

University of Cincinnati

Date: 6/25/2015

I, **Bethany A Stahl**, hereby submit this original work as part of the requirements for the degree of Doctor of Philosophy in Biological Sciences.

It is entitled:

Regressive evolution of pigmentation in the blind Mexican cavefish *Astyanax mexicanus*

Student's name: **Bethany A Stahl**

This work and its defense approved by:

Committee chair: Joshua Gross, Ph.D.



17352

Regressive evolution of pigmentation in the blind Mexican cavefish

Astyanax mexicanus

A dissertation submitted to the

Graduate School

of the University of Cincinnati

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in the Department of Biological Sciences

of the McMicken College of Arts and Sciences

by

Bethany A. Stahl

M.S., Biology, University of Northern Iowa, July 2010

Committee Chair: Dr. Joshua B. Gross

June 2015

ABSTRACT

The natural world reflects profound biodiversity in every corner of the globe. These phenotypes range from wing spot variation for camouflage in butterflies, to protective armored plates in marine stickleback fish, to extravagant mating displays in peacocks. For many phenotypes, the “selective” benefit is self-evident. However, some forms of phenotypic evolution are less obvious, such as regressed or “lost” characters, since it can be difficult to determine the association (if any) between a discarded trait and a selective advantage. Cave-dwelling animals such as the blind Mexican cavefish, *Astyanax mexicanus*, serve as excellent models to investigate regressive evolution. The surface form is extant, allowing for direct comparisons between river- and cave-dwelling conspecifics. Cavefish likely evolved from an “ancestral” surface-dwelling form, which invaded the caves of Northeastern Mexico. As a consequence of roughly 3 million years in constant darkness, these remarkable cavefish lost their coloration. Moreover, the recurrent loss of pigmentation in geographically isolated populations renders this system ideal for investigating the broader changes mediating regressive evolution in nature. To investigate these fundamental questions, we utilize an integrative approach to characterize genetic mechanisms contributing to coloration loss in nature.

For this, we aimed to describe multiple genetic components – simple traits, complex characters and global changes in gene expression – that may contribute to regressive pigmentation. Although some simple traits (albinism and *brown*) have been characterized, the roles of *cis*-regulatory mutations affecting these single locus traits have not been described. We investigated populations of cavefish that harbor “*brown*” yet depict an intact coding sequence. We discovered many sequence alterations present in the 5’ putative promoter region of the causative locus *Mc1r*, some of which co-localized to highly conserved non-coding elements that

may play a critical role in gene regulation. We next sought to characterize the complex trait of melanophore number variation. We scored pigment cell number variation in a large surface x Pachón cave F₂ pedigree, and association mapping identified multiple significant QTL that co-localize to 19 distinct regions of our linkage map. We then co-analyzed QTL based on positional information and available Gene Ontology, along with RNA-seq data, to identify prospective structural and expression alterations. This led to the identification of *Tyrp1b* and *Pmela* as genes that mediate complex pigmentation in *Astyanax*. Finally, we characterized extreme differences in gene expression that accompany colonization into the cave. We compared the transcriptomic profiles of two geographically distinct cavefish populations relative to surface fish. Strikingly, these studies revealed convergence of gene expression in many of the same loci, which may implicate certain genes as crucial for life in caves. Moreover, we performed a pathway analysis and discovered potential upstream regulators, including *Otx2* and *Mitf*, which appear to be similarly affected in these geographically distinct cavefish lineages. This dissertation suggests that regressive evolution of pigmentation results from diverse genetic mechanisms working in concert to give rise to reduced pigmentation phenotypes in cavefish. Broadly, this work provides insight to the mechanisms governing trait evolution in the wild, and further, characterizes genes that may have implications in human degenerative disorders.

ACKNOWLEDGEMENTS

As the time of my doctoral training is nearly complete, I would like to take a moment to express my sincerest gratitude and appreciation to the to the people that have significantly contributed to my success in completing this program, and further, those that have inspired me to grow and excel as a young scientist, mentor and as a person. I truly would not have been able to achieve this life goal without this network of supportive and inspiring folks.

Dr. Peter Berendzen, my master's advisor, and the rest of the UNI crew, for giving me my first opportunity for experience biological research. The hands on lab and fieldwork truly showed me first hand how exciting research could be. Moreover, the skills that I gained while training in your lab, clearly helped prepare me for my time as a doctoral student. Also to my best buddy Katie Berge, of whom we spent many long nights together in the lab that led us to be BFF's!

To the many members of the Gross lab, both past and present, for the laughs and fun times hanging out together. Of this group, I specifically want to thank Molly Perkins, Daniel Berning, Sam Onusko, Kate Curoe and Nuria Alonso for their dedicated assistance in the lab as they have contributed work that enhanced my overall dissertation project. I wish you all the best of luck in your future endeavors.

I particularly wish to thank my lab mates Brian Carlson and Mandy Powers, who have been there since the "beginning of time" in the Gross lab. Together we have managed to really get some amazing projects up and running in the lab, and have fun along the way as well. I

sincerely appreciate all the help that you have provided me during this time. You will both do fantastic with your forthcoming careers and I wish you the best.

To Drs. Stephanie Rollmann, Ken Petren, John Layne, and Heather Norton for being a fantastic and supportive committee – willing to go above and beyond average committee duties to enhance my graduate program during my time at UC. I have always appreciated your sincere interest in my work, and in helping me grow into a more well-rounded and confident scientist. I am also grateful to Drs. Mike Polak, Elke Buschbeck, Ilya Vilinsky and Joshua Benoit for going out of their way to provide me with opportunities to be involved in the department and community.

To the best group of graduate student friends that one would hope to encounter. These past five years have been a blast, from softball to our “old people” knitting night – having good friends when you are far away from family has been amazing. I am excited to see where everyone goes and we move onto the next stage of our careers. ...And I know these friendships will continue over time.

To my advisor, Dr. Joshua Gross, of whom I will be forever grateful for generously providing me this amazing opportunity to work in his lab. He is truly an outstanding advisor – he finds this magical way to both encourage us to work very hard, to help us maximize our “research” potential, and additionally, spends long hours helping us grow individually as young scientists. I have gained invaluable experience from Josh, as he is a superior model in real “life” lessons. I learned that it is possible to manage both a successful career yet still be very dedicated

to family – of which he is a pro. These past five years will continue to be indispensable as I hope to continue on this path of an academic researcher... and I hope that when I come back to visit, we can still give the ‘ol ball a toss.

To my Mom and Dad, and brother Todd, for their continued love and support as I seemingly try to carry on as a “perpetual” student. Although I can’t promise that I’ll have a quote “real job” before age 30, in the long run, this continued training will ultimately help me achieve my career goals. Thank you for selflessness in helping me pursue these opportunities, and for always being there to support in numerous ways. Moreover, I am truly fortunate to have an extended family full of loving in-laws, grandparents, aunts, uncles, cousins, and furry friends who have been there to support me in my throughout my life and career as well.

And finally, I extend my sincerest gratitude to the best husband and friend in the world, Aaron. Without his continued support and encouragement, I likely would not be in this position now. His kindness and selfless support of me as a person and a researcher is more than I could hope to ask for. As scientists, we can both be each other’s toughest critics and biggest cheerleaders, but this has dramatically impacted how we have grown together as scientists. It has truly been a remarkable experience for us to share in the pursuit of our dreams side-by-side... and I mean this literally in that we enrolled in nearly all courses together since freshman year of college. I am truly amazed at how far we have come in the past ten years and I can’t wait to continue to grow together in our lives and sharing our careers together.

TABLE OF CONTENTS

Introduction.....	1
References.....	9
CHAPTER 1	
Alterations in <i>Mclr</i> gene expression are associated with regressive pigmentation in <i>Astyanax</i> cavefish	15
Abstract.....	16
Introduction.....	17
Materials and Methods.....	20
Results and Discussion	23
References.....	36
Supplementary Tables.....	40
CHAPTER 2	
High-resolution genomic mapping reveals genes contributing to complex melanophore variation in <i>Astyanax mexicanus</i> cavefish	41
Abstract.....	42
Introduction.....	43
Materials and Methods.....	46
Results.....	57
Discussion	78
References.....	87
CHAPTER 3	
Developmental gene expression consequences of recurrent cave colonization	98
Abstract.....	99

Introduction.....	100
Materials and Methods.....	104
Results and Discussion	111
References.....	134
Supplementary Figures	145
Supplementary Tables.....	148
General Conclusions	155
References.....	159

INTRODUCTION

Overview

Regressive phenotypic evolution is a widespread phenomenon that affects a remarkably diverse number of organisms spanning most major phyla (Jeffery 2009). The most obvious consequences of life in perpetual darkness are the loss of pigmentation and reduction in eyes. The regression of morphological characters has long perplexed biologists, including classic evolutionists such as Charles Darwin, who attributed this loss of characters “to disuse” (Darwin 1859). The mechanisms driving regressive evolution remain largely unknown. Is the loss of pigmentation a result of selection for potential energy conservation in the cave environment? Or, are traits lost due to pleiotropic consequences of natural selection acting on an unrelated biological system (constructive traits)? Alternatively, does regressive evolution result from accumulation of neutral mutations (drift)?

Astyanax mexicanus: A model system for investigating regressive evolution

The blind Mexican cavefish *Astyanax mexicanus* is an excellent model for the investigation of novel traits because the ancestral, surface fish are extant, allowing direct comparisons with cave-dwelling morphs (Şadoğlu 1979; Borowsky 2008). The regression of pigmentation and eyes are representative of a suite of cave-associated (troglobitic) traits that consistently recur among broad taxa that have colonized the subterranean environment. In addition to “regressive” traits, cavefish exhibit constructive traits, including an enhanced lateral line system, increased numbers of taste buds, larger jaws, and keener auditory systems (Jeffery 2001). Finally, *Astyanax* is closely related to the teleost *Danio rerio* (zebrafish), which enables use of a number of similar molecular tools in *Astyanax*.

Freshwater cave forms of *Astyanax mexicanus* have been discovered in 29 separate cave localities in the Sierra de El Abra region of northeastern Mexico (Mitchell et al. 1977). Knowledge of the sequence of colonization, time since colonization, and relationships among various cave populations is critical for understanding the modes of trait acquisition through either convergent or parallel evolution (Gross 2012). Natural geographic barriers in northeastern Mexico imply independent colonization of, at least, several cave populations (Avisé and Selander 1972; Avisé et al. 1987; Espinasa and Borowsky 2001; Dowling et al. 2002; Gross 2012). The topography of this region includes several mountain ranges, including the Sierra de Colmena to the West, the Sierra de Guatemala to the North, and the Sierra de el Abra region to the East. With the knowledge of these distinct barriers, combined with morphological and population level analyses, the current consensus view suggests that cave forms arose from at least two separate ancestral surface (epigeal) stocks. The first “wave” of colonization seeded the El Abra cave system ~5 – 3 MYa (late Pliocene/early Pleistocene), and a second, “new” epigeal stock populated the Guatemala and Micos caves ~2 – 1 MYa (mid-Pleistocene; reviewed in Gross 2012). The widespread, and likely independent nature of these populations, render the blind Mexican cavefish an ideal system for the study of traits evolving in parallel (Jeffery 2001).

Pigmentation phenotypes in blind cavefish

One of the most conspicuous morphological changes associated with cave fauna is the loss of pigmentation (Wilkens 1988; Wilkens and Strecker 2003; Gross et al. 2009). Coloration in *Astyanax* is governed by several genes, however some aspects of skin color are inherited in a simple (monogenic) fashion (Şadoğlu 1957a; Borowsky and Wilkens 2002). To date, only two genes (*Oca2* and *Mc1r*) have been linked to specific pigmentation phenotypes. Complete

absence of pigment-producing melanophores (albinism) is governed by coding mutations in *Oca2* (Protas et al. 2005). The gene *Oca2* functions in melanogenesis as a L-tyrosine transporter into the melanosome (McCauley et al. 2004). The L-tyrosine is then converted into L-DOPA, and subsequently into L-DOPAquinone, which leads to a number of downstream reactions catalyzed by tyrosinase and related proteins, to produce melanin (McCauley et al. 2004). Lesions in *Oca2* impede L-tyrosine transport, and in turn, blocks overall eumelanin synthesis (Jeffery 2009).

A second regressive pigmentation phenotype, *brown*, leads to brownish eyes and reduced numbers of melanophores. Crosses between surface and cave morphotypes resulted in a near perfect 3:1 Mendelian ratio, suggesting that a single gene was responsible for the *brown* mutation (Şadoğlu and McKee 1969). Additionally, this phenotype failed to complement in crosses between different cave individuals, implying that *brown* is governed by the same locus across several populations (Wilkins and Strecker 2003). Consistent with these results, recent QTL analyses identified a single locus associated with the *brown* phenotype, later identified as the gene *Mc1r* (Gross et al. 2009). The Mc1r protein functions as a G-protein coupled receptor, which binds alpha-melanocyte-stimulating hormone (α MSH; Rees 2003). With ligand binding, adenylyl cyclase is activated, causing intracellular increases in cAMP (Widlund and Fisher 2003). Mc1r activation after binding stimulates downstream effectors, including the target gene *Mitf*, which regulates pigmentation (Gross et al. 2009; Widlund and Fisher 2003). Three of the seven caves manifesting the *brown* phenotype harbor coding sequence alterations in the gene *Mc1r*, however the other four populations in which the *brown* mutation occurs may be explained by *cis*-regulatory changes within the putative promoter region of *Mc1r* (Gross et al. 2009).

Though two genes affecting simple reduced pigment traits have been characterized, additional loci likely contribute to the broad range of *Astyanax* pigmentation phenotypes.

Melanophore number was previously scored across four body regions (head, dorsal, stripe, and anal) in a surface x Pachón cave F₂ pedigree (Protas et al. 2007). Subsequent analyses, based on a first-generation linkage map, identified 18 different QTL that affect melanophore-based pigmentation in *Astyanax mexicanus* (Protas et al. 2007). This suggests that some aspects of pigmentation are complex and are governed by several genes (polygenic), which have not yet been identified. To further investigate the genes underlying melanophore number, it will be critical to understand the pigment cells and prospective genes within the pigmentation pathway that may be targeted for mutation.

Pigment cells: A brief review of development and physiology

Astyanax surface fish harbor three types of pigmentation cells: yellow xanthophores, light-reflecting iridophores, and black/brown melanophores. An early comparison showed that xanthophores were present in similar numbers in surface versus Chica cavefish, with less known about the variation in iridophores (Wilkens 1988; McCauley et al. 2004; Jeffery 2009). Most knowledge among pigment cell types centers on melanophores, which are highly variable between surface and cave morphotypes.

All three pigmentation cell types arise from the neural crest, a set of migratory cells in vertebrates that give rise to numerous cell types including cranial cartilage and bone, peripheral neurons, fat cells, and pigment-producing melanophores (Erklickson and Perris 1993; Huang and Saint-Jeannet 2004). Due to the diversity of neural crest cell derivatives, it would be less likely to acquire mutations within genes of the neural crest pathway due to potential lethal consequences (Jeffery 2009). Labeling experiments revealed normal neural crest migration during cavefish development (McCauley et al. 2004). An alternative explanation for pigment cell regression is

cell apoptosis after neural crest-derived precursor migration. When this was evaluated, only a small number of dying neural crest cells were observed in cavefish, with comparable numbers in surface embryos (Jeffery 2006). These combined results suggest that evolutionary changes leading to pigment cell regression in cave morphs may be mediated by alterations late in melanogenesis (Jeffery 2009).

The physiology and regulation of pigmentation research is rather limited in fishes, in contrast to vast knowledge of known melanin functions in mammals (reviewed in Logan et al. 2006). The most notable role of melanin in teleost fish, more specifically in response to α MSH, is the ability to rapidly translocate melanin granules from the nucleus to the dendritic extensions of the pigment cells (Logan 2003a; Logan et al. 2003b; Metz et al. 2006). Rapid translocation of melanin allows the fish to quickly adapt to the local background environment (Logan et al. 2006). In addition to the benefits of camouflage, pigmentation functions in sexual preference and protects animals from damaging UV sunlight (Jeffery 2009). Melanin absorbs UV radiation and dissipates UV rays as heat (Riesz 2004). This same absorbing ability may also function to suppress sound with inner ear pigmentation (Robbins 1991) and protects the retina from overexposure by absorbing and scattering light (Sarna 1992). In general these melanin functions are beneficial to most organisms, therefore it is difficult to understand why pigment regresses in cave animals.

Current hypotheses on the evolutionary pressures facilitating regressive evolution

Why do cave-associated animals lose pigment in dark, subterranean environments? The cause of reduced pigmentation has long intrigued cave biologists. Darwin had difficulty reconciling the loss of traits evolving through selection. Currently, biologists recognize three

principal hypotheses to explain the regression of features in cave organisms. The first hypothesis suggests phenotypic loss evolves through direct selection, wherein regression confers some benefit for the animal. For instance, development of a feature likely imposes an energetic cost to the organism (Niven 2008). According to this hypothesis, pigment loss could be advantageous, such that pigmentation would be selected against in nutrient-poor environments. This energy conservation, evolving through selection, is frequently regarded as a primary mechanism leading to cave-associated features (Hüppop 1986; Jeffery 2005). Metabolic efficiency has also been observed in some cave-dwelling organisms, however this trait is not universal for troglobitic organisms, but evolves more frequently as a consequence of food scarcity (Gross 2012b).

A second hypothesis to explain the evolution of reduced phenotypes is indirect selection as a result of pleiotropy. In this case, regressive troglobitic features arise from association with an expanded trait under intense selection in cavefish. For example, the gene *Sonic hedgehog* (*Shh*) is known to play a role in both taste bud and eye development (Menuet et al. 2006; Rétaux and Pottin 2008; Yamamoto et al. 2004). Experiments in which *Shh* is overexpressed in the oral-pharyngeal region yielded both increased taste bud number and eye degeneration. These results indicate that selection for a constructive trait, such as increased taste bud number, in an environment with limited food resources may be responsible for reduced eyes through pleiotropic interactions with the *Shh* signaling pathway (Yamamoto et al. 2004). Recent evidence suggests a prospective benefit to loss of albinism through the *Oca2* pathway, wherein melanin precursors that are not converted to melanin in albino cavefish may instead be co-opted into the catecholamine pathway where they impact sleep loss. However, this more constant behavior of foraging for food could be of benefit to cavefish (Bilandžija et al. 2013).

The third hypothesis suggests reduced traits are caused by neutral mutation and genetic drift. This hypothesis proposes that genes associated with a character devoid of any benefit are free to accumulate mutations over time and drift in a population. Eventually, these genes can result in loss of function. Neutral mutation predicts that target genes are likely exclusive in function to decrease the possibility of harmful pleiotropic consequences (Wright 1964). In addition, neutral mutation/drift requires a significant time since isolation to occur, and populations experiencing longer periods since isolation would be predicted to demonstrate a higher degree of troglomorphy (Culver 1982; Wilkens 1988). Finally, neutral mutation assumes a certain level of randomness with respect to the genes that acquire mutations, yet some genes may be more susceptible to damage leading to recurrent phenotypic loss through the same gene (e.g., *Oca2*). With albinism, two independent cave populations acquired unique mutations in same locus *Oca2*, implying that some constraint mechanism is facilitating the complete loss of pigmentation in *Astyanax mexicanus* (Gross 2012b).

With these three competing hypotheses, how do we best determine the precise mechanism facilitating the evolution of cave-associated traits? Interconnection between the different evolutionary mechanisms makes it difficult to determine a single underlying cause. For example, genes of exclusive function are more likely targets to accumulate neutral mutations and drift in a population due to the reduced likelihood for harmful pleiotropic consequences. Yet, it is possible that some pleiotropic effects can be tolerated if the indirectly affected trait does not pose any reduced fitness outcomes in the cave environment (e.g., taste buds/eye degeneration; Gross 2012b). Thus, the evolution of a specific cave trait may need to be evaluated independently, and could be due to a variety of mechanisms (Poulson 1981; Gross 2012b). The investigation of these

diminished phenotypic traits will continue to clarify the mechanisms underlying regressive evolution.

Dissertation overview

This dissertation project adopts an integrative approach, incorporating multiple genetic and bioinformatic analyses, along with functional assays to characterize the genetic changes involved with cavefish evolution. This dissertation first identifies genetic changes in the 5' *Mc1r* putative promoter region leading to the *brown* phenotype in cave-dwelling fishes. Characterizing the nature of these sequence changes provides a clearer picture of how cave-dwelling animals evolve regressive features. Moreover, crosses between cave and surface forms of *Astyanax mexicanus* produce viable hybrid offspring and yield large F₂ pedigrees displaying a wide spectrum of phenotypic characters for quantitative trait locus (QTL) association mapping studies. The combination of QTL analyses, along with transcriptome profiling through RNA-seq, nominate several candidate genes involved with complex (polygenic) pigmentation traits. Finally, comparing the gene expression profiles of surface fish and two independent cavefish populations (Pachón and Tinaja) during early developmental stages clarify how distinct cavefish lineages have converged on depigmentation phenotypes, and other cave-associated traits, through both shared and different genetic mechanisms. The culmination of this work provides substantial insight into the genetic changes governing the evolution of regressive traits in natural populations, and identifies numerous genes that may be associated with degenerative disorders.

REFERENCES

- Avise, J. C., J. Arnold, R. M. Ball, Jr., E. Bermingham, T. Lamb, J. E. Neigel, C. A. Reed, and N. C. Saunders (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* 18: 489–522.
- Avise J.C. and R. K. Selander (1972) Evolutionary genetics of cave-dwelling fishes of the genus *Astyanax*. *Evolution* 26: 1–19.
- Bilandžija H., L. Ma, A. Parkhurst, and W. R. Jeffery (2013) A potential benefit of albinism in *Astyanax* cavefish: downregulation of the *Oca2* gene increases tyrosine and catecholamine levels as an alternative to melanin synthesis. *PLoS ONE* 8(11): e80823.
- Borowsky R. and H. Wilkens (2002) Mapping a cave fish genome: Polygenic systems and regressive evolution. *Journal of Heredity* 93: 19–21.
- Borowsky, R. (2008) Breeding *Astyanax mexicanus* through natural spawning. *Cold Harbor Spring Protocols*.
- Culver, D.C. (1982) *Cave Life: Evolution and Ecology*. Cambridge, MA: Harvard University Press. 189.
- Darwin, C. R. (1859) *On the Origin of Species*. 6th ed. London: John Murray.
- Dowling T.E., D. P. Martasian, and W. R. Jeffery (2002) Evidence for multiple genetic forms with similar eyeless phenotypes in the blind cavefish, *Astyanax mexicanus*. *Molecular Biology and Evolution* 19: 446–455.

- Erickson, C.A. and R. Perris (1993) The Role of Cell-Cell and Cell-Matrix Interactions in the Morphogenesis of the Neural Crest. *Developmental Biology*. 159(1): 60–74.
- Espinasa L. and R. B. Borowsky (2001) Origins and relationship of cave populations of the blind Mexican tetra, *Astyanax fasciatus*, in the Sierra de El Abra. *Environmental Biology of Fishes* 62: 233–237.
- Gross, J.B., R. Borowsky, and C. J. Tabin (2009) A Novel Role for *Mclr* in the Parallel Evolution of Depigmentation in Independent Populations of the Cavefish *Astyanax mexicanus*. *PLoS Genetics*. 5(1): 1–14.
- Gross, J.B. (2012a) The complex origin of *Astyanax* cavefish. *BMC Evolutionary Biology*. 12: 105.
- Gross, J.B. (2012b) *Cave evolution*. In: eLS. John Wiley & Sons, Ltd: Chichester.
- Huang, X. and J. P. Saint-Jeannet (2004) Induction of the neural crest and the opportunities of life on the edge. *Developmental Biology*. 275:1-11.
- Hüppop., K. (1986) Oxygen consumption of *Astyanax fasciatus* (Characidae, Pisces): A comparison of epigeal and hypogean populations. *Environmental Biology of Fishes*. 17(4): 299–308.
- Jeffrey, W.R. (2001) Cavefish as a model system in evolutionary developmental biology. *Developmental Biology*. 231: 1–12.
- Jeffery, W.R. 2005. Adaptive evolution of eye degeneration in the Mexican blind cavefish.

Journal of Heredity 96: 185–196

Jeffery, W.R. (2006) Regressive evolution of pigmentation in the cavefish *Astyanax*. *Israel*

Journal of Ecology and Evolution. 52: 405–422.

Jeffery, W.R. (2009a) Evolution and Development in the cavefish *Astyanax*. *Current topics in*

Developmental Biology. 86:191–221.

Jeffery, W.R. (2009b) Regressive evolution in *Astyanax* cavefish. *Annual Review of Genetics*.

43: 25–47.

Logan, D.W., R. J. Bryson-Richardson, M. S. Taylor, and I. J. Jackson (2003b) Sequence

characterization of the teleost fish melanocortin receptors. *Animals of the New York Academy of Sciences*. 994: 319–330.

Logan, D.W., S. F. Burn, and I. J. Jackson (2006) Regulation of pigmentation in zebrafish

melanophores. *Pigment Cell Research*. 19: 206–213.

McCauley, D.W., E. Hixon, and W. R. Jeffery (2004) Evolution of pigment cell regression in the

cavefish *Astyanax*: a late step in melanogenesis. *Evolution & Development*. 6:4, 209–218.

Menuet, A., A. Alunni, J. Joly, W. R. Jeffery, and S. Rétaux (2006) Expanded expression of

Sonic Hedgehog in *Astyanax* cavefish: multiple consequences on forebrain development and evolution. *Development*. 134: 845–855.

Metz, J.R., L. M. Peters, and G. Flik (2006) Molecular biology and physiology of the

- melanocortin system in fish: A review. *Genetics of Comparative Endocrinology* 148: 150–162.
- Mitchell, R.W., W. H. Russell, W. R. Elliot (1977) Mexican eyeless characin fishes, genus *Astyanax*: Environment, distribution, and evolution. Special Publications: The Museum, Texas Tech University. 12: 1–89.
- Niven., J.E. (2008) Evolution: Convergent eye losses in fishy circumstances. *Current Biology*. 18(1): 27–29.
- Poulson, T.L. (1981) Variations in life history of linyphiid cave spiders. *Proceedings of the Eighth International Congress of Speleology*, Bowling Green, KY. 1:60-62.
- Protas, M.E., C. Hersey, D. Kochanek, Y. Zhou, H. Wilkens, W. R. Jeffery, I. L. Zon, R. Borowsky, and C. J. Tabin (2006) Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nature Genetics*. 38: 107–111
- Protas, M., M. Conrad, J. B. Gross, C. Tabin, and R. Borowsky (2007) Regressive evolution in the Mexican tetra, *Astyanax mexicanus*. *Current Biology*. 17: 452–454.
- Rees, J.L. (2003) Genetics of Hair and Skin Color. *Annual Review Genetics* 37: 67–90.
- Rétaux, S. and K. Pottin (2008) Shh and forebrain evolution in the blind cavefish *Astyanax mexicanus*. *Biology of the Cell*. 100(3): 139–147.
- Riesz, M.P. (2004) Radiative relation quantum yields for synthetic eumelanin. *Photochemistry and photobiology* 79(2): 211–216.

- Robbins, A.H. (1991) *Biological perspectives on human pigmentation*. New York: Cambridge University Press.
- Şadoğlu, P. (1957a) A Mendelian gene for albinism in natural cave fish. *Experientia*. 13: 394.
- Şadoğlu, P. and A. McKee (1969) A second gene that effects eye and body color in Mexican blind cave fish. *Journal of Heredity* 60: 10–14.
- Şadoğlu, P. (1979) A breeding method for blind *Astyanax mexicanus* based on annual spawning patterns. *Copeia*. 1979: 369–371.
- Sarna, T. (1992) Properties and function of the ocular melanin – A photobiophysical view. *Journal of Photochemistry and Photobiology. B: Biology* 12: 215–258.
- Widlund H.R., and D. E. Fisher (2003) Microphthalmia-associated transcription factor: a critical regulator of pigment cell development and survival. *Oncogene* 22: 3035–3041
- Wilkens, H. (1988) *Evolution and genetics of epigeal and cave Astyanax fasciatus* (Characidae, Pisces): Support for the neutral mutation theory. In: Hecht MK, Wallace B. eds. *Evolutionary Biology*. New York: Plenum Publishing Corp. pp. 271–367.
- Wilkens, H. and U. Strecker (2003) Convergent evolution of the cavefish *Astyanax* (Characidae, Teleostei): genetic evidence from reduced eye-size and pigmentation. *Biological Journal of the Linnean Society*. 80: 545-554.
- Wright, S. (1964) Pleiotropy in the evolution of structural reduction and of dominance. *American Naturalist* 98: 65–70.

Yamamoto Y., D. W. Stock, and W. R. Jeffery WR. Hedgehog signalling controls eye degeneration in blind cavefish. *Nature* 431:844–473

CHAPTER 1

Alterations in *Mclr* gene expression are associated with regressive pigmentation in *Astyanax* cavefish¹

Bethany A. Stahl and Joshua B. Gross

Department of Biological Sciences, University of Cincinnati, 312 Clifton Court,
Cincinnati, Ohio, USA

¹In revision at *Development, Genes, and Evolution*

ABSTRACT

Diverse changes in coloration across distant taxa are mediated through alterations in certain highly conserved pigmentation genes. Among these genes, *Mclr* is a frequent target for mutation, and many documented alterations involve coding sequence changes. We investigated whether regulatory mutations in *Mclr* may also contribute to pigmentation loss in the blind Mexican cavefish, *Astyanax mexicanus*. This species comprises multiple independent cave populations that have evolved reduced (or absent) melanic pigmentation reduction as a consequence of living in darkness for millions of generations. Among the most salient cave-associated traits, complete absence (albinism) or reduced levels of pigmentation (*brown*) have long been the focus of degenerative pigmentation research in *Astyanax*. These two Mendelian traits have been linked to specific coding mutations in *Oca2* (albinism) and *Mclr* (*brown*). However, four of the seven caves harboring the *brown* phenotype exhibit unaffected coding sequences compared to surface fish. Thus, diverse genetic changes, involving the same genes, likely impact reduced pigmentation among cavefish populations. Using both sequence and expression analyses, we show that certain cave-dwelling populations harboring the *brown* mutation have substantial sequence alterations to the putative *Mclr* *cis*-regulatory region. Several of these sequence mutations in the *Mclr* 5' region were shared across multiple, independent cave populations. This study suggests that pigmentation reduction in *Astyanax* cavefish evolves through a combination of both coding and *cis*-regulatory mutations. Moreover, this study represents one of the first attempts to identify regulatory alterations linked to regressive changes in cave-dwelling populations of *Astyanax mexicanus*.

INTRODUCTION

Natural variants of the gene *Mc1r* explain several variable pigmentation phenotypes in natural populations. These phenotypes are often caused by coding sequence alterations, such as the *extension (e)* locus in mouse models. Recessive *Mc1r extension (e)* alleles produce a yellow phaeomelanin phenotype that impacts plumage in birds, and hair color variation in humans. At present, more than 60 coding sequence mutations have been identified in *Mc1r* (reviewed in García-Borrón et al. 2005).

Mc1r sequence alterations with functional impact often result from base pair deletions in the coding sequence, or amino acid substitutions (reviewed in Hoekstra 2006). Fewer studies have identified a role for regulatory mutations affecting pigmentation through altered *Mc1r* expression. However, *cis*-regulatory alterations affecting other pigmentation gene expression patterns have been identified in other systems. For instance, *Drosophila* wing spot coloration is influenced by *cis*-regulatory mutations impacting the *yellow* gene (Jeong et al. 2006).

Peromyscus deer mice demonstrating melanism (dark pigmentation) harbor a 125-kb deletion in the *Agouti* gene (antagonist of *Mc1r*) inclusive of a putative regulatory region, which reduces *Agouti* expression (Kingsley 2009). In stickleback fish, lighter pigmentation in the gills and ventrum of freshwater populations are caused by *cis*-regulatory changes affecting *Kitl* expression (Miller et al. 2007).

Cave animals are excellent models for investigating changes in coloration due to the recurrent loss of pigmentation in organisms living amidst the dark, subterranean environment. Among cave-dwelling models, the Mexican tetra *Astyanax mexicanus* is particularly powerful since closely related surface-dwelling fish are available, enabling direct comparisons between a surrogate ‘ancestral’ form and the 29 derived cave-dwelling populations. The genetic basis for

two key pigmentation phenotypes have been described in multiple independent populations, including absence of melanin (albinism; *Oca2*; Protas et al. 2006) and reduction in melanin content (*brown*; *Mc1r*; Gross et al. 2009).

For both albinism and *brown*, diverse loss-of-function or hypomorphic mutations have been characterized. For example, a 2-bp deletion ($\Delta 23,24$) in the *Mc1r* open reading frame of Pachón cavefish causes a frame-shift leading to a premature stop codon. A second (hypomorphic) mutation (C490T), identified in individuals from the Yerbaniz and Japonés cavefish populations, causes a cysteine substitution at position R164C homologous to the R160W mutant in humans (Gross et al. 2009). Although only three caves demonstrate coding sequence alterations in *Mc1r*, complementation tests suggested that the same gene might be responsible for *brown* in seven distinct cave populations (Table 1). However, in four of the populations harboring *brown*, the *Mc1r* coding sequence is identical to surface fish.

In this study, we investigated possible involvement of *cis*-regulatory mutations leading to the *brown* phenotype in cavefish compared to the (normally-pigmented) surface-dwelling form. We evaluated the upstream genomic sequence of *Mc1r* in several representative individuals drawn from multiple *brown Astyanax* populations. Indeed, when characterizing sequence variation in the 5' *Mc1r* region, we found a surprisingly high level of sequence diversity across cave-dwelling populations, with several instances of sequence variation within highly conserved non-coding elements. Quantitative gene expression analyses in adult fish revealed substantial changes in *Mc1r* expression differing between surface-dwelling fish and three cave populations – Pachón, Tinaja and Chica. There are no sequence differences in the *Mc1r* coding region of Tinaja or Chica compared to surface-dwelling fish based on a prior analysis in which the *Mc1r* open reading frame was sequenced and compared across multiple populations (Gross et al. 2009).

Table 1. Distribution of the *brown* phenotype among *Astyanax* cave populations.

Cave locality	Method of identification	Genetic alteration	References
<i>Brown</i> affected populations			
<i>Chica</i>	Observed in the wild	Unknown	Şadoğlu and McKee 1969
<i>Sabinos</i>	Observed in the wild	Unknown	Şadoğlu and McKee 1969 Şadoğlu and McKee 1969
<i>Pachón</i>	Observed in the wild Complementation tests	Δ23/24	Wilkins and Strecker 2003 Gross et al. 2009
<i>Curva</i>	Complementation tests	Unknown	Wilkins and Strecker 2003
<i>Piedras</i>	Complementation tests	Unknown	Wilkins and Strecker 2003
<i>Yerbaníz (Japonés)</i>	Complementation tests	C490T Δ23/24	Wilkins and Strecker 2003 Gross et al. 2009
Unaffected populations (confirmed absent)			
<i>Molino</i>	Complementation tests	None	Wilkins and Strecker 2003
Presence of <i>brown</i> unknown			
<i>Tinaja</i>	Unknown	Unknown	--

Previous functional analyses of *Mclr*, in which mRNA transcripts were abrogated using morpholino knock-down approaches, demonstrated that reduced levels of *Mclr* recapitulate the brownish eyes and reduced melanic content found in the classic *brown* phenotype (Gross et al. 2009). Thus, the functional role of *Mclr*, combined with the divergent expression patterns we report here, may explain the parallel evolution of reduced pigmentation in wild cavefish. Further, this study indicates that pigmentation reduction in geographically diverse cavefish populations may occur through both coding and regulatory mutations impacting the same gene.

MATERIALS AND METHODS

Sequence analysis of the 5' Mc1r region

Genomic DNA was isolated from fin clips derived from adult surface and nine cave populations including: Pachón, Molino, Japonés, Yerbaníz, Sabinos, Piedras, Tinaja, Curva, and Chica (Qiagen; Valencia, CA; n= 1 fin clip per locality). All cavefish used in this study were derived from wild-caught specimens collected and generously provided by Dr. Richard Borowsky (New York University). Lab-reared surface fish used in this study were derived from specimens collected in the wild at Arroyo Sarco (at the Río Sabinos drainage, Mexico) and the Río Valles drainage. Chica cavefish used in qPCR studies were obtained from a commercial supplier (Quality Marine; Los Angeles, CA). Roughly 2500 bp of the upstream genomic regions of *Mc1r* were isolated using the Universal GenomeWalker Kit (Clontech; Mountain View, CA) and amplified using overlapping primer sets in ~500 bp segments. PCR amplicons were subcloned into the pGEM-T Easy vector (Promega; Madison, WI) and sequenced (Operon; Huntsville, AL). Sequences were aligned using SeqMan Pro (DNASTAR.v.11.0; Madison, WI) and analyzed with MegAlign (DNASTAR.v.11.0; Madison, WI) using the Martinez-Needleman-Wunsch method for closely-related sequences.

Prediction of conserved non-coding elements

Consensus sequences were evaluated for conserved non-coding elements using the mVISTA online tool for comparative genomics (<http://genome.lbl.gov/vista>; Frazer et al. 2004) using the MLAGAN algorithm. Reference genomic sequence and annotation data from *Danio rerio* for *Mc1r* was retrieved using Ensembl (Zv9; www.ensembl.org). Program parameters were

adapted from similar VISTA studies (Frazer et al. 2004). Regulatory VISTA (rVISTA) predicted regulatory element motifs based on TRANSFAC Professional (v.9.2; 2005) to identify putative binding motifs for TFE, TFEB, Mitf and E-box transcription factors. We selected these specific transcription factors for our screen because they have previously been shown to interact with the putative promoter region of *Mclr* (Moro et al. 1999; Miccadei et al. 2008).

RNA-seq analyses

RNA was isolated from pools of 50 surface or Pachón cave embryos at 10 hours post-fertilization (hpf), 24 hpf, 36 hpf, and 72 hpf with the RNeasy Kit (Qiagen; Valencia, CA). RNA from juveniles was extracted from three individuals at four months of age. RNA-sequencing was performed in triplicate for 10 hpf – 72 hpf stages and in duplicate for the juveniles using Illumina HiSeq Technology (v.2 kit) at the Cincinnati Children’s Hospital Core Sequencing Facility (Cincinnati, OH). Sequencing reads were aligned to a comprehensive transcriptome template, inclusive of the *Mclr* sequence, in ArrayStar (DNASTar.v.11.0; Madison, WI) using RPKM normalization methods (Mortazavi et al. 2008; Gross et al. 2013) to calculate levels of gene expression. We tested for significant differences between surface and cavefish samples using a Student’s t-test with FDR (Benjamini Hochberg; DNASTAR.v.11.0). Sequencing reads are accessioned to the NCBI SRA (BioProject ID: PRJNA258661).

