DECLINING POPULATIONS IN CHANGING ENVIRONMENTS: ADAPTIVE RESPONSES, GENETIC DIVERSITY, AND CONSERVATION

by

Avril M. Harder

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Biological Sciences

West Lafayette, Indiana

December 2020

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Mark Christie, Chair
Department of Biological Sciences
Department of Forestry & Natural Resources

Dr. J. Andrew DeWoody
Department of Forestry & Natural Resources
Department of Biological Sciences

Dr. Michael Gribskov
Department of Biological Sciences

Dr. William Ardren
U.S. Fish & Wildlife Service

Approved by:
Dr. Janice Evans

ACKNOWLEDGMENTS

I am immensely grateful to my advisor, Dr. Mark Christie, for his patience, sense of humor, guidance, and mentorship throughout my time at Purdue. I will forever be thankful that I landed in his lab and for the support I found there. Thank you to my committee: Dr. J. Andrew DeWoody for welcoming me to his lab's community and for his guidance; Dr. William Ardren for facilitating and supporting all of our work in Lake Champlain and for countless helpful conversations; and Dr. Michael Gribskov for feedback on computational aspects of my research and for incredibly valuable genomics and bioinformatics courses.

My time at Purdue was made infinitely easier thanks to the friendship and support of Dr. Janna Willoughby; my academic twin brother, Alex Martinez; my academic sister from another advisor, Stephanie Gutierrez; my partners in commiseration, Morgan Sparks and Claire Schraidt; and everyone in the Christie and DeWoody labs over the past few years. I am grateful for the guidance and support of Purdue EEB faculty, including Dr. Dennis Minchella for annual TA dining hall lunches and for always having chocolate at the ready in times of both celebration and crisis; Dr. Ximena Bernal for welcoming me to her classroom for guest lectures and lively discussions and for setting example after example of advocacy in action; and Dr. Peter and Mary Sue Waser for generously funding the fellowship that enabled the writing of this dissertation during very strange times. The rest of the EEB community—graduate students, technicians, and postdocs—comprises some of the most helpful and supportive individuals I have met in my academic journey, and I thank them for comments and edits on proposals, inappropriately long lunch breaks, stimulating journal club presentations and conversations, and a decent fantasy football league. Thank you to Gina Rupp for keeping me on track and for always reminding me that everything would be alright. I would be remiss if I did not also revisit a much earlier stage in my academic history to thank Ms. Todd for always insisting that I "Prove it."

I feel so fortunate to have the most supportive parents. They have always encouraged my curiosity and believe everything I do is wonderful and impressive, even when it isn't.

Thank you to Ben for unflinching support over the past seven years, through two graduate programs and too many relocations. As a sign of my gratitude, I forgive you for having more citations than I do. I will be forever grateful to my work-from-home office mates, Ozzie and Pearl, who constantly remind me that sometimes all one needs is to just step away for a bit and go outside.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Many salmonid populations are supported through captive breeding programs in which hatchery production supplies fish for reintroduction or supplementation efforts. In Lake Champlain, Atlantic salmon (*Salmo salar*) are the subject of a reintroduction effort that is complicated by the occurrence of thiamine (vitamin B_1) deficiency in adult salmon returning to spawn. This deficiency results in high offspring mortality rates that must be mitigated by hatchery interventions (reviewed in Chapter 1). I used an experimental transcriptomics approach coupled with survival analyses to assess genetic variation in thiamine deficiency outcomes (i.e., survival at the family level) and identified candidate genes that may comprise a putatively adaptive response to selection imposed by thiamine deficiency (Chapter 2). Using sequence data from this study, I next compared patterns of genetic variation in the Lake Champlain population against two other populations to identify signatures of selection associated with hatchery rearing environment and differences in life history strategies (Chapter 3). Finally, I surveyed salmonid populations for density-dependent effects of adult spawning density on per capita fitness and found that in many cases, hatchery releases can contribute to decreased individual fitness. Using genotype data for returning adults in multiple populations, I also tested for reductions in effective population size (N_e) associated with hatchery supplementation and describe how increasing hatchery contribution to a population decreases N_e (Chapter 4). Together, my results demonstrate the powerful influences of hatchery supplementation on salmonid populations and suggest that specific modifications to hatchery practices can limit negative impacts of captive breeding on population genetic and demographic characteristics.

CHAPTER 1. THIAMINE DEFICIENCY IN FISHES: CAUSES, CONSEQUENCES, AND POTENTIAL SOLUTIONS

Reprinted by permission from Springer Nature: Springer. *Reviews in Fish Biology and Fisheries*. Thiamine deficiency in fishes: causes, consequences, and potential solutions. Avril M. Harder, William R. Ardren, Allison N. Evans, Matthew H. Futia, Clifford E. Kraft, J. Ellen Marsden, Catherine A. Richter, Jacques Rinchard, Donald E. Tillitt, and Mark R. Christie. 2018.

1.1 Abstract

Thiamine deficiency complex (TDC) is a disorder resulting from the inability to acquire or retain thiamine (vitamin B_1) and has been documented in organisms in aquatic ecosystems ranging from the Baltic Sea to the Laurentian Great Lakes. The biological mechanisms leading to TDC emergence may vary among systems, but in fishes, one common outcome is high mortality among early life stages. Here, we review the causes and consequences of thiamine deficiency in fishes and identify potential solutions. First, we examine the biochemical and physiological roles of thiamine in vertebrates and find that thiamine deficiency consistently results in impaired neurological function across diverse taxa. Next, we review natural producers of thiamine, which include bacteria, fungi, and plants, and suggest that thiamine is not currently limiting for most animal species inhabiting natural aquatic environments. A survey of historic occurrences of thiamine deficiency identifies consumption of a thiamine-degrading enzyme, thiaminase, as the primary explanation for low levels of thiamine in individuals and subsequent onset of TDC. Lastly, we review conservation and management strategies for TDC mitigation ranging from evolutionary rescue to managing for a diverse forage base. As recent evidence suggests occurrences of thiamine deficiency may be increasing in frequency, increased awareness and a better mechanistic understanding of the underlying causes associated with thiamine deficiency may help prevent further population declines.

1.2 Introduction

Recent studies have revealed an increasing number of wild fish populations are deficient in thiamine (vitamin B_1), an essential vitamin that is required for a range of metabolic functions (Bettendorff 2013). Low levels of thiamine result in a syndrome referred to as thiamine deficiency

complex (TDC; Riley and Evans 2008) that can result in high early life stage mortality and lead to population declines (Brown et al. 2005b; Balk et al. 2009, 2016). Thiamine deficiency complex encompasses all symptoms resulting from thiamine deficiency at all life stages, including sublethal effects and direct mortality, and has emerged as a possible contributor to decreased survival and reduced reproductive success in a variety of fish taxa (Ketola et al. 2000; Brown et al. 2005b; Balk et al. 2016). This disorder was first documented in salmonine fishes in the Laurentian Great Lakes in 1968 (Marcquenski and Brown 1997) and the Baltic Sea in 1974 (Bylund and Lerche 1995), but the cause of the observed signs was not recognized as thiamine deficiency until decades later (Bylund and Lerche 1995; Fitzsimons 1995; Fisher et al. 1995). Here, we examine the causes and consequences of TDC by reviewing i., the molecular and physiological functions of thiamine, ii., the natural and artificial sources of thiamine for fishes, iii., the possible causes and consequences of TDC, and iv., research needs and questions relevant to thiamine deficiency.

The geographic and taxonomic distributions of TDC occurrences have recently broadened, with low thiamine concentrations in tissues or eggs now being documented in previously unaffected populations. Thiamine deficiency has been most widely studied in salmonine fishes and has been identified in populations of lake trout (*Salvelinus namaycush*), Atlantic salmon (*Salmo salar*), Chinook salmon (*Oncorhynchus tshawytscha*), coho salmon (*Oncorhynchus kisutch*), steelhead trout (anadromous *Oncorhynchus mykiss*), and brown trout (*Salmo trutta*) from the Great Lakes and New York Finger Lakes (Fisher et al. 1995, 1996; Marcquenski and Brown 1997). Thiamine deficiency has also been widely reported in Baltic Sea populations of Atlantic salmon and brown trout (Bengtsson et al. 1999; Karlsson et al. 1999). Assessments of egg and tissue thiamine concentrations have been conducted in Alaskan Chinook salmon (Honeyfield et al. 2016) and American eels (*Anguilla rostrata*) from Lake Ontario (Fitzsimons et al. 2013), and the results suggest continued monitoring may be important in these populations. The broadening geographic distribution of TDC cases may indicate that thiamine availability is decreasing through decreases in thiamine production or increases in thiamine-degrading pathways in multiple aquatic systems. Alternatively, increased spatial distribution may reflect increased awareness of this issue, expanded research on thiamine, improved analytical methods for measurement of thiamine, and, consequently, increased sampling and detection.

Organisms susceptible to thiamine deficiency are thiamine auxotrophs (i.e., those that cannot synthesize thiamine on their own). In aquatic communities, thiamine is produced by prokaryotes,

algae, and plants, and fishes are presumed to obtain most of their thiamine via grazing and predation on lower trophic level sources. Fishes may also obtain thiamine from symbiotic gut microbes (Ji et al. 1998), as has been documented in ruminant mammals (Breves et al. 1980, 1981) and insects (Sannino et al. 2018), but the extent to which fishes exploit thiamine produced by resident microbiota has not been investigated.

Disruptions of the thiamine-producing community within an ecosystem could lead to decreased thiamine availability for consumers such as zooplankton, macroinvertebrates, and fishes. Changes in abiotic environmental factors including light availability, temperature, and salinity have been shown to affect the thiamine content of several phytoplankton species capable of synthesizing thiamine (Sylvander et al. 2013). Although diminished thiamine production at lower trophic levels could lead to thiamine deficiency in fishes, decreased production has not been implicated in documented cases of thiamine deficiency, and most research to date has focused on other causes.

In the Great Lakes, the most widely studied potential cause of TDC is an enzyme—thiaminase (specifically, thiaminase I, which is distinct from another thiamine-degrading enzyme referred to as thiaminase II or "tenA"; Jenkins et al. 2007)—that is present in some prey fishes. Thiaminase can degrade thiamine in the gastrointestinal (GI) tract of fishes, and can ultimately decrease the amount of thiamine available to individuals consuming thiaminase-containing prey (Honeyfield et al. 2005; Houde et al. 2015). Thiaminase activity is high in alewife (*Alosa pseudoharengus*; Tillitt et al. 2005), a ubiquitous clupeid that invaded the Great Lakes and forms the primary forage base for a number of vertebrates (including fishes and birds) (Stewart et al. 1981; Fox et al. 1990; Weseloh et al. 1995; Madenjian et al. 2002; Happel et al. 2017). Awareness of the consequences of thiaminase consumption has led to detection and analysis of this enzyme in a wide variety of other taxa within aquatic food webs, including round goby (*Neogobius melanostomus*) and dreissenid mussels (*Dreissena* spp.; Honeyfield at al. 2012, Tillitt et al. 2009). Thiaminase has also been a focus of TDC research in the Baltic Sea (Karlsson et al. 1999; Wistbacka et al. 2002), with some correlational data suggesting that other dietary factors (e.g., fat content, antioxidant availability) may exacerbate thiaminase-induced TDC (Lundström et al. 1999b; Pettersson and Lignell 1999; Keinänen et al. 2012). Regardless of the proximate cause of TDC, the physiological effects of thiamine deficiency are consistent among fishes across systems and are related to the crucial role thiamine plays in bioenergetic pathways.

1.3 Functions of thiamine and its derivatives

Thiamine synthesis by living organisms produces thiamine monophosphate (TMP) (Bettendorff 2013). TMP is then modified by the addition of phosphate groups to serve a variety of physiological functions. Free (dephosphorylated) thiamine and TMP have no known physiological function (Bettendorff et al. 2014), although free thiamine may act as an antioxidant (Lukienko et al. 2000). The addition of two phosphate groups to TMP yields thiamine triphosphate (TTP), which is constantly produced at a low rate in animal cells (Bettendorff et al. 2014). TTP has been suggested to be important for proper neuron function (Cooper and Pincus 1979; Bettendorff et al. 1993), but its precise role(s) in cellular pathways remain unclear (Bettendorff et al. 2014). By far, the most well-characterized thiamine derivative is thiamine diphosphate (TDP), which serves as a cofactor for 20 enzymes in humans (Widmann et al. 2010).

Enzymes requiring TDP as a cofactor catalyze reactions in the pentose phosphate pathway and tricarboxylic acid (TCA) cycle and are essential for metabolism and energy production (Fig. 1.1). With additional roles in the production of neurotransmitters, antioxidants, and myelin, TDP-dependent enzymes are crucial to proper neurological function (Bettendorff 2013). In humans, the earliest symptoms of thiamine deficiency are known as beriberi (Faigle et al. 2012). Severe and prolonged thiamine deficiency can lead to the development of Wernicke's encephalopathy, which is a neurological disorder characterized by weakness and involuntary movement of eye muscles, ataxia, and abnormal stance and gait (Sechi and Serra 2007). Wernicke's encephalopathy is often observed concurrently with Korsakoff's psychosis, as indicated by amnesia, disorientation, and disturbances and distortions of memory (Harper et al. 1986). These two disorders are often collectively referred to as Wernicke-Korsakoff syndrome and are definitively diagnosed postmortem based on characteristic brain lesions (Sechi and Serra 2007). Similar lesions have been described in thiamine-deficient Atlantic salmon and lake trout, with severity of brain tissue necrosis increasing with disease progression (Lundström et al. 1999a; Lee et al. 2009). Furthermore, the neurological signs of Wernicke-Korsakoff syndrome in humans (e.g., abnormal stance and gait) are analogous to behavioral signs of TDC in juvenile and adult salmonine fishes (e.g., inability to remain upright, ataxia, corkscrew swimming patterns, lethargy; Fisher et al. 1995; Brown et al. 2005c; Fitzsimons et al. 2005), suggesting that the pathology of thiamine deficiency may reflect conserved TDP-dependent enzyme functions across diverse vertebrate taxa.

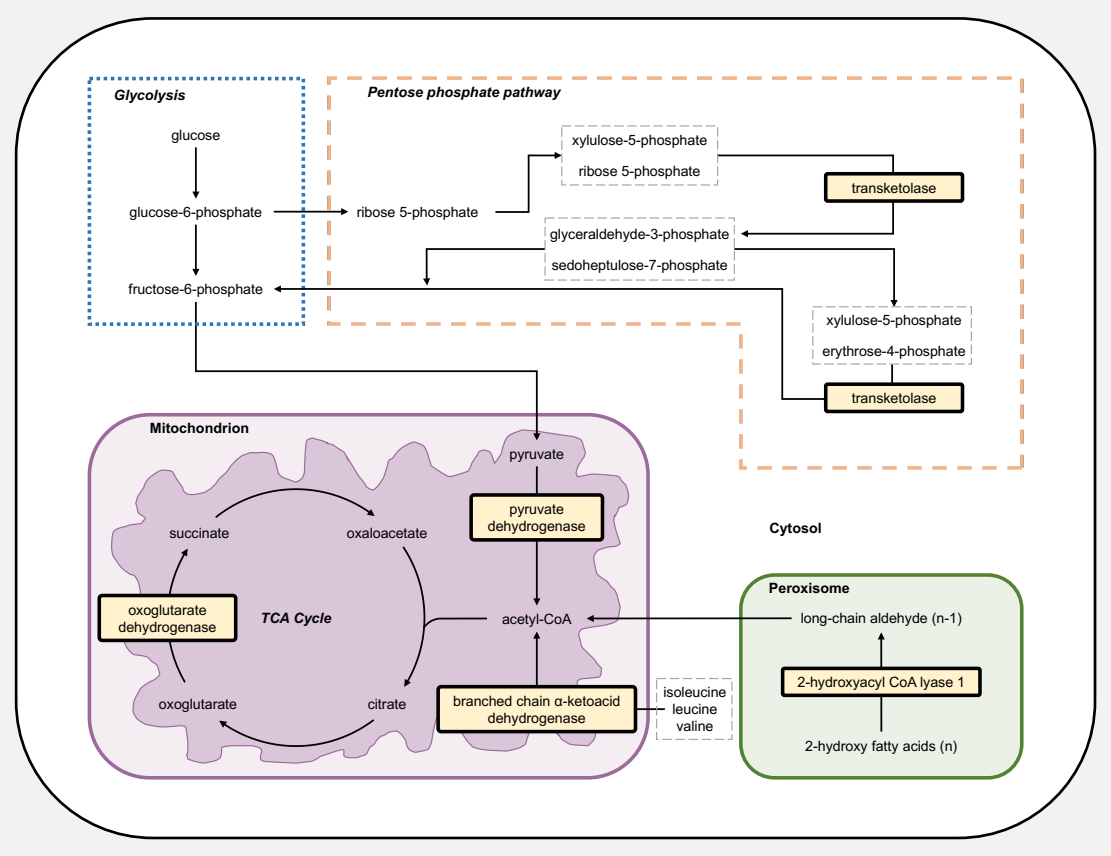


Figure 1.1. Pathways requiring thiamine diphosphate (TDP) with TDP-dependent enzymes highlighted in yellow boxes. Arrows in pathways may represent multiple reactions and intermediates for the purpose of simplification.

1.4 Natural and artificial thiamine sources

1.4.1 Production and availability of thiamine

All fishes acquire thiamine through intestinal absorption. This thiamine may be synthesized by bacteria and algae in the aquatic environment, consumed by planktivorous fishes, and later transferred via piscivory. Thiamine content in Great Lakes forage fishes varies seasonally and annually among species and locations, but does not seem to be limiting in any forage species under any particular environmental condition or set of conditions (Fitzsimons et al. 2005b; Tillitt et al. 2005). Thiamine may also be produced by resident microbiota in the GI tracts of fish. An experimental feeding study conducted in lake trout suggests that as much as 81% of thiamine present in the posterior portion of the intestine may be of non-dietary origin (i.e., synthesized by gut microbiota) (Ji et al. 1998). Gut microbes may therefore provide thiamine to their hosts, but

the degree to which non-dietary (synthesized) sources of thiamine are absorbed by posterior intestinal tissues is unknown. If fishes do receive substantial amounts of thiamine from resident microbes, perturbations of gut microbial communities could influence thiamine availability for the host fish, but little is known on this subject.

Environmental (i.e., non-dietary) sources of thiamine may also be available to fishes. Degradation of any cellular material likely releases intact cellular thiamine, but longevity of free thiamine in natural waters is unknown. Assessment of thiamine concentrations in a marine system suggest that concentrations are generally low in the water column (30–280 pM) but can be much higher in sediment pore water (30–770 pM) (Monteverde et al. 2015). In shallow aquatic environments where mixing is prevalent, thiamine may readily diffuse out of the sediment. High benthic concentrations of thiamine may be particularly important for salmonines and many other fish species that incubate their eggs in and on freshwater sediments. Whether thiamine can be absorbed across the chorion once the egg is water-hardened is also unknown, but free embryos are capable of absorbing thiamine from water (see *Thiamine therapies: hatchery treatments* below). Long overwinter incubation of salmonine eggs and 4-6 week residence of free embryos in spawning substrate after hatching may allow embryos to absorb sufficient thiamine to mitigate deficiency.

1.4.2 Thiamine therapies: hatchery treatments

Several effective methods of thiamine supplementation have been developed for multiple salmonine life stages. The method most commonly employed in hatcheries is the immersion of fertilized eggs in thiamine solution during water hardening (Wooster et al. 2000), the developmental stage where eggs absorb water and the egg membrane becomes rigid. Immersing Atlantic salmon eggs in a 1% thiamine hydrochloride (HCl) solution can yield a 14-fold increase in sac fry total thiamine concentration over untreated control fry (untreated: 0.625 nmol/g; treated: 8.706 nmol/g) (Wooster et al. 2000). The magnitude of increase in egg thiamine concentration is dosage-dependent; treating with higher concentrations of thiamine in an immersion bath results in a greater magnitude of increase in egg thiamine concentration (e.g., treatment with a 0.1% thiamine HCl solution results in a 3-fold increase in sac fry total thiamine; untreated: 0.625 nmol/g; treated: 1.845 nmol/g) (Wooster et al. 2000).

Treating individuals at the egg stage is more logistically tractable than at later developmental stages. In hatcheries that rear salmonines, fertilized eggs are often treated with an iodophor solution to prevent the spread of viral hemorrhagic septicemia (VHS) (Amend and Pietsch 1972). Hatchery biosafety protocols typically require that eggs be treated with iodophor solution first, followed by thiamine HCl (K. Kelsey, personal communication), and this series could lead to the degradation of thiamine by residual iodophor solution if the solution adheres to the eggs or if disinfected eggs are not thoroughly rinsed. Iodophor treatment also delays thiamine treatment and may limit the amount of thiamine that can be absorbed by eggs prior to chorion hardening. The possibility for iodophor-thiamine interactions has led to increased interest in thiamine treatments designed for fry and adult stages.

Thiamine-deficient fry may be treated with yolk sac thiamine injections or immersion in a 1% thiamine HCl bath for 1 hour; both approaches have been shown to effectively eliminate TDC-related mortalities (Fisher et al. 1996; Wooster et al. 2000; Lee et al. 2009). These studies show that fry absorb aqueous thiamine, and gill tissues are plausible sites of thiamine absorption from the environment. However, whether gill epithelial cells transport thiamine from the environment across the cell membrane is unknown, and this route of uptake has yet to be experimentally demonstrated. One risk of treatment at the sac fry stage is the potential for sublethal effects resulting from short-term thiamine deficiency, but this mechanism has not been investigated in fishes (Mimouni-Bloch et al. 2014) (see *Lethal and sublethal effects in fry* below).

Although egg or fry immersion is an easy method of treating large numbers of individuals simultaneously, injection of prespawn females can achieve far greater increases in offspring thiamine concentrations (Ketola et al. 2000). Injection of prespawn coho salmon females with 50 mg thiamine HCl/kg body mass (hereafter, mg/kg) ~1 month prior to spawning can result in a 28-fold increase in egg total thiamine concentration relative to eggs from untreated females (untreated: 0.767 nmol/g; treated: 21.925 nmol/g) (Fitzsimons et al. 2005a). In systems with small numbers of returning females, or where females will be released to spawn in the natural environment, injection of prespawn females is a viable alternative to egg immersion.

As with egg and fry immersion, the magnitude of increase in egg thiamine concentration resulting from prespawn female injection is dosage-dependent (Wooster et al. 2000). Across separate studies, injection concentrations of 50 and 100 mg/kg have resulted in significant increases in egg thiamine concentrations for Atlantic salmon, coho salmon, and steelhead trout

(Koski et al. 1999; Fitzsimons et al. 2005a; Futia et al. 2017). Injection of a much lower thiamine concentration (7 mg/kg) did not produce a significant effect on egg thiamine concentrations for Atlantic salmon females but did significantly decrease fry mortality from 98.6% in control females to 2.1% in thiamine-injected females (Ketola et al. 2000). This result demonstrates that large changes in egg thiamine concentration may not be required for increased fry survival in the short term, especially when egg thiamine concentrations are in the range of the dose-response curve where small increases in thiamine are expected to produce relatively large changes in survival. However, no study has continued to track offspring from treated and untreated females to detect potential sublethal effects of low egg thiamine at later life stages.

The amount of time between treatment and spawning may also impact the efficacy of treating prespawn females, although this effect has never been measured in a controlled manner. In two separate studies (both using a thiamine injection concentration of 50 mg/kg), treating steelhead trout females ~4 months prior to spawning resulted in a 6-fold increase in egg total thiamine concentration (untreated: 3.4 nmol/g; treated: 20.4 nmol/g) (Futia et al. 2017), whereas treating coho salmon females ~1.3 months prior to spawning yielded a 28-fold increase in egg total thiamine concentration (untreated: 0.767 nmol/g; treated: 21.925 nmol/g) (Fitzsimons et al. 2005a). The 28-fold increase in egg thiamine concentration observed in the coho salmon study may be due to females allocating thiamine to eggs in greater proportions when spawning is imminent as opposed to when spawning is 4 months in the future, though some of these differences may also reflect interspecific differences in how excess thiamine is allocated.

1.5 Causes of thiamine deficiency complex in natural populations

1.5.1 M74 in the Baltic Sea

Diverse causes of TDC, which may not be mutually exclusive, have been proposed. In 1974, anadromous Atlantic salmon sac fry in Baltic Sea hatcheries began exhibiting behavioral signs typical of thiamine deficiency, including ataxia and lethargy (Amcoff et al. 1998), and affected families experienced 100% mortality (Lundström et al. 1999b). Since the early 1990s, returning adult salmon have also displayed signs of thiamine deficiency, including lack of coordination and sideways swimming (Amcoff et al. 1998). This disorder—originally termed M74, for miljöbetingad (literally translated as “environmental”) and the year of its discovery

(Bylund and Lerche 1995)—was eventually linked to low levels of thiamine at the egg and fry stages of affected individuals (Amcoff et al. 1998). Using data spanning 25 years, Karlsson et al. (1999) identified correlations suggesting that over-consumption of native sprat (*Sprattus sprattus*) had led to thiamine deficiency in Atlantic salmon. For the salmon whose diet consisted primarily of sprat, frequency of TDC was positively correlated with sprat biomass (Karlsson et al. 1999).

Sprat, along with Atlantic herring (*Clupea harengus*), comprise a large proportion of Atlantic salmon diets in the Baltic Sea (Hansson 2001) and tissues of both forage species exhibit comparable degrees of thiaminase activity (Wistbacka et al. 2002; Wistbacka and Bylund 2008). Thiaminase consumption is hypothesized to be the cause of M74, but other dietary factors may exacerbate thiamine deficiency. For example, young, small sprat have higher fat and lower thiamine contents than older, larger sprat. Thiamine:fat content increases as fish age, and young sprat contain roughly half as much thiamine per gram of fat as older sprat (range for sprat aged 1-5 years: 16-122 nmol thiamine/g fat; aged 6-13 years: 48-248 nmol/g) (Keinänen et al. 2012). The combination of high fat and low thiamine concentrations is hypothesized to be particularly detrimental for thiamine-deficient fishes for two main reasons. First, consumption of high amounts of easily oxidized fatty acids can lead to oxidative stress in Atlantic salmon, and this stress can be exacerbated when thiamine—which can function as an antioxidant (Lukienko et al. 2000)—occurs in low concentrations relative to lipids (Keinänen et al. 2012). Development of thiamine deficiency has also been associated with low concentrations of other antioxidants in salmonine tissues, including astaxanthin (Pettersson and Lignell 1999) and vitamin E (Palace et al. 1998), but a correlation between TDC and low levels of other antioxidants is not always present (Brown et al. 2005a). Low availability of other antioxidant vitamins may also increase overall demand for thiamine. If an individual is already thiamine deficient, diverting thiamine from its role as a cofactor for thiamine-dependent enzymes to the role of an antioxidant may intensify the symptoms of thiamine deficiency. Second, oxidative stress may decrease the activity of thiamine-dependent enzymes (Park et al. 1999) as well as levels of available thiamine (Gibson and Zhang 2002).

As early as 1942, herring fat content was considered as a possible cause of a fatal disorder in herring-fed trout (later recognized as thiamine deficiency), but dietary fat content alone did not induce presentation of the disorder (Wolf 1942). Similarly, decreased antioxidant vitamin concentrations in salmonines (astaxanthin and vitamin E) and in forage species (low vitamin E in alewife; Honeyfield et al. 2012) have been associated with thiamine deficiency in salmonines, but

these correlations have all been detected in the presence of dietary thiaminase. Furthermore, treatment of thiamine-deficient eggs with astaxanthin does not affect fry mortality rate in steelhead trout (Hornung et al. 1998). Occurrences of TDC in the Baltic Sea may therefore be due to thiaminase consumption and compounded by consumption of low-thiamine prey or oxidative stress, but no evidence to date supports oxidative stress or low dietary antioxidants as direct, causative agents of TDC.

1.5.2 Thiamine deficiency complex in the Great Lakes

The first documentation of a thiamine deficiency in Great Lakes fishes was an investigation of a "dietary disease" observed in hatcheries and associated with the usage of particular forage fish species (Wolf 1942). Brown trout, brook trout (*Salvelinus fontinalis*) and rainbow trout (*Oncorhynchus mykiss*) fed diets containing varying proportions of herring (*Clupea* sp.) developed symptoms of thiamine deficiency (Wolf 1942), similar to feeding trials with farm-raised foxes that were fed raw fish diets and developed Chastek paralysis, a known thiamine deficiency disorder (Green and Evans 1940). Wolf (1942) concluded that certain forage fish contained a "vitamin B_1 destroying" capability (later identified as thiaminase) that was able to induce thiamine deficiency in hatchery reared trout fed raw forage fish.

In 1968, a few years prior to the first description of M74, emergence of TDC in populations of salmon and trout from the Great Lakes was documented for the first time in fishery agency reports. The condition manifested primarily as increased rates of mortality in hatchery culture prior to swim-up (the developmental stage where yolk sac absorption is complete and the fry become completely dependent on exogenous feeding) in coho and Chinook salmon, brown trout, and steelhead trout, and was originally named "early mortality syndrome" (EMS, now known as TDC). From 1968 through 1992, rates of TDC-induced mortality varied among species, but did not exceed 30% for families in hatchery culture. In 1993, mortality reached up to 90% for juvenile coho salmon in Lake Michigan hatcheries, and mortality also began to occur earlier at the sac fry stage in addition to mortality at the swim-up stage (Marcquenski and Brown 1997).

Initially, research in the Great Lakes aimed to determine whether environmental contaminants—such as polychlorinated biphenyls (PCBs), pesticides, dioxins, furans, and polynuclear aromatic hydrocarbons (PAHs)–were associated with the observed mortality patterns for affected species, which included lake trout, Chinook salmon, coho salmon, steelhead trout, and

brown trout (Skea et al. 1985; Mac et al. 1985, 1993; Fitzsimons et al. 1995; Marcquenski and Brown 1997). Maternal exposure of salmonine eggs to dioxins and other contaminants that activate the aryl hydrocarbon receptor causes a similar syndrome that results in mortality of sac fry (King-Heiden et al. 2012). However, egg contaminant levels in the Great Lakes decreased below levels expected to cause mortality by the early 1990s and could not explain the observed rates of juvenile mortality (see Tillitt et al. 2008 for a review). In 1995, attention turned to a nutritional basis for TDC in wild salmonine populations when thiamine treatments for sac fry (immersion and injection) were shown to significantly decrease mortality, whereas treatment with four other B-vitamins proved ineffective (Fitzsimons 1995). The identification of thiamine deficiency in sac fry effectively refocused the search for the cause of TDC (Fitzsimons 1995; Fisher et al. 1995).

