Morphological and Gene Expression Plasticity in Neotropical Cichlid Fishes

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Sharon Fern Clemmensen
December 2017

DEDICATION

For Santiago

ACKNOWLEDGEMENTS

I would like to thank my family for providing support throughout my time at UT, especially my husband Santiago. Without their encouragement I would never have completed my degree. I would also like to thank my Master's advisor from the University of Florida, Dan Hahn, who gave me emotional support during a particularly difficult time in the doctoral program.

I thank my committee: Ben Fitzpatrick, Brian O'Meara, and Elena Shpak who were very patient with me. I thank my advisor Jim Fordyce, who took me on late in my graduate career and helped me work through tricky analyses without complaint. Thank you to Meg Staton and Miriam Payà Milans, whose bioinformatics expertise was invaluable.

Thank you to the current and former members of the fish lab who assisted me in long rearing experiments: Ben Keck, Max Rupp, Phillip Hollingsworth, CJ Grabenstein, & Jasmine Vazin. Special thanks to Jennifer Joice for always helping out no matter the task. Very special thanks to Sam Borstein for not only helping me with the lab work but also helping me sound out ideas and work through analyses, and letting me use his awesome cichlid phylogeny.

Finally, thank you to the members of the EEB department, who have been friends and comrades throughout my time at UT. Thank you to the members of my cohort, especially Zach Marion, Christine Dumoulin, and Lacy Chick: without each other we all surely would have gone insane.

ABSTRACT

Trophic divergence in cichlid fish is linked to morphological shifts in the pharyngeal jaw apparatus. For instance, in the Heroine cichlids of Central America, the ability to crush hard-shelled mollusks is a convergent phenotype with multiple evolutionary origins. These durophagous species often have very similar pharyngeal jaw morphologies associated with the pharyngeal jaw apparatus and some of these similarities could be due to phenotypically plastic responses to mechanical stress. I examined the durophagous cichlid *Vieja maculicauda* for differences in pharyngeal osteology, dentition, and soft tissues when exposed to different diet regimes. Here I discuss the effect on the morphology and gene expression of the pharyngeal jaw apparatus of varying mechanical stress without varying nutrient content, place this in a comparative framework, and discuss the effect of plasticity on morphological diversity.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The overarching question of this dissertation is to ask to what extent is individual morphology shaped by phenotypic plasticity. Specifically, how components of a complex phenotype can be modified by plastic effects on multiple levels of biological organization – individual morphology, individual gene regulation, and gene regulation across species. I address this question using the pharyngeal jaw apparatus of cichlid fishes, a musculoskeletal complex that performs prey processing in this clade and is distinct from the oral jaws which perform prey capture. In the Central American Heroine cichlid clade there are multiple species that have independently evolved the ability to crush hard-shelled prey, also known as durophagy. These durophagous species have distinct modifications to their pharyngeal jaw apparatus that are known to be phenotypically plastic in other cichlid clades, making this system suitable to address the magnitude of effect diet may have on this complex phenotype.

In chapter 1, I evaluate morphological phenotypic plasticity in the Heroine cichlid *Vieja maculicauda* to determine if we can effectively explore the morphospace of the cichlid pharyngeal jaw apparatus. I evaluate the phenotypic plasticity of the pharyngeal jaw apparatus in *V. maculicauda* and compared the growth of replacement teeth between experimentally generated jaws to patterns seen in wild-caught *Herichthys minckleyi*, a polymorphic Heroine species that in the wild has both a durophagous and herbivorous morph. I evaluate if *V. maculicauda* can recapitualate the morphological variation observed in wild *H. minckleyi.*

In chapter 2, I determine if genes with plastic expression in a continuous trait experience greater adaptive sequence divergence, by raising *Vieja maculicauda* individuals on soft or hard experimental diets. Using RNAseq, I compare the rates of protein sequence evolution and differential expression in 4,751 genes transcribed in the muscular sling of the pharyngeal jaw apparatus. After examining expression levels, I subsequently calculate relative sequence

divergence for each gene between the cichlid species *V. maculicauda* and *Oreochromis niloticus* in order to determine if prey-induced gene expression evolves at a different rate than genes with constrained expression patterns.

In chapter 3, I investigate if there are constitutive differences in gene expression between durophagous and non-durophagous species when reared on a common diet. I address this question by rearing durophagous and non-durophagous Central American Heroine cichlids in a common garden experiment, and obtain transcriptomes for the pharyngeal jaw muscular sling. This experiment considers if gene expression has constitutive differences between durophagous and non-durophagous species, or if gene expression is invariant in the musclular sling the absence of prey-induced plasticity.

CHAPTER I
PHENOTYPIC PLASTICITY OF THE PHARYNGEAL JAW APPARATUS IN RESPONSE TO MECHANICAL STRAIN IN A NEOTROPICAL CICHLID FISH

A version of this chapter is being prepared for submission to *Animal Biology.* S.F. Clemmensen designed experiment and performed all analyses. Co-researchers C.D. Hulsey and J.T. Streelman provided resources and conceptual advice. Co-researcher C.D. Hulsey also collaborated on experimental design.

Abstract

Trophic divergence in cichlid fish is linked to shifts in morphology of the pharyngeal jaw apparatus. For instance, in the Heroine cichlids of Central America, the ability to crush hard-shelled mollusks is a convergent phenotype with multiple evolutionary origins. These durophagous species often have very similar pharyngeal jaw morphologies associated with the pharyngeal jaw apparatus and some of these similarities could be due to phenotypically plastic responses to mechanical stress. We examined the durophagous cichlid *Vieja maculicauda* for differences in pharyngeal osteology, dentition, and soft tissues when exposed to different diet regimes. Here we discuss the effect on the pharyngeal jaw of varying mechanical stress without varying nutrient content. We also examine plasticity in tooth growth in the lower pharyngeal jaw by comparing replacement and erupted teeth between *V. maculicauda* individuals on experimental diets and wild-caught polymorphic *Herichthys minckleyi.*

Introduction

Cranial morphology is directly linked to prey capture and processing in animals, and are especially important in species that lack robust limbs such as fishes. Changes to cranial morphology are therefore critical in the evolution of diet specializations. For example, cichlid species cover an enormous breadth of diet types, including piscivory hard-shelled prey, detritus, zooplankton, and plants. As a consequence, craniofacial morphology across cichlid lineages is extremely diverse (Hulsey & García de León 2005, Clabaut et al. 2007, López-Fernández et al. 2013). Modification of the feeding apparatus, specifically the fusion of the

lower pharyngeal jaw, has been implicated in the rapid diversification of cichlid fishes (Liem 1973, Hulsey et al 2006, Wainwright et al. 2012). Shelled organisms present particular challenges to vertebrate jaws, and durophagy requires a jaw apparatus that can withstand substantial mechanical strain in order to successfully ingest a prey item. Durophagy has evolved multiple times in New World cichlids, and in these lineages we can observe similar changes in morphological characters associated with the pharyngeal jaw apparatus relative to non-durophagous cichlid clades (Hulsey et al 2008). Durophagous cichlids have larger and more robust pharyngeal jaws, and associated muscles that are more massive in comparison to non-durophagous lineages (Hulsey 2006, Wainwright et al. 2012, Burress 2016).

Phenotypic plasticity has likely played a large role in cichlid jaw morphology. Behavioral and morphological plasticity may allow populations to take advantage of fluctuating or novel environments (Pfennig et al. 2010, Wund 2012). More specifically, phenotypic plasticity within a species allows individuals to explore morphological space in response to food, permitting specialization on disparate prey types depending on prey availability (Wimberger 1991, 1994). Plasticity in response to diet has been demonstrated in craniofacial traits in both Old and New World cichlids. These studies have examined oral jaws and head shape traits, and studies that have measured the plastic response in the pharyngeal jaw apparatus have done so primarily in Old World species (Meyer 1987, Huysseune 1995, Smits et al 1996, Gunter et al 2013). Muschick et al (2011) analyzed plasticity of the pharyngeal jaw in a New World cichlid using shape and overall size metrics. To our knowledge no detailed examination of the plasticity of pharyngeal jaw traits has been performed on New World cichlids (for example, midline suturing), which are separated from Old World lineages by roughly 60 million years (Friedman et al 2013).

The primary site of prey processing in cichlids is the pharyngeal jaw apparatus. An integral part of the pharyngeal jaw apparatus are the primary processing units, the pharyngeal teeth. These teeth vary in both size and shape

and these variants can be associated with different diet types. Cichlid pharyngeal teeth are polyphyodont, replaced continuously throughout their lives (Huysseune 1995). The ability of successive pharyngeal teeth to respond to mechanical stress has not been characterized under experimental conditions that control for diet quality, and Wimberger (1993) and Muschick et al. (2011) have demonstrated that nutrient deficiency can cause plastic effects on morphology independent of mechanical stress effects. If tooth development is responsive to mechanical strain, we can expect that cichlids exposed to a hard diet will develop larger teeth relative to individuals of the same species that consume softer prey items.

In this paper we evaluate phenotypic plasticity in the Heroine cichlid *Vieja maculicauda* to determine if we can effectively explore the morphospace of the cichlid pharyngeal jaw apparatus. *V. maculicauda* is a Neotropical molluskivorous cichlid species, with an estimated 20% of its diet consisting of hard-shelled prey (Winemiller et al 1995, Cochran-Biederman & Winemiller 2010). Pharyngeal jaws in this species are robust, with large molariform pharyngeal teeth (Hulsey et al 2008). We evaluated the phenotypic plasticity of the pharyngeal jaw apparatus in *V. maculicauda* and compared the growth of replacement teeth between experimentally generated jaws to patterns seen in wild-caught *Herichthys minckleyi*, a polymorphic Heroine species that in the wild has both a molluskivorous “molariform” morph and a herbivorous “papilliform” morph that have 28% and 0.5% mollusks in their diets, respectively (Hulsey et al 2005). *V. maculicauda* and *H. minckleyi* share a common ancestor 7mya (Hulsey et al 2010). Here we evaluated if *V. maculicauda* can recapitualate the morphological variation observed in wild *H. minckleyi*.

Materials and Methods

Vieja maculicauda *rearing and measurements*

This study was carried out in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The protocol was approved by the Institutional Animal Care and Use Committee of the University of Tennessee, Knoxville (Permit Number 1833-0412). *Vieja maculicauda* individuals were purchased commercially and fed flake food until they reached a minimum standard length of 49mm standard length (SL, measured from the tip of the maxilla to the base of the caudal fin). These fish were then individually housed for the duration of the feeding experiment. Individuals were randomly assigned to a "crushing" (N = 10) treatment or "non-crushing" treatment (N = 16) to induce differences in jaw morphology.
Both treatment diets contained the durable freshwater snail, *Melanoides tuberculata.* This invasive snail is found in high densities in several streams near Knoxville, TN. Snails were collected in large numbers, euthanized via freezing, and stored at -5^{o}C until needed for feeding experiments. These snails were generally larger than the gape size of the experimental cichlids, so we processed the snails in a blender to generate the feeding treatments. In the crushing treatment, snails were chopped to create pieces that could be taken into the buccal cavity, but were still sufficiently intact to require the fish to crush the snails with their pharyngeal jaws prior to ingestion. Snails in the non-crushing treatment were pureed until the snails were fragmented to the point that there was no need for crushing before ingestion. For both treatments, 50g of whole snails were processed as above and added to 14.4g of gelatin, 1.7g of pellet food, and 710mL of water. Each preparation of this mixture produced 60 samples of the 25mL experimental food items. The two experimental diet preparations ensured that no snail pieces were lost and the nutritional content did not differ between treatments. The treatments also created identical food presentation. Individuals were fed the experimental diet three times a week for six months to allow

sufficient time for divergence in the structure of the pharyngeal muscles (Huysseune et al 1989).

At the end of the six-month experimental period, we euthanized individuals with tricaine methanesulfonate (MS-222). Standard length (SL), the length from the tip of the premaxilla to the base of the caudal fin, was measured to the nearest 0.1mm for individuals both at the beginning of the feeding experiment and immediately after euthanization. Growth was measured as SLfinal – SLinitial. Individuals were preserved in 10% formalin, and transferred to 70% ethanol for long-term storage in the David A. Etnier Ichthyological Collection at the University of Tennessee. Following fixation, the *Adductor Mandibulae I-III* (AMI, AMII, AMIII), and the *Levator Posterior/Levator Externus IV* pharyngeal muscle complex (LP/LEIV) muscles on the right sides of the fish were dissected from the pharyngeal apparatus and weighed to the nearest 0.1mg with an electronic balance. The gut was dissected out and its length measured with calipers to the nearest 0.1mm. The lower pharyngeal jaw (LPJ) was dissected out, cleaned of soft tissue, and dried. Ten LPJs from each treatment (total 20) were imaged with a Scanco uCT50 x-ray microtomography (μCT) machine at the Georgia Institute of Technology. μCT images were analyzed using Osirix (Rosset et al. 2004).

We measured μCT images for length of the total LPJ and tooth plate (Figure 1a,c), and width of the insertion of the LP/LEIV ("horns") and anterior attachement point ("keel"). Depth of the jaw was measured at the midline for the first three tooth positions (Figure 1d,f), as well as the depth of the jaw suture at those positions. The perimeter of contact of the suture was calculated according to Hulsey 2006 (Figure 1a,b). We also counted the total number of erupted teeth. Cross-sectional area of erupted teeth was recorded for the first three posterior teeth along the midline of the LPJ on the left and right side, as well as for the most posterior tooth in the adjacent row (Figure 1d). These values were averaged across the left and right side. When present, cross-sectional area of the replacement tooth was recorded for those same positions (Figure 1e). Individuals generally had 2-3 replacement teeth present for these 8 tooth

positions, so proportional tooth growth for an individual was measured as the average of the ratio of replacement to erupted tooth size.

Redundancy analysis of plasticity in V. maculicauda

All statistical analyses were performed with Rv3.3.2 (R core team 2016). Traits measurements were scaled and centered. Some traits were highly correlated (rho >0.9): average horn width – tooth plate length, keel width – total LPJ length, and total number of teeth – jaw depth. We excluded the average horn width, the keel width, and the total number of teeth from further analyses (individual traits plotted in Figures 2-6). Plasticity effects were quantified using partial redundancy analysis using initial standard length as a conditioned variable and diet type as the constrained variable using the rda function in the vegan package (Oksanen et al. 2017). Analysis of variance was performed on the reduced model marginal means using anova.cca with 10,000 permutations.

Comparison of V. maculicauda *and* H. minckleyi *replacement teeth*

Herichtheys minckleyi were wild-caught, and LPJs were extracted, imaged, and measured using the same methods as for *V. maculicauda* (N = 25). For each individual, positions where the replacement tooth was absent were excluded. Z-transformed replacement tooth size of the different jaw types were compared using the lmer function of the lmerTest package, a linear mixed-effects model (Kuznetsova et al. 2016). The initial random intercept model used erupted tooth size, jaw type, species, and tooth position as fixed effects and individual as a random effect. The best-fit model, using AIC criterion, included only erupted tooth size and jaw type as fixed effects and individual as a random effect, where possible jaw types were *V. maculicauda* crushing treatment, *V. maculicauda* non-crushing treatment, *H. minckleyi* molariform type, and *H. minckleyi* papilliform type. Analysis of variance was performed on the best-fit model, and post-hoc Tukey comparison of the four groups was performed with the glht function in the multcomp package (Hothorn et al. 2008).

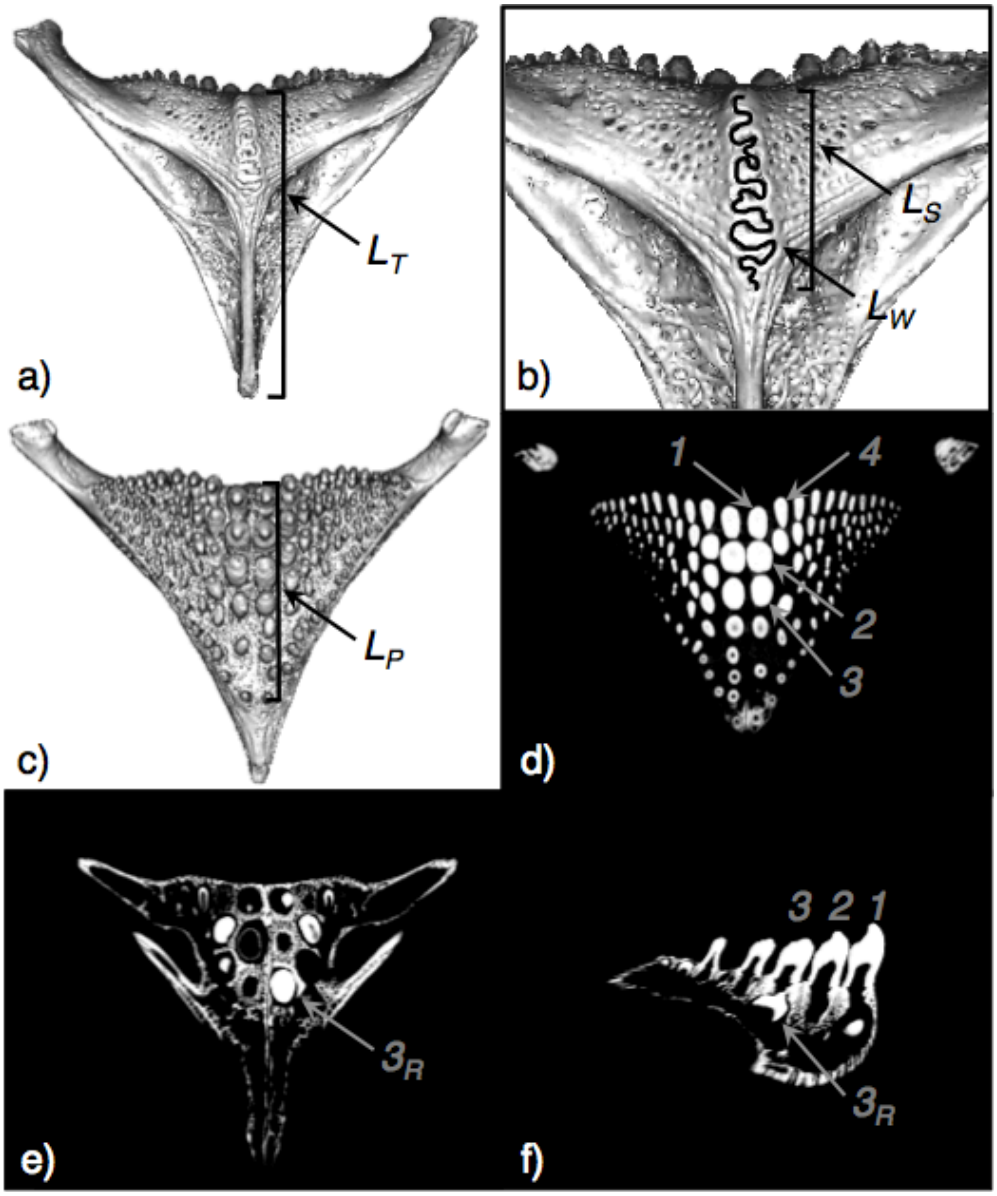

Figure 1. CT images of the cichlid lower pharyngeal jaw.

3D surface renderings (**a** – **c**) and µCT images (**d** – **f**) of a lower pharyngeal jaw (LPJ) from a Vieja maculicada individual. **a** and **b** show an inferior view. LT = total LPJ length, LS = length of LPJ occupied by the suture, and LW = winding length of the suture. The suture measurement, or perimeter of contact is PL = (LS/LT)(LW/LS) (Hulsey 2006). **c** – **e** show a superior view of the LPJ and **f** shows a parasagittal view, with LP = length of the tooth plate. 1 – 4 indicate the four positions of teeth measured for cross-sectional area on each side of the midline of the LPJ, and 3R indicates a replacement tooth growing in at the 3 position. Teeth 1 – 3 are numbered starting from the most posterior position.

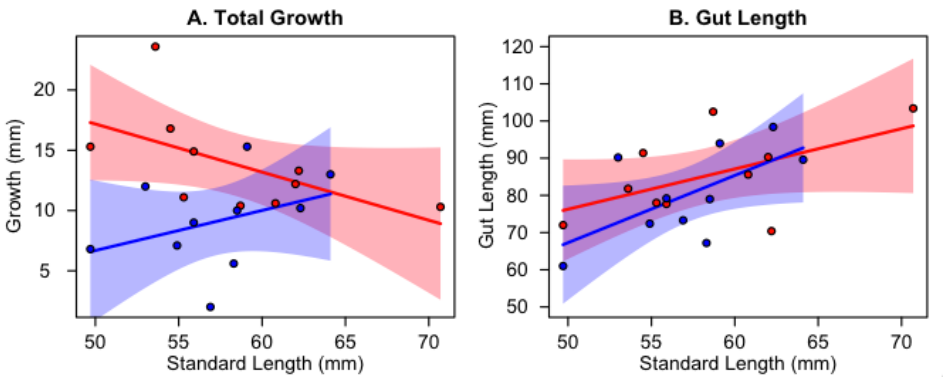

Figure 2. Total growth and gut length.

Individual measurements of growth and gut length plotted against body size (standard length). Individuals in the hard diet treatment are indicated in red, individuals in the soft diet treatment are indicated in blue.

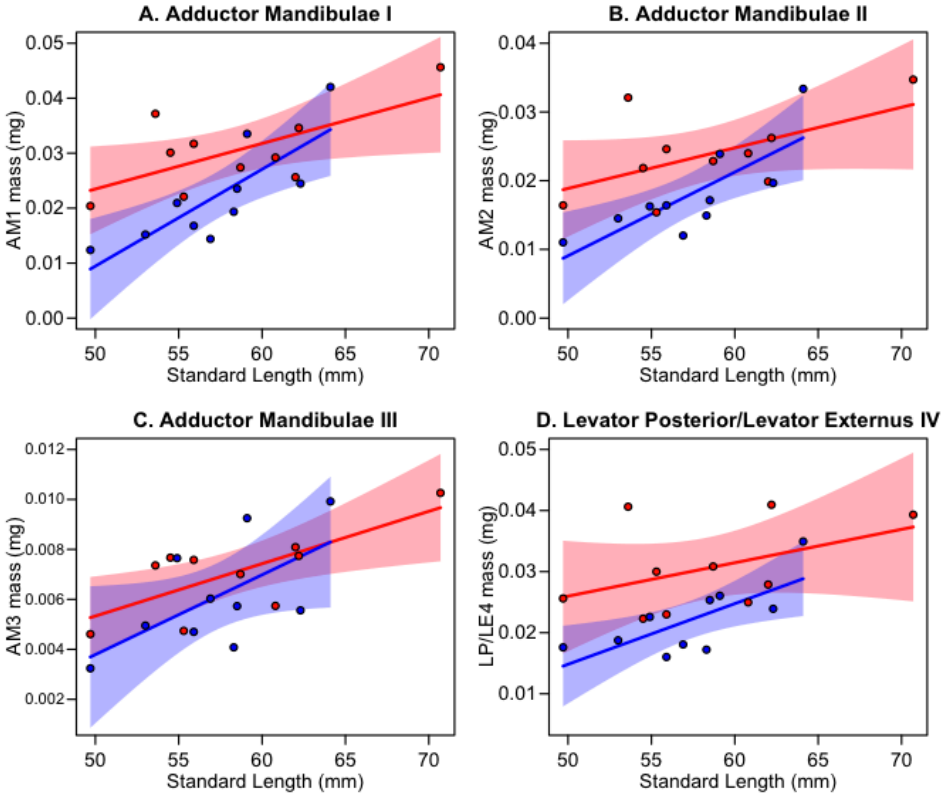

Figure 3. Muscle size.

Individual measurements of oral and pharyngeal jaw muscles plotted against body size (standard length). Individuals in the hard diet treatment are indicated in red, individuals in the soft diet treatment are indicated in blue.