Quantitative PCR expression analyses

RNA pools were generated from four populations (surface fish, Pachón cavefish, Tinaja cavefish, and Chica cavefish) from adult fin clips using the RNeasy Kit (Qiagen; Valencia, CA). Cavefish from Pachón and Chica harbor the *brown* mutation based on direct observations,

classical genetic crosses and complementation studies (Şadoğlu and McKee 1969; Wilkens and Strecker, 2003). We obtained Chica cavefish from a commercial supplier (Quality Marine; Los Angeles, CA), similarly described in prior studies (Espinasa and Jeffery 2006). Although Tinaja cavefish have not been formally reported to harbor *brown* in the literature, they display the phenotypic characteristics of *brown* based on direct observation (reduced eumelanin content). Further, the Tinaja cave entrance is ~2.5 km from the Sabinos cave which does harbor *brown*, and prior reports suggest migration has occurred between these populations (Bradic et al. 2012). Thus, Tinaja was used for this study because lab-reared specimens are available, this population breeds consistently, and is derived from the older “El Abra” lineage which is closely related to several reported *brown* populations (Şadoğlu and McKee 1969). Total RNA (1 µg) was reverse transcribed into cDNA pools (Invitrogen; Waltham, MA; Roche; Indianapolis, IN). qPCR primers designed to amplify ~100 bp fragment for detecting *Mclr* (forward: CACGTCCAGCTCACTCTTCA; reverse: TAGAGCCCGGCAGTGAATAC). Fragments were amplified using EvaGreen super mix (Bio-Rad; Hercules, CA) with the cycling parameters: step 1—95 °C for 30 s, step 2—95 °C for 5 s, step 3—55.1 °C for 10 s, plate read, go to step 2 for 39 additional cycles. Standard curves were generated to check amplification efficiency. Normalized expression ($\Delta\Delta C_q$) was calculated from samples in sextuplet with CFX Manager Software (BioRad.v.3.1) using reference gene *GAPDH* (forward: TGTGTCCGTGGTGGATCTTA; reverse: TGTCGCCAATGAAGTCAGAG-3’), and surface fish samples served as an inter-run calibration across replicate experiments. Significant differences in *Mclr* gene expression were calculated for each cavefish population compared to surface using a two-tailed Student’s t-test (CFX Manager Software, BioRad.v.3.1).

RESULTS AND DISCUSSION

5' Mc1r sequence analyses in cavefish identify multiple genetic alterations that co-localize with highly conserved non-coding regions

Previous studies in *Astyanax* have identified specific coding sequence alterations leading to regressive pigmentation traits, but this study is among the first to describe the involvement of prospective *cis*-regulatory alterations in degenerative, cave-associated phenotypes. Here, we characterized the putative regulatory region of *Mc1r* (2510 bp) – a gene with well-documented physiological roles governing melanin dynamics in vertebrates. *Mc1r* is a single-exon gene; hence it does not harbor introns that could otherwise include regulatory elements. In other systems, the minimal *Mc1r* promoter lies immediately upstream of the translational start codon (Moro et al. 1999; Miccadei et al. 2008). Moreover, the 5'-genomic region of *Mc1r* in humans harbors characteristics of G protein-coupled receptor (GPCR) promoters, including several GC-rich sequences, lack of a TATA box, and presence of numerous transcriptional start sites and E-box motifs (Moro et al. 1999). Though the gene *Mc1r* mediates reduced pigmentation in *brown* individuals derived from the Pachón, Yerbaníz and Japonés *Astyanax* cave populations, little is still known about how *Mc1r* governs reduced pigmentation the other cave lineages harboring the *brown* phenotype (Chica, Sabinos, Curva and Piedras caves). To explore this, we evaluated the 5' region of *Mc1r* in nine cavefish populations – the seven caves known to express *brown*, one lineage that likely harbors *brown* (Tinaja) and one additional cave where *brown* is confirmed to be absent (Molino; Wilkens and Strecker 2003).

Comparisons of cavefish and epigeal forms identified 42 unique mutations in the 5' region upstream of *Mc1r* (Fig. 1; Table S1). Three of the genetic alterations we discovered were

present in all eight caves including Molino, therefore these likely do not contribute to *brown*. Additionally, 22 mutations were only present in a single cavefish population and 6 mutations were present in two or more caves, but not across all of the “*brown*” caves. However, 11 mutations were fixed in all seven cavefish lineages harboring *brown*, excluding the Molino cave, rendering these alterations as intriguing candidates for future functional promoter studies.

Additional sequence analyses, including a pairwise alignment of sequences across representatives from nine caves revealed that the mean sequence identity shared between cave and surface morphs was 99.18% and the average sequence index, which includes a penalty for gaps/deletions, was 98.4% for the putative promoter region of *Mc1r*. Furthermore, distance comparisons indicated a similar level of sequence divergence across the nine cave lineages compared to surface in 5' *Mc1r* (mean = 0.49). These analyses indicated that, despite a moderate level of sequence variation across cave populations, the overall similarity of sequences between cave and surface fish was rather modest.

We then explored whether any of the precise mutations we identified may potentially impact the expression of *Mc1r* by altering highly conserved regions of non-coding DNA that harbor binding sites for regulatory elements. To identify these putative regulatory regions, we first compared the upstream *Mc1r* promoter region derived from surface and nine cave-dwelling *Astyanax* populations to the same genomic interval in zebrafish (*Danio rerio*) using VISTA alignments. Recent studies document significant genomic similarity shared between *Danio* and *Astyanax* (both are members of the Ostariophysian superorder) despite ~150 My of divergence (Gross 2012). We identified 16 highly conserved non-coding elements (CNEs) based on our comparisons between *Danio* and *Astyanax* for *Mc1r* (Fig. 1; gray, peach and yellow boxes). Interestingly, many of the 5' *Mc1r* mutations not shared with Molino co-localize to these CNE's:

5' *Mc1r*

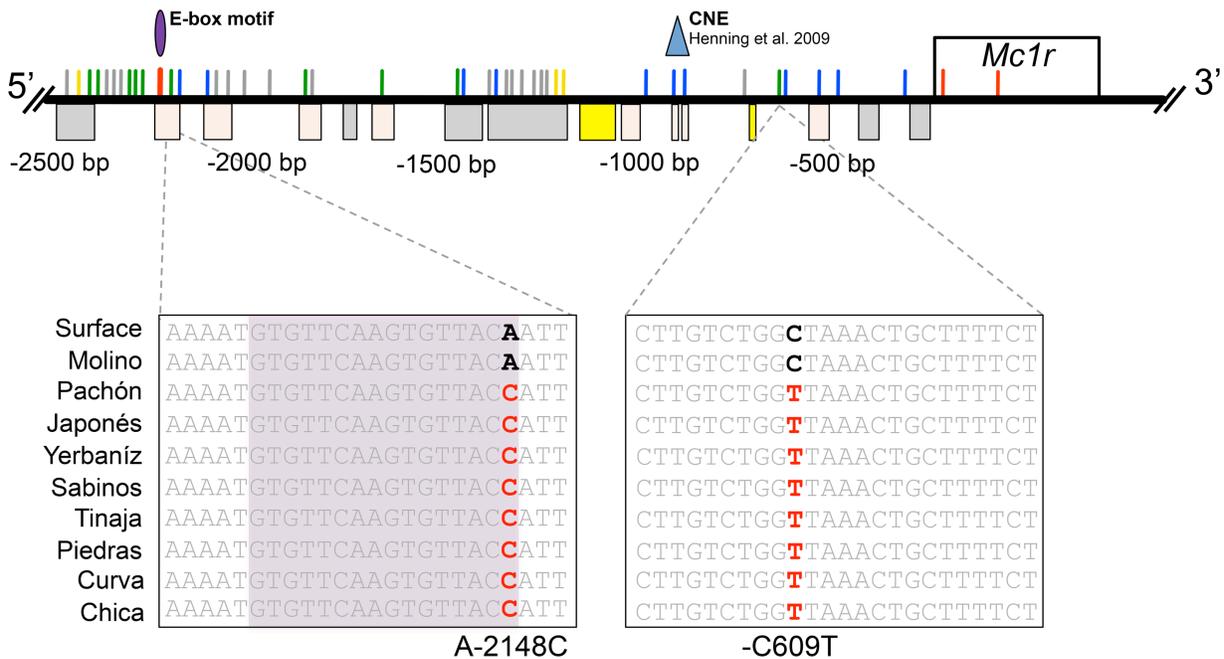


Figure 1. Multiple 5' *Mc1r* genetic mutations in cavefish co-localize to highly conserved non-coding elements and one transcription factor binding motif. Sequence analyses reveal the presence of numerous mutations throughout the 5' region of *Mc1r*. Any sequence alterations shared with Molino cave likely serve no functional significance (gray). Some mutations identified are only present in a single cavefish lineage (blue) and others are randomly present in two or more of the caves assayed (yellow). However, some alterations are present in all eight caves (excludes Molino; green). Boxes below the line indicate the highly conserved CNEs between *Danio* and *Astyanax* sequences detected with the VISTA alignments (Gray <55% conserved; Peach >55%; Yellow >75%). The highly conserved fish-specific CNE described in Henning et al. (2010) is marked with a blue triangle. Additionally, the E-box binding motif detected with rVISTA is represented by the purple oval, which co-localizes with the A-2148C mutation (red colored mutation) that is fixed in all caves except Molino.

6 alterations were found in one cavefish population, 3 mutations were found in 2 or more cave populations and 5 mutations were found in all eight cavefish lineages, except Molino (Table S1). Considering some of these different mutations reside in highly conserved non-coding regions, it is possible that distinct cavefish lineages are converging on the same *brown* phenotype via *Mclr*, but as a consequence of mutations that are specific to each cave. This would be similar to albinism, wherein Pachón and Molino caves demonstrate different coding sequence mutations (exon deletions) in the gene *Oca2* (Protas et al. 2006).

Next, we evaluated these conserved non-coding regions by identifying specific transcription factor binding sites within these regions, wherein cavefish-specific mutations that co-localize to these motifs could potentially hinder the binding efficiency, and ultimately, affect the expression of *Mclr*. We focused our search for binding motifs based on prior descriptions of transcription factors known to interact with the regulatory region of *Mclr*. In humans, the 5' region upstream of *Mclr* harbors several E-box motifs, and a band-shift assay demonstrated the binding of microphthalmia transcription factor (MITF), a melanocyte-specific regulator, to a 150bp region upstream of the *Mclr* start codon in pigment cells (Moro et al. 1999; Miccadei et al. 2008). MITF is a basic helix-loop-helix leucine zipper (bHLH-zip) transcription factor with highly conserved N-terminal domains and common putative phosphorylation sites; therefore, we searched for other MITF family members including TFE and TFEB. In sum, we focused our search for binding motifs to the E47 (binds E-boxes), MITF (binds M-boxes), TFE, and TFEB transcription factors.

The rVISTA analysis identified 8 putative binding motifs in the 5' *Mclr* region, all of which were conserved across both surface and hypogean populations with the exception of one motif present in a region 5' of *Mclr*. Among the putative transcription factor sites discovered,

one particular motif was mutated (nucleotide position -2148) in all *brown* cave populations, but the mutation was absent from both surface fish and the Molino cave population (which does not harbor *brown*; Strecker and Wilkens 2003). This polymorphism resides in a prospective E-box binding motif for the E47 transcription factor (Fig. 1; purple oval). The A-2148C mutation is not present in the core E47 motif (CAAGTG), however recent evidence shows that the genomic sequence immediately adjacent to E-box transcription factor motifs can influence binding specificity of bHLH factors by impacting the shape of the DNA molecule (Yasumoto et al. 1994; Gordan et al. 2013). It is possible that these binding motifs may have accumulated mutations in cavefish lineages, leading to a change in gene expression (and therefore the *brown* phenotype) in distinct subterranean populations. Moreover, sequence alignments between *Astyanax* and *Danio* demonstrate that the E-box site at position -2148 is present within a highly conserved non-coding element. This level of sequence similarity may indicate that this binding site is functional and serves a critical role in the regulation of *Mc1r*. To date, the *Mc1r* regulatory regions tested with gene deletion studies are limited to human-specific sequences which only tested functional binding of one prospective transcription factor (Mitf; Moro et al. 1999; Miccadei et al. 2008). Therefore, future functional analyses (e.g., promoter deletion constructs) will both identify the specific *Mc1r* minimal promoter region required to drive gene expression in teleosts, and determine whether the A-2148C is a causative regulatory lesion shared among *brown* cave populations.

Quantitative analyses reveal reduced Mc1r expression in later stages of development across independent populations cavefish

If sequence mutations in the putative *Mc1r* promoter confer reduced expression leading to loss of pigmentation in cave-dwelling forms of *Astyanax*, we would predict a reduction in transcript abundance of this gene. To examine this possibility, we first evaluated *Mc1r* gene expression during early development, wherein pigmentation begins to appear at ~24 hpf, using a next-generation sequencing approach comparing surface morphs and Pachón cavefish.

Differential gene expression of *Mc1r* was evaluated in triplicate across four early developmental stages (10 hpf, 24 hpf, 36 hpf, 72 hpf) and juvenile (~4 months old; duplicates), encompassing both early and juvenile stages of pigmentation development in *Astyanax*.

Interestingly, during early stages of development (10 hpf to 72 hpf), *Mc1r* is expressed slightly higher in Pachón cavefish compared to surface fish. However, by the juvenile stage, *Mc1r* demonstrates a moderate increase (~5.155 fold) in surface fish, compared to cavefish (Fig. 2). Although we detected subtle differences in *Mc1r* gene expression, none of the early time points we tested showed statistically significant differences between surface and Pachón cavefish. This may be due to the fact that many pigmentation genes are expressed at extremely low levels *in vivo*. Future studies increasing the mRNA-sequencing depth may help provide more sensitive measurements of expression differences of *Mc1r* between morphotypes during early developmental stages.

Although RNA-seq studies indicated subtle differences between surface and Pachón cavefish, the polarity of the *Mc1r* expression difference was consistent with our predictions at the juvenile stage. Therefore, we sought to further characterize *Mc1r* expression in later development (i.e., adulthood) in multiple *brown* cavefish lineages using qPCR. We quantified

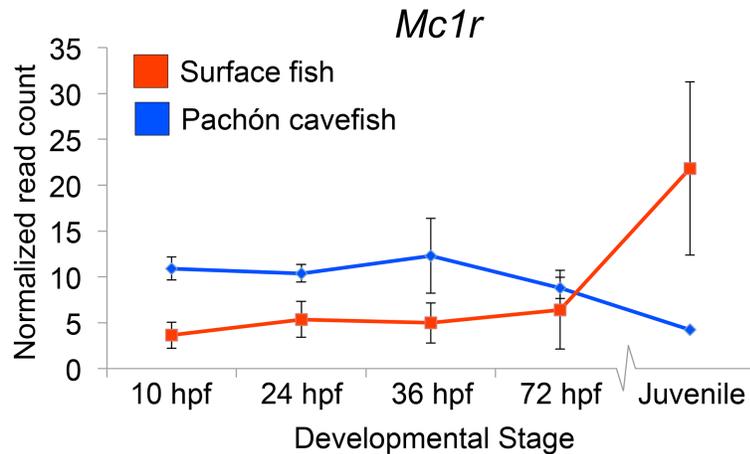


Figure 2. RNA-seq analyses demonstrate subtle differences in pigmentation gene expression between morphotypes in early development. Surprisingly, degrees of differential expression observed were relatively subtle between surface and Pachón cave-dwelling forms over early development. *Mc1r* appears slightly higher in cavefish than surface until juvenile hood.

mRNA transcriptional abundance in surface fish and cavefish derived from three depigmented populations: Pachón, Tinaja, and Chica using quantitative PCR. Pachon and Chica cavefish exhibit *brown*, however this phenotype has not been formally evaluated in Tinaja cavefish (Gross et al. 2009). Tinaja fish, however, exhibit dramatically reduced levels of melanic-based pigment (Protas et al. 2006; Fig. 3A-D). Our results demonstrated that *Mc1r* expression was significantly reduced in all three cavefish populations compared to surface fish (Fig. 3E; $p < 0.0001$).

Interestingly, *Mc1r* expression patterns were not identical across these independent troglomorphic localities. For example, the normalized expression in Tinaja and Chica cavefish was appreciably lower than the *Mc1r* levels detected in Pachón cavefish (Fig. 3E).

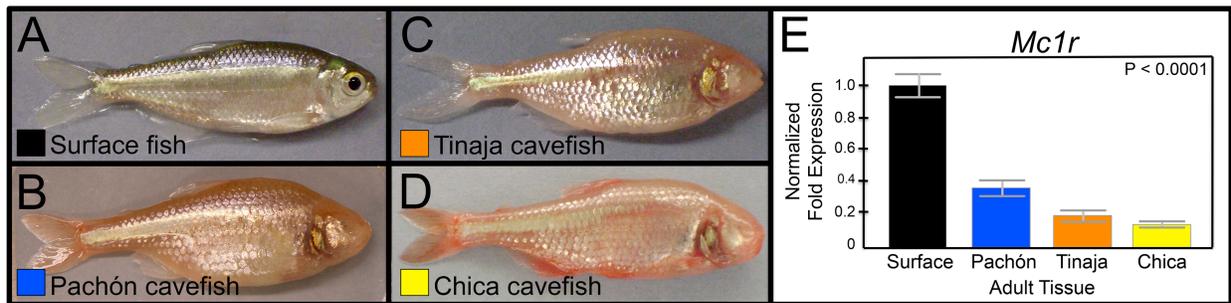


Figure 3. qPCR analyses reveal reduced *Mc1r* expression present in independent populations of adult cavefish. Varying levels of coloration are observed in *Astyanax mexicanus*. The surface-dwelling form is highly pigmented (black, A), compared to albino Pachón cavefish (blue, B). A severe reduction in pigmentation is also observed in both Tinaja (orange, C) and Chica cavefish (yellow, D). Normalized *Mc1r* gene expression levels for Pachón, Tinaja and Chica cave populations were significantly lower compared to the (ancestral) surface morphotype ($p < 0.0001$; E).

Since the Pachón *Mc1r* transcript harbors a destructive 2-bp deletion, an alternative explanation for its reduced transcriptional abundance may be through degradation via nonsense-mediated decay. This scenario could mean that upstream changes to the 5'-UTR may not have functional consequences in this cave population, but still be consistent with the qPCR results we report here. However, this mRNA deprivation could coincide with – or occur independently from – the upstream sequence alterations we discovered. Distinguishing between these two (not mutually exclusive) scenarios in Pachón cavefish cannot be easily answered with the approaches we employed in this study. However, it is interesting that two other (putatively) *brown* cave populations demonstrate reduced *Mc1r* expression despite having an intact *Mc1r* coding sequence.

Shedding light on the evolutionary and genetic mechanisms leading to pigmentation loss in the subterranean environment

Degenerative loss of pigmentation consistently recurs in disparate organisms that invade the cave environment. However, the precise evolutionary mechanism(s) leading to regressive loss has long perplexed cave biologists, including Charles Darwin, who attributed eye loss in cavefish to “disuse”. Indeed, the regression of characters in cave animals may evolve through *selection* (direct or indirect) or *neutral mutation* (through genetic drift). The precise benefit(s) of pigmentation loss in subterranean animals is poorly understood, however some other cave-related traits, namely eye loss, may be evolving through indirect selection through the expansion of non-visual sensory systems (Reviewed in Gross 2012). This concept is highly dependent upon whether the character is ecologically relevant. In cave-dwelling *Astyanax* populations, melanic-based pigment translocation (for background adaptation) is useless in the dark, subterranean environment. Thus, the absence of light may permit genes related to pigmentation to accumulate mutations following colonization of the cave habitat. A combination of previous QTL studies, and the presence of the *Oca2* loss-of-function allele in a non-albino cave population indicate that pigmentation loss in cave morphs may be evolving as a consequence of neutral forces (Gross and Wilkens 2013).

Alternatively, it may be that genes contributing to pigmentation may have pleiotropic consequences that are adaptive in cavefish. For example, a recent study showed experimentally reduced levels of *Oca2* in cavefish (using morpholino knock-down) influences levels of catecholamines, which in turn controls other behaviors in cavefish such as sleep and feeding (Bilandžija et al. 2013). These behavioral alterations may provide a benefit to cavefish in the nutrient-poor cave environment, however the precise mechanism through which this has evolved

remains unclear. Forthcoming genomic scans will clarify if certain genetic mutations, such as those discovered for *Mc1r* and *Oca2*, have risen to high frequencies in independent cavefish populations as a consequence of strong natural selection.

CONCLUSIONS

Little is known about the regulatory structure controlling *Mclr* expression. The regulatory structure of *Mclr* is best characterized in human and mouse. A luciferase promoter assay in two independent experiments suggested that the *Mclr* minimal promoter lies immediately 5' of the translational start and is approximately 500 bp in length (Moro et al. 1999; Miccadei et al. 2008). Interestingly, a VISTA alignment comparing putative regulatory regions of *Mclr* in several fish species including pufferfish, medaka, stickleback, and Midas cichlids identified a CNE present ~818 bp – 855 bp upstream of the translational start (Henning et al. 2010; Fig. 1; blue triangle). Together, these results suggest that sequence mutations present in the upstream *Mclr* cis-regulatory region may be associated with reduced pigmentation in *Astyanax* cavefish. Although seven caves harbor the *brown* mutation, four cave populations demonstrate no coding sequence alterations compared to surface fish (Gross et al. 2009). However, complementation crosses indicate that the same genomic region underlies the *brown* phenotype (Wilkens and Strecker 2003). Previous studies of albinism in cavefish similarly showed that independent cavefish populations likely acquired albinism through the same (or closely linked) locus. A subsequent study supported this notion – demonstrating that different coding sequence mutations in the same gene (*Oca2*) govern albinism in distinct populations (Protas et al. 2006). In this study, we similarly identified substantial sequence variation – some shared mutations and others unique – in the 5' *Mclr* region in independent cavefish lineages. Since experimentally reduced *Mclr* expression recapitulates the *brown* phenotype (Gross et al. 2009), this study suggests that regulatory changes may also contribute to the *brown* phenotype. The work presented here represents one of the first examples of a potential role for regulatory alterations mediating a classic regressive phenotype in *Astyanax* cavefish.

Alterations in key regulatory regions underlie morphological and physiological variation across broad taxa, including amphibians, birds, mammals, and insects (reviewed in Hoekstra 2006). For instance, pelvic spine reduction has evolved repeatedly in sticklebacks through deletion of a *Pel* enhancer, located directly upstream of the *Pitx1* promoter (Chan et al. 2010). The pattern of pigmentation gene expression demonstrated by *Mc1r* varied dynamically across development, but only demonstrated significantly different morphotype-specific expression at the adult stage. Forthcoming larger-scale expression analyses, evaluating more pigmentation-related genes across early development and juvenile hood, will better inform how particular genes impact complex pigmentation phenotypes in *Astyanax*.

In conclusion, *brown* cave populations that demonstrate no coding sequence errors in *Mc1r* still demonstrate reduced expression of this gene. This degenerative pigmentation phenotype may have evolved through the accumulation of sequence mutations affecting the 5' regulatory region. Thus, regressive pigmentation in *Astyanax* cavefish may have evolved through a combination of both coding and *cis*-regulatory alterations. The contribution of both forms of sequence variation implies a role for regulatory alterations, alongside coding sequence variation, in the evolution of cave-associated traits.

Funding

This study was funded by the National Science Foundation, Washington D.C., U.S.A. (grant number DEB-1457630).

Acknowledgements

The authors wish to thank Dr. Richard Borowsky (New York University) for generously providing fin clips used in this report. In addition, we thank the members of the Gross laboratory for helpful discussions. Two anonymous reviewers provided helpful criticisms on an earlier draft of this manuscript.

Conflict of Interest Statement

The authors declare that they have no conflict of interest.

Ethical Statement

All applicable international, national, and/or institutional guidelines for the care and use of animals were followed. All procedures performed in studies involving animals were in accordance with the ethical standards of the institution or practice at which the studies were conducted. The protocol was approved by the Institutional Animal Care and Use Committee (IACUC) of the University of Cincinnati (Protocol Number 10-01-21-01).

Informed Consent

This article does not contain any studies involving human participants performed by any of the authors. Therefore, for this type of study, formal consent is not applicable.

REFERENCES

- Bilandžija H., L. Ma, A. Parkhurst, and W. R. Jeffery (2013) A potential benefit of albinism in *Astyanax* cavefish: downregulation of the *Oca2* gene increases tyrosine and catecholamine levels as an alternative to melanin synthesis. *PLoS ONE* 8:e80823.
- Bradic M., P. Beerli, F. García-de León, S. Esquivel-Bobadilla, and R. L. Borowsky (2012) Gene flow and population structure in the Mexican blind cavefish complex (*Astyanax mexicanus*). *BMC Evolutionary Biology* 12: 9.
- Chan Y.F., M. E. Marks, F. C. Jones, G. Villarreal Jr., M. D. Shapiro, S. D. Brady, A. M. Southwick, D. M. Absher, J. Grimwood, J. Schmutz, R. M. Meyers, D. Petrov, B. Jónsson, D. Schluter, M. A. Bell, and D. M. Kingsley (2010) Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. *Science* 327: 302–305.
- Espinasa L. and W. R. Jeffery (2006) Conservation of retinal circadian rhythms during cavefish eye degeneration. *Evolution and Development* 8(1): 16–22.
- Frazer K.A., L. Pachter, A. Poliakov, E. M. Rubin, and I. Dubchak (2004) VISTA: computational tools for comparative genomics. *Nucleic Acids Research*. 2: W273–9.
- García-Borrón J.C., B. L. Sánchez-Laorden, and C. Jiménez-Cervantes (2005) Melanocortin-1 receptor structure and functional regulation. *Pigment Cell Research* 18: 393–410.
- Gordan R., N. Shen, I. Dror, T. Zhou, J. Horton, R. Rohs, and M. L. Bulyk (2013) Genomic regions flanking E-Box binding sites influence DNA binding specificity of bHLH transcription factors through DNA shape. *Cell Reports* 3: 1093–1104.

- Gross J.B., R. Borowsky, and C. J. Tabin (2009) A novel role for *Mclr* in the parallel evolution of depigmentation in independent populations of the cavefish *Astyanax mexicanus*. *PLoS Genetics* 5:e1000326.
- Gross J.B. (2012) *Cave evolution*. Encyclopedia of Life Sciences (eLS; Wiley Press). John Wiley & Sons Ltd, Chichester, UK.
- Gross J.B., A. Furterer, B. M. Carlson, and B. A. Stahl (2013) An integrated transcriptome-wide analysis of cave and surface dwelling *Astyanax mexicanus*. *PLoS ONE* 8(2):e55659.
- Gross J.B. and H. Wilkens H (2013) Albinism in phylogenetically and geographically distinct populations of *Astyanax* cavefish arises through the same loss-of-function *Oca2* allele. *Heredity* 111: 122–130.
- Henning F., A. J. Renz, S. Fukamachi, and A. Meyer (2010) Genetic, comparative genomic, and expression analyses of the *Mclr* locus in the polychromatic Midas cichlid fish (Teleostei, Cichlidae *Amphilophus* sp.) species group. *Journal of Molecular Evolution* 70: 405–412.
- Hoekstra H.E. (2006) Genetics, development and evolution of adaptive pigmentation in vertebrates. *Heredity* 97: 222–234.
- Jeong S., A. Rokas, and S. B. Carroll (2006) Regulation of body pigmentation by the abdominal-B hox protein and its gain and loss in *Drosophila* evolution. *Cell* 125: 1387–1399.
- Kingsley E.P., M. Manceau, C. D. Wiley, and H. E. Hoekstra (2009) Melanism in *Peromyscus* Is caused by independent mutations in *Agouti*. *PLoS ONE* 4:e6435.

- Miccadei S., B. Pascucci, M. Picardo, P. G. Natali, and D. Civitareale (2008) Identification of the minimal melanocyte-specific promoter in the *Melanocortin receptor 1* gene. *Journal of Experimental and Clinical Cancer Research* 27: 71.
- Miller C.T., S. Beleza, A. A. Pollen, D. Schluter, R. A. Kittles, M. D. Shriver, and D. M. Kingsley (2007) *cis*-regulatory changes in *Kit Ligand* expression and parallel evolution of pigmentation in sticklebacks and humans. *Cell* 131: 1179–1189.
- Moro O., R. Ideta, and O. Ifuku (1999) Characterization of the promoter region of the human *Melanocortin-1 Receptor (MC1R)* gene. *Biochemical and Biophysical Research Communications* 262: 452–460.
- Mortazavi A., B. A. Williams, K. McCue, L. Schaeffer, and B. Wold (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* 5: 621–628.
- Protas M.E., C. Hersey, D. Kochanek, Y. Zhou, H. Wilkens, W. R. Jeffrey, L. I. Zon, R. Bowowsky, and C. J. Tabin (2006) Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nature Genetics* 38: 107–111.
- Şadoğlu P. and A. McKee (1969) A second gene that affects eye and body color in Mexican blind cave fish. *Journal of Heredity* 60: 10–14.
- Wilkens H. and U. Strecker (2003) Convergent evolution of the cavefish *Astyanax* (Characidae, Teleostei): genetic evidence from reduced eye-size and pigmentation. *Biology Journal of the Linnean Society* 80: 545–554.

Yasumoto K., K. Yokoyama, K. Shibata, Y. Tomita, and S. Shibahara (1994) Microphthalmia-associated transcription factor as a regulator for melanocyte-specific transcription of the human tyrosinase gene. *Molecular and Cellular Biology* 14: 8058–8070.

SUPPLEMENTARY TABLES

Table S1. Detailed sequence analysis of mutations and conserved non-coding regions in the 5'*Mclr* region.

Description	Sequence alteration present in 5' region	Alteration co-localizes to CNE	Alteration co-localizes to predicted TF binding motif
Alterations present in all eight caves, including Molino (<i>brown</i> confirmed absent)	3	1	0
Alterations present in two or more of the eight caves (excludes Molino)	6	6	0
Alterations present in a single El Abra cavefish population	22	12	0
Alterations fixed in all seven El Abra caves (excludes Molino)	11	5	1
Total:	42	24	1

CHAPTER 2

High-resolution genomic mapping reveals genes contributing to complex melanophore variation in *Astyanax mexicanus* cavefish¹

Bethany A. Stahl, Molly Perkins and Joshua B. Gross

Department of Biological Sciences, University of Cincinnati, Cincinnati, Ohio 45221

¹The data collection, analyses and composition presented here are solely the work of Bethany A. Stahl. It is expected that the contents of this chapter will be adapted and submitted for publication with the listed co-authors pending assistance and input at that time.

ABSTRACT

Regressive phenotypic evolution is a widespread phenomenon that likely affects every living organism, yet the mechanisms underlying loss remain largely unknown. Cave animals represent excellent models to examine this phenomenon, owing to the common loss of eyes and pigmentation among lineages evolving in the subterranean habitat. The blind Mexican cavefish, *Astyanax mexicanus*, is a particularly powerful system because “ancestral” surface-dwelling forms are available, allowing for direct comparisons with cave-dwelling morphs. To date, only two genes (*Oca2*-albinism and *Mc1r*-brown) have been linked to specific pigmentation phenotypes in cavefish. However, pigment cell (melanophore) number is a complex trait governed by multiple genes, and the identities of these loci remain unidentified. To uncover these genes, we assembled a high-resolution, second-generation linkage map. We adapted an automated phenotypic scoring strategy, in which melanophore number is quantified across seven body regions in a surface x Pachón cave F₂ pedigree. QTL mapping analyses yielded several markers strongly associated with melanophore number variation in the dorsal mid-lateral stripe area and in the superior head region, which anchor to regions of the *Astyanax* draft genome and the zebrafish genome. Within these syntenic regions we identified two candidate genes, *Tyrp1b* and *Pmela*, with known roles in pigmentation-related processes based on Gene Ontology annotation. Mutant forms of these candidate genes in other systems cause global and regional pigmentation losses, respectively. In-depth analyses of these genes reveal coding sequence mutations and differential expression patterns in Pachón cavefish compared to surface morphs. This work uniquely identifies genes involved with complex aspects of *Astyanax* pigmentation, and provides insight to genetic mechanisms governing regressive phenotypic change.

INTRODUCTION

Pigmentation can range dramatically across the animal kingdom – from subtle crypsis to colorful ornamental displays – suggesting that coloration serves equally dynamic and adaptive functions. In many animals, these roles can range from mate choice selection (Protas and Patel 2008), cryptic coloration for defense from predators (Linnen et al. 2009), UV protection, structural support, and thermoregulation (Hubbard et al. 2010). Moreover, many pigmentation traits have served as a powerful model for linking specific genes to phenotypic characters (Hoekstra 2006). However, in most of these cases, the genes with known roles in pigmentation are often associated with simple Mendelian traits and knowledge of complex interactions among pigmentation genes is limited.

To investigate polygenic aspects of coloration, we turn to the emerging genetic model system *Astyanax mexicanus* (Jeffery 2001). This unique system harbors two distinct forms: a highly pigmented surface-dwelling form that populates the rivers of NE Mexico and a colorless cave-dwelling morphotype that resides exclusively in the subterranean environment. The blind Mexican cavefish is an excellent model for investigating this problem, owing to the recurrent loss of pigmentation and eyes in animals living in total darkness. Importantly, *Astyanax* forms have repeatedly colonized the cave, leading to at least 29 distinct cave-dwelling populations, providing numerous biological replicates (Gross 2012a).

Two genes affecting classic Mendelian color traits have been identified in *Astyanax*. The locus underlying the complete absence of pigmentation, albinism, is caused by the gene *Oca2* in two independent cavefish lineages (Protas et al. 2005). A second degenerative pigmentation phenotype, called *brown*, is associated with the gene *Mc1r* in three cavefish populations (Gross et al. 2009). Although these studies have discovered the genetic basis for simple pigmentation

loss, numerous other lineages show reduced pigmentation, yet the genes contributing to complex pigmentation loss in these caves have not been identified. An earlier mapping study aimed to characterize the variation in melanophore (pigment cell) number. This study detected 18 QTL associated with pigment cell number, suggesting melanophore number is indeed a complex trait involving multiple genes contributing to variation, yet the precise genes governing this trait remain unknown (Protas et al. 2007).

To uncover the identity of these genes, we employed a second generation linkage map (Carlson et al. 2015) inclusive of over 3,000 genomic markers to perform high-resolution mapping of melanophore number diversity in a large cave x surface F₂ pedigree (Fig. 1A-E). Our quantitative trait locus (QTL) mapping studies yielded numerous significant associations linked with 20 different regions of our linkage map. We then leveraged recently available genomics resources, specifically a draft genome (McGaugh et al. 2014), to nominate candidate genes involved with complex pigmentation. Comparative genomics identified the critical syntenic region for each QTL in the *Astyanax* draft genome and conserved intervals in the zebrafish genome. We nominated candidate genes by screening the genes within these syntenic regions for Gene Ontology (GO) terms related to pigmentation. These analyses yielded two intriguing candidate genes *Tyrp1b* and *Pmela* with previously known roles in melanin-based coloration in other systems. We further characterized the coding sequence and expression of candidate genes. Through these integrative studies, we identify *Tyrp1b* and *Pmela* as genes that likely contribute to complex pigmentation in *Astyanax*. Identification of additional genes will provide a clearer picture of the mechanisms contributing to regressive evolution in *Astyanax*, and inform our understanding of the broader principles governing trait loss in the natural world.

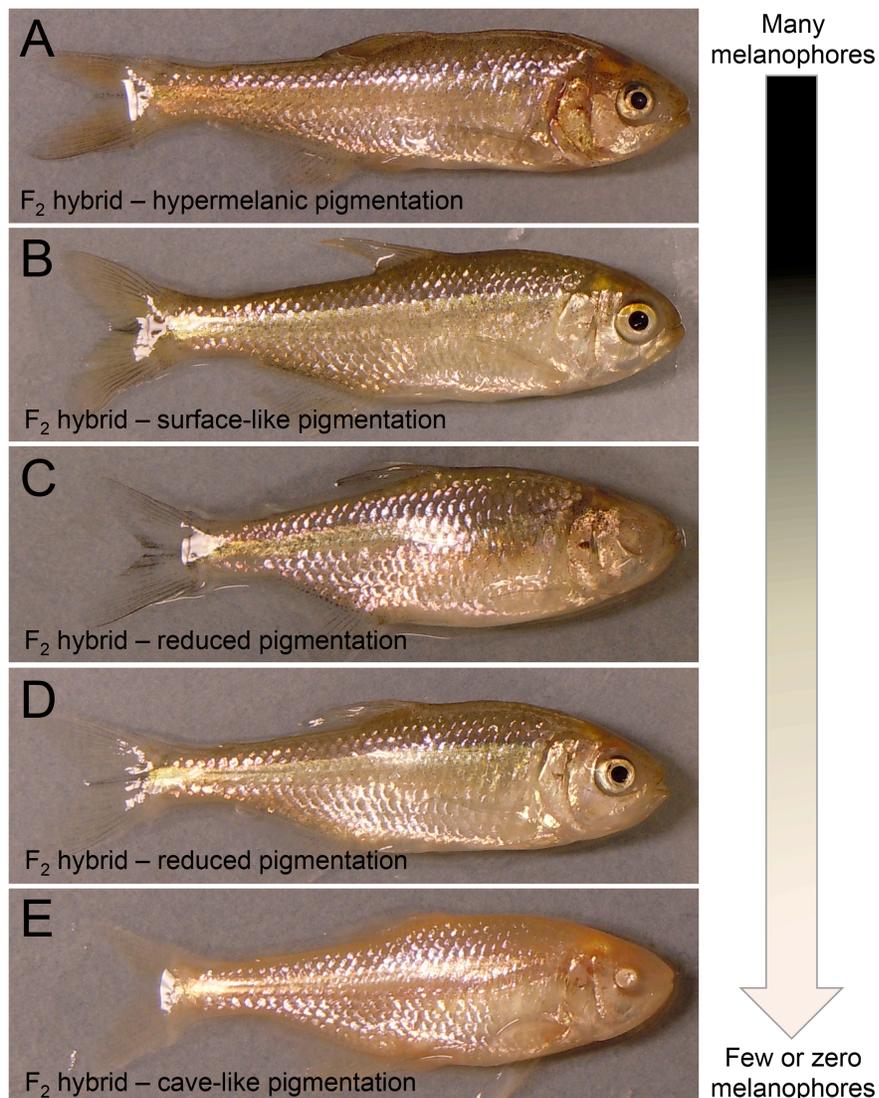


Figure 1. Surface x Pachón cave F₂ sibling hybrids display a vast array of coloration.

Hybrid offspring from a surface x Pachón cave cross reveal a varying degree of melanic-based pigmentation. The “dark” vs. “light” appearance is tightly linked with the number of melanophores (pigment cells) that an individual harbors (arrow, right side). Interestingly, we observe levels of pigmentation that are even higher (A) than the normally pigmented surface-like fish (B). Some hybrid individuals show dramatic reductions in pigmentation (C, D), while others demonstrate near or complete absence of pigmentation (cave-like, E).