1.5.3 Thiamine loss due to dietary thiaminase

By the 1960s, two nonnative species—alewife and rainbow smelt (*Osmerus mordax*)—were well established throughout the Great Lakes (Miller 1957; MacCallum and Regier 1970) and comprised a large portion of Great Lakes salmonine diets (Stewart et al. 1981; Jude et al. 1987), a trend which continues to the present (Ray et al. 2007; Jacobs et al. 2013; Roseman et al. 2014; Happel et al. 2018). Because alewife, rainbow smelt, and other thiaminase-containing prey species supply an otherwise sufficient amount of dietary thiamine to consumers (Tillitt et al. 2005; Honeyfield et al. 2012), TDC in the Great Lakes is likely caused by thiaminase consumption, rather than insufficient tissue thiamine concentrations in the salmonine forage base. Lake trout diets dominated by alewife and smelt have been correlated with an 89-94% decrease in lake trout egg thiamine concentrations (Fitzsimons and Brown 1998), indicating consumption of these two species—and thiaminase—as a potential cause of TDC in the Great Lakes.

Experimental studies have confirmed that consumption of thiaminase (Honeyfield et al. 2005; Houde et al. 2015) and diets comprised entirely of alewife (Honeyfield et al. 2005) can decrease tissue and egg thiamine concentrations and induce TDC. Salmonine egg thiamine concentrations have also been inversely correlated with alewife abundance (Riley et al. 2011), providing further evidence that the consumption of alewife (and thereby thiaminase) may be driving salmonine TDC occurrences in the Great Lakes.

1.5.4 Sources and activity of thiaminase

In the presence of an appropriate electron donor (nucleophile), thiaminase degrades thiamine into a thiazole-nucleophile moiety and a pyrimidine moiety (Fig. 1.2). The first bacterium to be isolated and identified as a thiaminase producer, *Bacillus thiaminolyticus*, was named to reflect this property (Kuno 1951). Now reclassified as *Paenibacillus thiaminolyticus*, this bacterium has been successfully cultured from entire homogenized viscera (intestines, liver, heart, and spleen) of alewife (Honeyfield et al. 2002), indicating intestinal tissues of forage fishes as potential sources of thiaminase. This hypothesis is consistent with later studies that demonstrated that thiaminase activity is highest in alewife splenic and bacteria-rich intestinal tissues compared to muscle and gill tissues (Zajicek et al. 2005; Kraft et al. 2014). Interestingly, a similar observation has been made for a plant species: bracken fern rhizome tissue, which contains much greater diversity and abundance of symbiotic bacteria than other plant tissues (Berendsen et al. 2012), exhibits higher thiaminase activity than stipe or leaf tissues, consistent with a potential microbial source of thiaminase (Kraft et al. 2014). Six other microbes are known to produce a thiaminase I-type enzyme (Kraft and Angert 2017). One of these (*Clostridium sporogenes*) is commonly found in the GI tracts of ruminant mammals that commonly die from thiamine deficiency associated with changes in their GI tracts (Hatheway 1990), but these microbes have not been described in fish intestinal tracts. Gene sequences of additional bacterial species indicate that they are capable of producing thiaminase I enzymes (C. E. Kraft, personal observation).

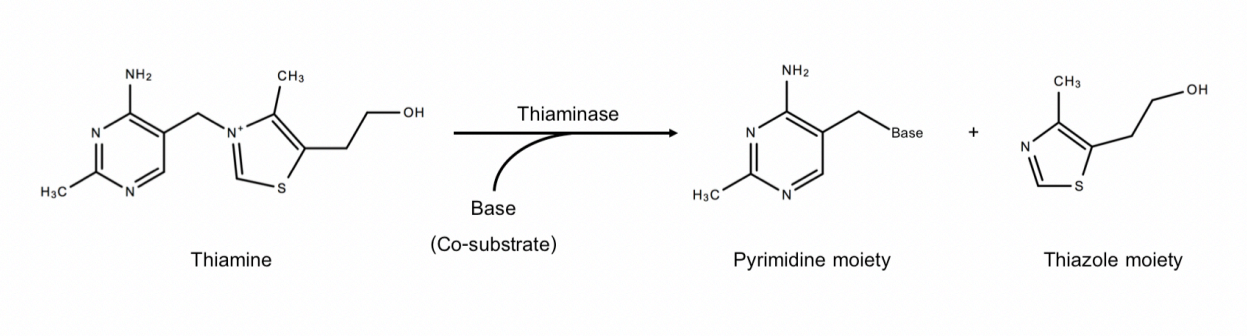


Figure 1.2. Degradation of thiamine by thiaminase into its constituent thiazole and pyrimidine moieties. The initiation of this reaction in food consumed by a thiamine auxotroph irreversibly decreases the amount of thiamine available to the consumer, as auxotrophs cannot recycle the products of the reaction for thiamine synthesis.

Thiaminase may also be produced by fishes endogenously. Identifications of putative thiaminase enzymes with distinct N-terminal protein sequences in carp (*Cyprinus* sp.; Boś and

Kozik 2000) and red cornetfish (*Fistularia petimba*; Nishimune et al. 2000) liver homogenates prompted further investigation into possible production of thiaminases by fishes. Richter et al. (2012) found that thiaminase activity in the tissues of multiple fish, zooplankton, and mussel species was not related to *P. thiaminolyticus* cell abundance or to the concentration of *P. thiaminolyticus* thiaminase. This result suggests that *P. thiaminolyticus* may not be the primary thiaminase producer in salmonine prey species. It is possible that other bacterial species contribute to thiaminase production or that alewife and other fishes produce their own thiaminase enzymes *de novo*. The hypothesis that fishes produce thiaminase is also interesting from an evolutionary perspective. Animals cannot produce their own thiamine and therefore cannot recycle the products of thiamine degradation for the purpose of thiamine synthesis. That any animal would produce an enzyme capable of destroying an essential nutrient seems counterintuitive. Highly speculative hypotheses include that thiaminase may play a role in innate immunity by restricting the thiamine available to bacteria in the gut or that thiaminase protects fishes from toxic thiamine analogs in the diet. Production of extracellular thiamine analogs is hypothesized to be important in ecological competition for thiamine (Kraft and Angert 2017).

Even though thiaminase activity has been detected in a variety of Great Lakes species, highly variable thiaminase activities among and within species complicate efforts to describe the origin(s) of thiaminase. In alewife tissues, thiaminase activity activity has been shown to vary by year, season, sampling location, and individual size (Ji and Adelman 1998; Tillitt et al. 2005; Honeyfield et al. 2012). In rainbow smelt, thiaminase activity varies significantly between years in some studied populations (Ji and Adelman 1998; Honeyfield et al. 2012), but not others (Tillitt et al. 2005). Activity has also been shown to vary by season and collection depth in invasive dreissenid mussels (Tillitt et al. 2009) and relative to reproductive status in native unionid mussels (Unionidae; Blakeslee et al. 2015). Atlantic herring and alewife have exhibited substantial increases in thiaminase activity levels when transferred to captivity from the wild, but these increases were not associated with any other explanatory variable (Lepak et al. 2008; Wistbacka and Bylund 2008). Despite descriptions of these patterns, no single environmental variable has been found to predict variation in thiaminase activity across multiple systems or time points for any species.

Although environmental predictors of thiaminase activity are not well established, trophic patterns have emerged across studies examining thiaminase activity in fishes. Fishes feeding at

lower trophic levels are more likely to exhibit thiaminase activity than fishes feeding at higher trophic levels (Riley and Evans 2008). Taxonomic patterns have also been described, with basal teleosts (i.e., Anguilliformes, Clupeiformes, Cypriniformes, and Siluriformes) more likely to exhibit thiaminase activity than basal euteleosts (i.e., Esociformes, Osmeriformes, and Salmoniformes) or neoteleosts (i.e., Gadiformes, Gasterosteiformes, Perciformes, and Scorpaeniformes; Riley and Evans 2008). Specifically, clupeids including Atlantic herring, alewife, and sprat have been implicated in emergence of increased TDC, and these species often exhibit high thiaminase activity in their tissues (Wistbacka and Bylund 2008; Kraft et al. 2014). Thiaminase activity has also been detected in other clupeids, including hickory shad (*Alosa mediocris*), threadfin shad (*Dorosoma petenense*), and gizzard shad (*Dorosoma cepedianum*; Tillitt et al. 2005; Honeyfield et al. 2007; Kraft et al. 2014), suggesting that either endogenous thiaminase production or symbiosis with thiaminase-producing bacteria is common to members of Clupeidae (Fitzsimons et al. 2012). Several members of Cyprinidae including the carps (*Cyprinus* spp.), blacknose dace (*Rhinichthys atratulus*), longnose dace (*Rhinichthys cataractae*), central stoneroller (Campostoma anomalum), common shiner (Luxilus cornutus), spottail shiner (*Notropis hudsonius*), creek chub (*Semotilus atromaculatus*), cutlips minnow (*Exoglossum maxilingua*), and fallfish (*Semotilus corporalis*) also exhibit thiaminase activity (Boś and Kozik 2000; Tillitt et al. 2005; Kraft et al. 2014). Thiaminase activity has also been observed in species belonging to Catostomidae (white sucker, *Catostomus commersonii*, Kraft et al. 2014) and Osmeridae (rainbow smelt, Tillitt et al. 2005), but the families Clupeidae and Cyprinidae are regularly recognized as exhibiting high thiaminase activities that are associated with thiamine deficiency in consumers.

1.6 Effects of thiaminase deficiency

The majority of research examining TDC in fishes has relied on data collected from the offspring of spawning adults caught in the wild and reared in hatcheries or laboratories. Observing the development of thiamine-deficient eggs and fry in culture or in the laboratory has allowed for detailed descriptions of the physical and behavioral signs of thiamine deficiency, but these data may not accurately describe developmental outcomes for deficient offspring in the wild. In hatcheries or laboratory environments, fry in culture are usually first given powdered feed near the time of yolk sac absorption (approximately 5 weeks after hatch, depending on water temperature). However, wild lake trout fry begin feeding within two weeks of hatching, with nearly all

individuals (98%) successfully feeding on live prey by the completion of yolk sac absorption (Swedberg and Peck 1984; Ladago et al. 2016). This behavior indicates that some offspring of thiamine-deficient parents may be able to mitigate deficiency by feeding early. In Lake Champlain, invasion of alewife led to significant declines in egg thiamine in lake trout and Atlantic salmon but was followed by the onset of substantial and sustained recruitment of wild juvenile lake trout (Marsden et al. 2017). However, in cases of severe thiamine deficiency, fry may be unable to forage regardless of food availability, as lake trout egg thiamine concentrations of 6.9 and 2.9 nmol thiamine/g egg are associated with 20% and 50% reductions in foraging rate, respectively (Fitzsimons et al. 2009).

1.6.1 Lethal and sublethal effects in fry

The physical and behavioral signs of thiamine deficiency are similar among salmonines. Thiamine-deficient fry exhibit a suite of physical signs, including hydrocephalus, vascular congestion, diminished yolk sac conversion efficiency, and large yolk sacs with opacities, edema, and hemorrhaging (Fisher et al. 1995; Lundström et al. 1999a; Fitzsimons et al. 2001b, a). Perhaps even more striking than the physical consequences of thiamine deficiency are the resulting behavioral abnormalities. These behaviors include lethargy, ataxia, and unusual swimming patterns such as bouts of uncoordinated swimming followed by variable intervals of passive drifting and swimming in a "corkscrew" or "wiggling" manner (Fisher et al. 1995; Marcquenski and Brown 1997; Fitzsimons et al. 2005a). Experimental induction of thiamine deficiency leads to similar physiological and behavioral shifts in juveniles of other species, including sterlet (*Acipenser ruthenus*) (Ghiasi et al. 2017), Nile tilapia (*Oreochromis niloticus*) (Lim et al. 2011), and Japanese eel (*Anguilla japonica*) (Hashimoto et al. 1970). In salmonine fry, death usually occurs after the onset of behavioral signs and prior to swim-up.

In addition to TDC-induced mortality, sublethal effects of TDC have been described in lake trout. For some of these effects, egg thiamine thresholds required to produce effects of specific magnitudes have been determined. For example, an egg thiamine concentration of 5.1 nmol/g was associated with a 50% reduction in growth rate of lake trout fry relative to thiamine-replete fry (Fitzsimons et al. 2009). Reduced growth due to thiamine deficiency can profoundly impact fry survival, as larger body size allows larvae to achieve and sustain faster swimming speeds, which may in turn decrease predation vulnerability and increase prey capture efficiency (Miller et al.

1988). The threshold concentration for 50% reduced growth is more than three times higher than the egg thiamine concentration associated with 50% TDC-induced mortality in lake trout (ED_{50} = 1.6 nmol/g, 95% CI [1.1, 2.1]) (Fitzsimons et al. 2007), suggesting that thiamine deficiency can impact long-term survival at egg thiamine concentrations far above those required to induce direct mortality. Additional sublethal effects of diminished egg thiamine concentrations on fry include reduced foraging rate (Carvalho et al. 2009; Fitzsimons et al. 2009), decreased visual acuity (Carvalho et al. 2009), and altered immune responses (Ottinger et al. 2012). Many of these sublethal effects likely translate into reduced survival and fitness later in life.

1.6.2 Thresholds for thiamine deficiency

Many studies have focused on determining how much thiamine must be allocated to an egg to prevent TDC-induced mortality of the hatched fry. Although egg threshold determination methods differ across studies, concentrations of free (unphosphorylated) egg thiamine required to avoid mortality from TDC during development range from 0.3 nmol/g (in Lake Michigan steelhead trout, Hornung et al. 1998) to 0.8-1.0 nmol/g (in Lake Michigan lake trout, Czesny et al. 2009). In many cases, variability has been observed among sampling locations, sampling years, and species for the relationship between egg thiamine concentrations and subsequent rates of mortality (Hornung et al. 1998; Honeyfield et al. 2005; Fitzsimons et al. 2007; Futia et al. 2017). This variation may result from differences in environmental characteristics during egg development and maturation (i.e., maternal effects), such as maternal diet, or may relate to uncharacterized patterns of adaptive genetic or epigenetic variation among families or populations.

Even when thiamine levels are not sufficiently low to result in TDC-related mortality, sub-lethal effects may result in demographic declines at the population level. Although increasing awareness of TDC has resulted in increased monitoring efforts for lake trout in the Great Lakes (Riley et al. 2011) and for other species in other systems (Honeyfield et al. 2016), thresholds for sublethal effects in other species have never been empirically established. Reliably assessing the risk of sublethal effects for species other than lake trout is not currently possible. Establishing lethal and sublethal effect thresholds for other species would expand the capacity for more accurate demographic projections for declining populations potentially affected by TDC.

Because of the negative relationship described between egg thiamine concentration and rates of TDC-related mortality for salmonines, applied studies have sought to define diagnostic maternal

characteristics that can be used to predict egg thiamine concentrations. For females that are sufficiently thiamine-depleted to produce offspring with TDC, no relationships have been detected between maternal tissue thiamine concentrations and egg thiamine concentration (Fisher et al. 1998; Futia et al. 2017). This absence of correlation may be common in a heavily impacted population, as maternal tissue and egg thiamine concentrations may be uniformly low with insufficient variation among individuals to describe a relationship that could be found if a population contained individuals with a wider range of thiamine statuses. Some data from affected females indicate an inverse relationship between maternal weight or length and egg thiamine concentrations (Brown et al. 2005a; Wolgamood et al. 2005; Werner et al. 2006; Honeyfield et al. 2008; Futia et al. 2017), suggesting possible variation in salmonine diet composition related to age and size (e.g., older, longer females may consume alewife in higher proportions than younger females). However, this hypothesis has not been experimentally tested, and questions regarding maternal thiamine allocation priorities and the physiological mechanisms of egg thiamine deposition still need to be addressed.

1.6.3 Thiamine deficiency in adults

While much research has focused on how developmental outcomes for eggs and fry are affected by TDC, fewer experiments have examined how thiamine deficiency directly impacts adult fish health. In Atlantic salmon sampled for hatchery rearing while returning to the River Luleälven from the Baltic Sea, females exhibiting "wiggling, sideways swimming, [and] a lack of coordination" were found to have significantly lower ovarian thiamine concentrations than returning females with normal behavior (Amcoff et al. 1998). Of 103 females displaying abnormal swimming behavior, 27 died before being spawned, with mortality attributed to TDC (Amcoff et al. 1998). While severe thiamine deficiency can directly lead to adult mortality (Amcoff et al. 1998; Brown et al. 2005c; Fitzsimons et al. 2005a), sublethal effects of deficiency may also reduce the fitness of affected populations. In two-year-old (subadult) Atlantic salmon, individuals fed a high-thiaminase diet for 8 months exhibited decreased tissue thiamine (as measured in red blood cells, white muscle, and liver), pronounced changes in body morphology and pigmentation, and decreased swimming performance (Houde et al. 2015). Although the experimental thiaminase diet did not impact survival over the course of the study, the traits that were affected by the thiaminase diet have previously been shown to impact Atlantic salmon survival. Thiamine deficiency can also

impact the ability of migrating salmonines to ascend cascades (Ketola et al. 2005) and decrease the number of attempts an individual fish will make to traverse challenging river reaches (Harbicht et al. 2018).

1.6.4 Thiamine deficiency in non-salmonine species

It may be particularly difficult to forecast demographic effects of reduced tissue thiamine concentrations for non-salmonine species that have not been frequent subjects of thiamine deficiency research, such as anguillids. An assessment of the Lake Ontario-upper St. Lawrence River population of American eels indicated that for some individuals, muscle thiamine concentrations approached thresholds associated with TDC-induced behaviors in Japanese eels (Hashimoto et al. 1970; Fitzsimons et al. 2013). However, thresholds may differ interspecifically for anguillids, as they do for salmonines, making direct comparisons between American and Japanese eels challenging. Resolving thresholds for American eels might be particularly important given that adult eels cease feeding and assimilating nutrients prior to migration (Pankhurst and Sorensen 1984) as do many salmonines (Quinn 2005). Migratory individuals must therefore rely on endogenous energy stores throughout most of migration and spawning—two metabolically demanding activities. The quantity of thiamine required for a migrating individual to successfully reproduce is unknown, but determination of species-specific thresholds could help predict overall spawning rates for populations affected by thiamine deficiency.

1.7 Conservation and management of fishes in TDC affected systems

1.7.1 Evolutionary rescue

In some systems, the introduction of a prey species known to cause TDC in consumers may threaten native populations or interfere with native species reintroduction efforts (Fisher et al. 1996; Marsden and Hauser 2009; Riley et al. 2011; but see Marsden et al. 2018). In these scenarios, it could be helpful to ascertain whether populations could evolve to tolerate low dietary thiamine availability (i.e., prey with thiaminase activity or with low thiamine:lipid ratios). Susceptibility to TDC is known to vary across salmonine species, with adult Chinook salmon remaining asymptomatic at lower muscle thiamine concentrations (Honeyfield et al. 2008) than coho salmon, lake trout, and steelhead trout (Brown et al. 2005c). Egg thiamine thresholds for fry thiamine

deficiency are also lower in Chinook than in coho salmon and lake trout (Fitzsimons et al. 2007), further suggesting that some species may be genetically adapted to require less thiamine than other species. Recent empirical evidence from a comparison of three Atlantic salmon strains fed diets high in thiaminase suggests that some locally adapted salmon populations may be able to tolerate lower thiamine concentrations than populations with less exposure to thiaminase-containing prey (Houde et al. 2015). Two landlocked strains that typically consume thiaminase-containing prey in their native habitats experienced smaller reductions in tissue thiamine than an anadromous strain with a more diverse diet lacking thiaminase (Houde et al. 2015). This result indicates that adaptation to low available thiamine may be possible at the population level.

One mechanism by which populations could adapt to low thiamine conditions is by responding to selection on genes associated with thiamine-dependent biochemical pathways. Sequence variation in the genes encoding the four TDP-dependent enzymes necessary for carbohydrate and branched-chain amino acid metabolism (Fig. 1.1) (Depeint et al. 2006) could result in enzymes with differing binding affinities for TDP. This variation could be subject to selection if altered cofactor affinities facilitate more efficient exploitation of limited thiamine resources. For example, genetic mutations in one thiamine-dependent enzyme complex (branched-chain α-ketoacid dehydrogenase, BCKDH; Fig. 1.1) results in maple syrup urine disease in humans, with symptoms including ataxia, neurologic impairments, and blindness (Ames et al. 2002). A missense mutation in the gene encoding one subunit of the BCKDH complex decreases the complex's affinity for TDP and, in turn, increases cellular requirement for TDP and thiamine (Chuang et al. 1982; Fisher et al. 1991; Ames et al. 2002). Variation in protein coding sequences (exons) of genes encoding TDP-dependent enzymes could therefore correspond to variation in thiamine requirements among individuals, populations, or species.

Non-protein coding (intronic) variation in thiamine-related genes could also contribute to differing patterns of thiamine use among individuals. For example, specific point mutations in an intron of the gene encoding thiamine pyrophosphokinase (*TPK*)—the enzyme that converts thiamine to TDP (Jurgenson et al. 2009)—are associated with variation in human birth weight (Fradin and Bougneres 2007). This relationship may be due to the ability of intron sequences to affect gene expression (Nott et al. 2003). Variation in *TPK* intron sequence could lead to varying patterns of TPK expression, altered availability of TDP, and metabolic shifts related to TDP abundance. Thiamine transporters, which are responsible for cellular and mitochondrial thiamine

uptake (Bettendorff 2013), could also be subject to selection. In humans, the *SLC23A1* gene encodes an active transporter protein that facilitates the absorption of L-ascorbic acid (vitamin C, an essential vitamin) in the intestine and kidneys. A single missense mutation in one *SLC23A1* exon is associated with decreased blood concentrations of vitamin C. This effect is hypothesized to result from a conformational change in the transporter protein that impairs uptake of vitamin C (Timpson et al. 2010). The efficacy of thiamine transporters could be similarly affected by mutations in corresponding genes.

Genes associated with gut microbiome composition could also play a role in host thiamine intake requirements. Host genotypes that promote the presence of beneficial microbes are plausible targets of positive selection if those microbes perform tasks (e.g., vitamin synthesis) that increase host fitness (Goodrich et al. 2016). If fishes exploit thiamine produced by the gut microbiome (Ji et al. 1998), fish genotypes associated with increased abundance of thiamine-producing microbial taxa could reasonably be subject to positive selection in ecosystems where thiamine uptake via piscivory is limited by the presence of thiaminase-containing prey. The majority of microbiome thiamine production occurs in the lower intestine in fishes (Ji et al. 1998), but the upper intestine is the primary site of thiamine absorption (Bettendorff 2013). Although epithelial cells in the human lower intestine are known to express thiamine transporter proteins and are capable of absorbing thiamine (Said et al. 2001; Nabokina and Said 2012), the degree to which fishes can absorb thiamine at the site of production in the lower intestine is unknown. Coprophagy may also provide fishes with access to thiamine produced in the lower intestine, and has been observed in Atlantic salmon smolts (Nylund et al. 1994), but the prevalence of this behavior among salmonine species and developmental stages is also unknown. Nevertheless, research dedicated to identifying associations between host genotype, diet, and microbiome composition (e.g., Goodrich et al. 2014; Carmody et al. 2015) could open promising avenues of research for investigating thiamine-related genetic adaptations in fishes.

In scenarios of evolutionary rescue, genetic adaptation allows for the recovery of a declining population that would have otherwise been extirpated due to environmental factors (Carlson et al. 2014). If individuals that can tolerate low available thiamine successfully reproduce at rates higher than those that cannot tolerate low thiamine, this process could lead to an increase in the proportion of individuals adapted to low available thiamine. The strength of the response to selection would depend on the degree to which low thiamine tolerance has a heritable, genetic basis. If selection

on thiamine requirement is possible, management efforts for an affected population should prioritize conservation of adaptive genetic variation in that population. In natural populations, this could be achieved by maximizing effective population sizes (Waples 1990), limiting additional sources of mortality (e.g., fishing pressure) and maximizing spawning and rearing opportunities (e.g., increasing or improving spawning habitat). Facilitating genetic adaptation to low available thiamine could be achieved in hatchery-dependent populations by modifying existing egg thiamine treatment regimes. For example, withholding supplemental thiamine from a portion of each family of eggs would allow for differential survival among untreated individuals with varying tolerances to low thiamine conditions. Over several generations, this approach could result in an increase in the proportion of the population capable of tolerating low thiamine availability. However, adaptation to low thiamine may not be possible for all populations and species and this procedure would likely result in an unintentional response to selection for other differences between the hatchery and wild environments (Christie et al. 2014). Thus, in most cases, it will be important to first identify the proximate cause of thiamine deficiency in the population.

1.7.2 Diagnosis and mitigation of TDC in fish populations

In some systems, identifying the root cause of TDC may be straightforward. In Lake Champlain, for example, TDC was not observed in lake trout or landlocked Atlantic salmon until after the introduction and establishment of alewife (Ladago et al. 2016), a species that had already been associated with occurrences of TDC in lake trout and Atlantic salmon in the Great Lakes. In other cases, the identification of TDC may not be as clear cut. The signs of severe thiamine deficiency described for salmonines seem to be consistent among other fishes but may also be initially interpreted as evidence of a novel pathogen or contaminant in a previously unaffected system. In less severe cases, sublethal effects of TDC may be influencing demographic trends in the population, but could easily remain undetected or be misattributed, at least in part, to any co-occurring biotic or abiotic change (e.g., eutrophication, temperature shifts, changes in prey species demographics).

Once TDC has been identified as the most likely cause of observed signs or population decline, establishing the proximate cause(s) of the deficiency can guide management approaches. For example, in systems where establishment of an introduced species has led to TDC emergence, as with alewife in the Great Lakes and Lake Champlain, suppression of the exotic species would

be expected to relieve the deficiency. Managing for diverse forage fish communities that include fishes with no or low thiaminase activity may help to displace thiaminase-containing prey from the affected species' diet, but as in the Great Lakes, such decisions are often complex and occur in the context of a wide range of economic and social issues (Dettmers et al. 2012).

1.8 Research questions and needs

The ability to detect TDC in aquatic organisms as early as possible would give managers and researchers greater capacity to identify and mitigate the proximate cause. However, detection of mild to moderate TDC is a challenging task, as "normal" tissue thiamine concentrations for fishes remain undefined (but see Table 1.1 for examples of tissue thiamine concentrations in wild caught fishes without apparent TDC). Thiamine deficiency research in fish culture environments has allowed for the determination of thresholds for detrimental effects of TDC at the population level using thiamine levels in eggs but has not produced analytical approaches for determining where an individual falls on a continuum from thiamine replete to thiamine deficient. Other research needs relevant to management include identification of environmental variables affecting thiaminase production in forage species and descriptions of thiamine deficient fry in natural environments (Table 1.2).

Several other research questions related to the molecular mechanisms and ecology of TDC are outlined in Fig. 1.3 and Table 1.2. For some topics, little to no research has been conducted in fishes. For example, changes in TDP-dependent enzyme activities following induction and reversal of thiamine deficiency have been documented in brown rats (*Rattus norvegicus*) (Gibson et al. 1984; Butterworth and Héroux 1989); similar experiments in fishes could help understand the mechanisms underlying sublethal effects in deficient fry and adults. Additional research needs, such as an efficient method of assessing thiamine concentrations in open water, are critical for understanding how rates of thiamine production impact nutritional status of species in higher trophic levels. Although not an exhaustive list, we present these questions as an outline for understanding ecological processes responsible for the development of TDC.

Despite extensive observational and experimental evidence of low thiamine in eggs of feral salmonines and TDC in fry derived from feral adults, the linkage of TDC to population-level effects in the wild is primarily supported by correlations between abundance of thiaminase-containing prey, egg thiamine data, and recruitment of wild salmonines (e.g., Fitzsimons et al.

Table 1.1. Example tissue thiamine concentrations in fishes captured in the wild from populations without apparent thiamine deficiency complex and tissue thiamine concentrations as reported for wild caught fishes in the US Department of Agriculture (USDA) National Nutrient Database. Total thiamine concentration is the sum of thiamine vitamer concentrations (i.e., free thiamine, thiamine monophosphate, and thiamine pyrophosphate). Total thiamine concentrations are supplied as means ± standard error. For wild caught samples, the method used to quantify thiamine is cited for each value.

Species	N	Sex	Tissue type	Total thiamine concentration nmol/g tissue	Waterbody	Reference
Wild caught fishes						
Alewife	403	N/A	Whole fish	15.5 ± 0.4[1]	Lake Michigan	Tillitt et al. 2005
Atlantic herring	65	M	Whole fish	7.7 ± 0.7[1]	Baltic Sea	Keinänen et al. 2012
Chinook salmon	2	F	Eggs	11.9 ± 2.2[1]	Harrison River/Pacific Ocean	Welch et al. 2018
Chum salmon	10	F	Eggs	15.6 ± 1.0[1]	Harrison River/Pacific Ocean	Welch et al. 2018
Lake trout	8	F	Eggs	22.9 ± 3.2[1]	Gull Island/Lake Superior	Fitzsimons and Brown 1998
Lake whitefish	20	F	Eggs	11.4 ± 1.4[2]	Lake Huron	Riley et al. 2011
Pink salmon	20	F	Eggs	19.5 ± 0.5[1]	Seton River/Pacific Ocean	Welch et al. 2018
Rainbow smelt	125	N/A	Whole fish	8.8 ± 0.6[1]	Lake Michigan	Tillitt et al. 2005
Sockeye salmon	5	F	Eggs	10.9 ± 0.5[1]	Harrison River/Pacific Ocean	Welch et al. 2018
Sockeye salmon	10	F	Muscle	12.2 ± 1.2[1]	Harrison and Adams Rivers/Pacific Ocean	Welch et al. 2018
USDA National Nutrient Database						
Atlantic herring	43	N/A	Muscle	3.5 ± 0.3	N/A	USDA 2018
Atlantic salmon	11	N/A	Muscle	8.5 ± 1.4	N/A	USDA 2018
Coho salmon	3	N/A	Muscle	4.3 ± 1.1	N/A	USDA 2018
Rainbow trout	5	N/A	Muscle	4.6 ± 0.5	N/A	USDA 2018

[1]*Brown et al. 1998;* [2]*Zajicek et al. 2005*

2010; Riley et al. 2011; He et al. 2012). For example, lake trout egg thiamine concentrations in Lake Huron have been negatively correlated with alewife abundance, and increasing thiamine concentrations (Riley et al. 2011) and natural lake trout recruitment (Riley et al. 2007) were observed following the 2002-2004 collapse of the Lake Huron alewife population. After 2004, Lake Huron lake trout diets were no longer alewife-dominated, and this shift—rather than reduced alewife predation on lake trout fry—seems to be responsible for the increase in lake trout

recruitment (Fitzsimons et al. 2010). However, testing causal relationships between ecological variables and rates of TDC in the field is more challenging because causal relationships cannot be evaluated by collecting correlational data. Capture or *in situ* observation of larval fishes and fry in the wild is difficult and biased—the individuals most susceptible to TDC are least likely to be caught or observable due to early life stage mortality. Consequently, studies of thiamine-deficient populations are conducted on eggs and fry in culture, and results are often extrapolated to wild environments (e.g., Fisher et al. 1996). Influence of TDC on mortality of wild fish is likely to be a result of a complex interaction of factors, including availability of thiamine-rich food, fat content of diet, density of predators, and availability of refuge (Czesny et al. 2009; Fitzsimons et al. 2009; Ladago et al. 2016). Continued research is needed to define the effect of TDC in wild populations and to determine whether TDC-related mortality is sufficient to affect recruitment and abundance of yearling and older year classes.