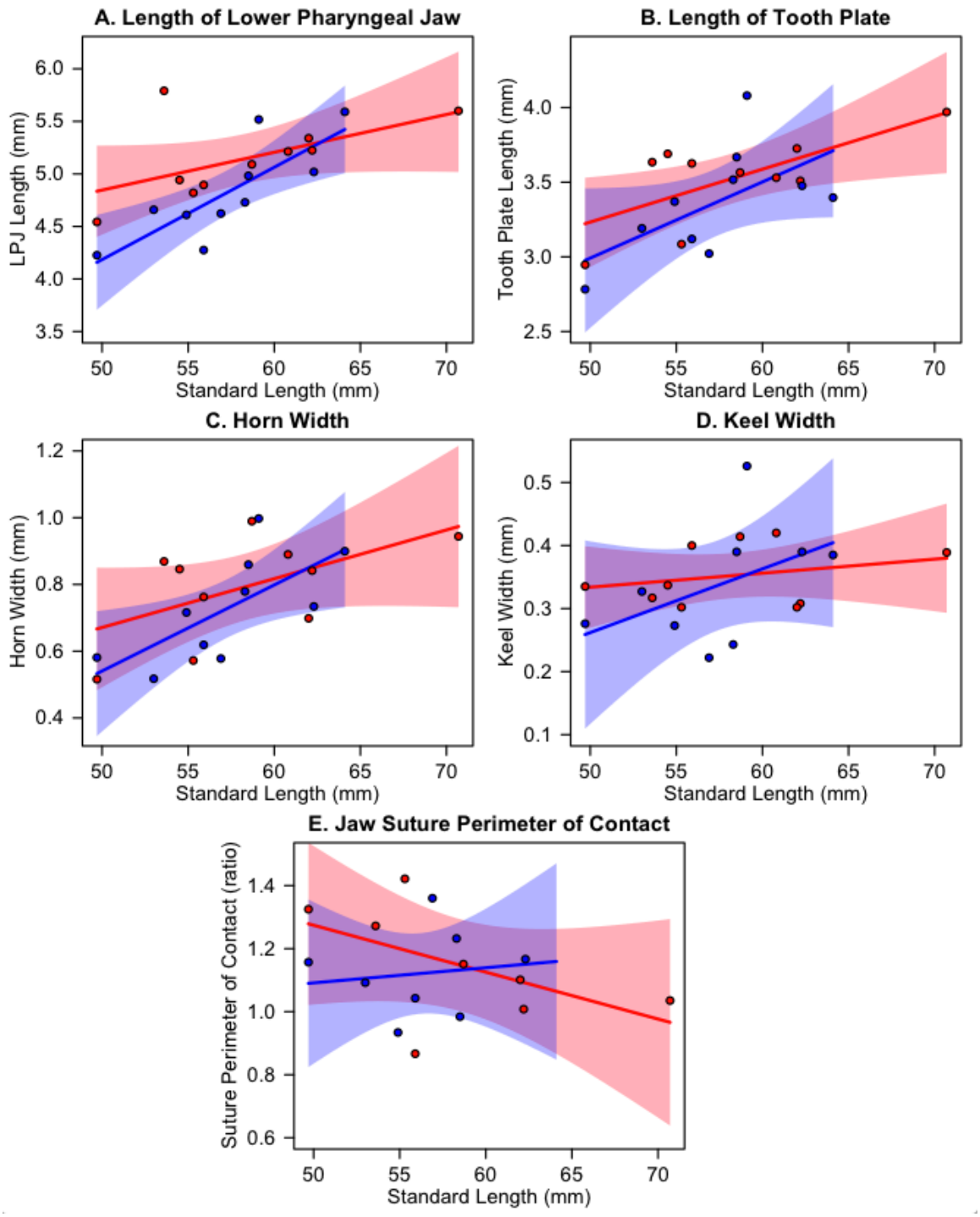

Figure 4. Lower pharyngeal jaw measurements.

Individual measurements of the lower pharyngeal jaw plotted against body size (standard length). Individuals in the hard diet treatment are indicated in red, individuals in the soft diet treatment are indicated in blue.

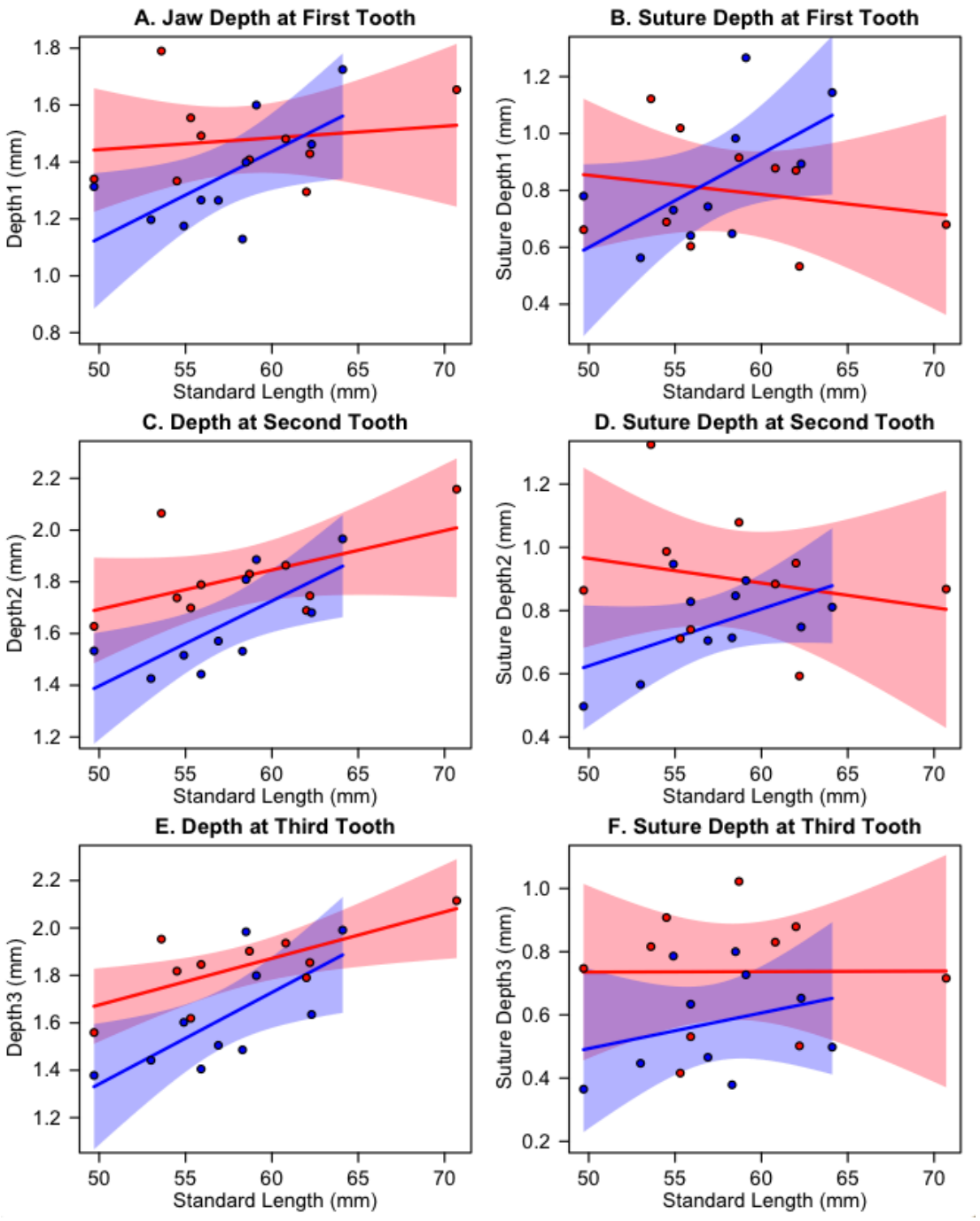

Figure 5. Pharyngeal jaw depth measurements.

Individual measurements of lower pharyngeal jaw (LPJ) depth and depth of LPJ suture plotted against body size (standard length). Individuals in the hard diet treatment are indicated in red, individuals in the soft diet treatment are indicated in blue.

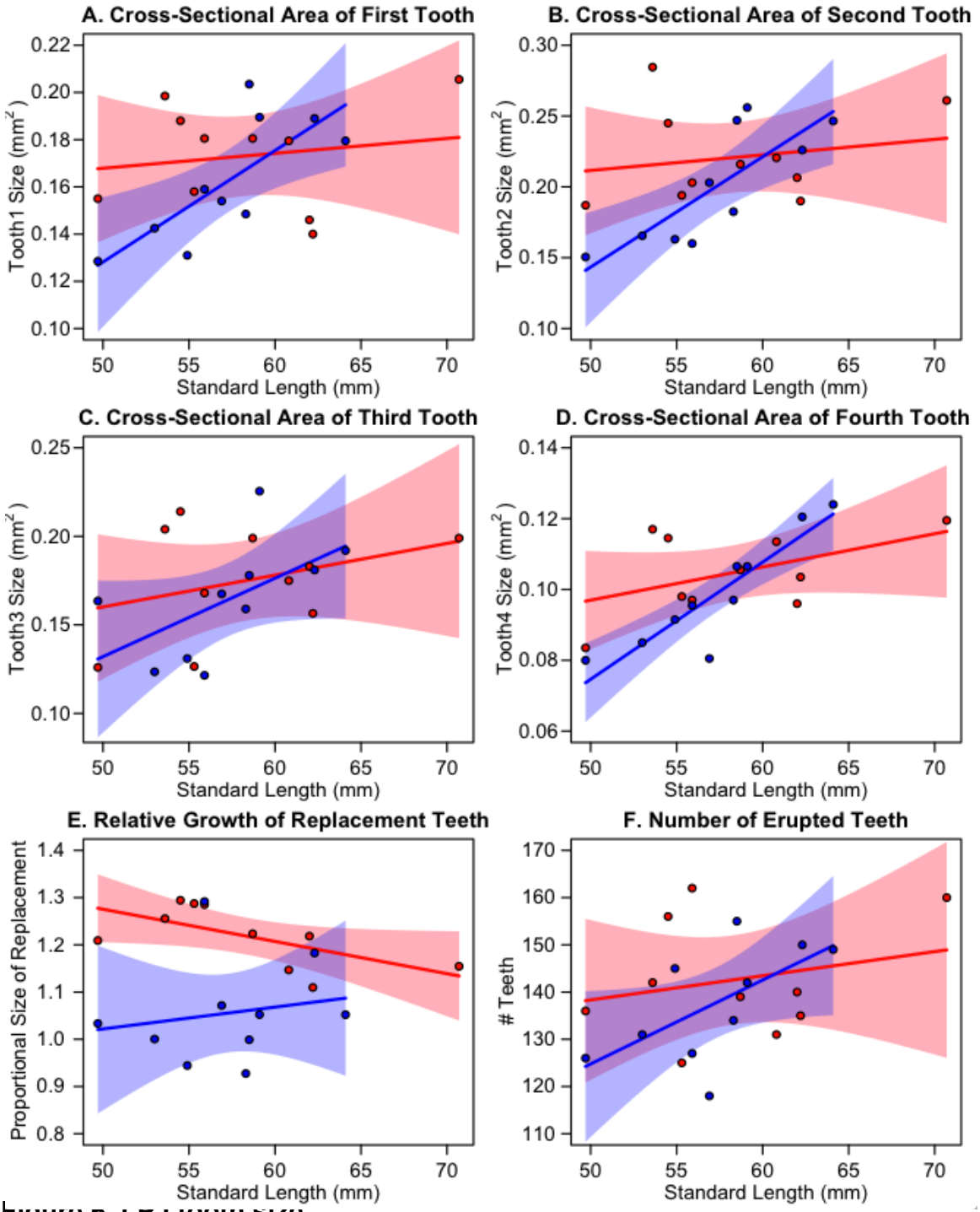

Figure 6. LPJ tooth size.

Individual measurements of pharyngeal teeth plotted against body size (standard length). Individuals in the hard diet treatment are indicated in red, individuals in the soft diet treatment are indicated in blue.

Results

Phenotypic plasticity of the pharyngeal jaw apparatus in V. maculicauda

Analysis of variance on the redundancy analysis (RDA) model found a significant effect of diet type on phenotype (p = 0.034). Within the RDA initial size accounted for 23% of the phenotypic variation. 11% of the phenotypic variation was constrained, or explained by the model, and 65% was unconstrained. After removing the contribution of the conditioning variable, diet type accounted for 15% of the model variation, the first principle components axis (PC1) explained 42%, and PC2 explained 10%. Loadings on the RDA axis indicated traits with the largest effects were growth, mass of the *Adductor Mandibulae III* muscle, the depth of the LPJ at the 2 most-posterior teeth, and the size of replacement teeth relative to erupted teeth (Figure 7). Overall, the RDA indicates that individuals in the crushing treatment had larger, deeper jaws with proportionally larger replacement teeth and larger jaw muscles (Table 1).

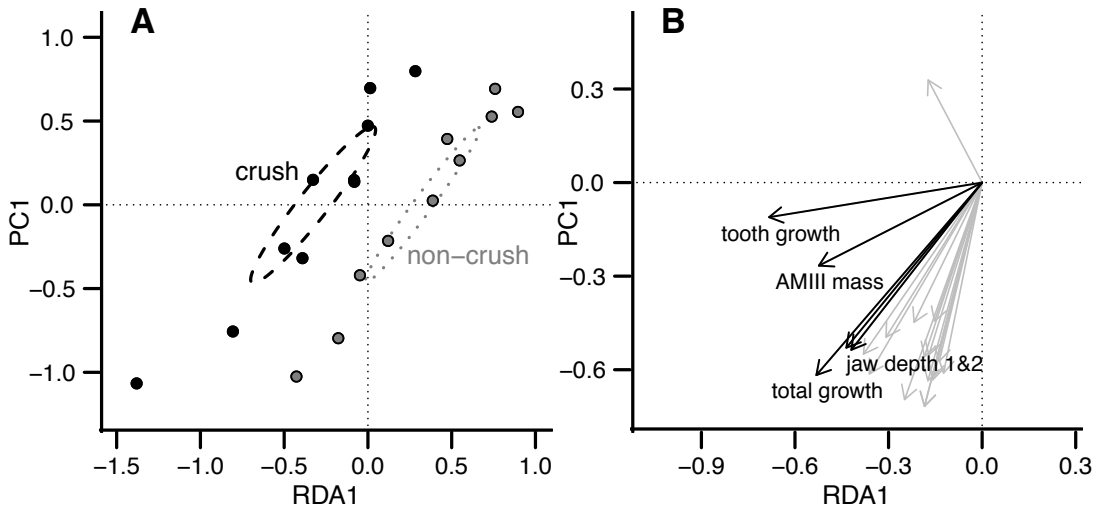

Figure 7. Redundancy analysis.

Redundancy analysis for *Vieja maculicauda* traits. Figure A shows scores for individual fish in the crushing (black) and non-crushing (gray) treatments, with scaling = 1. Figure B shows the trait scores with scaling = 2, the black arrows show the traits with the largest loadings on the RDA axis. Dashed lines in figure A shows 95% confidence ellipses around the treatment means.

Table 1. Morphology RDA results.

Means (unscaled/untransformed) and loadings for RDA (first entry crushing treatment, second non-crushing). NAs indicate traits that were omitted from the RDA due to high pairwise correlation with other variables.

Trait	Mean	Loading	Trait	Mean	Loading
SLinitial (mm)	58.3 57.3	*	LPJ depth (tooth 1) (mm)	1.478 1.353	-0.425
Growth (mm)	13.9 9.1	-0.538	Suture depth (tooth 1) (mm)	0.797 0.839	-0.366
AMI (g)	0.0303 0.0223	-0.221	LPJ depth (tooth 2) (mm)	1.821 1.636	-0.441
AMII (g)	0.0238 0.0179	-0.155	Suture depth (tooth 2) (mm)	0.900 0.756	-0.385
AMIII (g)	0.0071 0.0061	-0.530	LPJ depth (tooth 3) (mm)	1.839 1.623	-0.160
LP/LEIV (g)	0.0305 0.0220	-0.187	Suture depth (tooth 3) (mm)	0.737 0.576	-0.187
Gut (mm)	85.3 80.4	-0.312	Number of teeth	142.6 137.7	NA
LPJ length L_T (mm)	5.146 4.823	-0.176	Tooth 1 size (mm^2)	0.173 0.163	-0.187
Suture Perimeter of Contact P_L	1.148 1.121	-0.175	Tooth 2 size (mm^2)	0.221 0.200	-0.252
Tooth Plate length L_P (mm)	3.528 3.362	-0.139	Tooth 3 size (mm^2)	0.175 0.164	-0.125
Keel width (mm)	0.352 0.337	NA	Tooth 4 size (mm^2)	0.105 0.099	-0.173
Horn width (mm)	0.793 0.728	NA	Proportional tooth growth	1.219 1.056	-0.691

Replacement teeth of V. maculicauda *and* H. minckleyi

In the best-fit model, the size of replacement teeth is strongly predicted by the size of the erupted teeth. Jaw type had a significant effect, but we did not detect an interaction effect between erupted tooth size and jaw type (Table 2). Post-hoc comparison of means detects two groups: jaws with hard diets (*V. maculicauda* crushing & *H. minckleyi* molariform), and jaws with soft diets (*V. maculicauda* non-crushing & *H. minckleyi* papilliform) (Figure 8).

Table 2. ANOVA of redundancy analysis.

Analysis of Variance table, type III with Satterthwaite approximation for degrees of freedom.

	SS	Mean Sq	DenDF	F-value	P (>F)
Erupted tooth size	87.78	87.78	125.5	1064	<< 0.001*
Jaw type	1.440	0.480	55.03	5.82	0.002*
Erupt x type	0.587	0.196	114.5	2.37	0.074

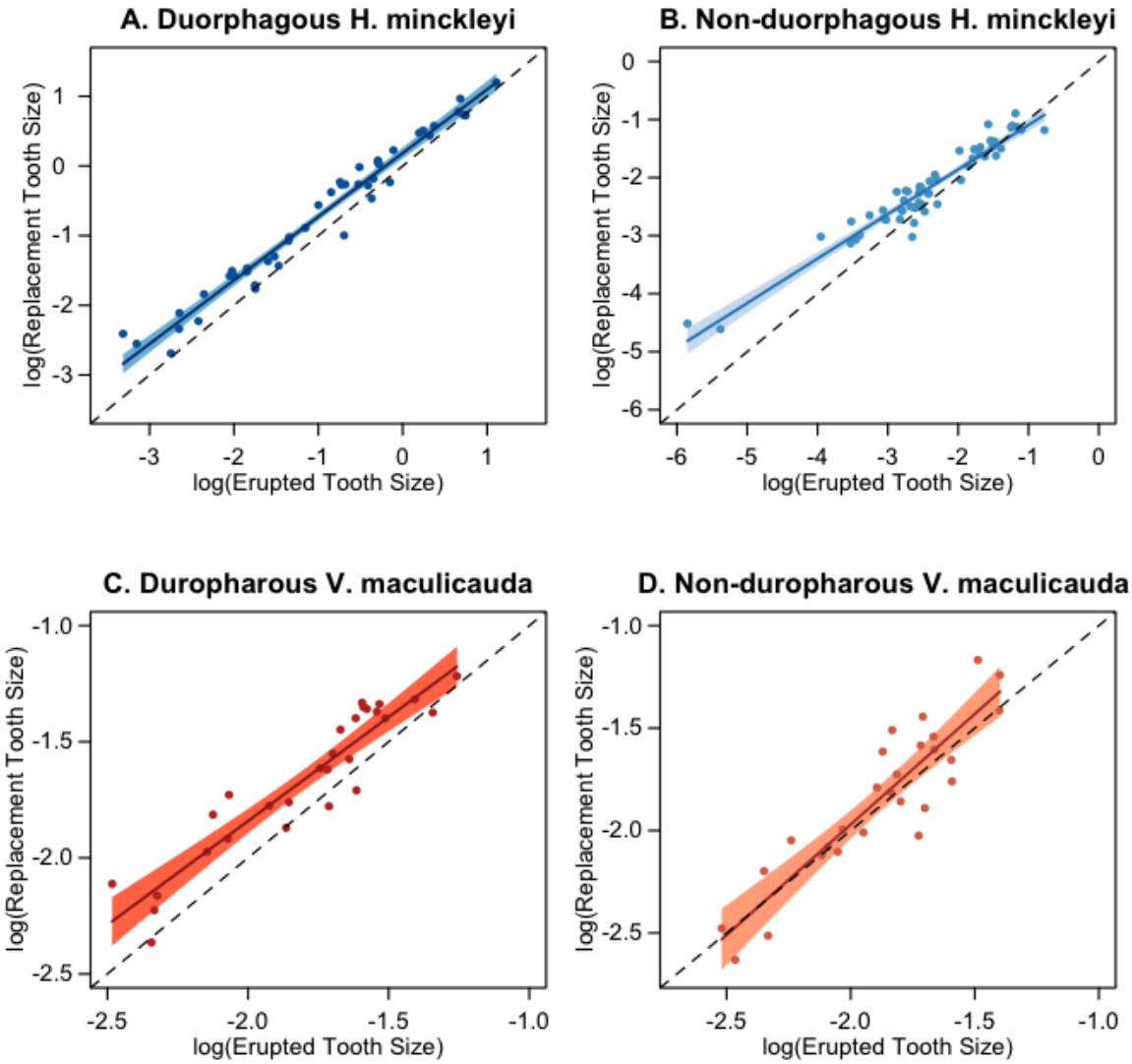

Figure 8. Erupted and replacement tooth comparison.

Replacement size relative to erupted tooth size and jaw type of wild-caught *Herichthys minckleyi* and lab-raised *Vieja maculicauda*. Dashed lines show one-to-one growth (on a log scale).

Discussion

Phenotypic plasticity in the pharyngeal jaw apparatus

Multivariate analysis of the pharyngeal jaw apparatus shows patterns similar to studies of phenotypic plasticity in Old World cichlids. Namely, individuals on more a durophagous diet had larger, deep lower pharyngeal jaws (LPJs), larger pharyngeal teeth, and larger muscles attached to the LPJ. 11% of the phenotypic variation can be attributed to diet change, meaning mechanical stress induced a plastic response that resulted in an 11% phenotypic shift.

Traits with the strongest response, indicated by the RDA loading values (Table 1), included some expected traits such as jaw depth and tooth growth but also traits that we did not expect to have strong plastic responses. The adductor mandibulae III (AMIII) muscle is not attached to the pharyngeal jaw, nor is it usually considered to be associated with prey processing in cichlids. The AMIII is thought to be involved in respiration in teleosts (Osse 1968, von Herbing et al 1996a,b), and we did not expect to find any relationship between diet type and AMIII size. Additionally, diet type can predict gut length in wild cichlids, species with more carnivorous diets have shorter gut lengths (Wagner et al 2009). We therefore predicted that the length of the gut in *V. maculicauda*, if it displayed any plasticity, would be shorter in our crushing treatment. Instead we found that the gut was longer in the crushing treatment (Table 1).

Plasticity in replacement teeth

In a multivariate framework we see that tooth growth, or the proportional increase of replacement teeth relative to the erupted teeth, are different in the crushing and non-crushing treatments. Crushing teeth had relatively larger replacements than non-crushing teeth. From that we expected to see a difference in the slope of an erupted to replacement tooth comparison, a significant interaction effect indicating that for a given tooth size the incoming replacement tooth should be relatively larger for individuals on more durophagous diets. We failed to detect

this difference in a linear mixed model. However, our power in this second analysis was 0.61, indicating that with a larger sample size this effect would likely be apparent.

What we do detect in our linear mixed model is a difference in replacement tooth size. When comparing *Vieja maculicauda* experimental groups to wild-caught *Herichthyes minckleyi*, we find that the replacement teeth of more durophagous types (molariform *H. minckleyi* and crushing *V. maculicauda*) group together and the less durophagous types (papilliform *H. minckley* and non-crushing *V. maculicauda*) group together. Species is not a significant factor, and diet predicts replacement tooth size better than relatedness. Critically, we find that the plastic response in *V. maculicauda* matches the phenotypic differences observed in the polymorphic *H. minckleyi*. We can infer from this that pharyngeal tooth replacement in cichlids is dependent on feedback processes and that erupted teeth are highly dependent on individual feeding history.

Conclusion

Vieja maculicauda, thought to be consistently durophagous in the wild, demonstrated considerable morphological plasticity in response to mechanical strain. Additionally tooth growth between the diet treatments showed similar variation as the trophically polymorphic species *Herichthyes minckleyi*, successfully demonstrating that phenotypic plasticity in the pharyngeal jaw apparatus of cichlids can recapitulate morphological varation observed in the wild in the critically important prey-processing unit of the pharyngeal jaw apparatus.

References

Burress ED (2016) Ecological diversification associated with the pharyngeal jaw diversity of Neotropical cichlid fishes. Journal of Animal Ecology 85: 302-313.

Clabaut C, Bunje PME, Salzburger W, Meyer A (2007) Geometric morphometric analyses provide evidence for the adaptive character of the Tanganyikan cichlid fish radiations. Evolution 61: 560-578.

Cochran-Biederman JL, Winemiller KO (2010) Relationships among habitat, ecomorphology and diets of cichlids in the Bladen River, Belize. Environmental Biology of Fishes 88: 143-152.

Friedman M, Keck BP, Dornburg A, Eytan RI, Martin CH, Hulsey CD, Wainwright PC, Near TJ (2013) Molecular and fossil evidence place the origin of cichlid fishes long after Gondwanan rifting. Proceedings of the Royal Society B. 280: 20131733.

Gunter HM, Fan S, Xiong F, Franchini P, Fruciano C, Meyer A (2013) Shaping development through mechanical strain: the transcriptional basis of diet-induced phenotypic plasticity in a cichlid fish. Molecular Ecology 22: 4516-4531.

von Herbing IH, Miyake T, Hall BK, Boutilier G (1996) Ontogeny of feeding and respiration in larval Atlantic cod *Gadus morhua* (Teleostei, Gadiformes): I. Morphology. Journal of Morphology 227: 15-35.

von Herbing IH, Miyake T, Hall BK, Boutilier G (1996) Ontogeny of feeding and respiration in larval Atlantic cod *Gadus morhua* (Teleostei, Gadiformes): II. Function. Journal of Morphology 227: 37-50.