MATERIALS AND METHODS

F₂ pedigree, husbandry, and genotyping

The linkage mapping and QTL association studies were performed using genotypic and phenotypic information derived from individuals of two separate pedigrees (n= 129; n= 41), each bred from a male surface x female Pachón cavefish cross, as described in Carlson et al. (2014). The adult parental specimens were laboratory strains that originated from the El Abra region of Northeastern Mexico and generously provided by Dr. Richard Borowsky (New York University). All live fish used in this study were reared as described in (Gross et al. 2013). Genomic DNA was isolated from surface (n= 4), Pachón cavefish (n= 4), surface x Pachón cave F₁ hybrids (n= 4) and F₂ individuals (n= 140) using the DNeasy Blood and Tissue Kit (Qiagen; Cat. No. 69506). The high-quality DNA was then processed (sequenced and library construction) using genotyping-by-sequencing (GBS) technologies at the Institute for Genomic Diversity (Cornell University) using previously described methods (Elshire et al. 2011; Lu et al. 2013), yielding a total of 7,656 GBS (SNP-based) markers (naming prefix “TP”). We also genotyped our pedigree for 959 additional genomic markers using Sequenom iPLEX Sequencing (Broad Institute). These markers were representative of both diverse positional and candidate loci and were generated during a previous SNP screen (Floragenex; Eugene, OR) and sequences generated with recent *Astyanax* transcriptomic study (naming prefix “ASTYANAX”; Gross et al. 2013). The genotypes of surface and cave individuals were used to determine the origin of each allele in the F₂ hybrid individuals, utilizing methodology as described in (Carlson et al. 2015). Moreover, markers were screened using stringent filtering and unsuitable markers were discarded from further analysis (Carlson et al. 2015). This yielded 3003 GBS and 917 Sequenom markers evaluated in 170 F₂ individuals.

Melanophore scoring

To quantify melanophore number for subsequent QTL analyses, we targeted numerous regions across the body where pigment cell number varied dramatically across our F₂ pedigree (Fig. 2A). This included seven regions: near the anal fin (MelAnalfinSquare; Fig. 2B), below the mid-lateral stripe (MelUnderStripe; Fig. 2C), dorsal square (MelDorsalSquare; Fig. 2D), area above the stripe (MelAboveStripe; Fig. 2E), above the eye (MelHeadSquare; Fig. 2F), full head (MelHead; Fig. 2G), and by the anal fin (MelAnalfinTriangle; not pictured). Regions were selected based on a consistent set of landmarks for reproducibility and were similar to the areas previously assayed (Protas et al. 2007). For counts, an automated system in ImageJ (v.1.6; National Institutes of Health, Bethesda, MD) quantified melanophore number, by first inverting the color in a selected area and then counting the light objects (e.g., pigment cells, now white) with a preset “noise tolerance.” The noise tolerance is set so only markings (melanophores in this case) above preset light pixel intensity are counted. The noise tolerance will be set slightly lower to avoid erroneous light pixels/debris from being included in the measurements. Each image was reviewed and any melanophores “missed” in the automatic quantification were added manually. When appropriate, we transformed the melanophore counts to log¹⁰ values to generate a normal distribution for association studies.

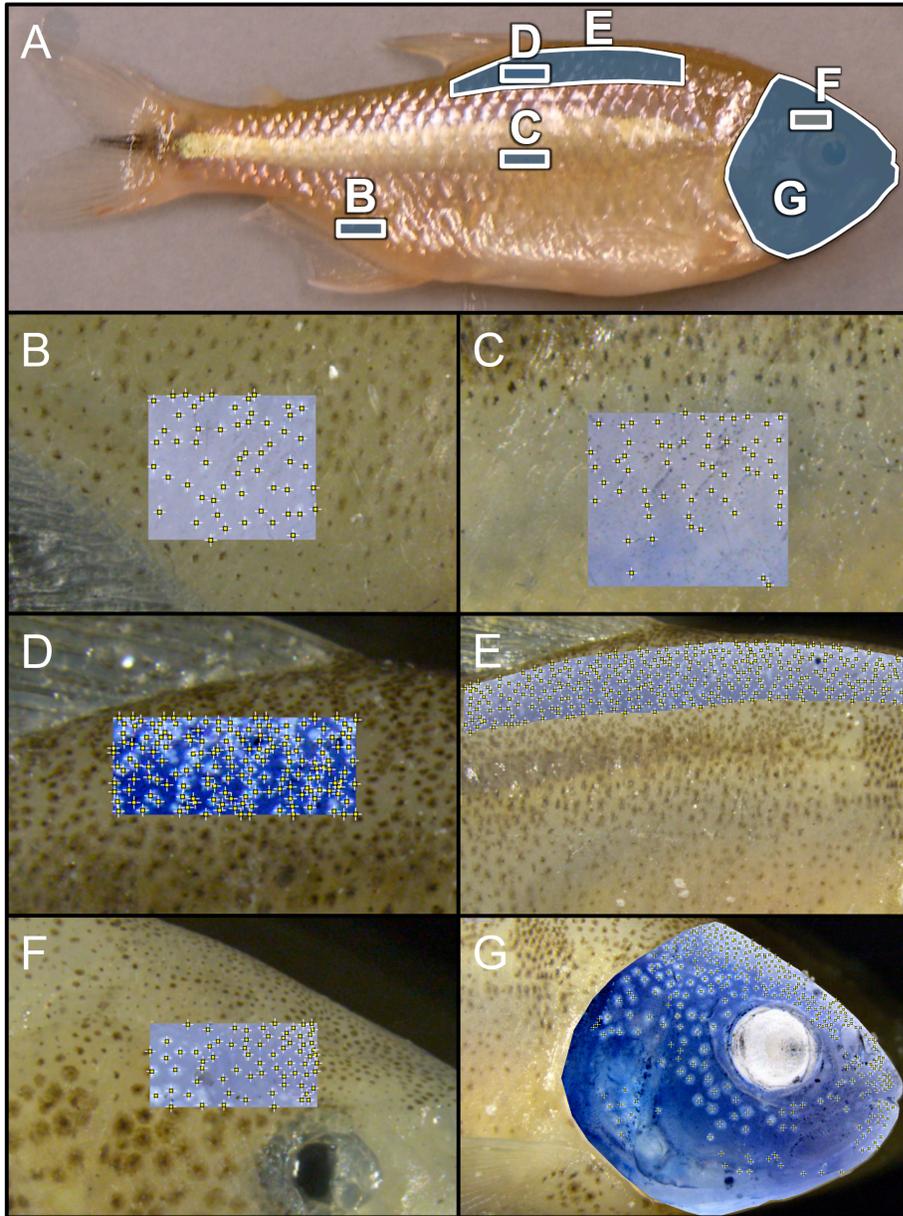


Figure 2. Quantifying melanophore number for quantitative trait locus (QTL) analyses.

Melanophore number was scored in a surface x Pachón F_2 pedigree for QTL association analyses to identify genes contributing to this complex trait. Seven regions spanning across the body were assayed (A). Consistency across individuals was maintained with the use of precise landmarks (fins, joints, etc.) and consistent box size per region (C, E, F, G). These regions included above the anal fin (B), below the stripe (C), below the dorsal fin (D), above the stripe (E), above the

eye (F), and the full head (G). Melanophore counting was performed with ImageJ software. Each selected region was outlined appropriately and color inverted. Since melanophores are normally dark in color and the program counts the bright spots, image inversion permitted the program to correctly identify the melanophores. Individual melanophores counted correctly by the program were automatically denoted with a yellow plus (+) sign (B, C, D, E, F, G).

Linkage map

For this study, we capitalized on the recently published high-density *Astyanax mexicanus* linkage map produced by (Carlson et al. 2015) that was generated using recombination frequencies (LOD score cutoff of 50) using the program JoinMap (v.4.1; Kyazma). This program uses a maximum-likelihood algorithm, appropriate for F₂ sibling populations, to arrange markers and calculate distances between loci by using multiple sampling methods. To further expand upon this map, we integrated 917 additional markers (described above) based on known sequences to ensure representation of markers across the entire genome. This yielded a densely populated genetic map constructed of 3398 markers comprising 26 linkage groups (*Astyanax* karyotypic number = 25; Kirby et al. 1977) with a mean inter-marker distance of 1.030 cM and a total map length of 3320.944 cM. Finally to ensure the reliability of this map, we scored for albinism (presence vs. absence) in our pedigree – a simple trait that is governed by a single gene (Protas et al. 2005) – of which we were able to successfully map near the known locus *Oca2*, confirming our ability to nominate candidate genes in subsequent analyses.

Quantitative trait locus (QTL) association mapping

To analyze the complex trait of melanophore variation, we employed the software program R/qtl (v.1.30; Broman et al. 2003) for all association analyses. We analyzed each trait using four mapping methods: marker regression (MR; Kearsey and Hyne 1994), expectation maximization (EM; Xu and Hu 2010), Haley-Knott (HK; Haley and Knott 1992) and non-parametric (NP; Kruglyak and Lander 1995) as described in Gross et al. (2014). With this approach, the limit for significant linkages was set at a LOD score of ≥ 4.0 – this is a threshold consistent with other similar QTL studies (Protas et al. 2007; Gross et al. 2009). To further confirm associations, permutation tests of 1000 iterations were performed to identify statistically significant QTL ($P < 0.05$). Moreover, the effect plots for associations were generated using the closest linked genetic marker. QTL regions were then anchored (~6-8 cM on each side of the top marker) using the NCBI BLAST Toolkit (v.2.28+) to the current *Astyanax* draft genome assembly (Ensembl build v.75; McGaugh et al. 2014). However, the *Astyanax* genome assembly is in its earlier stages (i.e., ~10,000 scaffolds rather than chromosomes), so we also determined the syntenic interval in the model fish *Danio rerio* – a system that has previously demonstrated large amounts of synteny with *Astyanax* (Gross et al. 2008; O'Quin et al. 2013; Carlson et al. 2015) – as a means to further support our candidate genes based on genomic position. Visual representations of synteny between our linkage map and the genomes were created with Circos (v.0.64; Krzywinski et al. 2009).

Gene Ontology (GO) term analysis

To nominate prospective candidate genes governing pigment cell variation, we interrogated all the genes present within the syntenic interval in the *Astyanax* draft genome for

gene ontology terms. The results of our BLAST search of genomic markers (described above) for a QTL determined the critical region of interest, yielding a set of *Astyanax* scaffolds as the syntenic region. From these comparisons, we collected the Gene Ontology (GO) terms for all genes located on the *Astyanax* genome scaffolds present within the syntenic interval using BioMart (v.0.8; Kasprzyk 2011). This approach yielded hundreds of prospective GO terms for each significant QTL. Within these datasets, we narrowed our search to identify genes terms with potential involvement in pigmentation-related processes by screening for terms such as: “pigment”, “pigmentation”, “melanin”, “eumelanin”, “phaeomelanin”, “melanophore”, “melanocyte”, “melanosome”, “xanthrophore”, “iridophore”, “chromatophore”, and “carotene”. This approach enabled the discovery of several genes that may be related to melanophore variation based on annotation information derived from known functions of the homologous genes in other systems.

RNA-seq, qualitative, and qPCR expression analyses

We also evaluated every gene located within the critical genomic regions for expression differences with RNA-seq analyses. For this, we isolated total RNA from either 50 surface or cave pooled individuals using the RNeasy Plus Mini Kit (Qiagen; Cat. No. 74134) at five developmental stages including: 10 hours post-fertilization (hpf), 24 hpf, 36 hpf, and 72 hpf, and from 3 individuals during juvenile hood (~4 months). Library preparation (TruSeq v.2 kit) and Illumina 2500 Hi-Seq for was performed either in triplicate (10 hpf – 72 hpf) or duplicate (juvenile) at the DNA Sequencing and Genotyping Core (Cincinnati Children’s Hospital and Medical Center). Samples were sequenced for a depth of 10 million reads each with 50-bp, single-end reads, totaling 28 separate sequencing runs. Normalized gene expression differences

were calculated with the software program ArrayStar (DNAStar.v.11.2a) using the default settings designed for analyzing “Illumina” based sequencing data. Moreover, this program quantifies expression through comparative read counts between cave and surface fish using the RPKM normalization method (Mortazavi et al. 2008) and detects significant differences between samples using a Student’s t-test with FDR (DNAStar.v.11.2a; Benjamini and Hochberg 2000). Raw sequencing reads have been deposited to the NCBI SRA under the BioProject: PRJNA258661.

To confirm the differences detected with RNA-seq, we validated the expression profiles at the 72 hpf stage using both qualitative and quantitative PCR analyses. Template cDNA from surface and cavefish RNA pools (n= 50 embryos each; RNeasy Plus Mini kit, Qiagen) was synthesized for both experiments using the Transcriptor RT kit (Roche). For this, 1 µg of RNA was hybridized to Oligo dT primers (Invitrogen) at 65 °C for 10 min, and cooled on ice. We added 4.0 µl 5x RT buffer, 0.5 µl Protector RNase Inhibitor, 2.0 µl dNTP mixture and 0.5 µl Transcriptor RT to each reaction and incubated at 50 °C for 1 hr, then inactivated at 85 °C for 5 min. Qualitative PCR experiments amplified fragments for the reference gene *Rps18* (forward primer: 5'- AGATCTGCGGGGTAGAGGAT-3', reverse primer: 5'- TGTCTGAGACCACGATGAGC-3'; 421 bp), *Tyrp1b* (forward primer: 5'- GTGGGGAAGACGTACATGGG -3', reverse primer: 5'- TGTGTCCAAGTCCAAGCAGT -3'; 441 bp), and *Pmela* (forward primer: 5'- CGTCCTTCCAGAAGCTCCAG-3', reverse primer: 5'- TCATTTGCAGCAGCCTCTGT-3'; 417 bp). Amplicons were visualized after gel electrophoresis on a 2% agarose gel. Additionally, quantitative PCR (qPCR) experiments were performed according to previously described methodology (Gross and Wilkens 2013). Fragments ~100 bp in length were amplified for the target genes *Tyrp1b* (forward primer: 5'-

TAGCGATTGTTGTGGTGGTG-3', reverse primer: 5'- TTCCAGGCCTTCGACTGAAC-3'; 92 bp) and *Pmela* (forward primer: 5'- TTGCCACAAATGCTCCAGTC-3', reverse primer: 5'- TGCAGCAGCCTCTGTATTTG-3'; 119 bp), and normalized to the reference gene *Rps18* (forward primer: 5'- ACACGAACATCGATGGTAGGAG-3', reverse primer: 5'- TTGTTGAGGTCGATGTCTGC-3'; 112 bp). Experiments were carried out with the Mini Opticon light-cycler (BioRad) using Evagreen supermix dye (BioRad) with the following cycling parameters: step 1—95 °C for 30 s, step 2—95 °C for 5 s, step 3—55.1 °C for 10 s, plate read, go to step 2 for 39 additional cycles. All samples were analyzed in sextuplet, and normalized expression values (C_q) and significant differences (two-tailed Student's t-test) were determined using the CFX Manager software program (v.3.1; BioRad).

Structural gene analyses

Genes residing within the syntenic region in the *Astyanax* genome were also analyzed for prospective sequence alterations. Sequencing reads derived from surface and Pachón cavefish (~280 million reads total) were aligned to the complete set of cDNA transcripts from the current draft *Astyanax* genome (Ensembl.v.75; McGaugh et al. 2014) using default parameters for Illumina 50-bp reads with the program SeqMan NGen (DNASTAR.v.11.2a). Utilizing the SNP report tool, we identified different types of mutations (SNP, Indel) segregating between surface and cave morphotypes (100% fixed differences), and further, noted the predicted effects on the gene structure (i.e., synonymous, nonsynonymous, frameshift, premature stop).

Expression analysis with whole mount in situ hybridization

The RNA probes for *in situ* hybridization generated by PCR amplifying the genes *Tyrp1b* (forward primer: 5'- GAAACAGCCCTCAGTTCGAG-3', reverse primer: 5'- AGGTGGGCCAGATTGTGTAG-3') and *Pmela* (forward primer: 5'- CTA CTGATGCTGCCACTGGA-3', reverse primer: 5'- AGAGCCGTAGCGGTAGATCA-3') previously subcloned into pGem- T Easy vector (Promega) using M13F and M13R primers. Gel purified DNA template was used to generate RNA probes using the DIG RNA Labeling Kit SP6/T7 (Roche, Cat. No. 11175025910) according to the manufacturer's protocol. The reaction mix included: 1 µg DNA, 2.0 µl 10x NTP labeling mix, 2.0 µl 10x transcription buffer, 1.0 µl protector RNase inhibitor, 2.0 µl SP6 or T7 RNA polymerase, and RNase-free water to a total volume of 13 µl, and was incubated at 37 °C for 2 hours. DNase was then added to the reaction and incubated for 15 min at 37 °C. To precipitate the RNA, 0.8 µl 0.5M EDTA, 2.5 µl 4.0M LiCl, and 75.0 µl of 100% chilled EtOH was added and incubated at -20 °C for 1 hr to overnight. The RNA was pelleted via centrifugation, washed with 70% EtOH, dried, and re-suspended in 50 µl RNase-free water. To confirm proper synthesis of probe, 3.0 µl of RNA was combined with 2.0 µl of water and 5.0 µl formamide, heated to 72 °C for 5 min, and then visualized after electrophoresis on a 2% agarose gel.

Whole mount *in situ* hybridizations were carried out using methodology as described in (Gross et al. 2011). Surface and cave individuals stage-matched at 72 hpf were fixed in MEMFA solution and de-hydrated in MeOH to 100% until use. Specimen were then bleached in a 6% H₂O₂/MeOH solution for ~12 hours, then re-hydrated in series (75%-25%) of MeOH/PBT rinses, washed in PBT (3x – 5 min), then digested with Proteinase K (10 mg/ml; Roche) at a dilution of 1:1000 for 15 min at room temperature. After digestion and additional PBT rinses (3x), samples

were fixed in a 4%PFA/0.2% Gluteraldehyde mixture for 20 min. Following additional washes (3x – 5 min), samples were incubated in a pre-warmed hybridization solution at 70 °C for 2 hours. Fresh hybridization solution was added to each vial, along with 10 µl of RNA probe, and then incubated overnight at 70 °C in a hybridization oven.

After overnight hybridization, the specimens were rinsed, incubated in Solution I (50% formamide/2× SSC [pH 4.5]/1% SDS; 4x 30 min) at 70 °C, and rinsed in a 50/50 SolutionI/MABT mixture for 5 min at room temperature. Embryos were rinsed (3x) in MABT (100 mM Maleic acid; 150 mM NaCl; 0.1% Tween-20; pH 7.5) and incubated in MABT for 30 min (2x). The vial solution was then replaced with a mixture of 20% heat-inactivated normal goat serum (HINGGS) + 2% blocking reagent (BR; Roche) in MABT for 2 hours. The 20%/HINGGS/2%BR/MABT was then replaced with new solution, anti-DIG-AP (1:2000 dilution; Roche) is added, and incubated overnight at 4 °C. Following the antibody incubation, the third day entails MABT rinses (3x) and washes (6x – 1 hr) at room temperature, then incubated overnight in MABT solution at 4 °C.

On the final day, specimens were rinsed (4x – 10 min) in NTM solution (100 mMNaCl; 100 mM Tris [pH 9.5]; 50 mM MgCl₂). Then, 1 ml of fresh NTM is added to each vial, combined with 4.5 µl NBT and 7.0 µl BCIP, and developed in the dark for ~2 hours (varies per gene) until the signal is visible. Post-development, embryos were washed in NTM (3x – 20 min), PBT (2x – 10 min), and TBST (TBS + 0.1% Tween-20; 3x – 20 min). Finally, specimens were fixed in 4% PFA/PBS for storage at 4 °C and visualized using Leica Microscope M205 FA (with LAS software v.3.8.0) montage imaging.

Functional validation in zebrafish with MO-knockdowns

A morpholino microinjection technique was used to characterize the precise pigmentation phenotype associated with each candidate gene in morphant individuals. Morpholino oligonucleotides were targeted to the first 25 base pairs of the zebrafish forms of *Pmela* (5'-GAGGAAGATGAGAGATGTCCACAT-3') and *Tyrp1b* (5'-GCACTAAACACACACTCTTCCACAT-3') (Gene Tools, LLC.). Morpholinos were injected into single-celled zebrafish embryos in a 1 nl volume at a 0.2 mM concentration – these oligonucleotides bind mRNA transcripts, and ultimately, prevent normal protein translation *in vivo*. Control individuals were administered a mock injection composed of phenol red and Danieaux's solution – protocol adapted from (Wingert et al. 2004). The qualitative “reduced pigmentation” phenotype for morphant individuals was determined after 48 hpf, when pigmentation normally appears in wild type zebrafish, yet remains within the timeframe for the morpholinos to maintain knockdown effects (~7 days post-injection). Imaging was performed using a Leica Microscope M205 FA (with LAS software v.3.8.0) stereoscope.

RESULTS

Cave x surface fish hybrid individuals demonstrate substantial diversity in melanophore numbers in distinct regions around the body.

Prior studies in *Astyanax* have identified the genetic basis for two pigmentation traits, namely *Oca2* (albinism; Protas et al. 2005) and *Mc1r* (*brown*; Gross et al. 2009), yet the genes governing complex coloration diversity have yet not been uncovered. However, direct observations demonstrate the wide array of color variation among F₂ siblings (Fig. 1A-E). Moreover, a prior report noted substantial pigment cell variation when assayed in four different body regions (Protas et al. 2007). Similarly, when we quantified melanophore number in seven areas our results revealed variation across our F₂ population (n= 170; Fig. 2A). Specifically, we discovered counts ranged uniquely with each region (measurements exclude individuals with albinism which do not have any melanophores): 1 – 99 melanophores in the square near the anal fin (MelAnalfinSquare; Fig. 2B), 1 – 113 pigment cells below the mid-lateral stripe (MelUnderStripe; Fig. 2C), 4 – 600 cells in the dorsal square (MelDorsalSquare; Fig. 2D), 1 – 2,762 melanophores in the area above the stripe (MelAboveStripe; Fig. 2E), 1 – 89 pigment cells above the eye (MelHeadSquare; Fig. 2F), 11 – 1,395 melanophores present in the full head region (MelHead; Fig. 2G), and 0 – 252 pigment cells by the anal fin (MelAnalfinTriangle; not pictured). Interestingly, melanophore counts did not simply increase or decrease proportionately across all regions of the body, instead pigment cell number in some areas seemed to vary in number independent of the other regions assayed. For instance, 14 F₂ individuals that harbored numerous melanophores in some areas (i.e., dorsal region or head) did not have any pigment cells (n= 0) present near the anal fin. We noted a similar finding for melanophore variation under

the mid-lateral stripe – again, over 20 individuals demonstrated pigmentation elsewhere on the body, but none in the area scored below the stripe.

QTL analysis revealed 19 genomic regions associated with complex melanophore variation

For this study, we scored melanophore variation in seven regions spanning the body in our F₂ pedigree to determine the genetic loci contributing to the complex pigmentation phenotypes observed in *Astyanax* cavefish. Of the regions analyzed, we detected a numerous QTL for six of the areas, including MelAnalfinSquare, MelUnderStripe, MelDorsalSquare, MelAboveStripe, MelHeadSquare, and MelAnalfinTriangle. We did not discover any QTL for melanophores scored for the full head (MelHead). This was surprising considering we detected associations for pigment cell counts in the head square. However, the head craniofacial shape, size and curvature varies dramatically in the F₂ hybrids (Gross et al. 2014), and it is possible that these differences mask our trait, or simply, the complex skull architecture makes it difficult to accurately and consistently score.

Additionally using four different mapping methods, our results yielded multiple significant (independent) associations (n= 41 markers) with the linkage map for the following pigment cell number traits: MelAnalfinSquare (n= 1 QTL), MelUnderStripe (n= 13 QTL), MelDorsalSquare (n= 3 QTL), MelAboveStripe (n= 1 QTL), MelHeadSquare (n= 3 QTL), and MelAnalfinTriangle (n= 6 QTL; see Table 1 for LOD scores and corresponding p-values). In some instances, the different mapping methods identified significant associations with exact same top marker (e.g., MR, EM, HK, and NP methods all noted the marker ASTYANAX_414 for MelUnderStripe) and in other cases, the different approaches yielded multiple significant

markers within a small positional window – for example, 3 markers for each MR, EM, and HK respective method for MelDorsalSquare on LG 20 (53 – 61 cM). Not surprisingly, some of these QTL co-localized near the same position in our linkage map. One “hotspot” in our map was on linkage group 1 from 75 – 93 cM, wherein associations with 5 traits (MelHeadSquare, MelUnderStripe, MelAnalfinTriangle, MelDorsalSquare, MelAnalfinSquare) yielded significant QTL. Moreover, 3 different regions assayed (MelDorsalSquare, MelAnalfinTriangle, MelAboveStripe) returned associations with linkage group (LG) 20 (53 – 61 cM), and 2 traits (MelAnalfinTriangle, MelUnderStripe) tended to co-localize together at several positions: LG 2 (29.8 – 33.6 cM), LG 7 (54.64 – 63.7 cM) and LG 26 (8.98 – 18 cM). Therefore in total, we discovered associations with 19 unique regions of the genome for melanophore variation.

Table 1. QTL association mapping identifies numerous significant associations for variation in melanophore number.

Trait	Marker	MR - LOD	EM - LOD	HK - LOD	NP - LOD	0.05 (LOD Threshold)	P-value	LG	Position (cM)
MelHeadSquare	ASTYANAX_681	5.11	--	--	--	4.27	0.008	1	75
MelUnderStripe	ASTYANAX_681	--	--	--	10.09	4.31	0	1	75
MelAnalTri	TP37468	5.47	--	--	--	4.81	0.012	1	82.5
MelDorsalSqu	TP43910	6.35	--	--	--	5.75	0.025	1	86.2
MelAnalSqu	TP52439	--	--	--	4.38	4.24	0.032	1	88.7
MelUnderStripe	TP56998	10.77	--	--	--	4.96	0	1	92.6
MelAnalTri	TP45463	--	5.04	5.08	6.56	4.27	0	1	93.1
MelUnderStripe	TP45463	--	11.67	12.04	--	4.85	0	1	93.1
MelUnderStripe	TP84626	8.53	--	--	4.33	4.31	0.045	2	29.8
MelAnalTri	TP77512	5	--	--	--	4.81	0.034	2	33.6
MelUnderStripe	TP91095	--	6	5.93	5.4	4.71	0.008	3	54
MelUnderStripe	TP63580	5.89	--	--	--	4.96	0.008	3	56.7
MelAnalTri	TP81818	--	--	--	4.29	4.27	0.046	7	54.64
MelUnderStripe	TP75360	5.91	--	--	--	4.96	0.007	7	59
MelUnderStripe	TP21243	--	--	4.8	--	4.71	0.04	7	63.7
MelUnderStripe	TP21243	--	--	--	4.7	4.31	0.025	7	63.7
MelUnderStripe	ASTYANAX_302	--	5.47	4.86	--	4.71	0.037	8	15.4
MelUnderStripe	TP36338	9.57	5.83	6.11	5.25	4.96	0	9	21.7
MelHeadSquare	ASTYANAX_580	--	--	4.79	--	4.75	0.048	10	138
MelUnderStripe	TP42078	4.98	--	--	--	4.96	0.049	11	134.8
MelUnderStripe	TP57531	5.29	--	--	--	4.96	0.023	12	53.8
MelAnalTri	TP27087	--	4.65	--	--	4.61	0.044	12	120.7
MelUnderStripe	TP30103	5.95	--	--	--	4.96	0.007	13	41.8
MelHeadSquare	ASTYANAX_705	4.53	--	--	--	4.27	0.03	14	69.8
MelUnderStripe	ASTYANAX_414	6.16	9.25	8.27	7.43	4.96	0.001	16	22.2
MelDorsalSqu	TP16875	8.06	13.8	--	--	5.75	0.003	19	-3.6
MelDorsalSqu	ASTYANAX_28	6.01	--	--	--	5.08	0.015	20	53
MelAnalTri	TP37042	--	--	--	6	4.27	0	20	53.94
MelAnalTri	TP23554	--	5.8	5.63	--	4.61	0.008	20	55
MelDorsalSqu	ASTYANAX_348	--	16.3	--	--	13.7	0.003	20	58
MelAboveStripe	ASTYANAX_348	--	21.7	--	--	17.8	0	20	58
MelAboveStripe	ASTYANAX_119	5.02	--	--	--	4.76	0.033	20	58.9
MelDorsalSqu	TP4784	--	--	6.41	--	5.62	0.021	20	61
MelUnderStripe	TP35845	7.05	6.51	6.37	--	4.96	0	20	99.4
MelUnderStripe	TP35845	--	--	--	5.61	4.31	0.001	20	99.4
MelUnderStripe	ASTYANAX_893	5.34	--	--	--	4.96	0.02	23	37.8
MelUnderStripe	TP67573	--	--	--	4.77	4.31	0.022	23	118.9
MelUnderStripe	TP67573	--	6.27	5.54	--	4.85	0.006	23	119
MelAnalTri	TP66518	--	--	--	4.32	4.27	0.038	26	8.98
MelUnderStripe	TP33309	5.08	--	--	--	4.96	0.041	26	18

Comparative analyses narrowed melanophore QTL positions to critical genomic regions in Astyanax and Danio

Although our linkage map is densely populated with over 3,000 markers, the vast majority of genes (>23,000) are not represented. Therefore, as a means to further our search for prospective genes residing near our top markers, we analyzed the positions of each QTL in the *Astyanax* cavefish draft genome (~10,000 scaffolds) and the current zebrafish genome. For this, we first queried marker sequences to the *Astyanax* genome (Ensembl.v.75) using the standard nucleotide BLAST algorithm to find the position of the top genomic marker, and the locations of the markers within the QTL interval (immediately adjacent ~6 – 8 cM to the top marker), to identify regions of synteny. Since these scaffolds varied in length (876 bp – 9,823,298 bp) the number of scaffolds reported back from our BLAST search ranged from 4 to 27 scaffolds for each association (Table 2).

However, since the current version of the *Astyanax* genome remains in draft form (~10,000 scaffolds rather than 25 consolidated chromosomes), we capitalized on the well-annotated genome of the model teleost system *Danio rerio* (Ensembl.Zv9). Three comparative studies have previously identified large syntenic regions shared between *Astyanax* and *Danio* (Gross et al. 2008; O'Quin et al. 2013; Carlson et al. 2015). These reports utilized similar approaches to identify marker position (SNP- or microsatellite-based) through sequence similarity. Moreover, zebrafish also harbor the same karyotypic number as *Astyanax* (n = 25 chromosomes) and diverged ~120 MYa (Gross 2012b). The results from our genomic comparisons revealed several syntenic blocks based on the positions of the BLAST hits in *Danio*, yielding regions of synteny on different zebrafish chromosomes (Table 2). We also sought to further support the syntenic block from our direct comparisons of marker sequences to

zebrafish, since usually < 20 markers yielded direct BLAST hits for each QTL. For this, we collected every predicted gene (often 100+ genes) from the respective *Astyanax* scaffolds associated with each QTL (described above), and queried all the full-length gene sequences to the zebrafish genome. These analyses proved immensely helpful in narrowing our search to specific regions of chromosomes in *Danio* by supporting putative syntenic blocks with hundreds of additional BLAST hits, and ultimately, reinforce the proximity of QTL markers to candidate genes based upon positional information.

Table 2. QTL regions of synteny reveal candidate genes based upon gene ontology of genes within syntenic interval.

Melanophore QTL			<i>Astyanax mexicanus</i>			<i>Danio rerio</i>			
LG	cM	Trait	Markers	Syntenic scaffolds (ID)	GO terms	Candidate genes	Syntenic regions (Chr:bp)	GO terms	Candidate genes
1	86.2	MelAnalTri MelDorsalSqu MelHeadSquare MelUnderStripe MelAnalSqu	TP45463 TP37468 TP43910 ASTYANAX_681 TP56998 ASTYANAX_68 TP52439	KB882299.1; KB882095.1; KB882299.1; KB882110.1; KB882299.1; KB871938.1; KB872295.1; KB882097.1; KB871594.1; KB872325.1; KB872362.1; KB882270.1; KB871831.1; KB882087.1; KB882097.1; KB871688.1; KB871831.1; KB872362.1; KB882213.1; KB882097.1; KB882160.1; KB872401.1 KB871761.1; KB871784.1; KB871788.1; KB871846.1; KB871993.1; KB872187.1; KB879810.1; KB882119.1; KB882130.1; KB882143.1; KB882198.1; KB882231.1; KB882270.1; KB882285.1; KB882310.1	melanocyte migration; late stripe melanocyte differentiation	<i>igsf11</i>	3:33154462-63835661 22:18153538-20153538	pigmentation; pigment granular localization; melanocyte differentiation; iridophore differentiation	<i>aldoa</i> <i>nsfa</i> <i>sox9b</i>
2	33.6	MelAnalTri MelUnderStripe	TP77512 TP84626	KB871628.1; KB871633.1; KB871887.1; KB872421.1; KB882103.1; KB882155.1; KB882237.1; KB882308.1 KB871655.1; KB871688.1; KB871698.1; KB871753.1; KB871785.1; KB871795.1; KB871797.1; KB871875.1; KB871900.1; KB871956.1; KB871970.1; KB872033.1; KB872295.1; KB882091.1; KB882097.1; KB882159.1; KB882160.1; KB882191.1; KB882216.1; KB882223.1; KB882249.1; KB882292.1	retinal pigment epithelium development; eye pigment granule organization	<i>atp6v0d1</i>	4:14683465-16683465	melanin-concentrating hormone activity	<i>pmch</i>
3	54	MelUnderStripe	TP63580 TP91095	KB871655.1; KB871688.1; KB871698.1; KB871753.1; KB871785.1; KB871795.1; KB871797.1; KB871875.1; KB871900.1; KB871956.1; KB871970.1; KB872033.1; KB872295.1; KB882091.1; KB882097.1; KB882159.1; KB882160.1; KB882191.1; KB882216.1; KB882223.1; KB882249.1; KB882292.1	beta-carotene monooxygenase activity; pigmentation	<i>bcmo1</i> <i>ctf9</i>	2:29618901-31618901 7:57352183-67982905 11:9612165-47461667 20:35870397-37870397 22:40377617-42377617 23:43356963-45356963	pigmentation	<i>ctf9</i>
7	59	MelAnalTri MelUnderStripe	TP81818 TP75360 TP21243	KB871684.1; KB872147.1; KB871839.1; KB871688.1; KB871836.1; KB882211.1; KB882105.1; KB871913.1; KB882183.1; KB871974.1; KB871607.1	developmental pigmentation; retinal pigment epithelium development; pigment granule aggregation in cell center	<i>dio2</i> <i>otx2</i> <i>ttc8</i>	8:14794876-54750014	melanocyte differentiation; melanocortin receptor activity; iridophore differentiation	<i>foxd2</i> <i>mc3r</i> <i>foxd5</i> <i>gfpt1</i>
8	15.4	MelUnderStripe	ASTYANAX_302	KB871655.1; KB871726.1; KB871991.1; KB872310.1; KB882117.1; KB882150.1; KB882158.1; KB882233.1	developmental pigmentation; retinal pigment epithelium development; pigment granule aggregation in cell center	<i>dio2</i> <i>otx2</i> <i>ttc8</i>	2:23713686-25713686 10:18549105-20549105 17:3004630-52756133 20:1-1874890 22:9388437-11388437	melanosome localization; regulation of melanocyte differentiation; developmental pigmentation; retinal pigment epithelium development; melanosome transport	<i>pornc</i> <i>ctbp2a</i> <i>dio2</i> <i>itk</i> <i>vps39</i> <i>otx1b</i> <i>otx2</i> <i>ippk</i>
9	21.7	MelUnderStripe	TP36338	KB871655.1; KB871726.1; KB871991.1; KB872310.1; KB882117.1; KB882150.1; KB882158.1; KB882233.1	developmental pigmentation; retinal pigment epithelium development; pigment granule aggregation in cell center	<i>dio2</i> <i>otx2</i> <i>ttc8</i>	2:23713686-25713686 10:18549105-20549105 17:3004630-52756133 20:1-1874890 22:9388437-11388437	melanosome localization; regulation of melanocyte differentiation; developmental pigmentation; retinal pigment epithelium development; melanosome transport	<i>pornc</i> <i>ctbp2a</i> <i>dio2</i> <i>itk</i> <i>vps39</i> <i>otx1b</i> <i>otx2</i> <i>ippk</i>

Table 2. (continued).

Melanophore QTL			<i>Astyanax mexicanus</i>			<i>Danio rerio</i>			
LG	cM	Trait	Markers	Syntenic scaffolds (ID)	GO terms	Candidate genes	Syntenic regions (Chr:bp)	GO terms	Candidate genes
19	-3.6	MelDorsalSqu	TP16875	KB871897.1; KB872777.1; KB873826.1; KB874764.1; KB882149.1; KB882219.1	--	--	13:40427566-42427566	none	--
20	58.9	MelAboveStripe MelDorsalSqu	TP4784 ASTYANAX_28 ASTYANAX_348 ASTYANAX_119	KB871605.1; KB871630.1; KB871854.1; KB872089.1; KB872138.1; KB872157.1; KB872462.1; KB873883.1; KB874320.1; KB878200.1	melanosome membrane melanin biosynthetic process	tyrp1b <i>tyrp1a</i>	1:10192286-12192286 1:11215779-20013359	pigmentation	tyrp1b
20	99.4	MelUnderStripe	TP35845	KB871605.1; KB882162.1	--	--	No direct marker hits within QTL	--	--
23	119	MelUnderStripe	TP67573	KB871711.1; KB871739.1; KB871920.1; KB882119.1	pigmentation	<i>adrb2a</i>	No direct marker hits within QTL	--	--
26	8.98	MelAnalTri	TP66518	KB871581.1; KB871584.1; KB871590.1; KB871614.1; KB871822.1; KB871987.1; KB876853.1; KB882111.1; KB882116.1; KB882145.1; KB882189.1; KB882218.1	eye pigmentation; developmental pigmentation; photoreceptor activity	<i>pmela</i> <i>lppk</i>	11:19150115-26754101	none	--
26	18	MelUnderStripe	TP33309	KB871581.1; KB871584.1; KB871590.1; KB871614.1; KB871641.1; KB871822.1; KB876853.1; KB877205.1; KB880823.1; KB882111.1; KB882116.1; KB882145.1	melanosome transport; developmental pigmentation; eye pigmentation	<i>lppk</i> <i>pmela</i>	9:24225715-26225715 11:19150115-44680161 11:21927-5980431	developmental pigmentation; eye pigmentation	<i>pmela</i>

Integrative analyses nominate Tyrp1b and Pmela as two candidate genes influencing melanophore trait diversity.

We developed an integrative approach to focus our attention on strong candidate genes by co-analyzing multiple pieces of evidence: 1) the effect plot for the top marker reveals a logical allelic-pattern for melanophore number (QTL association studies), 2) the top marker maps near the gene of interest (comparative genomics), 3) candidates demonstrate previously known roles in pigmentation development or maintenance based upon gene (GO) ontology terms, 4) alterations segregating between morphs are predicted to affect the coding sequence in cavefish (structural analyses), and 5) differential gene expression patterns are observed during early development (RNA-seq profiling).