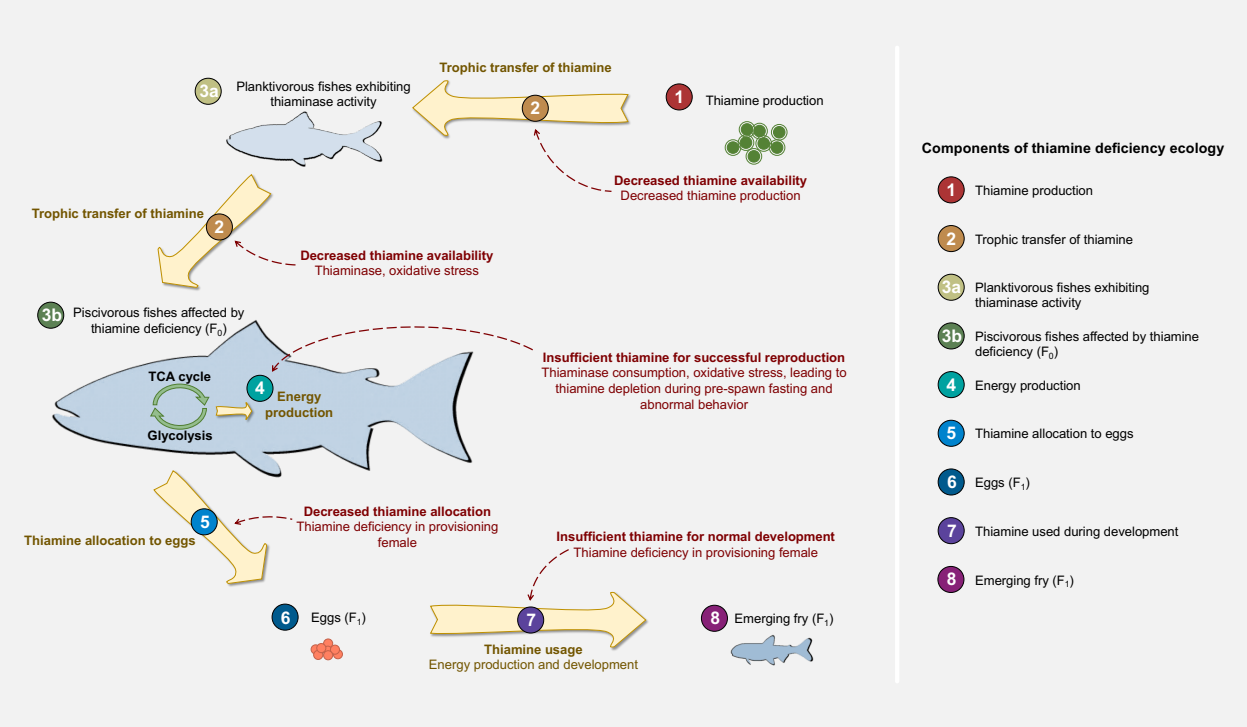


Figure 1.3. Summary of the processes and trophic levels involved in the ecology of thiamine deficiency. Processes are indicated by yellow arrows. Examples of how and why these processes may be impacted in an ecosystem where piscivorous fishes are thiamine deficient are indicated in red text. Numbers (1-8) correspond to table section headings in Table 1.2.

Table 1.2. Research questions and needs relevant to understanding the molecular mechanisms and ecology of thiamine deficiency (TD) in fishes. Numbered headings (1-8) correspond to the numbered processes and trophic levels in Fig. 1.3.

1. Thiamine production

- What is the variability in thiamine content among thiamine-producing species?
- Are rates of thiamine production changing in thiamine producers within aquatic ecosystems?
- How does production of thiamine vary within aquatic ecosystems, spatially and temporally?
- What environmental factors influence rates of thiamine production?
- How can thiamine concentrations be efficiently measured in abiotic environments (water and benthic substrate) in aquatic ecosystems?

2. Trophic transfer of thiamine

- What proportion of thiamine present in a prey item can be incorporated by a consumer?
- Do factors other than thiaminase inhibit transfer of thiamine from prey to consumer (e.g., transport competition by thiamine analogs)?

3. Fishes

- What are the pharmacokinetics (uptake, distribution, metabolism, storage, and elimination) of thiamine in fishes? What molecular mechanisms (e.g., thiamine transport proteins) control thiamine allocation within a fish?
- What role do microbial flora within the gut (i.e., microbiome) play in thiamine availability within a fish? To what extent do microbes in the fish gut influence production or degradation of thiamine?
- Can fishes obtain thiamine directly from the water? If so, what is the mechanism (passive or active) of thiamine transport? In what tissues does uptake occur?

3a. Planktivorous fishes exhibiting thiaminase activity

- Do planktivorous fishes modify feeding preferences or behaviors to compensate for changes in thiamine availability?
- Do any fish species produce thiaminase *de novo*? If so, what is the function of thiaminase production in fishes? How is *de novo* thiaminase production in fishes regulated?
- What is the tissue-specific and subcellular localization of thiaminase in fishes?
- How do fish species with high thiaminase activity protect their own thiamine stores?

Table 1.2 continued

3b. Piscivorous fishes affected by TD

- Are natural populations of fishes other than salmonines and anguillids currently affected by TD?
- What life history characteristics (physiological, behavioral, ecological) of a species predict its vulnerability to TD?
- How much thiamine is required for successful migration and spawning of diadromous species?
- Where and how does thiaminase break down thiamine in the consumer? How is thiaminase released from items in the diet?
- How long does thiaminase activity from a prey species persist within the gut of a predator fish species after consumption?
- What proportion of thiaminase in the diets of piscivorous fishes is of bacterial origin? What proportion is produced endogenously by prey fishes?
- What are the sublethal effects associated with TD? How do they ultimately affect success of a fish species or population?
- What are sufficient blood and tissue levels of thiamine vitamers?
- Does prey lipid content affect its consumer's tissue thiamine concentrations?

4. Energy production

- How does TD affect energy production and efficiencies of energy production in fish?
- What molecular tools or reagents are required to evaluate energy production in fish as affected by TD?

5. Thiamine allocation to eggs in TD-affected fishes

- How do thiamine-deficient females allocate thiamine during egg production? What are the mechanisms and controls?
- Do fish have specific thiamine transport proteins in eggs that regulate uptake during egg development and maturation?
- Do genetic or environmental factors influence thiamine allocation (including egg deposition) within fish?

6. Eggs produced by TD-affected fishes

- What factors influence the observed variation in egg thiamine thresholds for normal development vary among species?

7. Disposition of thiamine in eggs/fry/juveniles during early development

- What factors control thiamine pharmacokinetics during early development in fishes?
- Can eggs acquire thiamine from the water during incubation? At what rate?
- Do rates of thiamine metabolism or usage during development vary among species?
- What environmental factors influence thiamine uptake, distribution, metabolism, and elimination in developing fish?

Table 1.2 continued

8. Emerging fry produced by TD-affected fishes • How does TD affect fry under natural environmental conditions? Are observations of TD in hatchery-reared fry predictive of expected outcomes in the wild, under natural conditions? • Can fry hatched from thiamine deficient eggs avoid the lethal and sublethal effects of deficiency through early feeding? • What is the role of TD in limiting recruitment of salmonine fishes in wild populations?

1.9 Conclusion

A recent horizon scan (a method designed to identify emerging issues in conservation) included thiamine deficiency in wildlife populations as a potential threat to multiple taxonomic groups (Sutherland et al. 2018), indicating a need for increased research and awareness. Although extensive data support thiaminase consumption as a causative factor for TDC in fishes, research is still needed to identify and describe other ecological and environmental factors that may influence the manifestation of TDC. Although viable treatment options are available for mitigating existing thiamine deficiency in hatcheries, investigating why thiamine deficiency occurs in natural systems will allow development of predictive and preventative management strategies *in situ*. Increasing awareness of the potential causes of thiamine deficiency and of available treatment and management options represents a first step toward diminishing demographic losses in systems in which TDC occurs.

1.10 References

Amcoff P, Börjeson H, Lindeberg J, Norrgren L (1998) Thiamine concentrations in feral Baltic salmon exhibiting the M74 syndrome. In: Early life stage mortality syndrome in fishes of the Great Lakes and Baltic Sea. American Fisheries Society, Bethesda, Maryland, pp 82–89

Amend DF, Pietsch JP (1972) Virucidal activity of two iodophors to salmonid viruses. J Fish Res Board Can 29:61–65

Ames BN, Elson-Schwab I, Silver EA (2002) High-dose vitamin therapy stimulates variant enzymes with decreased coenzyme binding affinity (increased K_m): relevance to genetic disease and polymorphisms. Am J Clin Nutr 75:616–58

Balk L, Hägerroth P-Å, Åkerman G, et al (2009) Wild birds of declining European species are dying from a thiamine deficiency syndrome. Proc Natl Acad Sci 106:12001–12006

Balk L, Hägerroth P-Å, Gustavsson H, et al (2016) Widespread episodic thiamine deficiency in Northern Hemisphere wildlife. Sci Rep 6:38821

Bengtsson B-E, Hill C, Bergman Å, et al (1999) Reproductive disturbances in Baltic fish: a synopsis of the FiRe project. Ambio 28:2–8

Berendsen RL, Pieterse CMJ, Bakker PAHM (2012) The rhizosphere microbiome and plant health. Trends Plant Sci 17:478–486

Bettendorff L (2013) Thiamine. In: Handbook of Vitamins, Fifth Edition. CRC Press, pp 267–324

Bettendorff L, Hennuy B, Wins P, Schoffeniels E (1993) Thiamin and derivatives as modulators of rat brain chloride channels. Neuroscience 52:1009–1017

Bettendorff L, Lakaye B, Kohn G, Wins P (2014) Thiamine triphosphate: a ubiquitous molecule in search of a physiological role. Metab Brain Dis 29:1069–1082

Blakeslee CJ, Sweet SA, Galbraith HS, Honeyfield DC (2015) Thiaminase activity in native freshwater mussels. J Gt Lakes Res 41:516–519

Boś M, Kozik A (2000) Some molecular and enzymatic properties of a homogeneous preparation of thiaminase I purified from carp liver. J Protein Chem 19:75–84

Breves G, Brandt M, Hoeller H, Rohr K (1981) Flow of thiamin to the duodenum in dairy cows fed different rations. J Agric Sci 96:587–591

Breves G, Hoeller H, Harmeyer J, Martens H (1980) Thiamin balance in the gastrointestinal tract of sheep. J Anim Sci 51:1177–1181

Brown SB, Arts MT, Brown LR, et al (2005a) Can diet-dependent factors help explain fish-to-fish variation in thiamine-dependent early mortality syndrome? J Aquat Anim Health 17:36–47

Brown SB, Fitzsimons JD, Honeyfield DC, Tillitt DE (2005b) Implications of thiamine deficiency in Great Lakes salmonines. J Aquat Anim Health 17:113–124

Brown SB, Honeyfield DC, Hnath JG, et al (2005c) Thiamine status in adult salmonines in the Great Lakes. J Aquat Anim Health 17:59–64

Butterworth RF, Héroux M (1989) Effect of pyrithiamine treatment and subsequent thiamine rehabilitation on regional cerebral amino acids and thiamine-dependent enzymes. J Neurochem 52:1079–1084

Bylund G, Lerche O (1995) Thiamine therapy of M74 affected fry of Atlantic salmon *Salmo salar*. Bull Eur Assoc Fish Pathol 15:93–97

Carlson SM, Cunningham CJ, Westley PAH (2014) Evolutionary rescue in a changing world. Trends Ecol Evol 29:521–530

Carmody RN, Gerber GK, Luevano JM, et al (2015) Diet dominates host genotype in shaping the murine gut microbiota. Cell Host Microbe 17:72–84

Carvalho PSM, Tillitt DE, Zajicek JL, et al (2009) Thiamine deficiency effects on the vision and foraging ability of lake trout fry. J Aquat Anim Health 21:315–325

Casteels M, Sniekers M, Fraccascia P, et al (2007) The role of 2-hydroxyacyl-CoA lyase, a thiamin pyrophosphate-dependent enzyme, in the peroxisomal metabolism of 3-methyl-branched fatty acids and 2-hydroxy straight-chain fatty acids. Biochem Soc Trans 35:876–880

Christie MR, Ford MJ, Blouin MS (2014) On the reproductive success of early-generation hatchery fish in the wild. Evol Appl 7:883–896

Chuang DT, Ku LS, Cox RP (1982) Thiamin-responsive maple-syrup-urine disease: decreased affinity of the mutant branched-chain alpha-keto acid dehydrogenase for alpha-ketoisovalerate and thiamin pyrophosphate. Proc Natl Acad Sci 79:3300–3304

Cooper JR, Pincus JH (1979) The role of thiamine in nervous tissue. Neurochem Res 4:223–239

Czesny S, Dettmers JM, Rinchard J, Dabrowski K (2009) Linking egg thiamine and fatty acid concentrations of Lake Michigan lake trout with early life stage mortality. J Aquat Anim Health 21:262–271

Depeint F, Bruce WR, Shangari N, et al (2006) Mitochondrial function and toxicity: role of the B vitamin family on mitochondrial energy metabolism. Chem Biol Interact 163:94–112

Dettmers JM, Goddard CI, Smith KD (2012) Management of alewife using Pacific salmon in the Great Lakes: whether to manage for economics or the ecosystem? Fisheries 37:495–501

Faigle R, Mohme M, Levy M (2012) Dry beriberi mimicking Guillain-Barre syndrome as the first presenting sign of thiamine deficiency. Eur J Neurol 19:e14–e15

Fisher CW, Lau KS, Fisher CR, et al (1991) A 17-bp insertion and a Phe215→Cys missense mutation in the dihydrolipoyl transacylase (E2) mRNA from a thiamine-responsive maple syrup urine disease patient WG-34. Biochem Biophys Res Commun 174:804–809

Fisher JP, Brown SB, Wooster GW, Bowser PR (1998) Maternal blood, egg and larval thiamin levels correlate with larval survival in landlocked Atlantic salmon (*Salmo salar*). J Nutr 128:2456–2466

Fisher JP, Fitzsimons JD, Combs GF, Spitsbergen JM (1996) Naturally occurring thiamine deficiency causing reproductive failure in Finger Lakes Atlantic salmon and Great Lakes trout. Trans Am Fish Soc 125:167–178

Fisher JP, Spitsbergen JM, Iamonte T, et al (1995) Pathological and behavioral manifestations of the "Cayuga syndrome," a thiamine deficiency in larval landlocked Atlantic salmon. J Aquat Anim Health 7:269–283

Fitzsimons J, Brown SB (1998) Reduced egg thiamine levels in inland and Great Lakes lake trout and their relationship with diet. In: McDonald G, Fitzsimons JD, Honeyfield (eds) Early life stage mortality syndrome in fishes of the Great Lakes and Baltic Sea. American Fisheries Society, Bethesda, Maryland, pp 160–171

Fitzsimons JD (1995) The effect of B-vitamins on a swim-up syndrome in Lake Ontario lake trout. J Gt Lakes Res 21:286–289

Fitzsimons JD, Brown S, Brown L, et al (2010) Increase in lake trout reproduction in Lake Huron following the collapse of alewife: relief from thiamine deficiency or larval predation? Aquat Ecosyst Health Manag 13:73–84

Fitzsimons JD, Brown SB, Brown LR, et al (2013) Impacts of diet on thiamine status of Lake Ontario American eels. Trans Am Fish Soc 142:1358–1369

Fitzsimons JD, Brown SB, Williston B, et al (2009) Influence of thiamine deficiency on lake trout larval growth, foraging, and predator avoidance. J Aquat Anim Health 21:302–314

Fitzsimons JD, Huestis S, Williston B (1995) Occurrence of a swim-up syndrome in Lake Ontario lake trout in relation to contaminants and cultural practices. J Gt Lakes Res 21:277–285

Fitzsimons JD, Ketola G, Wooster GW, Brown SB (2001a) Use of a thiamine antagonist to induce Cayuga-Syndrome-like mortalities in larval Atlantic salmon. J Aquat Anim Health 13:151–157

Fitzsimons JD, Vandenbyllaardt L, Brown SB (2001b) The use of thiamine and thiamine antagonists to investigate the etiology of early mortality syndrome in lake trout (*Salvelinus namaycush*). Aquat Toxicol 52:229–239

Fitzsimons JD, Williston B, Amcoff P, et al (2005a) The effect of thiamine injection on upstream migration, survival, and thiamine status of putative thiamine-deficient coho salmon. J Aquat Anim Health 17:48–58

Fitzsimons JD, Williston B, Williston G, et al (2007) Egg thiamine status of Lake Ontario salmonines 1995-2004 with emphasis on lake trout. J Gt Lakes Res 33:93–103

Fitzsimons JD, Williston B, Zajicek JL, et al (2005b) Thiamine content and thiaminase activity of ten freshwater stocks and one marine stock of alewives. J Aquat Anim Health 17:26–35

Fitzsimons JD, Wolgamood M, Madenjian CP, Bunnell DB (2012) Thiamine deficiency in aquatic food chains: the cumulative result of ecosystem disruption by clupeids? In: Norrgren L, Levengood J (eds) Ecology and Animal Health. Baltic University Press, Uppsala, pp 167–180

Fox GA, Allan LJ, Weseloh DV, Mineau P (1990) The diet of herring gulls during the nesting period in Canadian waters of the Great Lakes. Can J Zool 68:1075–1085

Fradin D, Bougneres P (2007) Three common intronic variants in the maternal and fetal thiamine pyrophosphokinase gene (TPK1) are associated with birth weight. Ann Hum Genet 71:578–585

Futia MH, Hallenbeck S, Noyes AD, et al (2017) Thiamine deficiency and the effectiveness of thiamine treatments through broodstock injections and egg immersion on Lake Ontario steelhead trout. J Gt Lakes Res 43:352–358

Ghiasi S, Falahatkar B, Arslan M, Dabrowski K (2017) Physiological changes and reproductive performance of Sterlet sturgeon *Acipenser ruthenus* injected with thiamine. Anim Reprod Sci 178:23–30

Gibson GE, Ksiezak-Reding H, Sheu K-FR, et al (1984) Correlation of enzymatic, metabolic, and behavioral deficits in thiamin deficiency and its reversal. Neurochem Res 9:803–814

Gibson GE, Zhang H (2002) Interactions of oxidative stress with thiamine homeostasis promote neurodegeneration. Neurochem Int 40:493–504

Goodrich JK, Davenport ER, Waters JL, et al (2016) Cross-species comparisons of host genetic associations with the microbiome. Science 352:532–535

Goodrich JK, Waters JL, Poole AC, et al (2014) Human genetics shape the gut microbiome. Cell 159:789–799

Green RG, Evans CA (1940) A deficiency disease of foxes. Science 92:154–155

Hansson S (2001) Stomach analyses of Baltic salmon from 1959–1962 and 1994–1997: possible relations between diet and yolk-sac-fry mortality (M74). J Fish Biol 58:1730–1745

Happel A, Jonas JL, McKenna PR, et al (2018) Spatial variability of lake trout diets in Lakes Huron and Michigan revealed by stomach content and fatty acid profiles. Can J Fish Aquat Sci 75:95–105

Happel A, Pattridge R, Walsh M, Rinchard J (2017) Assessing diet compositions of Lake Ontario predators using fatty acid profiles of prey fishes. J Gt Lakes Res 43:838–845

Harbicht AB, Castro-Santos T, Gorsky D, et al (2018) Environmental, anthropogenic, and dietary influences on fine scale movement patterns of Atlantic salmon through challenging waters. Can J Fish Aquat Sci 75:2198-2210

Harper CG, Giles M, Finlay-Jones R (1986) Clinical signs in the Wernicke-Korsakoff complex: a retrospective analysis of 131 cases diagnosed at necropsy. J Neurol Neurosurg Psychiatry 49:341–345

Hashimoto Y, Arai S, Nose T (1970) Thiamine deficiency symptoms experimentally induced in the eel. Bull Jpn Soc Sci Fish 36:791–797

He JX, Ebener MP, Riley SC, et al (2012) Lake trout status in the main basin of Lake Huron, 1973–2010. North Am J Fish Manag 32:402–412

Honeyfield D, Murphy J, Howard K, et al (2016) An exploratory assessment of thiamine status in western Alaska Chinook salmon (*Oncorhynchus tshawytscha*). North Pac Anadromous Fish Comm Bull 6:21–31

Honeyfield D, Peters A, Jones M (2008) Thiamine and fatty acid content of Lake Michigan Chinook salmon. J Gt Lakes Res 34:581–589

Honeyfield DC, Daniels ME, Brown LR, et al (2012) Survey of four essential nutrients and thiaminase activity in five Lake Ontario prey fish species. J Gt Lakes Res 38:11–17

Honeyfield DC, Hinterkopf JP, Brown SB (2002) Isolation of thiaminase-positive bacteria from alewife. Trans Am Fish Soc 131:171–175

Honeyfield DC, Hinterkopf JP, Fitzsimons JD, et al (2005) Development of thiamine deficiencies and early mortality syndrome in lake trout by feeding experimental and feral fish diets containing thiaminase. J Aquat Anim Health 17:4–12

Honeyfield DC, Vandergoot CS, Bettoli PW, et al (2007) Thiamine and fatty acid content of walleye tissue from three southern U.S. reservoirs. J Aquat Anim Health 19:84–93

Hornung MW, Miller L, Peterson RE, et al (1998) Efficacy of thiamine, astaxanthin, β-carotene, and thyroxine treatments in reducing early mortality syndrome in Lake Michigan salmonid embryos. In: McDonald G, Fitzsimons JD, Honeyfield DC (eds) Early life stage mortality syndrome in fishes of the Great Lakes and Baltic Sea. American Fisheries Society, Bethesda, Maryland, pp 124–134

Houde ALS, Saez PJ, Wilson CC, et al (2015) Effects of feeding high dietary thiaminase to sub-adult Atlantic salmon from three populations. J Gt Lakes Res 41:898–906

Jacobs GR, Madenjian CP, Bunnell DB, et al (2013) Chinook salmon foraging patterns in a changing Lake Michigan. Trans Am Fish Soc 142:362–372

Jenkins AH, Schyns G, Potot S, et al (2007) A new thiamin salvage pathway. Nat Chem Biol 3:492–497

Ji YQ, Adelman IR (1998) Thiaminase activity in alewives and smelt in Lakes Huron, Michigan, and Superior. In: McDonald G, Fitzsimons JD, Honeyfield (eds) Early life stage mortality syndrome in fishes of the Great Lakes and Baltic Sea. American Fisheries Society, Bethesda, Maryland, pp 154–159

Ji YQ, Warthesen JJ, Adelman IR (1998) Thiamine nutrition, synthesis, and retention in relation to lake trout reproduction in the Great Lakes. In: McDonald G, Fitzsimons JD, Honeyfield (eds) Early life stage mortality syndrome in fishes of the Great Lakes and Baltic Sea. American Fisheries Society, Bethesda, Maryland, pp 99–111

Jude DJ, Tesar FJ, Deboe SF, Miller TJ (1987) Diet and selection of major prey species by Lake Michigan salmonines, 1973-1982. Trans Am Fish Soc 116:677–691

Jurgenson CT, Begley TP, Ealick SE (2009) The structural and biochemical foundations of thiamin biosynthesis. Annu Rev Biochem 78:569–603

Karlsson L, Ikonen E, Mitans A, Hansson S (1999) The diet of salmon (*Salmo salar*) in the Baltic Sea and connections with the M74 syndrome. Ambio 28:37–42

Keinänen M, Uddström A, Mikkonen J, et al (2012) The thiamine deficiency syndrome M74, a reproductive disorder of Atlantic salmon (*Salmo salar*) feeding in the Baltic Sea, is related to the fat and thiamine content of prey fish. ICES J Mar Sci 69:516–528

Ketola HG, Bowser PR, Wooster GA, et al (2000) Effects of thiamine on reproduction of Atlantic salmon and a new hypothesis for their extirpation in Lake Ontario. Trans Am Fish Soc 129:607–612

Ketola HG, Chiotti TL, Rathman RS, et al (2005) Thiamine status of Cayuga Lake rainbow trout and its influence on spawning migration. North Am J Fish Manag 25:1281–1287

King-Heiden TC, Mehta V, Xiong KM, et al (2012) Reproductive and developmental toxicity of dioxin in fish. Mol Cell Endocrinol 354:121–138

Kitamura T, Seki N, Kihara A (2017) Phytosphingosine degradation pathway includes fatty acid α-oxidation reactions in the endoplasmic reticulum. Proc Natl Acad Sci 114:E2616–E2623

Koski P, Pakarinen M, Nakari T, et al (1999) Treatment with thiamine hydrochloride and astaxanthine for the prevention of yolk-sac mortality in Baltic salmon fry (M74 syndrome). Dis Aquat Organ 37:209–220

Kraft CE, Angert ER (2017) Competition for vitamin B_1 (thiamin) structures numerous ecological interactions. Q Rev Biol 92:151–168

Kraft CE, Gordon ER, Angert ER (2014) A rapid method for assaying thiaminase I activity in diverse biological samples. PloS One 9:e92688

Kuno Y (1951) *Bacillus thiaminolyticus*, a new thiamin-decomposing bacterium. Proc Jpn Acad 27:362–365

Ladago BJ, Marsden JE, Evans AN (2016) Early feeding by lake trout fry. Trans Am Fish Soc 145:1–6

Lee B-J, Jaroszewska M, Dabrowski K, et al (2009) Effects of vitamin B_1 (thiamine) deficiency in lake trout alevins and preventive treatments. J Aquat Anim Health 21:290–301

Lepak JM, Kraft CE, Honeyfield DC, Brown SB (2008) Evaluating the effect of stressors on thiaminase activity in alewife. J Aquat Anim Health 20:63–71

Lim C, Yildirim-Aksoy M, Barros MM, Klesius P (2011) Thiamin requirement of Nile tilapia, *Oreochromis niloticus*. J World Aquac Soc 42:824–833

Lukienko PI, Mel'nichenko NG, Zverinskii IV, Zabrodskaya SV (2000) Antioxidant properties of thiamine. Bull Exp Biol Med 130:874–876

Lundström J, Börjeson H, Norrgren L (1999a) Histopathological studies of yolk-sac fry of Baltic salmon (*Salmo salar*) with the M74 syndrome. Ambio 28:16–23

Lundström J, Carney B, Amcoff P, et al (1999b) Antioxidative systems, detoxifying enzymes and thiamine levels in Baltic salmon (*Salmo salar*) that develop M74. Ambio 28:24–29

Mac MJ, Edsall CC, Seelye JG (1985) Survival of lake trout eggs and fry reared in water from the upper Great Lakes. J Gt Lakes Res 11:520–529

Mac MJ, Schwartz TR, Edsall CC, Frank AM (1993) Polychlorinated biphenyls in Great Lakes lake trout and their eggs: relations to survival and congener composition 1979-1988. J Gt Lakes Res 19:752–765

MacCallum WR, Regier HA (1970) Distribution of smelt, *Osmerus mordax*, and the smelt fishery in Lake Erie in the early 1960's. J Fish Board Can 27:1823–1846

Madenjian CP, Fahnenstiel GL, Johengen TH, et al (2002) Dynamics of the Lake Michigan food web, 1970-2000. Can J Fish Aquat Sci 59:736–753

Marcquenski S, Brown S (1997) Early mortality syndrome (EMS) in salmonid fishes from the Great Lakes. In: Rolland RM, Gilbertson M, Peterson RE (eds) Chemically induced alterations in functional development and reproduction of fishes. SETAC Press, Pensacola, FL, pp 135–152

Marsden JE, Hauser M (2009) Exotic species in Lake Champlain. J Gt Lakes Res 35:250–265

Marsden JE, Kozel CL, Chipman BD (2018) Recruitment of lake trout in Lake Champlain. J Gt Lakes Res 44:166–173

Miller RR (1957) Origin and dispersal of the alewife, *Alosa pseudoharengus*, and the gizzard shad, *Dorosoma cepedianum*, in the Great Lakes. Trans Am Fish Soc 86:97–111

Miller TJ, Crowder LB, Rice JA, Marschall EA (1988) Larval size and recruitment mechanisms in fishes: toward a conceptual framework. Can J Fish Aquat Sci 45:1657–1670

Mimouni-Bloch A, Goldberg-Stern H, Strausberg R, et al (2014) Thiamine deficiency in infancy: long-term follow-up. Pediatr Neurol 51:311–316

Monteverde DR, Gómez-Consarnau L, Cutter L, et al (2015) Vitamin B_1 in marine sediments: pore water concentration gradient drives benthic flux with potential biological implications. Front Microbiol 6:434

Nabokina SM, Said HM (2012) A high-affinity and specific carrier-mediated mechanism for uptake of thiamine pyrophosphate by human colonic epithelial cells. AJP Gastrointest Liver Physiol 303:G389–G395

Nishimune T, Watanabe Y, Okazaki H, Akai H (2000) Thiamin is decomposed due to *Anaphe* spp. entomophagy in seasonal ataxia patients in Nigeria. J Nutr 130:1625–1628

Nott A, Meislin SH, Moore MJ (2003) A quantitative analysis of intron effects on mammalian gene expression. RNA 9:607–617

Nylund A, Hovland T, Hodneland K, et al (1994) Mechanisms for transmission of infectious salmon anaemia (ISA). Dis Aquat Organ 19:95–100

Ottinger CA, Honeyfield DC, Densmore CL, Iwanowicz LR (2012) Impact of thiamine deficiency on T-cell dependent and T-cell independent antibody production in lake trout. J Aquat Anim Health 24:258–273