Hothorn T, Bretz F, Westfall P (2008). Simultaneous inference in general parametric models. Biometrical Journal 50(3), 346--363.

Hulsey CD (2006) Function of a key morphological innovation: fusion of the cichlid pharyngeal jaw. Proceedings of the Royal Society B. 273: 669-675.

Hulsey CD, García de León F (2005) Cichlid jaw mechanics: linking morphology to feeding specialization. Functional Ecology 19: 487-494.

Hulsey CD, García de León F, Rodiles-Hernandez R (2006) Micro- and macroevolutionary decoupling of cichlid jaws: a test of Liem's key innovation hypothesis. Evolution 60: 2096-2109.

Hulsey CD, Hendrickson DA, Garcia de Leon FJ (2005) Trophic morphology, feeding performance and prey use in the polymorphic fish *Herichthys minckleyi*. Evolutionary Ecology Research 7: 1-22.

Hulsey, CD, Roberts RJ, Lin ASP, Guldberg R, Streelman JT (2008) Convergence in a mechanically complex phenotype: detecting structural adaptations for crushing in cichlid fish. Evolution 62: 1587-1599.

Hulsey CD, Hollignsworth P, Fordyce J (2010) Temporal diversification of Central American cichlids. BMC Evolutionary Biology 10: 279.

Huysseune A, Aerts P, Verraes W (1989) Pharyngeal tooth replacement in *Astatotilapia elegans*. Progress in Zoology 35: 480–481.

Kuznetsova A, Brockhoff PB, Christensen RHB (2016). lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-33. https://CRAN.R-project.org/package=lmerTest

Liem K (1973) Evolutionary strategies and morphological innovations: cichlid pharyngeal jaws. Systematic Zoology 22: 425-441.

López-Fernández H, Arbour JH, Winemiller KO, Honeycutt RL (2013) Testing for ancient adaptive radiations in neotropical cichlid fishes. Evolution 67: 1321-1337.

Mayer A (1987) Phenotypic plasticity and heterochrony in Cichlasoma managuense (Pisces, Cichlidae) and their implications for speciation in cichlid fishes. Evolution 41: 1357-1369.

Muschick M, Barluenga M, Salzburger W, Meyer A (2011) Adaptive phenotypic plasticity in the Midas cichlid fish pharyngeal jaw and its relevance in adaptive radiation. BMC Evolutionary Biology 11: 116.

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H (2017) vegan: Community Ecology Package. R package version 2.4-3. https://CRAN.R-project.org/package=vegan

Osse JWM (1968) Functional morphology of the head of the perch (Perca Fluviatilis L.): an electromyographic study. Netherlands Journal of Zoology 19: 289-392.

Pfennig DW (1992) Proximate and functional causes of polyphenism in an anuran tadpole. Functional Ecology 6: 167-174.

Pfennig DW, Wund M, Snell-Rood E, Cruickshank T, Schlichting CD, Moczek AP (2010) Phenotypic plasticity's impacts on diversification and speciation. Trends in Ecology & Evolution. 25: 459-467.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rosset A, Spadola L, Ratib, O (2004) OsiriX: an open-source software for navigating in multidimensional DICOM images. Journal of Digital Imaging 17: 205-216.

Smits JD, Witte F, VanVeen FG (1996) Functional changes in the anatomy of the pharyngeal jaw apparatus of *Astatoreochromis alluaudi* (Pisces, Cichlidae), and their effects on adjacent structures. Biological Journal of the Linnean Society 59: 389-409.

Wagner CE, McIntyre PB, Buels KS, Gilbert DM, Michel E (2009) Diet predicts intestine length in Lake Tanganyika's cichlid fishes. Functional Ecology. 23: 1122–1131.

Wainwright PC, Smith WL, Price SA, Tang KL, Sparks JS, Ferry LA, Kuhn KL, Eytan RI, Near TJ (2012) The evolution of pharyngognathy: the phylogenetic and functional appraisal of the pharyngeal jaw key innovation in labroid fishes and beyond. Systematic Biology 61: 1001-1027.

Wimberger PH (1991) Plasticity of jaw and skull morphology in the neotropical cichlid *Geophagus brasiliensis* and *G. steindachneri*. Evolution 45: 1545-1563.

Wimberger PH (1993) Effects of vitamin C deficiency on body shape and skull osteology in *Geophagus brasiliensis*: implications for interpretations of morphological plasticity. Copeia 1993: 343-351.

Wimberger PH (1994) Trophic polymorphisms, plasticity, and speciation in vertebrates. In: Theory & Application of Fish Feeding Ecology. (Eds)

Stouder DJ, Fresh KL, Feller RJ. University of South Carolina Press. Columbia, SC.

Winemiller KO, Kelso-Winemiller LC, Brenkert AL (1995) Ecological diversification and convergence in fluvial cichlid fishes. Environmental Biology of Fishes 44: 235-261.

Wund M (2012) Assessing the impacts of phenotypic plasticity on evolution. Integrative and comparative biology. 52: 5-15.

CHAPTER II
RATES OF MOLECULAR EVOLUTION CORRELATE TO PLASTICITY IN GENE EXPRESSION IN A COMPONENT OF THE PHARYNGEAL JAW APPARATUS OF A NEOTROPICAL CICHLID FISH

A version of this chapter is being prepared for submission to *Evolution & Development.* S.F. Clemmensen designed experiment and performed all analyses. Co-researcher C.D. Hulsey provided resources and collaborated on experimental design.

Abstract

Plasticity in gene expression is linked to relaxation of purifying selection in discrete polyphenisms, but the relationship between gene expression and evolution of continuous traits is unknown. We gathered transcriptome data from the muscular sling of the pharyngeal jaw apparatus in *Vieja maculicauda* individuals reared on diets with different hardness. We calculated differential gene expression between genes in the different diet treatments, peptide divergence rate (k_A) of each gene, and selection intensity (k_A/k_S) of each gene. We find that non-muscle genes that are under weaker purifying selection also have more peptide divergence from our reference species, and muscle genes that are under stronger purifying selection also have less peptide divergence from our reference species. In our trait with a continuously plastic response to diet (i.e. size), we find that some, but not all genes match expectations from previous work on discrete characters.

Introduction

Phenotypic plasticity, through its combined influence on both morphological novelty and molecular evolution, could strongly impact organismal diversification (Pavey et al. 2010, West-Eberhard 2005, Wund 2012). For instance, traits that are phenotypically plastic could evolve faster than non-plastic traits (Leichty et al. 2012, Wund 2012). Advantageously, quantitative estimates of gene expression now permit us to test whether plastic phenotypes are associated with faster rates of molecular evolution. If genes that have experienced less purifying selection are more often co-opted into novel plastic phenotypes or deleterious mutations in

these genes are less likely to be selected against, genes with biased expression should exhibit greater divergence during macroevolution (Hunt et al. 2011, Leichty et al. 2012, Snell-Rood et al. 2010, 2011). Understanding the relationships between environmentally induced plasticity, gene expression, and molecular evolution could clarify the role that phenotypic plasticity plays in the evolution of highly successful groups like cichlid fishes.

In discrete environmentally-induced polyphenisms, it has been repeatedly found that genes showing biased expression between alternative morphotypes show signatures of relaxed purifying selection. Because genes with biased expression in discrete morphs generally exhibit more peptide divergence, it has been inferred that selection does not act as strongly against genes expressed in only one phenotypic context (Hunt et al. 2011, Leichty et al. 2012, Snell-Rood et al. 2011). These genes expressed in only one context are thought to be "masked" from selection and are therefore less conserved (Snell-Rood et al. 2011). However, discrete environmentally induced alternate phenotypes are relatively rare evolutionary phenomena and are much less common than continuously plastic traits (West-Eberhard 2003). Furthermore, it is unclear if plastically expressed genes associated with continuously variable traits should commonly experience relaxed selection. For instance, quantitatively variable traits like jaw muscles that change continuously in response to stimuli like prey hardness could exhibit a fundamentally different relationship between gene expression and molecular evolution.

Cichlid fishes likely exhibit exceptional phenotypic plasticity in their pharyngeal jaw morphology (Muschick et al. 2011, Meyer 1987, Stauffer and van Snick 2004, Wimberger 1991). Additionally, the musculoskeletal structure of the cichlid pharyngeal apparatus allows these fish to efficiently crush hard-shelled mollusks, a prey type unavailable to many other fish groups (Hulsey et al. 2006)(Fig. 9). This novel configuration has been described as a key innovation, specifically a trait that facilitated an increase in rates of diversification in its clades (Liem 1973, Wainwright et al. 2012). Environmentally induced phenotypic

divergence in traits like the pharyngeal jaw musculature could also have determined which genes have been exposed to selection during cichlid macroevolutionary history (Huysseune 1995, Muschick et al. 2011). Plasticity could commonly dictate expression levels of genes, that when induced during the crushing of hard-shelled prey, were subject to greater adaptive molecular evolution. This prey-induced adaptive gene divergence could readily have translated into more efficient exploitation of a novel adaptive zone that has been essential to the evolutionary success of cichlid fishes.

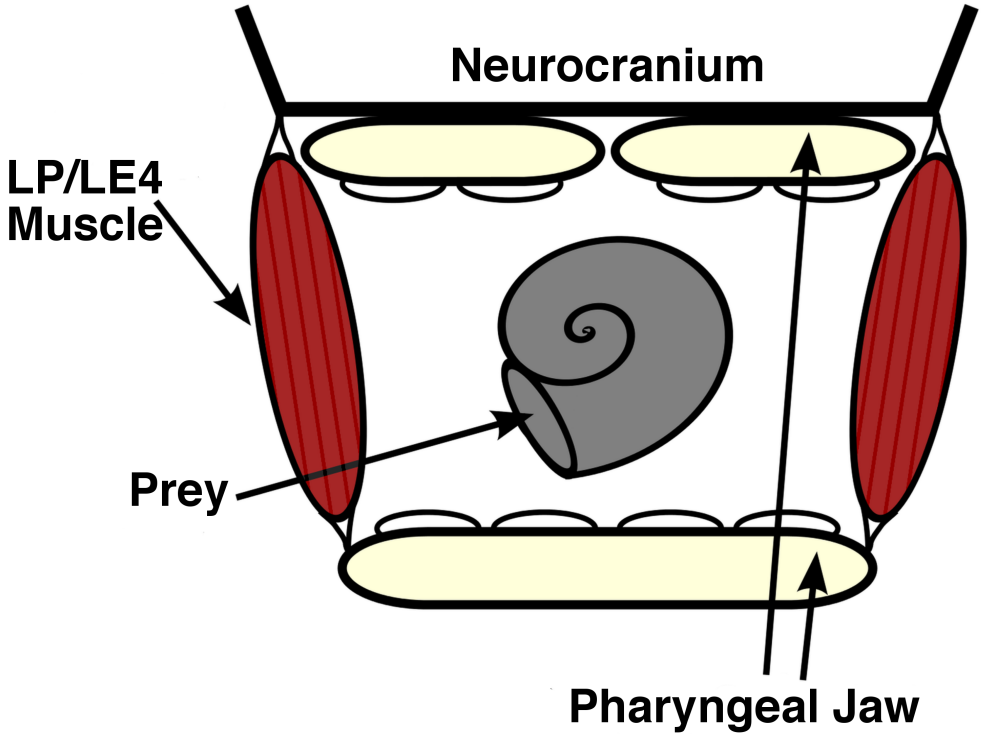

Figure 9. The cichlid pharyngeal jaw apparatus.
Diagram of the cichlid pharyngeal jaw apparatus from an anterior view, with general morphological features labeled.

A number of genes could show differential expression in response to diet hardness in the *Levator Posterior/Levator Externus IV* (LP/LEIV), the two muscles that constitute the muscular sling of the cichlid pharyngeal jaw

apparatus (Liem 1973, Hulsey et al. 2005). For example, genes that code for muscle fibers such as myosin, troponin, and actin have been shown to be up-regulated as muscles grow (Lieber and Lieber 2002, Takada et al. 2001). If increased differential expression in muscle-associated genes correlates to increased sequence diveregence in a continuous trait such as muscle size, these differentially expressed genes may be operating under similar relaxed selection regiemes as has been inferred in genes associated with discrete morphological changes (Hunt et al. 2011, Leichty et al. 2012, Snell-Rood et al. 2011). Examination of the expression and molecular evolution of genes modulated in response to environmentally induced differences in the cichlid LP/LEIV complex could provide insight into the similarities of muscle transcriptome plasticity across all vertebrates.

To determine if genes with plastic expression in a continuous trait experience greater adaptive sequence divergence, we raised individuals of the Central American cichlid *Vieja maculicauda* on soft or hard experimental diets. Using RNAseq, we compared the rates of protein sequence evolution and differential expression in 4,751 genes transcribed in the LP/LEIV complex. After examining expression levels, we subsequently calculated relative sequence divergence for each gene between the cichlid species *V. maculicauda* and *Oreochromis niloticus*. Divergence between these Neotropical and African species subtends the origin of virtually all cichlid diversity (Friedman et al. 2013), and allowed us to generate integrated estimates of molecular evolution across the common ancestor of over 2000 cichlid species.

Materials and Methods

Generation of plastic muscle phenotypes

This study was carried out in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The protocol was approved by the Institutional Animal Care and Use

Committee of the University of Tennessee, Knoxville (Permit Number 1833-0412).

Vieja maculicauda individuals were purchased commercially and fed flake food until they reached a minimum standard length of 49mm standard length (SL, measured from the tip of the maxilla to the base of the caudal fin). These fish were then individually housed for the duration of the feeding experiment. Individuals were randomly assigned to a hard diet treatment (n = 10) or soft diet treatment (n = 16) to induce differences in muscle morphology. Both treatment diets contained the durable freshwater snail, *Melanoides tuberculata*. This invasive snail is found in high densities in several steams near Knoxville, TN. Snails were collected in large numbers, euthanized via freezing, and stored at -5^{o}C until needed for feeding experiments. These snails were generally larger than the gape size of the experimental cichlids, so we processed the snails in a blender to generate the feeding treatments. In the hard diet treatment, snails were “chopped” to create pieces that could be taken into the buccal cavity, but were still sufficiently intact to require the fish to crush the snails with their pharyngeal jaws prior to ingestion. Snails in the soft diet treatment were “pureed” until the snails were fragmented to the point that there was no need for crushing before ingestion. For both treatments, 50g of whole snails were processed as above and added to 14.4g of gelatin, 1.7g of pellet food, and 710mL of water. Each preparation of this mixture produced 60 samples of the 25mL experimental food items. The two experimental diet preparations ensured that no snail pieces were lost and the nutritional content did not differ between treatments. The treatments also created identical food presentation. Individuals were fed the experimental diet three times a week for six months to allow sufficient time for divergence in the structure of the pharyngeal muscles (Huysseune et al 1989).

At the end of the six-month experimental period, we euthanized individuals with tricaine methanesulfonate (MS-222). The *Levator Posterior/Levator Externus IV* pharyngeal muscle complex (LP/LEIV, Figure 9) was immediately dissected from the left side of the fish and stored in RNAlater. Individuals were then

preserved in 10% formalin, and transferred to 70% ethanol for long-term storage. Following fixation, the remaining LP/LEIV muscles on the right sides of the fish were dissected from the pharyngeal apparatus and weighed to the nearest 0.1mg with an electronic balance. To determine if the crushing treatments induced differences in LP/LEIV mass, the muscle masses in hard diet and soft diet groups were compared using a linear model in R Rv3.3.2 with standard length as a covariate (R core team 2016).

RNA alignment and expression

Total RNA was extracted from three LP/LEIV muscles in each treatment group using a Qiagen RNeasy Fibrous Tissue Mini Kit (n = 6). Transcriptomes were generated at the Yale Center for Genome Analysis. Samples were bar-coded and multiplexed on a single lane run on an Illumina HiSeq 2000 (1 x 75bp). Quality filtering using trimmomatic v0.36 sequentially removed adapter sequences, trimmed reads if average quality across 4 bases fell below a threshold phred score of 15, and removed reads with fewer than 30 bases (Bolger et al. 2014). The Illumina run generated ~23million single-end reads for each sample after quality filtering.

Reads were aligned to a reference genome for *Oreochromis niloticus* using TopHat v2.1.1 (Trapnell et al. 2012), using parameters designed to permit cross-species alignments (see Supp. Materials). Reads were processed with the default htseq-count script in HTSeq v0.7.2 to calculate the number of reads that had a single unique alignment for each gene (Anders et al. 2015). Differential expression was calculated from the HTSeq output using default parameters with the *DESeq* function in the DESeq2 package in R (Love et al. 2014).

RNA Assembly, annotation, and divergence from Tilapia

Transcripts were assembled with a genome-guided assembly in Trinity v2.2.0 with default settings (Haas et al. 2013), using the alignment *.bam* file from one

Vieja maculicauda sample. Assembly quality statistics were computed with TransRate v1.0.3 and (Smith-Unna et al. 2016). Individual *V. maculicauda* contigs were aligned to *O. niloticus* RefSeq sequences using the *pairwiseAlignment* function in the Biostrings package in R (Pagès et al. 2016). Amino acid divergence, the number of non-synonymous substitutions per non-synonymous site (k_A), and selection, the ratio of non-synonymous substitution rate and synonymous substitution rate (k_A/k_S), were calculated using the *kaks* function in the seqinr package in R (Charif and Lobry 2007).

Alignments and divergence estimates for each sequence were generated with a custom R script that first identified matching *V. maculicauda* and *O. niloticus* sequences and created a preliminary alignment, identified the proper frame shift to match the *O. niloticus* amino acid sequence, created a protein-guided alignment of the transcripts, and finally calculated k_A and k_A/k_S between *V. maculicauda* and *O. niloticus* (see Appendix).

Our initial data set had 11,327 *Vieja maculicauda* transcripts that matched *Oreochromis niloticus* reference sequences. However, some *V. maculicauda* transcripts mapped to the same *O. niloticus* reference, so these we selected the longest *V. maculicauda* transcript as the best match to *O. niloticus* and discarded the replicate matching transcripts, resulting in 9,926 unique alignments. Because our calculations of divergence were automated, we used conservative filtering on our dataset. First, we removed genes that had a read count less than 5 after normalization in the DESeq2 output (174 genes). To eliminate genes from our dataset that had unrealistic k_A or k_S estimates, we only retained genes from our dataset with $0 \leq k_S$ or $k_A \leq 1$, which removed 355 genes. We then removed 3 genes with $k_A/k_S > 4$.

Categorizing and comparing muscle and non-muscle genes

Gene names for our transcripts were generated from *O. niloticus* RefSeq protein IDs using the *biomaRt* package in R (Durinck et al. 2009). Genes that we were unable to assign a name to were eliminated from the dataset. We identified

genes associated with skeletal muscle by combining a search of the ZFIN database (www.zfin.org) with a list of muscle genes identified by the Gene Ontology Consortium that had literature support for expression in skeletal muscle (Feltrin et al. 2009, Table 3). Out of a putative 288 genes associated with skeletal muscle, we identified 146 in our filtered data set (Table 3). Genes in our dataset that were not part of the list of muscle genes were categorized as "non-muscle" for our analyses. Our final dataset consisted of 4,751 genes categorized as either "muscle" or "non-muscle" gene types, each with values for k_A and k_A/k_S estimates of divergence between *O. niloticus* and *V. maculicauda*, differential expression between diet treatments for *V. maculicauda*, and number of bases (length) in the *O. niloticus* reference sequence.

For each numerical variable, data were log-transformed to meet assumptions of normality in ANOVA residuals. Data were fit to a linear model and then evaluated with an *Anova* test from the car package in R (Fox and Weisberg 2011). We tested the following models:

[1] $$y_{KAi} = \beta_0 + \beta_1 Z_{Ei} + \beta_2 Z_G i + \beta_3 Z_{Ei} Z_{Gi} + \beta_4 Z_{Li} + e_i$$

[2] $$y_{KA/KSi} = \beta_0 + \beta_1 Z_{Ei} + \beta_2 Z_G i + \beta_3 Z_{Ei} Z_{Gi} + \beta_4 Z_{Li} + e_i$$

Where *E* = differential expression; *G* = gene type (muscle or non-muscle); and *L* = length of reference transcript. The reference length was included as a covariate to account for a known positive correlation between the length of a gene and expression levels (Oshlack and Wakefield 2009). We initially included an additional interaction term between reference length and gene type, but this term was not significant in either model and was therefore dropped from our final analysis.

Results

Using a genome reference of related species for transcriptome assembly have comparable error rates to *de novo* assembly followed by annotation via BLAST if

Table 3. Muscle genes.

Full list of muscle-associated genes from the ZFIN database and the Gene Ontology Consortium. Asterisks indicated genes that had hits in our dataset.

Gene	Reference	Gene	Reference
abch1	Popovic et al. 2010	*myh10	Keller et al. 2011
*acta1	Hwang and Sykes 2015	*myh11	Palstra et al. 2014
*acta2	Palstra et al. 2014	myh13	Hwang and Sykes 2015
actg2	Palstra et al. 2014	myh16	Schiaffino and Reggiani 2011
actn2b	Gupta et al. 2012	myh2	Palstra et al. 2014
actn3	Hwang and Sykes 2015	myh3	Hwang and Sykes 2015
adamts10	Brunet et al. 2015	*myh4	Schiaffino and Reggiani 2011
*adamts12	Brunet et al. 2015	myh6	Hwang and Sykes 2015
*adamts14	Brunet et al. 2015	myh7	Chopra et al. 2010
*adamts15a	Brunet et al. 2015	myh8	Hwang and Sykes 2015
adamts16	Brunet et al. 2015	myhz1.1	Sztal et al. 2011
adamts17	Brunet et al. 2015	*myl1	Hwang and Sykes 2015
adamts18	Brunet et al. 2015	myl10	Jackson et al. 2015
adamts2	Brunet et al. 2015	*myl12	Palstra et al. 2014
adamts3	Brunet et al. 2015	*myl13	Schiaffino and Reggiani 2011
*adamts5	Brunet et al. 2015	*myl2	Hwang and Sykes 2015
adamts6	Brunet et al. 2015	myl3	Hwang and Sykes 2015
*adamts7	Brunet et al. 2015	myl4	Hwang and Sykes 2015
adamts8a	Brunet et al. 2015	myl5	Bottinelli and Reggiani 2000
adamts9	Brunet et al. 2015	*myl6	Schiaffino and Reggiani 2011
*adap2	Venturin et al. 2014	myl6b	Schiaffino and Reggiani 2011
*ampd1	Norman et al. 2001	myl7	Hwang and Sykes 2015
*ank3	Kordeli et al. 1998	myl9	Palstra et al. 2014
ankdd1b	Vaz et al. 2017	*mylpfa	Houbrechts et al. 2016
ankrd1	Barash et al. 2007	*myod1	Andrews et al. 2010
ankrd2	Barash et al. 2007	*myog	Minchin et al. 2013
ankrd23	Barash et al. 2007	*myom1	Bottinelli and Reggiani 2000
ap1m1	Zizioli et al. 2010	myom2	Schiaffino and Reggiani 2011
ap1m2	Zizioli et al. 2010	myot	Hwang and Sykes 2015
*apobec2a	Etard et al. 2010	myoz1	Palstra et al. 2014
apobec2b	Etard et al. 2010	myoz2	Palstra et al. 2014
aspolin	Kinoshita et al. 2011	myoz3	Hwang and Sykes 2015
*atp2a1	Palstra et al. 2014	mypn	Hwang and Sykes 2015
*atp2a2	Vangheluwe et al. 2005	*nbr1	Murgiano et al. 2010
*cacng1	Palstra et al. 2014	neb	Hwang and Sykes 2015
cacng5	Klugbauer et al. 2000	*neo1a	Gibert et al. 2011
calm1	Palstra et al. 2014	nes	Vaittinen et al. 2001

Table 3 continued.