As described earlier, nearly all of the assayed pigment cell traits yielded multiple QTL associated with multiple distinct regions of the linkage map (Fig. 3A). Among these traits, we characterized the wide range of melanophore number variation observed in our pedigree, including the regions scored in the dorsal square (Fig. 3B) and under the mid-lateral stripe (Fig. 3D). Our association studies revealed multiple QTL for these two traits, including significantly associated markers ASTYANAX_28 on LG 20 ($\text{LOD}^{\text{MR}} = 6.01$, $p = 0.015$) and TP33309 on LG 26 ($\text{LOD}^{\text{MR}} = 5.08$, $p = 0.041$), respectively. The corresponding effect plots demonstrated a significant effect of genotype for each marker, wherein individuals homozygous for surface alleles “S/S” harbor more melanophores and those homozygous for cave alleles “C/C” have fewer pigment cells (Fig. 3C,E). The heterozygous genotype for ASTYANAX_28 presented an intermediate level of pigment cells at this locus, whereas the heterozygous form for TP33309 appeared to have melanophore numbers similar to the homozygous cave genotype, possibly indicating dominance of the cave-allele in hybrid individuals.

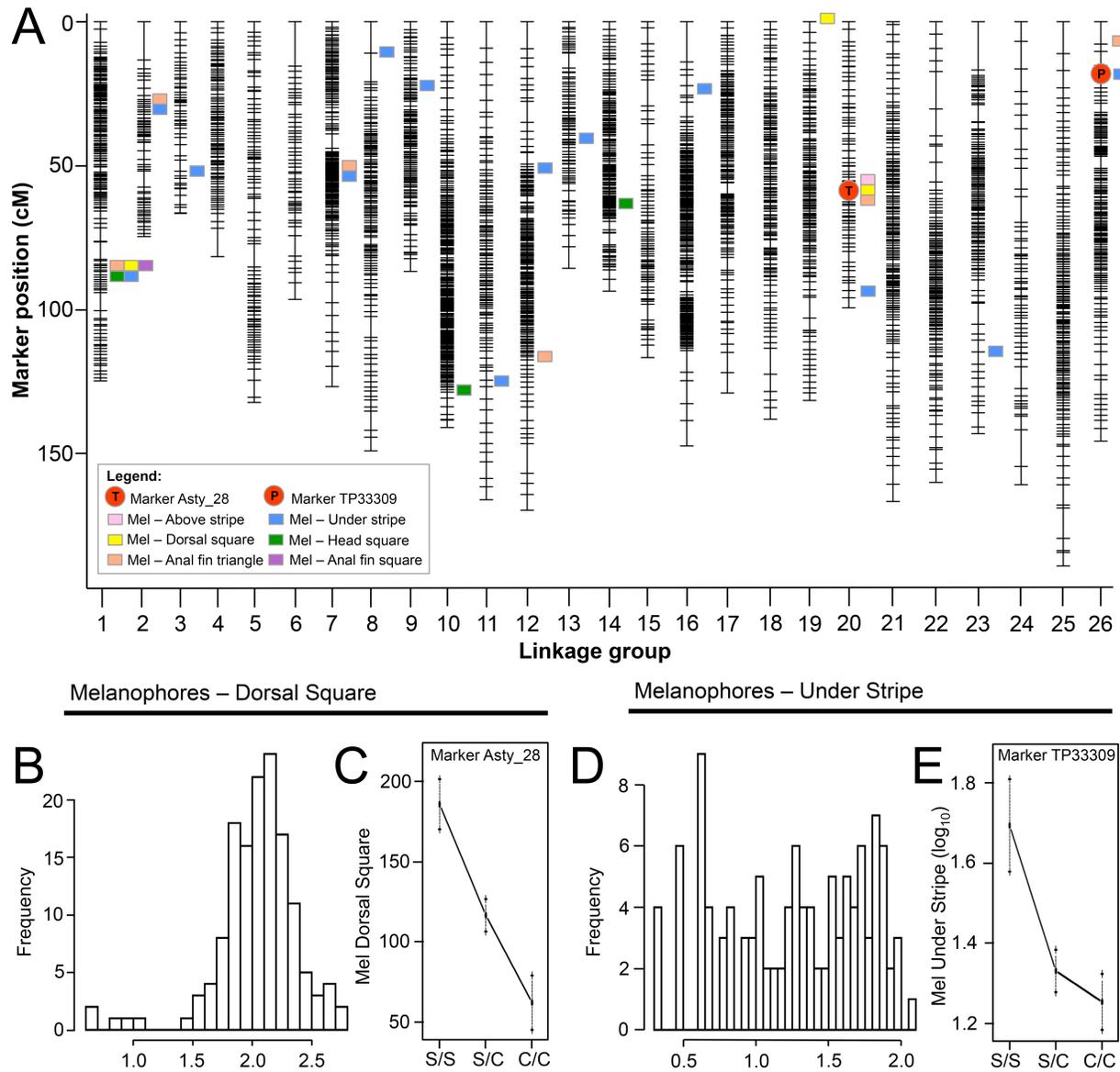


Figure 3. High-resolution mapping identified multiple QTL associations with melanophore number variation. Our quantitative trait locus (QTL) study detected numerous associations ($n=40$) with pigment cell variation (A). Many of these co-localized to similar positions on our linkage map. With this approach, we were able to capture the dramatic variation of pigment cell number, including the wide range of melanophores detected in the dorsal square (B) and under the mid-lateral stripe (D). The markers underlying these two QTL demonstrated a significant

effect of phenotype, where individuals harboring two copies of the surface allele “S/S” demonstrated increased numbers of melanophores, and hybrids with two cavefish alleles “C/C” exhibited fewer pigment cells. This effect is evident in both the dorsal (C) and stripe (E) regions assayed.

To pursue prospective candidate genes involved with pigmentation-related processes, we further explored the syntenic regions that resulted from our genomic comparisons between *Astyanax* and *Danio* using gene ontology (GO) term analyses. For this, we collected the available ontology information from all genes within the syntenic blocks. For the QTL association on LG 20, we identified 951 terms affiliated with 157 genes on 10 *Astyanax* scaffolds, and 868 terms associated with 193 genes from the interval chromosome 1 in zebrafish (see Table 2 for scaffold IDs and *Danio* genomic intervals). Within this dataset, we identified two GO terms related to pigmentation called “melanosome membrane” (GO:0033162) and “melanin biosynthetic process” (GO:0042438) of which both terms are associated with two genes *Tyrp1b* and *Tyrp1a* in *Astyanax*, and the terms “pigmentation” (GO:0043473) with the gene *Tyrp1b* within syntenic block in zebrafish. We note that the *Tyrp1a* paralog resides on a different chromosome (Chr7) in *Danio rerio*. Additionally, we discovered 1,136 GO terms for 409 genes associated with 12 scaffolds for ‘MelUnderStripe’ on LG 26, and 3,017 terms for the 654 genes within the critical block on chromosome 11 in *Danio rerio*. This analysis revealed three terms related to pigmentation in *Astyanax* including “melanosome transport” (GO:0032402) linked to *Ippk* and “developmental pigmentation” (GO:0048066) and “eye pigmentation” (GO:0048069) assigned to *Pmela*. Within the region in *Danio*, we identified the

terms “developmental pigmentation” (GO:0048066) and “eye pigmentation” (GO:0048069) associated with *Pmela*. By co-analyzing the gene ontology and positional information from both systems, we nominated candidate genes for each QTL: *Tyrp1b* (LG20) and *Pmela* (LG26). Moreover, these two genes reside within close proximity of the top two QTL markers, ~7 MB and ~0.05 MB, respectively (Fig. 4A, D).

After identification of these two candidate genes, we interrogated the coding sequences in cavefish. We aligned our surface- and cave-tagged RNA-sequencing reads to the *Astyanax* draft genome, inclusive of our candidate loci. The gene *Tyrp1b*, near the marker where multiple melanophore traits mapped, harbored three mutations. One SNP was a G-to-A substitution in exon 2 at position 630 (Fig. 4B). This alteration is predicted to be a synonymous change that does not appear to affect the protein sequence (F210F). However, this alteration does fall within a documented tyrosinase copper-binding domain (205 – 222 bp; Ensembl.v.75) which could affect interactions with tyrosinase – an enzyme that catalyzes reactions in the melanin biosynthesis pathway (Oetting 2000) – as several synonymous mutations have been shown to affect splicing, stability, structure and protein folding (Hunt et al. 2009). Moreover, our analysis identified two putative splice variants for *Tyrp1b*. We also characterized the sequence differences between surface and cavefish for *Pmela*. Interestingly, the gene sequence revealed numerous alterations, such as a G-to-A change at position 1738 in exon 7 (Fig. 4E). This non-synonymous mutation does alter the amino acid sequence, causing a change from a hydrophobic alanine residue to a hydrophilic threonine in cavefish (A580T). In addition, *Pmela* demonstrates several more mutations including two other non-synonymous changes, two silent SNPs, and three potential splice variants.

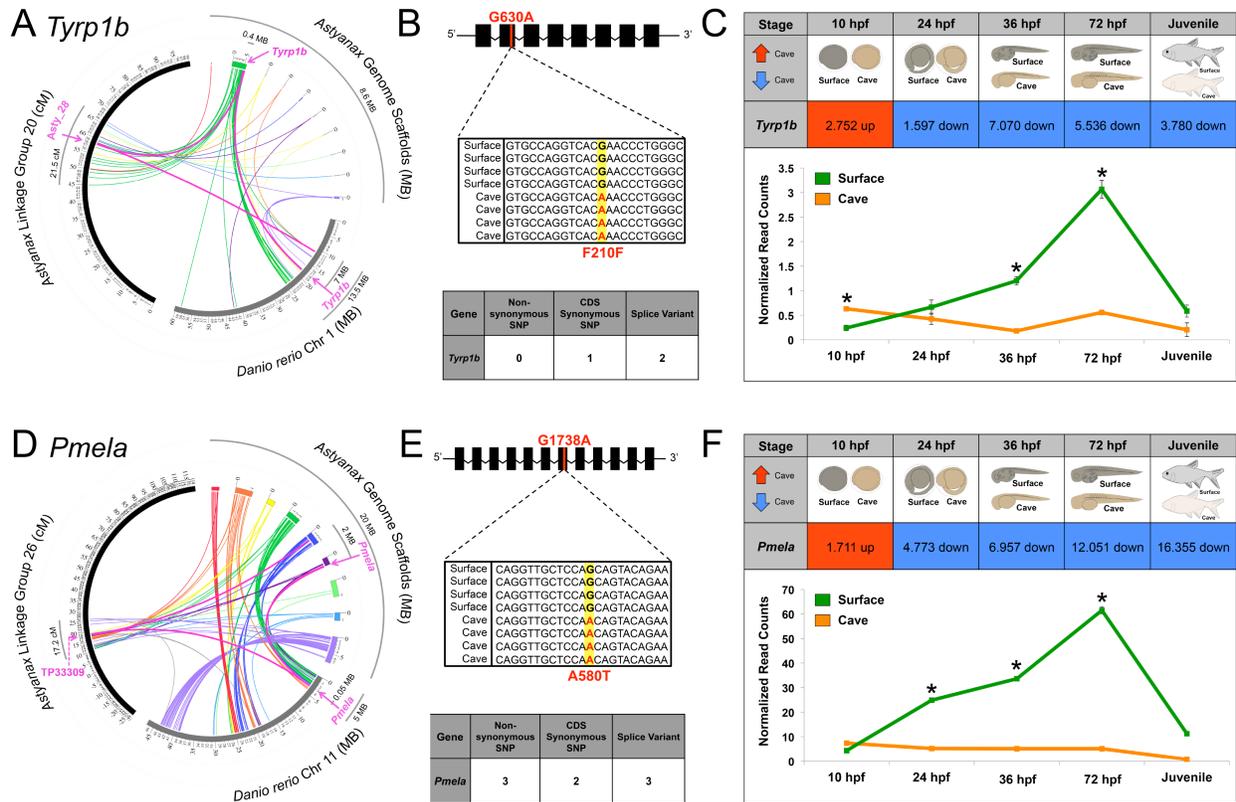


Figure 4. Integrative analyses reveal two candidate genes for melanophore variation. The evaluation of melanophore variation using QTL analyses identified two top candidate genes that may be associated with pigmentation loss in *Astyanax mexicanus* cavefish. Pigment cell variation under the mid-dorsal stripe region showed a significant association with marker ASTYANAX_28 on LG 20 ($\text{LOD}^{\text{MR}} = 6.01$, $p = 0.015$), which anchors to genomic regions in *Astyanax* and *Danio rerio* inclusive of *Tyrosinase-related protein 1b* (*Tyrp1b*; A). A second melanophore trait in the superior dorsal region demonstrates a significant association with marker TP33309 on linkage group 26 ($\text{LOD}^{\text{MR}} = 5.08$, $p = 0.041$; D). Regions of synteny were identified between *Astyanax* and *Danio rerio*. The syntenic region was mined for any genes with potential roles in the pigmentation pathway using Gene Ontology (GO) terms. This led to the discovery of the gene *Premelanosome protein a* (*Pmela*). These two candidates, and any other pigment-related genes that exist within the syntenic region, were evaluated for prospective

coding mutations and expression alterations using RNA-seq technologies (B-C, E-F). This combined approach has led to discovery of candidate genes that may contribute to loss of pigmentation in cave-dwelling fish.

We then aimed to discover dramatic differences in gene expression between surface- and cave-dwelling morphotypes during early development. Strikingly, *Tyrp1b* demonstrated reduced expression in cave relative to surface beginning at 24 hpf (Fig. 4C). During this time, pigmentation normally begins to appear in surface, but not in cave. We observed dynamic fold changes differences of 7-fold down (36 hpf) and 5.5-fold down (72 hpf; Fig. 4C). Further, expression profiles of cave vs. surface forms demonstrated significant differences at 10 hpf ($p=0.0472$), 36 hpf ($p=0.0312$), and 72 hpf ($p=0.00167$). The gene *Pmela* similarly revealed a severe reduction in expression at the onset of 24 hpf (Fig. 4F). The profile for *Pmela* demonstrated remarkable fold change differences such as 12-fold down at 72 hpf throughout embryonic development. Levels of gene expression were significantly different between morphs at 4 stages assayed: 10 hpf ($p=0.00607$), 24 hpf ($p=0.00117$), 36 hpf ($p=0.0311$), and 72 hpf ($p=0.000239$).

Since *Tyrp1b* and *Pmela* yielded significant differences in expression in cavefish relative to surface at 72 hpf using RNA-seq (Fig. 5A-B), we generated RNA pools at this developmental stage to validate transcript levels with two different approaches. We first carried out qualitative expression analyses using standard PCR. As anticipated, we observed slightly weaker bands in Pachón cavefish compared to surface for *Tyrp1b* and *Pmela* (gel images not shown). We further confirmed reduced expression of these genes in cavefish at 72 hpf with quantitative PCR

(qPCR). Again, our results demonstrated significant decreases in expression in Pachón cave-dwelling fish in *Tyrp1b* ($p = 0.005967$; Fig. 5C) and *Pmela* ($p < 0.000001$; Fig. 5D).

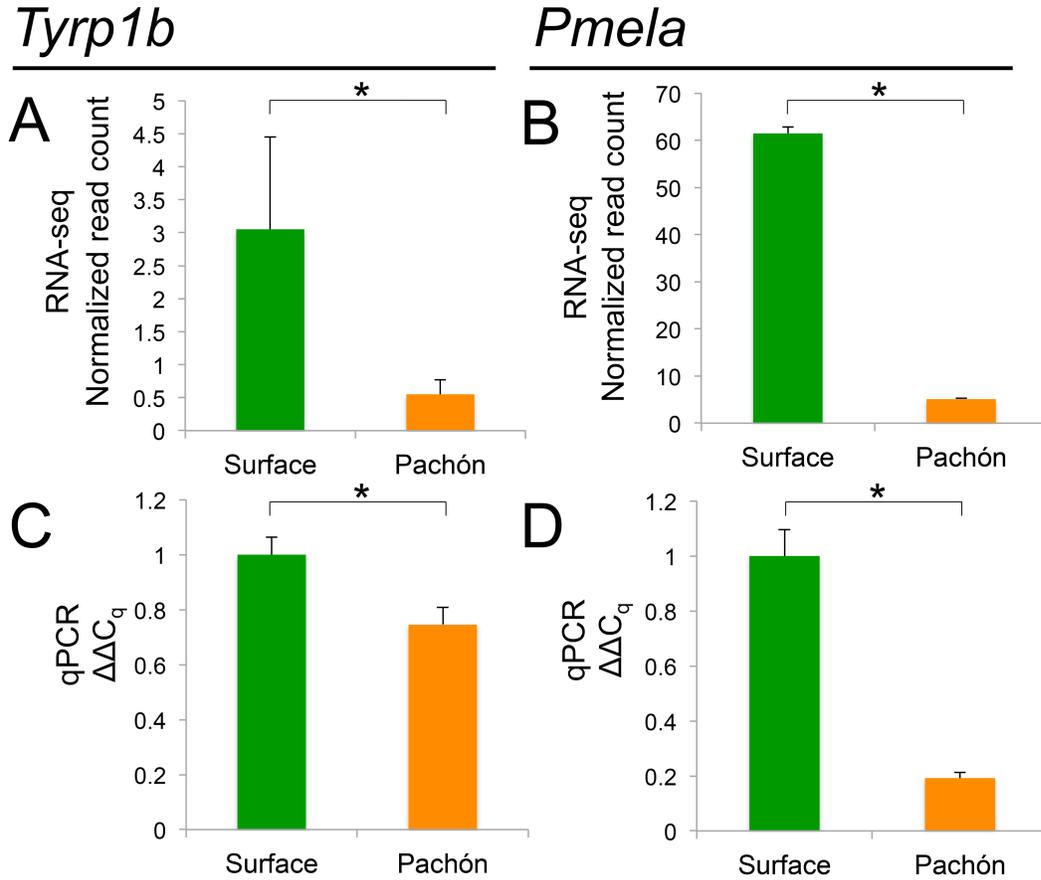


Figure 5. Quantitative (qPCR) confirm reduced expression of *Tyrp1b* and *Pmela* in Pachón cavefish. To further validate reduced expression patterns from our RNA-seq expression profiling (A, B), we performed quantitative PCR on surface and Pachón cavefish at 72 hpf. These analyses confirmed dramatic reduction of expression in cave-dwelling morphotypes for both *Tyrp1b* ($p = 0.005967$) and *Pmela* ($p < 0.000001$).

Tyrp1b and Pmela demonstrate distinct melanophore-specific expression patterns between cave and surface morphotypes

Tyrp1b and *Pmela* have previously appreciated functions contributing to pigmentation in other systems, and we predicted similar involvement in coloration; however, their precise roles have not been characterized in *Astyanax mexicanus*. To test if these genes are expressed in melanophores, we utilized whole-mount *in situ* hybridization (WISH) to visualize expression patterns in 72 hpf surface and Pachón cave specimens. Images of the WISH individuals demonstrate positive staining for *Tyrp1b* and *Pmela* in pigment cells – dark, purple marks across the body – in both surface (Fig. 6B,C,H,I) and cave individuals (Fig. 6E,F,K,L), compared to negative control specimens (Fig. 6A,D,G,J). Melanophores still persist in albino cavefish, but the cells simply do not contain any melanin – the gene *Oca2* that governs albinism only affects the eumelanin biosynthesis pathway (Protas et al. 2005; Jeffery 2013). This melanophore-specific expression confirms that these two genes likely serve homologous roles in pigmentation in *Astyanax*.

As expected, surface fish appear to have increased numbers of pigment cells compared to cavefish. In particular, we discovered numerous melanophores on the head, dorsal and inferior regions of surface specimen stained for *Tyrp1b* (Fig. 6B) and *Pmela* (Fig. 6H). Interestingly, “*Tyrp1b*” cave individuals show *absence* of expression in the dorsal area and is maintained along the entire top of the specimen (Fig. 6E). Additionally, fewer cells are detected in the region near the developing anal fin (Fig. 6K). The signal for *Pmela* in surface fish is robust, revealing expression in numerous melanophores, especially throughout the region surrounding the mid-lateral stripe (Fig. 6C,I). In cavefish, we did not detect *Pmela* expression in this area near the stripe (Fig. 6F,L).

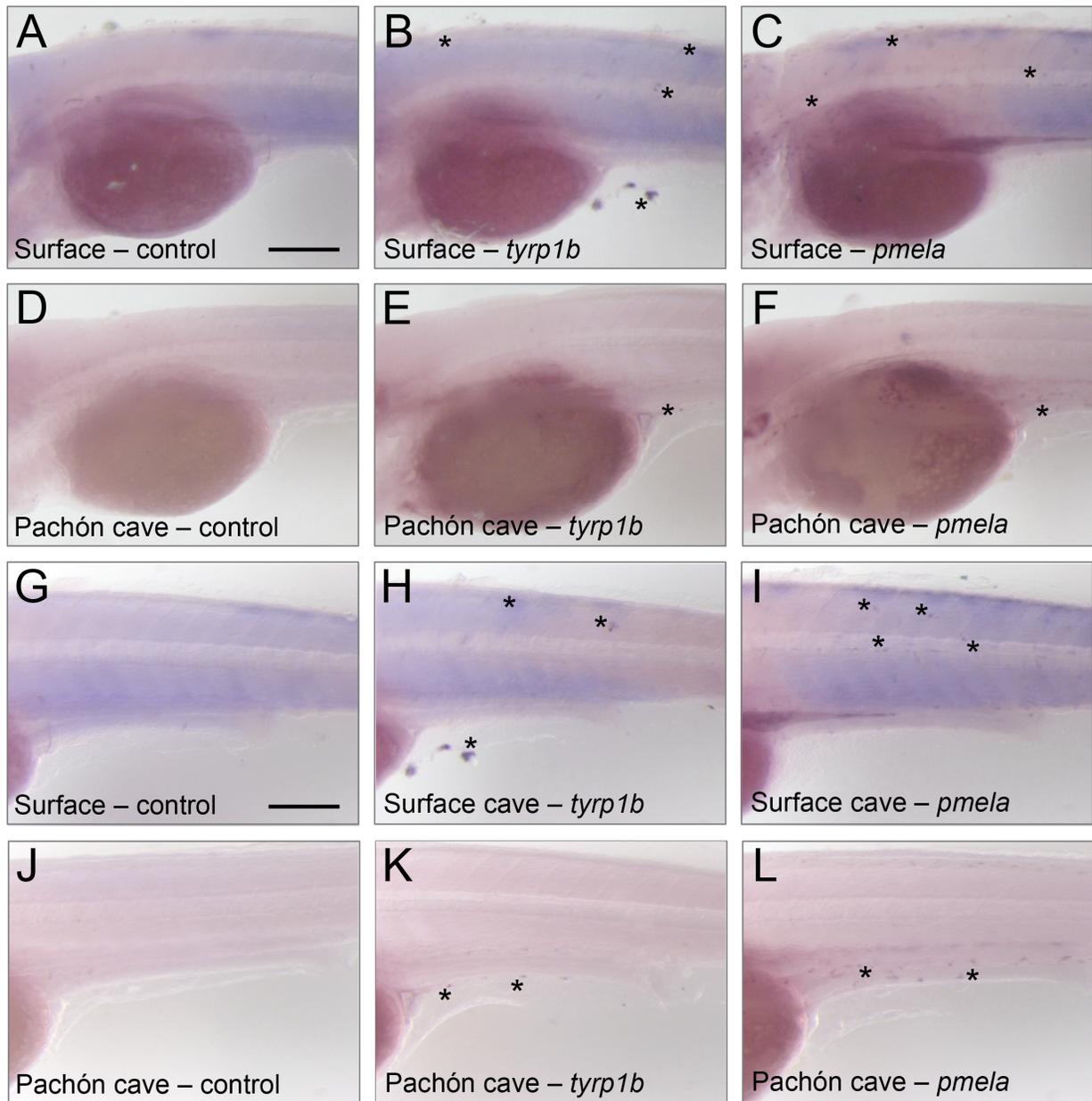


Figure 6. Melanophore-specific expression detected for *Tyrp1b* and *Pmela* in *Astyanax*.

Although pigmentation-related genes are generally conserved across taxa, we performed a series of whole mount *in situ* to characterize the localized expression patterns of *Tyrp1b* and *Pmela* in *Astyanax*. As anticipated, we detected a melanophore-specific pattern for these two candidate genes (dark-purple speckles) compared to negative control specimens (A, D, G, J). We placed

asterisks in the regions showing a positive signal in the pigment cells. Further, we discovered that expression of *Tyrp1* was detected in the dorsal region surface fish (B, H), yet not in this region in the cave individuals (E, K). *Pmela* expression is fairly robust in surface (C, I), but only modestly expressed in cavefish, primarily localized to the ventral region (F, L). Interestingly, *Pmela* expression in melanophores is absent in the area near the mid-lateral stripe. Scale bars = 100 μm .

Functional analyses of Tyrp1b and Pmela reveal altered melanophore dispersion and structure in morphants

To further characterize the role of *Tyrp1b* and *Pmela*, we performed gene knockdown studies. Although *Astyanax* serves as an excellent non-model system for uncovering genes related to pigmentation, the breeding schedule in *Astyanax* is not ideal for collecting single-celled embryos for injections – typically most breeding occurs at night from ~11 PM to 4 AM. Therefore, we utilized the closely-related (120 MYa diverged) model fish system *Danio rerio* to carry out these experiments. For this, we injected single-celled embryos with morpholino oligonucleotides targeted against the first 25-bp of the open reading frame (ORF). The binding of the morpholino specifically blocks the “ATG” start site, and inhibits binding of translational machinery, ultimately causing “knockdown” of the target mRNA. Consistent with our expression studies, we discovered that knockdown of either *Tyrp1b* and *Pmela* in zebrafish causes an overall, qualitative reduction of pigmentation in morphant individuals (Fig. 7B,C) compared to the controls (Fig. 7A). Yet, the reduced pigmentation phenotype varied among individuals.

Interestingly, the *Tyrp1b* mutants (n= 6) often showed patches where melanophores were missing on the dorsal region of head, below the mid-lateral stripe, and/or on the yolk sac (Fig. 7B). The *Tyrp1b* morphants also demonstrated fewer melanophores compared to the control or wild-type individuals. Further, the eye appeared smaller when *Tyrp1b* was inhibited. Similar results were observed in *Pmela* mutant zebrafish (n= 7). Generally, we noted a qualitative decrease in pigment cell number –this was apparent in dorsal region of the head and on the yolk sac (Fig. 7C). Another interesting observation was the altered morphology of the melanophores in *Pmela* mutants. As seen in the control individual, a typical melanophore is reasonably symmetric in shape; however, pigment cells in the *Pmela* morphant exhibit an odd, non-radial morphology. Moreover, the melanin itself appears to be aggregated within the melanophore, evident as dark spots and inconsistent color throughout the pigment cells. Forthcoming studies, including genome-editing to induce lesions in *Astyanax*, combined with high resolution scoring, to will continue to clarify the precise pigmentations phenotypes associated with the genes *Tyrp1b* and *Pmela*.

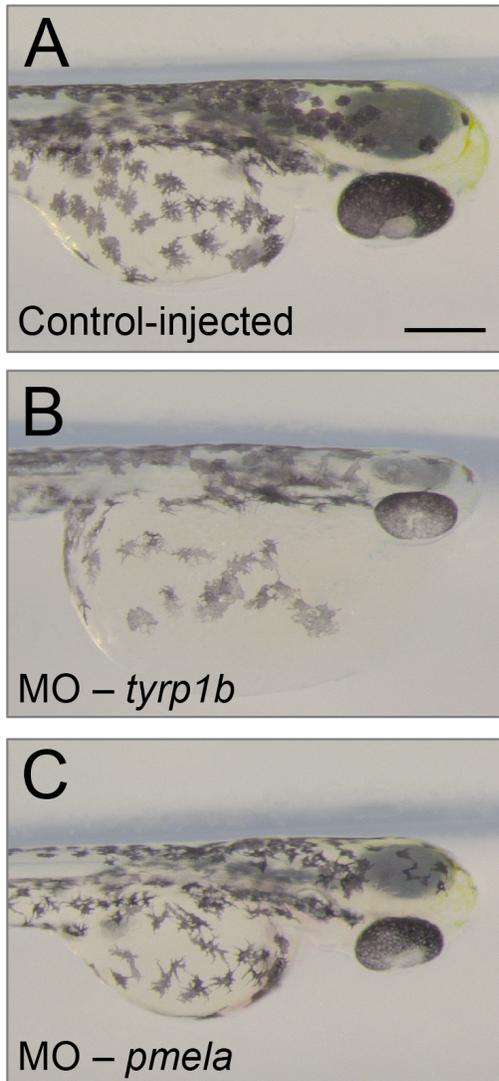


Figure 7. Reduced pigmentation phenotypes are recapitulated in *Tyrp1b*-MO and *Pmela*-MO experimental knockdowns in zebrafish. To characterize the phenotypes associated with inhibition of our two candidates, *Tyrp1b* and *Pmela*, we turn to the model fish system *Danio rerio* to carry out morpholino knockdown experiments. Strikingly, these results demonstrated decreased pigmentation in both *Tyrp1b* (B) and *Pmela* (C) morphant individuals compared to the control (A). These phenotypes were consistent with degenerative pigmentation in other systems, namely global reductions in pigment and altered shape of melanophores, respectively. Scale bar = 200 μm .

DISCUSSION

Tyrp1b and Pmela contribute to melanophore number variation in cavefish

Many pigmentation studies focus on traits mediated by a single (or few) genes. Yet, overall coloration is quite complex – often the result of hundreds or thousands of genes. However, the task of identifying multiple genes of varying or modest effect can prove challenging, therefore our knowledge of the genes governing complex traits is limited. To approach this, we capitalized on an emerging genetics model system *Astyanax mexicanus* cavefish (Jeffery 2001; Gross 2012a; Gross et al. 2014). Notably, cave-dwelling *Astyanax* has evolved an array of diverse phenotypes upon colonization of the cave, including reduction in pigmentation and eyes, and several expanded characters such as taste and an enhance lateral line relative to the extant, surface-dwelling form (Montgomery et al. 2001; Gross 2012a; Kowalko et al. 2013). Many of these characters also vary in degree of severity by cave population, further suggesting that numerous different genes underlie certain phenotypes (Borowsky 2008; Gross 2012a).

Here, we utilized an integrative approach to identify two genes – *Tyrp1b* and *Pmela* – that play a role in complex melanophore number diversity in *Astyanax*. In characterizing these genes, we capitalized on a wealth of knowledge that suggests deeply conserved sets of pigmentation-related processes are governed by *Tyrp1b* and *Pmela*. *Tyrosinase-related protein 1* (*Tyrp1*) serves as a stabilizing protein for tyrosinase which functions in melanin production, and when the *Tyrp1b* protein is absent, tyrosinase rapidly degrades (Müller et al. 1988; Kobayashi and Hearing 2007). The role of *Tyrp1* in eumelanin (black/brown color) pigmentation appears to be shared across very distant taxa. Two distinct *Tyrp1* alleles in transgenic mice yield degenerative pigmentation phenotypes specific to each variant: one causing either brown coat

coloration or nearly white fur relative to the normally black mice (b allele) and another form could induce albinism (c allele; Kwon et al. 1989). Similarly, variants of *Tyrp1* are the cause of “chocolate” and “cinnamon” coat colors in the domestic cat (Lyons et al. 2005; Schmidt-Küntzel et al. 2005). In addition, numerous alleles have been identified in humans of various geographic origins to cause oculocutaneous albinism type III, red hair or blonde hair along with changes in skin pigmentation, which are often caused by small deletions or single bp changes that alter the amino acid sequence (Forsheew et al. 2005; Rooryck et al. 2006; Chiang et al. 2008; Kenny et al. 2012).

Although most mutations characterized in *Tyrp1* are the result of non-synonymous changes, we instead discovered a synonymous alteration in the coding sequence (F210F; Fig. 4B). Yet, some studies of *Tyrp1b* also identify synonymous mutations – for example, six SNPs described are silent mutations in the domestic cat, wherein the precise impact of these sequence alterations has not yet been explained (Lyons et al. 2005). But notably, the results of our RNA-seq and qPCR studies showed significantly reduced *Tyrp1b* expression in Pachón cavefish compared to the surface-dwelling form, suggesting that this gene is indeed altered and associated with reduced pigmentation in zebrafish knockdown experiments. Therefore, the decrease in expression we detect in cavefish may contribute to insufficient stability of tyrosinase and could be responsible, in part, for the reduced numbers of melanophores observed in Pachón cave morphotypes.

The gene *Premelanosome protein (Pmel)* – also known as “*Silv*” or “*Pmel17*” – is a similarly well-described, integral gene in pigmentation development (Theos et al. 2005). This gene was first identified as the locus responsible for “recessive dilution” of coat color that continues to deteriorate with age in a traditional mouse model (Dunn and Thigpen 1930), but

recently, has been implicated in a full spectrum of coloration phenotypes in diverse non-model systems. Variants of *Pmel* cause dilution or hypotrichosis of coat color in the domestic yak (Zhang et al. 2014) and cattle (Jolly et al. 2011; Schmutz and Dreger 2013) and the “silver” phenotype in horse (Brunberg et al. 2006) and zebrafish (Schonthaler et al. 2005) that result as a consequence of small deletions or SNPs affecting the amino acid sequence.

Interestingly, we discovered many SNPs in the *Pmela* coding sequence, including three non-synonymous changes. One mutation we highlight is the G-to-A in cavefish that causes an alanine to threonine at residue 580 in the protein sequence. This change proves quite intriguing, considering that *Pmel* undergoes substantial post-translational modification and processing that is imperative for proper functioning of the protein (Theos et al. 2013). In this instance, the dramatic change in the “class” of amino acid from a hydrophobic alanine (non-polar side chains) to a hydrophilic threonine (polar side chains) may alter the *Pmel* protein. In fact, Ala-to-Thr residue changes that are caused by SNPs in diverse proteins have been shown to induce self-aggregation into amyloids (i.e., an insoluble β -pleated sheet form due to an alteration in secondary structure). The accumulation of these amyloidogenic proteins occurs as a consequence of the preferential switch of alanine (common in helices) to threonine (supports β -sheet structures), which are the basis for amyloid fibril formation (Sunde and Blake 1998; Podoly et al. 2010). Although some amyloids are native to *Pmel*, additional changes in this structure-sensitive protein may prove detrimental to normal functioning (Fowler et al. 2005).

The function of *Pmel* in melanin biosynthesis is currently debated, but it has been described as a melanocyte-specific type 1 transmembrane encoded protein that is enriched in melanosomes (i.e., the pigment-producing organelles within the melanocytes/melanophores) (Reviewed in Theos et al. 2005). In particular, *Pmel* plays a critical role in the premelanosome

“fibril” ultrastructure – Pmel-targeted antibodies are evident in fibrous, stage II melanosomes as Pmel is thought to polymerize fibrillar arrays that ultimately form the backbone of eumelanosomes (Spanakis et al. 1992; Solano et al. 2000; Raposo et al. 2001; Berson et al. 2001; Raposo and Marks 2002). Moreover, additional studies revealed that high levels of *Pmel* expression in non-agouti (solid dark) mice is necessary to construct the fibrils associated with the shape of normal eumelanosomes (Theos et al. 2005).

Remarkably, we detected substantial reduction in *Pmela* expression across early development in Pachón cavefish relative to the surface form using both RNA-seq and quantitative PCR analyses. It is possible that the reduced expression of *Pmela* could impede fibril formation, leading to the aberrant melanophore morphology observed in our knockdown studies of *Pmela* in *Danio*. This is consistent with other reports that similarly describe changes in melanosome shape but modest effects on the overall appearance of body pigmentation in mice (Hellström et al. 2011). Also of note, *Pmel* mutants often have reduced or scattered pigment granules in individual hairs (Dunn and Thigpen 1930). Furthermore, the “merle” phenotype in domestic dogs (e.g., Australian Shepherd breed) is caused by codominance of the merle (M) and the non-merle (m) *Pmel* alleles in heterozygous individuals that collectively yield inconsistent patches of dark- vs. light-colored fur due to protein instability in some regions at random during pigmentation development (Clark et al. 2006; Schmutz and Berryere 2007). Curiously, these findings could explain the increases and decreases of melanophores in the regions we scored in our F₂ hybrid individuals. These previously described degenerative pigmentation phenotypes, combined with the structural changes and severely reduced expression in colorless cavefish, suggest that *Tyrp1b* and *Pmela* contribute to the complex trait of melanophore variation in *Astyanax mexicanus*.

Analysis of complex pigmentation provides insight on regressive evolution in cavefish

The loss or regression of phenotypic characters is a remarkable evolutionary process, though the genetic mechanisms mediating this phenomenon remain unclear. *Astyanax mexicanus* cavefish have undergone intense losses – namely eye degeneration and reduction of pigmentation – upon colonization of the subterranean environment (Jeffery 2001; Gross 2012a). Although these traits are rendered “useless” in the total darkness of the cave, it is difficult to understand “why” certain characters are lost, and further, determine the evolutionary and genetic processes through which traits are discarded. To investigate this notion, we utilize the striking regression of pigmentation that is evident across multiple independent lineages of *Astyanax* cave-dwelling fish to understand the mechanisms governing trait losses (Borowsky 2008).

Within the literature, three competing hypotheses for regressive evolution include natural selection, neutral mutation/genetic drift, and pleiotropy (Culver 1982). A previous QTL study, mapping melanophore number in four regions spanning the body, yielded 18 associations that varied in their phenotypic effect. In some cases, the homozygous genotype for cave was associated with having *increased* numbers of melanophores (Protas et al. 2007). Our results revealed a similar finding, with genotypic effect plots with varying directionality for surface and cave morphs. Some argue this is evidence for neutral mutation/genetic drift or pleiotropy because it would be unlikely that depigmented cavefish (if under selection) would harbor alleles that increase pigment cell numbers (Protas et al. 2007). However, one recent study discovered a prospective “selective” benefit for loss of pigmentation in cave-dwelling morphs. The gene *Oca2*, known for governing albinism in cavefish, normally functions to transport a melanin precursor (L-DOPA) into the melanosome for melanin biosynthesis. However, in colorless individuals this intermediate may be instead utilized in the catecholamine pathway, and

potentially cause sleep loss – this altered behavior may benefit cavefish by enabling constant foraging for food in the nutrient poor subterranean environment (Bilandžija et al. 2013).

Although the evidence remains unclear as to the precise evolutionary pressures governing regressive traits in caves, continued studies will enlighten the genetic mechanisms that accompany phenotypic loss in the wild.

Astyanax mexicanus serves as an excellent model for degenerative pigmentation disorders

Various levels of degenerative pigmentation are found in a wide array of natural systems spanning the globe. This includes numerous pigment-related diseases that directly affect humans including albinism, vitiligo, and skin melanoma (Oetting et al. 1996; Agarwal 1998; Hocker and Tsao 2007). In many cases, the etiologies for these disorders remain unknown. To investigate this, the blind Mexican cavefish serves as an ideal system to model such losses, since these fish have recurrently accumulated extreme reductions or absence of pigmentation upon colonization of the cave, and have achieved this through natural environmental pressures (Jeffery 2005; Gross 2012a). Moreover, the functions of many pigmentation genes are shared broadly across taxa. For example, the gene *Oca2* causes albinism in three cavefish populations, and similarly, the homologous form causes oculocutaneous albinism type II in humans (Protas et al. 2005). Additionally, the *Mclr* locus causes *brown* in cavefish (Gross et al. 2009), and the orthologous gene in humans leads to red hair and fair skin (Valverde 1995; Flanagan et al. 2000; Rees 2003).