Palace VP, Brown SB, Baron CL, et al (1998) An evaluation of the relationships among oxidative stress, antioxidant vitamins and early mortality syndrome (EMS) of lake trout (*Salvelinus namaycush*) from Lake Ontario. Aquat Toxicol 43:195–208

Pankhurst NW, Sorensen PW (1984) Degeneration of the alimentary tract in sexually maturing European *Anguilla anguilla* (L.) and American eels *Anguilla rostrata* (LeSueur). Can J Zool 62:1143–1149

Park LCH, Zhang H, Sheu K-FR, et al (1999) Metabolic impairment induces oxidative stress, compromises inflammatory responses, and inactivates a key mitochondrial enzyme in microglia. J Neurochem 72:1948–1958

Pettersson A, Lignell Å (1999) Astaxanthin deficiency in eggs and fry of Baltic salmon (*Salmo salar*) with the M74 syndrome. Ambio 28:43–47

Quinn TP (2005) The behavior and ecology of Pacific salmon and trout, 1st edn. University of Washington Press, Seattle, WA

Ray BA, Hrabik TR, Ebener MP, et al (2007) Diet and prey selection by Lake Superior lake trout during spring, 1986–2001. J Gt Lakes Res 33:104–113

Richter CA, Evans AN, Wright-Osment MK, et al (2012) *Paenibacillus thiaminolyticus* is not the cause of thiamine deficiency impeding lake trout (*Salvelinus namaycush*) recruitment in the Great Lakes. Can J Fish Aquat Sci 69:1056–1064

Riley SC, Evans AN (2008) Phylogenetic and ecological characteristics associated with thiaminase activity in Laurentian Great Lakes fishes. Trans Am Fish Soc 137:147–157

Riley SC, He JX, Johnson JE, et al (2007) Evidence of widespread natural reproduction by lake trout *Salvelinus namaycush* in the Michigan waters of Lake Huron. J Gt Lakes Res 33:917–921

Riley SC, Rinchard J, Honeyfield DC, et al (2011) Increasing thiamine concentrations in lake trout eggs from Lakes Huron and Michigan coincide with low alewife abundance. North Am J Fish Manag 31:1052–1064

Roseman EF, Schaeffer JS, Bright E, Fielder DG (2014) Angler-caught piscivore diets reflect fish community changes in Lake Huron. Trans Am Fish Soc 143:1419–1433

Said HM, Ortiz A, Subramanian VS, et al (2001) Mechanism of thiamine uptake by human colonocytes: studies with cultured colonic epithelial cell line NCM460. AJP Gastrointest Liver Physiol 281:G144–G150

Sannino DR, Dobson AJ, Edwards K, et al (2018) The *Drosophila melanogaster* gut microbiota provisions thiamine to its host. mBio 9:e00155-18

Sechi G, Serra A (2007) Wernicke's encephalopathy: new clinical settings and recent advances in diagnosis and management. Lancet Neurol 6:442–455

Skea JC, Symula J, Miccoli J (1985) Separating starvation losses from other early feeding fry mortality in steelhead trout *Salmo gairdneri*, Chinook salmon *Oncorhynchus tshawytscha*, and lake trout *Salvelinus namaycush*. Bull Environ Contam Toxicol 35:82–91

Stewart DJ, Kitchell JF, Crowder LB (1981) Forage fishes and their salmonid predators in Lake Michigan. Trans Am Fish Soc 110:751–763

Sutherland WJ, Butchart SHM, Connor B, et al (2018) A 2018 horizon scan of emerging issues for global conservation and biological diversity. Trends Ecol Evol 33:47–58

Swedberg DV, Peck JW (1984) Food of young-of-the-year lake trout (*Salvelinus namaycush*) in Presque Isle Harbor, Lake Superior. J Gt Lakes Res 10:280–285

Sylvander P, Häubner N, Snoeijs P (2013) The thiamine content of phytoplankton cells is affected by abiotic stress and growth rate. Microb Ecol 65:566–577

Tillitt DE, Cook PS, Giesy JP, et al (2008) Reproductive impairment of Great Lakes lake trout by dioxin-like chemicals. In: The toxicology of fishes. CRC Press, Boca Raton, FL, pp 819–876

Tillitt DE, Riley SC, Evans AN, et al (2009) Dreissenid mussels from the Great Lakes contain elevated thiaminase activity. J Gt Lakes Res 35:309–312

Tillitt DE, Zajicek JL, Brown SB, et al (2005) Thiamine and thiaminase status in forage fish of salmonines from Lake Michigan. J Aquat Anim Health 17:13–25

Timpson NJ, Forouhi NG, Brion M-J, et al (2010) Genetic variation at the SLC23A1 locus is associated with circulating concentrations of L-ascorbic acid (vitamin C): evidence from 5 independent studies with >15,000 participants. Am J Clin Nutr 92:375–382

U.S. Department of Agriculture, Agricultural Research Service. 2018. USDA National Nutrient Database for Standar Reference, Legacy. Version Current: April 2018. Nutrient Data Laboratory Home Page, http://www.ars.usda.gov/nutrientdata

Vogel C, Pleiss J (2014) The modular structure of ThDP-dependent enzymes: the modular structure of ThDP-dependent enzymes. Proteins Struct Funct Bioinforma 82:2523–2537

Waples RS (1990) Conservation genetics of Pacific salmon. II. Effective population size and the rate of loss of genetic variability. J Hered 81:267–276

Werner RM, Rook B, Greil R (2006) Egg-thiamine status and occurrence of early mortality syndrome (EMS) in Atlantic salmon from the St. Marys River, Michigan. J Gt Lakes Res 32:293–305

Weseloh DV, Ewins PJ, Struger J, et al (1995) Double-crested cormorants of the Great Lakes: Changes in population size, breeding distribution and reproductive output between 1913 and 1991. Colon Waterbirds 18:48

Widmann M, Radloff R, Pleiss J (2010) The Thiamine diphosphate dependent Enzyme Engineering Database: A tool for the systematic analysis of sequence and structure relations. BMC Biochem 11:9

Wistbacka S, Bylund G (2008) Thiaminase activity of Baltic salmon prey species: a comparision of net- and predator-caught samples. J Fish Biol 72:787–802

Wistbacka S, Heinonen A, Bylund G (2002) Thiaminase activity of gastrointestinal contents of salmon and herring from the Baltic Sea. J Fish Biol 60:1031–1042

Wolf LE (1942) Fish-diet disease of trout: a vitamin deficiency produced by diets containing raw fish. New York State Conservation Department, Bureau of Fish Culture

Wolgamood M, Hnath JG, Brown SB, et al (2005) Temporal and spatial variation of early mortality syndrome in salmonids from Lakes Michigan and Huron. J Aquat Anim Health 17:65–76

Wooster GA, Bowser PR, Brown SB, Fisher JP (2000) Remediation of Cayuga Syndrome in landlocked Atlantic salmon Salmo salar using egg and sac-fry bath treatments of thiamine-hydrochloride. J World Aquac Soc 31:149–157

Zajicek JL, Tillitt DE, Honeyfield DC, et al (2005) A method for measuring total thiaminase activity in fish tissues. J Aquat Anim Health 17:82–94

CHAPTER 2. AMONG-FAMILY VARIATION IN SURVIVAL AND GENE EXPRESSION UNCOVERS ADAPTIVE GENETIC VARIATION IN A THREATENED FISH

This work has been published previously by John Wiley and Sons: Harder et al. 2020. Among-family variation in survival and gene expression uncovers adaptive genetic variation in a threatened fish. *Molecular Ecology* 29:1035-1049.

2.1 Abstract

Variation in among-family transcriptional responses to different environmental conditions can help to identify adaptive genetic variation, even prior to a selective event. Coupling differential gene expression with formal survival analyses allows for the disentanglement of treatment effects, required for understanding how individuals plastically respond to environmental stressors, from the adaptive genetic variation responsible for differential survival. We combined these two approaches to investigate responses to an emerging conservation issue, thiamine (vitamin B_1) deficiency, in a threatened population of Atlantic salmon (*Salmo salar).* Thiamine is an essential vitamin that is increasingly limited in many ecosystems. In Lake Champlain, Atlantic salmon cannot acquire thiamine in sufficient quantities to support natural reproduction; fertilized eggs must be reared in hatcheries and treated with supplemental thiamine. We evaluated transcriptional responses (RNA-seq) to thiamine treatment across families and found 3,616 genes differentially expressed between control (no supplemental thiamine) and treatment individuals. Fewer genes changed expression equally across families (i.e., additively) than exhibited genotype x environment interactions in response to thiamine. Differentially expressed genes were related to known physiological effects of thiamine deficiency, including oxidative stress, cardiovascular irregularities, and neurological abnormalities. We also identified 1,446 putatively adaptive genes that were strongly associated with among-family survival in the absence of thiamine treatment, many of which related to neurogenesis and visual perception. Our results highlight the utility of coupling RNA-seq with formal survival analyses to identify candidate genes that underlie the among-family variation in survival required for an adaptive response to natural selection.

2.2 Introduction

Understanding if and how species can adapt to rapidly changing environmental conditions is a primary goal of modern conservation biology (Bernatchez, 2016; Stockwell, Hendry, & Kinnison, 2003). One of the key challenges in meeting this goal is uncovering the adaptive genetic variation required for a response to selection and deciphering whether this adaptive genetic variation will be sufficient to respond to anthropogenically-induced agents of selection. Contemporary genomic approaches have revolutionized our ability to identify regions of the genome responding to selection, even over relatively short time periods (Franks, Kane, O'Hara, Tittes, & Rest, 2016; van't Hof et al., 2016; Willoughby, Harder, Tennessen, Scribner, & Christie, 2018). However, such methods often lack sufficient power to detect rapid responses to selection, especially when examining polygenic traits shaped by large numbers of loci of small effect (Pritchard, Pickrell, & Coop, 2010; Wellenreuther & Hansson, 2016). Furthermore, genomic approaches can only provide insights after selection has already occurred; thus their utility for predicting responses to selection requires appropriate study systems or long-term experimental breeding designs. One alternative to these approaches is experimental transcriptomics. By carefully designing treatments, rearing F_1 offspring in a common environment, and deeply sequencing mRNA, it is possible to uncover an adaptive, genetic response to selection (Christie, Marine, Fox, French, & Blouin, 2016; Passow et al., 2017; Uusi-Heikkilä, Sävilammi, Leder, Arlinghaus, & Primmer, 2017). Coupling family-level replication and formal survival analyses allows for the disentanglement of treatment effects, required for understanding how individuals plastically respond to environmental stressors, and among-family variation in survival and gene expression. Here, we apply these techniques to a threatened fish population, whose successful reintroduction will require an adaptive response to an emerging conservation issue, thiamine deficiency.

Thiamine (vitamin B_1) is an essential vitamin that is synthesized by prokaryotes, plants, and fungi; animals are incapable of producing thiamine and primarily acquire the vitamin through their diets (Bettendorff, 2013). The physiological manifestations of thiamine deficiency are directly related to thiamine's roles in bioenergetic, neurological, and cardiovascular pathways. Thiamine primarily serves as a cofactor for enzymes in metabolism and energy production pathways (i.e., pentose phosphate pathway and tricarboxylic acid cycle) (Brown et al., 2005; Fitzsimons, Brown, Honeyfield, & Hnath, 1999). Thiamine is also required for the production of neurotransmitters,

antioxidants, and myelin (Bettendorff, 2013), consistent with the neurological and behavioral signs of thiamine deficiency, including brain lesions (Butterworth, 2009; Honeyfield et al., 2008; Lee, Jaroszewska, Dabrowski, Czesny, & Rinchard, 2009) and uncoordinated movements (Brown et al., 2005; Fisher, Spitzbergen, et al., 1995; Fitzsimons et al., 2005; Sechi & Serra, 2007). Thiamine deficiency can also impair cardiovascular function, leading to low blood pressure, irregular heart rate, pulmonary edema, and circulatory collapse (Essa et al., 2011; Sechi & Serra, 2007). Because thiamine plays a central role in growth, development, and proper neurological function (Bettendorff, 2013), thiamine deficiency can impair an individual's capacity to forage, avoid predation, and reproduce (Carvalho et al., 2009; Fisher, Spitzbergen, et al., 1995; Fitzsimons et al., 2009), all of which can contribute to large reductions in survival and reproductive success (Ketola, Bowser, Wooster, Wedge, & Hurst, 2000; Mörner et al., 2017).

Evidence is mounting that populations of diverse taxa are becoming increasingly deficient in thiamine (Balk et al., 2009, 2016). For example, high rates of mortality or reduced reproductive success associated with thiamine deficiency have been observed in invertebrates (Balk et al., 2016), fishes (Futia et al., 2017), reptiles (Honeyfield et al., 2008; Ross et al., 2009), and birds (Balk et al., 2009), and many cases of thiamine deficiency may remain undetected. The underlying causes of thiamine deficiency vary among taxa and environments. In fishes, the emergence of thiamine deficiency is largely attributed to diet. For example, thiamine deficiency has often been observed in salmonids with diets containing alewife (*Alosa pseudoharengus*) and rainbow smelt (*Osmerus mordax*), both of which contain high levels of thiaminase, a thiamine-degrading enzyme (reviewed in Harder et al., 2018). In the Baltic Sea, the occurrence of thiamine deficiency in Atlantic salmon (*Salmo salar*) also coincides with the consumption of fishes with low thiamine to fat content ratios, including Atlantic herring (*Clupea harengus*) and sprat (*Sprattus sprattus*) (Hansson et al., 2001; Keinänen et al., 2012). However, these fishes also contain thiaminase, making it difficult to establish low thiamine to fat content ratios as direct, causative agents of thiamine deficiency. For adult salmon returning to spawn, the most obvious signs of thiamine deficiency are uncoordinated, "wiggling" swimming patterns and an inability to remain upright in the water column (Fisher, Spitsbergen, Iamonte, Little, & Delonay, 1995; Fitzsimons et al., 2005). Individuals hatching from thiamine deficient eggs do not survive for more than a few weeks and exhibit physical signs of deficiency, such as hemorrhaging and large yolk sacs with opacities and edema, prior to death (Fig. 2.1A; Fisher, Spitsbergen, et al., 1995). The inability of thiamine deficient salmon to successfully

reproduce is an emerging conservation and management issue (reviewed in Harder et al., 2018), and impedes reintroduction efforts throughout their native range.

One such reintroduction effort occurs in Lake Champlain (Canada and USA), where Atlantic salmon were extirpated from the lake in the early 1800s (Marsden & Langdon, 2012; Appendix A Introduction). Diversifying the forage base or controlling the non-native alewife population in Lake Champlain could alleviate thiamine deficiency in Atlantic salmon, but efforts to eradicate invasive species after population establishment are often prohibitively expensive and the possibility of reinvasion cannot be eliminated (Myers, Simberloff, Kuris, & Carey, 2000). Alternatively, recent research suggests that Atlantic salmon populations with diets high in thiaminase may have genetically adapted to low thiamine availability (Houde, Saez, Wilson, Bureau, & Neff, 2015). This rapid genetic adaptation could be the result of selection on genes associated with thiamine-dependent pathways. For example, conformational changes in enzymes requiring thiamine as a cofactor could increase the binding affinity for thiamine or, alternatively, variation in regulatory sequences could modify the expression of genes involved in thiamine uptake and intracellular transport. However, the application of supplemental thiamine to all fertilized eggs reared in the Lake Champlain hatchery relaxes all selection related to thiamine deficiency at early life stages, and it is currently unknown whether genetic variation in this population could support a response to such selection. By coupling thiamine treatments, RNA-seq, and survival analyses on F_1 offspring from 9 families raised in a common environment, we identified an among-family adaptive response in a thiamine-deficient population of Atlantic salmon and identified pathways and functions impacted by thiamine deficiency.

2.3 Methods

2.3.1 Study system and experimental crosses

We collected gametes from 35 pairs of adult male and female Atlantic salmon returning to the Ed Weed Fish Culture Station (Grand Isle, Vermont, USA) across two spawning seasons: 17 pairs in November 2016 and 18 pairs in November 2017. We immediately froze approximately 50 eggs from each female on dry ice for total thiamine concentration analysis whereby two, 1-g biological replicates of unfertilized egg tissue were analyzed via high performance liquid chromatography (*sensu stricto* Futia et al. 2017). We also performed total thiamine concentration

analyses on 2-g samples of maternal muscle tissue sampled from each female (two samples per female) during U.S. Fish and Wildlife Service disease testing procedures. We transported the remaining gametes at 4 °C to the White River National Fish Hatchery (Bethel, Vermont, USA), where we systematically combined milt and eggs to generate 35 families (limited egg and milt availability precluded reciprocal crosses and in 2016, five families were half-sibling families; see Appendix A Methods for crossing details). We divided fertilized eggs from each family into two groups, placing one group into a 1% thiamine mononitrate solution (hereafter, "treated") and the other into a control water bath (hereafter, "untreated"). After 30 minutes, we rinsed all eggs with fresh water and transferred them to heath trays with one tray per family and treatment combination. We left the eggs undisturbed until they reached the eyed stage (when individuals exhibit retinal pigmentation, approximately 50 days post fertilization), at which point we counted and removed inviable eggs. After the eyed stage was reached, we recorded mortality and removed inviable eggs from all trays each week. Hatching occurred approximately 75 days post fertilization and we continued to monitor and remove dead individuals from all trays each week for the remainder of the experiments. We concluded the experiments after surviving fry had absorbed their yolk sacs (~130 days post fertilization) and just prior to initiation of exogenous feeding. For families spawned in 2016, rearing and collection procedures were approved by the Purdue Animal Care and Use Committee (PACUC #: 1612001514). Families spawned in 2017 were collected and sampled by U.S. Fish and Wildlife Service (USFWS) according to accepted guidelines for the use of fish in research (AFS, AIFRB, ASIH, 2014).

2.3.2 Sampling for RNA-seq

At 95 days post fertilization, we sampled a total of 36 individuals for gene expression analyses from 9 of the 18 families spawned in 2017. Due to hatchery broodstock quotas for treated individuals, we were limited to sampling these 9 families. To control for variation in development, we only sampled from families that were spawned on the same day. We froze individuals from each family and treatment group in dry cryogenic shipping dewars charged with liquid nitrogen and shipped them to Purdue University for storage at -80 °C. We subsequently placed two frozen individuals from each family and treatment group (n = 9 families * 2 treatments (+/- supplemental thiamine) * 2 individuals = 36) into 10 volumes of RNAlater-ICE (Invitrogen) pre-chilled to -80 °C and allowed the samples to reach -20 °C overnight. We then homogenized samples using a

TissueRuptor II (Qiagen) and extracted total RNA from each homogenate using an RNeasy kit (Qiagen).

2.3.3 Survival analysis

We generated a dose-response curve for egg thiamine concentration and proportion of untreated individuals surviving at the end of the experiments in each of the 35 families with the *drc* package (Ritz, Baty, Streibig, & Gerhard, 2015) in R version 3.5.3 (R Core Team, 2019). We selected the appropriate model by using the mselect function to calculate AIC values, with the two-parameter log-logistic function having the lowest AIC value. We next calculated effective concentrations of egg thiamine required for 25% and 50% survival (EC_{25} and EC_{50}, respectively) from the resulting logistic curve.

For the 9 families used in RNA-seq, we conducted survival analyses to determine whether treatment affected survival within a family over time and to determine the relative risks of death associated with belonging to each family according to survivorship of untreated individuals. We constructed Kaplan-Meier survival distributions for each family and treatment combination and used a log-rank test to determine whether treatment status significantly affected survival within each family (Kleinbaum & Klein, 2012). We then compared survival distributions for untreated individuals from each family against the survival distribution for untreated individuals from family A (the family with the highest survival rate of all families) and used Cox proportional hazards regressions to calculate hazard ratio values for all families (Cox, 1972) using the *survival* package (Therneau, 2015) in R. We censored individuals removed for RNA-seq in the analysis. For each family, the calculated hazard ratio represents the probability of mortality associated with belonging to that family, compared to family A. We also conducted a linear regression to test for a relationship between hazard ratio value and egg thiamine concentration. To meet assumptions of normality, we log-transformed hazard ratio values prior to all regression analyses.

Because USFWS gamete availability prevented us from creating reciprocal crosses, we could not formally test for maternal effects (*sensu* Christie et al., 2016). However, female size is often correlated with offspring size, and larger offspring frequently exhibit higher fitness than smaller offspring in a common environment (reviewed in Marshall, Heppell, Munch, & Warner, 2010). We therefore performed linear regressions to test for relationships between maternal physical characteristics (i.e., standard length and weight) and proportion of untreated offspring

surviving at the end of the experiment. We also plotted maternal muscle and egg thiamine concentration against proportion of untreated offspring surviving for the 9 families sampled for RNA-seq. A strong association between maternal characteristics and untreated offspring survival might indicate that maternal condition plays a larger role in determining thiamine deficiency outcomes than among-family genetic variation.

2.3.4 RNA-seq and sequence processing

We assessed total RNA concentration and quality on an Agilent BioAnalyzer at the Purdue Genomics Core Facility, with sample RIN scores ranging from 9.3-10.0. One library was prepared for each individual using the TruSeq Stranded mRNA protocol (Illumina) and cDNA was sequenced on an Illumina NovaSeq 6000 to generate an average of 87 million 150 bp paired-end reads per library. We removed adapter sequences and clipped poor quality bases (quality score < 20) from both ends of reads using *Trimmomatic* (Bolger, Lohse, & Usadel, 2014) and aligned reads to the annotated Atlantic salmon reference genome (*S. salar* ICSASG_v2 assembly, NCBI accession GCA_000233375.4; Lien et al., 2016) using *HISAT2* (Kim, Langmead, & Salzberg, 2015) with the --downstream-transcriptome-assembly option and reporting primary alignments. We next assembled transcripts for each sample using *StringTie* (Pertea et al., 2015) default parameters and the Atlantic salmon reference annotation file (ICSASG_v2) to guide assembly, and merged sample transcripts using *StringTie*. A transcript count matrix was next created with *featureCounts* (Liao, Smyth, & Shi, 2014), excluding chimeric fragments (-C option) and requiring that both reads in a pair be successfully mapped (-B option). By default, *featureCounts* does not count reads with multiple alignments (i.e., a single read aligned to multiple locations in the reference) or read pairs that overlap multiple features, and we retained these settings in our analyses.

2.3.5 Treatment effects (environmental and genetic effects)

We first made comparisons between treated and untreated individuals using both an *a priori* list of reference genes and a standard discovery-based gene identification pipeline. We generated a list of *a priori* genes predicted to be differentially expressed between treated and untreated samples using 4 criteria: (1) genes associated with thiamine-related biological process

gene ontology (GO) terms (any line containing "thiamine" in Ssal_ICSASG_v2_GOAccession.txt downloaded from SalmoBase (Samy et al., 2017) on June 28, 2018), (2) genes encoding thiamine diphosphate (TDP) dependent enzymes, (3) genes encoding enzymes that contain a TDP binding site (NCBI conserved protein domain family "TPP_enzymes"), and (4) genes included in the *S. salar* thiamine metabolism pathway in the NCBI BioSystems Database (BSID: 1429556).

We conducted differential gene expression analyses separately in *DESeq2* (Love, Huber, & Anders, 2014) in R for: (1) the *a priori* list of predicted differentially expressed genes (DEGs) and (2) the list of all assembled transcripts. We applied independent filtering to remove transcripts with low read counts, and identified DEGs associated with thiamine treatment status while controlling for the effects of family. We considered genes with an FDR-adjusted *p*-value (p_{adj}) < 0.05 to be differentially expressed and used the prcomp command in R to conduct a principal component analysis for DEGs identified from the list of all assembled transcripts.

Using the count matrix for all samples, we identified modules of co-expressed genes by calculating pair-wise Pearson correlations between each pair of genes to created signed networks using the *WGCNA* package (Langfelder & Horvath, 2008). We set the minimum module size to 30 genes and merged correlated modules ($r^2 > 0.9$). Each module comprised genes that showed similar expression patterns across samples within a treatment. Following the approach outlined in Langfelder and Horvath (2008) we performed the following steps. First, we summarized module expression using a principal component analysis (PCA) and calculated eigengenes as the first principal component (PC1) for each module. Second, we used the Pearson correlation to search for associations between module eigengenes and treatment status, and calculated *p*-values for correlations using a Student's asymptotic test. Finally, we applied a Bonferroni correction to account for multiple testing.

For each module significantly associated with treatment status, we performed a gene ontology (GO) enrichment analysis to identify which Biological Process GO terms associated with the DEGs were overrepresented compared to the genome-wide complement of S. salar GO terms ($p < 0.001$). We used the *TopGo* package in R (Alexa & Rahnenfuhrer, 2016), which is less biased towards the most general GO terms because it employs a hierarchical methodology, and chose the 'weight01' algorithm because this method efficiently identifies enriched terms at all levels of the GO hierarchy while limiting the proportion of false positives (Alexa, Rahnenfuhrer, & Lengauer, 2006). After identifying overrepresented GO terms in each module, we created a list of terms

unique to each module. For each module, we created a list of the top 20 genes ranked by gene significance (a value calculated in *WGCNA* that indicates the biological significance of a module gene with respect to the explanatory variable of interest). We used unique GO terms associated with the top 20 genes to construct a network of GO terms for each module, and the *metacoder* package (Foster, Sharpton, & Grünwald, 2017) to visualize networks in R. We pruned internal nodes from each network for ease of visualization.

2.3.6 Categorizing treatment effects: additive vs. genotype x environment interactions

To categorize treatment effects, we first limited our analyses to genes previously identified as differentially expressed with respect to thiamine treatment (see section above). Additive effects occur when the response to the thiamine treatment is equal across families. When gene expression does not change across families within both treatment and control groups, this pattern represents an environmental response to thiamine treatment, with no genetic variation in response across families. For genes exhibiting this pattern, genetic variation in response may be possible, but this variation is not present in our sampled families and may not be present in the Lake Champlain population. By contrast, one family may respond to thiamine treatment differently than another, and this pattern can result in a genotype x environment interaction (i.e., family x treatment effect). Both additive effects and genotype x environment interactions are identifiable by regressing among-family variation in survival against gene expression.

We first calculated regressions separately for each treatment group with log(hazard ratio) as the explanatory variable and fragments per million mapped fragments (FPM) as the response variable. To account for the fact that we sequenced two siblings from each family, we conducted a bootstrapping approach where we randomly selected one individual from each family, performed a linear regression between log(hazard ratio) and FPM, and repeated this process 1,000 times per gene. We identified interactions (i.e., significant differences) between the treatment and control groups by comparing the coefficient estimates for slope and intercept. Specifically, additive effects occur when *y*-intercepts differ between the treated and untreated groups, but slopes do not. Genotype x environment interactions occur when the slopes differ. To approximate a significance cut off of $\alpha = 0.05$, for slope and intercepts separately, we identified genes where the mean coefficient estimates +/- 1 standard error (approximated by 83% quantiles; Payton, Greenstone, & Schenker, 2003) between the treatment and control groups did not overlap.

2.3.7 Identifying putatively adaptive genes (genetic effects)

By examining how gene expression changes in untreated individuals across families with varying survival, we can identify genes that underlie adaptive differences in how families respond to thiamine deficiency. To accomplish this task, we first conducted differential gene expression analysis on untreated individuals in *DESeq2* with family hazard ratio value as the explanatory variable. To identify genes where changes in expression were directly associated with changes in survival we performed 3 steps: First, we considered genes with $p_{adj} < 0.05$ and with a fold-change > 1 ($\log_2$ fold-change > 0.5) between the families with the lowest and highest hazard ratio values to be putatively adaptive. Second, we filtered genes by applying a linear regression approach to each gene, with log(hazard ratio) as the explanatory variable and overall gene expression (fragments per million mapped fragments, FPM) as the response variable. To account for the fact that we sequenced two siblings from each family, we again conducted a bootstrapping approach for each gene where we randomly selected one individual from each family, performed the linear regression, and repeated this process 1,000 times. For each slope, we calculated means and 95% confidence intervals. We discarded genes from further analyses if the slope of the regression did not differ from 0 or if the adjusted r^2 of the regression was < 0.3. Finally, we further categorized the putatively adaptive genes by whether gene expression increased or decreased with increasing hazard ratio. For the putatively adaptive genes identified using the above criteria, we performed a gene ontology (GO) enrichment analysis (see "Treatment effects (environmental and genetic effects)"), ranked GO terms by *p*-value for each category, and retained the top 50 terms from each group ($p < 0.001$ for all retained terms). Because we are identifying associations between differential survival among families (i.e., genotypes) and phenotypic variation (i.e., differential gene expression), among-family variation in expression of a subset of putatively adaptive genes is predicted to play a causative role in determining survival outcomes (A. P. Brown, Arias-Rodriguez, Yee, Tobler, & Kelley, 2018). Genes upregulated in families with high survival relative to families with low survival are predicted to play roles in maintenance of normal physiological conditions that individuals from low survival families cannot maintain. By contrast, genes upregulated in families with low survival are predicted to comprise responses to physiological stress induced by severe thiamine deficiency.

2.4 Results

2.4.1 Thiamine concentrations and survival analyses

In 2016 and 2017, the proportion of untreated individuals surviving within each family varied widely and ranged from 0 to 0.99 (mean = 0.25, SD = 0.34) (Fig. 2.1B). Total thiamine concentrations in unfertilized eggs were also variable and ranged from 0.98 to 12.71 nmol total thiamine/g unfertilized egg tissue (mean = 3.09 nmol/g, SD = 2.23 nmol/g). Fitting a dose-response curve to the relationship between egg thiamine concentration (nmol/g) and proportion of untreated individuals surviving at the end of each experiment resulted in an EC_{25} of 2.89 nmol/g and an EC_{50} of 5.46 nmol/g (i.e., 5.46 nmol/g of thiamine is required for 50% survival) (Fig. 2.1C).