Gene	Reference	Gene	Reference
calm2	Toutenhoofd et al. 1998	*nexn	Hassel et al. 2009
*calm3	McGivney et al. 2009	*nfatc1	Ehlers et al. 2014
*camk1db	Senga et al. 2012	*nfatc2	Michel et al. 2004
camk2a	Palstra et al. 2014	notum2	Cantu et al. 2013
*camk2b	Rose & Hargreaves 2003	nppa	Bendig et al. 2006
camk2d	Palstra et al. 2014	*obscn	Raeker et al. 2006
*camk2g	Rose & Hargreaves 2003	*optn	Schilter et al. 2015
*canx	Hung et al. 2013	*park7	Bretaud et al. 2007
*cap1	Effendi et al 2013	pax3	Palstra et al. 2014
*cap2	Effendi et al 2013	pax7	Otten et al. 2012
*capn3	Palstra et al. 2014	pbx1a	Teoh et al. 2010
*capza1	Papa et al. 1999	*pdlim1	Rodriguez et al. 2007
capza2	Papa et al. 1999	pdlim2	Carraro et al. 2009
*capzb	Palstra et al. 2014	*pdlim3	Parker et al. 2008
casq1a	Furlan et al. 2015	pdlim4	Raymond et al. 2010
casq1b	Furlan et al. 2015	*pdlim5	Keller et al. 2011
*cbfb	Meder et al. 2010	*pitx2	Braun and Gautel 2011
*chrna1	Wen et al. 2010	pln	Jorgensen and Jones 1986
*chrnb1	Papke et al. 2012	pml	Kojic et al. 2004
ckm	Palstra et al. 2014	popdc2	Kirchmaier et al. 2012
*ckmt2	Palstra et al. 2014	*ppargc1a	Sandri et al. 2006
*cox4i1	Little et. al 2010	*ppp3ca	Palstra et al. 2014
cox4i2	Little et. al 2010	ppp3r1	He et al. 2011
cox5aa	Little et. al 2010	ppp3r2	He et al. 2011
*cox5ab	Little et. al 2010	*pvalb	Palstra et al. 2014
cox5b1	Little et. al 2010	*rbfox1	Frese et al. 2015
cox5b2	Little et. al 2010	rdh12	Nadauld et al. 2006
cox6a1	Little et. al 2010	rdh5	Nadauld et al. 2006
cox6a2	Little et. al 2010	*rdh8a	Nadauld et al. 2006
*cox6b1	Little et. al 2010	*rgma	Gibert et al. 2011
*cox6b2	Little et. al 2010	rgmb	Gibert et al. 2011
cox8a	Little et. al 2010	rgmd	Gibert et al. 2011
cox8b	Little et. al 2010	*rhoa	Carson and Wei 2000
cryab	Mao & Shelden 2006	ryr1a	Hirata et al. 2007
*csrp3	Kostek et al. 2007	*ryr1b	Hirata et al. 2007
*dag1	Gupta et al. 2011	*ryr3	Darbandi and Franck 2009
dcc	Shen et al. 2002	scn4a	Chopra et al 2010
*des	Palstra et al. 2014	sgca	Lui and Engvall 1999

Table 3 continued.

Gene	Reference	Gene	Reference
*dio1	Houbrechts et al. 2016	sgcb	Lui and Engvall 1999
dio2	Houbrechts et al. 2016	*sgcd	Guyon et al. 2005
dio3a	Houbrechts et al. 2016	*sgce	Lui and Engvall 1999
dio3b	Houbrechts et al. 2016	*sgcg	Lui and Engvall 1999
*dlst	Keßler et al. 2015	slco1d1	Popovic et al. 2010
*dmd	Berger et al. 2010	slco1e1	Popovic et al. 2010
dmn	Carlsson et al. 2000	*slco2a1	Popovic et al. 2010
*dnajb6a	Ding et al. 2016	slco2b1	Popovic et al. 2010
dpf3	Lange et al. 2008	*slco3a1	Popovic et al. 2010
*dtna	Nawrotzki et al. 1998	slco3a2	Popovic et al. 2010
emd	Frock et al. 2006	*slco5a1	Popovic et al. 2010
eno3	Chao et al. 2007	sln	Odermatt et al. 1998
etfa	Kim et al. 2013	smpx	Kemp et al. 2001
etv7	Qunitana et al. 2013	*smyd1b	Du et al. 2006
*fbxo32	Bühler et al. 2016	*smyd5	Fujii et al. 2016
fgf10	Mitchell et al. 1999	smyhc1	Elworthy et al. 2008
*fhl1	Cowling et al. 2008	*snta1	Abramovici et al. 2003
*fhl2	Chu et al. 2000	sntb1	Palstra et al. 2014
*fhl3	Chu et al. 2000	*sntb2	Abramovici et al. 2003
*flnb	Zhou et al. 2007	sntg1	Abramovici et al. 2003
*flnc	Palstra et al. 2014	sntg2	Abramovici et al. 2003
fxr1	Engels et al. 2004	*snx13	Li et al. 2014
gja5a	Tao and Valdimarsson 2010	sod1	Ramesh et al. 2010
gnrh2	Tello et al. 2008	*sox6	Jackson et al. 2015
gnrh3	Tello et al. 2008	spta1	Craig and Pardo 1983
gnrhr3	Tello et al. 2008	sptb	Craig and Pardo 1983
gnrhr4	Tello et al. 2008	*sqstm1	Braun and Gautel 2011
hfe2	Bian et al. 2011	srfa	Davis et al. 2008
*hsp90a	Etard et al. 2010	*sspn	Cohn et al. 2002
*hspb1	Mao & Shelden 2006	sync1	Newey et al. 2001
htr2cl1	Schneider et al. 2012	syne1	Grady et al. 2005
ilk	Meder et al. 2010	*synpo2	Parker et al. 2008
*inppl1	Schurmans et al. 1999	syt2	Wen et al. 2010
*itga7	Mayer 2003	syt7	Wen et al. 2010
*itgb1	Keller et al. 2011	*tbx1	Braun and Gautel 2011
itgb1bp2	Middelbos et al. 2009	*tcap	Hwang and Sykes 2015
itgb2	Marino et al. 2008	tcf21	Lu et al. 2002
junba	Meder et al. 2010	thraa	Houbrechts et al. 2016

Table 3 continued.

Gene	Reference	Gene	Reference
klhl40a	Ravenscroft et al. 2013	*thrb	Houbrechts et al. 2016
*klhl40b	Ravenscroft et al. 2013	*tln1	Belkin et al. 1986
krt19	Lovering et al. 2011	*tmod1	Gokhin et al. 2010
lama2	Sztal et al. 2011	tmod4	Berger et al. 2014
*lamb2	Sztal et al. 2011	*tnnc1	Genge et al. 2013
*lamc1	Sztal et al. 2011	*tnnc2	Hwang and Sykes 2015
ldb3	Hwang and Sykes 2015	tnni1	Hwang and Sykes 2015
lgi1a	Teng et al. 2011	tnni2	Palstra et al. 2014
*lgi1b	Teng et al. 2011	tnni3	Bottinelli and Reggiani 2000
*lmna	Koshimizu et al. 2011	tnnt1	Hwang and Sykes 2015
*lmnb2	Frock et al. 2006	tnnt2	Hwang and Sykes 2015
lmnl1	Frock et al. 2006	*tnnt3	Palstra et al. 2014
lmnl2	Frock et al. 2006	*tp53	Bartlett et al. 2014
lss	Qunitana et al. 2013	*tpm1	Palstra et al. 2014
*lum	Yeh et al. 2010	*tpm2	Jackson et al. 2015
*map1lc3b	Yabu et al. 2012	tpm3	Hwang and Sykes 2015
*mb	Cossins et al. 2009	*tpm4	Palstra et al. 2014
*msc	Lu et al. 1999	*trim54	Perera et al. 2012
mstna	Helterline et al. 2007	*trim55a	Macqueen et al. 2014
mstnb	Helterline et al. 2007	trim55b	Macqueen et al. 2014
*murca	Housley et al. 2016	trim63a	Macqueen et al. 2014
murcb	Housley et al. 2016	trim63b	Macqueen et al. 2014
*mybpc1	Li et al. 2016	*ttn	Hanashima et al. 2016
mybpc2a	Li et al. 2016	*ube4b	Mammen et a.l 2011
mybpc2b	Li et al. 2016	*unc45b	Etard et al. 2010
*mybpc3	Li et al. 2016	*utrn	Helliwell et al. 1992
mybph	Palstra et al. 2014	*vcl	Craig and Pardo 1983
mycbp2	Wang 2010	vim	Palstra et al. 2014
myf5	Palstra et al. 2014	*xirp1	Otten et al. 2012
myf6	Hinist et al. 2007	*ybx1	Kojic et al. 2004
myh1	Hwang and Sykes 2015	zyx	Nix et al. 2001

the species are less than 100my distant (Hornett and Wheat 2012). The origin of the cichlid clade is 65-57mya (Friedman et al. 2013), putting *Vieja maculicauda* and *Oreochromis niloticus* well within an acceptable range for genome-guided analysis of the *V. maculicauda* transcriptome using *O. niloticus* as a reference

(see Appendix for TopHat alignment parameters). For our 6 samples, 52-61% (mean 56.6%) of reads mapped to the reference genome once, and 2.8-3.7% (mean 3.2%) mapped to multiple sites. Differential expression analysis using only diet type as a factor identified 10 genes with significantly different expression levels, all non-muscle genes (Table 4, Figure 10).

The genome-guided transcriptome assembly generated 40,729 contigs, ranging from 224 – 8,326bp (N50 = 673). Quality assessment by Transrate calculated 14,042 contigs with an open reading frame (ORF), 88% of the contig was covered by the ORF (averaged across contigs with ORFs present).

As expected, muscles were larger in the hard diet treatment (Fig. 11). Individuals in the hard diet treatment had LP/LEIV muscles that were ~28% larger than muscles from individuals in the soft diet treatment ($F = 10.78$, $R^2 = 0.55$, $p = 0.004$).

Table 4. Significant differential expression between diet types.

Genes with significant differential expression between hard and soft diet treatments. Positive differential expression (log_2foldchange) indicates over-expression in the hard diet treatment, negative fold-change indicates over-expression in the soft diet treatment

ID	Gene name	Matching Tilapia ID	DiffExp
coch	Cochlin 2C	XM_005455947.2	2.01
gdi1	Rab GDP dissociation inhibitor alpha	XM_003438820.3	1.46
nceh1a	Neutral cholesterol ester hydrolase 1a	XM_005477316.2	1.44
phgdh	D-3-phosphoglycerate dehydrogenase	XM_003447386.3	-1.27
top1	DNA topoisomerase 1	XM_003439033.3	0.86
sult6b1	Sulfotransferase 6B1	XM_003455667.3	-0.86
arhgap26	Rho GTPase-activating protein 26	XM_005452580.2	0.78
qars	Glutamine--tRNA ligase	XM_003439003.2	-0.66
zfpm1	Zinc finger protein ZFPM1	XM_005449222.1	0.62
sema4a	Semaphorin-4A	XM_005476151.2	-0.62

There was a significant interaction between gene type and differential expression in our full model of peptide divergence rate (k_A), so we analyzed each gene type separately and compared slope estimates for muscle and non-muscle genes (Table 5). Muscle genes decreased k_A with increasing differential

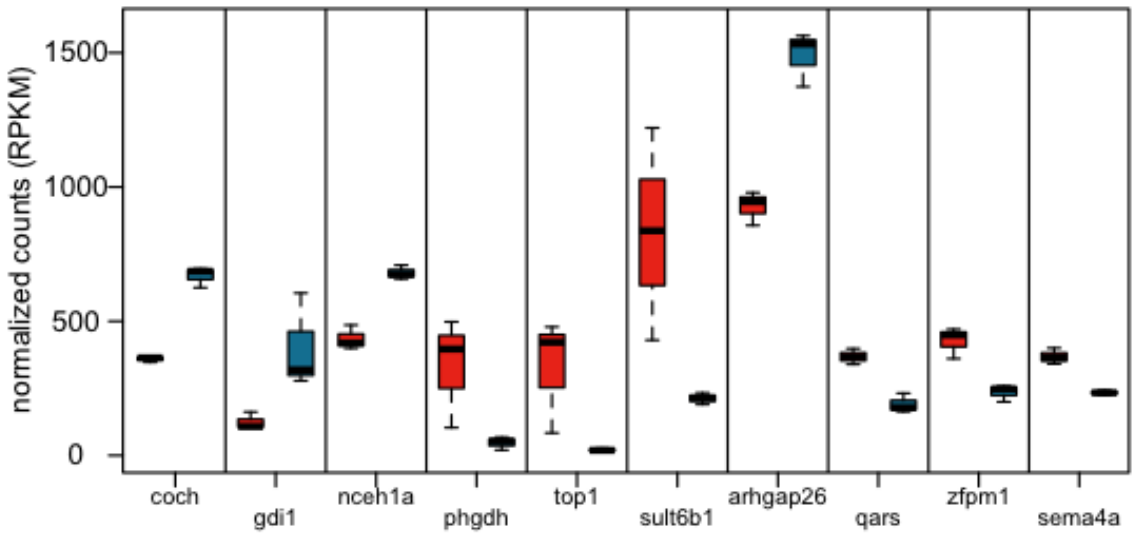

Figure 10. Differential expression between diet treatments.

Genes with significant differential expression between hard and soft diet treatments. Red indicates individuals in the hard diet treatment, blue indicated individuals in the soft diet treatment.

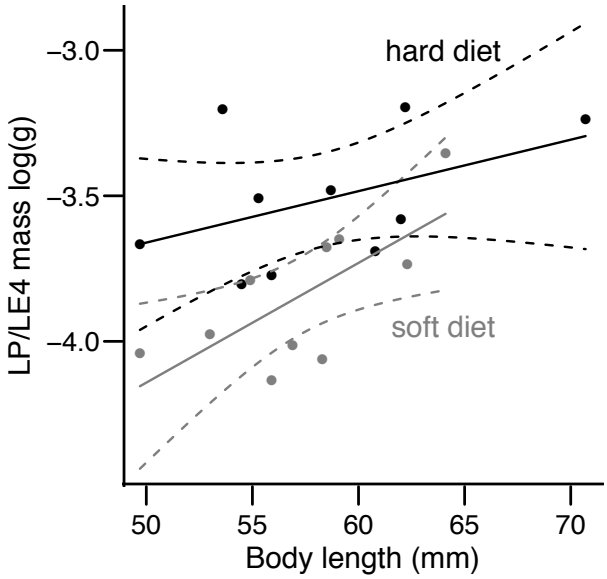

Figure 11. Muscle response to diet type.

LP/LEIV muscles mass plotted against standard length. Individuals in the hard diet treatment (black) had larger muscles than individuals in the soft diet treatment (gray).

Table 5. Peptide divergence rate (k_A).

Results for k_A analysis. All numeric variables were log-transformed and then z-transformed prior to analysis – all estimates are in units of standard deviations. Asterisks indicate significant effects. The estimates of differential expression for muscle and non-muscle genes were significantly different when evaluated with a z-test (Z = -2.943, p = 0.002).

Effect	Estimate	Std Error	SS	Df	F-value	P-value
Full Model (k_A)						
Differential Expression	-0.128	0.082	17.7	1	18.48	< 0.001*
Gene Type	-0.099	0.093	0.0	1	0.001	0.981
Sequence Length	-0.191	0.014	172.1	1	179.5	< 0.001*
DiffExp x Gene Type	0.195	0.083	5.3	1	5.501	0.019*
Residuals			4550.5	4746		
Muscle Genes (k_A)						
Differential Expression	-0.133	0.069	2.458	1	3.663	0.058
Sequence Length	-0.213	0.051	11.60	1	17.29	< 0.001*
Residuals			95.93	143		
Non-muscle Genes (k_A)						
Differential Expression	0.067	0.015	20.7	1	21.35	< 0.001*
Sequence Length	-0.190	-0.015	160.6	1	165.2	< 0.001*
Residuals			4454.5	4602		

expression (β = -0.133, se = 0.069), but we did not detect a significant correlation between differential expression and peptide divergence rate (R^2 = 0.04, p = 0.058). Divergence rate of non-muscle genes increased with increasing differential expression (β = 0.067, se = 0.015). These slopes were significantly different from one another (Z = -2.943, p = 0.002).

For our measure of selection intensity (k_A/k_S) we find that the majority of genes showed signatures of purifying selection: 4,727 of 4,751 genes had k_A/k_S < 1. Overall genes in our data set, k_A/k_S decreased with increasing differential expression (Fig. 12B, Table 6). Our full model did not detect a significant interaction between gene type and differential expression (R^2 = 0.02, p = 0.068), however ANOVA tests assume homoscedasticity between groups while our samples were extremely uneven. A more robust z-test comparing the slopes of muscle and non-muscle genes was significant (Z = -2.254, p = 0.012). Muscle genes decreased in k_A/k_S with increasing differential expression (β = -0.106, se =

Table 6. Selection intensity (k_A/k_S).

Results for k_A/k_S analysis. All numeric variables were log-transformed and then z-transformed prior to analysis – all estimates are in units of standard deviations. Asterisks indicate significant effects. The estimates of differential expression for muscle and non-muscle genes were significantly different when evaluated with a z-test (Z = -2.254, p = 0.012).

Effect	Estimate	Std Error	SS	Df	F-value	P-value
Full Model (k_A/k_S)						
Differential Expression	-0.093	0.083	15.1	1	15.40	< 0.001*
Gene Type	-0.262	0.094	4.8	1	4.883	0.027*
Sequence Length	-0.115	0.014	62.9	1	64.01	< 0.001*
DiffExp x Gene Type	0.154	0.084	3.3	1	3.34	0.068
Residuals			4660.8	4746		
Muscle Genes (k_A/k_S)						
Differential Expression	-0.106	0.073	1.562	1	2.119	0.148
Sequence Length	-0.175	0.054	7.863	1	10.66	0.001*
Residuals			105.43	143		
Non-muscle Genes (k_A/k_S)						
Differential Expression	0.062	0.015	17.2	1	17.43	< 0.001*
Sequence Length	-0.112	0.015	56.0	1	56.57	< 0.001*
Residuals			4554.4	4602		

0.073), but we did not detect a significant correlation (p = 0.148). Non-muscle genes increased in k_A/k_S with (β = 0.062, se = 0.015). Independent of differential expression, muscle genes had overall higher k_A/k_S values than non-muscle genes Z = 2.625, p = 0.004).

Discussion

We set out to determine if there was a relationship within the muscular sling of the pharyngeal jaw in gene expression differences between diet treatments within our study species (differential expression), and metrics of gene evolution between our study species and *Oreochromis niloticus* (k_A and k_A/k_S). Peptide divergence (k_A) is a measurement of the rate of non-synonymous, or mutations within a codon that alter the resulting amino acid, mutations in a gene sequence. Larger k_A values indicate greater peptide divergence between the two species. Selection intensity (k_A/k_S) is a measurement of the ratio of the non-synonymous

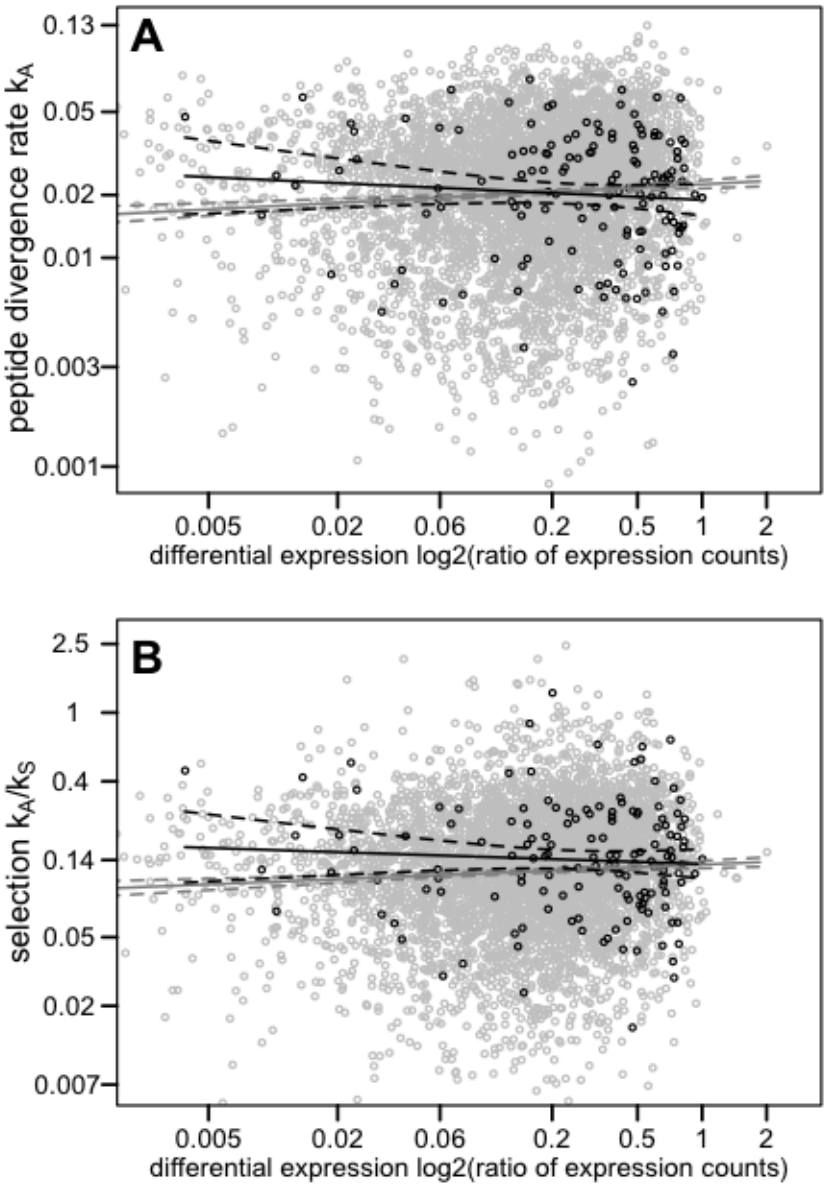

Figure 12. Differential expression plotted against k_A and k_A/k_S.

Differential expression of transcripts in LP/LEIV muscles between hard and soft diets in *Vieja maculicauda* plotted against (A) peptide divergence rate, k_A and (B) selection intensity, k_A/k_S. between *O. niloticus* and *V. maculicauda*. Axes are on a log scale. Muscles-associated genes are in black, all other genes are in gray. Differential gene expression was calculated as the absolute value of $\log_2$(fold-change), resulting in a measure of the magnitude of differential expression between treatments. All data were log-transformed to fit model assumptions, and axes are cropped to maximize visibility (34 and 39 non-muscle genes are not visible in A and B, respectively).

to synonymous (mutations within a codon that do not alter an amino acid) substitution rate, and is a metric of the type of selection pressures a gene has undergone. Values of selection intensity close to zero (few non-synonymous substitutions relative to synonymous substitutions) indicate the gene in our target and reference species are very similar and likely experienced purifying selection, values close to one (non-synonymous and synonymous substitutions are equal) indicate the gene has undergone neutral drift, and values greater than one (more non-synonymous substitutions than synonymous substitutions) indicate a gene has undergone diversifying selection between the two species.

Our goal was to evaluate whether differential gene expression in muscle mass, a continuously variable and highly plastic trait, was associated with patterns in peptide divergence or selection intensity in the pharyngeal jaw muscles associated with prey processing in cichlids. A positive relationship between differential expression and gene evolution is consistent with the hypothesis that genes associated with phenotypically plastic traits should be masked from purifying selection and therefore be more divergent from a reference species than genes that are not associated with plastic traits (Hunt et al. 2011, Leichty et al. 2012, Snell-Rood et al. 2011). While almost all genes in our dataset are characterized by purifying selection (values closer zero than to one), a positive association between k_A/k_S and differential expression may indicate that genes with more plastic expression patterns are under less intense purifying selection than genes with constrained expression patterns, which would be consistent with this hypothesis.

Previous research on discrete traits such as caste differentiation and trait presence/absence predicts a positive relationship between differential expression and both peptide divergence and relaxed purifying selection (Hunt et al. 2011, Leichty et al. 2012, Snell-Rood et al. 2011). Genes that are not associated with skeletal muscle (our “non-muscle” genes) show increasing peptide divergence between *V. maculicauda* and *O. niloticus* with increased differential expression in the pharyngeal jaw muscles. We also found that in this muscle tissue, genes that

are associated with skeletal muscles (our “muscle” genes) show *less* peptide divergence between *V. maculicauda* and *O. niloticus* with increasing differential expression. Our muscle genes were therefore more conserved as differential expression increased. In muscle tissue, our non-muscle genes had a positive relationship between differential expression and k_A/k_S, indicating that as genes became more differentially expressed they were under less purifying selection between *V. maculicauda* and *O. niloticus*. Our muscle gene group had no detectable relationship, with increased differential expression (not significantly associated with k_A/k_S) indicating these genes are in general under purifying selection but are not associated with gene expression plasticity.

Our results are internally consistent, with non-muscle genes that are under weaker purifying selection also have more peptide divergence from our reference species, and muscle genes that are under stronger (but not significantly so) purifying selection also have less peptide divergence from our reference species. Our non-muscle gene results are consistent with the hypothesis that genes that are more plastic in their expression pattern are also masked from purifying selection. However, our muscle genes are inconsistent with this prediction, showing the opposite pattern. With a gene-guided assembly, we might expect a small positive bias between expression and peptide divergence from *O. niloticus* (Hornett and Wheat 2012). This makes our observed negative correlation within the muscle gene group all the more remarkable. This analysis does not consider the effect of within-species variability in genetic sequences such as single-nuceleotide polymorphisms (SNPs). The McDonald-Kreitman (MK) test and its derivatives compare synonymous and non-synonymous SNPs to k_A and k_A/k_S estimates within a gene (Mcdonald and Kreitman 1991, Messer and Petrov 2013), and would improve our estimates of genes under positive selection it attempts to account for genes fixed by genetic drift. However, our individual fish were purchased commercially and we don’t know the relatedness of our individuals, which would compromise our ability to detect intraspecific variation as SNPs are less likely within a brood.