Although, we describe the role of *Tyrb1b* and *Pmela* in melanophore variation in cavefish, these genes also are also critical for normal production of melanin in humans. When mutated in human populations, *Tyrb1* is responsible for oculocutaneous albinism type III due to mutations (usually small deletions) in the coding sequence. In a scenario similar to cavefish,

Tyrp1 has been repeatedly subject to mutation with several known variants, including specific alleles associated with different geographic origins (Manga et al. 1997; Forshew et al. 2005; Rooryck et al. 2006). Yet, several common *Tyrp1* alleles can instead contribute to normal hair, skin and iris coloration in humans (Frudakis et al. 2003; Sulem et al. 2007; Han et al. 2008; Liu and Fisher 2010; Eriksson et al. 2010). However, knowledge of *Pmel* variants in humans is limited. Yet, *Pmel* warrants additional investigation considering some of the disorders present in other systems, such as double merle dogs, cause hearing and ocular deficiencies (due to pigmentation losses in the ears and eyes) alongside pigmentation variation (Clark et al. 2006). Evaluation of pigmentation genes in cavefish will continue to identify vulnerable genes, shared broadly across taxa, and may ultimately have implications in human, pigmentation-related diseases.

CONCLUSIONS

To investigate the dynamic range of coloration observed in *Astyanax*, we quantified melanophore number variation in several regions of the body in a large cave x surface F₂ pedigree. Our QTL studies identified associations with 20 distinct regions of our linkage map. Utilizing comparative genomics and gene ontology analyses, we narrowed our search for candidate genes based on the physical position of the prospective loci relative to the top marker location, and further, prioritized genes with previously appreciated roles in pigmentation-related processes using GO terms. This yielded two intriguing genes *Tyrp1b* and *Pmela* for our subsequent analyses. For this, we interrogated our candidates to discover prospective structural changes to the coding sequence and dramatic differences in gene expression in cavefish relative to surface morphs. Remarkably, *Tyrp1b* and *Pmela* did indeed harbor mutations in the coding region, and demonstrated significant differences in expression between morphotypes. We also validated that these candidate genes are expressed specifically in melanophores, and further, that inhibition via knockdowns leads to reduced levels of pigmentation in morphants. In conclusion, we identified two genes *Tyrp1b* and *Pmela* that are involved in the complex trait of melanophore variation, which may contribute to regressive pigmentation in *Astyanax* cavefish.

Funding

This study was funded by the National Science Foundation, Washington D.C., U.S.A. (grant number DEB-1457630).

Acknowledgements

The authors wish to thank Brian Carlson for assistance with the linkage map assembly, Amanda Powers for help imaging fish, and Kate Curoe and Nuria Alonso for aiding with melanophore scoring. The authors also thank Aaron Stahl for guidance with *in situ* hybridization methods and members of the Gross lab for helpful discussions.

Ethical Statement

All applicable international, national, and/or institutional guidelines for the care and use of animals were followed. All procedures performed in studies involving animals were in accordance with the ethical standards of the institution or practice at which the studies were conducted. The protocol was approved by the Institutional Animal Care and Use Committee (IACUC) of the University of Cincinnati (Protocol Number 10-01-21-01).

REFERENCES

- Agarwal G, A. (1998) Vitiligo: an under-estimated problem. *Family Practice* 15 Suppl 1: S19–23.
- Benjamini, Y., and Y. Hochberg (2000) On the Adaptive Control of the False Discovery Rate in Multiple Testing With Independent Statistics. *Journal of Educational and Behavioral Statistics* 25(1): 60–83.
- Berson, J. F., D. C. Harper, D. Tenza, G. Raposo, and M. S. Marks (2001) *Pmel17* Initiates Premelanosome Morphogenesis within Multivesicular Bodies. *Molecular Biology of the Cell* 12: 3451–3464.
- Bilandžija, H., L. Ma, A. Parkhurst, and W. R. Jeffery (2013) A Potential Benefit of Albinism in *Astyanax* Cavefish: Downregulation of the *oca2* Gene Increases Tyrosine and Catecholamine Levels as an Alternative to Melanin Synthesis (H. Escriva, Ed.). *PLoS ONE* 8: e80823–14.
- Borowsky, R. (2008) *Astyanax mexicanus*, the Blind Mexican Cave Fish: A Model for Studies in Development and Morphology. *Cold Spring Harbor Protocols* 2008: pdb.emo107–pdb.emo107.
- Broman, K. W., H. Wu, S. Sen, and G. A. Churchill (2003) R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19: 889–890.
- Brunberg, E., Anderson, L., Cothran, G., Sandberg, K., Mikko S., and G. Lindgren (2006) A missense mutation in *PMEL17* is associated with the Silver coat color in the horse. *BMC Genetics* 7:46.
- Carlson, B. M., S. W. Onusko, and J. B. Gross (2015) A High-Density Linkage Map for

Astyanax mexicanus Using Genotyping-by-Sequencing Technology. *G3* 5: 241–251.

Chiang, P. W., A. B. Fulton, E. Spector, and F. M. Hisama (2008) Synergistic interaction of the *OCA2* and *OCA3* genes in a family. *American Journal of Medical Genetics Part A* 146A: 2427–2430.

Clark, L. A., J. M. Wahl, C. A. Rees, and K. E. Murphy (2006) Retrotransposon insertion in *SILV* is responsible for merle patterning of the domestic dog. *Proceedings of the National Academy of Sciences* 103: 1376–1381.

Culver, D. C. (1982) *Cave life: evolution and ecology*. Cambridge: Harvard University Press 189 p.

Dunn, L. C., and L. W. Thigpen (1930) The silver mouse a recessive color variation. *Journal of Heredity* 21 (12): 495-498.

Elshire, R. J., J. C. Glaubitz, Q. Sun, J. A. Poland, K. Kawamoto, E. S. Buckler, and S. E. Mitchell (2011) A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species (L. Orban, Ed.). *PLoS ONE* 6: e19379.

Eriksson, N. J. M. Macpherson, J. Y. Tung, L. S. Hon, B. Naughton, S. Saxonov, L. Avey, A. Wojcicki, I. Pe'er, and J. Mountain (2010) Web-based, Participant-Driven Studies Yield Novel Genetic Associations for Common Traits. *PLoS Genetics* 6(6): e1000993.

Flanagan, N, E. Healy, A. Ray, S. Philips, C. Todd, I. J. Jackson, M. A. Birch-Machin, and J. L. Rees (2000) Pleiotropic effects of the melanocortin 1 receptor (*Mclr*) gene on human pigmentation. *Human Molecular Genetics* 9(17): 2531–2537.

- Forsheaw, T., S. Khaliq, L. Tee, U. Smith, C. A. Johnson, S. Q. Mehdi, and E. R. Maker (2005) Identification of novel *TYR* and *TYRP1* mutations in oculocutaneous albinism. *Clinical genetics* 68: 182–184.
- Fowler, D. M., A. V. Koulov, C. Alory-Jost, M. S. Marks, W. E. Balch, and J. W. Kelly (2005) Functional Amyloid Formation within Mammalian Tissue. *PLoS Biology* 4(1): e6.
- Frudakis, T., M. Thomas, Z. Gaskin, K. Venkateswarlu, K. Suresh Chandra, Siva Ginjupalli, Sitaram Gunturi, Sivamani Natrajan, Viswanathan K. Ponnuswamy and K. N. Ponnuswamy (2003) Sequences Associated with Human Iris Pigmentation. *Genetics* 165(4): 2071–2083.
- Gross, J. B. (2012a) *Cave Evolution*. John Wiley & Sons, Ltd, Chichester, UK.
- Gross, J. B. (2012b) The complex origin of *Astyanax* cavefish. *BMC Evolutionary Biology* 12: 105–22.
- Gross, J. B., and H. Wilkens (2013) Albinism in phylogenetically and geographically distinct populations of *Astyanax* cavefish arises through the same loss-of-function *Oca2* allele. *Heredity* 1–9.
- Gross, J. B., R. Borowsky, and C. J. Tabin (2009) A Novel Role for *Mcl1* in the Parallel Evolution of Depigmentation in Independent Populations of the Cavefish *Astyanax mexicanus* (G. S. Barsh, Ed.). *PLoS Genetics* 5: e1000326–14.
- Gross, J. B., A. Furterer, B. M. Carlson, and B. A. Stahl (2013) An Integrated Transcriptome-Wide Analysis of Cave and Surface Dwelling *Astyanax mexicanus*. *PLoS ONE* 8: e55659–14.

- Gross, J. B., R. Kerney, J. Hanken, and C. J. Tabin (2011) Molecular anatomy of the developing limb in the coquí frog, *Eleutherodactylus coqui*. *Evolution & Development* 13: 415–426.
- Gross, J. B., A. J. Krutzler, and B. M. Carlson (2014) Complex Craniofacial Changes in Blind Cave-Dwelling Fish Are Mediated by Genetically Symmetric and Asymmetric Loci. *Genetics* 196: 1303–1319.
- Gross, J. B., M. Protas, M. Conrad, P. E. Scheid, O. Vidal, W. R. Jeffery, R. Borowsky, and C. J. Tabin (2008) Synteny and candidate gene prediction using an anchored linkage map of *Astyanax mexicanus*. *Proceedings of the National Academy of Sciences* 105: 20106–20111.
- Haley, C. S., and S. A. Knott (1992) A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* 69, 315–324.
- Han, J., P. Kraft, H. Nan, Q. Guo, C. Chen, A. Qureshi, S. E. Hankinson, F. B. Hu, D. L. Duffy, Z. Z. Zhao, N. G. Martin, G. W. Montgomery, N. K. Hayward, G. Thomas, R. N. Hoover, S. Chanock, and D. J. Hunter (2008) A Genome-Wide Association Study Identified Novel Alleles Associated with Hair Color and Skin Pigmentation. *PLoS Genetics* 4(5): e1000074.
- Hellström, A. R., B. Watt, S. S. Fard, D. Tenza, P. Mannström, K. Mannström, B. Ekesten, S. Ito, K. Wakamatso, J. Larsson, M. Ulfendahl, K. Kullander, G. Raposo, S. Kerje, F. Hallböök, M. S. Marks, L. Andersson (2011) Inactivation of *Pmel* Alters Melanosome Shape But Has Only a Subtle Effect on Visible Pigmentation (I. J. Jackson, Ed.). *PLoS Genetics* 7: e1002285.
- Hocker, T., and H. Tsao (2007) Ultraviolet radiation and melanoma: a systematic review and analysis of reported sequence variants. *Human Mutation* 28: 578–588.

- Hoekstra, H. E. (2006) Genetics, development and evolution of adaptive pigmentation in vertebrates. *Heredity* 97: 222–234.
- Hubbard, J. K., J. A. C. Uy, M. E. Hauber, H. E. Hoekstra, and R. J. Safran (2010) Vertebrate pigmentation: from underlying genes to adaptive function. *Trends in Genetics* 26: 231–239.
- Hunt, R., Z. E. Sauna, S. V. Ambudkar, M. M. Gottesman, and C. Kimchi-Sarfaty (2009) Silent (Synonymous) SNPs: Should We Care About Them?, pp. 23–39 in *Single Nucleotide Polymorphisms*, Methods in Molecular Biology, Humana Press, Totowa, NJ.
- Jeffery, W. R. (2005) Adaptive Evolution of Eye Degeneration in the Mexican Blind Cavefish. *Journal of Heredity* 96: 185–196.
- Jeffery, W. R. (2001) Cavefish as a Model System in Evolutionary Developmental Biology. *Developmental Biology* 231: 1–12.
- Jeffery, W. R. (2013) Regressive Evolution of Pigmentation in the Cavefish *Astyanax*. *Israel Journal of Ecology & Evolution* 52: 405–422.
- Jolly, R. D., J. L. Wills, J. E. Kenny, J. I. Cahill, and L. Howe (2011) Coat-colour dilution and hypotrichosis in Hereford crossbred calves. *New Zealand Veterinary Journal* 56: 74–77.
- Kasprzyk, A. (2011) BioMart: driving a paradigm change in biological data management. *Database* 2011:bar049–bar049.
- Kearsey, M. J. and V. Hyne (1994) QTL analysis: a simple “marker-regression” approach. *Theoretical Applied Genetics* 89: 698–702.
- Kenny, E. E., N. J. Timpson, M. Sikora, M.-C. Yee, A. Moreno-Estrada, C. Eng, S. Huntsman,

- E. González Burchard, M. Stoneking, C. D. Bustamante, S. Myles (2012) Melanesian Blond Hair Is Caused by an Amino Acid Change in *TYRP1*. *Science* 336: 554–554.
- Kirby, R. F., K. W. Thompson, and C. Hubbs (1977) Karyotypic Similarities between the Mexican and Blind Tetras. *Copeia* 1977: 578.
- Kobayashi, T., and V. J. Hearing (2007) Direct interaction of tyrosinase with *Tyrp1* to form heterodimeric complexes *in vivo*. *Journal of Cell Science* 120: 4261–4268.
- Kowalko, J. E., N. Rohner, T. A. Linden, S. B. Rompani, W. C. Warren, R. Borowsky, C. J. Tabin, W. R. Jeffery, and M. Yoshizawa (2013) Convergence in feeding posture occurs through different genetic loci in independently evolved cave populations of *Astyanax mexicanus*. *Proceedings of the National Academy of Sciences* 110: 16933–16938.
- Kruglyak, L., and E. S. Lander (1995) A nonparametric approach for mapping quantitative trait loci. *Genetics* 139: 1421–1428.
- Krzywinski, M., J. Schein, I. Birol, J. Connors, R. Gascoyne, D. Horsman, S. J. Jones, and M. A. Marra (2009) Circos: An information aesthetic for comparative genomics. *Genome Research* 19: 1639–1645.
- Kwon, B. S., R. Halaban, and C. Chintamaneni (1989) Molecular basis of mouse Himalayan mutation. *Biochemical and Biophysical Research Communications* 161: 252–260.
- Linnen, C. R., E. P. Kingsley, J. D. Jensen, and H. E. Hoekstra (2009) On the Origin and Spread of an Adaptive Allele in Deer Mice. *Science* 325: 1095–1098.
- Liu, J. J., and D. E. Fisher (2010) Lighting a path to pigmentation: mechanisms of *MITF*

induction by UV. *Pigment Cell and Melanoma Research* 23: 741–745.

Lu, F., A. E. Lipka, J. Glaubitz, R. Elshire, J. H. Cherney, M. D. Casler, E. S. Buckler, and D. E. Costich (2013) Switchgrass Genomic Diversity, Ploidy, and Evolution: Novel Insights from a Network-Based SNP Discovery Protocol (G. P. Copenhaver, Ed.). *PLoS Genetics* 9: e1003215.

Lyons, L. A., D. L. Imes, H. C. Rah, and R. A. Grahn (2005) Tyrosinase mutations associated with Siamese and Burmese patterns in the domestic cat (*Felis catus*). *Animal Genetics* 36: 119–126.

Manga, P., J. G. R. Kromberg, N. F. Box, R. A. Sturm, T. Jenkins, and M. Ramsay (1997) Rufous Oculocutaneous Albinism in Southern African Blacks Is Caused by Mutations in the *TYRP1* Gene. *The American Journal of Human Genetics* 61: 1095–1101.

McGaugh, S. E., J. B. Gross, B. Aken, M. Blin, R. Borowsky, D. Chalopin, H. Hinaux, W. R. Jeffery, A. Keene, L. Ma, P. Minx, D. Murphy, K. E. O'Quinn, S. Rétaux, N. Rohner, S. M. J. Searle, B. A. Stahl, C. Tabin, J. Volff, M. Yoshizawa, and W. C. Warren (2014) The cavefish genome reveals candidate genes for eye loss. *Nature Communications* 5: 5307.

Montgomery, J. C., S. Coombs, and C. F. Baker (2001) The mechanosensory lateral line system of the hypogean form of *Astyanax fasciatus*, pp. 87–96 in *The biology of hypogean fishes*, Developments in environmental biology of fishes, Springer Netherlands, Dordrecht.

Mortazavi, A., B. A. Williams, K. McCue, L. Schaeffer, and B. Wold (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* 5: 621–628.

- Müller, G., S. Ruppert, E. Schmid, and G. Schütz (1988) Functional analysis of alternatively spliced tyrosinase gene transcripts. *The EMBO Journal* 7: 2723.
- O'Quin, K. E., M. Yoshizawa, P. Doshi, and W. R. Jeffery (2013) Quantitative Genetic Analysis of Retinal Degeneration in the Blind Cavefish *Astyanax mexicanus* (A. Lewin, Ed.). *PLoS ONE* 8: e57281–11.
- Oetting, W. S., 2000 The Tyrosinase Gene and Oculocutaneous Albinism Type 1 (*OCA1*): A Model for Understanding the Molecular Biology of Melanin Formation. *Pigment Cell Research* 13: 320–325.
- Oetting, W. S., M. H. Brilliant, and R. A. King (1996) The clinical spectrum of albinism in humans. *Molecular Medicine Today* 2: 330–335.
- Podoly, E., G. Hanin, and H. Soreq (2010) Alanine-to-threonine substitutions and amyloid diseases: Butyrylcholinesterase as a case study. *Chemico-Biological Interactions* 187: 64–71.
- Protas, M., M. Conrad, J. B. Gross, C. Tabin, and R. Borowsky (2007) Regressive Evolution in the Mexican Cave Tetra, *Astyanax mexicanus*. *Current Biology* 17: 452–454.
- Protas, M. E., and N. H. Patel (2008) Evolution of Coloration Patterns. *Annual Review of Cell and Developmental Biology* 24: 425–446.
- Protas, M. E., C. Hersey, D. Kochanek, Y. Zhou, H. Wilkens, W. R. Jeffery, L. I. Zon, R. Borowsky, and C. J. Tabin (2005) Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nature Genetics* 38: 107–111.
- Raposo, G., and M. S. Marks (2002) The Dark Side of Lysosome-Related Organelles:

- Specialization of the Endocytic Pathway for Melanosome Biogenesis. *Traffic* 3: 237–248.
- Raposo, G., D. Tenza, D. M. Murphy, J. F. Berson, and M. S. Marks (2001) Distinct Protein Sorting and Localization to Premelanosomes, Melanosomes, and Lysosomes in Pigmented Melanocytic Cells. *Journal of Cell Biology* 152: 809–824.
- Rees, J. L. (2003) Genetics of Hair and Skin Color. *Annual Review of Genetics* 37:67–90.
- Rooryck, C., C. Roudaut, E. Robine, J. Müsebeck, and B. Arveiler (2006) Oculocutaneous albinism with *TYRP1* gene mutations in a Caucasian patient. *Pigment Cell Research* 19: 239–242.
- Schmidt-Küntzel, A., E. Eizirik, S. J. O'Brien, and M. Menotti-Raymond (2005) *Tyrosinase* and *Tyrosinase Related Protein 1* Alleles Specify Domestic Cat Coat Color Phenotypes of the albino and brown Loci. *Journal of Heredity* 96: 289–301.
- Schmutz, S. M., and T. G. Berryere (2007) Genes affecting coat colour and pattern in domestic dogs: a review. *Animal Genetics* 38: 539–549.
- Schmutz, S. M., and D. L. Dreger (2013) Interaction of *MC1R* and *PMEL* alleles on solid coat colors in Highland cattle. *Animal Genetics* 44: 9–13.
- Schonthaler, H. B., J. M. Lampert, J. von Lintig, H. Schwarz, R. Geisler, S. C. Neuhaus (2005) A mutation in the *silver* gene leads to defects in melanosome biogenesis and alterations in the visual system in the zebrafish mutant fading vision. *Developmental Biology* 284: 421–436.
- Solano, F., M. Martínez Esparza, C. Jimenez-Cervantes, S. P. Hill, J. A. Lozano, J. C. García-

- Borrón (2000) New Insights on the Structure of the Mouse Silver Locus and on the Function of the Silver Protein. *Pigment Cell Research* 13: 118–124.
- Spanakis, E., P. Lamina, and D. C. Bennett (1992) Effects of the developmental colour mutations silver and recessive spotting on proliferation of diploid and immortal mouse melanocytes in culture. *Development* 114: 675–680.
- Sulem, P., D. F. Gudbjartsson, S. N. Stacey, A. Helgason, T. Rafnar, M. Jakobsdottir, S. Steinberg, S. A. Gudjonsson, A. Palsson, G. Thorleifsson, S. Pálsson, B. Sigurgeirsson, K. Thorisdottir, R. Ragnarsson, K. R. Benediksdottir, K. K. Aben, S. H. Vermeulen, A. M. Goldstein, M. A. Tucker, L. A. Kiemeny, J. H. Olafsson, J. Gulcher, A. Kong, U. Thorsteinsdottir and K. Stefansson (2008) Two newly identified genetic determinants of pigmentation in Europeans. *Nature Genetics* 40: 835–837.
- Sunde, M., and C. C. F. Blake (1998) From the globular to the fibrous state: protein structure and structural conversion in amyloid formation. *Quarterly Reviews of Biophysics* 31: 1–39.
- Theos, A. C., S. T. Truschel, G. Raposo, and M. S. Marks (2005) The *Silver* locus product *Pmel17/gp100/Silv/ME20*: controversial in name and in function. *Pigment Cell Research* 18: 322–336.
- Theos, A. C., B. Watt, D. C. Harper, K. J. Janczura, S. C. Theos, K. E. Herman, and M. S. Marks (2013) The PKD domain distinguishes the trafficking and amyloidogenic properties of the pigment cell protein *PMEL* and its homologue *GPNMB*. *Pigment Cell & Melanoma Research* 26: 470–486.
- Valverde, P., E. Healy, I. Jackson, J. L. Rees, and A. J. Thody (1995) Variants of the

melanocyte-stimulating hormone receptor gene are associated with red hair and fair skin in humans. *Nature Genetics* 11: 328–330.

Wingert, R. A., A. Brownlie, J. L. Galloway, K. Dooley, P. Fraenkel, J. L. Axe, A. J. Davidson, B. Barut, L. Noriega, X. Sheng, Y. Zhou, and L. I. Zon (2004) The chianti zebrafish mutant provides a model for erythroid-specific disruption of *transferrin receptor 1*. *Development* 131: 6225–6235.

Xu, S., and Z. Hu (2010) Mapping Quantitative Trait Loci Using Distorted Markers. *International Journal of Plant Genomics* 2009: 1–11.

Zhang, M. Q., X. Xu, and S. J. Luo (2014) The genetics of brown coat color and white spotting in domestic yaks (*Bos grunniens*). *Animal Genetics* 45: 652–659.

CHAPTER 3

Developmental gene expression consequences of recurrent cave colonization¹

Bethany A. Stahl and Joshua B. Gross

Department of Biological Sciences, University of Cincinnati, Cincinnati, Ohio 45221

¹It is expected that the information contained in the chapter will be published elsewhere with Dr. Joshua B. Gross serving as co-author. However, the narrative, analyses and representations of results presented here were developed solely by Bethany A. Stahl.

ABSTRACT

Organisms that colonize extreme environments often evolve equally dramatic phenotypes. However, our knowledge of the gene expression changes that accompany these colonization events remains limited. To investigate this phenomenon, we study organisms that reside in the unusual environment of a cave, which is marked by constant darkness and limited food availability. The blind Mexican cavefish, *Astyanax mexicanus*, has evolved an array of cave-associated phenotypes including regression of pigmentation and eyes, sleep loss, craniofacial aberrations, increased fat storage and expanded non-visual sensory systems. Further, *Astyanax* has repeatedly colonized separate caves throughout Mexico, yet the “ancestral” surface form is extant, allowing for direct comparison to the “derived” cave-dwelling conspecific. In this study, we examine the numerous changes associated with subterranean colonization by measuring differential gene expression between two geographically isolated cavefish populations (Pachón and Tinaja), compared to surface morphs. We performed this analysis using RNA-seq profiling across four critical stages of early development wherein traits appear and others regress. We discovered certain genes that may play a role in cave evolution, such as *Otx2* and *Mitf*, demonstrate similar patterns of expression across development in both cavefish populations. However, other genes including *Olfm4*, *Tas2r202*, *Sagb* and *Nrl* show unique expression profiles specific to each cave-dwelling lineage. This work reveals that cave-adapted traits arise through a combination of both shared and unique patterns of gene expression. Shared expression profiles may signal common environmental pressures driving the evolution of cavefish, yet patterns that are specific to each cave indicate that similar adaptive traits can arise through diverse genetic mechanisms.

INTRODUCTION

In spite of minimal nutrition and constant darkness, cave environments frequently attract inhabitants to avoid predation or expand an unexploited niche. Diverse taxa have thrived in this extreme environment; however the factors determining their success have remained largely unknown. The recurrent evolution of certain stereotypical phenotypes, such as eye/pigmentation loss and extra-visual sensory expansion, hint at potential selective forces present across different cave biomes (Jeffery 2001). However, a role for accumulation of deleterious alleles over long stretches of time cannot be ruled out as a mechanism driving morphological (visual system loss) degeneration.

The blind Mexican cavefish, *Astyanax mexicanus*, has emerged as a powerful model for understanding cave adaptation and evolution. This species includes a common surface-dwelling form and numerous independently-derived cave adapted populations (Jeffery 2001). Population genetic studies reveal a complex evolutionary history resulting from several cave invasions from an “old” stock of surface fish, and a “younger” stock that invaded NE Mexico, ~2–5My and ~1–2My, respectively (Gross 2012a; Bradic et al. 2012). This demographic history provides a powerful opportunity to understand the genetic changes accompanying repeated colonization of the subterranean environment.

Prior complementation cross studies between distinct eyeless populations resulted in offspring with functional visual systems (Wilkens and Strecker 2003). This would suggest different cave populations produce the same cave phenotype through different genetic pathways. Conversely, certain simple phenotypes (albinism and *brown*) are caused by mutations to the same genes (*Oca2* and *Mc1r*) in geographically distinct populations (Protas et al. 2005; Gross et

al. 2009). These disparate results beg the question of whether replicated instances of cave adaptation evolve through changes to the same sets of genes, or rather, if cave evolution involves evolution of particular phenotypes through distinct genetic mechanisms.

An historical lack of genomic resources in *Astyanax* limited prior genetic analyses to candidate gene studies. However, the availability of a draft *Astyanax* genome now enables quantification of genome-wide expression patterns in cave and surface morphs (McGaugh et al. 2014). Using this genomic reference template, we performed deep mRNA-sequencing (RNA-seq) to directly compare transcript abundance between surrogate “ancestral” surface-dwelling fish and two independent cave populations, Pachón and Tinaja (Fig. 1). Both El Abra region caves were colonized by the “older” invasion, and therefore should reflect gene expression patterns associated with extreme levels of cave adaptation. Many phenotypic differences between surface and cavefish emerge during early development, including eye and pigmentation regression. Therefore we profiled expression at four embryonic time points to understand dynamic expression changes associated with these morphological changes across development.

Through expression analyses of >23,000 transcripts from the current draft genome, we discovered unexpectedly diverse patterns global expression changes. For instance, through direct comparisons to the surface-dwelling form, we found that Tinaja and Pachón cavefish demonstrate distinct transcriptomic responses to cave adaptation. At the global expression level, Tinaja cavefish harbor more up-regulated genes than Pachón cavefish. In contrast, Pachón cavefish demonstrate nearly twice as many down-regulated genes compared to Tinaja. Given these dramatic differences, we next evaluated gene ontology “GO” term enrichment to

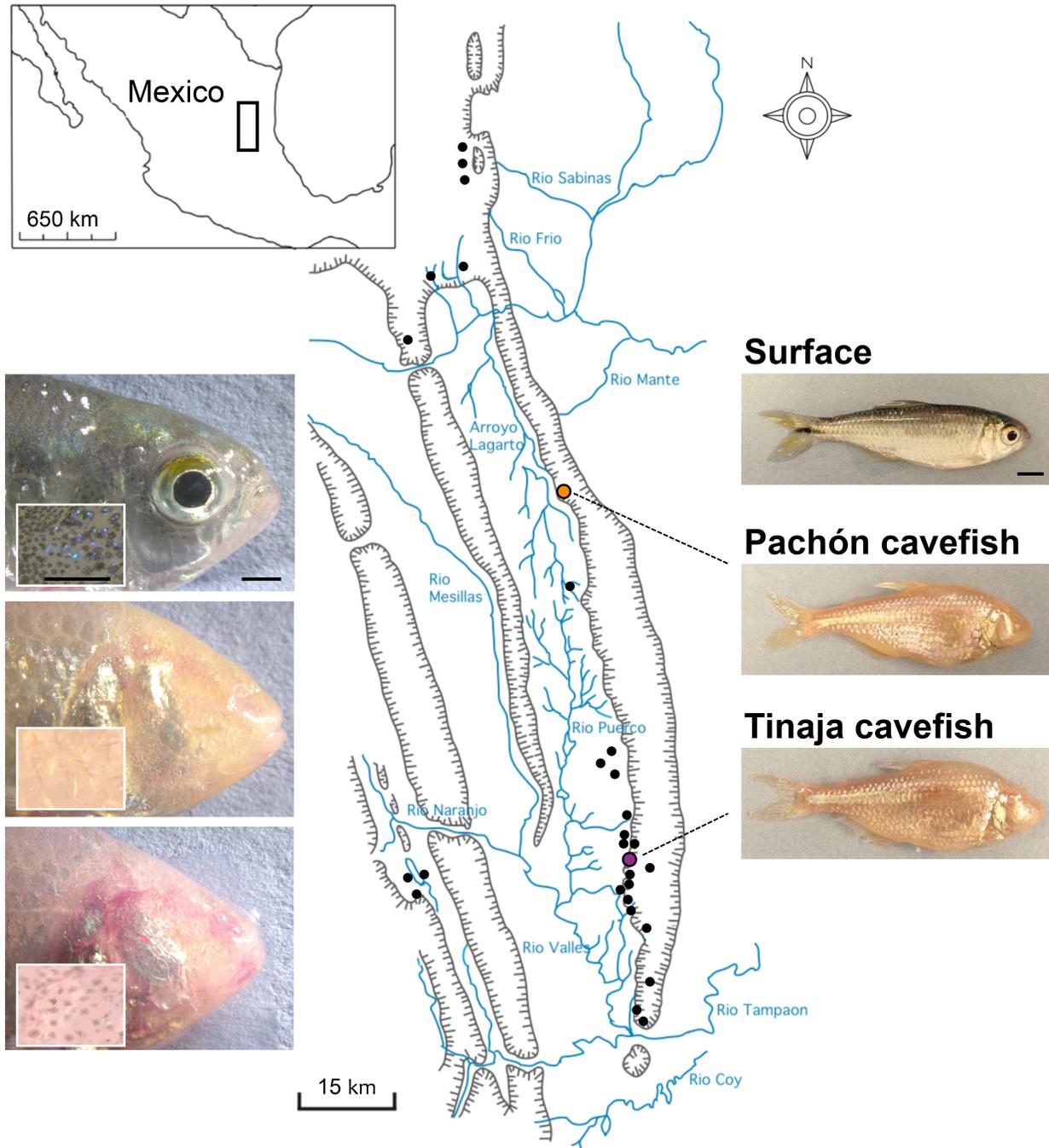


Figure 1. Two geographically isolated *Astyanax* cavefish populations. To characterize prospective gene expression changes governing cave adaptation, we utilize cavefish derived from two independent lineages – Pachón and Tinaja localities – to compare to surface-dwelling individuals that reside in adjacent rivers and streams throughout the El

Abra region of Northeastern Mexico (whole body images, right panels). The overall physical appearance of Pachón and Tinaja appears quite similar, however, higher magnification images reveal apparent differences between caves. For example, Pachón cavefish harbor albinism (complete absence of pigmentation), yet Tinaja cavefish still maintain a very low level of melanin coloration but it is severely reduced compared to surface fish (head and dorsal square inset, left panels). Scale bars = 0.5 cm.

understand if similar gene functions have been selected in these environments. We discovered a complex mosaic of shared and divergent gene processes evident for each cave population. Surprisingly, several of the ‘shared’ GO term functions were associated with different genes in Pachón and Tinaja.

Our results imply selection at the level of genetic processes associated with cave evolution, based on similarities in Gene Ontology functions. Interesting, however, these shared terms are mediated, in part, by different genes. Moreover, specific gene studies reveal a complex pattern of shared and divergent expression patterns. These shared patterns may reflect broad similarity in environmental pressures of the two caves (reduced light and nutrition). Divergent patterns may reflect subtle ecological differences between Pachón and Tinaja, or demographic differences (timing of colonization, migration rates) between the two populations. In sum, this analysis suggests that the process of cave adaptation has unfolded differently in two related cavefish populations.

MATERIALS AND METHODS

Animal husbandry and rearing

Adult surface, Pachón and Tinaja cavefish were reared in a husbandry unit (Aquaneering; San Diego, CA) outfitted with automatic dosers to control pH (~7.4) and conductivity (~800 $\mu\text{S/m}$) levels. Animals were maintained under a 12 hr light:12 hr dark cycle at ~24°C in 10-gallon glass tanks. Natural breedings were induced from groups of 3 – 6 adults with an overnight increase (+2-3°C) in tank water temperature, following 2 days of increased feeding and diversity of nutrition. Adults are routinely reared on TetraMin Pro flake food. A week prior to breeding, adults were fed a diverse diet of liver paste, frozen blood worms, frozen daphnia, frozen brine shrimp, and live black worms. All adult fish were generously provided to us by Dr. Richard Borowsky (New York University).

mRNA extraction and purification

Embryos were reared to the appropriate embryonic stage according to the *Astyanax* developmental staging series (Hinaux et al. 2011). Each of four sets of stage-matched clutches from surface fish, Pachón cavefish and Tinaja cavefish were collected for 10 hours post-fertilization (hpf), 24 hpf, 36 hpf and 72 hpf. Under sterile conditions, 50 embryos from each stage were pooled and processed for RNA extraction (RNeasy Plus Mini Kit; Qiagen, Cat. No. 74134). Following purification, all samples were quantified using a NanoVue Plus Spectrophotometer (GE Healthcare Life Sciences, Product No. 28-9569-62), diluted to a final concentration of 40 $\mu\text{g/mL}$ in 50 μL of RNase-free water (Table S1), and stored at -80°C until sample submission.

mRNA-sequencing

RNA sequencing was carried out at the Cincinnati Children's Hospital Medical Center DNA Sequencing and Genotyping Core (dna.cchmc.org) using an Illumina Hi-Seq 2500 sequencing instrument. Libraries were generated from ~1 µg of purified total RNA using the Illumina TruSeq (v.2) kit. Samples were then sequenced for up to 10 million single-end 50-bp reads, in technical triplicate. In total, we performed 36 sequencing experiments: four developmental stages (each sequenced in triplicate) for surface fish, Pachón cavefish and Tinaja cavefish. Following sequencing, raw data was retrieved (fastq-formatted files) and used for subsequent sequence alignment and expression analyses.

RNA-seq expression analyses

To determine differences in gene expression in distinct cavefish lineages, we utilized a next generation sequencing approach. Raw RNA-sequencing reads from each of four developmental stages (10 hpf, 24 hpf, 36 hpf, and 72, hpf) were aligned to the draft *Astyanax* genome template (Ensembl Genome Browser v.75) comprising a comprehensive set of predicted cDNA sequences (n=23,719). Normalized read counts were calculated for surface, Pachón cavefish, and Tinaja cavefish at each of the four time points using RPKM methods (Mortazavi et al. 2008) in ArrayStar (v.12.0, DNASTar, Madison, WI; see Gross et al. 2013) – a software package that has been widely accepted for analysis of RNA-sequencing data (Leyva-Pérez et al. 2014; Youngblood and MacDonald 2014; Guaiquil et al. 2014). Further, expression differences (i.e., fold change) were determined for every gene across development in each cavefish population relative to the surrogate “ancestral” surface-dwelling form (ArrayStar;

DNASTAR.v.12.0). The significance of differential expression for each comparison was determined using a Student's t-test with FDR (Benjamini and Hochberg 2000) using ArrayStar (DNASTAR.v.12.0). Each subsequent analysis stemmed from the transcriptomic comparisons described above, wherein we aimed to characterize the most *extreme* differences in gene expression that were both shared and unique for each cavefish population. For our subsequent comparisons, we qualified the most dramatic differences as genes expressed at ≥ 10 -fold in each cavefish lineage compared to surface fish. This 10-fold threshold yielded a more manageable dataset: Pachón (n= 1,248 genes) and Tinaja (n= 1,776 genes). For genes that attained this heightened level of differential expression, we performed a cross-comparison (Pachón vs. Tinaja) using ensembl transcript IDs for every stage measured. From this, we denoted “convergent” expression as genes (based on matching transcript IDs) that reached this high threshold at a particular time point in both Pachón and Tinaja cavefish. Conversely, genes expressed at ≥ 10 -fold difference, but only in one cavefish lineage, were categorized as a gene demonstrating “divergent” expression.

Gene ontology enrichment analysis

Gene Ontology enrichment studies were carried out using Blast2GO v.2 (www.blast2go.com). We evaluated multiple thresholds to estimate an appropriate number of genes over- (+) and under-expressed (–) in each cavefish population (Pachón and Tinaja) compared to surface fish. Since each cave population demonstrated different numbers (and fold-change levels) of over- and under-expressed genes, we sought to define a set number of genes for comparison. This approach enabled us to directly compare the “most” over- or under-expressed

genes in each cave population, irrespective of the *level* of differential expression. The thresholds we tested included the top 1% (n=237), 2.5% (n=592), 5% (n= 1,185) and 10% (n= 2,371) most differentially expressed genes. For each percent threshold, genes were collated and assembled into “test sets”, and enrichment of terms in each set were determined through comparison with all ontology terms in the *Astyanax* genome (v.75) reference, using Fischer’s Exact Test (p< 0.05) (Blast2GO). The 2.5% threshold (n=592 genes) provided the most biologically informative terms and provided a manageable number of terms for downstream analyses.

Qualitative and quantitative PCR validation

Quantitative PCR analysis was used to validate representative gene expression patterns observed in RNA-seq studies using methods (described in Gross and Wilkens 2013). Total RNA was extracted from stage-matched Pachón (n=50), Tinaja (n=50) and surface fish (n=50) embryos at 72hpf using the RNeasy extraction kit (Qiagen). cDNA templates were synthesized from 1 µg of high quality RNA following hybridization with oligo dT primers (Invitrogen) at 65 °C for 10 min. After cooling for 5 min on ice, a mixture of 4 µl 5 × RT buffer, 0.5 µl Protector RNase Inhibitor, 2 µl dNTP mixture and 0.5 µl Transcriptor RT was added to each reaction, incubated at 50 °C for 1 h, and then inactivated at 85 °C for 5 min.