Within each of the 9 families sampled for gene expression analyses, Kaplan-Meier survival distributions for treated and untreated individuals were significantly different (log-rank test; family A: $p < 0.01$, families B-I: $p < 0.0001$), indicating that thiamine treatment significantly and positively impacted survival over time for all families (Fig. A.1). Hazard ratios ranged from 1 (for reference family A) to 80.12 (family I). Hazard ratio values > 1 indicate that a higher risk of death is associated with belonging to a particular family (i.e., the risk of death associated

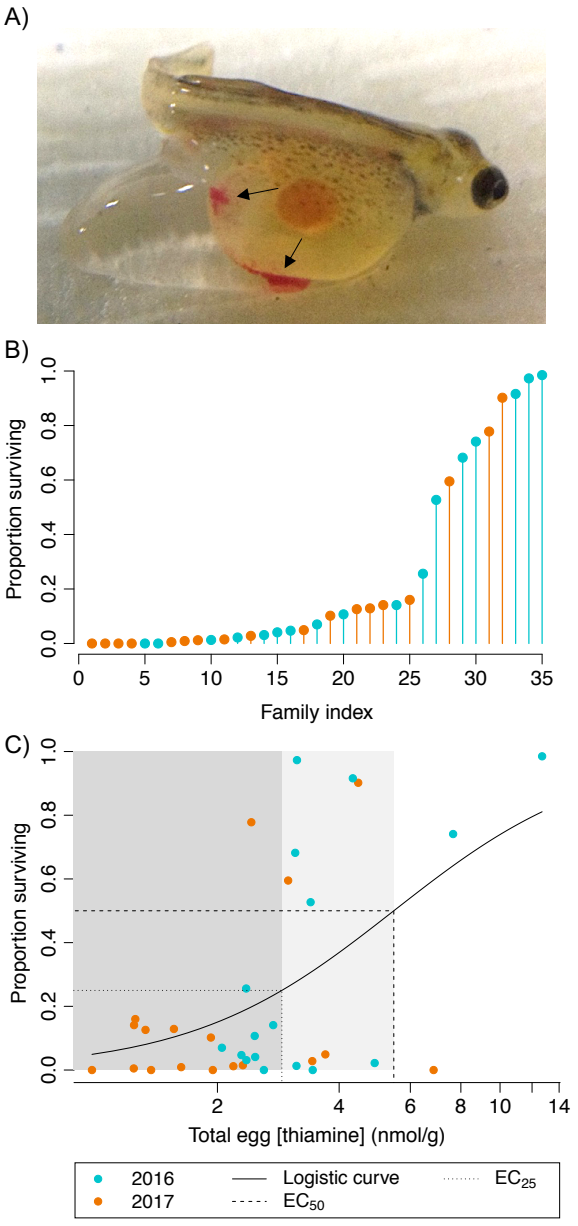


Figure 2.1. A) Atlantic salmon fry exhibiting characteristic signs of thiamine deficiency, including hemorrhaging (indicated by arrows) and edema in the posterior portion of the yolk sac. B) Proportion of untreated fry surviving to the onset of exogenous feeding for 35 families spawned in 2016 and 2017. C) Dose-response curve illustrating the relationship between total egg thiamine concentration (nmol/g) and proportion of untreated fry surviving to yolk sac absorption. Shaded grey areas highlight families with egg [thiamine] <EC_{25} (dark grey) and with egg [thiamine] >EC_{25} but <EC_{50} (light grey), where EC_{25} and EC_{50} equal the effective concentrations required for 25% and 50% survival, respectively.

with belonging to family I is 80.12 times greater than the risk of death associated with belonging to family A).

2.4.2 Treatment effects (environmental and genetic effects)

Across all 36 individuals sequenced, the average rate of single concordant alignment for read pairs per sample was 80.9% and 62.2% of read pairs were successfully assigned to annotated features with *featureCounts*. The final list of *a priori* genes included 106 unique genes. Of these genes, 17 were differentially expressed between treated and untreated individuals after controlling for false discovery ($p_{adj} < 0.05$; Table 2.1). Three of these genes—which encode adenylate kinase and reduced folate carrier—are involved in regulating intracellular concentrations of TDP (Fig. A.2). Most of the remaining *a priori* DEGs comprise TDP-dependent enzymes or kinases that control TDP-dependent enzyme activity (Table 2.1). Differential expression analysis conducted using the full list of assembled transcripts resulted in the identification of 3,616 DEGs after controlling for false discovery ($p_{adj} < 0.05$). A principal component analysis conducted with these DEGs showed treated samples clustering closely together, with PC1 differentiating treated and untreated individuals within each family and describing 59% of the variation (Fig. 2.2). After Bonferroni correction, 3 *WGCNA* modules of co-expressed genes were significantly correlated with treatment status (corrected $p < 0.05$). Module A contained 667 genes and these genes were associated with 647 significantly overrepresented GO terms; 46 GO terms were unique to Module A and associated with the top genes in the module when genes were ranked by gene significance (terminal nodes in Fig. 2.3A, Table A.1). Many GO terms associated with genes in Module A were related to neurological function and development, including regulation of long- term neuronal synaptic plasticity, neurotransmitter secretion, and neuromuscular junction development (Fig. 2.3A). Differential expression of genes involved in neurological function may underlie the abnormal locomotion patterns observed in thiamine deficient fry. Module B contained 355 genes associated with 261 significantly overrepresented GO terms; 17 GO terms were unique to Module B and associated with the top genes in the module. Of these 17 GO terms, 8 were associated with metabolism, including positive regulation of insulin secretion, glutamine metabolic process, and tricarboxylic acid metabolic process (Fig. 2.3B). Differential expression of genes related to these terms is likely related to diminished metabolic rates in untreated individuals.

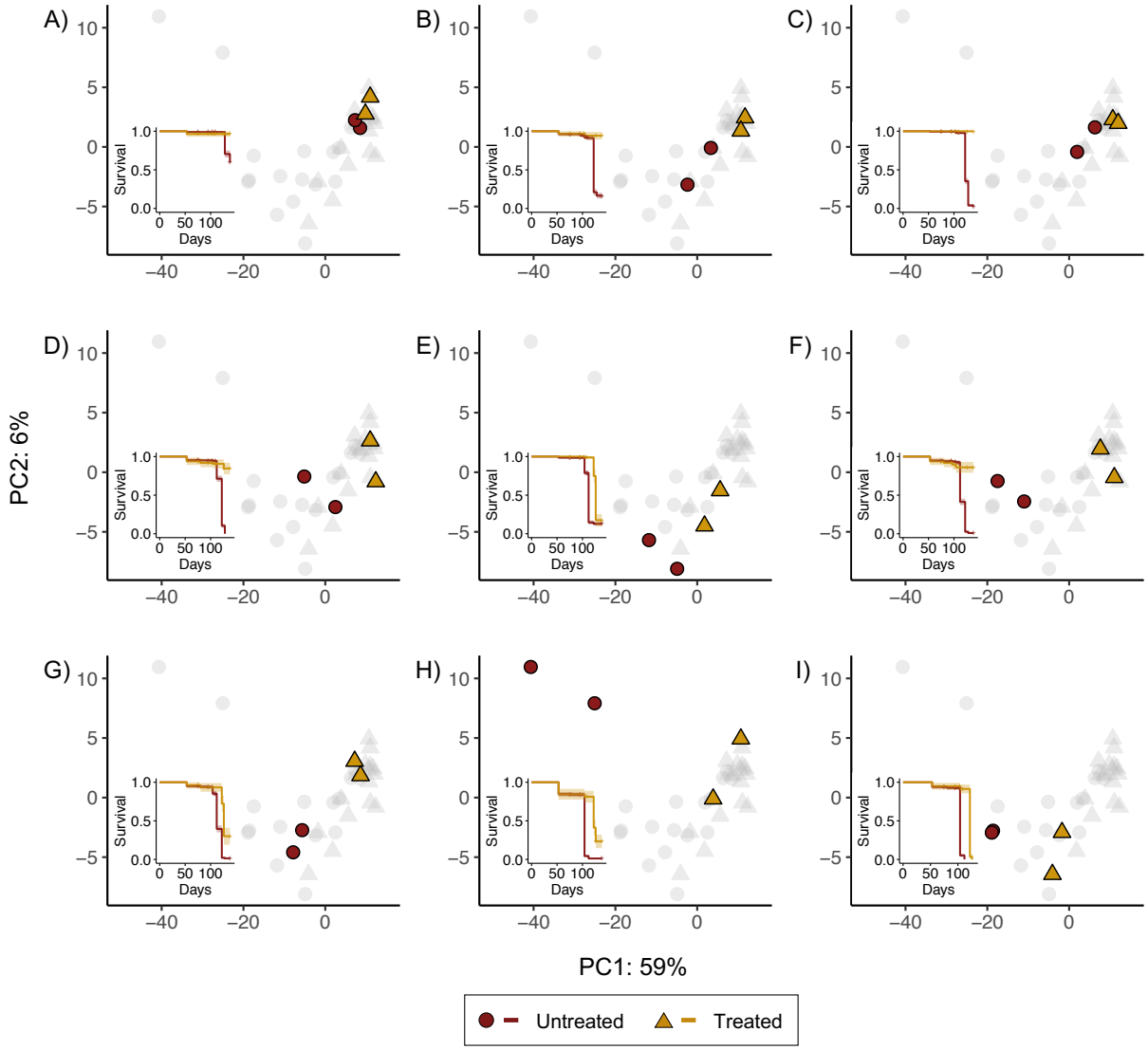


Figure 2.2. Principal component analysis performed using genes differentially expressed with respect to treatment (n = 3,616). Results are presented by family (A-I), such that the four colored points plotted in each panel are full siblings. PC1 explains 59% of the variation and distinguishes among treated and untreated samples. Triangles represent treated individuals and circles represented untreated individuals. Insets are Kaplan-Meier survival distributions for treated (yellow) and untreated (red) individuals from each family. Inset *x*-axes represent time in days post fertilization, whereas the *y*-axes represent survival probability. Hatch marks on survival distributions indicate censored individuals (i.e., samples removed for RNA-seq sampling or disease testing). Stand-alone survival distributions are presented in Fig. A.1 along with family thiamine concentrations and hazard ratio values.

Module C contained 470 genes associated with 768 significantly overrepresented GO terms; 51 GO terms were unique to Module C and associated with the top genes in the module. Many of these GO terms were related to cardiovascular function and development, such as oxygen transport, endocardium formation, and blood vessel maturation (Fig. 2.3C).

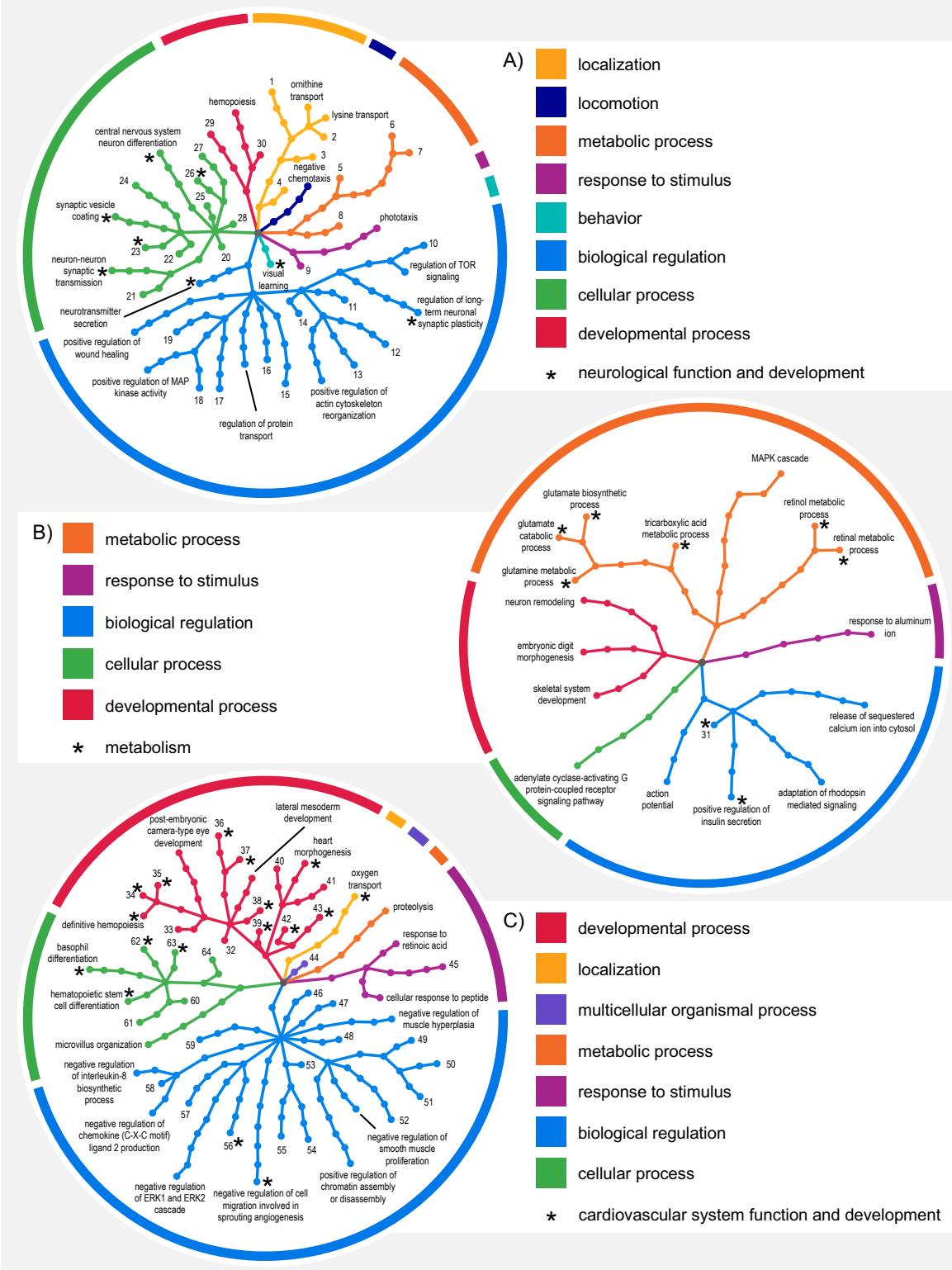
Additionally, all 3 modules contained terms related to vision, including visual learning, retinal metabolic process, adaptation of rhodopsin mediated signaling, and post-embryonic camera-type eye development. Differential expression of genes related to these terms in untreated individuals is likely associated with decreased visual acuity documented in thiamine deficient fry (Carvalho et al., 2009). Each module also contained DEGs identified through differential expression analysis (representing 23.4%, 18.3%, and 24.5% of genes in each module, respectively). The DEGs assigned to module A were downregulated in treated individuals, while the DEGs assigned to modules B and C were upregulated in treated individuals (Fig. A.3).

2.4.3 Categorizing treatment effects: additive vs. genotype x environment interactions

The differential expression of 114 genes in response to thiamine treatment was driven by additive effects, meaning that the response to treatment was equal among families (Fig. 2.4A). Of these 114 genes, 84 genes also showed no among-family variation in gene expression, suggesting that the response to thiamine in this group of genes is largely environmental in this study (i.e., the sampled families are genetically invariant in their responses). For 30 genes exhibiting additive environmental effects, we also identified significant among-family variation in expression, suggesting that these genes are both putatively adaptive and respond equally across families (Fig. 2.4A,B). For example, expression level of popeye domain-containing protein 2 (*popdc2*; Appendix A Discussion) decreases with increasing hazard ratio rank for both treatment groups, with equal slopes between treatments (Fig. 2.4B).

Figure 2.3. Gene ontology (GO) hierarchy networks constructed using the *metacoder* package in R for three modules (A, B, and C) of co-expressed genes significantly associated with treatment status. Included GO terms were unique to each module and were associated with at least one of the top 20 genes in that module when genes were ranked by *WGCNA* gene significance. Branch and node colors indicate the biological process child term to which distal nodes belong, with the central grey node representing the biological process level of the GO hierarchy. Terms associated with numbered terminal nodes are provided in Table A.1. Terminal nodes marked with an * indicate nodes related to a function or process commonly represented in that module network (e.g., several terminal nodes in Module A are related to neurological function and development).

Figure 2.3 continued



For 597 genes differentially expressed with respect to treatment, the slopes of the regressions for each treatment group differed, indicating a genotype x environment interaction. The vast majority (460/720) of genes that exhibited genotype x environment interactions fell into two categories (Fig. 2.4A), both of which had a shared *y*-intercept (see Fig. A.4 for all identified categories). Gene expression between treated and untreated individuals was most similar at lower hazard ratio values (i.e., higher survival), with expression levels of treated and untreated individuals diverging with increasing hazard ratio. For example, expression levels of the optineurin and gamma-crystallin M2 genes did not differ between treatments for the family with the highest survival (lowest hazard ratio value) (Fig. 2.4C,D). These two genes differ in their responses to treatment; thiamine treatment decreases expression of optineurin in low survival families, whereas treatment increases expression of gamma-crystallin M2 in low survival families. Because the slopes of the treatment group regressions differ, treatment does not evenly affect gene expression across families; the families with the lowest survival (highest hazard ratio values) experienced the largest shifts in expression levels in response to treatment.

2.4.4 Putatively adaptive genes (genetic effects)

Maternal effects may influence among-family variation in survival and gene expression. However, we could not identify any maternal characteristics, including maternal thiamine concentrations, that were associated with survival of untreated offspring. Specifically, maternal size and weight were not correlated with untreated offspring survival rate (standard length: $F_{1,32} = 0.15$, $p = 0.70$; weight: $F_{1,31} = 0.10$, $p = 0.75$), indicating that differences in survival among families is not simply a function of maternal condition (Fig. 2.5A,B). Furthermore, for the 9 families sampled for RNA-seq, no relationship appears to exist between maternal muscle or egg thiamine concentrations and proportion of untreated offspring surviving (n = 8 and n = 9, respectively; Fig. 2.5C,D). Results of linear regressions also indicated that egg thiamine concentration was not a significant predictor of log(hazard ratio) ($F_{1,7} = 2.18$, $p = 0.18$; Fig. A.5). Although we cannot entirely rule out the influence of maternal effects, these results suggest that maternal effects are not driving all of the among-family variation in survival.

Differential expression analyses conducted using a count matrix for untreated individuals and with family hazard ratio as the explanatory variable yielded 1,656 DEGs. Of these DEGs, 471 were discarded because the adjusted r^2 of the regressions for these genes were < 0.3, and an

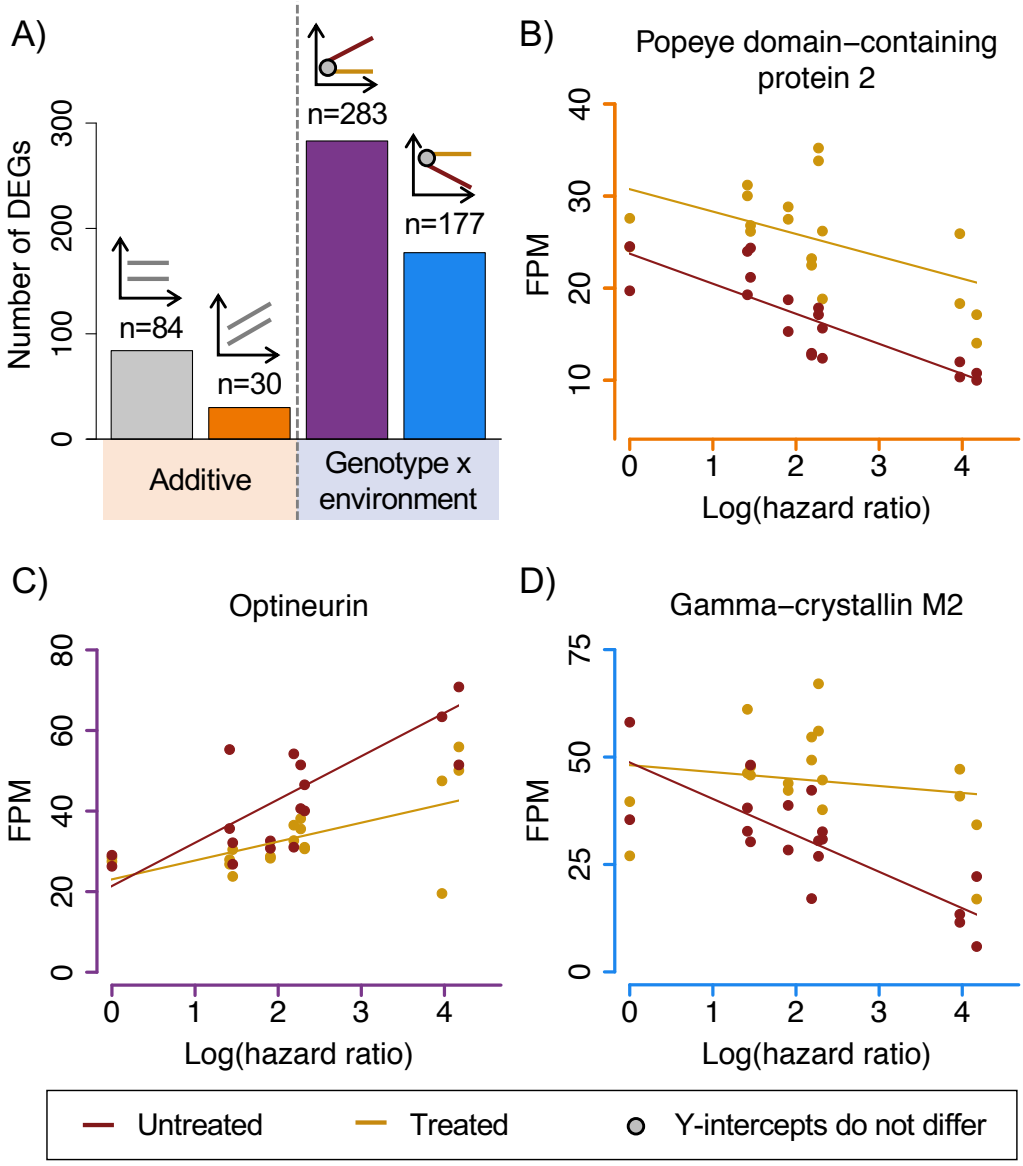


Figure 2.4. A) The number of genes differentially expressed in response to treatment and among-family differences in 4 categories (from left to right): 1) regression slopes are equal between treatment groups and expression is equal across families, indicating a purely environmental effect of treatment on differential gene expression; 2) regression slopes are equal, indicating that families respond evenly to treatment, but among-family differences suggest putatively adaptive responses; 3) and 4) expression levels are equal between treatment groups in high-surviving families with expression patterns diverging in low-surviving families, indicating genotype x environment interactions. Grey lines in example plots in (A) may represent regression lines for either treated or untreated individuals. B-D) Genes exhibiting expression patterns represented in (A). Axes colors correspond to bar colors in (A). Relative expression of B) popeye domain-containing protein 2 (*popdc2*), C) optineurin (*optn*), and D) gamma-crystallin M2 (LOC106575874) in fragments per million mapped (FPM) across log(hazard ratio) values.

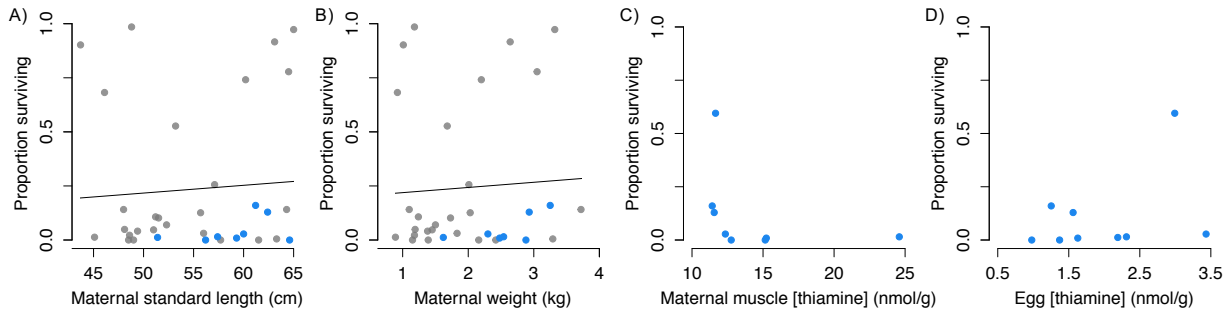


Figure 2.5. No associations were found between maternal characteristics (weight (kg) and standard length (cm), respectively) and proportion of untreated offspring surviving at the end of the experiment, when tested with linear regressions (A,B). There also does not appear to be any relationship between maternal muscle or egg thiamine concentrations (nmol/g) and proportion of offspring surviving for the 9 families sample for RNA-seq (C,D). In A-D, families sampled for RNA-seq are indicated by blue points; note that data were unavailable for some families in panels A-C (i.e., 1 family from panels A and C and 2 families from panel B).

additional 210 were discarded because the regression slopes did not significantly differ from 0. The remaining 1,446 putatively adaptive DEGs were divided into 812 genes positively associated with increased risk of mortality (Fig. 2.6A,C) and 634 genes negatively associated with increased risk of mortality (Fig. 2.6B,C). A total of 720 putatively adaptive DEGs were also differentially expressed in response to thiamine treatment.

Adaptively expressed DEGs with positive and negative slopes were associated with 870 and 741 overrepresented GO terms, respectively ($p < 0.001$). Of the top 50 terms associated with genes in each slope category, 17 terms were shared between the categories. Shared terms were related to a variety of processes, including regulation of transcription, response to glucose, aging, and oxidation-reduction process. Terms associated with genes with negative slopes (i.e., genes upregulated in families with high survival) relate to growth and developmental processes, including cellular proliferation, DNA replication, embryo development, neurogenesis, and visual perception. Genes with positive slopes (i.e., genes upregulated in families with low survival) were associated with terms that indicate stressful physiological conditions, including response to hydrogen peroxide, response to hypoxia, response to toxic substance, and several terms related to toll-like receptor signaling pathways.

2.5 Discussion

Among all families and across both years, we found a high degree of variation in survival (Fig. 2.1B) and egg thiamine concentrations, with most females producing eggs that cannot survive

without supplemental treatment (Fig. 2.1C). Although thiamine concentration thresholds for mortality associated with thiamine deficiency vary across species and populations, the nature of the relationship between egg thiamine concentration and offspring survival is consistently sigmoidal in shape (Fig. 2.1C; Fitzsimons et al., 2007; Futia & Rinchard, 2019; Werner, Rook, & Greil, 2006). As egg thiamine concentration increases, a threshold is eventually reached where small increases in thiamine concentration coincide with large increases in offspring survival. All 9 of the families sampled for RNA-seq exhibited egg thiamine concentrations below this threshold (EC_{50} = 5.46 nmol/g). The threshold for this population is slightly higher than thresholds calculated for other landlocked and anadromous Atlantic salmon populations (Amcoff, Börjeson, Lindeberg, & Norrgren, 1998; Werner et al., 2006), but this difference may be due to the rarity of thiamine replete individuals in the Lake Champlain population. The ubiquity of low egg thiamine in these samples is consistent with extremely limited reproductive success documented in Lake Champlain tributaries (Prévost, Hill, Grant, Ardren, & Fraser, in press).

Treatment with supplemental thiamine does improve survival outcomes for all families, but does not guarantee survival in families with higher hazard ratios (Fig. A.1D-I). Egg thiamine concentration could not predict survival (hazard ratio value) for the 9 families included in gene expression analyses (Fig. A.5), and no relationships appear to exist between maternal muscle or egg thiamine concentrations and proportion of untreated offspring surviving (Fig. 2.5C,D). Across all families, maternal length and weight also do not predict offspring survival (proportion surviving; Fig. 2.5A,B). Because all offspring were raised in a common environment, the absence of this relationship coupled with the high variation in among-family survival indicates that family identity (i.e., genetic background) plays an important role in determining whether an individual will survive thiamine deficiency. Furthermore, in the face of thiamine deficiency, certain families are better able to maintain gene expression profiles that approximate expression profiles under thiamine-rich conditions without the aid of supplemental thiamine (Fig. 2.4), a pattern consistent with a genetic basis for tolerance to low thiamine availability.

Across all families, we found that a large number of genes responded to supplemental thiamine treatment. Of the genes hypothesized to be differentially expressed between treated and untreated individuals *a priori*, two gene products perform functions that balance relative intracellular concentrations of thiamine and its various derivatives (see Appendix A text for Discussion; Fig. A.2).

Table 2.1. Genes identified as differentially expressed that were hypothesized *a priori* to be implicated in thiamine deficiency. Gene symbols correspond to those used in *S. salar* NCBI assembly GCA_000233375.4 (Lien et al., 2016). Direction of log_2 fold-change values indicate direction of regulation in the treated group relative to the untreated group.

Gene symbol	Gene description	Log_2 fold-change	FDR corrected *p*-value
bckdha	branched chain keto acid dehydrogenase E1, alpha polypeptide	-0.166	1.17E-08
LOC106600850	pyruvate dehydrogenase (acetyl-transferring) kinase isozyme 2, mitochondrial-like	0.476	6.09E-05
kad	adenylate kinase 1-2	0.185	6.44E-04
kad	adenylate kinase	0.235	1.70E-03
ilvbl	ilvB (bacterial acetolactate synthase)-like	0.165	1.70E-03
LOC106581299	alkaline phosphatase, tissue-nonspecific isozyme-like	0.638	1.70E-03
LOC106583968	DET1- and DDB1-associated protein 1-like	0.260	1.70E-03
LOC106569142	2-oxoglutarate dehydrogenase, mitochondrial-like	-0.268	3.44E-03
LOC106561658	probable 2-oxoglutarate dehydrogenase E1 component DHKTD1, mitochondrial	-0.288	1.63E-02
LOC106587025	2-oxoglutarate dehydrogenase, mitochondrial	-0.226	1.63E-02
hacl1	2-hydroxyacyl-CoA lyase 1	0.241	2.20E-02
LOC106579320	pyruvate dehydrogenase (acetyl-transferring) kinase isozyme 2, mitochondrial-like	-0.489	2.20E-02
LOC106578452	pyruvate dehydrogenase (acetyl-transferring) kinase isozyme 3, mitochondrial	-0.214	2.28E-02
LOC106580908	thiamine transporter 2-like	0.969	2.28E-02
LOC106560358	thiamine transporter 2-like	1.862	2.85E-02
LOC106563967	dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex, mitochondrial-like	-0.182	3.93E-02
LOC106575607	folate transporter 1-like	0.253	4.36E-02

From the full list of transcript counts, we identified three modules of co-expressed genes associated with treatment status (Fig. 2.3). For each of these modules, clear themes emerged from their unique lists of overrepresented GO terms. Module A's association with neurological function and development identified genes related to specific signs of thiamine deficiency in Atlantic salmon fry, such as uncoordinated swimming patterns, inability to maintain an upright position in the water column, and absence of avoidance behavior in response to light exposure (Fisher, Spitsbergen, et al., 1995). These signs of thiamine deficiency may also be related to other overrepresented terms unique to Module A, including responses to stimuli, such as phototaxis and negative chemotaxis. Overrepresented terms in Module B identified genes associated with metabolism, and differential expression of these genes likely underlies slower rates of development under thiamine deficient conditions, with treated individuals achieving larger body sizes than untreated individuals of the same age (Fitzsimons et al., 2009). In Module C, overrepresented

terms identified genes related to cardiovascular function and development and may drive vascular dysfunction observed in untreated individuals, as evidenced by hemorrhaging, vascular congestion, and irregular heart rate (Fig. 2.1A; Fisher, Spitsbergen, et al., 1995). The ubiquity of terms related to vision and eye development shared across all three modules of co-expressed genes demonstrates the complexity of relationships among genes that influence proper development of the visual system (e.g., A: visual learning and phototaxis; B: retinol and retinal metabolic processes; C: post-embryonic camera-type eye development).