Our observation that muscle genes with more differential expression show lower rates of peptide divergence (k_A) stands in contrast to previous work on discrete polyphenisms (Hunt et al. 2011, Leichty et al. 2012, Snell-Rood et al. 2011). One possible explanation for these different patterns may be that genes associated with pharyngeal jaw muscles will also be associated with other skeletal muscles, and therefore a positive correlation between gene expression plasticity and sequence divergence will depend heavily on those genes being associated primarily and specifically with tissues that exhibit plastic phenotypes. In this case, we might expect to find increased divergence in the regulatory regions of muscle genes, rather than the protein-coding sequences. Non-coding regions would be undetectable in analyses like this one because they are not part of the transcriptome. Additionally, in order for our selection criteria to be reproducible both transcription factors and genes that encode the structural constituents of muscles were binned in the “muscle” groups. Coding regions of these structural genes are often highly conserved (Weiss et al. 1999) and grouping them with transcription factors in our analyses may conflate different relationships between gene expression plasticity and peptide divergence. Expression patterns for some transcription factors in our muscle group are shown in Figure 13, and their effects described in Table 7.

Gene expression plasticity may be associated with peptide divergence and relaxed purifying selection in some cases, but it is unclear at this time if this pattern is restricted to regulatory elements or if this relationship is restricted to discrete environmentally induced alternate phenotypes that are relatively rare compared to continuously variable traits. Nevertheless, we have detected patterns in peptide divergence from *Oreochromis niloticus* and selection intensity associated with gene expression plasticity in *Vieja maculicauda*, indicating that this plasticity has likely played a role in organismal diversification.

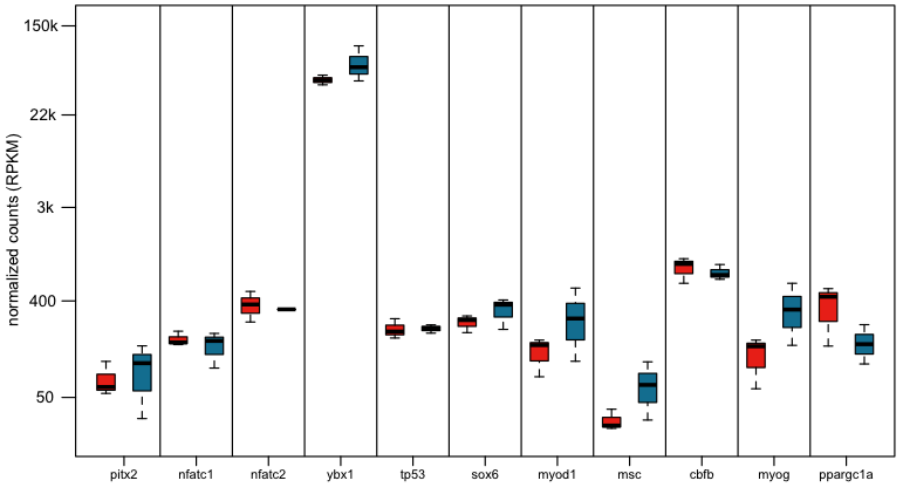

Figure 13. Differential expression in transcriptions factors.

Differential expression of transcripts in LP/LEIV muscles between hard (red) and soft (blue) diets in *Vieja maculicauda* for different transcription factors.

Table 7. Transcription factors.

Some transcription factors found in our dataset that regulate muscle growth, and their known regulatory functions in vertebrates. If the transcription factor has regulatory activity in both embryo and post-embryonic muscles, only the post-embryonic actions are listed.

ID	Gene	Regulatory Action (in muscles)
pitx2	pituitary homeobox 2	In embryo, activates genes that induce terminal differentiation in muscle cells (Braun and Gautel 2011)
nfact1	nuclear factor of activated T-cells 1	Downregulates myod and involved in fast- to slow-twitch muscle fiber-type differentiation (Ehlers et al. 2014)
nfact2	nuclear factor of activated T-cells 2	Involved in muscle fiber-type differentiation and hypertrophy by activating myosin heavy chain genes (Michel et al. 2004)
ybx1	y-box binding protein 1	Downregulates tp53 (Kojic et al. 2004)
tp53	cellular tumor antigen p53	Induces mitochondrial biogenesis (Bartlett et al. 2014)
sox6	Transcription factor SOX-6	In embryo, promotes differentiation of fast-twitch muscle fiber type (Jackson et al. 2015)
myod1	myoblast-determination protein 1	Induces myogenesis, involved in circadian maintenance of muscle tissue (Andrews et al. 2010)
msc	musculin	Inhibits myogenesis, myod antagonist (Lu et al. 1999)
cbfb	core-binding factor β	Maintains sarcomere structure (Meder et al. 2010)
myog	myogenin	Involved in muscle fiber-type differentiation by increasing protein degradation (Braun and Gautel 2011)
ppargc1	peroxisome proliferator-activated receptor γ, coactivator 1	Suppresses atrophy in fast-twitch fibers, promotes fast- to slow-twitch muscle fiber-type differentiation (Sandri et al. 2006)

References

Abramovici H, Hogan AB, Obagi C (2003) Diacylglycerol kinase-ζ localization in skeletal muscle is regulated by phosphorylation and interaction with syntrophins. Molecular Biology of the Cell 14: 4499-4511.

Anders S, Pyl PT, Huber W (2015) HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics 31: 166-169.

Andrews JL, Zhang X, McCarthy JJ, McDearmon EL, Hornberger TA, Russell B, Campbell KS et al. (2010) CLOCK and BMAL1 regulate MyoD and are necessary for maintenance of skeletal muscle phenotype and function. Proceedings of the National Academy of Sciences 107:19090-19095.

Barash IA, Bang M-L, Mathew L, Greaser ML, Chen J, Lieber RL (2007) Structural and regulatory roles of muscle ankyrin repeat protein family in skeletal muscle. American Journal of Physiology – Cell Physiology 293: C218-C227.

Bartlett JD, Close GL, Drust B, Morton JP (2014) The emerging role of p53 in exercise metabolism. Sports medicine 44:303-309.

Belkin AM, Zhidkova NI, Koteliansky VE (1986) Localization of talin in skeletal and cardiac muscles. FEBS letters 200: 32-36.

Berger J, Berger S, Hall TE, Lieschke GJ, Currie PD (2010) Dystrophin-deficient zebrafish feature aspects of the Duchenne muscular dystrophy pathology. Neuromuscular Disorders 20: 826-832.

Berger J, Tarakci H, Berger S, Li M, Hall TE, Arner A, Currie PD (2014) Loss of tropomodulin4 in the zebrafish mutant träge causes cytoplasmic rod formation and muscle weakness reminiscent of nemaline myopathy. Disease Models and Mechanisms 7: 1407-15.

Bendig G, Grimmler M, Huttner IG, Wessels G, Dahme T, Just S, Trano N et al. (2006) Integrin-linked kinase, a novel component of the cardiac mechanical stretch sensor, controls contractility in the zebrafish heart. Genes and Development 20: 2361-2372.

Bian YH, Xu C, Li J, Xu J, Zhang H, Du SJ (2011) Development of a transgenic zebrafish model expressing GFP in the notochord, somite and liver directed by the hfe2 gene promoter. Transgenic Research 20: 787-98.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30: 2114-2120.

Braun T, Gautel M. (2011) Transcriptional mechanisms regulating skeletal muscle differentiation, growth and homeostasis. Nature Reviews Molecular Cell Biology 12: 349-361.

Bretaud S, Allen C, Ingham PW, Bandmann O (2007) p53-dependent neuronal cell death in a DJ-1-deficient zebrafish model of Parkinson's disease. Journal of Neurochemistry 100: 1626-1635.

Brunet FG, Fraser FW, Binder MJ, Smith AD, Kintakas C, Dancevic CM, Ward AC, McCulloch DR (2015) The evolutionary conservation of the A Disintegrin-like and Metalloproteinase domain with Thrombospondin-1 motif metzincins across vertebrate species and their expression in teleost zebrafish. BMC Evolutionary Biology 15: 22.

Bottinelli R, Reggiani C (2000) Human skeletal muscle fibers: molecular and functional diversity. Progress in Biophysics and Molecular Biology 73: 195-262.

Bühler A, Kustermann M, Bummer T, Rottbauer W, Sandri M, Just S (2016) Atrogin-1 deficiency leads to myopathy and heart failure in zebrafish. International Journal of Molecular Sciences 17.

Cantu JA, Flowers GP, Topczewski J (2013) Notum homolog plays a novel role in primary motor innervation. Journal of Neuroscience 33: 2177-2187.

Carlsson L, Li ZL, Paulin D, Price MG, Breckler J, Robson RM, Wiche G, Thornell LE (2000) Differences in the distribution of synemin, paranemin, and plectin in skeletal muscles of wild-type and desmin knock-out mice. Histochemistry and Cell Biology 114: 39-47.

Carraro L, Ferraresso S, Cardazzo B, Romualdi C, Montesissa C, Gottardo F, Patarnello T, Castagnaro M, Bargelloni L (2009) Expression profiling of skeletal muscle in young bulls treated with steroidal growth promoters. Physiological Genomics 38: 138-148.

Carson JA, Wei L. Integrin signaling’s potential for mediating gene expression in hypertrophying skeletal muscle (2000) Journal of applied physiology 88: 337-343.

Chao LC, Zhang Z, Pei L, Saito T, Tontonoz P, Pilch PF (2007) Nur77 coordinately regulates expression of genes linked to glucose metabolism in skeletal muscle. Molecular Endocrinology 21: 2152-2163.

Charif D, Lobry JR (2007) SeqinR 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. Eds: Bastolla U, Porto M, Roman HE, Vendruscolo M. Springer, New York pp 207-232.

Chopra SS, Stroud DM, Watanabe H, Bennett JS, Burns CG, Wells KS, Yang T, Zhong TP, Roden DM (2010) Voltage-Gated Sodium Channels Are Required for Heart Development in Zebrafish. Circulation Research 106: 1342-1350.

Chu P-H, Ruiz-Lozano P, Zhou Q, Cai C, Chen J (2000) Expression patterns of FHL/SLIM family members suggest important functional roles in skeletal muscle and cardiovascular system. Mechanisms of Development 1-2: 259-265.

Cohn RD, Henry MD, Michele DE, Barresi R, Saito F, Moore SA, Flanagan JD, Skwarchuk MW, Robbins ME, Mendell JR, Williamson RA (2002) Disruption of DAG1 in differentiated skeletal muscle reveals a role for dystroglycan in muscle regeneration. Cell 110: 639-648.

Cossins AR, Williams DR, Foulkes NS, Berenbrink M, Kipar A (2009) Diverse cell-specific expression of myoglobin isoforms in brain, kidney, gill and liver of the hypoxia-tolerant carp and zebrafish. Journal of Experimental Biology 212: 627-638.

Cowling BS, McGrath MJ, Nguyen M-A, Cottle DL, Kee AJ, Brown S, Schessi J, et al. (2008) Identification of FHL1 as a regulator of skeletal muscle mass: implications for human myopathy. Journal of Cell Biology 183: 1033

Craig SW, Pardo JV (1983) Gamma actin, spectrin, and intermediate filament proteins colocalize with vinculin at costomeres, myofibril-to-sarcolemma attachment sites. Cytoskeleton 3: 449-62.

Darbandi S, Franck JP (2009) A Comparative Study of Ryanodine Receptor (RyR) Gene Expression Levels in a Basal Ray-Finned Fish, Bichir (*Polypterus ornatipinnis*) and the Derived Euteleost Zebrafish (*Danio rerio*). Comparative Biochemistry and Physiology Part B Biochemistry and Molecular Biology 154: 443-448.

Davis JL, Long X, Georger MA, Scott IC, Rich A, Miano JM (2008) Expression and comparative genomics of two serum response factor genes in zebrafish. International Journal of Developmental Biology 52: 389-396.

Ding Y, Long PA, Bos JM, Shih YH, Ma X, Sundsbak RS, Chen J et al. (2016) A modifier screen identifies DNAJB6 as a cardiomyopathy susceptibility gene. JCI Insight 1.

Du SJ, Rotllant J, Tan X (2006) Muscle-specific expression of the smyd1 gene is controlled by its 5.3-kb promoter and 5'-flanking sequence in zebrafish embryos. Developmental Dynamics 235: 3306-3315.

Durinck S, Spellman PT, Birney E, Huber W (2009) Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nature Protocols 4: 1184-1191.

Effendi K, Yamazaki K, Mori T, Masugi Y, Makino S, Sakamoto M (2013) Involvement of hepatocellular carcinoma biomarker, cyclase-associated protein 2 in zebrafish body development and cancer progression. Experimental Cell Research 319: 35-44.

Ehlers ML, Celona B, Black BL (2014) NFATc1 controls skeletal muscle fiber type and is a negative regulator of MyoD activity. Cell reports. 8:1639-1648.

Elworthy S, Hargrave M, Knight R, Mebus K, Ingham PW (2008) Expression of multiple slow myosin heavy chain genes reveals a diversity of zebrafish slow twitch muscle fibres with differing requirements for Hedgehog and Prdm1 activity. Development 135: 2115-2126.

Engels B, van't Padje S, Blonden L, Severijnen LA, Oostra BA, Willemsen R (2004) Characterization of Fxr1 in Danio rerio; a simple vertebrate model to study costamere development. Journal of Experimental Biology 207: 3329-3338.

Etard C, Roostalu U, Strähle U (2010) Lack of Apobec2-related proteins causes a dystrophic muscle phenotype in zebrafish embryos. Journal of Cell Biology 189: 527-539.

Feltrin E, Campanaro S, Diehl AD, Ehler E, Faulkner G, Fordham J, Gardin C, et al. (2009) Muscle research and gene ontology: new standards for improved data integration. BMC Medical Genomics 2: 6.

Fox J, Weisberg S (2011) An R companion to applied regression, 2nd Ed. Thousand Oaks, CA: Sage. URL: http://socserv.socsci.mcmaster.ca/jfox/Books/Companion.

Frese KS, Meder B, Keller A, Just S, Haas J, Vogel B, Fischer S et al. (2015) RNA splicing regulated by RBFOX1 is essential for cardiac function in zebrafish. Journal of Cell Science 128: 3030-3040.

Friedman M, Keck BP, Dornburg A, Eytan RI, Martin CH, et al. (2013) Molecular and fossil evidence place the origin of cichlid fishes long after Gondwanan rifting. Proc Biol Sci 280: 20131733.

Frock RL, Kudlow BA, Evans AM, Jameson SA, Hauschka SD, Kennedy BK (2006) Lamin A/C and emerin are critical for skeletal muscle satellite cell differentiation. Genes and Development 20: 486-500.

Fujii T, Tsunesumi SI, Sagara H, Munakata M, Hisaki Y, Sekiya T, Furukawa Y et al. (2016) Smyd5 plays pivotal roles in both primitive and definitive hematopoiesis during zebrafish embryogenesis. Scientific Reports 6: 29157.

Furlan S, Mosole S, Murgia M, Nagaraj N, Argenton F, Volpe P, Nori A (2015) Calsequestrins in skeletal and cardiac muscle from adult *Danio rerio*. Journal of Muscle Research and Cell Motility 37: 27-39.

Genge CE, Davidson WS, Tibbits GF (2013) Adult teleost heart expresses two distinct troponin C paralogs: cardiac TnC and a novel and teleost-specific ssTnC in a chamber- and temperature-dependent manner. Physiological Genomics 45: 866-875.

Gibert Y, Lattanzi VJ, Zhen AW, Vedder L, Brunet F, Faasse SA, Babitt JL et al. (2011) BMP signaling modulates hepcidin expression in zebrafish embryos independent of hemojuvelin. PLoS One 6: e14553.

Gokhin DS, Lewis RA, McKeown CR, Nowak RB, Kim NE, Littlefield RS, Lieber RL, Fowler VM (2010) Tropomodulin isoforms regulate thin filament pointed-end capping and skeletal muscle physiology. The Journal of Cell Biology. 189: 95-109.

Grady RM, Starr DA, Ackerman GL, Sanes JR, Han M (2005) Syne proteins anchor muscle nuclei at the neuromuscular junction. Proceedings of the National Academy of Sciences of the United States of America 102: 4359-4364.

Gupta V, Discenza M, Guyon JR, Kunkel LM, Beggs AH (2012) α-Actinin-2 deficiency results in sarcomeric defects in zebrafish that cannot be rescued by α-actinin-3 revealing functional differences between sarcomeric isoforms. FASEB journal 26: 1892-1908.

Gupta V, Kawahara G, Gundry SR, Chen AT, Lencer WI, Zhou Y, Zon LI, Kunkel LM, Beggs AH (2011) The zebrafish dag1 mutant: A novel genetic model for dystroglycanopathies. Human Molecular Genetics 20: 1712-1725.

Guyon JR, Mosley AN, Jun SJ, Montanaro F, Steffen LS, Zhou Y, Nigro V, Zon LI, Kunkel LM (2005) delta-Sarcoglycan is required for early zebrafish muscle organization. Experimental Cell Research 304: 105-115.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB et al. (2013) *De novo* transcript sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity. Nature Protocols 8: 1494-1512.

Hanashima A, Hashimoto K, Ujihara Y, Honda T, Yobimoto T, Kodama A, Mohri S (2016) Complete primary structure of the I-band region of connectin at which mechanical property is modulated in zebrafish heart and skeletal muscle. Gene 596: 19-26.

Hassel D, Dahme T, Erdmann J, Meder B, Huge A, Stoll M, Just S et al. (2009) Nexilin mutations destabilize cardiac Z-disks and lead to dilated cardiomyopathy. Nature Medicine 15: 1281-1288.

He ZH, Hu Y, Li YC, Yvert T, Santiago C, Gomez-Gallego F, Ruiz JR, Lucia A (2011) Are calcineurin genes associated with athletic status? A function, replication study. Medicine & Science in Sports & Exercise 43: 1433-1440.

Helliwell TR, Morris GE, Davies KE (1992) The dystrophin-related protein, utrophin, is expressed on the sarcolemma of regenerating human skeletal muscle fibres in dystrophies and inflammatory myopathies. Neuromuscular Disorders 2: 177-184.

Helterline DL, Garikipati D, Stenkamp DL, Rodgers BD (2007) Embryonic and tissue-specific regulation of myostatin-1 and -2 gene expression in zebrafish. General and Comparative Endocrinology 151: 90-97.

Hinits Y, Osborn DP, Carvajal JJ, Rigby PW, Hughes SM (2007) Mrf4 (myf6) is dynamically expressed in differentiated zebrafish skeletal muscle. Gene Expression Patterns 7: 738-745.

Hirata H, Watanabe T, Hatakeyama J, Sprague SM, Saint-Amant L, Nagashima A, Cui WW, Zhou W, Kuwada JY (2007) Zebrafish relatively relaxed mutants have a ryanodine receptor defect, show slow swimming and provide a model of multi-minicore disease. Development 134: 2771-2781.

Horrnett EA, Wheat CW (2012) Quantitative RNA-Seq analysis in non-model species: assessing transcriptome assemblies as a scaffold and the utility of evolutionary divergent genomic reference species. BMC Genomics 13: 361.

Houbrechts AM, Delarue J, Gabriëls IJ, Sourbron J, Darras VM (2016) Permanent deiodinase type 2 deficiency strongly perturbs zebrafish development, growth, and fertility. Endocrinology 157: 3668-3681.

Housley MP, Njaine B, Ricciardi F, Stone OA, Hölper S, Krüger M, Kostin S, Stainier DY (2016) Cavin4b/Murcb is required for skeletal muscle development and function in zebrafish. PLoS Genetics 12: e1006099.

Hulsey CD, García de León FJ, Rodiles-Hernandez R (2006) Micro- and macroevolutionary decoupling of cichlid jaws: a test of Liem's key innovation hypothesis. Evolution 60: 2096–2109.

Hung IC, Cherng BW, Hsu WM, Lee SJ (2013) Calnexin is required for zebrafish posterior lateral line development. International Journal of Developmental Biology 57: 427-438.

Hunt BG, Ometto L, Wurm Y, Shoemaker D, Soojin VY, Keller L, Goodisman MA (2011) Relaxed selection is a precursor to the evolution of phenotypic plasticity. Proceedings of the National Academy of Sciences 108: 15936–15941.

Huysseune A (1995) Phenotypic plasticity in the lower pharyngeal jaw dentition of Astatoreochromis alluaudi (Teleostei: Cichlidae). Arch. Oral. Biol. 40:1005–1014.

Huysseune A, Aerts P, Verraes W (1989) Pharyngeal tooth replacement in *Astatotilapia elegans*. Progress in Zoology 35: 480–481.

Hwang PM, Sykes BD (2015) Targeting the sarcomere to correct muscle function. Nature Reviews Drug Discovery 14: 313-328.

Jackson HE, Ono Y, Wang X, Elworthy S, Cunliffe VT, Ingham PW (2015) The role of Sox6 in zebrafish muscle fiber type specification. Skeletal Muscle 5: 2.

Jorgensen AO, Jones LR (1986) Localization of phospholamban in slow but not fast canine skeletal muscle fibers. An immunocytochemical and biochemical study. Journal of Biological Chemistry 261: 3775-3781.

Keßler M, Berger IM, Just S, Rottbauer W (2015) Loss of dihydrolipoyl succinyltransferase (DLST) leads to reduced resting heart rate in the zebrafish. Basic Research in Cardiology 110: 468.

Keller P, Vollaard NBJ, Gustafsson T, Gallagher IJ, Sundberg CJ, Rankinen T, Britton SL, Bouchard C, Koch LG, Timmons JA (2011) A transcriptional map of the impact of endurance exercise training on skeletal muscle phenotype. Journal of Applied Physiology 110: 46-59.

Kemp TJ, Sadusky TJ, Simon M, Brown R, Eastwood M, Sassoon DA, Coulton GR (2001) Identification of a novel stretch-responsive skeletal muscle gene (smpx). Genomics 72: 260-271.

Kim SH, Scott SA, Bennett MJ, Carson RP, Fessel J, Brown HA, Ess KC (2013) Multi-organ abnormalities and mTORC1 activation in zebrafish model of multiple Acyl-CoA dehydrogenase deficiency. PLoS Genetics 9: e1003563.

Kinoshita S, Katsumi E, Yamamoto H, Takeuchi K, Watabe S. Molecular and functional analyses of aspolin, a fish-specific protein extremely rich in aspartic acid. Marine Biotechnology 13: 517-526.

Kirchmaier BC, Poon KL, Schwerte T, Huisken J, Winkler C, Jungblut B, Stainier DY, Brand T (2012) The Popeye domain containing 2 (popdc2) gene in zebrafish is required for heart and skeletal muscle development. Developmental Biology 363: 438-450.

Klugbauer N, Dai S, Specht V, Lacinová L (2000) A family of γ-like calcium channel subunits. FEBS Letters 470: 189-197.

Kojic S, Medeot E, Guccione E, Krmac H, Zara I, Martinelli V, Valle G, Faulkner G (2004) The Ankrd2 protein, a link between the sarcomere and the nucleus in skeletal muscle. Journal of Molecular Biology 339: 313-325.

Kordeli E, Ludosky MA, Deprette C, Frappier T, Cartaud J (1998) AnkyrinG is associated with the postsynaptic membrane and the sarcoplasmic reticulum in the skeletal muscle fiber. Journal of Cell Science 111: 2197-2207.

Koshimizu E, Imamura S, Qi J, Toure J, Valdez DM, Carr CE, Hanai J Kishi S (2011) Embryonic senescence and laminopathies in a progeroid zebrafish model. PLoS One 6: e17688.

Kostek MC, Chen Y-W, Cuthbertson DJ, Shi R, Fedele MJ, Esser KA, Rennie MJ (2007) Gene expression responses over 24 h to lengthening and shortening contractions in human muscle: major changes in CSRP3, MUSTN1, SIX1, and FBXO32. Physiological Genomics 31: 42-52.

Lange M, Kaynak B, Forster UB, Tönjes M, Fischer JJ, Grimm C, Schlesinger J et al. (2008) Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex. Genes and Development 22: 2370-2384.

Leichty AR, Pfennig DW, Jones CD, Pfennig KS (2012) Relaxed genetic constraint is ancestral to the evolution of phenotypic plasticity. Integrative and Comparative Biology 52:16–30.

Li M, Andersson-Lendahl M, Sejersen T, Arner A (2016) Knockdown of fast skeletal myosin-binding protein C in zebrafish results in a severe skeletal myopathy. Journal of General Physiology 147: 309-22.

Li J, Li C, Zhang D, Shi D, Qi M, Feng J, Yuan T et al. (2014) SNX13 reduction mediates heart failure through degradative sorting of apoptosis repressor with caspase recruitment domain. Nature Communications 5: 5177.

Lieber R, Lieber R (2002) Skeletal muscle structure, function, and plasticity: the physiological basis of rehabilitation. Lippincott Williams & Wilkins. 448pp.