We amplified 6 test genes in sextuplet across using a MiniOpticon light cycler (BioRad). The quantification cycle (C_q) values were determined from each sample and normalized by comparison to the reference gene *ribosomal protein s18* (*Rps18*; ENSAMXT00000008147) (primers: forward=5'-ACACGAACATCGATGGTAGGAG-3', reverse=5'-TTGTTGAGGTCGATGTCTGC-3', 112 bp). Experimental gene primer sets included: *slc5a8*

(ENSAMXT00000021731) (forward=5'-AGGCACACTTTCTGGACTATCC-3', reverse=5'-AACAGGGACTTGCAAGACAC-3', 107 bp); *Or115-15* (ENSAMXT00000025780) (forward=5'-CAGCGGTTTTACTTGGATCTTCC-3', reverse=5'-TGCTGCAGGTTTTCAAAGCC-3', 120 bp); *itln1* (ENSAMXT00000003604) (forward=5'-AACCGGGCTACATTTGGAAC-3', reverse=5'-TTATTGGGAACGTGCCACAC-3', 110 bp); *sb:cb1081* (ENSAMXT00000018849) (forward=5'-GCAGCAGCTGTTTGCATATG-3', reverse=5'-AAGCCAAGGAAGAAGGAAGC-3', 99 bp); and *Or101-1* (ENSAMXT00000025777) (forward=5'-GAACCGACTCCAAGCTGAATTC-3', reverse=5'-TGAAAGCATGTTGGGGACTG-3', 101 bp). Primer sets amplified fragments using EvaGreen supermix dye (BioRad) using the following cycling parameters: step 1—95 °C for 30 s, step 2—95 °C for 5 s, step 3—55.1 °C for 10 s, plate read, repeat step 2 for 39 additional cycles. qPCR data were collected and analyzed for normalized fold expression values ($\Delta\Delta C_q$) using Bio-Rad CFX Manager v.3.1 software (BioRad). Qualitative expression was also measured using routine PCR on cDNA libraries also derived from 72 hpf RNA (as described above). We amplified fragments in our target genes ~300-600 bp in length for *Rps18* (forward=5'-TGGCTGGTTAGATGGGTAGC-3', reverse=5'-AGCCCTTGGCGGTTTACTAT-3', 470 bp), *slc5a8* (forward=5'-AGATCTCTGGGGAGCTGTGA-3', reverse=5'-GGCATTAGCTGGTCTTGAGC-3', 492 bp), *Or115-15* (forward=5'-CAGTGGAAAGGAACCTGCAT-3', reverse=5'-TGTGCTGCAGGTTTTCAAAG-3', 584 bp), *itln1* (forward=5'-CTACCTGACAACGGAGAGTGG-3', reverse=5'-GTCTCGGCGTTATTGGGAAC-3', 306 bp), *sb:cb1081* (forward=5'-CCTCACTGAGCACACCAAGA-3', reverse=5'-GCTGGAGCAGGAGTTACAGG-3', 643 bp), and *Or101-1* (forward=5'-AGAACCGACTCCAAGCTGAA-3', reverse=5'-

GCTGATGGCTTTCCCAATTA-3', 531 bp). Bands were visualized using standard (2%) gel electrophoresis under UV transillumination.

Ingenuity pathway analysis (IPA)

We performed analyses to uncover relevant functions and pathways hidden within our complex RNA-seq dataset using the Ingenuity Pathway Analysis (IPA) Suite (Qiagen.v.1.0; www.ingenuity.com/products/ipa). Currently, the IPA analysis package only provides annotation for a few mammalian genomes, so we first converted our *Astyanax* Ensembl transcript IDs to the corresponding human, mouse or rat Ensembl protein ortholog via BioMart (www.ensembl.org/biomart/). Using these identifiers, we then uploaded the output from our RNA-seq analysis (see Methods; fold change and p-value for each population – Pachón and Tinaja relative to surface – by developmental stage) using the 'flexible format' with default settings and a fold change cutoff of 2.0.

For our analyses, we utilized the 'IPA Upstream Regulator' tool, which predicts prospective gene regulators based upon two statistical measures. The first 'overlap p-value' identifies if a statistically significant overlap exists between the observed expression changes in our RNA-seq dataset and the genes known to be governed by the regulator (from evidence available in the Ingenuity® Knowledge Base) using Fisher's Exact Test ($p < 0.01$; IPA Qiagen.v.1.0). In addition, the second metric termed the 'activation z-score' aims to quantify the regulation direction (i.e., activating or inhibiting) based on relationships and experimentally observed results within a given molecular network. With this combined information, the program can predict the state of the transcriptional regulator. However, since some of the genes within a

network will likely show expression patterns inconsistent with the predictions, the program provides a statistical measure – an activation z-score – that determines if a regulator has significantly more “activated” predictions than inhibited predictions ($z\text{-score} > 0$). Conversely, a regulator that has a predicted inhibition state should have significantly more “inhibited” downstream targets ($z\text{-score} < 0$; IPA Qiagen.v.1.0). Moreover, z-scores provide a metric for “sorting” within each dataset to identify the upstream regulators that are most affected and best supported by the observed RNA-seq expression changes in the putative gene targets. Regulators yielding significant results, and their corresponding gene networks, were visualized using the ‘My Pathway’ analysis tool (IPA Qiagen.v.1.0).

RESULTS AND DISCUSSION

Developmental RNA-seq profiling in two distinct populations reveals subterranean adaptation proceeds through surprisingly different gene expression changes

Oftentimes in nature, pressures in a surface environment (e.g., predation or climatic changes) drive animals to pursue the subterranean habitat. Consequently, surface-dwelling relatives will frequently become extinct. Since related surface-dwelling fish persist in the rivers and streams surrounding the El Abra cave network, *Astyanax* cavefish provide the powerful opportunity to directly compared derived cave forms to a putative “ancestral” form. While the historical literature has largely focused on morphological differences between morphs, here we explore transcriptomic alterations in two distinct cave populations.

Each cave population arose from the same ancestral epigeal stock that populated the cavernous El Abra region of NE Mexico ~3–5 MYa (Gross 2012b). Comparisons across the broader landscape have revealed remarkable convergence at the level particular genes for the albinism and *brown* phenotypes (Protas et al. 2005; Gross et al. 2009). This could indicate adaptation to the extreme cave environment is accomplished ‘preferentially’ through a limited set of genetic loci. To evaluate this notion, we characterized gene expression differences in two *Astyanax* cave populations – Pachón and Tinaja – and compared them to extant surface-dwelling fish (Fig. 1). We aligned and quantified reads derived from stage-matched Pachón, Tinaja and surface embryos at four key developmental stages (10 hpf, 24 hpf, 36 hpf, 72 hpf) to a comprehensive genomic template (n=23,719 transcripts). We discovered both substantial differential expression (including many genes demonstrating >10 fold-level changes between morphotypes) alongside many genes expressed at roughly equivalent levels (< 1 fold difference).

We first filtered our dataset by determining the average fold change across all four developmental stages between cave- and surface-dwelling forms. Interestingly, the vast majority of profiled genes exhibited minimal differences in gene expression across all three populations. However, roughly 5,000 genes demonstrated a fold change difference of 2 or greater in each cavefish population (Table S1). More stringent filtering, as expected, revealed fewer genes demonstrating dramatic expression differences in both cave populations: including 4-fold (Pachón: 2,729 vs. Tinaja: 3,352), 6-fold (Pachón: 1,948 vs. Tinaja: 2,553), 8-fold (Pachón: 1,516 vs. Tinaja: 2,080) and 10-fold (Pachón: 1,248 vs. Tinaja: 1,776).

Interestingly, when comparing each cave population to surface fish, Pachón and Tinaja cavefish “over-express” similar numbers of genes at each threshold level: 2-fold (Pachón: 3,000 vs. Tinaja: 2,032), 4-fold (Pachón: 1,536 vs. Tinaja: 1,124), 6-fold (Pachón: 1,087 vs. Tinaja: 858), 8-fold (Pachón: 829 vs. Tinaja: 698) and 10-fold (Pachón: 683 vs. Tinaja: 594). However, Tinaja cavefish show reduced expression of approximately twice as many genes at each fold change threshold when compared to Pachón: 2-fold (Pachón- 2,115 vs. Tinaja- 3,674), 4-fold (Pachón- 1,193 vs. Tinaja- 2,228), 6-fold (Pachón- 861 vs. Tinaja- 1,695), 8-fold (Pachón- 687 vs. Tinaja- 1,382) and 10-fold (Pachón- 565 vs. Tinaja- 1,182; Table S2).

This result indicates that recurrent cave colonization does not proceed through identical patterns of gene expression change. Although the Tinaja and Pachón populations are descended from the same “older” El Abra stock of epigeal ancestors, Pachón fish are regarded as older and more troglomorphic compared to Tinaja fish (Bradic et al. 2012; Gross 2012b). Therefore, our results may indicate that cavefish resulting from younger – and perhaps less isolated – colonization events like Tinaja may manifest regressive changes through alterations in expression patterns. In contrast, the older and more isolated Pachón fish may lose traits through

the accumulation of loss-of-function (LOF) mutations as a consequence of relaxed selection, due to significant time since isolation, rather than reductions gene expression (Culver 1982; Protas et al. 2005; Gross 2012b).

Given the large number of differences in gene expression in both cave forms compared to surface fish, we next evaluated genes demonstrating the *most* extreme expression differences in each cave population. We first identified all genes differentially expressed ≥ 10 -fold (+ or -) at each of four developmental stages, as well as the average fold expression difference across development. Within this dataset, we identified those genes expressed ≥ 10 -fold in either a convergent or divergent fashion in Pachón and Tinaja cavefish. Surprisingly, while some orthologous genes are expressed at the same level of up- or down-regulation in both lineages, many genes were expressed in divergent, cave-specific manners (Table S2). For example, of the 10-fold or greater ‘under-expressed’ genes in Pachón (n= 724) and Tinaja (n= 597) at 72 hpf, 188 of these genes are the same in both populations. This indicates a remarkable level of convergence at the level of individual genes that may play common roles in subterranean colonization. We also discovered, however, a large number of genes at the 72 hpf stage, which demonstrate cave-specific patterns of expression (at ≥ 10 -fold) in Pachón (n= 536) and Tinaja (n= 409). Surprisingly, at 72 hpf, only about 30% of differentially genes (≥ 10 -fold) in Pachón and Tinaja show convergent levels of reduced expression. The remaining majority (~70%) comprise genes with divergent expression patterns (Table S3). These patterns of convergence and divergence were observed at every stage of development evaluated, for both under- and over-expressed genes (Tables S2 and S3). Future studies incorporating multiple surface fish populations may help eliminate noise (i.e., expression differences that vary naturally across

populations, not necessarily involved with cave adaptation), which will continue to help narrow the search for genes contributing to cave-specific changes.

The basis for divergent gene expression differences is unclear. One explanation may be that the same cave-associated phenotype is evolving through distinct genetic mechanisms, such as the mechanisms of eye loss in different caves. Alternatively, nuanced differences in environmental pressures in each cave microenvironment may drive the evolution of distinct phenotypes. Gene expression changes associated with these phenotypes may be manifested over the course of development, explaining in part the differential expression at 72 hpf. Although orthologous genes are not always recruited during independent cave evolution, genes of similar function may explain, in part, recurrent phenotypic adaptation to the cave environment. To evaluate this possibility, we explored the gene ontology terms associated with genes demonstrating extreme differences between cave-dwelling populations and surface fish.

GO enrichment analyses reveal convergence and divergence for cave-associated traits

We characterized prospective biological functions associated with differentially expressed genes by first calculating the *average* fold change difference for every reference transcript (n= 23,719) across all four stages in early development (10 hpf, 24 hpf, 36 hpf and 72 hpf) for Pachón and Tinaja cavefish compared to surface fish. To filter genes demonstrating the most radical expression differences in cave forms, we tested several thresholds (Materials and Methods), and found that genes present in the top 2.5% of over- or under-expressed genes reported a manageable number of genes, inclusive of those with morphological and/or behavioral relevance (Strickler and Jeffery 2009; Gross 2012). Subsequent GO enrichment analyses (Fig. 2)

revealed a suite of terms associated with regressive characters (loss of pigmentation, eye reduction, altered circadian rhythms), and constructive changes (metabolic efficiency, acute non-visual sensation, starvation resistance).

Several cave-associated terms were present in both Pachón and Tinaja enrichment sets. Among those found in the convergent under-expressed category (Fig. 2, bottom), included GO terms relating to vision loss such as, “absorption of visible light”, “G-protein coupled photoreceptor activity” and “photoreceptor outer segment”. Enrichment studies of those genes dramatically over-expressed in both populations revealed convergent GO terms including “antigen processing and presentation”, “determination of left/right asymmetry in lateral mesoderm” and “sensory perception of taste”. These terms are consistent with prior studies reporting difference in immune response (Gross et al. 2013), lateral symmetry (Gross et al. 2014), and enhanced tastant detection (Kowalko et al. 2013) in obligate cave-dwellers.

Genes demonstrating severely reduced expression uniquely in Pachón cavefish were enriched for the term “eye pigmentation”, consistent with the complete absence of pigmentation and reduced eyes in this population (Jeffery 2001; Jeffery 2005). In the over-expressed genes unique to Pachón cavefish, we discovered a number of metabolism and appetite-related processes (e.g., “adult feeding behavior”, “energy reserve metabolic process”, “negative regulation of appetite”, “cellular response to starvation”). This finding is consistent with several reports showing Pachón cavefish demonstrate remarkably efficient metabolic rates in comparison to surface fish as well as other caves (Hüppop 1989).



Figure 2. GO enrichment for extreme expression differences reveals shared and unique functions detected in Pachón and Tinaja cavefish lineages. RNA-seq studies yielded dramatic gene expression differences in Pachón and Tinaja cavefish relative to the surface form, wherein we sought to describe the biological functions and processes associated with the genes most differentially expressed in our dataset. For this, we included genes that fell into the top 2.5% of genes most differentially over- or under-expressed in Pachón cavefish (left side) and Tinaja (right side). We then performed a gene ontology (GO) enrichment test on these rarified gene sets that yielded GO term functions that were both shared (middle) and unique to Pachón and Tinaja cavefish lineages (left and right sides, respectively).

The most under-expressed genes in Tinaja revealed GO terms for the “sensory perception of smell”, “melanocyte-stimulating hormone receptor activity”, “lens fiber cell differentiation” and “lysozyme activity”. In general, these results could potentially indicate different mechanisms through which Tinaja cavefish have acquired cave-associated traits. In contrast, terms derived from our Tinaja over-expressed gene set demonstrated enrichment for “adipose tissue development”, “negative regulation of intrinsic apoptotic signaling pathway in response to oxidative stress” and “9,10 (9', 10')-carotenoid-cleaving dioxygenase activity”. These terms further recapitulate several well-characterized cave-associated traits such as increased fat content, lens apoptosis during early development and abundance of yellow-orange carotenoids, respectively, in cave lineages (Culver and Wilkens 2000).

Interestingly, convergence in GO terminology did not always indicate gene expression level changes in the same genes. For instance, we discovered both Pachón and Tinaja cavefish showed enrichment for “detection of chemical stimulus” and “pheromone receptor activity”. Interestingly, a gene encoding a taste receptor, *Tas21rl*, was upregulated in both caves. However, the gene *V1r1l*, which encodes a putative pheromone receptor, contributed to this enrichment only in Pachón cavefish. Conversely, upregulation of two other genes associated with tastant detection, *Cysltr2* and *Tas2r202*, were only detected in Tinaja cave (Table S4).

Finally, we discovered several unexpected GO terms enriched in both cave populations. For instance, in the under-expressed gene sets, we observed “negative regulation of type B pancreatic cell development”, “heme binding”, tetrahydrofolate biosynthetic process” and “lactose binding”. In the over-expressed gene sets, we observed “DNA (cytosine-5)-methyltransferase activity”, “histone H3-K36 methylation”, “thyroid-stimulating hormone

signaling pathway” and “neurotransmitter: sodium symporter activity”. Future studies may clarify what role these various functions play in cave evolution.

Convergence at the level of individual genes shared between both Pachón and Tinaja cavefish

We sought to identify shared genes demonstrating the highest fold-change differences (≥ 10 -fold) in both cave populations compared to surface fish (Materials and Methods). Several of the genes we identified impact on eye development, maintenance and visual reception. Interestingly, in both cave populations, an eye begins to form, but begins to degrade at ~ 24 hpf (Jeffery 2009). Our results mirror this degeneration. For instance, early on (~ 10 hpf) only three eye-related genes *Saga*, *Bcmo1* and *Cryba2a* demonstrate severely reduced expression. However, beginning at 24 hpf an additional 17 vision-related genes demonstrate dramatic reduction in expression (Table 1; Table S4), including *Pde6h*, *Gnat2*, *Crygn2* and *Cryba111*. By 36 hpf, the genes *Cryaa*, *Lim2*, *Mipa*, *Pax6a* and *Rho* similarly become downregulated. By the latest developmental timepoint we evaluated (72 hpf), the genes *Crx*, *Hsp5*, *Vsx1*, *Nrl*, *Rx3*, *Impg2* and *Prph2l* demonstrate shared reduction in both cave populations. Several of these vision genes (*Pde6*, *Gnat2*, *Crygn2*, *Cryaa*, *Lim2*, *Mipa*, *Pax6a*, *Rho*, *Crx*, *Rx3*, and *Prph2l*) have been previously identified in the context of altered eye development/size in cavefish (Strickler et al. 2002; Gross et al. 2008; Strickler and Jeffery 2009; Gross et al. 2013; O'Quin et al. 2013; McGaugh et al. 2014; Hinaux et al. 2015), this study is the first to determine that these genes are reduced in expression in two geologically distinct eyeless cavefish populations.

Table 1. Genes demonstrating convergent and divergent patterns of expression in both Pachón and Tinaja cavefish compared to surface-dwelling fish.

Expression pattern	Gene	Transcript ID	Physiological function(s)*	Pachón cavefish (fold change)**	Tinaja cavefish (fold change)**
10 hpf					
Convergent reduced	<i>Pmchl</i>	ENSAMXT00000026184	Pigmentation	32.492 down	32.492 down
Convergent reduced	<i>Bcmo1</i>	ENSAMXT00000010290	Eye photoreceptor; hindbrain	21.458 down	10.005 down
Convergent increased	<i>Tas2r11</i>	ENSAMXT00000006165	Sensory perception of taste	70.497 up	234.314 up
Convergent increased	<i>Bco2a</i>	ENSAMXT00000003245	Pigmentation	14.021 up	17.053 up
Divergent – Pachón reduced	<i>Cry1b</i>	ENSAMXT00000020701	Circadian clock	287.900 down	7.724 down
Divergent – Pachón reduced	<i>Btk</i>	ENSAMXT00000002400	Eye; whole organism	92.641 down	1.551 down
Divergent – Tinaja reduced	<i>Npas2</i>	ENSAMXT00000007111	Circadian rhythm; photoperiodism	5.043 down	292.627 down
Divergent – Tinaja reduced	<i>Unc119b</i>	ENSAMXT00000025106	Visual perception	6.511 down	49.682 down
Divergent – Pachón increased	<i>Trpa1a</i>	ENSAMXT00000008706	Sensory perception	115.042 up	3.185 down
Divergent – Pachón increased	<i>Dfnb59</i>	ENSAMXT00000004928	Auditory system; inner ear	37.118 up	no change
Divergent – Tinaja increased	<i>Hhatla</i>	ENSAMXT00000015498	Acyltransferase	9.003 up	308.924 up
Divergent – Tinaja increased	<i>Adcyap1a</i>	ENSAMXT00000007454	eye development; brain development	4.948 up	18.826 up
24 hpf					
Convergent reduced	<i>Crygn2</i>	ENSAMXT00000018569	Eye; lens constituent	76.026 down	2418.442 down
Convergent reduced	<i>Oca2</i>	ENSAMXT00000013137	Pigmentation	16.356 down	11.361 down
Convergent increased	<i>Top1</i>	ENSAMXT00000019646	Oxidative stress; heme binding	89.712 up	50.468 up
Convergent increased	<i>Cd40lg</i>	ENSAMXT00000020533	Immune response	16.574 up	20.235 up
Divergent – Pachón reduced	<i>Vsx2</i>	ENSAMXT00000011986	Eye; retina	17.223 down	6.633 down
Divergent – Pachón reduced	<i>Cryba1a</i>	ENSAMXT00000018270	Lens constituent	12.311 down	1.465 up
Divergent – Tinaja reduced	<i>Pmchl</i>	ENSAMXT00000026184	Pigmentation	2.445 down	492.004 down
Divergent – Tinaja reduced	<i>Pax6l</i>	ENSAMXT00000010326	Retina development camera eye	2.245 up	73.524 down
Divergent – Pachón increased	<i>Dlg4b</i>	ENSAMXT00000020555	Retina; olfactory bulb; midbrain	245.106 up	no change
Divergent – Pachón increased	<i>Myo6b</i>	ENSAMXT00000017674	Auditory receptor; sensory	15.681 up	1.058 up
Divergent – Tinaja increased	<i>Hla-F10al</i>	ENSAMXT00000003617	Immune response	2.726 up	88.106 up
Divergent – Tinaja increased	<i>Mtnr1a</i>	ENSAMXT00000006184	Melatonin receptor	4.581 up	57.666 up
36 hpf					
Convergent reduced	<i>Pax6a</i>	ENSAMXT00000013384	Eye; retina; brain	27.878 down	21.645 down
Convergent reduced	<i>Lim2</i>	ENSAMXT00000004471	Lens constituent	21.287 down	25.195 down
Convergent increased	<i>Alpi.2</i>	ENSAMXT00000001748	Metabolic process	245.367 up	132.464 up
Convergent increased	<i>Plb1</i>	ENSAMXT00000025263	Lipid metabolic process	179.225 up	76.688 up
Divergent – Pachón reduced	<i>Six7</i>	ENSAMXT00000016345	Embryonic camera-type eye	59.027 down	4.105 down
Divergent – Pachón reduced	<i>Atoh7</i>	ENSAMXT00000026693	Retina layer; swimming behavior	38.701 down	2.920 down
Divergent – Tinaja reduced	<i>Chordc1a</i>	ENSAMXT00000003933	Hsp90 protein binding	4.997 down	3610.523 down
Divergent – Tinaja reduced	<i>Lim2.4</i>	ENSAMXT00000009826	Lens constituent	8.052 down	285.606 down
Divergent – Pachón increased	<i>Catsper1</i>	ENSAMXT00000009853	Sperm associated; ion transport	85.391 up	3.365 down
Divergent – Pachón increased	<i>Olfm4</i>	ENSAMXT00000014729	Olfaction	17.626 up	3.768 up
Divergent – Tinaja increased	<i>Endo1</i>	ENSAMXT00000004210	Nucleic acid binding; hydrolase activity	4.124 up	466.036 up
Divergent – Tinaja increased	<i>Cysltr2</i>	ENSAMXT00000026838	Sensory perception of taste	1.104 up	65.981 up
72 hpf					
Convergent reduced	<i>Vsx1</i>	ENSAMXT00000011754	Visual system	32.800 down	15.161 down
Convergent reduced	<i>Crx</i>	ENSAMXT00000001606	Visual perception	31.387 down	13.068 down
Convergent increased	<i>Or115-15</i>	ENSAMXT00000025780	Sensory perception of smell	204.065 up	323.993 up
Convergent increased	<i>Ogn</i>	ENSAMXT00000013874	Ectopic bone formation	11.901 up	10.820 up
Divergent – Pachón reduced	<i>Rgrb</i>	ENSAMXT00000004427	Retinal protein	20.154 down	5.311 down
Divergent – Pachón reduced	<i>Tyr</i>	ENSAMXT00000008998	Pigmentation	12.772 down	3.126 down
Divergent – Tinaja reduced	<i>Crygn2</i>	ENSAMXT00000018569	Lens constituent	2.900 up	1791.267 down
Divergent – Tinaja reduced	<i>Irx7</i>	ENSAMXT00000000343	Retina morphogenesis	1.526 down	17.011 down
Divergent – Pachón increased	<i>Sb:cb1081</i>	ENSAMXT00000018849	Eye; lens constituent	880.522 up	2.117 down
Divergent – Pachón increased	<i>Pde6g</i>	ENSAMXT00000001177	Visual perception	35.934 up	3.174 up
Divergent – Tinaja increased	<i>Olfm4</i>	ENSAMXT00000014744	Olfaction	2.633 down	164.740 up
Divergent – Tinaja increased	<i>Gh1</i>	ENSAMXT00000015091	Adipose tissue development	5.171 down	143.272 up

* Physiological functions are determined by a combination of gene ontology and known literature. Functions in bold may be associated with troglomorphic traits.

** Genes are under- or over-expressed 10-fold or greater in both Pachón and Tinaja cavefish compared to surface. Blue indicates under-expressed, Red represents over-expression.

Two of the genes we identified, *Nrl* and *Impg2*, may have previously unappreciated roles in *Astyanax* eye loss. Interestingly, the gene *Nrl*, which is involved in opsin expression, is also downregulated in the Chinese cavefish species *Sinocyclocheilus* (Meng 2013). Similarly, the retinal gene *Impg2* is inactivated (a pseudogene harboring two splice acceptor mutations and a frameshift deletion) in the subterranean-dwelling Cape golden mole, *Chrysochloris asiatica* (Emerling and Springer 2014). Further, *Impg2* variants are associated with autosomal recessive forms of retinitis pigmentosa in humans (Bandah-Rozenfeld et al. 2010). These results suggest that many of the same genes can contribute to eye loss across disparate lineages of subterranean animals, despite deep divergence times.

Additionally, six genes demonstrating functions related to pigmentation were convergently under-expressed, including *Pmchl* at 10 hpf, *Oca2* at 24 hpf, *Mitfa* and *Pmch* at 36 hpf, and *Mchlrb* and *Hps5* at 72 hpf. The reduced-expression of *Oca2* was anticipated as it was previously identified as the locus governing albinism in Pachón and Molino cave populations (Protas et al. 2005). It is fascinating that Pachón (16.356 down) and Tinaja (11.361 down) have potentially evolved degenerative pigmentation through the same gene, but through different mutations. These include large coding sequence deletions and prospective changes in gene regulation, respectively. Further, the reduction in *Mitfa* expression across both cave populations is curious, considering that this is a highly conserved transcription factor specific to the melanocortin signaling pathway, whereas reduced expression of *Mitfa* could contribute to overall reduction of downstream pigmentation genes (Shibahara et al. 2001; Levy et al. 2006; Fig. S2). Finally, we observed convergent, reduced expression in genes related to processes where the functional role in cave evolution is unknown. One example includes the gene *Soul4* (predicted to function in heme binding), which demonstrates significantly lower expression maintained across

the four stages of development in both Pachón and Tinaja. Future studies will clarify the precise role of *Soul4* in cavefish.

In contrast to the striking regressive traits, cave populations have evolved constructive characters that help them thrive and forage for food. Our RNA-seq studies across early development captured this convergence by detecting several genes similarly over-expressed in Pachón and Tinaja cavefish (Table 1, Table S5). These constructive characters include the dramatic expansion of several non-visual sensory systems such as enhanced lateral line, increased numbers of taste buds, heightened odor sensitivity, increased fat reserves and resistance to starvation (Jeffery 2009; Gross 2012). Our analyses revealed increased expression with a gene involved with taste (*Tas2r11*) at 10 hpf. Studies in other systems such as humans, *Drosophila* and zebrafish suggest that *Tas2r* genes are related to bitter taste receptors and further evidence suggests that T2Rs have the ability to detect aversive tastants and diverse organic compounds (Oike et al. 2007; Aihara et al. 2008). Additionally, two loci contributing to sensory perception of smell were over-expressed in cavefish lineages at 72 hpf include the fish-specific odorant receptors *Or107-1* and *Or115-15*. The heterologous expression of receptor proteins in a baculovirus/Sf9 cell system showed that Or115 receptor types responded to the potent odorant pyrazine (Breer et al. 1998; Lévai et al. 2006). Additionally, increased deposition of yellow-orange carotenoid pigments has been previously documented in Pachón (Culver and Wilkens 2000) and observed at a modest level in Tinaja cavefish relative to surface fish. This is supported by our study as we detected convergent patterns of increased expression in *Bco2a* across all four stages and *Bco2l* at 72 hpf. With this, numerous “yellow” mutant phenotypes are observed in cow, sheep and chicken as a consequence of alterations affecting *Bco2* (Reviewed in Wang 2012). Another notable character exhibited in subterranean fish populations is the accumulation

of subdermal adipose/fat tissue – this modification is so remarkable among fish living in total darkness that it is accompanied by a change to the overall body contour (Rasquin 1947). We observed a dramatic over-expression in the gene *Alpi.2* in Pachón (245.347 up) and Tinaja (132.464 up) cavefish compared to the surface form. Remarkably, *Alpi* has been implicated as a candidate gene in The Human Obesity Gene Map project with an association of “accelerated weight gain on a high-fat diet” (Narisawa et al. 2003; Rankinen et al. 2006). The changes related to fat content is likely a consequence of their high-fat diet (e.g., invertebrates; Mitchell et al. 1977; Hüppop 1987; Hüppop 2000), which help cave-dwelling fish resist starvation due to low nutrient/food availability of the subterranean environment. Interestingly, Pachón and Tinaja cave morphs demonstrated convergent levels of increased expression of numerous immune-related genes, such as *Ccl14l*, *Hla-F10al*, *Tnfa*, *Cxcl10l*, *Mr1* and *Ifnphil*, possibly in response to the differences in the cave habitat relative to the surface rivers and streams. Additionally, distinct cave-dwelling *Astyanax* populations have similarly evolved craniofacial aberrations including fragmentation and fusion of various bones of the suborbital (SO) series (Gross et al. 2014). Our results identified a single gene related to bone formation that is over-expressed in both Pachón (11.901 up) and Tinaja (10.820 up) cave populations at 72 hpf called *Osteoglycin*. This gene *Ogn* functions as a glycoprotein inhibiting osteoclast formation and has been shown to induce ectopic bone formation when interacting with transforming growth factor beta (Kukita et al. 1990; Hu et al. 2005; Hynek et al. 2011). Although we describe a number of genes depicting convergent expression profiles in both Pachón and Tinaja cavefish, we discovered that each individual population uniquely expresses certain transcripts at dramatically different levels.

Many genes demonstrate divergent patterns of expression specific to each cavefish population

The overall appearance of many different *Astyanax* cavefish populations seem similar, yet apparent differences such as distinct evolutionary histories, complementation tests for eye loss and fine-scale analyses (e.g., reveal modest melanophore pigmentation) render each cavefish lineage unique. With this, notable phenotypic and physiological variation is often accompanied by genetic changes, including but not limited to, alterations in gene expression specific to each cave. Among these distinct differences, Pachón and Tinaja cave populations demonstrated divergent expression patterns in numerous eye and vision-related genes. Across the four stages assayed, over 20 genes that showed severely reduced expression unique to Pachón cavefish, including *Btk* and *Shisa6* at 10 hpf, *Tmtops2a*, *Cryba1a*, *Vsx2*, and *Lim2.2* at 24 hpf, *Pde6h*, *Atoh7*, *Nrl*, *Irbp*, and *Six7* at 36 hpf, and *Opn4.1*, *Rgrb*, *Impg1a* and *Cryball2* at 72 hpf (Table 1, Table S6). Of these, *Cryba1a* and *Pde6* also showed reduced expression in Pachón cavefish in a prior 454 transcriptome analysis (Gross et al. 2013). Additionally, *Lim2* demonstrated was closely associated with vision QTL in a 2013 mapping study (O'Quin et al. 2013). Yet, our analysis did discover many new candidate genes yielding intriguing expression profiles that may help uncover the genetic basis for eye loss in cavefish, which is still unknown. Our findings include the gene *sine oculis homeobox homolog 7* (*Six7*), which depicts divergent reduced expression in Pachón cavefish (59.027 down at 36 hpf) and has been shown to be expressed during embryonic retina morphogenesis in camera-type eyes, and further, in the retinal outer layer (retina rod cells) in zebrafish (Saade et al. 2013). Moreover, the simultaneous loss of both *Six7* and *Six3b* resulted in microphthalmia or anophthalmia in zebrafish mutants (Inbal et al. 2007). Further, the reduction in multiple lens-related genes is interesting, it is important to note that some of the observed reduced expression of crystallins (e.g., *Cryball*) in cavefish may not necessarily be the

cause of eye defects in cavefish, but instead, could be a consequence of the of the reduced size of the lens due to the onset of apoptosis at ~36 hpf (Hinaux et al. 2015). Interestingly, Tinaja cavefish also showed dramatically reduced expression in eye-related genes, however these differences were in 15 *different* loci that varied by developmental stage. These genes included *Unc119b*, *Opn5*, *Socs1a*, *Tmtops2b*, and *Vsx2* at 10 hpf, *Opn4b*, *Hsf4*, *Opn1mw4*, *Crybb111*, *Pde6g* and *Pax6l* at 24 hpf, *Sagb*, *Lim2.4*, *Crygm5* and *Pde6c* at 36 hpf, and *Rp1*, *Lim2.1*, *Elavl1*, *Crygn2*, *Irx7* and *Ndr2* at 72 hpf (Table 1; Table S7). Some of these genes (*Pde6*, *Pax6*, *Lim2* and *Crygn2*) have been previously implicated in Pachón-related *Astyanax* studies (Gross et al. 2008; Jeffery and Strickler 2010; Gross et al. 2013; O'Quin et al. 2013). However, some of the fascinating genes that we found include *Vsx2* and *Sagb*, of which are associated with microphthalmia and retinitis pigmentosa, respectively, in several vertebrate systems (Phelan and Bok 2000; Dejneka and Bennett 2001; Thisse and Thisse 2004; Iseri et al. 2010). Curiously, the results of our expression analyses further support previous complementation studies, including a Pachón x Tinaja cross (both eyeless cavefish) that yielded “eyed” hybrid offspring (Borowsky 2008), thus confirming that *Astyanax* cave-dwelling populations converge on eye loss through different genetic mechanisms.

As noted, the severe reduction in body pigmentation is among the most conspicuous and classically studied traits *Astyanax* cave-dwelling populations. Additionally, the genetic basis for two degenerative pigmentation traits has been identified as *Oca2* (albinism) in the Pachón and Molino caves and *Mc1r* (*brown*) in Pachón and Yerbaníz/Japonés subterranean lineages (Protas et al. 2005; Gross et al. 2009). However, the loci contributing to reduced color in other subterranean populations (e.g., Tinaja) remains unknown. With this, we characterized the divergent expression patterns (≥ 10 fold differences relative to surface) in numerous color-related

genes. In Pachón cavefish, we detected reduced expression in seven pigmentation genes, including *Pmela* and *Csflra* at 24 hpf, *Smtlb* and *Pomca* at 36 hpf, and *Mc4r*, *Agrp* and *Tyr* at 72 hpf (Table 1; Table S6). As mentioned, Pachón cavefish express albinism as the result of a large exon deletion in *Oca2*, but interestingly, our findings show that many other genes such as *Tyr* (12.772 down) involved in the melanocortin signaling pathway also seem affected (García-Borron et al. 2005; Protas et al. 2005). These results may suggest the accumulation of genetic alterations in multiple loci contributing to pigmentation in Pachón cavefish as a consequence neutral mutation (Protas et al. 2007; Gross 2012). Yet in parallel, Tinaja cave-dwelling fish similarly harbor severely reduced expression but in eight *different* genes. These distinct loci involve *Mchlra* at 10 hpf, *Pomcb*, *Mlpha* and *Pmchl* at 24 hpf, *Mchlrb*, *Agrp* and *Mclr* at 36 hpf, and *Mchr2* at 72 hpf (Table 1; Table S7). This includes the gene *Mclr* (58.100 down), which has a known function in *Astyanax* degenerative pigmentation in other cave populations (Gross et al. 2009), but a previously unappreciated role in Tinaja. One intriguing gene that we identified called *Mlpha* (102.456 down) participates in melanocyte differentiation and melanosome localization and is linked to specific pigmentation phenotypes in other systems such as aggregated melanosomes/disrupted pigment granule dispersion in zebrafish (Thisse and Thisse 2004), dilute coat color in cats (Ishida et al. 2006), and Griscelli syndrome type 3 (light hair and skin color) in humans (Ménasché et al. 2005). Nonetheless, many additional loci contributing to other physiological functions, such as circadian rhythms, immune system, cartilage development, etc., depict divergent expression profiles unique to each cave. Additional interesting genes are summed in Table 1 and Supplementary Tables S3-S4.

Coordinately, Pachón and Tinaja cavefish also demonstrate divergent patterns of over-expression that are specific to each subterranean population. Pachón cave-dwelling fish reveal

increased expression in many non-visual sensory traits, such as taste: *Vn1r1* at 10 hpf and *Tas2r11* at 72 hpf, mechanical stimulus/neuromast development: *Trpa1a* at 10 hpf and *Wnt5b* at 24 hpf, olfaction: *Dlg4b* at 24 hpf and *Olfm4* at 36 hpf, and auditory sensation: *Dfnb59* at 10 hpf and *Myo6b* at 24 hpf (Table 1; Table S8). Curiously, *Trpa1a* (115.042 up) has previously been shown to contribute to chemosensation in zebrafish (Prober et al. 2008), thermal nociception in *Drosophila* (Neely et al. 2011) and episodic pain syndrome familial 1 in rats and humans (Kremeyer et al. 2010; Chen et al. 2011). Moreover, the gene *Dfnb59* (37.118 up) is uniquely over-expressed in Pachón, yet when inhibited in humans and mice, causes hearing loss due hair cell dysfunction of the inner ear (Delmaghani et al. 2006; Ebermann et al. 2007). Similarly, Tinaja cavefish demonstrate dramatically heightened expression in extrasensory processes, but in a unique set of genes varying by developmental timepoint (Table 1; Table S9). With this, we detected divergent over-expression in many olfaction genes, including *Or104-2* at 36 hpf and *Or101-1*, *Olfm4*, *Or132-5*, *Or115-15* and *Ora4* at 72 hpf, several taste loci, such as *Cystltr2* at 36 hpf and *Tas2r202*, *Tas1r3*, and *Tas1r2.1* at 72 hpf, and one mechanosensory associated gene *Scn8aa* at 72 hpf. A noteworthy gene called *Ora4* (21.935 up) exhibits increased expression specific to Tinaja cavefish, and prior studies using *in situ* hybridization show that *Ora4* expression is detected in zebrafish olfactory epithelium (Saraiva and Korsching 2007). Additionally, *Scn8aa* (80.723 up) is an intriguing gene considering the enhanced lateral line system in *Astyanax* cavefish, and when disrupted, demonstrates altered mechanosensory sensation and decreased larval locomotory behavior in zebrafish (Chen et al. 2008; Low et al. 2010). Finally as anticipated, other genes related to additional biological processes such as adipose tissue storage, response to starvation, visual perception, etc. show divergent profiles of increased expression respective to each cave, as summarized in Table 1 and Tables S8-S9.

Moreover, to confirm that the diverse expression profiles observed in our RNA-seq comparisons do indeed reflect the appropriate gene expression pattern, we validated the normalized expression pattern for a subset of our experimental genes at 72 hpf using both qualitative and qPCR analyses (Fig. S1). With this, our study aims to highlight an array of intriguing genes with prospective roles in cave evolution based on the compelling convergent and divergent expression profiles in Pachón and Tinaja cavefish populations. Yet, future functional studies may continue to clarify the precise role of the genes we describe in cave-dwelling *Astyanax*. Further, it is important to note that this report simply characterizes a small subset of genes, however it is likely that many additional loci and gene expression changes contribute the diverse cave-associated traits observed in *Astyanax mexicanus*.