Of the 3,616 treatment-effect DEGs, 84 displayed evidence of an additive effect of treatment that was not associated with among-family survival (Fig. 2.4A). In other words, these 84 genes responded to thiamine treatment equally across the sampled families and represent a consistent environmental response to the treatment condition. Thus, we would not predict these genes to respond to selection in a thiamine-poor environment, although a response may be possible in a more genetically variable population. By contrast, 597 DEGs provided evidence of a genotype x environmental interaction (i.e., family x treatment effect). Most of the genes exhibiting genotype x environment interactions (460/720 or 64%) followed predictable patterns of expression that can be broadly grouped into two out of six identified categories (Fig. 2.4A; see Fig. A.4 for all family x treatment patterns). The observation that the majority (64%) of genotype x environment interaction genes occur in the two categories with shared *y*-intercepts indicates that a successful response to thiamine-poor conditions involves the maintenance of homeostatic conditions.

Differential responses to thiamine deficiency among families comprise 1,446 putatively adaptive genes. The large number of putatively adaptive genes with expression patterns that may be subject to selection is consistent with previous findings describing major roles for regulatory changes in adaptation (Jones et al., 2012; Schoville, Barreto, Moy, Wolff, & Burton, 2012; Uusi-Heikkilä et al., 2017). A total of 812 genes were significantly upregulated in untreated individuals from high-survival families (Fig. 2.6A,C). This result suggests that the increased expression of these genes is associated with higher survival and that these genes, or the various cis- or trans-acting regulatory elements that influence their expression, could respond to selection in a thiamine-poor environment. Of course, these putatively adaptive genes could also be affected by maternal effects (though survival was not correlated with any maternal or egg traits that we measured; Fig. 2.5), heritable epigenetic effects (Le Luyer et al., 2017; Rodriguez Barreto et al., 2019), or may be influenced by epistatic interactions with other genes (Hemani et al., 2014). Furthermore, while the

putatively adaptive genes may not be the direct targets of selection, their expression patterns are directly affected by the genes (i.e., pathways) or regulatory elements that selection would act upon. Thus, we are not suggesting that all of these genes would underlie an adaptive response to selection, but rather that this list represents a suite of candidate genes that would likely respond to selection. Confirmation of the roles of candidate genes in facilitating adaptation could be achieved in future studies through mapping of quantitative trait loci (QTLs), reciprocal crosses or extending the current experimental design through the F_2 generation to formally control for maternal effects, or through a genome-wide association study.

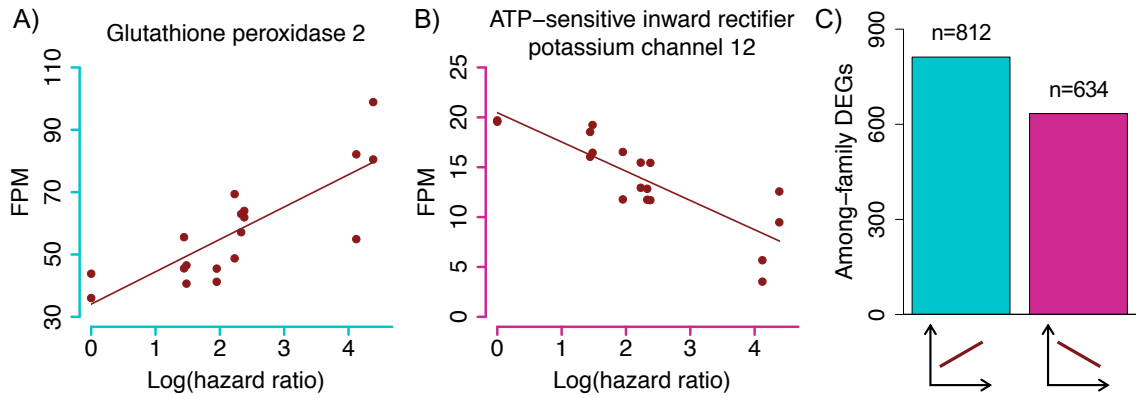


Figure 2.6. Relative expression of A) glutathione peroxidase (LOC106583190) and B) ATP-sensitive inward rectifier potassium channel 12 (LOC106600689) in fragments per million mapped (FPM) across log(hazard ratio) values. C) The number of putatively adaptive genes positively and negatively associated with increasing risk of mortality. Putatively adaptive genes positively associated with increasing risk of mortality are largely associated with gene ontology terms related to physiological stress, whereas genes negatively associated with risk are associated with terms related to growth and developmental processes.

Of the 35 families sampled in this study, the majority (25/35) experienced >75% mortality of untreated individuals, indicating that thiamine deficiency would exert strong selection on the Lake Champlain population in the absence of thiamine treatments (Fig. 2.1B). Variability in gene expression measured across families with varying survival rates suggests that evolutionary rescue could occur in response to the withdrawal of thiamine treatments in this population (Carlson, Cunningham, & Westley, 2014), but more genetic, experimental, and theoretical work is warranted. It is worth noting that there is minimal natural reproduction in Lake Champlain (Prévost et al., in press) and that all released salmon are treated with supplemental thiamine; this relaxed natural selection could be limiting the successful reintroduction of salmon into the wild.

2.6 Conclusion

Across families, we found no relationship between egg thiamine concentration, maternal traits, and the risk of mortality due to thiamine deficiency indicating that among-family genetic variation plays an important role in determining thiamine deficiency outcomes. Through gene co-expression network analyses, we determined that many GO terms associated with top DEGs are consistent with observed behavioral and physical signs of thiamine deficiency. Specifically, terms related to neurological function and development, metabolism, cardiovascular function, and visual system development parallel signs of deficiency including uncoordinated swimming patterns, stunted growth, irregular heart rate, and decreased visual acuity. We also described two broad categories of gene expression patterns in response to thiamine deficiency: (1) putatively adaptive genes, which underlie family-level differences in tolerance to low thiamine availability and represent candidate genes likely to respond to selection, and (2) treatment effect genes, which comprise additive effects and genotype x environment interactions in response to changes in available thiamine. An additive response coupled with no among-family variation in expression identifies genes that are plastic and respond largely as a function of the treatment condition (i.e., different environments). Such genes would be useful to identify in scenarios where a response to selection was not desirable (e.g., captive breeding programs). Genes exhibiting genotype x environment interactions, on the other hand, illustrate that genetic background differentially affects patterns of gene expression. Our results also identified putatively adaptive genes that would likely respond to selection and that are directly associated with among-family variation in survival. Precisely how much of a response to selection in the wild could occur remains unknown, but uncovering the adaptive genetic variation required for a response to selection represents the first step towards the successful management and conservation species threatened by changing environmental conditions.

2.7 References

AFS, AIFRB, ASIH (American Fisheries Society, American Institute of Fishery Research Biologists, and American Society of Ichthyologists and Herpetologists). (2014). Guidelines for the use of fishes in research. Bethesda, Maryland: American Fisheries Society.

Alexa, A., Rahnenfuhrer, J., & Lengauer, T. (2006). Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics, 22(13), 1600–1607. doi: 10.1093/bioinformatics/btl140

Alexa, A., & Rahnenfuhrer, J. (2016). topGO: Enrichment analysis for gene ontology (Version R package v2.30.1).

Amcoff, P., Börjeson, H., Lindeberg, J., & Norrgren, L. (1998). Thiamine concentrations in feral Baltic salmon exhibiting the M74 syndrome. In Early life stage mortality syndrome in fishes of the Great Lakes and Baltic Sea (pp. 82–89). Bethesda, Maryland: American Fisheries Society.

Balk, L., Hägerroth, P.-Å., Åkerman, G., Hanson, M., Tjärnlund, U., Hansson, T., … Sundberg, H. (2009). Wild birds of declining European species are dying from a thiamine deficiency syndrome. Proceedings of the National Academy of Sciences, 106(29), 12001–12006.

Balk, L., Hägerroth, P.-Å., Gustavsson, H., Sigg, L., Åkerman, G., Ruiz Muñoz, Y., … Hansson, T. (2016). Widespread episodic thiamine deficiency in Northern Hemisphere wildlife. Scientific Reports, 6, 38821.

Bernatchez, L. (2016). On the maintenance of genetic variation and adaptation to environmental change: Considerations from population genomics in fishes. Journal of Fish Biology, 89(6), 2519–2556. doi: 10.1111/jfb.13145

Bettendorff, L. (2013). Thiamine. In Handbook of Vitamins, Fifth Edition (pp. 267–324). CRC Press.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics, 30(15), 2114–2120. doi: 10.1093/bioinformatics/btu170

Brown, A. P., Arias-Rodriguez, L., Yee, M.-C., Tobler, M., & Kelley, J. L. (2018). Concordant changes in gene expression and nucleotides underlie independent adaptation to hydrogen-sulfide-rich environments. Genome Biology and Evolution. doi: 10.1093/gbe/evy198

Brown, S. B., Honeyfield, D. C., Hnath, J. G., Wolgamood, M., Marcquenski, S. V., Fitzsimons, J. D., & Tillitt, D. E. (2005). Thiamine status in adult salmonines in the Great Lakes. Journal of Aquatic Animal Health, 17(1), 59–64.

Butterworth, R. F. (2009). Thiamine deficiency-related brain dysfunction in chronic liver failure. Metabolic Brain Disease, 24(1), 189–196. doi: 10.1007/s11011-008-9129-y

Carlson, S. M., Cunningham, C. J., & Westley, P. A. H. (2014). Evolutionary rescue in a changing world. Trends in Ecology & Evolution, 29(9), 521–530.

Carvalho, P. S. M., Tillitt, D. E., Zajicek, J. L., Claunch, R. A., Honeyfield, D. C., Fitzsimons, J. D., & Brown, S. B. (2009). Thiamine deficiency effects on the vision and foraging ability of lake trout fry. Journal of Aquatic Animal Health, 21(4), 315–325.

Christie, M. R., Marine, M. L., Fox, S. E., French, R. A., & Blouin, M. S. (2016). A single generation of domestication heritably alters the expression of hundreds of genes. Nature Communications, 7, 10676. doi: 10.1038/ncomms10676

Cox, D. R. (1972). Regression models and life tables. Journal of the Royal Statistical Society. Series B (Methodological), 34, 187–220.

Essa, E., Velez, M. R., Smith, S., Giri, S., Raman, S. V., & Gumina, R. J. (2011). Cardiovascular magnetic resonance in wet beriberi. Journal of Cardiovascular Magnetic Resonance, 13(1), 41. doi: 10.1186/1532-429X-13-41

Fisher, J. P., Spitsbergen, J. M., Iamonte, T., Little, E. E., & Delonay, A. (1995). Pathological and behavioral manifestations of the "Cayuga syndrome," a thiamine deficiency in larval landlocked Atlantic salmon. Journal of Aquatic Animal Health, 7(4), 269–283.

Fisher, J. P., Spitzbergen, J. M., Getchell, R., Symula, J., Skea, J., Babenzein, M., & Chiotti, T. (1995). Reproductive failure of landlocked Atlantic Salmon from New York's Finger Lakes: Investigations into the etiology and epidemiology of the "Cayuga Syndrome." Journal of Aquatic Animal Health, 7, 81–94.

Fitzsimons, J. D., Brown, S. B., Honeyfield, D. C., & Hnath, J. G. (1999). A review of early mortality syndrome (EMS) in Great Lakes salmonids: Relationship with thiamine deficiency. Ambio, 28, 9–15.

Fitzsimons, J. D., Brown, S. B., Williston, B., Williston, G., Brown, L. R., Moore, K., … Tillitt, D. E. (2009). Influence of thiamine deficiency on lake trout larval growth, foraging, and predator avoidance. Journal of Aquatic Animal Health, 21(4), 302–314.

Fitzsimons, J. D., Williston, B., Amcoff, P., Balk, L., Pecor, C., Ketola, H. G., … Honeyfield, D. C. (2005). The effect of thiamine injection on upstream migration, survival, and thiamine status of putative thiamine-deficient coho salmon. Journal of Aquatic Animal Health, 17(1), 48–58.

Fitzsimons, J. D., Williston, B., Williston, G., Brown, L., El-Shaarawi, A., Vandenbyllaardt, L., … Brown, S. B. (2007). Egg thiamine status of Lake Ontario salmonines 1995-2004 with emphasis on lake trout. Journal of Great Lakes Research, 33, 93–103.

Foster, Z. S. L., Sharpton, T. J., & Grünwald, N. J. (2017). Metacoder: An R package for visualization and manipulation of community taxonomic diversity data. PLOS Computational Biology, 15.

Franks, S. J., Kane, N. C., O'Hara, N. B., Tittes, S., & Rest, J. S. (2016). Rapid genome-wide evolution in *Brassica rapa* populations following drought revealed by sequencing of ancestral and descendant gene pools. Molecular Ecology. doi: 10.1111/mec.13615

Futia, Matthew H., & Rinchard, J. (2019). Evaluation of adult and offspring thiamine deficiency in salmonine species from Lake Ontario. Journal of Great Lakes Research, 45(4), 811–820. doi: 10.1016/j.jglr.2019.05.010

Futia, M.H., Hallenbeck, S., Noyes, A. D., Honeyfield, D. C., Eckerlin, G. E., & Rinchard, J. (2017). Thiamine deficiency and the effectiveness of thiamine treatments through broodstock injections and egg immersion on Lake Ontario steelhead trout. Journal of Great Lakes Research, 43(2), 352–358.

Hansson, S., Karlsson, L., Ikonen, E., Christensen, O., Mitans, A., Uzars, D., … Ragnarsson, B. (2001). Stomach analyses of Baltic salmon from 1959-1962 and 1994-1997: Possible relations between diet and yolk-sac-fry mortality (M74). Journal of Fish Biology, 58(6), 1730–1745.

Harder, A. M., Ardren, W. R., Evans, A. N., Futia, M. H., Kraft, C. E., Marsden, J. E., … Christie, M. R. (2018). Thiamine deficiency in fishes: Causes, consequences, and potential solutions. Reviews in Fish Biology and Fisheries, 28(4), 865–886. doi: 10.1007/s11160-018-9538-x

Hemani, G., Shakhbazov, K., Westra, H.-J., Esko, T., Henders, A. K., McRae, A. F., … Powell, J. E. (2014). Detection and replication of epistasis influencing transcription in humans. Nature, 508(7495), 249–253. doi: 10.1038/nature13005

Honeyfield, D. C., Ross, J. P., Carbonneau, D. A., Terrell, S. P., Woodward, A. R., Schoeb, T. R., … Hinterkopf, J. P. (2008a). Pathology, physiologic parameters, tissue contaminants, and tissue thiamine in morbid and healthy central Florida adult American alligators. Journal of Wildlife Diseases, 44(2), 280–294.

Honeyfield, D. C., Ross, J. P., Carbonneau, D. A., Terrell, S. P., Woodward, A. R., Schoeb, T. R., … Hinterkopf, J. P. (2008b). Pathology, physiologic parameters, tissue contaminants, and tissue thiamine in morbid and healthy central Florida adult American alligators (*Alligator mississippiensis).* Journal of Wildlife Diseases, 44(2), 280–294. doi: 10.7589/0090-3558-44.2.280

Houde, A. L. S., Saez, P. J., Wilson, C. C., Bureau, D. P., & Neff, B. D. (2015). Effects of feeding high dietary thiaminase to sub-adult Atlantic salmon from three populations. Journal of Great Lakes Research, 41, 898–906.

Jones, F. C., Grabherr, M. G., Chan, Y. F., Russell, P., Mauceli, E., Johnson, J., … Kingsley, D. M. (2012). The genomic basis of adaptive evolution in threespine sticklebacks. Nature, 484(7392), 55–61. doi: 10.1038/nature10944

Keinänen, M., Uddström, A., Mikkonen, J., Casini, M., Pönni, J., Myllylä, T., … Vuorinen, P. J. (2012). The thiamine deficiency syndrome M74, a reproductive disorder of Atlantic salmon (*Salmo salar*) feeding in the Baltic Sea, is related to the fat and thiamine content of prey fish. ICES Journal of Marine Science, 69(4), 516–528.

Ketola, H. G., Bowser, P. R., Wooster, G. A., Wedge, L. R., & Hurst, S. S. (2000). Effects of thiamine on reproduction of Atlantic salmon and a new hypothesis for their extirpation in Lake Ontario. Transactions of the American Fisheries Society, 129(2), 607–612.

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. Nature Methods, 12(4), 357–360. doi: 10.1038/nmeth.3317

Kleinbaum, D. G., & Klein, M. (2012). Survival analysis: A self-learning text (3rd ed). New York: Springer.

Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. BMC Bioinformatics, 9(1), 559. doi: 10.1186/1471-2105-9-559

Le Luyer, J., Laporte, M., Beacham, T. D., Kaukinen, K. H., Withler, R. E., Leong, J. S., … Bernatchez, L. (2017). Parallel epigenetic modifications induced by hatchery rearing in a Pacific salmon. Proceedings of the National Academy of Sciences, 114(49), 12964–12969. doi: 10.1073/pnas.1711229114

Lee, B.-J., Jaroszewska, M., Dabrowski, K., Czesny, S., & Rinchard, J. (2009). Effects of vitamin B_1 (thiamine) deficiency in lake trout alevins and preventive treatments. Journal of Aquatic Animal Health, 21(4), 290–301.

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics, 30(7), 923–930. doi: 10.1093/bioinformatics/btt656

Lien, S., Koop, B. F., Sandve, S. R., Miller, J. R., Kent, M. P., Nome, T., … Davidson, W. S. (2016). The Atlantic salmon genome provides insights into rediploidization. Nature, 533, 200-205. doi: 10.1038/nature17164

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology, 15(12), 550. doi: 10.1186/s13059-014-0550-8

Marsden, J. E., & Langdon, R. W. (2012). The history and future of Lake Champlain's fishes and fisheries. Journal of Great Lakes Research, 38, 19–34. doi: 10.1016/j.jglr.2011.09.007

Marshall, D. J., Heppell, S. S., Munch, S. B., & Warner, R. R. (2010). The relationship between maternal phenotype and offspring quality: Do older mothers really produce the best offspring? Ecology, 91(10), 2862–2873. doi: 10.1890/09-0156.1

Mörner, T., Hansson, T., Carlsson, L., Berg, A.-L., Ruiz Muñoz, Y., Gustavsson, H., … Balk, L. (2017). Thiamine deficiency impairs common eider (Somateria mollissima) reproduction in the field. Scientific Reports, 7(1). doi: 10.1038/s41598-017-13884-1

Myers, J. H., Simberloff, D., Kuris, A. M., & Carey, J. R. (2000). Eradication revisited: Dealing with exotic species. Trends in Ecology & Evolution, 15(8), 316–320. doi: 10.1016/S0169-5347(00)01914-5

Passow, C. N., Henpita, C., Shaw, J. H., Quackenbush, C. R., Warren, W. C., Schartl, M., … Tobler, M. (2017). The roles of plasticity and evolutionary change in shaping gene expression variation in natural populations of extremophile fish. Molecular Ecology, 26(22), 6384–6399. doi: 10.1111/mec.14360

Payton, M. E., Greenstone, M. H., & Schenker, N. (2003). Overlapping confidence intervals or standard error intervals: What do they mean in terms of statistical significance? Journal of Insect Science, 3, 1–6.

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., & Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nature Biotechnology, 33(3), 290–295. doi: 10.1038/nbt.3122

Prévost, A., Hill, N., Grant, J., Ardren, W., & Fraser, D. (in press). Patterns of reproductive success among reintroduced Atlantic salmon in two Lake Champlain tributaries. Conservation Genetics, in press.

Pritchard, J. K., Pickrell, J. K., & Coop, G. (2010). The genetics of human adaptation: Hard sweeps, soft sweeps, and polygenic adaptation. Current Biology, 20(4), R208–R215. doi: 10.1016/j.cub.2009.11.055

R Core Team. (2019). R: A language and environment for statistical computing. Retrieved from https://www.R-project.org/

Ritz, C., Baty, F., Streibig, J. C., & Gerhard, D. (2015). Dose-response analysis using R. PLOS ONE, 10(12), e0146021. doi: 10.1371/journal.pone.0146021

Rodriguez Barreto, D., Garcia de Leaniz, C., Verspoor, E., Sobolewska, H., Coulson, M., & Consuegra, S. (2019). DNA methylation changes in the sperm of captive-reared fish: a route to epigenetic introgression in wild populations. Molecular Biology and Evolution, 36(10), 2205–2211. doi: 10.1093/molbev/msz135

Ross, J. P., Honeyfield, D. C., Brown, S. B., Brown, L. R., Waddle, A. R., Welker, M. E., & Schoeb, T. R. (2009). Gizzard shad thiaminase activity and its effect on the thiamine status of captive American alligators *Alligator mississippiensis*. Journal of Aquatic Animal Health, 21(4), 239–248. doi: 10.1577/H08-002.1

Samy, J. K. A., Mulugeta, T. D., Nome, T., Sandve, S. R., Grammes, F., Kent, M. P., … Våge, D. I. (2017). SalmoBase: An integrated molecular data resource for salmonid species. BMC Genomics, 18(1). doi: 10.1186/s12864-017-3877-1

Schoville, S. D., Barreto, F. S., Moy, G. W., Wolff, A., & Burton, R. S. (2012). Investigating the molecular basis of local adaptation to thermal stress: Population differences in gene expression across the transcriptome of the copepod *Tigriopus californicus*. BMC Evolutionary Biology, 12(1), 170. doi: 10.1186/1471-2148-12-170

Sechi, G., & Serra, A. (2007). Wernicke's encephalopathy: New clinical settings and recent advances in diagnosis and management. The Lancet Neurology, 6(5), 442–455.

Stockwell, C. A., Hendry, A. P., & Kinnison, M. T. (2003). Contemporary evolution meets conservation biology. Trends in Ecology & Evolution, 18(2), 94–101.

Therneau, T. (2015). A package for survival analysis in R (Version 2.38). Retrieved from https://CRAN.R-project.org/package=survival

Uusi-Heikkilä, S., Sävilammi, T., Leder, E., Arlinghaus, R., & Primmer, C. R. (2017). Rapid, broad-scale gene expression evolution in experimentally harvested fish populations. Molecular Ecology. doi: 10.1111/mec.14179

van't Hof, A. E., Campagne, P., Rigden, D. J., Yung, C. J., Lingley, J., Quail, M. A., … Saccheri, I. J. (2016). The industrial melanism mutation in British peppered moths is a transposable element. Nature, 534(7605), 102–105. doi: 10.1038/nature17951

Wellenreuther, M., & Hansson, B. (2016). Detecting polygenic evolution: Problems, pitfalls, and promises. Trends in Genetics, 32(3), 155–164. doi: 10.1016/j.tig.2015.12.004

Werner, R. M., Rook, B., & Greil, R. (2006). Egg-thiamine status and occurrence of early mortality syndrome (EMS) in Atlantic salmon from the St. Marys River, Michigan. Journal of Great Lakes Research, 32(2), 293–305.

Willoughby, J. R., Harder, A. M., Tennessen, J. A., Scribner, K. T., & Christie, M. R. (2018). Rapid genetic adaptation to a novel environment despite a genome-wide reduction in genetic diversity. Molecular Ecology, 27(20), 4041–4051. doi: 10.1111/mec.14726

CHAPTER 3. SIGNALS OF SELECTION ASSOCIATED WITH DIFFERENCES IN LIFE HISTORY STRATEGY AND HATCHERY REARING IN LANDLOCKED AND ANADROMOUS ATLANTIC SALMON (*SALMO SALAR*)

3.1 Abstract

Adaptation of marine fishes to novel, freshwater environments requires major physiological shifts in functions related to osmoregulation, immunity, and metabolism. For Atlantic salmon (*Salmo salar)* populations, such changes have occurred independently in many landlocked populations that were formed as a result of extensive hydrological shifts in North America around 10,000 years ago. We compared patterns of genetic variation between two landlocked and one anadromous population of Atlantic salmon to identify loci that may have played a role in adaptation from anadromy to an entirely freshwater life cycle. Outlier loci included SNPs in genes related to expected functions, including immunity, growth, and osmoregulation. We also used these same populations to characterize loci that may be associated with hatchery rearing practices. This additional comparison identified outlier SNPs in genes related to wound healing and regulation of circadian rhythm, consistent with findings from other genetic studies of domestication selection. Together, our results highlight putative responses to both natural selection imposed by major environmental changes and artificial selection levied by differing hatchery environments.

3.2 Introduction

Hatchery programs are often used to support imperiled fish populations—particularly for salmonids—either via the restoration of previously extirpated populations or by supplementing declining populations. Initiation of hatchery rearing necessarily comprises a genetic bottleneck event, increasing the effects of genetic drift and decreasing heterozygosity and effective population size in subsequent hatchery-supplemented generations (Ryman and Laikre 1991; Christie et al. 2012). Overall, the effects of drift are expected to reduce genetic diversity throughout the genome due to the random sampling of individuals for breeding. More targeted reductions in genetic diversity can result from new patterns of selection imposed by hatchery rearing practices, often called domestication selection, affecting genes involved in metabolism, behavior, immunity, and

wound healing (Christie et al. 2016; Naval-Sanchez et al. 2020; Jin et al. 2020). Both genetic drift and domestication selection can reduce genetic variation at loci under selection in the natural environment, diminishing individual and population fitness after release into the wild (Araki et al. 2007; Milot et al. 2013).

Atlantic salmon (*Salmo salar*) populations are declining throughout their native range, with many landlocked (i.e., potamodromous) and anadromous populations supported by hatchery supplementation (Parrish et al. 1998; Naish et al. 2007; Lehnert et al. 2019). In North America, landlocked populations of salmon likely formed around 10,000 years ago, following the end of the Last Glacial Maximum and the initiation of isostatic rebound, or the gradual rise of land masses previously depressed by the weight of glaciers (Yesner et al. 1984; Lumme et al. 2016). In most cases, these populations became physically isolated from their anadromous progenitor populations, completing their life cycle in entirely freshwater systems. Adaptation to this new environment required shifts in a broad suite of functions, including osmoregulatory, immunological, behavioral, and metabolic traits, as has been described for other fish species with populations secondarily adapted to freshwater (Deagle et al. 2013; Kozak et al. 2014; Willoughby et al. 2018).

To identify candidate loci associated with freshwater adaptation, we examined genome-wide patterns of genetic variation for landlocked and anadromous populations of Atlantic salmon. Two landlocked populations, Lake Champlain and Sebago Lake, were compared against the anadromous LaHave River population, and outlier loci represent candidate adaptive loci with respect to the marine to freshwater transition. We used the same populations and another set of comparisons to identify loci putatively under selection in different hatchery rearing environments. Individuals from the Lake Sebago and LaHave River populations were maintained in the same hatchery facility for >1 generation prior to sampling, and therefore may have already begun to respond to selection imposed by their shared captive environment. We compared these two populations against the Lake Champlain population to determine whether distinct hatchery rearing environments may have led to detectable patterns of selection. We demonstrate that putative signals of selection are present across the three populations with respect to both life history strategy and hatchery rearing environment and we highlight strongly differentiated SNPs in genes with physiological functions relevant to freshwater adaptation and to potential hatchery-imposed selection.

3.3 Methods

3.3.1 Population histories

Lake Champlain

In Lake Champlain (VT, NY, QC; Fig. 3.1A), landlocked Atlantic salmon were extirpated in the mid-1800s and reintroduction efforts began with sustained stocking goals in the 1970s (Marsden and Langdon 2012). Despite these efforts, there has been little to no documented natural reproduction, although 123 natural-origin young-of-the-year were observed in 2016—the first in the lake since the late 1800s (Prévost et al. 2020). Individuals from potential source populations have been experimentally introduced to the lake by New York and Vermont state agencies and the U.S. Fish and Wildlife Service, including landlocked salmon sourced from Sebago Lake (ME), Lake Memephremagog (VT, QC), Little Clear Lake (NY), and West Grand Lake (ME). Of these sources, the Vermont Department of Fish and Wildlife determined that fish from Sebago Lake exhibited the strongest returns and continued introducing them to Lake Champlain (Dimond and Smitka 2005). Descendents of wild Sebago Lake Atlantic salmon therefore likely comprise a large portion of the Lake Champlain gene pool.

Sebago Lake

Landlocked Atlantic salmon in Sebago Lake (ME) comprise one of only four natural Atlantic salmon populations remaining in Maine, with 62% of the population designated as wild-born in 2013 (Pellerin and Pierce 2015). The Sebago Lake individuals used in this study are descended from fertilized eggs of natural-origin Sebago Lake salmon received by the Ontario Ministry of Natural Resources and Forestry's Harwood Fish Culture Station (FCS) in 2006 (Houde 2015). Specifically, single-pair matings were conducted in November 2011 at Harwood FCS (He et al. 2018). Based on this timeline and a generation time of approximately 4 years, we estimate that the Sebago Lake fish sequenced likely had 1-2 generations of hatchery ancestry.

LaHave River

Fertilized eggs from the anadromous LaHave River (NS) population were received by Normandale FCS from 1989 to 1995, and these eggs were sourced from a naturally reproducing population (Houde 2015). Fish were maintained at Normandale FCS from 1996 until 2006 when they were moved to Harwood FCS. Individuals sequenced by He et al. (2018) were produced in November 2011 at Harwood FCS. Because the study fish were produced 16 years after the last wild egg collection, we estimate that the LaHave fish had 3-4 generations of hatchery ancestry, including one generation at Harwood FCS.

3.3.2 Sampling and sequenced individuals

Thirty-six Lake Champlain Atlantic salmon were sampled in February 2018 at 95 days post-fertilization, and comprise four offspring from each of nine single-pair matings conducted in November 2017. Whole fry were sampled and sequenced as described in Harder et al. (2020). Briefly, mRNA was extracted using an RNeasy kit (Qiagen), separate libraries were prepared for all individuals, and each library was sequenced at the Purdue Genomics Core.

Sebago Lake and LaHave River fish were produced from crosses in fall 2011 and were sampled in July 2013 by He et al. (2018). RNA was extracted from spleen tissue of 20 subadult Atlantic salmon from each of the two populations. For each of these populations, equal concentrations of RNA from four individuals were combined into five pools (n=20 individuals per population) and sequenced on an Illumina HiSeq 2000. We obtained these reads from the NCBI Sequence Read Archive (BioProject PRJNA335629).

