology & Therapy **10**:391-396.

Wright, A. E., M. Fumagalli, C. R. Cooney, N. I. Bloch, F. G. Vieira, S. D. Buechel, N. Kolm, and J. E. Mank. 2018. Male-biased gene expression resolves sexual

conflict through the evolution of sex-specific genetic architecture. Evolution Letters **2**:52-61.

Zeigler, M., V. Meiner, J. P. Newman, B. Steiner-Birmanns, R. Bargal, V. Sury, G. Mengistu, O. Kakhlon, I. Leykin, Z. Argov, O. Abramsky, and A. Lossos. 2014. A novel SCARB2 mutation in progressive myoclonus epilepsy indicated by reduced β-glucocerebrosidase activity. Journal of the Neurological Sciences **339**:210-213.

Zheng, W., H. Xu, S. H. Lam, H. Luo, R. K. M. Karuturi, and Z. Gong. 2013. Transcriptomic Analyses of Sexual Dimorphism of the Zebrafish Liver and the Effect of Sex Hormones. PLoS One **8**:e53562.

Zhong, L., L. Yuan, Y. Rao, Z. Li, X. Zhang, T. Liao, Y. Xu, and H. Dai. 2014. Distribution of vitellogenin in zebrafish (Danio rerio) tissues for biomarker analysis. Aquatic Toxicology **149**:1-7.

Zigmond, S. H. 2004. Formin-induced nucleation of actin filaments. Current Opinion in Cell Biology **16**:99-105.

CHAPTER V

CONCLUSIONS

In the course of this thesis, I was able to provide evidence for lack of age influence on sexual selection in *S. scovelli*, demonstrate the use of an algorithm I developed to locate genomic regions of high estrogen receptor binding, and find genes putatively involved in secondary sex traits and an abundance of EREs near these genes. The non-existent effect of age on banding patterns and male mate choice allows me to consider only the function of genes involved in the maintenance and development of the these traits without having to consider the age of individual in those expression patterns. The efficacy of my algorithm provides more certainty that it can correctly identify putative estrogen-binding regions thereby enabling me to quickly scan the *S. scovelli* genome to locate EREs. Finally, the transcriptomic work illuminated genes that were both estrogen sensitive and sex-biased in the dimorphic skin and muscle tissue. Taking a closer look at these genes, I was able to demonstrate an abundance of estrogen response regions thereby meeting my expectation that genes under sexual selection in sex-role reversed species would have an excess of EREs.

I conducted a male mate-choice experiment using *Syngnathus scovelli*, a sex-role reversed pipefish with male brood care and female ornaments. There was no effect of female age on male mate choice detected, thereby assuaging concerns over previously described sexual selection in the natural population. I further confirmed that males prefer females with greater courtship effort and body depth and both of these traits correlated with size of bands. Bands may therefore be an honest indicator of female fecundity and

quality that males use to assess females. Lastly, I find that ornaments may be potentially low-cost and that fast-maturing females have similar size bands and can successfully compete with their slower-maturing competitors. With evidence for both body depth and banding as sexually selected secondary sex traits, a study on the transcriptome of the dimorphic tissues would reveal putative genes as targets of sexual selection.

Before moving into the transcriptome of the secondary sex traits, I needed to successfully identify EREs in the *S. scovelli* genome. In developing EREFinder, I was able to demonstrate several aspects of its utility, particularly 1) how altering the window size can affect the interpretation of the output, 2) that results from EREFinder are consistent with expectations from empirical studies of ER binding *in vivo*, and 3) that EREFinder is probably most useful for comparative investigations of ERE gains and losses. While window-size and concordance with ChIP studies is essential for validation of the algorithm, the comparative studies are excellent demonstrations of the utility of EREFinder. Identifying high estrogen receptor binding regions proximal to genes is necessary to confirm ERE's role in sexual selection.

Female bands in *S. scovelli* are believed to be estrogen mediated as males exposed to estradiol show banding patterns. Using RNA-seq, I show most genes differentially expressed between males and females are not estrogen-mediated, but when they are it is because 1) they are involved in the extracellular matrix of skin for arrangement of cells, 2) they are involved muscular growth and structural arrangement as well as potentially allowing for longer muscle use, or 3) they are part of the extrahepatic system involved in fat storage. I suggest the banding pattern is generated by

arranging pigment cells in a certain way as supported by iridophore production genes that are expressed in higher amounts in females. Further, females may have more energetic demands for courtship since they usually initiate courtship and are polyandrous.

With confirmation of bands and body depth as targets of sexual selection, a set of genes involved in the production of those traits, and an algorithm for detecting EREs I was able to test the roles EREs play in sexual selection in *S. scovelli*. Because EE2 induces female traits in males, the genes up-regulated by estrogen are more likely to be intertwined in banding and body depth than purely sex-biased genes. Taking the sets of trait-related genes and general sex-difference genes I was able to compare the relative amount of estrogen receptor binding in each respective set. I demonstrate that the set putatively involved in the production of the secondary sex traits had an excess of estrogen binding compared to the set comprising just sex-biased expression. This evidence of high estrogen binding on genes putatively under sexual selection supports the hypothesis that EREs, and by extension sex-biased HREs, play in the evolution of sexually selected traits.