Pathway analysis reveals Otx2 and Mitf as possible upstream regulators contributing to regressive cave-associated traits

Although our comparative expression analyses identified many individual gene candidates that demonstrate convergent or divergent expression patterns, we sought to characterize key genes that may regulate other downstream loci, and further, could help explain the broad-scale expression changes observed in our dataset. For this analysis, we utilized the well-documented program Ingenuity Pathway Analysis. Of our 23,719 transcripts analyzed, a total of 17,897 genes were able to map IDs with extensive annotation information. Although the possibilities of analysis using this program are endless, we focus our attention on the ‘IPA Upstream Regulator’ analytic to help reveal the cascade of upstream transcriptional regulators that may explain the broader expression changes observed in our dataset. This particular analysis

capitalizes on extensive prior literature denoting expected effects – both direct and indirect relationships – between regulators and their downstream target genes. Moreover, the upstream regulator analysis can statistically predict the ‘activation state’ of the transcriptional regulator as “activated” or “inhibited” based upon the expression patterns exhibited by target genes in dataset (Methods).

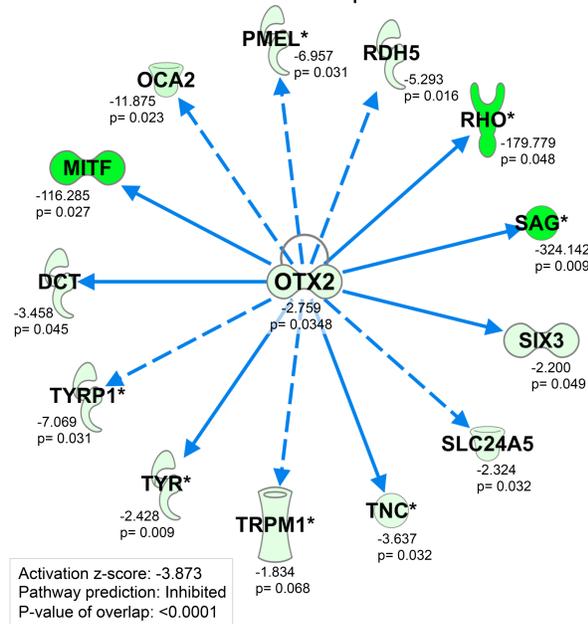
We concentrated our efforts in identifying potential upstream regulators affecting eye degeneration and reduced pigmentation. We characterized regulators based upon expression patterns in Pachón and Tinaja cave lineages, relative to surface at 36 hpf. At this critical time point, melanin coloration begins to appear in surface fish but not in cave, and cavefish eyes actively regress (Jeffery 2005; Jeffery 2009; Hinaux et al. 2011; Hinaux et al. 2015). Therefore, the any effects of an altered upstream regulator at this developmental stage may impact the observed morphotype-specific phenotypes. In Pachón cavefish, our analysis identified 40 prospective upstream regulators as “activated” (activation z-score > 2.0) and 53 potential “inhibited” regulators (activation z-score < -2.0), and in parallel, implicated 86 activated and 56 inhibited regulators in Tinaja cavefish.

Interestingly, the overall top upstream “inhibited” regulator in Pachón was *orthodenticle homeobox 2* (*Otx2*; fold change = -2.759, predicted activation state = inhibited, z-score = -3.873, p-value < 0.0001; Fig. 3A). Similarly, we also detected *Otx2* as an “inhibited” regulator in Tinaja cavefish (fold change = -1.543, predicted activation state = inhibited, z-score = -2.673, p-value < 0.0001; Fig. 3B). Curiously, *Otx2* is well-documented transcriptional regulator, and when altered in other systems such as humans and zebrafish, the mutated form causes retinal dystrophy and microphthalmia (decreased eye size; Wyatt et al. 2008; Chassaing et al. 2012; Vincent et al. 2014). A whole-mount *in situ* hybridization series (10 hpf – 48 hpf) in *Astyanax mexicanus*

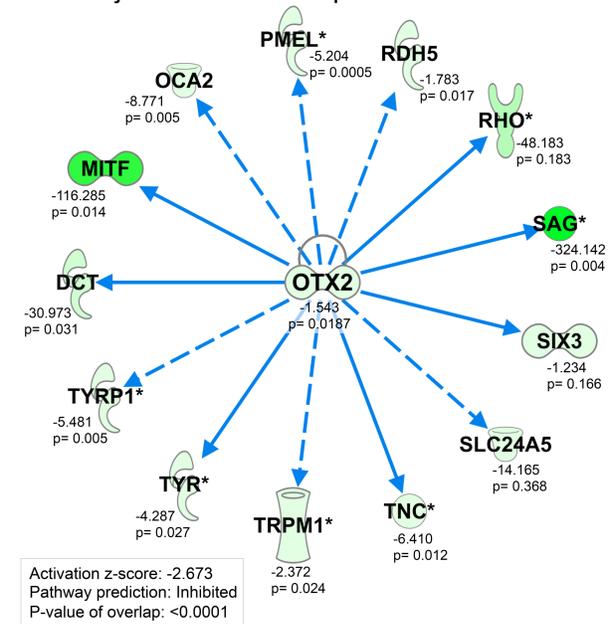
demonstrated an overall decrease of *Otx2* gene expression in Pachón cavefish compared to surface fish. Further, *in situ* sections of both morphotypes at 48 hpf reveal severely reduced *Otx2* expression in the cavefish lens (McGaugh et al. 2014). Some of the downstream targets of *Otx2* include additional eye-related genes and several loci with known roles in pigmentation processes, including *Oca2* (Fig. 3A-B).

We also screened our results to identify inhibited targets with pigmentation-related functions. In our analysis, we uncovered *Mitf* – a transcription factor known to govern diverse pigment-related processes – in both Pachón (fold change = -116.285, predicted activation state = inhibited, z-score = 2.237, p-value < 0.0001; Fig. S2A) and Tinaja (fold change = -116.285, predicted activation state = inhibited, z-score = -2.937, p-value < 0.0001; Fig. S2B), which includes many downstream targets such as *Mclr*. Together, these results suggest that certain regulators and pathways may be repeatedly altered in independent cavefish lineages. Further, the pleiotropic consequences of inhibited upstream factors could contribute to the numerous cave-adapted traits observed in *Astyanax* cave-dwelling fish.

A. Pachón cavefish – 36 hpf



B. Tinaja cavefish – 36 hpf



Legend:
 More extreme (green box) Modest changes (light green box) Down-regulated (blue box) Predicted inhibition (solid blue arrow) Predicted inhibition – direct (dashed blue arrow) Predicted inhibition – indirect (dotted blue arrow)

Figure 3. Pathway analysis suggests *Otx2* as a potential upstream regulator governing

cavefish evolution. Due to the complexity of these large RNA-seq datasets, it can be difficult to explain broad-scale expression changes present in a specific population at a given time point. To address this, we employ the Ingenuity Pathway Analysis (IPA) – a program that capitalizes on the extensive knowledge of known gene functions and relationships – to help identify upstream gene regulators that may justify the observed expression patterns of downstream target loci within the experimental dataset. Using this analysis, we independently detected *Otx2* as a potential upstream regulator that is inhibited in both Pachón (A) and Tinaja (B) cavefish at 36 hpf.

CONCLUSIONS

This study aimed to identify potential genes contributing to cave adaptation by characterizing measurable changes in gene expression across four stages of early development in two geographically isolated cavefish populations compared to surface. From these comparisons, we discovered that Pachón and Tinaja cavefish have evolved dramatic changes in gene expression in some of the same loci, possibly as a result of either gene flow or selection. This may suggest a common genetic basis for the overall similarity in morphology observed in *Astyanax* cavefish. However, our analyses also revealed expression differences exclusive to each cavefish population, suggesting subtle differences in subterranean habitats and phylogenetic histories may lead to cave-specific gene expression differences. This study further indicates that these remarkable animals manifest the spectrum of cave-adapted traits through a combination of both convergent and divergent genetic mechanisms, and identified numerous genes that may contribute to trait evolution in natural populations.

FUNDING

This study was funded by the National Science Foundation, Washington D.C., U.S.A. (DEB-1457630) and the National Institutes of Health (NIDCR; R01DE025033).

ACKNOWLEDGEMENTS

The authors wish to thank Brian Carlson and Amanda K. Powers of the Gross lab for assistance with embryo collection and helpful discussions. We also thank Dr. Richard Borowsky (New York University) for generously providing the adults used for breeding in this study. In addition, we are grateful for the guidance of Dr. Suzanne McGaugh (University of Minnesota) with the pathway analysis and for the resources of the University of Minnesota Supercomputing Institute.

ETHICAL STATEMENT

All applicable international, national, and/or institutional guidelines for the care and use of animals were followed. All procedures performed in studies involving animals were in accordance with the ethical standards of the institution or practice at which the studies were conducted. The protocol was approved by the Institutional Animal Care and Use Committee (IACUC) of the University of Cincinnati (Protocol Number 10-01-21-01).

INFORMED CONSENT

This article does not contain any studies involving human participants performed by any of the authors. Therefore, for this type of study, formal consent is not applicable.

REFERENCES

- Aihara, Y., A. Yasuoka, S. Iwamoto, Y. Yoshida, T. Misaka, and K. Abe (2008) Construction of a taste-blind medaka fish and quantitative assay of its preference–aversion behavior. *Genes, Brain and Behavior* 7: 924–932.
- Bandah-Rozenfeld, D., R. W. J. Collin, E. Banin, L. Ingeborgh van den Born, K. L. M. Coene, A. M. Siemiatkowska, L. Zelinger, M. I. Khan, D. J. Lefeber, I. Erdinest, F. Testa, F. Simonelli, K. Voesenek, E. A.W. Blokland, T. M. Strom, C. C.W. Klaver, R. Qamar, S. Banfi, F. P.M. Cremers, D. Sharon, and A. I. den Hollander (2010) Mutations in *IMPG2*, Encoding Interphotoreceptor Matrix Proteoglycan 2, Cause Autosomal-Recessive Retinitis Pigmentosa. *The American Journal of Human Genetics* 87: 199–208.
- Benjamini, Y., and Y. Hochberg (2000) The adaptive control of the false discovery Rate in Multiple Testing With Independent Statistics. *Journal of Educational and Behavioral Statistics* 25(1): 60–83.
- Borowsky, R. (2008) Restoring sight in blind cavefish. *Current Biology* 18: R23–R24.
- Bradic, M., P. Beerli, F. J. G.-D. León, S. Esquivel-Bobadilla, and R. L. Borowsky (2012) Gene flow and population structure in the Mexican blind cavefish complex (*Astyanax mexicanus*). *BMC Evolutionary Biology* 12: 9.
- Breer, H., J. Krieger, C. Meinken, H. Kiefer, and J. Strotmann (1998) Expression and Functional Analysis of Olfactory Receptors. *Annals of the New York Academy of Sciences* 855: 175–181.

- Chassaing, N., S. Sorrentino, E. E. Davis, D. Martin-Coignard, A. Iacovelli, W. Paznekas, B. D. Webb, O. Faye-Petersen, F. Encha-Razavi, L. Lequeux, A. Vigouroux, A. Yesilyurt, S. A. Boyadjiev, H. Kayserili, P. Loget, D. Carles, C. Sergi, S. Puvabanditsin, C. Chen, H. C. Etchevers, N. Katsanis, C. L. Mercer, P. Calvas, and E. W. Jabs (2012) *OTX2* mutations contribute to the otocephaly-dysgnathia complex. *Journal of Medical Genetics* 49: 373–379.
- Chen, J., S. K. Joshi, S. DiDomenico, R. J. Perner, J. P. Mikusa, D. M. Gauvin, J. A. Segreti, P. Han, X. Zhang, W. Niforatos, B. R. Bianchi, S. J. Baker, C. Zhong, G. H. Simler, H. A. McDonald, R. G. Schmidt, S. P. McGaraughty, K. L. Chu, C. R. Faltynek, M. E. Kort, R. M. Reilly, and P. R. Kym (2011) Selective blockade of *TRPA1* channel attenuates pathological pain without altering noxious cold sensation or body temperature regulation. *Pain* 152: 1165–1172.
- Chen, Y.-H., F.-L. Huang, Y.-C. Cheng, C.-J. Wu, C.-N. Yang, and H.-J. Tsay (2008) Knockdown of zebrafish *Nav1.6* sodium channel impairs embryonic locomotor activities. *Journal of Biomedical Science* 15: 69–78.
- Culver, D. C. (1982) *Cave life: evolution and ecology*. Cambridge: Harvard University Press 189 p.
- Culver, D. C., and H. Wilkens (2000) *Critical review of the relevant theories of the evolution of subterranean animals*. Ecosystems of the world.
- Dejneka, N. S., and J. Bennett (2001) Gene therapy and retinitis pigmentosa: advances and future challenges. *BioEssays* 23: 662–668.
- Delmaghani, S., F. J. del Castillo, V. Michel, M. Leibovici, A. Aghaie, U. Ron, L. Van Laer, N.

- Ben-Tal, G. Van Camp, D. Weil, F. Langa, M. Lathrop, P. Avan and C. Petit (2006) Mutations in the gene encoding pejvakin, a newly identified protein of the afferent auditory pathway, cause DFNB59 auditory neuropathy. *Nature Genetics* 38: 770–778.
- Ebermann, I., M. Walger, H. P. N. Scholl, P. C. Issa, C. Lüke. G. Nürnberg, R. Lang-Roth, C. Becker, P. Nürnberg and H. J. Bolz (2007) Truncating mutation of the *DFNB59* gene causes cochlear hearing impairment and central vestibular dysfunction. *Human Mutation* 28: 571–577.
- Emerling, C. A., and M. S. Springer (2014) Eyes underground: Regression of visual protein networks in subterranean mammals. *Molecular Phylogenetics and Evolution* 78: 260–270.
- Garcia-Borron, J. C., B. L. Sanchez-Laorden, and C. Jimenez-Cervantes (2005) Melanocortin-1 receptor structure and functional regulation. *Pigment Cell Research* 18(6): 393–410.
- Gross, J. B. (2012a) *Cave Evolution*. John Wiley & Sons, Ltd, Chichester, UK.
- Gross, J. B. (2012b) The complex origin of *Astyanax* cavefish. *BMC Evolutionary Biology* 12: 105–22.
- Gross, J. B., and H. Wilkens (2013) Albinism in phylogenetically and geographically distinct populations of *Astyanax* cavefish arises through the same loss-of-function *Oca2* allele. *Heredity* 111: 122–130.
- Gross, J. B., R. Borowsky, and C. J. Tabin (2009) A Novel Role for *Mcl1r* in the Parallel Evolution of Depigmentation in Independent Populations of the Cavefish *Astyanax mexicanus*. *PLoS Genetics* 5(1): e1000326.

- Gross, J. B., A. Furterer, B. M. Carlson, and B. A. Stahl (2013) An Integrated Transcriptome-Wide Analysis of Cave and Surface Dwelling *Astyanax mexicanus*. *PLoS ONE* 8: e55659–14.
- Gross, J. B., A. J. Krutzler, and B. M. Carlson (2014) Complex Craniofacial Changes in Blind Cave-Dwelling Fish Are Mediated by Genetically Symmetric and Asymmetric Loci. *Genetics* 196: 1303–1319.
- Gross, J. B., M. M. Protas, M. Conrad, P. E. Scheid, O. Vidal, W. R. Jeffery, R. Borowsky, and C. J. Tabin (2008) Synteny and candidate gene prediction using an anchored linkage map of *Astyanax mexicanus*. *Proceedings of the National Academy of Sciences* 105: 20106–20111.
- Guaiquil, V. H., Z. Pan, N. Karagianni, S. Fukuoka, G. Alegre, and M. I. Rosenblatt (2014) VEGF-B selectively regenerates injured peripheral neurons and restores sensory and trophic functions. *Proceedings of the National Academy of Sciences* 111: 17272–17277.
- Hinaux, H., M. Blin, J. Fumey, L. Legendre, A. Heuzé, D. Casane and S. Rétaux (2015) Lens defects in *Astyanax mexicanus* Cavefish: Evolution of *crystallins* and a role for *alphaA-crystallin*. *Developmental Neurobiology* 75: 505–521.
- Hinaux, H., K. Pottin, H. Chalhoub, S. Père, Y. Elipot, L. Legendre, and S. Rétaux (2011) A Developmental Staging Table for *Astyanax mexicanus* Surface Fish and Pachón Cavefish. *Zebrafish* 8(4): 155–165.
- Hu, Z., S. A. F. Peel, S. K. C. Ho, G. K. B. Sándor, and C. M. L. Clokie (2005) Role of Bovine Bone Morphogenetic Proteins in Bone Matrix Protein and Osteoblast-Related Gene Expression During Rat Bone Marrow Stromal Cell Differentiation. *Journal of Craniofacial*

Surgery 16: 1006–1014.

Hüppop, K. (1987) Food-finding ability in cave fish (*Astyanax fasciatus*). *International Journal of Speleobiology* 16: 59–66.

Hüppop, K. (1989) Genetic analysis of oxygen consumption rate in cave and surface fish of *Astyanax fasciatus* (Characidae, Pisces), further support for the neutral mutation theory. *Mémoires de Biospéologie* 16: 163-168.

Hüppop, K. (2000) *How do cave animals cope with the food scarcity in caves?* Ecosystems of the world.

Hynek, R., S. Kuckova, P. Konik, and R. Prchlikova (2011) A novel approach to protein analysis in hard tissues. *Journal of Surface Analysis* 17(3): 310–313.

Inbal, A., S-H. Kim, J. Shin, and L. Solnica-Krezel (2007) *Six3* Represses Nodal Activity to Establish Early Brain Asymmetry in Zebrafish. *Neuron* 55: 407–415.

Iseri, S. U., A. W. Wyatt, G. Nürnberg, C. Kluck, P. Nürnberg, C. Kluck, P. Nürnberg, G. E.

Holder, E. Blair, A. Salt, and N. K. Ragge (2010) Use of genome-wide SNP homozygosity mapping in small pedigrees to identify new mutations in *VSX2* causing recessive microphthalmia and a semidominant inner retinal dystrophy. *Human Genetics* 128: 51–60.

Ishida, Y., V. A. David, E. Eizirik, A. A. Schäffer, B. A. Neelam, M. E. Roelke, S. S. Hannah, S.

J. O'Brien, and M. Menotti-Raymond (2006) A homozygous single-base deletion in *MLPH* causes the dilute coat color phenotype in the domestic cat. *Genomics* 88: 698–705.

Jeffery, W. R. (2005) Adaptive Evolution of Eye Degeneration in the Mexican Blind Cavefish.

Journal of Heredity 96: 185–196.

Jeffery, W. R. (2001) Cavefish as a Model System in Evolutionary Developmental Biology.

Developmental Biology 231: 1–12.

Jeffery, W. R. (2009) *Evolution and Development in the Cavefish Astyanax*. Evolution and

Development, Current Topics in Developmental Biology, Elsevier. p 191–221.

Jeffery, W. R., and A. G. Strickler (2010) *Development as an evolutionary process in Astyanax*

cavefish. Biology of Subterranean Fishes. Enfield: Science Publishers. p. 141-168

Kowalko, J. E., N. Rohner, S. B. Rompani, B. K. Peterson, T. A. Linden, M. Yoshizawa, E. H.

Kay, J. Weber, H. E. Hoekstra, W. R. Jeffery, R. Borowsky, and C. J. Tabin (2013) Loss of

Schooling Behavior in Cavefish through Sight-Dependent and Sight-Independent

Mechanisms. *Current Biology* 23: 1874–1883.

Kremeyer, B., F. Lopera, J. J. Cox, A. Momin, F. Rugiero, S. Marsh, C. G. Woods, N. G. Jones,

K. J. Paterson, F. R. Fricker, A. Villegas, N. Acosta, N. G. Pineda-Trujillo, J. D. Ramírez, J.

Zea, M.-W. Burley, G. Bedoya, D. L. H. Bennett, J. N. Wood, and A. Ruiz-Linares (2010) A

Gain-of-Function Mutation in *TRPA1* Causes Familial Episodic Pain Syndrome. *Neuron* 66:

671–680.

Kukita, A., L. Bonewald, D. Rosen, S. Seyedin, G. R. Mundy, and G. D. Roodman (1990)

Osteoinductive factor inhibits formation of human osteoclast-like cells. *Proceedings of the*

National Academy of Sciences 87: 3023–3026.

la O Leyva-Pérez, de, M., A. Valverde-Corredor, R. Valderrama, J. Jiménez-Ruiz, A. Muñoz-

- Merida, O. Trelles, J. B. Barroso, J. Mercado-Blanco, and F. Luque (2014) Early and delayed long-term transcriptional changes and short-term transient responses during cold acclimation in olive leaves. *DNA Research* 22: dsu033–11.
- Levy, C., M. Khaled, and D. E. Fisher (2006) MITF: master regulator of melanocyte development and melanoma oncogene. *Trends in Molecular Medicine* 12: 406–414.
- Lévai, O., T. Feistel, H. Breer, and J. Strotmann (2006) Cells in the vomeronasal organ express odorant receptors but project to the accessory olfactory bulb. *Journal of Comparative Neurology* 498: 476–490.
- Low, S. E., W. Zhou, I. Choong, L. Saint Amant, S. M. Sprague, H. Hirata, W. W. Cui, R. I. Hume, and J. Y. Kuwada (2010) *NaVI.6a* is required for normal activation of motor circuits normally excited by tactile stimulation. *Developmental Neurobiology* 70: 508–522.
- McGaugh, S. E., J. B. Gross, B. Aken, M. Blin, R. Borowsky, D. Chalopin, H. Hinaux, W. R. Jeffery, A. Keene, L. Ma, P. Minx, D. Murphy, K. E. O'Quin, S. Rétaux, N. Rohner, S. M. J. Searle, B. A. Stahl, C. Tabin, J.-N. Volff, M. Yoshizawa, W. C. Warren (2014) The cavefish genome reveals candidate genes for eye loss. *Nature Communications* 5: 5307.
- Meng, F., I. Braasch, J. B. Phillips, X. Lin, T. Titus, C. Zhang, and J. H. Postlethwait (2013) Evolution of the eye transcriptome under constant darkness in *Sinocyclocheilus* cavefish. *Molecular biology and evolution* 30(7): 1527–1543.
- Ménasché, G., C. H. Ho, O. Sanal, J. Feldmann, I. Tezcan, F. Ersoy, A. Houdusse, A. Fischer, and G. de Saint Basile (2005) Griscelli syndrome restricted to hypopigmentation results from a *melanophilin* defect (GS3) or a *MYO5A* F-exon deletion (GS1). *Journal of Clinical*

Investigation 12(3), 450.

Mitchell, R. W., W. H. Russell, and W. R. Elliott (1977) *Mexican eyeless characin fishes, genus Astyanax: environment, distribution, and evolution*. Texas Tech Press.

Mortazavi, A., B. A. Williams, K. McCue, L. Schaeffer, and B. Wold (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* 5: 621–628.

Narisawa, S., L. Huang, A. Iwasaki, H. Hasegawa, D. H. Alpers, and J. L. Millán (2003) Accelerated Fat Absorption in Intestinal Alkaline Phosphatase Knockout Mice. *Molecular and Cellular Biology* 23: 7525–7530.

Neely, G. G., A. C. Keene, P. Duchek, E. C. Chang, Q.-P. Wang, Y. A. Aksoy, M. Rosenzweig, M. Costigan, C. J. Woolf, P. A. Garrity, and J. M. Penninger (2011) *TrpA1* Regulates Thermal Nociception in *Drosophila*. *PLoS ONE* 6: e24343.

O'Quin, K. E., M. Yoshizawa, P. Doshi, and W. R. Jeffery (2013) Quantitative Genetic Analysis of Retinal Degeneration in the Blind Cavefish *Astyanax mexicanus*. *PLoS ONE* 8: e57281–11.

Oike, H., T. Nagai, A. Furuyama, S. Okada, Y. Aihara, Y. Ishimaru, T. Marui, I. Matsumoto, T. Misaka, and K. Abe (2007) Characterization of Ligands for Fish Taste Receptors. *Journal of Neuroscience* 27: 5584–5592.

Phelan, J. K., and D. Bok (2000) A brief review of retinitis pigmentosa and the identified retinitis pigmentosa genes. *Molecular Vision* 6: 116-124

Prober, D. A., S. Zimmerman, B. R. Myers, J. B. M. McDermott, S.-H. Kim, S. Caron, J. Rihel,

- L. Solnica-Krezel, D. Julius, A. J. Hudspeth, and A. F. Schier (2008) Zebrafish *TRPA1* Channels Are Required for Chemosensation But Not for Thermosensation or Mechanosensory Hair Cell Function. *Journal of Neuroscience* 28: 10102–10110.
- Protas, M., M. Conrad, J. B. Gross, C. Tabin, and R. Borowsky (2007) Regressive Evolution in the Mexican Cave Tetra, *Astyanax mexicanus*. *Current Biology* 17: 452–454.
- Protas, M. E., C. Hersey, D. Kochanek, Y. Zhou, H. Wilkens, W. R. Jeffery, L. I. Zon, R. Borowsky, and C. J. Tabin (2005) Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nature Genetics* 38: 107–111.
- Rankinen, T., A. Zuberi, Y. C. Chagnon, S. J. Weisnagel, G. Argyropoulos, B. Walts, L. Perusse, and C. Bouchard (2006) The Human Obesity Gene Map: The 2005 Update. *Obesity* 14: 529–644.
- Rasquin, P. (1947) Progressive pigmentary regression in fishes associated with cave environments. *Zoologica; scientific contributions of the New York Zoological Society* 32(1-8): 35.
- Saade, C. J., K. Alvarez-Delfin, and J. M. Fadool (2013) Rod Photoreceptors Protect from Cone Degeneration-Induced Retinal Remodeling and Restore Visual Responses in Zebrafish. *Journal of Neuroscience* 33: 1804–1814.
- Saraiva, L. R., and S. I. Korsching (2007) A novel olfactory receptor gene family in teleost fish. *Genome Research* 17: 1448–1457.
- Shibahara, S., K. Takeda, K.-I. Yasumoto, T. Udono, K.-I. Watanabe, H. Saito, and K. Takahashi

- (2001) Microphthalmia-Associated Transcription Factor (MITF): Multiplicity in Structure, Function, and Regulation. *Journal of Investigative Dermatology Symposium Proceedings* 6: 99–104.
- Strickler, A. G., and W. R. Jeffery (2009) Differentially expressed genes identified by cross-species microarray in the blind cavefish *Astyanax*. *Integrative Zoology* 4: 99–109.
- Strickler, A. G., K. Famuditimi, and W. R. Jeffery (2002) Retinal homeobox genes and the role of cell proliferation in cavefish eye degeneration. *International Journal of Developmental Biology* 46(3): 285–294.
- Thisse, B., and C. Thisse (2004) *Fast Release Clones: A High Throughput Expression Analysis* ZFIN direct data submission.
- Vincent, A., N. Forster, J. T. Maynes, T. A. Paton, G. Billingsley, N. M. Roslin, A. Ali, J. Sutherland, T. Wright, C. A. Westall, A. D. Paterson, C. R. Marshall, FORGE Canada Consortium, and E. Héon (2014) *OTX2* mutations cause autosomal dominant pattern dystrophy of the retinal pigment epithelium. *Journal of Medical Genetics* 51: 797–805.
- Wang, X.-D. (2012) Lycopene metabolism and its biological significance. *American Journal of Clinical Nutrition* 96: 1214S–1222S.
- Wilkens, H., and U. Strecker (2003) Convergent evolution of the cavefish *Astyanax* (Characidae, Teleostei): genetic evidence from reduced eye-size and pigmentation. *Biological Journal of the Linnean Society* 80: 545–554.
- Wyatt, A., P. Bakrania, D. J. Bunyan, R. J. Osborne, J. A. Crolla, A. Salt, C. Ayuso, R.

Newbury-Ecob, Y. Abou-Rayyah, J. R. O. Collin, D. Robinson, and N. Ragge (2008) Novel heterozygous *OTX2* mutations and whole gene deletions in anophthalmia, microphthalmia and coloboma. *Human Mutations* 29(11): E278–E283.

Youngblood, B. A., and C. C. MacDonald (2014) *CstF-64* is necessary for endoderm differentiation resulting in cardiomyocyte defects. *Stem Cell Research* 13: 413–421.

SUPPLEMENTARY FIGURES

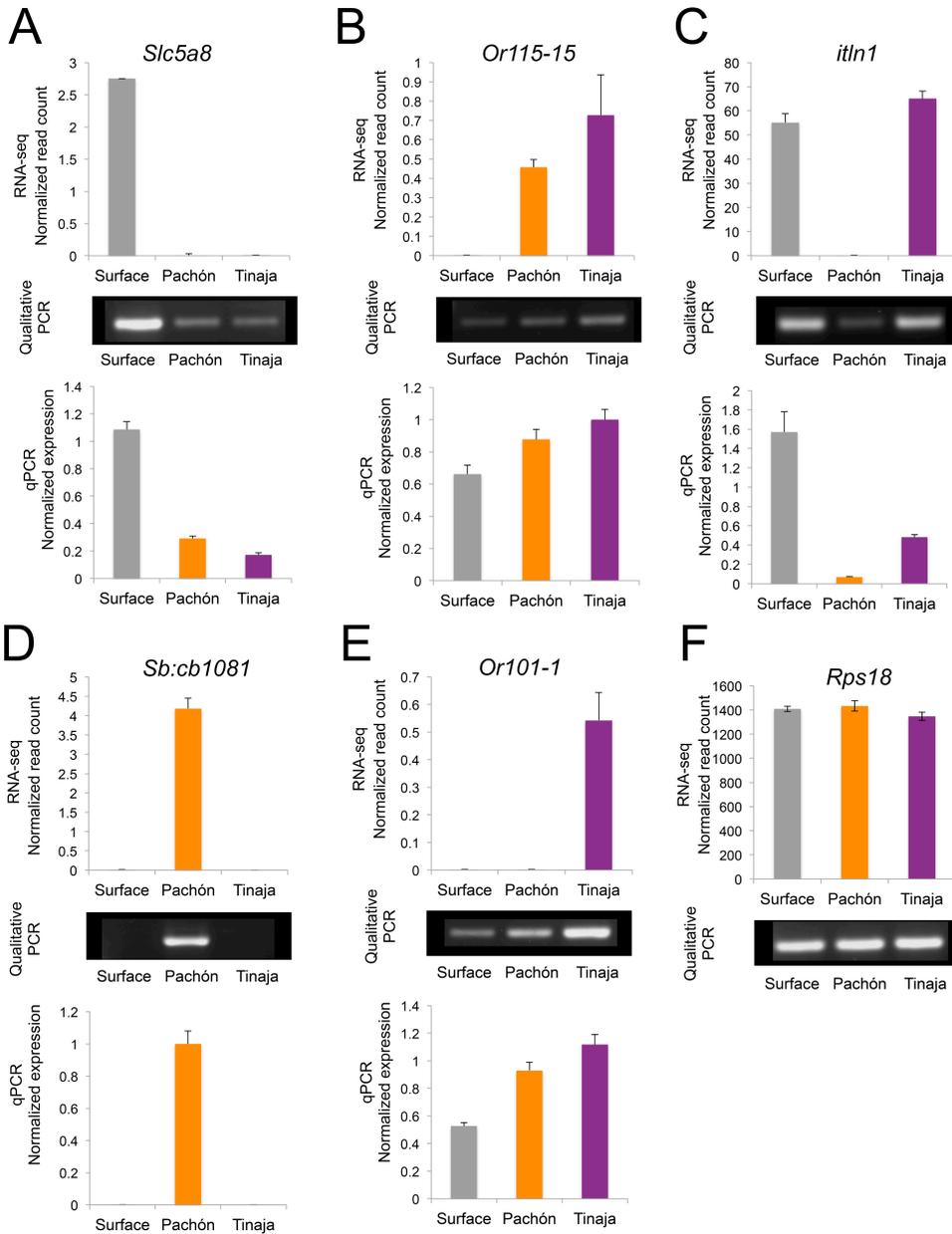
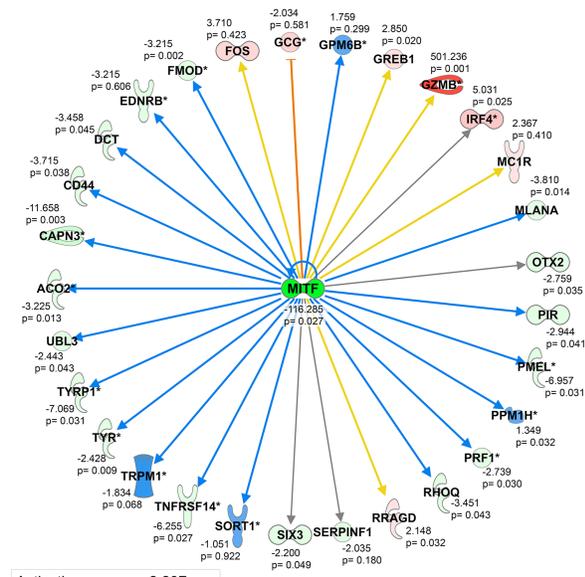


Figure S1. Qualitative and quantitative PCR analyses validate a subset of genes representing the diverse expression profiles at 72 hpf. To confirm that the diverse expression profiles observed in our RNA-seq comparison do indeed reflect the appropriate gene expression

pattern, we validate the normalized expression pattern (top graphs) for a subset of our experimental genes at 72 hpf using both qualitative (gel images) and qPCR analyses (bottom graphs). Two genes *slc5a8* (A) and *or115-15* (B) demonstrate convergent expression in Pachón and Tinaja. In addition, some genes including *itln1* (C) and *sb:cb1081* (D) and *or101-1* (E) depict profiles specific to Tinaja cavefish. We utilize the reference gene *Rps18* for comparison (F).

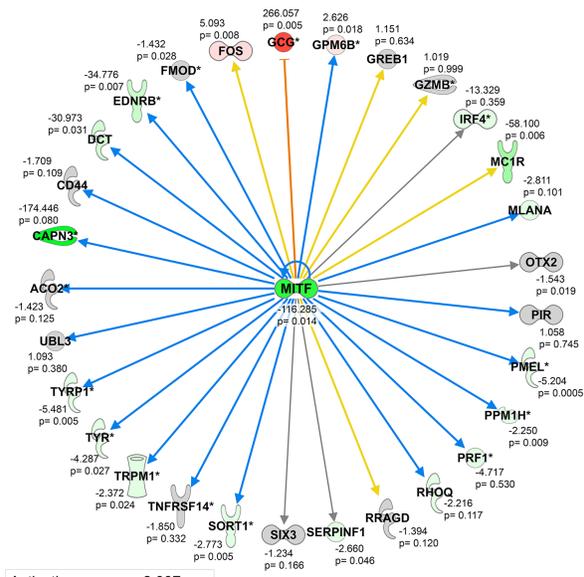
A. Pachón cavefish– 36 hpf



Activation z-score: -2.237
 Pathway prediction: Inhibited
 P-value of overlap: <0.0001



B. Tinaja cavefish– 36 hpf



Activation z-score: -2.937
 Pathway prediction: Inhibited
 P-value of overlap: <0.0001



Figure S2. Pathway analysis implicates Mitf as a potential upstream regulator in Pachón and Tinaja cavefish at 36 hpf. To understand the complex nature of dynamic gene expression changes in our RNA-seq dataset, we utilize Ingenuity Pathway Analysis (IPA) to nominate potential upstream genetic regulators that could explain the observed expression changes in the downstream target loci. In our parallel analyses, we detected Mitf – a known factor that governs many pigmentation-related processes – as a potential regulator that is inhibited in both Pachón (A) and Tinaja (B) cavefish lineages at 36 hpf.

SUPPLEMENTARY TABLES

Table S1. Number of genes differentially expressed in Pachón and Tinaja cavefish relative to surface at varying thresholds.

	Pachón cavefish		Tinaja cavefish	
	Over	Under	Over	Under
10 fold -/+	565	683	1,182	594
8 fold -/+	687	829	1,382	698
6 fold -/+	861	1,087	1,695	858
4 fold -/+	1,193	1,536	2,228	1,124
2 fold -/+	2,115	3,000	3,674	2,032
0.001 -/+	10,896	12,605	14,631	8,847

* Blue indicates under-expressed genes, red represents over-expression.

Table S2. Genes that harbor ≥ 10 -fold expression differences depict both convergent and divergent patterns across development.

	Under-expressed			Over-expressed		
	Pachón	Both	Tinaja	Pachón	Both	Tinaja
10 hpf	120	150	860	1269	365	271
24 hpf	222	299	910	304	101	187
36 hpf	255	291	760	209	67	207
72 hpf	536	188	409	183	136	638
Average	286	279	903	483	200	394

*Values indicate the number of genes differentially expressed ≥ 10 -fold either similarly or unique at each stages.

**Blue indicates under-expressed genes, red represents over-expression.

Table S3. The percentage of genes that harbor ≥ 10 -fold expression differences at each stage depict both convergent and divergent patterns across development.

	Under-expressed				Over-expressed			
	Pachón		Tinaja		Pachón		Tinaja	
	Divergent	Convergent	Divergent	Convergent	Divergent	Convergent	Divergent	Convergent
10 hpf	44.4%	55.6%	85.1%	14.9%	77.7%	22.3%	42.6%	57.4%
24 hpf	42.6%	57.4%	75.3%	24.7%	75.0%	24.9%	64.9%	35.0%
36 hpf	46.7%	53.3%	72.3%	27.7%	75.7%	24.3%	75.5%	24.5%
72 hpf	74.0%	26.0%	68.5%	31.5%	57.4%	42.6%	82.4%	17.6%
Average	50.6%	49.4%	76.4%	23.6%	70.7%	29.3%	66.3%	33.7%

*Values indicate the percentage of genes differentially expressed ≥ 10 -fold for each category across development.

**Blue indicates percentage of under-expressed genes, red represents percentage of over-expression.

Table S4. Genes demonstrating convergent patterns of reduced expression in both Pachón and Tinaja cavefish compared to surface-dwelling fish.