Following the same protocol as in Harder et al. (2020), we removed adapter sequences and clipped poor quality bases (quality score < 20) from both ends of reads for all three populations using Trimmomatic (Bolger et al. 2014). Lake Champlain cleaned reads are available via the NCBI Sequence Read Archive (BioProject PRJNA587762) and all scripts used for analyses presented in this manuscript are available at https://github.com/ChristieLab/Ssalar_3_pop_comp.

3.3.3 Lake Champlain variant calling and filtering

We aligned cleaned reads to the annotated Atlantic salmon reference genome (S. salar ICSASG_v2 assembly, NCBI accession GCA_000233375.4; Lien et al. 2016) using *HISAT2*

v2.1.0 (Kim et al. 2015) and reporting primary alignments. The BAM files of mapped reads were sorted using SAMtools v1.8 (Li et al. 2009) and single nucleotide polymorphisms (SNPs) were called using GATK v3.8.0 and v4 (see gatk.sh in the Ssalar_3_pop_comp GitHub repository for version application details) (McKenna et al. 2010). We used the *HaplotypeCaller* algorithm in Genomic Variant Call Format (GVCF) mode and called genotypes jointly with *GenotypeGVCFs* as recommended for RNA-seq data (Brouard et al. 2019). *VariantFiltration* was applied to the resulting VCF files to flag SNPs that occurred in clusters (≥3 SNPs in a 35-bp window), exhibited strand bias (FisherStrand values > 30), or had low call confidence (QualByDepth <2).

VCF files were concatenated using BCFtools (Li et al. 2009), imported into R v3.6.2 using the *vcfR* package (Knaus and Grünwald 2017), and a total of 2,112,910 SNPs were included in the final concatenated VCF file. Next, we removed SNPs that did not pass GATK filters, indels, loci with >2 alleles, and loci for which <18 samples had read depth ≥5. Individual genotypes were extracted for the remaining 246,627 SNPs.

3.3.4 Sebago Lake and LaHave River variant calling and filtering

Cleaned reads for Sebago Lake and LaHave River samples were aligned to the Atlantic salmon reference genome using the same criteria as described above. The resulting BAM files (1 file per pool; n=10) were sorted and SAMtools was used to generate a pileup file for each pool. Pileup files were converted to sync files using PoPoolation2 (Kofler et al. 2011), requiring a minimum base quality of 20. Loci were filtered such that all loci in the final sync file had read depths ≥1 in all 10 pools, leaving 271,692 SNPs for further filtering.

We read this sync file into R (R Core Team 2019) using the *poolfstat* package (Hivert et al. 2018) and applied a more stringent read depth requirement, keeping loci with read depth ≥8 in at least 4 pools per population. To prevent any single pool within a population from having an outsized effect on determining population allele frequencies, we calculated the coefficient of variation (CV) in read depth for each locus and flagged loci with a CV greater than 3 standard deviations from the mean calculated for all 10 pools. Loci with variation greater than this threshold were included in estimation of population genetic parameters but excluded from interpretation if identified as outlier SNPs.

3.3.5 Calculating allele frequencies and estimating population genetic parameters

Genotypes for Lake Champlain individuals were used to calculate population allele frequencies, allowing ≤5 samples to have a missing genotype at each locus and only retaining SNPs mapped to assembled chromosomes (i.e., excluding those mapped to scaffolds in the reference genome). For Sebago Lake and LaHave River, we estimated population allele frequencies by first calculating pooled allele frequencies (reference allele depth / total read depth at a locus) and then averaging these frequencies within each population. Because *poolfstat* randomly designates an observed allele as the reference allele, we checked all reference allele identities in the Sebago Lake/LaHave River data and corrected these assignments to match allele identities in the published reference genome and in the Lake Champlain data. We combined the calculated allele frequencies for all three populations into a single data set, only retaining SNPs included in both final sets of SNPs (i.e., Lake Champlain and Sebago Lake/LaHave River).

To visualize relationships among the populations, we first replicated the structure of the Sebago Lake/LaHave River data for the non-pooled Lake Champlain population by randomly assigning individuals to one of nine pools (replicating the four individuals per pool) and calculating pooled allele frequencies using the same approach as for Sebago Lake/LaHave River (i.e., reference allele depth for the pool/total read depth at a locus). Using allele frequencies for the nine bioinformatically pooled Lake Champlain samples and the ten physically pooled Sebago Lake/LaHave River samples, we applied the *dist* function in R to calculate a Euclidean distance matrix. We supplied this matrix to the *hclust* function to perform hierarchical clustering using the unweighted pair group method with arithmetic mean (UPGMA) approach and plotted the resulting dendrogram.

To estimate genetic diversity within each population, we calculated pooled heterozygosity (H_P) at each locus as $H_P = 2 \times \Sigma n_{MAJ} \times \Sigma n_{MIN}/(\Sigma n_{MAJ} + \Sigma n_{MIN})^2$, where Σn_{MAJ} and Σn_{MIN} are the sums of the major and minor allele counts at a locus (Rubin et al. 2010). For pooled genomic data, H_P is typically calculated in windows, summing major and minor allele counts across each window. However, because RNA-seq reads produce uneven read depths across the genome due to variation in transcript expression levels, we calculated H_P for each locus separately. We then averaged locus-specific H_P over each chromosome and across the whole genome. To statistically test for differences in H_P between populations, we used a randomization test. Specifically, we pooled all chromosome-specific H_P values for the two populations being compared and randomly

drew 29 values (i.e., the number of nuclear chromosomes) and calculated the mean value. We repeated this process 10,000 times per population comparison and compared the distribution of sample mean values against the known mean chromosome-specific H_P values for the two populations being compared. Because the Lake Champlain individuals comprise nine sets of four full siblings, we evaluated the effect of relatedness among individuals on H_P by sampling one individual per family (n=9), calculating whole-genome H_P, repeating this process 1,000 times, and calculating the mean and standard deviation of H_P across all subsamples.

To identify SNPs differentiated among populations, we estimated pooled F_{ST} from population allele frequencies, substituting number of individuals per population for read depth at each locus because read depth always exceeded number of individuals sampled (Willoughby et al. 2018; https://github.com/jacobtennessen/GOPOPS/blob/master/FstFromPooledFreqs.pl). To estimate genome-wide differentiation among populations, we averaged pairwise locus F_{ST} values for each population comparison. We also examined the effects of relatedness among Lake Champlain individuals on the estimation of pooled F_{ST} by again randomly sampling one individual per family (n=9) and calculating allele frequencies for this subsample. We repeated this process 10,000 times, each time calculating the absolute difference between F_{ST} values for the full and subsampled datasets. Finally, to examine any potential effects of relatedness, we determined the mean difference in F_{ST} between the full and subsampled data sets at each locus.

3.3.6 Identifying SNPs putatively under selection

To identify SNPs putatively under selection across the three populations, we first used a *k*-nearest neighbor (kNN) approach. Implemented in the R *PopGenome* package (Pfeifer et al. 2014) and relying on ELKI software (Schubert and Zimek 2019), this method uses pairwise F_{ST} values from all population comparisons to define the locations of all loci in multidimensional space (Pfeifer et al. 2020). Loci in putative neutrally evolving genomic regions cluster together in this space and loci identified as global outliers represent candidate SNPs subject to selection. We used the *kNN_tau_window* function to determine the appropriate value for *k* and performed the kNN scan with the *kNNCallElkiSINGLE* function. We applied a 0.01 quantile to identify SNPs putatively under selection. We further filtered this list of SNPs by applying a *Z*-transformation to pairwise F_{ST} values and only retaining loci for which F_{ST} was ≥3 standard deviations from the mean (i.e., $Z(F_{ST}) \geq 3$) in at least one pairwise comparison between two of the three populations.

Outlier SNPs were annotated and features were identified using the reference genome and SnpEff v4.3 (Cingolani et al. 2012). We retained all error-free annotations comprising unique SNP:gene combinations, excluding SNPs mapped to pseudogenes. We categorized annotations as coding (missense variants), regulatory (synonymous, intron, and untranslated region (UTR) variants), or distant (within 5kb of a gene, but not located within a transcribed sequence). We categorized synonymous SNPs as 'regulatory' rather than 'coding' because while synonymous SNPs do not affect the amino acid sequence of a protein, they may have consequences for translation and post-translational processes (reviewed in Shabalina et al. 2013).

Finally, we partitioned the outlier SNPs by considering patterns of differentiation across multiple population comparisons. Sebago Lake and LaHave River fish had been reared in the same hatchery facility for >1 generation prior to sampling, and shared differences between these two populations and Lake Champlain may be due to differences in hatchery practices. To identify outlier SNPs that may be associated with the effects of hatchery rearing practices, we filtered loci with $Z(F_{ST}) \geq 3$ in both the Lake Champlain-Sebago Lake and Lake Champlain-LaHave River comparisons. We used a similar approach to identify SNPs putatively associated with differences in life history strategy (i.e., anadromous LaHave River population vs. landlocked Lake Champlain and Sebago Lake populations) by requiring $Z(F_{ST}) \geq 3$ in both the Lake Champlain-LaHave River and Sebago Lake-LaHave River comparisons.

3.4 Results

3.4.1 Sequence processing and variant filtering

Our final sets of cleaned reads comprised an average of 39,348,214 read pairs per Lake Champlain individual and an average of 20,377,740 read pairs per pool for the Sebago Lake/LaHave River data sets. After all SNP filtering steps, we retained 246,627 SNPs for Lake Champlain and 135,839 SNPs for Sebago Lake/LaHave River. Of the filtered SNP set for Sebago Lake and LaHave River, 1,727 SNPs were flagged for high among-pool variation in read depth (with CV > 3 SD above the mean value; mean=0.355, SD=0.119; Fig. B.1). A total of 25,282 SNPs were covered in both the Lake Champlain and Sebago Lake/LaHave River data sets and these loci were used for all downstream analyses. Differences between Lake Champlain and Sebago Lake/LaHave River samples with respect to developmental stage and tissue identity likely resulted

in distinct transcriptional profiles between the two data sets. While these differences probably affect the identify and abundance of expressed transcripts, impacts on allele frequency estimation are expected to be minimal.

3.4.2 Relationships among populations and genetic diversity

The dendrogram constructed using allele frequencies for all three populations illustrates that Lake Champlain and Sebago Lake are more closely related than either population is to LaHave River (Fig. 3.1B). Estimates of genome-wide genetic differentiation are consistent with this pattern, with the lowest pairwise F_{ST} estimate for the Lake Champlain-Sebago Lake comparison (F_{ST}=0.03) and higher estimates for Lake Champlain-LaHave River (F_{ST}=0.13) and Sebago Lake-LaHave River (F_{ST}=0.12). Genome-wide pooled heterozygosity (H_P) was highest in the Lake Champlain population (H_P=0.268), followed by Sebago Lake (H_P=0.242) and LaHave River (H_P=0.235) (Fig. 3.1C). As determined through a randomization test conducted using chromosome-specific H_P values, H_P was significantly higher in Lake Champlain than in both Sebago Lake and LaHave River (p<0.0001), which did not significantly differ from one another (p>0.05) (Fig. B.2).

Relatedness among Lake Champlain individuals had minimal impact on estimation of H_P, with a mean H_P value for family-sampling permutations of 0.260 (SD=0.001) (Fig. B.3). We also found relatedness to have little effect on estimation of F_{ST}. For the Lake Champlain-Sebago Lake population comparison, the mean difference between F_{ST} estimated for nine and 36 individuals was ≤0.05 for 79% of loci (median value = 0.005) and was ≤0.05 for 49% of loci for the Lake Champlain-LaHave River comparison (median value = 0.051) (Fig. B.4).

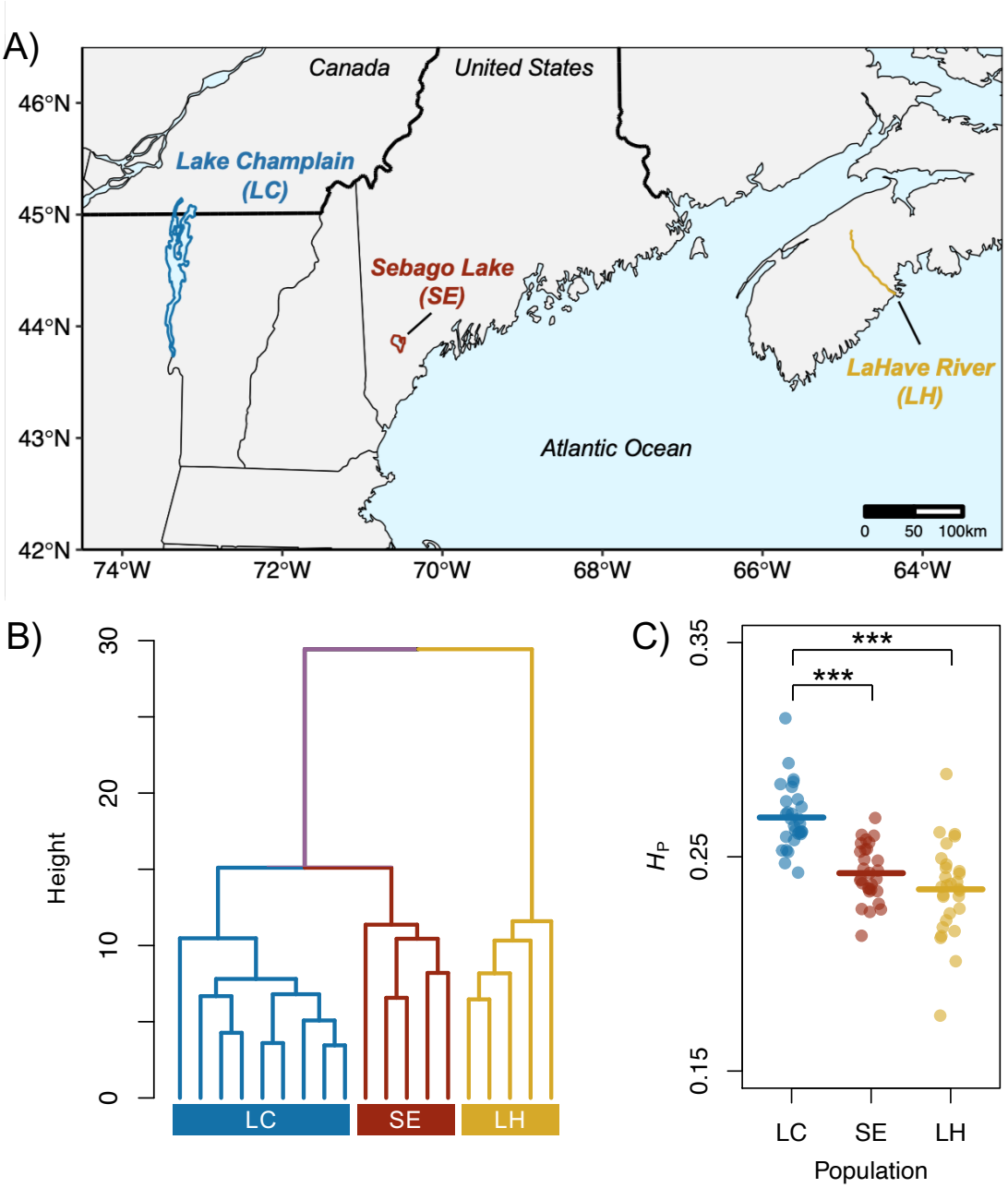


Figure 3.1. (A) Locations of sampled populations. (B) Dendrogram visualizing relationships among the three populations; relationships were determined based on a Euclidian distance matrix calculated from pooled allele frequencies. Lake Champlain and Sebago Lake individuals are more closely related to each other than either population is to the LaHave River. (C) Pooled heterozygosity (H_P) estimated for each population. Horizontal lines correspond to genome-wide estimates of H_P and points correspond to chromosome-specific H_P estimates. Lake Champlain H_P is significantly higher than H_P in Sebago Lake and LaHave River, as determined via randomization tests (***p<0.0001), with H_P not differing between Sebago Lake and LaHave River (p>0.05).

3.4.3 Outlier SNPs putatively under selection

In total, the kNN approach identified 252 SNPs putatively under selection and there were 1,114 SNPs with $Z(F_{ST}) \geq 3$ in at least one population comparison. The intersection of these two lists of SNPs comprises 228 SNPs and represents the final set of outlier loci. Of these SNPs, four exhibited variation in read depth across the Sebago Lake/LaHave River pools greater than 3 SD above the mean CV and were excluded from downstream considerations (Fig. B.1A). The Lake Champlain-Sebago Lake comparison included the fewest outlier SNPs (40 SNPs; Fig. 3.2A), with similar numbers of outlier SNPs for the Sebago Lake-LaHave River and Lake Champlain-LaHave River comparisons (159 and 178 SNPs, respectively; Fig. 3.2B,C).

3.4.4 Outlier SNP annotations

Of the 228 outlier SNPs, 214 were successfully annotated to functional protein-coding features in the reference genome. Some SNPs were annotated to more than one feature (for example, a SNP can occur in the 3' UTR of a gene and also within 5kb of another gene), resulting in a total of 329 unique annotations (Fig. 3.3A). Two genes contained multiple missense outlier variants; one of these genes is uncharacterized in the reference annotation and the other is a transcriptional repressor known as zinc finger and BTB domain-containing protein 18-like (LOC106610379; Fig. 3.4).

We identified 23 SNPs as outliers in both the Lake Champlain-Sebago Lake and Lake Champlain-LaHave River comparisons, and these loci may be associated with differences in hatchery rearing environments (i.e., putative hatchery-associated SNPs). Error-free annotations were produced for 22/23 of hatchery-associated SNPs, with 17 SNPs annotated to transcribed sequences (Fig. 3.3B) and three of these SNPs were identified as missense variants (Fig. 3.3B). We also identified 120 SNPs that were outliers in both the Lake Champlain-LaHave River and Sebago Lake-LaHave River comparisons and that may be associated with adaptation to freshwater in Lake Champlain and Sebago Lake (i.e., potential life history-associated SNPs). For life history-associated SNPs, 111/120 were annotated. Twelve of these variants were classified as missense (Fig. 3.3C), including one SNP in a gene involved in osmoregulation (Fig. 3.5).

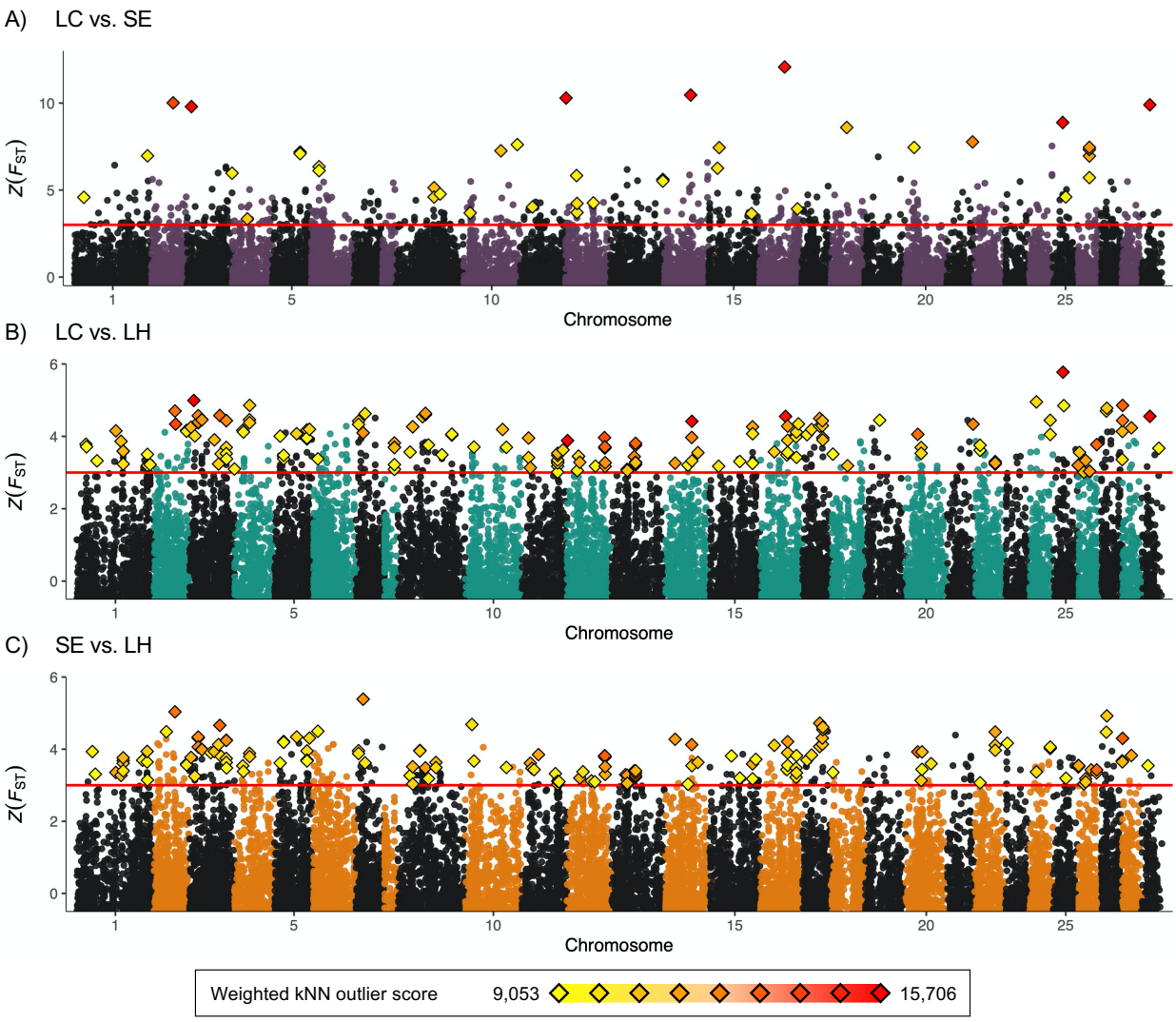


Figure 3.2. *Z*-transformed F_{ST} estimates for final set of 25,282 SNPs for each pairwise population comparison. Outlier SNPs are represented by diamond shapes with weighted kNN outlier scores indicated by color. SNPs were designated as outliers if they had kNN outlier scores in the top 1% of all values and $Z(F_{ST}) \geq 3$ in at least one population comparison. Alternating point colors across *x*-axis correspond to nuclear chromosomes. The Lake Champlain-Sebago Lake comparison includes the fewest outlier loci (n=40), with comparable numbers of outlier loci for the Lake Champlain-LaHave River (n=178) and Sebago Lake-LaHave River comparisons (n=159).

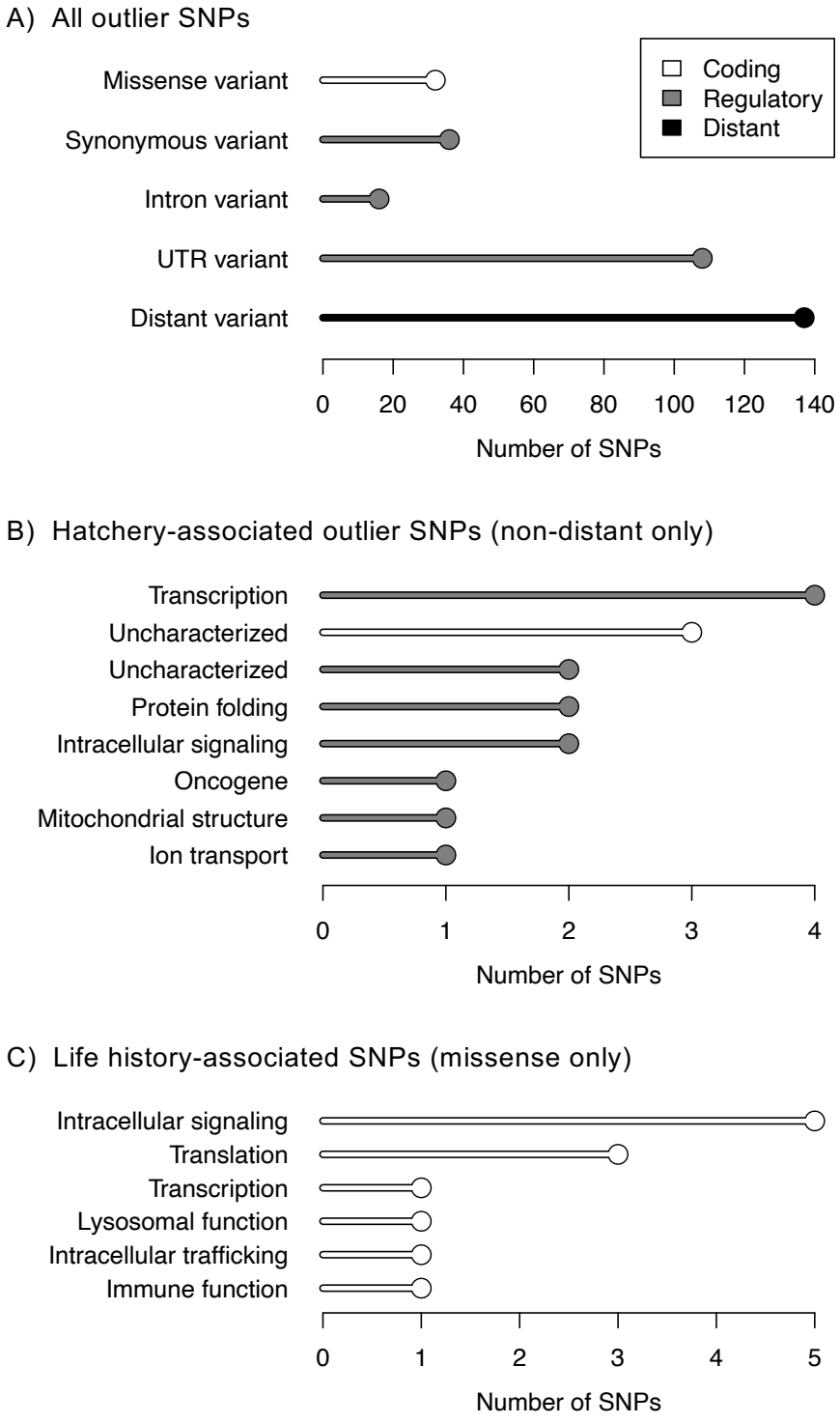


Figure 3.3. Summary of annotations for: (A) 214 outlier SNPs (i.e., loci within top 1% of kNN scores and with $Z(F_{ST}) \geq 3$ in at least one population comparison); (B) 17 hatchery-associated outlier SNPs located within transcribed sequences (5 additional outlier SNPs were classified as distant loci relative to all genes annotated in the reference genome); and (C) 12 life history-associated missense SNPs (in total, 111 outlier SNPs were successfully annotated for this set of comparisons).

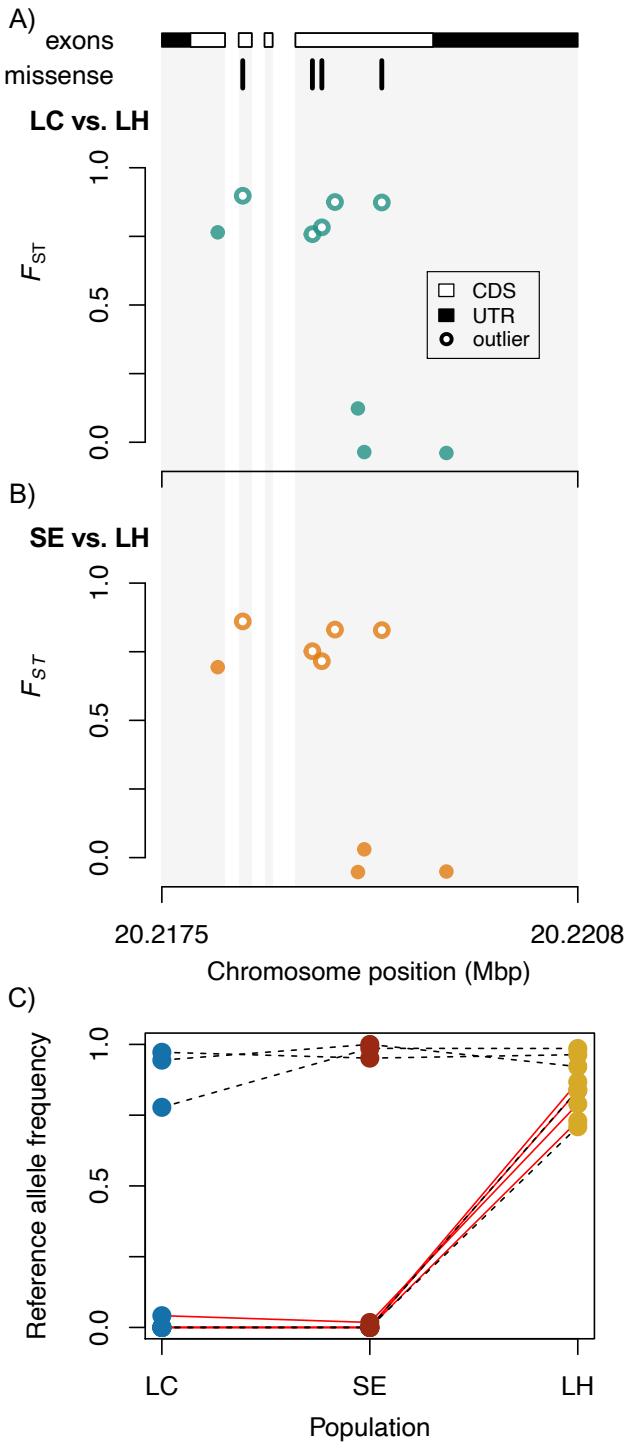


Figure 3.4. Diagram of *ZBTB18* gene, with gene structure and positions of four missense outlier variants indicated at the top of panel A. F_{ST} values for all SNPs identified within the gene are provided for the Lake Champlain-LaHave River (A) and Sebago Lake-LaHave River (B) population comparisons. Outlier SNPs are indicated by empty circles and include four missense SNPs and one synonymous SNP. (C) Differences in population allele frequencies for all SNPs within the gene, with missense variants represented by red lines and all other SNPs indicated by grey dashed lines. Note that outlier SNPs in panels (A) and (B) were identified as putatively under selection using the kNN approach but have $Z(F_{ST}) \geq 3$ only in the Lake Champlain-LaHave River comparison, and not in the Sebago Lake-LaHave River comparison (Sebago Lake-LaHave River $Z(F_{ST})$ range: 2.09-2.67).