Liem KF (1973) Evolutionary strategies and morphological innovations: cichlid pharyngeal jaws. Syst Zool 22: 425–441.

Little AG, Kocka KM, Lougheed SC, Moyes CD (2010) Evolution of the nuclear-encoded cytochrome oxidase subunits in vertebrates. Physiological Genomics 42: 76-84.

Liu LA, Engvall E (1999) Sarcoglycan isoforms in skeletal muscle. Journal of Biological Chemistry 274: 38171-38176.

Love M, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15: 550.

Lovering RM, O'Neill A, Muriel JM, Prosser BL, Strong J, Bloch RJ (2011) Physiology, structure, and susceptibility to injury of skeletal muscle in mice

lacking keratin 19-based and desmin-based intermediate filaments. American Journal of Physiology 300: C803-C813.

Lu JR, Bassel-Duby R, Hawkins A, Chang P, Valdez R, Wu H, Gan L, Shelton JM, Richardson JA, Olson EN 2002) Control of facial muscle development by MyoR and capsulin. Science 298: 2378-2381.

Lu J, Webb R, Richardson JA, Olson EN (1999) MyoR: A muscle-restricted basic helix–loop–helix transcription factor that antagonizes the actions of MyoD. Proceedings of the National Academy of Sciences. 96:552-557.

Macqueen DJ, Fuentes EN, Valdés JA, Molina A, Martin SA (2014) The vertebrate muscle-specific RING finger protein family includes MuRF4 - A novel, conserved E3-ubiquitin ligase. FEBS Letters 588: 4390-4397.

Mammen AL, Mahoney JA, St Germain A, Badders N, Taylor JP, Rosen A, Spinette S (2011) A Novel Conserved Isoform of the Ubiquitin Ligase UFD2a/UBE4B Is Expressed Exclusively in Mature Striated Muscle Cells. PLoS One 6: e28861.

Mao L, Shelden EA (2006) Developmentally regulated gene expression of the small heat shock protein Hsp27 in zebrafish embryos. Gene Expression Patterns 6: 127-133.

Marino JS, Tausch BJ, Dearth CL, Manacci MV, McLoughlin TJ, Rakyta SJ, Linsenmayer MP, Pizza FX (2008) β_2-Integrins contribute to skeletal muscle hypertrophy in mice. Cell Physiology 295: C1026-C1036.

Mayer U (2003) Integrins: redundant or important players in skeletal muscle? Journal of Biological Chemistry 278: 14587-14590.

McDonald JH, Kreitman M (1991) Adaptive protein evolution at the Adh locus in Drosophila. Nature 351: 652-654.

McGivney BA, Eivers SS, MacHugh DE, MacLeod JN, O'Gorman GM, Park SDE, Katz LM, Hill EW (2009) Transcriptional adaptations following exercise in Thoroughbred horse skeletal muscle highlights molecular mechanisms that lead to muscle hypertrophy. BMC Genomics 10:638.

Meder B, Just S, Vogel B, Rudloff J, Gärtner L, Dahme T, Huttner I et al. (2010) JunB-CBFβ signaling is essential to maintain sarcomeric Z-disc structure and when defective leads to heart failure. Journal of Cell Science 123: 2613-2620.

Messer PW, Petrov DA (2013) Frequent adaptation and the McDonald-Kreitman test. Proceedings of the National Academy of Sciences 110: 8615-8620.

Meyer A (1987) Phenotypic plasticity and heterochrony in *Cichlasoma managuense* (Pisces, Chichlidae) and their implications for speciation in cichlid fishes. Evolution 41: 1357–1369.

Michel RN, Dunn SE, Chin ER (2004) Calcineurin and skeletal muscle growth. Proceedings of the Nutrition Society. 63:341-349.

Middelbos IS, Vester BM, Karr-Lilienthal LK, Schook LB, Swanson KS (2009) Age and diet affect gene expression profile in canine skeletal muscle. PLoS One 4: e4481.

Minchin JE, Williams VC, Hinits Y, Low S, Tandon P, Fan CM, Rawls JF, Hughes SM (2013) Oesophageal and sternohyal muscle fibres are novel Pax3-dependent migratory somite derivatives essential for ingestion. Development 140: 2972-2984.

Mitchell P, Steenstrup T, Hannon K (1999) Expression of fibroblast growth factor family during postnatal skeletal muscle hypertrophy. Journal of Applied Physiology 86: 313-319.

Murgiano L, D'Alessandro A, Egidi MG, Crisà A, Prosperini G, Timperio M, Valentini A, Zolla L (2010) Proteomics and transcriptomics investigation on longissimus muscles in large white and casertana pig breeds. Journal of Proteome Research 9: 6450-6466.

Muschick M, Barluenga M, Salzburger W, Meyer A (2011) Adaptive phenotypic plasticity in the Midas cichlid fish pharyngeal jaw and its relevance in adaptive radiation. BMC Evolutionary Biology 11: 116.

Nadauld LD, Chidester S, Shelton DN, Rai K, Broadbent T, Sandoval IT, Peterson PW et al. (2006) Dual roles for adenomatous polyposis coli in regulating retinoic acid biosynthesis and Wnt during ocular development. Proceedings of the National Acadadmy of Sciences 103: 13409-13414.

Nawrotzki R, Loh NY, Ruegg MA, Davies KE, Blake DJ (1998) Characterisation of alpha-dystrobrevin in muscle. Journal of Cell Science 111: 2595-2605.

Newey SE, Howman EV, Ponting CP, Benson MA, Nawrotzki R, Loh NY, Davies KE, Blake DJ (2001) Syncoilin, a novel member of the intermediate filament superfamily that interacts with α-dystrobrevin in skeletal muscle. Journal of Biological Chemistry 276: 6645-6655.

Nix DA, Fradelizi J, Bockholt S, Menichi B, Louvard D, Friederich E, Beckerle MC (2001) Targeting of zyxin to sites of actin membrane interaction and to the nucleus. Journal of Biological Chemistry 276: 34759-34767.

Norman B, Sabina RL, Jansson E (2001) Regulation of skeletal muscle ATP catabolism by AMPD1 genotype during sprint exercise in asymptomatic subjects. Journal of Applied Physiology 91: 258-264.

Odermatt A, Becker S, Khanna VK, Kurzydlowski K, Leisner E, Pette D, MacLennan DH (1998) Sarcolipin regulates the activity of SERCA1, the fast-twitch skeletal muscle sarcoplasmic reticulum Ca2+-ATPase. Journal of Biological Chemistry 273: 12360-12369.

Oshlack A, Wakefield MJ (2009) Transcript bias in RNA-seq data confounds systems biology. Biology Direct 4:14.

Otten C, van der Ven PF, Lewrenz I, Paul S, Steinhagen A, Busch-Nentwich E, Eichhorst J et al. (2012) Xirp proteins mark injured skeletal muscle in zebrafish. PLoS One 7: e31041.

Pagès H, Aboyoun P, Gentleman R, DebRoy S (2016). Biostrings: String objects representing biological sequences, and matching algorithms. R package version 2.42.1.

Palstra AP, Rovira M, Rizo-Roca D, Torrella JR, Spaink HP, Planas JV (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.

Papa I, Astier C, Kwiatek O, Raynaud F, Bonnal C, Lebart MC, Roustan C, Benyamin Y (1999) Alpha actinin-CapZ, an anchoring complex for thin filaments in Z-line. Journal of Muscle Research & Cell Motility 20: 187-197.

Papke RL, Ono F, Stokes C, Urban J, Boyd RT (2012) The nicotinic acetylcholine receptors of zebrafish and an evaluation of pharmacological tools used for their study. Biochemical Pharmacology 84: 352-365.

Parker KC, Walsh RJ, Salajegheh M, Amato AA, Krastins B, Sarracino DA, Greenberg SA. (2009) Characterization of human skeletal muscle biopsy samples using shotgun proteomics. Journal of Proteome Research 8: 3265-3277.

Pavey SA, Collin H, Nosil P, Rogers SM (2010) The role of gene expression in ecological speciation. Ann. N.Y. Acad. Sci. 1206:110–129.

Perera S, Mankoo B, Gautel M (2012) Developmental regulation of MURF E3 ubiquitin ligases in skeletal muscle. Journal of Muscle Research and Cell Motility 33: 107-122.

Popovic M, Zaja R, Loncar J, Smital T (2010) A novel ABC transporter: the first insight into zebrafish (*Danio rerio*) ABCH1. Marine Environmental Research 69: S11-S13.

Popovic M, Zaja R, Smital T (2010) Organic anion transporting polypeptides (OATP) in zebrafish (*Danio rerio*): Phylogenetic analysis and tissue distribution. Comparative Biochemistry and Physiology Part A Molecular and Integrative Physiology 155: 327-335.

Qunitana AM, Picchione F, Klein Geltink RI, Taylor MR, Grosveld GC (2013) Zebrafish etv7 regulates red blood cell development through the cholesterol synthesis pathway. Disease Models & Mechanisms 7: 265-270.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Raeker MO, Su F, Geisler SB, Borisov AB, Kontrogianni-Konstantopoulos A, Lyons SE, Russell MW (2006) Obscurin is required for the lateral alignment of striated myofibrils in zebrafish. Developmental Dynamics 235: 2018-2029.

Ramesh T, Lyon AN, Pineda RH, Wang C, Janssen PM, Canan BD, Burghes AH, Beattie CE (2010) A genetic model of amyotrophic lateral sclerosis in zebrafish displays phenotypic hallmarks of motoneuron disease. Disease Models and Mechanisms 3: 652-662.

Ravenscroft G, Miyatake S, Lehtokari VL, Todd EJ, Vornanen P, Yau KS, Hayashi YK, et al. (2013) Mutations in KLHL40 Are a frequent cause of severe autosomal-recessive nemaline myopathy. American Journal Human Genetics 93: 6-18.

Raymond F, Métairon S, Kussmann M, Colomer J, Nascimento A, Mormeneo E, García-Martínez C, Gómez-Foix AM (2010) Comparative gene expression profiling between human cultured myotubes and skeletal muscle tissue. BMC Genomics 11: 125.

Rodriguez A, Hilvo M, Kytömäki L, Fleming RE, Britton RS, Bacon BR, Parkkila S (2007) Effects of iron loading on muscle: genome-wide mRNA expression profiling in the mouse. BMC Genomics 8: 379.

Rose AJ, Hargreaves M (2003) Exercise increases Ca2+-Calmodulin-Dependent Protein Kinase II activity in human skeletal muscle. Journal of Physiology 553: 303-309.

Sandri M, Lin J, Handschin C, Yang W, Arany ZP, Lecker SH, Goldberg AL, Spiegelman BM. (2006) PGC-1α protects skeletal muscle from atrophy by suppressing FoxO3 action and atrophy-specific gene transcription. Proceedings of the National Academy of Sciences. 103:16260-16265.

Schiaffino S, Reggiani C (2011) Fiber types in mammalian skeletal muscles. Physiological Reviews 91: 1447-1531.

Schilter KF, Reis LM, Sorokina EA, Semina EV (2015) Identification of an Alu-repeat-mediated deletion of OPTN upstream region in a patient with a complex ocular phenotype. Molecular Genetics Genomic Medicine 3: 490-499.

Schneider H, Fritzky L, Williams J, Heumann C, Yochum M, Pattar K, Noppert G, Mock V, Hawley E (2012) Cloning and expression of a zebrafish 5-HT(2C) receptor gene. Gene 502: 108-117.

Schurmans S, Carrió R, Behrends J, Pouillon V, Merino J, Clément S (1999) The mouse SHIP2 (Inppl1) gene: complementary DNA, genomic structure, promoter analysis, and gene expression in the embryo and adult mouse. Genomics 62: 260-271.

Senga Y, Nagamine T, Kameshita I, Sueyoshi N (2012) Knockdown of two splice variants of Ca(2+)/calmodulin-dependent protein kinase Iδ causes developmental abnormalities in zebrafish, *Danio rerio*. Archives of Biochemistry and Biophysics 517: 71-82.

Shen H, Illges H, Reuter A, Stürmer C (2002) Cloning, expression, and alternative splicing of neogenin1 in zebrafish. Mechanisms of Development 118: 219-223.

Smith-Unna R, Boursnell C, Patro R, Hibberd JM, Kelly S (2016) TransRate: reference-free quality assessment of de novo transcriptome assemblies. Genome Research 26: 1134-1144.

Snell-Rood EC, Cash A, Han MV, Kijimoto T, Andrews J, Moczek AP (2011) Developmental decoupling of alternative phenotypes: insights from the transcriptomes of horn-polyphenic beetles. Evolution 65: 231–245.

Snell-Rood, E. C., Van Dyken, J. D., Cruickshank, T., Wade, M. J., Moczek, A. P. 2010. Toward a population genetic framework of developmental evolution: the costs, limits, and consequences of phenotypic plasticity. Bioessays 32:71–81.

Stauffer JR, van Snick GE (2004) Phenotypic plasticity: its role in trophic radiation and explosive speciation in cichlids (Teleostei: Cichlidae). Animal Biology 54: 137–158.

Sztal T, Berger S, Currie PD, Hall TE (2011) Characterization of the laminin gene family and evolution in zebrafish. Dev Dyn 240(2): 422-431.

Takada F, Vander Woude DL, Tong H-Q, Thomson TG, Watkins SC, et al. (2001) Myozenin: an alpha-actinin- and gamma-filamin-binding protein of skeletal muscle Z lines. Proc Natl Acad Sci 98: 1595–1600.

Tao L, Valdimarsson G (2010) Tandem alternative splicing of zebrafish connexin45.6. Genomics 96: 112-118.

Tello JA, Wu S, Rivier JE, Sherwood NM (2008) Four functional GnRH receptors in zebrafish: Analysis of structure, signaling, synteny and phylogeny. Integrative and Comparative Biology 48: 570-587.

Teng Y, Xie X, Walker S, Saxena M, Kozlowski DJ, Mumm JS, Cowell JK (2011) Loss of zebrafish igi1b leads to hydrocephalus and sensitization to pentylenetetrazol induced seizure-like behavior. PLoS One 6: e24596.

Teoh PH, Shu-Chien AC, Chan WK (2010) Pbx1 is essential for growth of zebrafish swim bladder. Developmental Dynamics 239: 865-874.

Toutenhoofd SL, Foletti D, Wicki R, Rhyner JA, Garcia F, Tolon R, Strehler EE (1998) Characterization of the human CALM2 calmodulin gene and comparison of the transcriptional activity of CALM1, CALM2 and CALMS. Cell Calcium 23: 323-338.

Trapnell C, Roberts A, Goff L, Pertea G, Daehwan K, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature Protocols 7: 562-578.

Vaittinen S, Lukka R, Sahlgren C, Hurme T, Rantanen J, Lendahl U, Eriksson JE, Kalimo H (2001) The expression of intermediate filament protein nestin as related to vimentin and desmin in regenerating skeletal muscle. Journal of Neuropathy and Experimental Neurology 60: 588-597.

Vangheluwe P, Schuermans M, Zádor E, Waelkens E, Raeymaekers L, Wuytack F (2005) Sarcolipin and phospholamban mRNA and protein expression in cardiac and skeletal muscle of different species. Biochemical Journal 389: 151-159.

Vaz R (2017) *ankdd1b* expression pattern in early stages of zebrafish development. ZFIN Direct Data Submission (http://zfin.org)

Venturin M, Carra S, Gaudenzi G, Brunelli S, Gallo GR, Moncini S, Cotelli F, Riva P (2014) ADAP2 in heart development: a candidate gene for the occurrence of cardiovascular malformations in NF1 microdeletion syndrome. Journal of Medical Genetics 51: 436-443.

Wainwright PC, Smith WL, Price SA, Tang KL, Sparks JS, et al. (2012) The evolution of pharyngognathy: a phylogenetic and functional appraisal of the pharyngeal jaw key innovation in labroid fishes and beyond. Syst Biol 61: 1001–1027.

Wang H (2010) Characterization of zebrafish Esrom (Myc-binding protein 2) RCC1-like domain splice variants. Molecular and Cellular Biochemistry 339: 191-199.

Weiss A, Schiaffino S, Leinwand LA (1999) Comparative sequence analysis of the complete human sarcomeric myosin heavy chain family: implications for functional diversity. Journal of molecular biology. 290:61-75.

Wen H, Linhoff MW, McGinley MJ, Li GL, Corson GM, Mandel G, Brehm P. Distinct roles for two synaptotagmin isoforms in synchronous and asynchronous transmitter release at zebrafish neuromuscular junction. Proceedings of the National Academy of Science 107: 13906-13911.

West-Eberhard MJ (2003) Developmental plasticity and evolution. Oxford University Press, Oxford.

West-Eberhard MJ (2005) Developmental plasticity and the origin of species differences. Proceedings of the National Academy of Sciences 102 Suppl:6543–6549.

Wimberger P (1991) Plasticity of jaw and skull morphology in the neotropical cichlid *Geophagus brasiliensis* and *G. steindachneri*. Evolution 45: 1545–1563.

Wund MA (2012) Assessing the impacts of phenotypic plasticity on evolution. Integrative and Comparative Biology 52: 5–15.

Yabu T, Imamura S, Mizusawa N, Touhata K, Yamashita M (2012) Induction of autophagy by amino acid starvation in fish cells. Marine Biotechnology 14: 491-501.

Yeh LK, Liu CY, Kao WW, Huang CJ, Hu FR, Chien CL, Wang IJ (2010) Knockdown of zebrafish lumican gene (ZLUM) causes scleral thinning and increased size of scleral coats. Journal of Biological Chemistry 285: 28141-28155.

Zhou X, Tian F, Sandzén J, Cao R, Flaberg E, Szekely L, Cao Y, Ohlsson C, Bergo MO, Borén J, Akyürek LM (2007) Filamin B deficienct in mice results in skeletal malformations and impaired microvascular development. Proceedings of the National Academy of Science 104: 3919-3924.

Zizioli D, Forlanelli E, Guarienti M, Nicoli S, Fanzani A, Bresciani R, Borsani G et al. (2010) Characterization of the AP-1 mu1A and mu1B adaptins in zebrafish (*Danio rerio*). Developmental Dynamics 239: 2404-2412.

Appendix

TopHat alignment parameters

This example script was parameterized to allow reads from *V. maculicauda* to align to the *O. niloticus* referece, after consultation with Margaret Staton & Miriam Payá Milans in the Department of Entomology and Plant Pathology at the University of Tennessee – Knoxville.

Parameters:

read-mismatches [10]: maximum number of mismatches permitted in the final read alignment
read-gap-length [10]: total length of gaps permitted in the final read alignment
read-edit-dist [10]: total length of edit distance permitted in the final alignment (maximum number of editing operations)
splice-mismatches [2]: maximum number of mismatches that appear in the anchor region (region crossing a junction) of a spliced alignment
b2-D [20]: number of times seed extension is permitted to fail to find a better alignment after the initial match
b2-R [3]: maximum times re-seeding is permitted with repetitive reads (where reads are repetitive if # seed hits / # seeds that aligned at least once > 300)
b2-N [1]: number of mismatches allowed per seed
b2-L [20]: seed length (length of initial search string for read)
b2-I [S,1,0.5]: setting for “very sensitive” alignments (at the expense of computational time)

Script:

```
tophat \
-o 6402_tophat \
--transcriptome-index ../raw_files/gene_index \
--read-mismatches 10 \
--read-gap-length 10 \
--read-edit-dist 10 \
--splice-mismatches 2 \
--b2-D 20 --b2-R 3 --b2-N 1 --b2-L 20 --b2-i S,1,0.50 \
../raw_files/Orenil_genome \
../2_trimmomatic/6402_trimmed.fastq.gz
```

Script for automated calculation of k_A and k_A/k_S

```
library(Biostrings)
library(seqinr)

#read data --> rest sequences
refseq<-readBStringSet("Orenil1.1_rna4.fasta", format="fasta") #first of 4 files (split for parallel
 runs)
newseq<-readDNAStringSet("Trinity_6402.fasta", format="fasta")
protseq<-readAAStringSet("Orenil1.1_protein.faa", format="fasta")
featuretable<-read.csv("Orenil1.1_feature_table.csv", header=F)

testblast<-read.csv("6402_refRNAblast.csv", header=F)
colnames(testblast)<-c("query_id", "subject_id", "percent_id", "alignment_length", "mismatches",
 "gap_opens", "query_start", "query_end", "subject_start", "subject_end", "evalue", "bit_score")

seq.div1<-as.data.frame(matrix(data=NA,nrow=length(refseq),ncol=8))
colnames(seq.div1)<-
 c("mac_transcriptID","Orenil_refID","length_ref","length_align","align_pid","ka","ks","kaks")

for (i in 1:length(refseq)){

refseq1<-refseq[i]
name1<-strsplit(names(refseq1)," ")[[1]][1] #transcriptID
transcript.length<-as.vector(featuretable$V19[featuretable$V11==as.vector(name1)])

seq.div1[i,2]<-name1
seq.div1[i,3]<-transcript.length

genename<-as.vector(featuretable$V7[featuretable$V11==as.vector(name1)])
protname<-as.vector(featuretable$V13[featuretable$V11==as.vector(name1)])
protseq1<-protseq[grep(protname, names(protseq))]

if(protname!=""){

matches<-testblast[which(testblast$subject_id==as.vector(name1)),]
singlematch<-matches[which(matches$percent_id==max(matches$percent_id)),]
if(length(singlematch$query_id)!=1){
singlematch<-singlematch[1,]
}

newseqname<-as.vector(singlematch$query_id)

if(is.na(newseqname)!="TRUE"){

newseq1<-newseq[grepl(gsub("\\|","\\\\|",newseqname),names(newseq))]

seq.div1[i,1]<-newseqname

# pairwise alignment of transcripts, (check fwd and revcomplement), trim to align

newseq2<-DNAString(toString(newseq1),start=1)
newseqrev<-reverseComplement(newseq2)
```

```
align.rev<-pairwiseAlignment(refseq1, newseqrev, "local", patternQuality=PhredQuality(22L),
 subjectQuality=PhredQuality(22L))
align.fwd<-pairwiseAlignment(refseq1, newseq1, "local", patternQuality=PhredQuality(22L),
 subjectQuality=PhredQuality(22L))

align.prelim<-c(align.fwd,align.rev)[[which(c(length(aligned(align.fwd, degap=TRUE)[1][[1]]),
 length(aligned(align.rev, degap=TRUE)[1][[1]]))==max(length(aligned(align.fwd,
 degap=TRUE)[1][[1]]), length(aligned(align.rev, degap=TRUE)[1][[1]])))]]

seq.div1[i,4]<-length(aligned(align.prelim,degap=TRUE)[1][[1]]) seq.div1[i,5]<-pid(align.prelim)

refseq1.trim<-aligned(pattern(align.prelim))
newseq1.trim<-aligned(subject(align.prelim))

#sub 'N' for '-'
refseq1.trim<-c2s(replace(s2c(toString(refseq1.trim)),which(s2c(toString(refseq1.trim))=="-"),"N"))
newseq1.trim<-c2s(replace(s2c(toString(newseq1.trim)),which(s2c(toString(newseq1.trim))=="-
 "),"N"))

# determine frameshift for reference
reftrans0<-AAString(c2s(translate(s2c(toString(refseq1.trim)), sens="F", frame=0,
 ambiguous=TRUE, numcode=1)))
reftrans1<-AAString(c2s(translate(s2c(toString(refseq1.trim)), sens="F", frame=1,
 ambiguous=TRUE, numcode=1)))
reftrans2<-AAString(c2s(translate(s2c(toString(refseq1.trim)), sens="F", frame=2,
 ambiguous=TRUE, numcode=1)))
reftrans<-c(toString(reftrans0),toString(reftrans1),toString(reftrans2))

try0<-pairwiseAlignment(reftrans0, protseq1, "local", patternQuality=PhredQuality(22L),
 subjectQuality=PhredQuality(22L), gapOpening=0, gapExtension=1)
try1<-pairwiseAlignment(reftrans1, protseq1, "local", patternQuality=PhredQuality(22L),
 subjectQuality=PhredQuality(22L), gapOpening=0, gapExtension=1)
try2<-pairwiseAlignment(reftrans2, protseq1, "local", patternQuality=PhredQuality(22L),
 subjectQuality=PhredQuality(22L), gapOpening=0, gapExtension=1)

refmatch<-c(pid(try0),pid(try1),pid(try2))
refshift<-which(refmatch==max(refmatch)) # 1 = no clip, 2 = need to clip 1, 3 = need to clip 2

if(length(refshift)==1){
if(max(refmatch)>90){

# clip sequences (ref and new) to match frameshift
is.whole <- function(x)
 is.numeric(x) && floor(x)==x

refclip<-DNAString(toString(refseq1.trim), start=refshift)

if (is.whole(length(s2c(toString(refclip)))/3)==TRUE) {
 refclip<-refclip
} else {
 refclip<-DNAString(toString(refclip), start=1, nchar=length(s2c(toString(refclip)))-1)
 if (is.whole(length(s2c(toString(refclip)))/3)==TRUE) {
```

```
 refclip<-refclip
} else {
 refclip<-DNAString(toString(refclip), start=1, nchar=length(s2c(toString(refclip)))-1)
}
}


newclip<-DNAString(toString(newseq1.trim), start=refshift)
if (is.whole(length(s2c(toString(newclip)))/3)==TRUE) {
 newclip<-newclip
} else {
 newclip<-DNAString(toString(newclip), start=1, nchar=length(s2c(toString(newclip)))-1)
 if (is.whole(length(s2c(toString(newclip)))/3)==TRUE) {
 newclip<-newclip
} else {
 newclip<-DNAString(toString(newclip), start=1, nchar=length(s2c(toString(newclip)))-1)
}
}

# get sequence divergence
align.final<-as.alignment(nb=2, nam=c(names(refseq1[1]), names(newseq1[1])),
 seq=c(c2s(replace(s2c(toString(refclip)),which(s2c(toString(refclip))=="-"),"N")),
 c2s(replace(s2c(toString(newclip)),which(s2c(toString(newclip))=="-"),"N"))))

diverge<-kaks(align.final)
seq.div1[i,6]<-diverge$ka
seq.div1[i,7]<-diverge$ks
seq.div1[i,8]<-diverge$ka[1]/diverge$ks[1]
}
}
}
}
}
write.csv(seq.div1, file="kaks_calcs4.csv")
```

CHAPTER III
COMPARATIVE GENE EXPRESSION OF HEROINE CICHLIDS ASSOCIATED WITH DIET

A version of this chapter is being prepared for submission to *BMC Evolutionary Biology.* S.F. Clemmensen designed experiment and performed all analyses. Co-researchers C.D. Hulsey & S. Borstein provided resources and collaborated on experimental design.