Gene	Transcript ID	Physiological function(s)*	Pachón cavefish (fold change)**	Tinaja cavefish (fold change)**
10 hpf				
<i>Saga</i>	ENSAMXT00000019744	Retina and pineal gland	46.500 down	46.500 down
<i>Pmchl</i>	ENSAMXT00000026184	Pigmentation	32.492 down	32.492 down
<i>Bcmo1</i>	ENSAMXT00000010290	Eye photoreceptor; hindbrain	21.458 down	10.005 down
<i>Mtnr1c</i>	ENSAMXT00000007073	Circadian rhythms; melatonin	14.360 down	14.360 down
<i>Cryba2a</i>	ENSAMXT00000002261	Lens constituent	23.185 down	23.175 down
<i>Or120-1</i>	ENSAMXT00000026305	Odorant receptor	15.869 down	15.869 down
<i>PTGR1</i>	ENSAMXT00000005117	Oxidation-reduction process; zinc ion binding	195.314 down	1130.146 down
<i>Lmo3</i>	ENSAMXT00000011740	Zinc ion binding	224.732 down	187.472 down
<i>Mep1b</i>	ENSAMXT00000009834	Proteolysis; metalloendopeptidase activity	120.050 down	362.535 down
<i>Nr0b1</i>	ENSAMXT00000008950	Transcription; steroid hormone pathway	23.229 down	464.010 down
24 hpf				
<i>Ccl-c5a</i>	ENSAMXT00000001272	Immune response	25.847 down	25.847 down
<i>Pde6h</i>	ENSAMXT00000011074	Visual perception	170.196 down	31.341 down
<i>Gnat2</i>	ENSAMXT00000014437	Visual perception	16.647 down	13.621 down
<i>Oca2</i>	ENSAMXT00000013137	Pigmentation	16.356 down	11.361 down
<i>Soul4</i>	ENSAMXT00000004688	Digestion	367.405 down	3192.621 down
<i>Crygn2</i>	ENSAMXT00000018569	Eye; lens constituent	76.026 down	2418.442 down
<i>Cryba111</i>	ENSAMXT00000013439	Eye; lens constituent	33.936 down	1181.636 down
<i>SPP2</i>	ENSAMXT00000000810	Bone remodeling	10.382 down	229.771 down
<i>Cry1b</i>	ENSAMXT00000020701	Circadian rhythm	142.600 down	24.708 down
<i>Irbp</i>	ENSAMXT00000003213	Photoreceptor	44.374 down	44.374 down
<i>Apip</i>	ENSAMXT00000003914	L-methionine salvage from methylthioadenosine	4571.555 down	298.500 down
36 hpf				
<i>Cryaa</i>	ENSAMXT00000008315	Eye development	250.495 down	45.556 down
<i>Lim2</i>	ENSAMXT00000004471	Lens constituent	21.287 down	25.195 down
<i>Mitfa</i>	ENSAMXT00000004064	Melanocyte differentiation	116.285 down	116.285 down
<i>Mipa</i>	ENSAMXT00000004665	Lens development	29.444 down	2691.257 down
<i>Pax6a</i>	ENSAMXT00000013384	Eye; retina; brain	27.878 down	21.645 down
<i>Rho</i>	ENSAMXT00000027072	Absorption of visible light	179.779 down	48.183 down
<i>Or129-1</i>	ENSAMXT00000026728	Odorant receptor	79.498 down	13.164 down
<i>Or117-1</i>	ENSAMXT00000025723	Odorant receptor	25.139 down	25.139 down
<i>Pmch</i>	ENSAMXT00000012159	Pigmentation	20.950 down	20.950 down
<i>Mybl1</i>	ENSAMXT00000014174	Regulation of transcription	250.736 down	1747.910 down
<i>PNLIPRP2</i>	ENSAMXT00000002038	Protein binding	107.075 down	104.339 down
72 hpf				
<i>Crx</i>	ENSAMXT00000001606	Visual perception	31.387 down	13.068 down
<i>Mch1rb</i>	ENSAMXT00000026699	Pigmentation	12.359 down	24.573 down
<i>Hps5</i>	ENSAMXT00000019612	Melanocyte; iridophore; retina	27.819 down	27.819 down
<i>Vsx1</i>	ENSAMXT00000011754	Visual system	32.800 down	15.161 down
<i>Ocstamp</i>	ENSAMXT00000011253	Osteoclast differentiation	15.983 down	15.983 down
<i>Nrl</i>	ENSAMXT00000011120	Retina; lens	60.957 down	384.303 down
<i>Rx3</i>	ENSAMXT00000008550	Retina	37.910 down	203.650 down
<i>IMPG2</i>	ENSAMXT00000004773	Photoreceptor	20.224 down	22.888 down
<i>Prph2l</i>	ENSAMXT00000007155	Visual perception	13.167 down	19.627 down
<i>SLC5A8</i>	ENSAMXT00000021731	Transmembrane transport	1226.358 down	337.560 down
<i>Vbp1</i>	ENSAMXT00000003797	Protein folding; prefoldin complex	156.980 down	156.980 down

* Physiological functions are determined by a combination of gene ontology and known literature. Functions in bold may be associated with troglomorphic traits.

** Genes are under-expressed 10-fold or greater in both Pachón and Tinaja cavefish compared to surface. Blue indicates under-expressed, Red represents over-expression.

Table S5. Genes demonstrating convergent patterns of increased expression in both Pachón and Tinaja cavefish compared to surface-dwelling fish.

Gene	Transcript ID	Physiological function(s)*	Pachón cavefish (fold change)**	Tinaja cavefish (fold change)**
10 hpf				
<i>Ccl14l</i>	ENSAMXT0000005669	Immune response	1139.197 up	77.691 up
<i>Nrl</i>	ENSAMXT00000011120	Lens; retina	59.789 up	299.183 up
<i>Tas2r11</i>	ENSAMXT00000006165	Sensory perception of taste	70.497 up	234.314 up
<i>Whsc1</i>	ENSAMXT00000012843	Embryonic skeletal development	81.970 up	131.517 up
<i>Slc45a2</i>	ENSAMXT00000011610	Response to light stimulus	11.977 up	24.740 up
<i>Bco2a</i>	ENSAMXT00000003245	Pigmentation	14.021 up	17.053 up
<i>Elavl1</i>	ENSAMXT00000006559	Vision	11.742 up	12.751 up
<i>Hla-F10a1</i>	ENSAMXT00000003617	Immune response; antigen processing	68.677 up	132.038 up
<i>Rpe65c</i>	ENSAMXT00000000054	Retina pigment epithelium	13.902 up	34.777 up
<i>Npffr11l</i>	ENSAMXT00000021195	G-protein coupled receptor signaling activity	69.538 up	3905.906 up
<i>Mxtx2</i>	ENSAMXT00000011414	Transcription; actin; gastrulation	1247.060 up	268.733 up
24 hpf				
<i>Trim35</i>	ENSAMXT00000002908	Zinc ion binding	624.919 up	184.620 up
<i>Cald1</i>	ENSAMXT00000001463	Vasculogenesis; heart looping	10.096 up	552.200 up
<i>Dand5</i>	ENSAMXT00000000429	Mesoderm/endoderm formation; heart looping	174.737 up	71.102 up
<i>Top1</i>	ENSAMXT00000019646	Response to oxidative stress; heme binding	89.712 up	50.468 up
<i>Cd40lg</i>	ENSAMXT00000020533	Immune response	16.574 up	20.235 up
<i>Ela2l</i>	ENSAMXT00000000165	Elastase; proteolysis	139.754 up	282.115 up
<i>Tdrd1</i>	ENSAMXT000000015791	mRNA processing	119.836 up	98.049 up
<i>MUC5B</i>	ENSAMXT000000019204	Negative regulation of translation	28.852 up	34.320 up
<i>Sult5a1</i>	ENSAMXT00000007262	Sulfotransferase activity	26.405 up	76.762 up
<i>Gtf3ab</i>	ENSAMXT00000014879	Metal ion binding	52.035 up	41.623 up
36 hpf				
<i>Plb1</i>	ENSAMXT00000025263	Lipid metabolic process	179.225 up	76.688 up
<i>No name</i>	ENSAMXT00000003595	Immune response	48.232 up	25.539 up
<i>Nans</i>	ENSAMXT000000015834	Carbohydrate biosynthetic process	82.080 up	931.723 up
<i>Nphs1</i>	ENSAMXT000000011235	Kidney; skeletal muscle fiber development	11.441 up	23.736 up
<i>Tnfa</i>	ENSAMXT000000011932	Immune response	11.307 up	19.287 up
<i>Alpi.2</i>	ENSAMXT000000001748	Metabolic process	245.367 up	132.464 up
<i>Nod1</i>	ENSAMXT000000008263	Regulation of apoptotic process	71.810 up	129.108 up
<i>Arl15a</i>	ENSAMXT000000020246	Small GTPase mediated signal transduction	59.315 up	22.815 up
<i>Klhl10a</i>	ENSAMXT000000018852	Protein binding	47.545 up	15.691 up
<i>Ccr6b</i>	ENSAMXT000000027055	G-protein coupled receptor pathway	17.513 up	24.564 up
72 hpf				
<i>Ovgp1</i>	ENSAMXT000000010150	Carbohydrate metabolic process	113.094 up	1033.057 up
<i>Or115-15</i>	ENSAMXT000000025780	Sensory perception of smell	204.065 up	323.993 up
<i>Bco2l</i>	ENSAMXT00000005225	Pigmentation	29.474 up	103.832 up
<i>Or107-1</i>	ENSAMXT000000011497	Sensory perception of smell	13.243 up	27.711 up
<i>Ogn</i>	ENSAMXT000000013874	Ectopic bone formation	11.901 up	10.820 up
<i>CXCL10l</i>	ENSAMXT000000009192	Immune response; chemokine activity	33.566 up	57.712 up
<i>Mr1</i>	ENSAMXT000000009606	Immune response; antigen processing	68.836 up	62.073 up
<i>Ifnphi1</i>	ENSAMXT000000000458	Defense response; chemokine receptor binding	18.720 up	19.751 up
<i>Fabp6</i>	ENSAMXT000000005871	Bile acid binding	18.754 up	4640.367 up
<i>Ache</i>	ENSAMXT000000005815	Metabolic process	68.660 up	367.686 up
<i>Rxfp3</i>	ENSAMXT000000026167	G-protein coupled receptor signaling pathway	40.948 up	52.722 up

* Physiological functions are determined by a combination of gene ontology and known literature. Functions in bold may be associated with troglomorphic traits.

** Genes are over-expressed 10-fold or greater in both Pachón and Tinaja cavefish compared to surface. Blue indicates under-expressed, Red represents over-expression.

Table S6. Genes demonstrating divergent patterns of reduced expression of 10-fold or greater in Pachón cavefish compared to surface-dwelling fish.

Gene	Transcript ID	Physiological function(s)*	Pachón cavefish (fold change)**	Tinaja cavefish (fold change)
10 hpf				
<i>Ora6</i>	ENSAMXT00000006293	Olfaction	12.160 down	3.145 up
<i>Gh1</i>	ENSAMXT00000015091	Adipose tissue development	16.448 down	2.444 down
<i>Ora3</i>	ENSAMXT00000004874	Olfaction; Sensory perception of taste	12.836 down	2.655 down
<i>Cart2</i>	ENSAMXT00000001741	Adult feeding behavior	41.140 down	3.427 up
<i>Cart4</i>	ENSAMXT00000021105	Adult feeding behavior	163.894 down	1.766 down
<i>Wnt5b</i>	ENSAMXT00000011118	Neuromast development	19.447 down	1.831 up
<i>Cry1b</i>	ENSAMXT00000020701	Circadian clock	287.900 down	7.724 down
<i>Btk</i>	ENSAMXT00000002400	Eye; whole organism	92.641 down	1.551 down
<i>Shisa6</i>	ENSAMXT00000016152	Eye; retina	78.748 down	2.057 up
<i>Ghra</i>	ENSAMXT00000006038	Protein binding; whole organism	265.952 down	3.004 down
<i>Aifm2</i>	ENSAMXT00000002011	Oxidation-reduction process	64.994 down	1.098 up
24 hpf				
<i>Pmela</i>	ENSAMXT00000006340	Pigmentation	46.581 down	1.012 down
<i>Tmtops2a</i>	ENSAMXT00000009817	Photoreceptor activity	10.002 down	1.052 up
<i>Mhc1zba</i>	ENSAMXT00000005234	Immune response	16.669 down	2.095 up
<i>Cryba1a</i>	ENSAMXT00000018270	Lens constituent	12.311 down	1.465 up
<i>Vsx2</i>	ENSAMXT00000011986	Eye; retina	17.223 down	6.633 down
<i>Csf1ra</i>	ENSAMXT00000017605	Pigmentation; xanthophore	11.912 down	1.129 up
<i>Lim2.2</i>	ENSAMXT00000022220	Lens protein	12.428 down	1.441 up
<i>C1galt1</i>	ENSAMXT00000016160	Transferase activity	528.520 down	1.710 up
<i>Prlh2</i>	ENSAMXT00000002578	Hormone activity	253.411 down	2.758 up
<i>Lss</i>	ENSAMXT00000006248	Transferase activity	252.579 down	1.117 down
36 hpf				
<i>Pde6h</i>	ENSAMXT00000011755	Visual perception	43.409 down	5.656 down
<i>Atoh7</i>	ENSAMXT00000026693	Retina layer formation; swimming behavior	38.701 down	2.920 down
<i>Clock3</i>	ENSAMXT00000022444	Circadian rhythm	47.465 down	1.390 up
<i>Nrl</i>	ENSAMXT00000011120	Lens; retina	20.492 down	2.574 down
<i>Myo6b</i>	ENSAMXT00000017674	Auditory receptor; sensory perception	63.424 down	1.031 down
<i>Smtlb</i>	ENSAMXT00000006546	Pigment granule aggregation	17.174 down	2.126 down
<i>Pomca</i>	ENSAMXT00000012191	Melanosome localization	23.553 down	1.899 down
<i>Irbp</i>	ENSAMXT00000003213	Photoreceptor activity	27.892 down	2.035 down
<i>Six7</i>	ENSAMXT00000016345	Embryonic camera-type eye morphogenesis	59.027 down	4.105 down
<i>Itn2</i>	ENSAMXT00000003593	Defense response	470.848 down	1.063 up
72 hpf				
<i>Opn4.1</i>	ENSAMXT00000026354	Light-induced release of calcium ion	33.247 down	8.402 down
<i>Or130-1</i>	ENSAMXT00000026727	Olfaction	12.529 down	1.716 up
<i>Mc4r</i>	ENSAMXT00000027076	Pigmentation	126.414 down	6.274 down
<i>Rgrb</i>	ENSAMXT00000004427	Retinal protein	20.154 down	5.311 down
<i>Olfml1</i>	ENSAMXT00000027193	Olfaction	29.240 down	1.079 down
<i>Per1b</i>	ENSAMXT00000019299	Response to light and temperature stimulus	11.676 down	5.744 up
<i>Agrp</i>	ENSAMXT00000005977	Pigmentation	382.851 down	1.038 up
<i>Irx4a</i>	ENSAMXT00000013029	Lateral line; neuromast; photoreceptor	15.695 down	5.391 down
<i>Mhcf10al</i>	ENSAMXT00000003605	Immune response	81.509 down	4.380 up
<i>Impg1a</i>	ENSAMXT00000010962	Photoreceptor	100.072 down	3.291 down
<i>Cryba1l2</i>	ENSAMXT00000009610	Lens constituent	23.979 down	4.517 down
<i>Tyr</i>	ENSAMXT00000008998	Pigmentation	12.772 down	3.126 down
<i>Laptm4b</i>	ENSAMXT00000014869	Lysosomal protein	1077.861 down	1.021 up
<i>Itn2</i>	ENSAMXT00000003593	Defense response	756.587 down	1.047 down
<i>Itn1</i>	ENSAMXT00000003604	Defense response	847.209 down	1.179 up

* Physiological functions are determined by a combination of gene ontology and known literature. Functions in bold may be associated with troglomorphic traits.

** Genes are under-expressed 10-fold or greater in Pachón cavefish compared to surface. These genes show a modest level of differential expression in Tinaja cavefish. Blue indicates under-expressed, Red represents over-expression.

Table S7. Genes demonstrating divergent patterns of reduced expression of 10-fold or greater in Tinaja cavefish compared to surface-dwelling fish.

Gene	Transcript ID	Physiological function(s)*	Pachón cavefish (fold change)	Tinaja cavefish (fold change)**
10 hpf				
<i>Clock3</i>	ENSAMXT00000022444	Circadian rhythm	9.052 up	50.092 down
<i>Unc119b</i>	ENSAMXT00000025106	Visual perception	6.511 down	49.682 down
<i>Npas2</i>	ENSAMXT00000007111	Circadian rhythm; photoperiodism	5.043 down	292.627 down
<i>Pcsk5a</i>	ENSAMXT00000006794	Neuromast; lateral line development	4.432 down	22.071 down
<i>Mhcf10la</i>	ENSAMXT00000012479	Immune response	4.165 up	118.172 down
<i>Opn5</i>	ENSAMXT00000010451	Photoreceptor activity	3.679 down	23.466 down
<i>Socs1a</i>	ENSAMXT00000026498	Retina morphogenesis in camera-type eye	2.486 up	20.388 down
<i>Tmtops2b</i>	ENSAMXT00000003950	Photoreceptor activity	2.252 up	24.402 down
<i>Vsx2</i>	ENSAMXT00000011986	Visual system	2.088 down	18.708 down
<i>Mchr1a</i>	ENSAMXT00000009888	Pigmentation	1.713 up	23.202 down
<i>And2</i>	ENSAMXT00000014657	Cartilage development; fin development	1.651 down	17.097 down
24 hpf				
<i>Opn4b</i>	ENSAMXT00000001657	Photoreceptor activity	1.345 down	43.101 down
<i>Pomcb</i>	ENSAMXT00000006682	Pigmentation	2.726 down	77.225 down
<i>Mlpha</i>	ENSAMXT00000009802	Melanocyte; melanosome; pigment dispersion	5.027 down	102.456 down
<i>Pmchl</i>	ENSAMXT00000026184	Pigmentation	2.445 down	492.004 down
<i>Hsf4</i>	ENSAMXT00000021068	Lens morphogenesis in camera-type eye	3.417 down	60.965 down
<i>Opn1mw4</i>	ENSAMXT00000001292	Photoreceptor activity	1.299 down	163.081 down
<i>Crybb1ll</i>	ENSAMXT00000013433	Lens constituent	8.014 down	280.841 down
<i>Ora3</i>	ENSAMXT00000004874	Olfaction; sensory perception of taste	4.906 down	57.792 down
<i>Pde6g</i>	ENSAMXT00000001177	Visual perception	3.934 up	52.127 down
<i>Pax6l</i>	ENSAMXT00000010326	Retina development in camera-type eye	2.245 up	73.524 down
36 hpf				
<i>Or128-10</i>	ENSAMXT00000026303	Olfaction	1.295 up	183.637 down
<i>Sagb</i>	ENSAMXT00000018289	Retina; pineal gland	1.210 down	21.903 down
<i>Lim2.4</i>	ENSAMXT00000009826	Lens constituent	8.052 down	285.606 down
<i>Crygm5</i>	ENSAMXT00000019695	Lens constituent	2.928 down	56.559 down
<i>Pde6c</i>	ENSAMXT00000005049	Photoreceptor; embryonic retina	8.195 down	70.634 down
<i>Mchlrb</i>	ENSAMXT00000026699	Pigmentation	1.719 down	23.770 down
<i>Agrp</i>	ENSAMXT00000005977	Pigmentation	6.697 down	41.822 down
<i>Mc1r</i>	ENSAMXT00000026536	Pigmentation	2.367 up	58.100 down
<i>Mtnr1aa</i>	ENSAMXT00000025669	Melatonin receptor	1.331 up	87.679 down
<i>Chordc1a</i>	ENSAMXT00000003933	Hsp90 protein binding	4.997 down	3610.523 down
72 hpf				
<i>Rp1</i>	ENSAMXT00000018756	Retina	1.028 up	21.102 down
<i>Lim2.1</i>	ENSAMXT00000004471	Lens constituent	1.083 up	16.093 down
<i>Elavl1</i>	ENSAMXT00000006559	Vision	1.515 up	209.712 down
<i>Mchr2</i>	ENSAMXT00000009124	Pigmentation	3.445 down	48.298 down
<i>Crygn2</i>	ENSAMXT00000018569	Lens constituent	2.900 up	1791.267 down
<i>Irx7</i>	ENSAMXT00000000343	Retina morphogenesis in camera-type eye	1.526 down	17.011 down
<i>Ndr2</i>	ENSAMXT00000016182	Retina development; embryonic eye	6.363 down	74.534 down
<i>Mtnr1bb</i>	ENSAMXT00000009025	Melatonin receptor	1.493 up	16.571 down
<i>Arl17a</i>	ENSAMXT00000004078	GTP binding	3.622 up	556.616 down
<i>Dmcl</i>	ENSAMXT00000008883	DNA repair	1.438 down	303.721 down

* Physiological functions are determined by a combination of gene ontology and known literature. Functions in bold may be associated with troglomorphic traits.

** Genes are under-expressed 10-fold or greater in Tinaja cavefish compared to surface. These genes show a modest level of differential expression in Pachón cavefish. Blue indicates under-expressed, Red represents over-expression.

Table S8. Genes demonstrating divergent patterns of increased expression of 10-fold or greater in Pachón cavefish compared to surface-dwelling fish.

Gene	Transcript ID	Physiological function(s)*	Pachón cavefish (fold change)**	Tinaja cavefish (fold change)
10 hpf				
<i>Vn1r1</i>	ENSAMXT00000025599	Sensory perception of taste	382.812 up	3.369 down
<i>Gnrh2</i>	ENSAMXT00000014502	Camera-type eye development; brain	299.455 up	no change
<i>Pax6l</i>	ENSAMXT00000010326	Retina development in camera-type eye	176.212 up	4.500 down
<i>Trpa1a</i>	ENSAMXT00000008706	Sensory perception; chemical/mechanical stimulus	115.042 up	3.185 down
<i>Rom1b</i>	ENSAMXT00000010795	Eye; retina; photoreceptor layer	79.760 up	3.683 up
<i>Sagb</i>	ENSAMXT00000018289	Eye; retina; photoreceptor layer	71.854 up	1.706 up
<i>Lim2.3</i>	ENSAMXT00000003097	Eye; lens constituent	46.413 up	5.594 down
<i>Mc5rb</i>	ENSAMXT00000025661	Pigmentation	44.587 up	4.034 down
<i>Dfnb59</i>	ENSAMXT00000004928	Auditory system; inner ear nerves	37.118 up	no change
<i>Chat</i>	ENSAMXT00000012506	Locomotory behavior	35.994 up	3.751 up
24 hpf				
<i>Dlg4b</i>	ENSAMXT00000020555	Retina; olfactory bulb/epithelium; midbrain	245.106 up	no change
<i>Lyz</i>	ENSAMXT00000020555	Lysozyme	209.467 up	6.142 up
<i>Crygm4</i>	ENSAMXT00000008090	Eye; lens constituent	177.556 up	no change
<i>Crygm3</i>	ENSAMXT00000011158	Eye; lens constituent	19.649 up	1.280 up
<i>Myo6b</i>	ENSAMXT00000017674	Auditory receptor; sensory perception	15.681 up	1.058 up
<i>Elavl1</i>	ENSAMXT00000006559	Visual perception	13.673 up	43.151 down
<i>Wnt5b</i>	ENSAMXT00000011118	Otolith development; neuromast development	12.436 up	5.739 down
<i>Tmtops2b</i>	ENSAMXT00000003950	Eye; photoreceptor	12.258 up	20.008 down
<i>Esrp2</i>	ENSAMXT00000006045	Nucleic acid binding	503.162 up	no change
<i>Apnl</i>	ENSAMXT00000021007	Cation transport; pore complex; hemolysis	188.626 up	no change
36 hpf				
<i>Ccl18</i>	ENSAMXT00000009248	Immune response; chemokine activity	27.926 up	no change
<i>Lim2.2</i>	ENSAMXT00000022220	Eye; lens constituent	18.190 up	4.117 up
<i>Olfm4</i>	ENSAMXT00000014729	Olfaction	17.626 up	3.768 up
<i>Grid2tpa</i>	ENSAMXT00000007995	Regulation of transcription	261.730 up	1.029 up
<i>Phf20a</i>	ENSAMXT00000002218	Zinc ion binding	234.814 up	5.453 up
<i>Gxylt1b</i>	ENSAMXT00000017058	Transferase activity, transfer glycosyl groups	151.277 up	no change
<i>Arpc4</i>	ENSAMXT00000002298	Arp2/3 complex-mediated actin nucleation	92.820 up	9.483 down
<i>Catsper1</i>	ENSAMXT00000009853	Sperm associated; ion transport	85.391 up	3.365 down
<i>Tjp3</i>	ENSAMXT00000016403	Protein binding	48.074 up	no change
<i>Pcd5</i>	ENSAMXT00000003666	DNA binding	53.517 up	no change
72 hpf				
<i>Sb:cb108l</i>	ENSAMXT00000018849	Eye; lens constituent	880.522 up	2.117 down
<i>H2-IEal</i>	ENSAMXT00000021452	Immune response; MHC class II complex	186.416 up	no change
<i>Mipb</i>	ENSAMXT00000002019	Lens development in camera-type eye	171.014 up	5.080 down
<i>Ccl-11l</i>	ENSAMXT00000005317	Immune response; chemokine activity	44.778 up	3.634 up
<i>Tas2r1l</i>	ENSAMXT00000006165	Sensory perception of taste	38.398 up	5.496 up
<i>Pde6g</i>	ENSAMXT00000001177	Visual perception	35.934 up	3.174 up
<i>Foxe3</i>	ENSAMXT00000017091	Lens morphogenesis in camera-type eye	24.211 up	9.269 down
<i>Bglap</i>	ENSAMXT00000011048	Bone; calcium ion binding	22.571 up	5.429 up
<i>Mettl5</i>	ENSAMXT00000005950	rRNA methylation; protein methylation	341.201 up	no change
<i>H2b</i>	ENSAMXT00000018389	Nucleosome assembly; DNA binding	199.164 up	no change

* Physiological functions are determined by a combination of gene ontology and known literature.

** Genes are over-expressed 10-fold or greater in Pachón cavefish compared to surface. These genes show a modest level of differential expression in Tinaja cavefish. Blue indicates under-expressed, Red represents over-expression.

Table S9. Genes demonstrating divergent patterns of increased expression of 10-fold or greater in Tinaja cavefish compared to surface-dwelling fish.

Gene	Transcript ID	Physiological function(s)*	Pachón cavefish (fold change)	Tinaja cavefish (fold change)**
10 hpf				
<i>Impg2</i>	ENSAMXT0000002036	Photoreceptor	5.696 up	170.409 up
<i>Bco2l</i>	ENSAMXT0000005225	Pigmentation	no change	111.277 up
<i>Hla-F10a1</i>	ENSAMXT0000003605	Immune response; antigen processing	1.110 up	105.602 up
<i>Gnat2</i>	ENSAMXT00000014437	Visual perception; detection of light stimulus	8.129 up	95.238 up
<i>Adcyap1a</i>	ENSAMXT0000007454	Camera-type eye development; brain development	4.948 up	18.826 up
<i>Mtnr1a1</i>	ENSAMXT00000011185	Melatonin receptor	2.875 up	17.387 up
<i>Alox5a</i>	ENSAMXT0000001500	Oxidation-reduction process	3.286 down	1320.028 up
<i>Tspan7b</i>	ENSAMXT0000004820	Integral component of membrane	no change	529.021 up
<i>Hhatla</i>	ENSAMXT00000015498	Acyltransferase	9.003 up	308.924 up
<i>Trim35</i>	ENSAMXT00000018321	Protein ubiquitination	3.312 up	253.744 up
24 hpf				
<i>H2-IEbl</i>	ENSAMXT0000005358	Immune response; MHC class II protein	no change	700.042 up
<i>Hla-F10a1</i>	ENSAMXT0000003617	Immune response; antigen processing	2.726 up	88.106 up
<i>Rx3</i>	ENSAMXT00000008550	Retina development	no change	63.414 up
<i>Mtnr1a</i>	ENSAMXT00000006184	Melatonin receptor	4.581 up	57.666 up
<i>Lepr</i>	ENSAMXT00000010903	Camera-type eye; inner ear development	6.365 up	28.204 up
<i>Mtnr1c</i>	ENSAMXT00000007073	Melatonin receptor	1.229 up	22.801 up
<i>Mtnr1bb</i>	ENSAMXT00000009025	Melatonin receptor	6.223 up	20.374 up
<i>Ofec1</i>	ENSAMXT00000006139	Facial development; protein binding	1.872 up	15.277 up
<i>Rbpja</i>	ENSAMXT00000011495	Transcription; neurogenesis; somite specification	no change	663.454 up
<i>Flt1</i>	ENSAMXT00000014767	Regulation of heart contraction; vasculogenesis	6.452 up	252.886 up
36 hpf				
<i>Or104-2</i>	ENSAMXT00000021297	Odorant receptor	no change	91.607 up
<i>Cysltr2</i>	ENSAMXT00000026838	Sensory perception of taste	1.104 up	65.981 up
<i>Calm14b</i>	ENSAMXT00000000837	Calcium ion binding	1.191 down	63.339 up
<i>Per1b</i>	ENSAMXT00000019299	Response to light and temperature stimulus	8.135 down	63.281 up
<i>Cxcl11l</i>	ENSAMXT00000009188	Immune response; chemokine activity	2.262 down	50.321 up
<i>Lepb</i>	ENSAMXT00000016791	Response to starvation	4.080 up	31.219 up
<i>Ywhae2</i>	ENSAMXT00000007058	Immature eye; optic vesicle; hindbrain	no change	10.224 up
<i>Endo1</i>	ENSAMXT00000004210	Nucleic acid binding; hydrolase activity	4.124 up	466.036 up
<i>Ctrl</i>	ENSAMXT00000004736	Proteolysis; endopeptidase activity	2.297 up	296.230 up
<i>Gcgb</i>	ENSAMXT00000010272	Hormone activity	3.894 up	266.057 up
72 hpf				
<i>Epd</i>	ENSAMXT00000001956	Chondrocranium; calcium ion binding	no change	34096.052 up
<i>Azgp11</i>	ENSAMXT00000004667	Immune response; antigen processing	no change	780.671 up
<i>Or101-1</i>	ENSAMXT00000025777	Odorant receptor	no change	241.668 up
<i>Prg4b</i>	ENSAMXT00000010540	Immune response; scavenger receptor	4.035 up	210.527 up
<i>Olfm4</i>	ENSAMXT00000014744	Olfaction	2.633 down	164.740 up
<i>Gh1</i>	ENSAMXT00000015091	Adipose tissue development	5.171 down	143.272 up
<i>Or132-5</i>	ENSAMXT00000026673	Odorant receptor	no change	141.565 up
<i>Tas2r202</i>	ENSAMXT00000026707	Sensory perception of taste	no change	123.288 up
<i>Cryabb</i>	ENSAMXT00000007540	Structural constituent of eye lens	3.630 down	106.407 up
<i>Sen8aa</i>	ENSAMXT00000004768	Mechanosensory; larval locomotory behavior	7.063 up	80.723 up
<i>Ccl18</i>	ENSAMXT00000009248	Immune response; chemokine activity	no change	56.564 up
<i>Tas1r3</i>	ENSAMXT00000003856	Sensory perception of taste	2.369 down	42.454 up
<i>Or115-15</i>	ENSAMXT00000025782	Odorant receptor	1.008 up	38.161 up
<i>Il22</i>	ENSAMXT00000004278	Immune response; response to bacterium	4.422 up	24.939 up
<i>Tas1r2.1</i>	ENSAMXT00000014814	Sensory perception of taste	no change	22.066 up
<i>Ora4</i>	ENSAMXT00000008926	Odorant receptor	5.050 up	21.935 up
<i>Astl</i>	ENSAMXT00000021359	Proteolysis; metalloendopeptidase activity	1.701 down	14244.55 up
<i>Cela1</i>	ENSAMXT00000014505	Proteolysis; serine-type endopeptidase activity	4.189 down	9968.700 up

* Physiological functions are determined by a combination of gene ontology and known literature.

** Genes are over-expressed 10-fold or greater in Tinaja cavefish compared to surface. These genes show a modest level of differential expression in Pachón cavefish. Blue indicates under-expressed, Red represents over-expression.

GENERAL CONCLUSIONS

Regressive evolution or phenotypic “loss” of characters affects a remarkable number of organisms throughout nature. However, the evolutionary and genetic mechanisms mediating trait loss are still poorly understood (Jeffery 2009). For these studies, the blind Mexican cavefish is an ideal system for investigating this phenomenon as it exhibits extreme losses of pigmentation upon invasion of the subterranean environment (Gross 2012). The overall goal of this research was to uncover some of the broader genetic mechanisms mediating regressive pigmentation in cavefish. To approach this bigger question, my work characterized different genetic mechanisms – *simple traits, complex traits and global transcriptomic changes* – that could contribute to the observed regressive pigmentation in cavefish. The work presented here is novel because it integrates diverse molecular, functional, and bioinformatics approaches to better understand the complex genetic underpinnings governing the evolution of unique phenotypic characters in natural populations.

The first chapter presents the analysis of a *simple* pigmentation trait in cavefish called *brown*. This classic pigmentation phenotype was first characterized in wild cavefish in the mid-1950’s, where initial crosses of (parental: surface x cave) segregated into a perfect Mendelian ratio, suggesting that this trait is caused by a single locus (Sadoglu and Mckee 1969). Complementation crosses demonstrated that the same gene appeared to cause *brown* in seven cave populations (Wilkens and Strecker 2003). However, it was not until recently that an association mapping study identified *Mc1r* as the gene mediating *brown* due to coding sequence alterations in three populations (Gross et al. 2009). Interestingly, four of the seven caves harboring *brown* have an intact coding sequence in cavefish. Therefore, we surmised that *cis*-regulatory alterations at this locus may also cause *brown*. We interrogated the 5’ putative

regulatory region of *Mc1r* for any mutations that could affect expression of this gene. Sequence analyses revealed 42 mutations that segregated between surface and cave morphotypes. Some of these mutations co-localized to highly conserved non-coding elements (CNEs), with one mutation that is fixed in all *brown* caves and also falls within a predicted E-box motif transcription factor-binding site. We also identified significantly reduced expression of *Mc1r* in adults of two caves confirmed to express *brown* (Pachón and Chica) and another population that likely harbors *brown* (Tinaja). From these studies, we confirmed that *brown* cavefish populations that demonstrate no coding sequence mutations still exhibit reduced expression of *Mc1r*. This suggests that *brown* may have evolved through a combination of both coding and *cis*-regulatory alterations of the same locus.

My second chapter investigates pigmentation phenotypes in *Astyanax* that are *complex*, resulting from the influence of multiple genes. Prior to this dissertation, the only genes known to govern specific traits in *Astyanax* were *Oca2* (albinism; Protas et al. 2005) and *Mc1r* (*brown*; Gross et al. 2009). Yet, an earlier mapping study assaying melanophore number in four regions of the body yielded 18 QTL associations, suggesting numerical pigment cell variation is complex (Protas et al. 2007). The identities of the genes underlying these loci was not identified. To characterize genes involved in melanophore variation, we capitalized on a recently published, densely populated linkage map (Carlson et al. 2015). This enabled high-resolution mapping of melanophore number variation scored in seven regions of the body in our large surface x cave F₂ pedigree. Our association analyses yielded numerous significant associations, which localized to 20 distinct regions of our linkage map. Using comparative genomics, we identified the syntenic regions for each QTL in both the *Astyanax* draft genome and the model fish system *Danio rerio*. We then screened genes residing within syntenic intervals for all gene ontology (GO) annotation

that indicated a prospective role in pigmentation-related functions. The positional information, combined with previously known roles in pigmentation, narrowed our search for candidate genes to two intriguing loci, *Tyrp1b* and *Pmela*. Further integrative analyses of both genes identified multiple SNPs and differential expression during early development in Pachón cavefish compared to surface. We also visualized these genes using *in situ* hybridization to confirm melanophore-specific expression in *Astyanax*. We validated that inhibition of *Tyrp1b* and *Pmela* using morpholino knockdowns in zebrafish, which yielded qualitative reduction of pigmentation and altered melanophore shape, respectively. These observed phenotypes could explain differing melanophore numbers in various body regions in F₂ hybrids. Together, this provides evidence that *Tyrp1b* and *Pmela* likely contribute to the complex trait of pigment cell number diversity in *Astyanax*.

The final dissertation chapter aimed to characterize *global transcriptomic changes* in pigmentation genes upon colonization into the cave habitat using RNA-seq expression profiling. We worked to identify genes contributing to cave adaptation by searching for genes exhibiting dramatic changes in expression across four stages of early development. We aimed to further bolster support for genes we identified by assaying *two geographically distinct* cavefish populations (Pachón and Tinaja) compared to surface morphs. We discovered that Pachón and Tinaja cavefish evolved convergent expression in the some of the same loci, possibly as a result of gene flow or selection. However, a common genetic basis shared among cavefish populations could explain the overall morphological appearance. Yet, some genes were expressed in a divergent pattern, which could suggest modest differences among cave localities could confer cave-specific expression modifications. Our pathway analysis revealed two intriguing genes *Otx2* and *Mitf*. The reduced expression of these two genes could explain some of the global,

downstream effects of reduced expression observed in our dataset. This study further indicated a wide array of expression patterns affecting genes of diverse functions. This study showed how genes governing different functions (eyes and pigmentation) can involve the same genetic pathway, which could further support the hypothesis of pleiotropy in the evolution of cave-adapted animals. This study demonstrated that striking characters observed across different cavefish populations evolve through a combination of both convergent and divergent genetic mechanisms.

In summary, the work presented in this dissertation includes an integrative analysis of regressive pigmentation studies in *Astyanax mexicanus*. Investigating multiple aspects of pigmentation clearly demonstrated that complex genetic mechanisms –including *simple traits*, *complex traits* and *global expression changes* – work in concert to cause regressive pigmentation in geographically distinct populations of cavefish. Ultimately, this dissertation provides deeper insight to the genetic mechanisms governing regressive evolution in natural populations, and identifies several candidate genes that may play a role in human degenerative disorders.

REFERENCES

- Carlson, B. M., S. W. Onusko, and J. B. Gross (2015) A High-Density Linkage Map for *Astyanax mexicanus* Using Genotyping-by-Sequencing Technology. *G3* 5: 241–251.
- Gross, J. B. (2012) *Cave Evolution*. John Wiley & Sons, Ltd, Chichester, UK.
- Gross, J. B., R. Borowsky, and C. J. Tabin (2009) A Novel Role for *Mclr* in the Parallel Evolution of Depigmentation in Independent Populations of the Cavefish *Astyanax mexicanus*. *PLoS Genet* 5: e10x300326–14.
- Jeffery, W. R. (2009) *Evolution and Development in the Cavefish Astyanax*. Evolution and Development, Current Topics in Developmental Biology, Elsevier p. 191–221.
- Protas, M., M. Conrad, J. B. Gross, C. Tabin, and R. Borowsky (2007) Regressive Evolution in the Mexican Cave Tetra, *Astyanax mexicanus*. *Current Biology* 17: 452–454.
- Protas, M. E., C. Hersey, D. Kochanek, Y. Zhou, H. Wilkens, W. R. Jeffery, L. I. Zon, R. Borowsky, and C. J. Tabin (2005) Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nature Genetics* 38: 107–111.
- Sadoglu, P., and A. Mckee (1969) A Second Gene That Affects Eye and Body Color in Mexican Blind Cave Fish. *Journal of Heredity* 60: 10–14.
- Wilkens, H., and U. Strecker (2003) Convergent evolution of the cavefish *Astyanax* (Characidae, Teleostei): genetic evidence from reduced eye-size and pigmentation. *Biological Journal of the Linnean Society* 80: 545–554.