Abstract

Gene regulation shapes morphology as much or more than mutations in the coding regions of genes, and changes to gene regulation such as gene expression likely shapes phenotypic evolution. Expression changes have been extensively documented in the bony elements of the pharyngeal jaw apparatus (PJA) of cichlid fishes, but the PJA muscular sling has not been as well examined. We reared 19 species of Central American cichlids in a common garden experiment and obtained transcriptomes from their muscular sling. We compared gene expression between durophagous and non-durophagous species in a phylogenetic framework. We find that some genes associate with the different diet groups, indicating that differences in gene expression between species groups is due in part to constitutive differences in gene expression, and this relationship can be detected across species that have independently evolved durophagy.

Introduction

The interaction between genes and development are essential building blocks for generating morphological traits. While changes to the protein-coding regions of genes are an important way mutations can generate phenotypic diversity, it is not the only mechanism. Changes in gene expression can functionally alter the effect of a gene without changing its coding sequence, and evolution of gene expression might drive phenotypic divergence among species (Carroll 2008). Species-specific differences in gene regulation, such as expression, have been observed over long evolutionary time scales (Brawand et al. 2011). However, the

extent to which observable differences in complex adult phenotypes of closely related species are caused by constitutive differences in gene expression, rather than due to plasticity in gene expression, is largely unexplored.

Cichlid fishes (Family Cichlidae) are group of labrid fishes with diverse craniofacial morphology associated with diverse trophic specializations (Hulsey and García de León 2005, Clabaut et al. 2007, López-Fernández et al. 2013). One of these specializations is the ability to consume hard-shelled prey, durophagy, and has evolved multiple times across the Heroines, a Central American clade (Hulsey 2008). Durophagous species have similar modifications to their pharyngeal jaw apparatus (PJA), a complex of bone, teeth, and muscle that allow the pharyngeal jaw to crush hard-shelled prey (Hulsey 2006, Wainwright et al. 2012).

Craniofacial development in teleost fishes is well characterized (Ahi 2016), making the PJA amenable to expression analysis of candidate genes (Schneider et al. 2014, Hulsey et al. 2016). Analysis of post-larval stages in six African cichlid species have demonstrated species- and trophic-specific patterns in gene expression, though this was considered a weaker effect than alternative splicing (Singh et al. 2017). Additionally, shifts in gene expression in skeletal PJA elements of an African cichlid have been detected in association with diet shifts associated with durophagy (Gunter et al. 2013, Gunter and Meyer 2014). However, studies of gene expression in the cichlid PJA have primarily been limited to skeletal elements, leaving the muscular sling largely unexamined (but see Schneider et al 2014). Elements of the PJA are phenotypically plastic in response to diet, altering the size and shape of both skeletal and muscular elements (Huysseune 1995, Muschick et al. 2011, Chapter 1). We have observed that muscle mass and gene expression in the muscular sling of the cichlid PJA is determined in part by a plastic response to diet (Chapter 1, 2).

Here we investigate if there is a relationship between durophagous and non-durophagous species in patterns of gene expression in the absence of experimental diet differences. We address this question by rearing durophagous

and non-durophagous Central American Heroine cichlids in a common garden experiment, and obtaining transcriptomes for the pharyngeal jaw muscular sling. If gene expression is completely plastic in response to diet, i.e. differences in gene expression of cichlid PJA muscles are determined only by environmental effects, we expect to find no association between durophagous and non-durophagous species when they have been reared on the same experimental diet. However, if there are species-specific differences in gene expression between durophagous and non-durophagous species, we expect to find similar expression phenotypes between these species groups in comparative analyses.

Materials and Methods

Rearing

This study was carried out in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The protocol was approved by the Institutional Animal Care and Use Committee of the University of Tennessee, Knoxville (Permit Number 1833-0412).

Heroine cichlid fish were purchased commercially and housed individually for the duration of the experiment. Individuals were fed flake food daily for six months. At the end of the six-month experimental period, we euthanized individuals with tricaine methanesulfonate (MS-222) and measured standard length (SL, measured from the tip of the maxilla to the base of the caudal fin). The *Levator Posterior/Levator Externus IV* pharyngeal muscle complex was immediately dissected from the left side of the fish and stored in RNAlater. Individuals were then preserved in 10% formalin, and transferred to 70% ethanol for long-term storage. Following fixation, the remaining LP/LEIV muscles on the right sides of the fish were dissected from the pharyngeal apparatus and weighed to the nearest 0.1mg with an electronic balance. To determine if the crushing treatments induced differences in LP/LEIV mass, the muscle masses in hard diet

and soft diet groups were compared using a linear model in R Rv3.3.2 with standard length as a covariate (R core team 2016).

RNA alignment, expression, and normalization

Total RNA was extracted from LP/LEIV muscles using a Qiagen RNeasy Fibrous Tissue Mini Kit (n = 24). Transcriptomes were generated at the Yale Center for Genome Analysis. Samples were bar-coded and multiplexed on a single lane run on an Illumina HiSeq 2000 (1 x 75bp). Quality filtering using trimmomatic v0.36 sequentially removed adapter sequences, trimmed reads if average quality across 4 bases fell below a threshold phred score of 15, and removed reads with fewer than 30 bases (Bolger et al. 2014).

Reads were aligned to a reference genome for *Oreochromis niloticus* using TopHat v2.1.1 (Trapnell et al. 2012), using parameters designed to permit cross-species alignments (see Chapter 2 Appendix). Reads were processed with the default htseq-count script in HTSeq v0.7.2 to calculate the number of reads that had a single unique alignment for each gene (Anders et al. 2015). Estimated counts from HTSeq were normalized using the number of aligned reads in each sample, as well as transcript length and GC content of reference *O. niloticus* transcripts calculated using bedtools. We used the R package *cqn* to normalize counts, which calculates a normalized reads per kilobase per million mapped reads (RPKM) value for each gene in each sample (Hansen et al. 2012). All statistical analyses in this study were performed with Rv3.3.2 (R core team 2016).

Analysis of gene expression without phylogenetic correction

We selected candidate genes subsets using differential expression analysis methods that do not correct for phylogeny. These methods are subject to bias due to species relatedness, but allowed us to identify groups of genes that can then be used in more statistically rigorous analyses. First, differential expression was calculated from the raw HTSeq output using default parameters with the

DESeq function in the DESeq2 package in R (Love et al. 2014). DESeq2 uses a negative binomial distribution to perform Wald significance tests to identify genes with significant differential expression between durophagous and non-durophagous individuals. These genes we classified as our “group 1” genes.

Additionally, we performed a partial redundancy analyses (RDA) on normalized gene expression that had been scaled and centered, in order to identify genes with the largest differences in expression between diet types. We used standard length (SL) as the conditioned variable and diet as the constrained variable, testing for the effect of diet after removing size effects. For species where we had biological replicates, we used the species average of the normalized RPKM estimate for each gene. The analysis was performed with the *rda* function in the vegan package (Oksanen et al. 2017). Analysis of variance was performed on the reduced model marginal means using *anova.cca* with 999 permutations. The 48 genes with the highest loading values (the top 1%) in the RDA ordination were classified as our “group 2” genes.

Phylogeny

We used a phylogeny of cichlids assembled by Borstein et al. (in prep), which used 16 mitochondrial loci and 13 nuclear protein coding genes for 1,015 cichlid taxa, constrained by previously well-supported groups and using fossil calibrations. Using the R packages *ape* and *phytools* (Paradis et al. 2004, Revell 2012), we dropped species that we lacked samples for, leaving us a phylogeny with 19 tips from a clade of 151 species (from MRCA of sampled species), representing the majority of Central American Heroine cichlids (Figures 14, 15).

Analysis of gene expression with phylogenetic correction

Analysis of gene regulation is difficult due to the high-dimensionality of the data. Traditional gene expression analysis involves pairwise comparisons and cannot robustly evaluate hypotheses for more than a few species, making comparative analysis across many species impossible. Methods that incorporate phylogenies

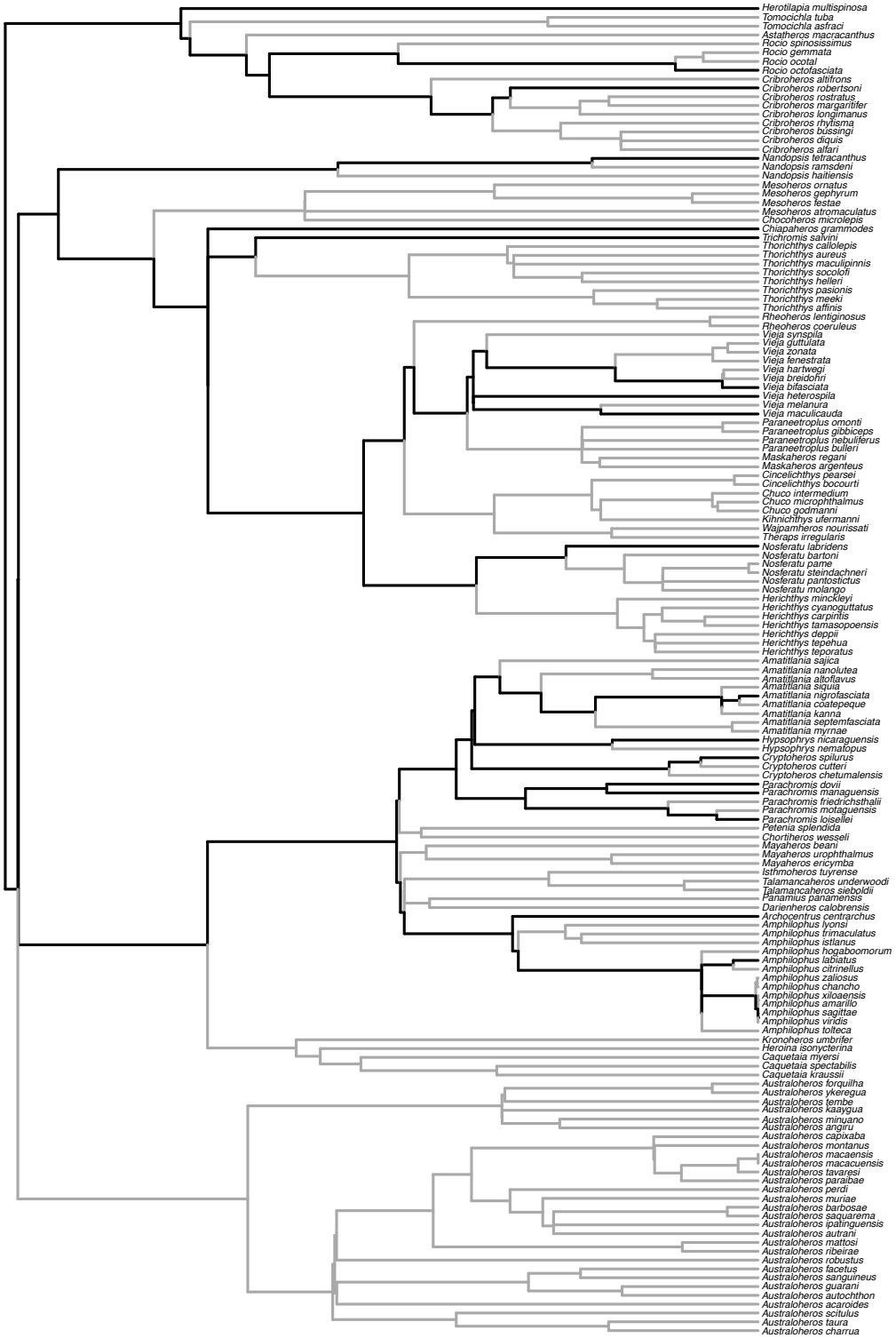

Figure 14. Heroine phylogeny.

All species descended from most recent common ancestor (MRCA) of experimental species (black).

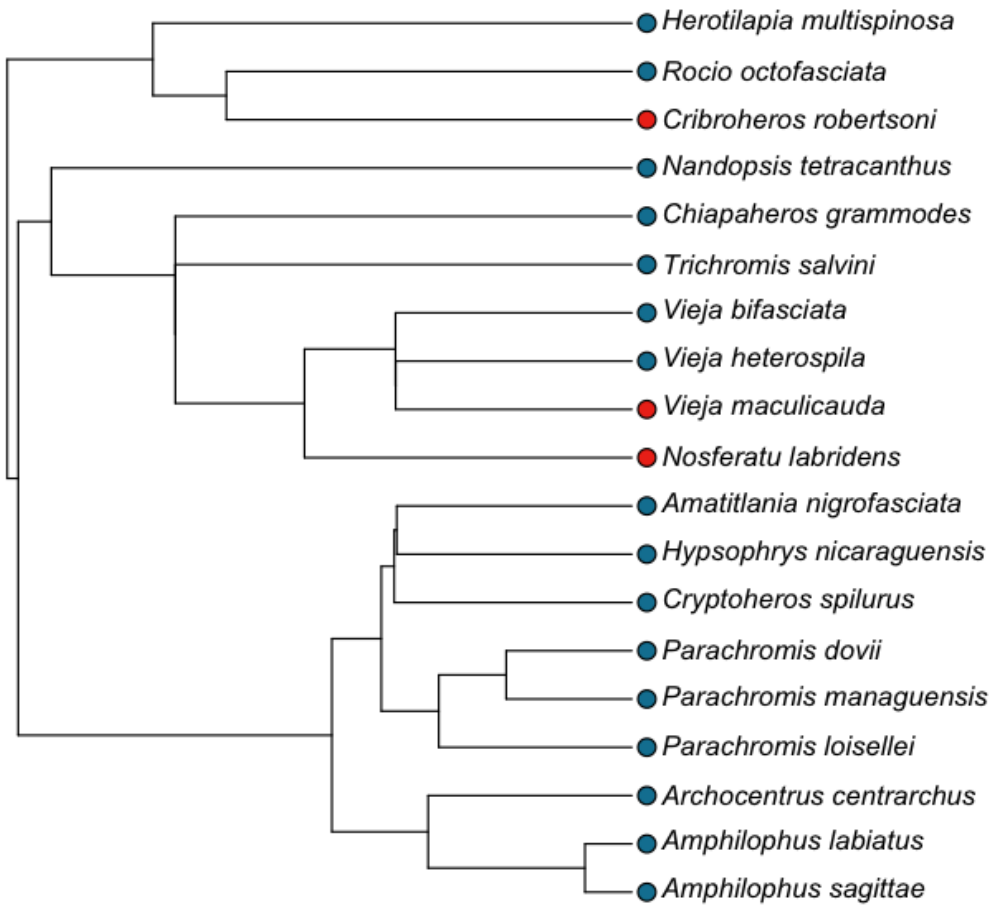

Figure 15. Phylogeny of experimental species.

Trimmed phylogeny, showing experimental species. Durophagous species are marked in red, non-durophagous species in blue.

have their own limitations, primarily revolving around reducing dimensionality of their trait data to satisfy the mathematical requirements of a phylogenetic ANOVA (Adams and Collyer 2017). However, methods introduced by Adams (2014) use distance-based matrices to perform phylogenetic least-squares analyses (D-PGLS) without the loss of power in traditional PGLS analyses associated with increased dimensionality of data. These methods were introduced to address high-dimensional morphology data but can be broadly applied to any multidimensional data, such as gene expression data.

Gene expression was analyzed using distance-based phylogenetic least-squares analysis with the R package geomorph v3.0.4 (Adams and Otarola-Castillo 2013, Adams 2014). For each test, we used the model: `procD.pgls(normalized_counts~SL+diet, tree, iter=999)`, where SL = log-transformed standard length. This analysis tested if there was a significant association between species diet (durophagous or non-durophagous) and gene expression in individuals with identical diets, i.e. testing gene expression in the absence of diet-induced plasticity. D-PGLS analyses uses species averages, so we used the mean normalized RPKM values for each gene for species where we had biological replicates. We analyzed 4,719 genes that we could match to well-annotated transcripts from chapter 2. We also tested 4 subsets of this data:

1. "muscle genes": 120 genes associated with muscles that were present in all species of our dataset (Chapter 2).
2. "diet genes": 10 genes that showed significant changes in expression in response to an experimental diet manipulation in *Vieja maculicauda* (Chapter 2).
3. "group 1": 8 genes that showed significant changes in expression between durophagous and non-durohagous species (from DESeq2 analysis with no phylogenetic correction).
4. "group 2": 48 genes that comprise the top 1% loading values from partial RDA ordination with no phylogenetic correction (3 of these genes overlap with the "group 1" genes).

We estimated the power for each of our D-PGLS tests. To do so, we simulated expression data in *OUwie.sim* for *n* genes (Beaulieu and O'Meara 2016), where *n* is the number of traits (genes) in the original analysis, θ_{DURO} and θ_{NON} for each gene equaled the group mean, and $\theta_0 = \theta_{NON}$. We used the *compare.evol.rates* function in geomorph to estimate overall σ^2 rates for the different diet types. We used these methods to simulate 1000 datasets for each of our four subgroups, and performed D-PGLS analysis on each simulated

dataset. The proportion of significant results was treated as an estimate of statistical power, similar to methods used in Adams (2014).

Results

Mass of the muscular sling

Mass of the LE/LPIV muscular sling was significantly associated with body length (SL), but did not vary between durophagous and non-durophagous species after 6 months of identical diet in the lab (R^2 = 0.824, Table 8). These results confirm previous research that the muscular sling has a plastic response to diet (Huysseune 1995, Muschick et al. 2011, Chapter 1), and indicate that any differences in gene expression we detect here are due to constitutive differences and not diet-induced plasticity in gene expression.

Table 8. ANOVA table for size of muscular sling.

Analysis of variance of LE/LPIV muscular sling size (log(mass)).

	SS	**Df**	**F-value**	**p-value**
Log(Standard Length)	7.326	1	79.86	< 0.001
Diet	0.2861	1	3.118	0.092
Interaction term	0.0406	1	0.443	0.513
Residuals	1.8348	20		

Alignment statistics

The Illumina run generated16.4 million single-end reads for each sample after quality filtering (sd = 5.1 million reads). For each sample, 57.7% of reads aligned to the *Oreochromis niloticus* reference (sd = 2.5%).

Analysis of gene expression without phylogenetic correction

Our non-phylogenetic methods identified candidate genes that may be associated with diet type. First, a Wald test with *DESeq* and found 8 genes with significantly different expression levels between durophagous and non-durophagous individuals (Table 9). These genes form the “group 1” genes.

Table 9. Group 1 genes.

Genes with significantly different expression between durophagous and non-durophagous species, calculated in DESeq. Differential expression (DiffExp) = $\log_2(\text{expression}_{DURO}/\text{expression}_{NON})$.

Gene ID	Gene name	Matching Tilapia ID	DiffExp
steap4	STEAP4 metalloreductase	XM_005472762.2	2.321
grina	protein lifeguard 1	XM_013272739.1	2.001
mta1	metastasis associated 1	XM_013270066.1	-0.848
fam151b	family 151 member B	XM_005461817.2	1.660
rdm1	RAD52 motif containing 1	XM_003442415.3	1.317
elp4	elongator acetyltransferase complex subunit 4	XM_003442311.3	-0.502
smarcb1b	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily B member 1	XM_003439639.2	1.268
tha1	probable low-specificity L-threonine aldolase 2	XM_003441873.3	1.757

In our second non-phylogenetic method, ANOVA testing on the redundancy analysis (RDA) model failed to find a significant effect of diet on gene expression (p = 0.337, Table 10). Within the RDA, size (SL) accounted for 3% of the variation in gene expression. After removing the contribution of the conditioning variable (SL), the diet RDA axis accounted for 2% of the model variation (Figure 16). The first principle components axis (PC1), or the variation not accounted for by the RDA axis, explained 60% of the model variation, and PC2 explained 10%. We selected genes with the top 1% of RDA scores (n = 48) as the genes with the largest effect on the RDA axis, i.e. the genes most strongly associated with diet; these genes form the “group 2” genes (Table 11). Normalized counts of the different groups are visualized in Figure 17.

Table 10. ANOVA of Redundancy analysis (RDA) of all genes.

Permutational test for partial RDA, marginal effects of terms with 999 permutations.

	Df	Variance	F	Pr(<F)
Diet	1	96.3	0.3431	0.91
Residual	16	4491.8		

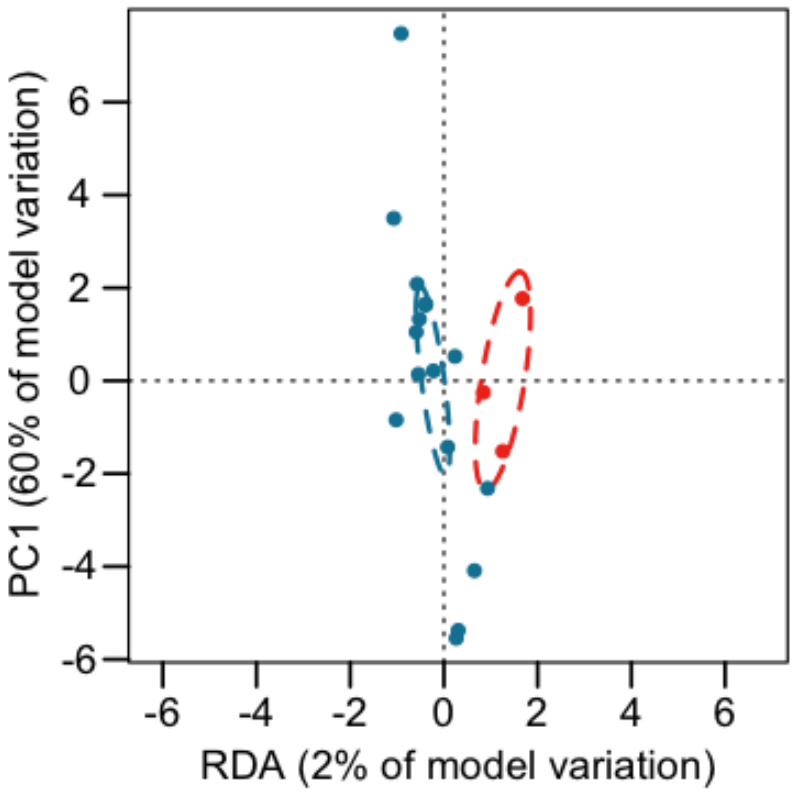

Figure 16. RDA results by individual.

RDA ordination results (scaling = 1) with durophagous individuals in red, non-durophagous individuals in blue. RDA axis indicates association with either durophagy or non-durophagy, conditioned on body size (SL). Dashed lines indicate 95% confidence ellipses around the centroid.

Analysis of gene expression with phylogenetic correction

Distance-based phylogenetic least squares (D-PGLS) analyses tested whether there was significant association between species diet and gene expression when diet is not varied within individuals (individual PGLS plotted in Figure 18). We failed to detect a significant association between diet and gene expression within our full dataset (Table 12). There was a significant association between the diet and “group 1” genes, the subset of genes with significantly different expression levels detected by a non-phylogenetically corrected Wald test (p = 0.042, R^2 = 0.129), with power = 0.164. We failed to detect an association

Table 11. Group 2 genes.

Genes with the largest effect on diet (upper 1% of calculated loading values) in a redundancy analysis. Asterisks(*) indicate genes that are also present in the “diet genes 1” group.

Gene ID	Gene name	Matching Tilapia ID	Loading
lipe	lipase, hormone-sensitive	XM_013266511.1	0.0972
fkbp5	peptidyl-prolyl cis-trans isomerase FKBP5	XM_003448347.3	0.0972
gne	glucosamine (UDP-N-acetyl)-2-epimerase/N-acetylmannosamine kinase,	XM_005454768.2	0.0972
pde8a	phosphodiesterase 8A	XM_003458202.3	0.0975
thbs4	thrombospondin 4	XM_003451759.3	0.0975
tubb	tubulin beta chain	XM_005458819.2	0.0975
agpat4	1-acylglycerol-3-phosphate O-acyltransferase 4	XM_005474042.2	0.0976
add3	gamma-adducin	XM_013271819.1	0.0977
cat	catalase	XM_005453108.2	0.0977
hsd17b12	hydroxysteroid (17-beta) dehydrogenase 12	XM_003455397.3	0.0976
srd5a2	steroid-5-alpha-reductase 2	XM_003441028.3	0.0990
fn1	fibronectin	XM_005450408.2	0.0990
rbm38	RNA binding motif protein 38	XM_003442657.3	0.0993
gata2	GATA-binding factor 2	XM_003442634.3	0.1003
*rdm1	RAD52 motif containing 1	XM_003442415.3	0.1004
cd302	CD302 molecule	XM_005452988.2	0.1007
msra	methionine sulfoxide reductase A	XM_005461509.2	0.1009
herc1	HECT and RLD domain containing E3 ubiquitin protein ligase family member 1	XM_013266333.1	0.1015
ca1	carbonic anhydrase 1	XM_003439275.3	0.1017
slc43a2	large neutral amino acids transporter small subunit 4	XM_005472225.1	0.1028
tkt	transketolase	XM_003444759.3	0.1035
rhcg	Rh family, C glycoprotein	XM_003440579.2	0.1039
*mta1	metastasis associated 1	XM_013270066.1	0.1046
tspan13	tetraspanin-13	XM_003451215.3	0.1053
selenbp1	selenium binding protein 1	XM_003457584.3	0.1055
cd59	CD59 glycoprotein	XM_003460266.3	0.1064
sostdc1	sclerostin domain containing 1	XM_003450293.3	0.1065
cebpd	CCAAT/enhancer binding protein (C/EBP), delta	XM_003438106.3	0.1080
rgl1	ral guanine nucleotide dissociation stimulator-like 1	XM_003443160.3	0.1082
ddr2	discoidin domain-containing receptor 2	XM_013276415.1	0.1090
myof	myoferlin	XM_013275556.1	0.1112
inip	SOSS complex subunit C	XM_005463559.2	0.1113
rhag	ammonium transporter Rh type A	XM_003446231.3	0.1114
creb1	cAMP responsive element binding protein 1	XM_003454909.3	0.1117
hipk2	homeodomain interacting protein kinase 2	XM_003447065.3	0.1136
mief1	mitochondrial dynamics protein MID51	XM_005460721.2	0.1140
prom1	prominin-1-A	XM_003441087.3	0.1141
znf143	zinc finger protein 143	XM_003453807.3	0.1166

Table 11 continued.

Gene ID	Gene name	Matching Tilapia ID	Loading
aqp1	aquaporin-1	XM_003438085.3	0.1171
atp1a1	sodium/potassium-transporting ATPase subunit alpha-1	XM_003446605.2	0.1181
add2	adducin 2 (beta)	XM_013273381.1	0.1217
tpr	protamine-like protein	XM_003446064.3	0.1221
metap2	methionine aminopeptidase 2	XM_005450940.2	0.1228
ass1	argininosuccinate synthase 1	XM_013268308.1	0.1248
lipt2	lipoyl(octanoyl) transferase 2	XM_013265808.1	0.1278
*steap4	STEAP family member 4	XM_005472762.2	0.1340
ca	carbonic anhydrase	XM_003456728.3	0.1361
hbab	hemoglobin subunit alpha-B	XM_005468759.2	0.1511

between diet and gene expression in our other subsets (Table 13) – genes associated with muscles (R^2 = 0.012, power = 0.09), genes with a plastic response to diet (R^2 = 0.019, power = 0.152), or the 1% of genes with the highest effects in a non-phylogenetically corrected redundancy analysis (R^2 = 0.084, power = 0.004).

Discussion

Constitutive differences in gene expression between species might drive phenotypic diversity among even closely related species. If this is the case, we expected to find differences between durophaogus and non-durophagous species in our “group 1” and “group 2” genes, but no difference among our “diet genes” group. Our results were consistent with this prediction - we found a significant difference between durophagous and non-durophagous species in the “group 1” and “group 2” genes, but not in any other gene sets or in the analysis of all 4,719 genes in our dataset. As expected, muscle size correlates with body size but not diet type. As demonstrated in Chapter 1, muscle size shows a plastic response to diet, and we would not expect to see a difference in muscle size of these fishes after 6 months of identical diet in the lab. From this we assume that observed differences in gene expression observed between durophagous and non-durophagous species in the “group 1” and “group 2” gene sets are candidate

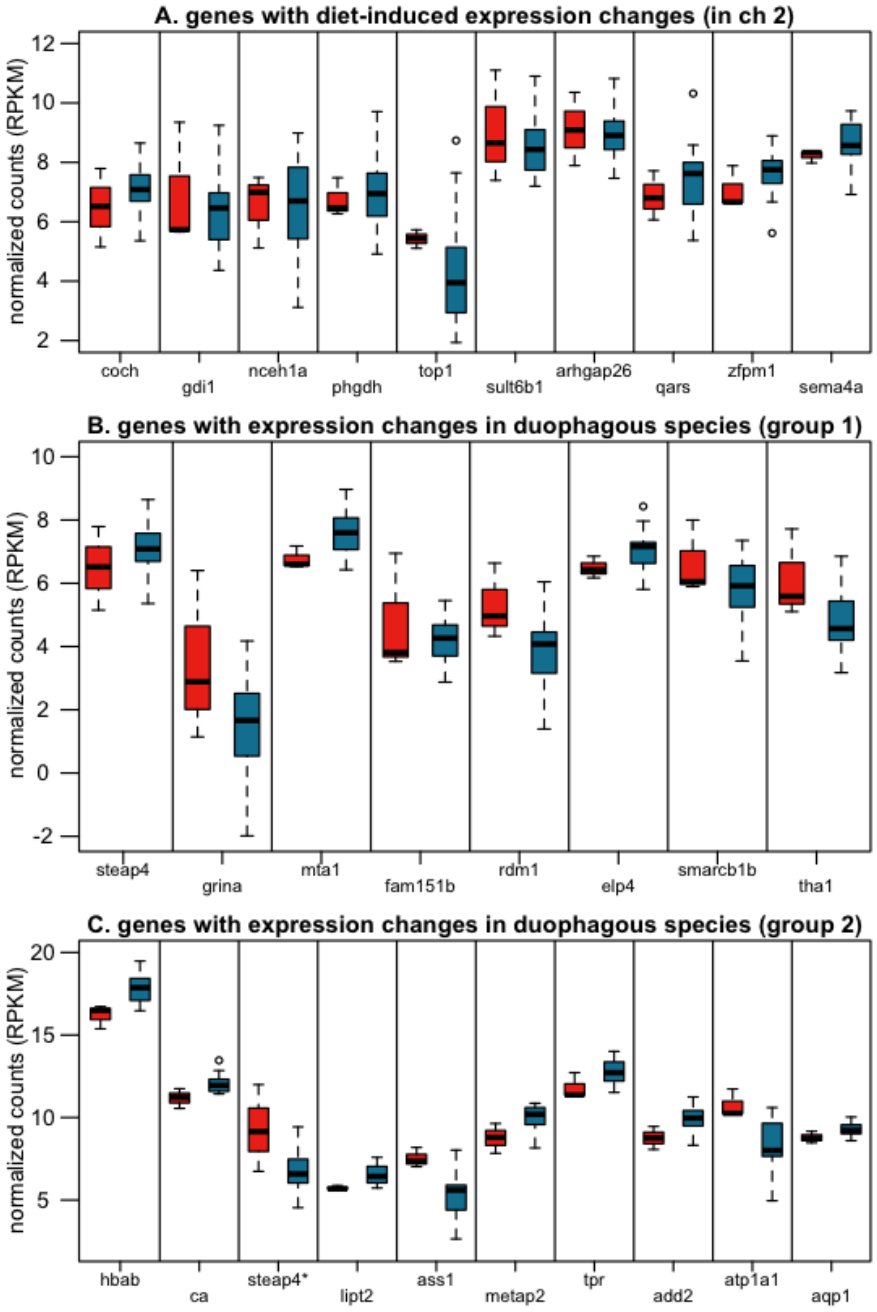

Figure 17. Count data for durophagous and non-durophagous species.

RPKM values (normalized in *cqn*) for different genes, with durophagous species indicated in red and non-durophagous species in blue. (A) shows genes that had plastic expression patterns in response to diet in *Vieja maculicauda* (Chapter 2). (B) shows genes with significant differential expression between durophagous and non-duropahgous species in DESeq2 (group 1 genes). (C) shows the genes with the top 10 loading values in a redundancy analysis (the first 10 of the group 2 genes).

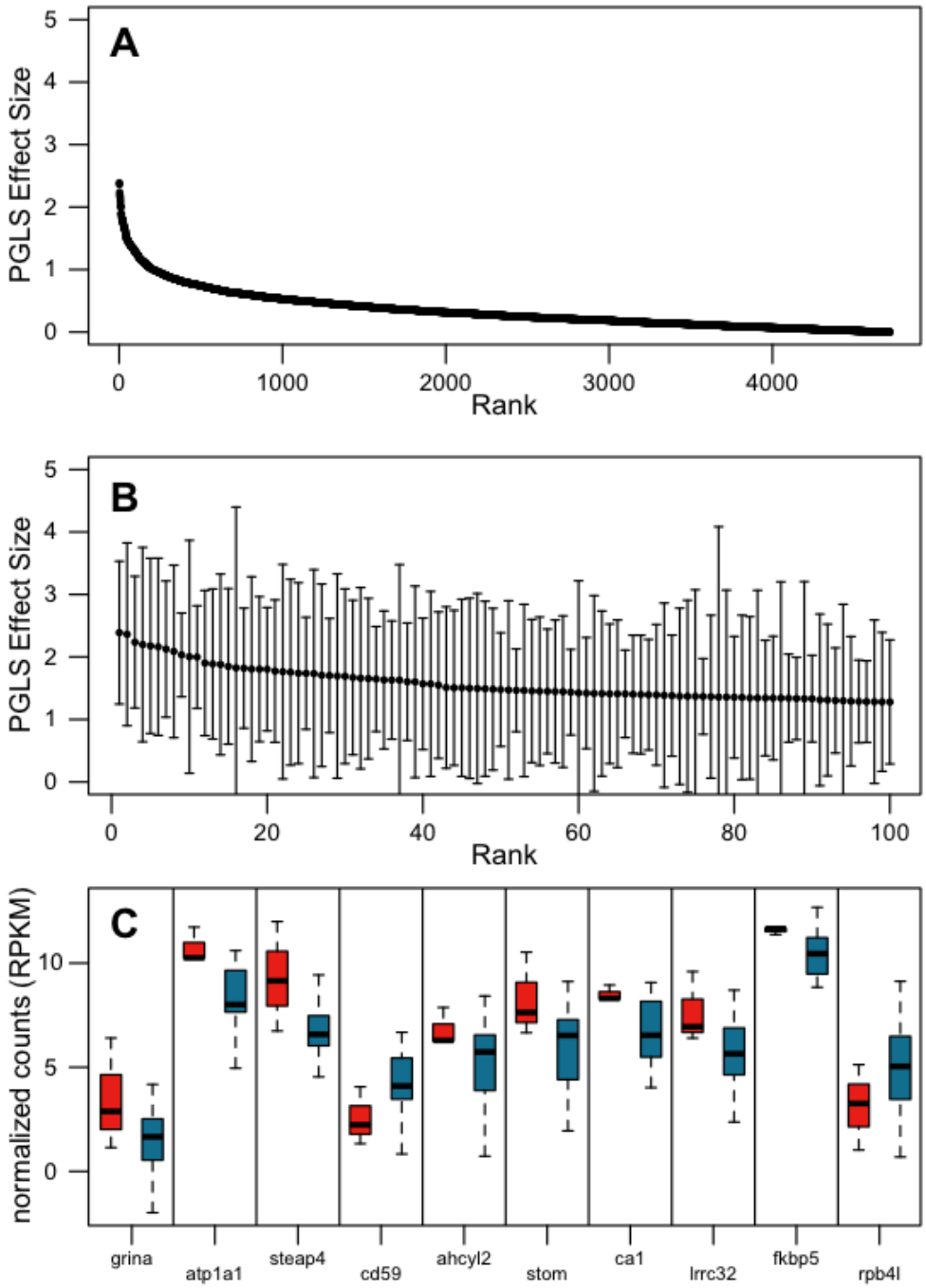

Figure 18. PGLS of individual genes.

PGLS analysis of each gene in our full data set, ranked by effect size. (A) shows effect sizes of all genes, and (B) shows effect size and standard error of the 100 genes with largest effect sizes. (C) shows normalized count data for the 10 genes with largest effect sizes in individual PGLS analyses, with durophagous species in red and non-durophagous species in blue.

Table 12. Distance-based phylogenetic least squares (D-PGLS) analysis of all genes.

Sequential Sums of Squares and cross-products, using randomized residual permutation procedure (999 permutations).

	Df	SS	MS	R^2	F	Z	p-value
SL	1	635.2	635.3	0.058	1.015	-1.424	0.956
Diet	1	207.0	207.0	0.019	0.331	-1.197	0.896
Residuals	16	10014	625.9				
Total	18	10889					

genes for constitutive effects on diet-associated muscle morphology.

We do not have enough power to be certain that there was no difference between durophagous and non-durophagous species in the “muscle genes” group or the “diet genes” group (the genes observed to have plastic expression patterns in Chapter 2). Overall patterns observed here might indicate that there are constitutive gene expression patterns associated with trophic specialization – in this case durophagy. Suggestive associations with diet in our analysis could indicate that gene effects may be common across durophagous species, despite durophagy being independently evolved in these lineages. This phenomenon would be consistent with the hypothesis that similar genes are co-opted to achieve similar results. We infer from these analyses that phenotypic plasticity is likely not the only determinant of gene expression differences in the muscular sling of the pharyngeal jaw apparatus (PJA) between species with different trophic specializations.

Evolution of *cis*-regulatory elements can change both gene expression and alternative splicing, functionally altering the effect of a gene without changing its coding sequence (Carroll 2008). Evolution of gene expression and alternative splicing has been observed across many vertebrate lineages, though alternative splicing has been predicted to have a stronger effect on phenotype evolution than changes in gene expression (Brawand et al. 2011, Barbosa-Morais et al. 2012, Merkin et al. 2012, Necsulea and Kaessmann 2015). Given results from studies of bony elements of the PJA (Singh et al. 2016), we can expect that other

Table 13. D-PGLS analysis of all gene groups.

Sequential Sums of Squares and cross-products, using randomized residual permutation procedure (999 permutations).

Muscle Genes							
	Df	**SS**	**MS**	**R^2**	**F**	**Z**	**p-value**
SL	1	18.33	18.33	0.054	0.923	-1.588	0.962
Diet	1	4.13	4.129	0.012	0.208	-1.880	0.981
Residuals	16	317.6	19.85				
Total	18	341.6					
Diet Genes							
	Df	**SS**	**MS**	**R^2**	**F**	**Z**	**p-value**
SL	1	2.098	2.098	0.050	1.180	-0.989	0.837
Diet	1	0.787	0.787	0.019	0.443	-0.654	0.749
Residuals	16	28.45	1.778				
Total	18	42.18					
Group 1 gene (Wald test)							
	Df	**SS**	**MS**	**R^2**	**F**	**Z**	**p-value**
SL	1	0.757	0.757	0.027	0.624	-1.416	0.926
Diet	1	3.592	3.592	0.129	2.960	1.703	0.042*
Residuals	16	19.42	1.214				
Total	18	27.90					
Group 2 genes (Redundancy Analysis)							
	Df	**SS**	**MS**	**R^2**	**F**	**Z**	**p-value**
SL	1	11.18	11.18	0.088	1.785	-0.544	0.686
Diet	1	15.65	15.65	0.123	2.498	1.978	0.023*
Residuals	16	100.2	6.264				
Total	18	127.3					

cis-regulatory elements, such as alternative splicing, are likely to have stronger associations with diet than gene expression alone. The nature of our sequence data was incompatible with accurate analysis of alternative splicing because it was single-end and mapped to a relatively distant reference so we cannot address alternative splicing in this study (Trapnell et al. 2012). However, previous research has indicated that the regulation of the muscular and bony elements of the PJA are linked within the African cichlid *Astatoreochromis alluaudi* (Schnieder et al. 2014), and taken together with this work, it is likely that the difference in *cis*-regulation of the PJA is altered in similar ways across independently-durophagous cichlid lineages. With these candidate genes identified, more in-

depth analysis of their splicing patterns and regulatory elements would confirm this prediction.

Previous research has focused on the bony elements of the PJA, but the muscular sling is an integral component of the PJA and deserves close examination. We have identified candidate genes that may be associated with muscle differences in durophagous and non-durophagous species. We also find here that gene expression in the muscular sling can be putatively associated with durophagy in the absence of diet-induced plasticity, indicating that there may be constitutive differences in gene expression associated with diet type. That these differences in can be detected in independently evolved trophic specializations provides additional evidence that regulatory elements shape phenotypic evolution.

References

Adams DC (2014) A method for assessing phylogenetic least squares models for shape and other high-dimensional multivariate data. Evolution 68: 2675-2688.

Adams DC, Otarloa-Castillo E (2013) geomorph: an R package for the collection and analysis of geometric morphometric shape data. Methods in Ecology and Evolution 4: 393-399.

Adams DC, Collyer ML (2017) Multivariate phylogenetic comparative methods: evalutations, comparisons, and recommendations. Systematic Biology 0: 1-18.

Ahi EP (2016) Signalling pathways in trophic skeletal development and morphogenesis: insights from studies on Teleost fish. Developmental Biology 420: 11-31.

Anders S, Pyl PT, Huber W (2015) HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics 31: 166-169.

Barbosa-Morais NL, Irimia M, Pan Q, Xiong HY, Gueroussov S, Lee LJ, Slobodeniuc V, Kutter C, Watt S et al. (2012) The evolutionary landscape of alternative splicing in vertebrate species. Science 338: 1587-1593.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30: 2114-2120.

Carroll SB (2008) Evo-Devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell 134: 25-36.

Clabaut C, Bunje PME, Salzburger W, Meyer A (2007) Geometric morphometric analyses provide evidence for the adaptive character of the Tanganyikan cichlid fish radiations. Evolution 61: 560-578.

Gunter HM, Fan S, Xiong F, Franchini P, Fruciano C, Meyer A (2013) Shaping development through mechanical strain: the transcriptional basis of diet-induced phenotypic plasticity in a cichlid fish. Molecular Ecology 22: 4516-4531.

Gunter HM, Meyer A (2014) Molecular investigation of mechanical strain-induced phenotypic plasticity in the ecologically important pharyngeal jaws of cichlid fish. Journal of Applied Ichthyology 30: 630-635.

Hansen KD, Irizarry RA, Wu, Z (2012) Removing technical variability in RNA-seq data using conditional quantile normalization. Biostatistics 13: 204-216.

Hulsey CD (2006) Function of a key morphological innovation: fusion of the cichlid pharyngeal jaw. Proceedings of the Royal Society B. 273: 669-675.

Hulsey CD, Fraser GJ, Meyer A (2016) Biting into the genome to phenome map: developmental genetic modularity of cichlid fish dentitions. Integrative and Comparative Biology 56: 373-388.

Hulsey CD, García de León F (2005) Cichlid jaw mechanics: linking morphology to feeding specialization. Functional Ecology 19: 487-494.

Hulsey, CD, Roberts RJ, Lin ASP, Guldberg R, Streelman JT (2008) Convergence in a mechanically complex phenotype: detecting structural adaptations for crushing in cichlid fish. Evolution 62: 1587-1599.

Huysseune A, Aerts P, Verraes W (1989) Pharyngeal tooth replacement in *Astatotilapia elegans*. Progress in Zoology 35: 480–481.

López-Fernández H, Arbour JH, Winemiller KO, Honeycutt RL (2013) Testing for ancient adaptive radiations in neotropical cichlid fishes. Evolution 67: 1321-1337.

Love M, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15: 550.

Merkin J, Russell C, Chen P, Burge CB. (2012) Evolutionary dynamics of gene and isoform regularion in mammalian tissues. Science 338: 1593-1599.

Muschick M, Barluenga M, Salzburger W, Meyer A (2011) Adaptive phenotypic plasticity in the Midas cichlid fish pharyngeal jaw and its relevance in adaptive radiation. BMC Evolutionary Biology 11: 116.

Necsulea A, Kaessmann H (2015) Evolutionary dynamics of coding and non-coding transcriptomes. Nature Reviews Genestics 15: 734-748.

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H (2017) vegan: Community Ecology Package. R package version 2.4-3. https://CRAN.R-project.org/package=vegan

Paradis E, Claude J, Strimmer K (2004) APE: analyses of phylogenetics and evoltion in R language. Bioinformatics 20: 289-290.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Revell LJ (2012) phytools: an R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution 3: 217-223.

Schneider RF, Li Y, Meyer A, Gunter HM (2014) Regulatory gene networks that shape the development of adaptive phenotypic plasticity in a cichlid fish. Molecular Ecology 23: 4511-4526.

Singh P, Börger C, More H, Sturmbauer C (2017) The role of alternative splicing and differential gene expression in cichlid adaptive radiation. Genome Biology and Evolution accepted manuscript evx204, http://doi.org/10.1093/gbe/evx204

Trapnell C, Roberts A, Goff L, Pertea G, Daehwan K, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature Protocols 7: 562-578.

Wainwright PC, Smith WL, Price SA, Tang KL, Sparks JS, Ferry LA, Kuhn KL, Eytan RI, Near TJ (2012) The evolution of pharyngognathy: the phylogenetic and functional appraisal of the pharyngeal jaw key innovation in labroid fishes and beyond. Systematic Biology 61: 1001-1027.

CONCLUSION

The overarching question of this dissertation is to ask to what extent is individual morphology shaped by phenotypic plasticity, using the pharyngeal jaw apparatus in Heroine cichlids. In chapter 1 I found that *Vieja maculicauda*, thought to be consistently durophagous in the wild, demonstrated considerable morphological plasticity in response to mechanical strain. Additionally tooth growth between the diet treatments showed similar variation as the trophically polymorphic species *Herichthyes minckleyi*, successfully demonstrating that phenotypic plasticity in the pharyngeal jaw apparatus of cichlids can recapitulate morphological variation observed in the wild in the critically important prey-processing unit of the pharyngeal jaw apparatus.

In chapter 2, I found that gene expression plasticity in the muscular sling of the pharyngeal jaw apparatus may be associated with peptide divergence and relaxed purifying selection in some cases, but not in genes that are specifically associated with skeletal muscles. Nevertheless, I detected patterns in peptide divergence and selection intensity associated with gene expression plasticity, indicating that this plasticity has likely played a role in shaping individual morphology.

In chapter 3 I found that that gene expression in the muscular sling can be associated with species diet in the absence of diet-induced plasticity, indicating that there are constitutive differences in gene expression associated with diet type. That these differences in can be detected in independently evolved trophic specializations provides additional evidence that regulatory elements shape phenotypic evolution, and that plasticity alone does not drive differences among species in these phenotypes.

In this dissertation, I found that diet-induced phenotypic plasticity is able to generate morphologies observed in the wild. Additionally, I found plasticity in gene expression associated with a component of this complex phenotype can be associated with increased peptide divergence in some cases, indicating that

plasticity may promote phenotypic evolution. However, I also found that some patterns of gene expression are consistent in durophagous species in the absence of diet-induced plasticity, indicating that while plasticity in the expression of some genes may promote morphological diversity, that diversity is also supported by conserved patterns of expression in other genes.

VITA

Sharon Fern Clemmensen was born in Arlington,VA. She graduated from McGill University in 2007 with a Bachelor of Science, majoring in Biology and minoring in Education. Sharon spent a little over a year working as an intern with the U.S. Fish & Wildlife Service in Warm Springs, GA for the Student Conservation Association. Sharon graduated from the University of Florida in 2010 with a Master of Science in Entomology, where she was awarded the Pauline O. Lawrence Scholarship in Physiology/Biochemistry/Toxicology.

When Sharon started at UT, she was awarded the J. Wallace and Katie Dean Multi-Year Graduate Fellowship, an intra-university fellowship. After completing her comprehensive exams, Sharon was applied for and was awarded a Doctoral Dissertation Improvement Grant (DDIG) from the National Science Foundation, and subsequently won a departmental award for “Exceptional Performance as a Graduate Student”.

Sharon is currently Instructor-of-Record for EEB 240 – Human Anatomy. She plans to pursue teaching at the collegiate level.