3.5 Discussion

Of the three populations compared in this study, we found the closest genetic relationship between Lake Champlain and Sebago Lake individuals. This result likely reflects the inclusion of Sebago Lake salmon in the Lake Champlain breeding program. This history of shared ancestry also likely explains the relatively low number of outlier SNPs shared between Lake Champlain and Sebago Lake, but this trend may also be influenced by environmental similarities between the two populations. Both Lake Champlain and Sebago Lake Atlantic salmon are landlocked, returning to lake tributaries to spawn and both lakes also contain prey species with high concentrations of thiaminase in their tissues (Dimond and Smitka 2005). Thiaminase is an enzyme that degrades thiamine (vitamin B_1) and consumption of this enzyme has been linked to thiamine deficiency in salmonids (reviewed in Harder et al. 2018). Thiamine deficiency is prevalent in Lake Champlain salmon and requires treatment with supplemental thiamine at the egg stage to mitigate otherwise high rates of deficiency-induced juvenile mortality (Ladago et al. 2020). In Sebago Lake, however, thiamine deficiency has not been observed in salmon despite the presence of dietary thiaminase (Dimond and Smitka 2005). Relative to Sebago Lake salmon, LaHave River fish (with a natural diet lacking thiaminase) have been shown to experience greater decreases in liver tissue thiamine concentrations when fed an experimental diet containing thiaminase, suggesting that the Sebago Lake population may have genetically adapted to dietary thiaminase (Houde et al. 2015). The genetic variation associated with this putative adaptation may be advantageous in Lake Champlain, where hatchery thiamine treatments prevent any potential natural selection on such a trait (Marsden and Langdon 2012). Indeed, some evidence suggests that the genetic variation required for adaptation to low thiamine availability is already present in the Lake Champlain population (Harder et al. 2020).

In addition to receiving genetic input from Sebago Lake individuals, the Lake Champlain population has also been supplemented by other source populations (Dimond and Smitka 2005). Although salmon originating from Sebago Lake were deemed most successful in Lake Champlain by management agencies, the high genetic diversity (i.e., H_P) in Lake Champlain relative to both Sebago Lake and the LaHave River suggests that fish sourced from other populations may have also been reproductively successful in the Lake Champlain system and genetically contributed to the contemporary population. Alternatively, elevated H_P in Lake Champlain may be due to differences in breeding cross design or variance in family size between Lake Champlain hatcheries

and hatcheries that have propagated Sebago Lake and LaHave River individuals (i.e., the Normandale and Harwood Fish Culture Stations). Testing these hypotheses would require a more extensive genetic survey of all populations and hatcheries that have served as sources for Lake Champlain reintroduction efforts.

Of the few outlier SNPs differentiated between Lake Champlain and Sebago Lake, over half are also outliers in the Lake Champlain-LaHave River comparison (23/40 Lake Champlain-Sebago Lake outlier SNPs). From 2006 until sampling in 2011, the Sebago Lake and LaHave River fish were reared in the same hatchery (Houde 2015), and shared differences between these two populations and Lake Champlain may reflect parallel patterns of genetic drift or hatchery-imposed selection. Responses to selection in hatcheries can occur over short timescales (e.g., a single generation; Christie et al. 2016) and may also vary among specific rearing strategies (Berejikian et al. 2000; Venney et al. 2020). We identified putative hatchery-associated SNPs annotated to genes with roles in transcription, protein folding, and intracellular signaling, among other functions (Fig. 3.3B). All three outlier SNPs encoding missense mutations are localized to uncharacterized genes in the reference annotation, but two SNPs in regulatory regions are annotated to genes important for circadian rhythm regulation across eukaryotes (probable ATP-dependent RNA helicase DDX5; LOC106588346) and epithelial cell adhesion (chloride intracellular channel 4; *CLIC4*) (Padmanabhan et al. 2012; Argenzio et al. 2014; Emerson et al. 2015). Specifically, *CLIC4* knockdown and suppression experiments have demonstrated its critical roles in skin and corneal wound healing (Padmakumar et al. 2012). Differentiation at these genes is consistent with previous findings that the earliest stages of domestication in salmon can lead to new patterns of gene expression in circadian and wound healing pathways (Christie et al. 2016; Jin et al. 2020).

We used a second set of comparisons to identify loci that may be associated with life history by comparing the landlocked Lake Champlain and Sebago Lake populations with the anadromous LaHave River population. For putative life history-associated SNPs located in regulatory regions, 86 SNPs were annotated to untranslated portions of 73 unique gene sequences. Of these genes, one was identified as differentially expressed in response to supplemental thiamine treatment and with respect to family-level survival in the face of thiamine deficiency in Lake Champlain Atlantic salmon by Harder et al. (2020). This gene, suppressor of cytokine signaling 2-like (LOC106576177; *SOCS2*), is a negative regulator of growth hormone signaling and may also be important in adaptation to salinity (Greenhalgh et al. 2005; Komoroske et al. 2016; Dalongeville

2018). In anadromous Arctic charr (*Salvelinus alpinus*), *SOCS2* is strongly upregulated in response to fasting, indicating a role for this gene in maintenance of energy homeostasis in another migratory fish species (Jørgensen et al. 2013). The expression patterns of *SOCS2* associated with thiamine deficiency in Lake Champlain coupled with strong differentiation at a 3' UTR locus between the two lake populations and the LaHave River suggest that different diets and life histories in the two types of environments may have lead to adaptive differences in gene expression regulation. Two additional outlier SNPs in the 3' UTR of the glutathione peroxidase 1-like gene (LOC106566059) may also be related to dietary differences between environment types given that thiamine deficiency is often accompanied by oxidative stress, perhaps due to thiamine's role as an antioxidant (Gibson and Zhang 2002; Bettendorff 2013). When thiamine concentrations are low, enzymes such as glutathione peroxidases can increase in expression in response to the demand for antioxidants (Lundström et al. 1999; Harder et al. 2020). The limited number of SNPs in this study precludes definitive assessment of genetic adaptation to dietary thiaminase, but greater marker density (including SNPs in intergenic regions that may be linked to loci under selection) could be used to investigate genetic variants associated with adaptation.

Twelve life-history associated SNPs were missense variants annotated to genes with roles in intracellular signaling, translation, and immune function (Fig. 3.3C), including one MHC class I gene (major histocompatibility complex class I-related gene protein-like; LOC106588402). Of the genes involved in intracellular signaling, one encodes adenylate cyclase type 6-like (LOC106571946)—a protein that promotes increased water reabsorption from urine in kidney collecting ducts (Konno et al. 2010; Jung and Kwon 2016; Olesen and Fenton 2017) (Fig. 3.5) Specifically, activated adenylate cyclase type 6 initiates a signaling cascade that leads to insertion of water channel proteins (aquaporin-2) into collecting duct cell membranes. This allows for increased water reabsorption from urine, a process critical for maintenance of tissue water homeostasis (Jung and Kwon 2016). Variation in the coding sequence of the adenylate cyclase type 6-like gene between the two landlocked populations and the LaHave River population is likely driven by different osmotic demands in freshwater and marine environments, as has been described for other osmoregulatory genes in fish populations secondarily adapted to entirely freshwater life cycles (Deagle et al. 2013; Kozak et al. 2014; Willoughby et al. 2018).

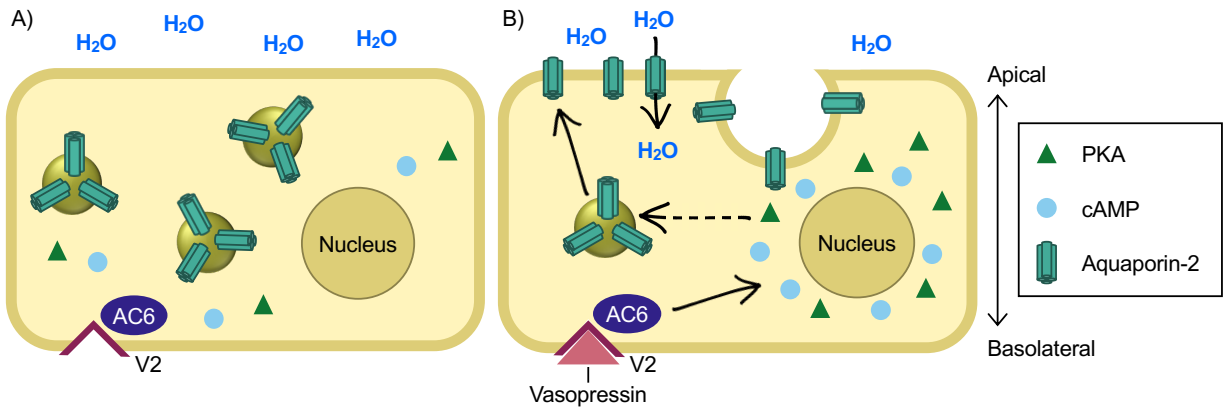


Figure 3.5. Diagram illustrating the role of adenylate cyclase 6 (AC6) in water reabsorption by kidney collecting duct cells. (A) Cellular resting state, with aquaporin-2 localized to intracellular vesicles and low levels of cyclic adenosine monophosphate (cAMP) and protein kinase A (PKA) in the cytosol. (B) Vasopressin receptor 2 (V2) is stimulated by vasopressin, leading to activation of AC6, and elevation of cAMP levels and activation of PKA (predominantly near the nucleus). Activated PKA interacts with PKA-anchoring proteins located on intracellular vesicles, leading these vesicles to fuse with the apical cell membrane and resulting in translocation of aquaporin-2 to the cell surface. Insertion of aquaporin-2 facilitates increased osmotic permeability of the membrane. Adapted from Olesen and Fenton 2017.

When considering all outlier SNPs (i.e., with $Z(F_{ST})$ ≥3 in at least one population comparison), only two genes contained multiple missense variants. Only one of these genes is characterized in the reference annotation: zinc finger and BTB domain-containing protein 18-like (LOC106610379; *ZBTB18*). Four missense SNPs located in exons of *ZBTB18* are $Z(F_{ST})$ outliers in the Lake Champlain-LaHave River comparison (Fig. 3.4A,B). Allele frequency shifts were also strong between Sebago Lake and LaHave River at these four missense SNPs (mean F_{ST}=0.79), but they were not identified as $Z(F_{ST})$ outliers (Fig. 3.4B,C). *ZBTB18* is a transcriptional repressor involved in a wide variety of functions including brain and muscle development (Yokoyama et al. 2009; Okado 2019) and is also differentially expressed in response to hyperosmotic conditions in cell culture (Boyd et al. 2005; Ni et al. 2017). Differentiation at multiple SNPs within this gene between landlocked and anadromous populations may be related to differences in growth and development (e.g., age and size at maturity; Riley and Power 1987; Hutchings et al. 2019) or to differences in salinity between the two environments.

By comparing landlocked and anadromous Atlantic salmon populations, we identified outlier SNPs in genes related to immune, osmoregulatory, and metabolic functions, consistent with previous investigations of adaptive differences between freshwater and marine fish populations. We also noted outlier SNPs in genes that may be associated with the presence of thiaminase in the

diets of Lake Champlain and Sebago Lake salmon. In addition to SNPs that may be subject to selection in the natural environment, we also identified outlier loci putatively associated with hatchery rearing conditions by comparing Lake Champlain against Sebago Lake and LaHaveRiver. These SNPs were located in genes related to functions such as regulation of circadian rhythm and wound healing, similar to findings of other genetic studies examining domestication selection. That the Sebago Lake population was used as a source population for Lake Champlain reintroductions strengthens the supposition that outlier loci observed between these two, genetically similar populations and between Lake Champlain and LaHave River may be associated with hatchery practices. Together, these two sets of results highlight putative responses to both natural selection imposed by major environmental changes and artificial selection levied by differing hatchery environments.

3.6 References

Araki, H., Cooper, B., and Blouin, M.S. 2007. Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. Science 318(5847): 100–103. doi:10.1126/science.1145621.

Argenzio, E., Margadant, C., Leyton-Puig, D., Janssen, H., Jalink, K., Sonnenberg, A., and Moolenaar, W.H. 2014. CLIC4 regulates cell adhesion and 1 integrin trafficking. Journal of Cell Science 127(24): 5189–5203. doi:10.1242/jcs.150623.

Berejikian, B.A., Tezak, E.P., Flagg, T.A., LaRae, A.L., Kummerow, E., and Mahnken, C.V. 2000. Social dominance, growth, and habitat use of age-0 steelhead (*Oncorhynchus mykiss*) grown in enriched and conventional hatchery rearing environments. Can. J. Fish. Aquat. Sci. 57(3): 628–636. doi:10.1139/f99-288.

Bettendorff, L. 2013. Thiamine. In Handbook of Vitamins, Fifth Edition. CRC Press. pp. 267–324.

Bolger, A.M., Lohse, M., and Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30(15): 2114–2120. doi:10.1093/bioinformatics/btu170.

Boyd, L.M., Richardson, W.J., Chen, J., Kraus, V.B., Tewari, A., and Setton, L.A. 2005. Osmolarity regulates gene expression in intervertebral disc cells determined by gene array and real-time quantitative RT-PCR. Ann Biomed Eng 33(8): 1071–1077. doi:10.1007/s10439-005-5775-y.

Brouard, J.-S., Schenkel, F., Marete, A., and Bissonnette, N. 2019. The GATK joint genotyping workflow is appropriate for calling variants in RNA-seq experiments. J Animal Sci Biotechnol 10(1): 44. doi:10.1186/s40104-019-0359-0.

Christie, M.R., Marine, M.L., Fox, S.E., French, R.A., and Blouin, M.S. 2016. A single generation of domestication heritably alters the expression of hundreds of genes. Nature Communications 7: 10676. doi:10.1038/ncomms10676.

Christie, M.R., Marine, M.L., French, R.A., Waples, R.S., and Blouin, M.S. 2012. Effective size of a wild salmonid population is greatly reduced by hatchery supplementation. Heredity 109(4): 254–260. doi:10.1038/hdy.2012.39.

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w^{1118}; *iso*-2; *iso*-3. Fly 6(2): 80–92. doi:10.4161/fly.19695.

Dalongeville, A. 2018. Combining six genome scan methods to detect candidate genes to salinity in the Mediterranean striped red mullet (*Mullus surmuletus*). BMC Genomics 19: 217.

Deagle, B.E., Jones, F.C., Absher, D.M., Kingsley, D.M., and Reimchen, T.E. 2013. Phylogeography and adaptation genetics of stickleback from the Haida Gwaii archipelago revealed using genome-wide single nucleotide polymorphism genotyping. Mol Ecol 22(7): 1917–1932. doi:10.1111/mec.12215.

Dimond, P., and Smitka, J. 2005. Evaluation of selected strains of Atlantic salmon as potential candidates for the restoration of Lake Ontario. Trout Unlimited Canada Technical Report ON-012. Guelph: Trout Unlimited Canada. Available from http://tucanada.ibcnetwork.com/files/1/ON-012_AtlanticSalmonStrains.pdf [accessed 8 December 2015].

Emerson, J.M., Bartholomai, B.M., Ringelberg, C.S., Baker, S.E., Loros, J.J., and Dunlap, J.C. 2015. *period*-1 encodes an ATP-dependent RNA helicase that influences nutritional compensation of the *Neurospora* circadian clock. Proc Natl Acad Sci USA 112(51): 15707–15712. doi:10.1073/pnas.1521918112.

Gibson, G.E., and Zhang, H. 2002. Interactions of oxidative stress with thiamine homeostasis promote neurodegeneration. Neurochemistry international 40(6): 493–504.

Greenhalgh, C.J., Rico-Bautista, E., Lorentzon, M., Thaus, A.L., Morgan, P.O., Willson, T.A., Zervoudakis, P., Metcalf, D., Street, I., Nicola, N.A., Nash, A.D., Fabri, L.J., Norstedt, G., Ohlsson, C., Flores-Morales, A., Alexander, W.S., and Hilton, D.J. 2005. SOCS2 negatively regulates growth hormone action *in vitro* and *in vivo*. J. Clin. Invest. 115(2): 397–406. doi:10.1172/JCI200522710.

Harder, A.M., Ardren, W.R., Evans, A.N., Futia, M.H., Kraft, C.E., Marsden, J.E., Richter, C.A., Rinchard, J., Tillitt, D.E., and Christie, M.R. 2018. Thiamine deficiency in fishes: causes, consequences, and potential solutions. Reviews in Fish Biology and Fisheries 28(4): 865–886. doi:10.1007/s11160-018-9538-x.

Harder, A.M., Willoughby, J.R., Ardren, W.R., and Christie, M.R. 2020. Among-family variation in survival and gene expression uncovers adaptive genetic variation in a threatened fish. Mol Ecol 29: 1035–1049. doi:10.1111/mec.15334.

He, X., Houde, A.L.S., Neff, B.D., and Heath, D.D. 2018. Transcriptome response of Atlantic salmon (*Salmo salar*) to competition with ecologically similar non-native species. Ecology and Evolution 8(3): 1769–1777. doi:10.1002/ece3.3798.

Hivert, V., Leblois, R., Petit, E.J., Gautier, M., and Vitalis, R. 2018. Measuring genetic differentiation from Pool-seq data. Genetics 210: 315–330.

Houde, A.L.S. 2015. Restoration of native biodiversity in altered environments: Reintroduction of Atlantic salmon into Lake Ontario. University of Western Ontario, London, Ontario, Canada.

Houde, A.L.S., Saez, P.J., Wilson, C.C., Bureau, D.P., and Neff, B.D. 2015. Effects of feeding high dietary thiaminase to sub-adult Atlantic salmon from three populations. Journal of Great Lakes Research 41: 898–906.

Hutchings, J.A., Ardren, W.R., Barlaup, B.T., Bergman, E., Clarke, K.D., Greenberg, L.A., Lake, C., Piironen, J., Sirois, P., Sundt-Hansen, L.E., and Fraser, D.J. 2019. Life-history variability and conservation status of landlocked Atlantic salmon: an overview. Can. J. Fish. Aquat. Sci. 76(10): 1697–1708. doi:10.1139/cjfas-2018-0413.

Jin, Y., Olsen, R.E., Harvey, T.N., Østensen, M., Li, K., Santi, N., Vadstein, O., Bones, A.M., Vik, J.O., Sandve, S.R., and Olsen, Y. 2020. Comparative transcriptomics reveals domestication-associated features of Atlantic salmon lipid metabolism. Mol Ecol 29(10): 1860–1872. doi:10.1111/mec.15446.

Jørgensen, E.H., Martinsen, M., Strøm, V., Hansen, K.E.R., Ravuri, C.S., Gong, N., and Jobling, M. 2013. Long-term fasting in the anadromous Arctic charr is associated with downregulation of metabolic enzyme activity and upregulation of leptin A1 and SOCS expression in the liver. Journal of Experimental Biology 216(17): 3222–3230. doi:10.1242/jeb.088344.

Jung, H.J., and Kwon, T.-H. 2016. Molecular mechanisms regulating aquaporin-2 in kidney collecting duct. American Journal of Physiology-Renal Physiology 311(6): F1318–F1328. doi:10.1152/ajprenal.00485.2016.

Kim, D., Langmead, B., and Salzberg, S.L. 2015. HISAT: a fast spliced aligner with low memory requirements. Nat Methods 12(4): 357–360. doi:10.1038/nmeth.3317.

Knaus, B.J., and Grünwald, N.J. 2017. vcfR: a package to manipulate and visualize variant call format data in R. Mol Ecol Resour 17(1): 44–53. doi:10.1111/1755-0998.12549.

Kofler, R., Pandey, R.V., and Schlotterer, C. 2011. PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics 27(24): 3435–3436. doi:10.1093/bioinformatics/btr589.

Komoroske, L.M., Jeffries, K.M., Connon, R.E., Dexter, J., Hasenbein, M., Verhille, C., and Fangue, N.A. 2016. Sublethal salinity stress contributes to habitat limitation in an endangered estuarine fish. Evol Appl 9(8): 963–981. doi:10.1111/eva.12385.

Konno, N., Kurosawa, M., Kaiya, H., Miyazato, M., Matsuda, K., and Uchiyama, M. 2010. Molecular cloning and characterization of V2-type receptor in two ray-finned fish, gray bichir, *Polypterus senegalus* and medaka, *Oryzias latipes*. Peptides 31(7): 1273–1279. doi:10.1016/j.peptides.2010.04.014.

Kozak, G.M., Brennan, R.S., Berdan, E.L., Fuller, R.C., and Whitehead, A. 2014. Functional and population genomic divergence within and between two species of killifish adapted to different osmotic niches. Evolution 68(1): 63–80. doi:10.1111/evo.12265.

Ladago, B.J., Futia, M.H., Ardren, W.R., Honeyfield, D.C., Kelsey, K.P., Kozel, C.L., Riley, S.C., Rinchard, J., Tillitt, D.E., Zajicek, J.L., and Marsden, J.E. 2020. Thiamine concentrations in lake trout and Atlantic salmon eggs during 14 years following the invasion of alewife in Lake Champlain. Journal of Great Lakes Research: S0380133020301404. doi:10.1016/j.jglr.2020.06.018.

Lehnert, S.J., Kess, T., Bentzen, P., Kent, M.P., Lien, S., Gilbey, J., Clément, M., Jeffery, N.W., Waples, R.S., and Bradbury, I.R. 2019. Genomic signatures and correlates of widespread population declines in salmon. Nat Commun 10(1): 2996. doi:10.1038/s41467-019-10972-w.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25(16): 2078–2079. doi:10.1093/bioinformatics/btp352.

Lien, S., Koop, B.F., Sandve, S.R., Miller, J.R., Kent, M.P., Nome, T., Hvidsten, T.R., Leong, J.S., Minkley, D.R., Zimin, A., Grammes, F., Grove, H., Gjuvsland, A., Walenz, B., Hermansen, R.A., von Schalburg, K., Rondeau, E.B., Di Genova, A., Samy, J.K.A., Olav Vik, J., Vigeland, M.D., Caler, L., Grimholt, U., Jentoft, S., Inge Våge, D., de Jong, P., Moen, T., Baranski, M., Palti, Y., Smith, D.R., Yorke, J.A., Nederbragt, A.J., Tooming-Klunderud, A., Jakobsen, K.S., Jiang, X., Fan, D., Hu, Y., Liberles, D.A., Vidal, R., Iturra, P., Jones, S.J.M., Jonassen, I., Maass, A., Omholt, S.W., and Davidson, W.S. 2016. The Atlantic salmon genome provides insights into rediploidization. Nature 533: 200–205. doi:10.1038/nature17164.

Lumme, J., Ozerov, M.Y., Veselov, A.E., and Primmer, C.R. 2016. The formation of landlocked populations of Atlantic salmon. In Evolutionary biology of the Atlantic salmon. Edited by T. Vladić and E. Petersson. CRC Press, New York. pp. 26–43.

Lundström, J., Carney, B., Amcoff, P., Pettersson, A., Börjeson, H., Förlin, L., and Norrgren, L. 1999. Antioxidant systems, detoxifying enzymes and thiamine levels in Baltic salmon (*Salmo salar*) that develop M74. Ambio 28: 24–29.

Marsden, J.E., and Langdon, R.W. 2012. The history and future of Lake Champlain's fishes and fisheries. Journal of Great Lakes Research 38: 19–34. doi:10.1016/j.jglr.2011.09.007.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research 20(9): 1297–1303. doi:10.1101/gr.107524.110.

Milot, E., Perrier, C., Papillon, L., Dodson, J.J., and Bernatchez, L. 2013. Reduced fitness of Atlantic salmon released in the wild after one generation of captive breeding. Evol Appl 6(3): 472–485. doi:10.1111/eva.12028.

Naish, K.A., Taylor, J.E., Levin, P.S., Quinn, T.P., Winton, J.R., Huppert, D., and Hilborn, R. 2007. An evaluation of the effects of conservation and fishery enhancement hatcheries on wild populations of salmon. In Advances in Marine Biology. Elsevier. pp. 61–194. doi:10.1016/S0065-2881(07)53002-6.

Naval-Sanchez, M., McWilliam, S., Evans, B., Yáñez, J.M., Houston, R.D., and Kijas, J.W. 2020. Changed patterns of genomic variation following recent domestication: selection sweeps in farmed Atlantic salmon. Front. Genet. 11: 264. doi:10.3389/fgene.2020.00264.

Ni, G., Liu, G., and Yu, K. 2017. Identification of key genes associated with the effect of osmotic stimuli on intervertebral discs using microarray analysis. Oncology Letters 14(4): 4249–4255. doi:10.3892/ol.2017.6657.

Okado, H. 2019. Regulation of brain development and brain function by the transcriptional repressor RP58. Brain Research 1705: 15–23. doi:10.1016/j.brainres.2018.02.042.

Olesen, E.T.B., and Fenton, R.A. 2017. Aquaporin-2 membrane targeting: still a conundrum. American Journal of Physiology-Renal Physiology 312(4): F744–F747. doi:10.1152/ajprenal.00010.2017.

Padmakumar, V.C., Speer, K., Pal-Ghosh, S., Masiuk, K.E., Ryscavage, A., Dengler, S.L., Hwang, S., Edwards, J.C., Coppola, V., Tessarollo, L., Stepp, M.A., and Yuspa, S.H. 2012. Spontaneous skin erosions and reduced skin and corneal wound healing characterize CLIC4^{NULL} mice. The American Journal of Pathology 181(1): 74–84. doi:10.1016/j.ajpath.2012.03.025.

Padmanabhan, K., Robles, M.S., Westerling, T., and Weitz, C.J. 2012. Feedback regulation of transcriptional termination by the mammalian circadian clock PERIOD complex. Science 337(6094): 599–602. doi:10.1126/science.1221592.

Parrish, D.L., Behnke, R.J., Gephard, S.R., McCormick, S.D., and Reeves, G.H. 1998. Why aren't there more Atlantic salmon (*Salmo salar*)? Canadian Journal of Fisheries and Aquatic Sciences 55(S1): 281–287.

Pellerin, J., and Pierce, U.D. 2015. The Crooked River: characteristics, history, and fisheries management. Available from https://www.maine.gov/ifw/docs/Crooked%20Report%20final5_15.pdf [accessed 16 September 2020].

Pfeifer, B., Alachiotis, N., Pavlidis, P., and Schimek, M.G. 2020. Genome scans for selection and introgression based on *k*-nearest neighbour techniques. Mol Ecol Resour: 1755–0998.13221. doi:10.1111/1755-0998.13221.

Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S.E., and Lercher, M.J. 2014. PopGenome: an efficient Swiss army knife for population genomic analyses in R. Molecular Biology and Evolution 31(7): 1929–1936. doi:10.1093/molbev/msu136.

R Core Team. 2019. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from https://www.R-project.org/.

Riley, S.C., and Power, G. 1987. Age at maturity in landlocked and anadromous Atlantic salmon parr from two Québec rivers. Environmental Biology of Fishes 19(3): 223–228.

Rubin, C.-J., Zody, M.C., Eriksson, J., Meadows, J.R.S., Sherwood, E., Webster, M.T., Jiang, L., Ingman, M., Sharpe, T., Ka, S., Hallböök, F., Besnier, F., Carlborg, Ö., Bed'hom, B., Tixier-Boichard, M., Jensen, P., Siegel, P., Lindblad-Toh, K., and Andersson, L. 2010. Whole-genome resequencing reveals loci under selection during chicken domestication. Nature 464(7288): 587–591. doi:10.1038/nature08832.

Ryman, N., and Laikre, L. 1991. Effects of supportive breeding on the genetically effective population size. Conservation Biology 5(3): 6.

Schubert, E., and Zimek, A. 2019. ELKI: a large open-source library for data analysis - ELKI Release 0.7.5 "Heidelberg." arXiv: 1902.03616 [cs.LG].

Shabalina, S.A., Spiridonov, N.A., and Kashina, A. 2013. Sounds of silence: synonymous nucleotides as a key to biological regulation and complexity. Nucleic Acids Research 41(4): 2073–2094. doi:10.1093/nar/gks1205.

Venney, C.J., Wellband, K.W., and Heath, D.D. 2020. Rearing environment affects the genetic architecture and plasticity of DNA methylation in Chinook salmon. Heredity. doi:10.1038/s41437-020-0346-4.

Willoughby, J.R., Harder, A.M., Tennessen, J.A., Scribner, K.T., and Christie, M.R. 2018. Rapid genetic adaptation to a novel environment despite a genome-wide reduction in genetic diversity. Mol Ecol 27(20): 4041–4051. doi:10.1111/mec.14726.

Yesner, D.R., Hamilton, N.D., and Doyle, R.A. 1984. "Landlocked" salmon and early Holocene lacustrine adaptation in southwestern Maine. North American Archaeologist 4(4): 307–333. doi:10.2190/Q11L-2CJ9-MKKH-NXM0.

Yokoyama, S., Ito, Y., Ueno-Kudoh, H., Shimizu, H., Uchibe, K., Albini, S., Mitsuoka, K., Miyaki, S., Kiso, M., Nagai, A., Hikata, T., Osada, T., Fukuda, N., Yamashita, S., Harada, D., Mezzano, V., Kasai, M., Puri, P.L., Hayashizaki, Y., Okado, H., Hashimoto, M., and Asahara, H. 2009. A systems approach reveals that the myogenesis genome network is regulated by the transcriptional repressor RP58. Developmental Cell 17(6): 836–848. doi:10.1016/j.devcel.2009.10.011.

CHAPTER 4. DENSITY-DEPENDENT FITNESS LIMITS THE EFFECTIVENESS OF HATCHERY SUPPLEMENTATION

4.1 Abstract

In natural populations, density-dependent processes help to regulate population size via changes in one or more vital rates. As population sizes near carrying capacity, compensatory density dependence acts to stabilize population size by decreasing population growth rate. In some salmonid populations, environmental changes that decrease spawning habitat availability combined with the addition of hatchery-origin fish can push population size above carrying capacity. This can result in diminished recruitment due to density-dependent decreases in individual fitness. Here, we examined six studies of hatchery-supplemented salmonid populations to test for the presence of density-dependent effects on fitness. For a subset of these studies, we also use genotypic data to identify decreases in effective population size (N_e) resulting from hatchery supplementation, or Ryman-Laikre effects. Our results suggest that both density-dependent fitness and Ryman-Laikre effects can limit the effectiveness of hatchery supplementation programs. Specifically, we find that in many scenarios, the addition of greater numbers of hatchery-origin fish to a population does not increase recruitment and greatly reduces N_e. The prevalence of density-dependent effects on fitness and Ryman-Laikre effects in supplemented populations indicates that both of these potential outcomes should be taken into consideration when designing or modifying hatchery supplementation strategies.

4.2 Introduction

Wild populations are declining at increasing rates due to habitat destruction, climate change, and myriad environmental changes (Butchart 2010; Chase et al. 2020). One widely used stop-gap measure is to release captive-born individuals into the wild in attempts to increase the number of breeding individuals in a given locality. However, these efforts may not be effective, and could even be counterproductive, if populations are strongly regulated by density-dependent processes. Density dependence refers to any relationship between population density or size and a per capita rate of population change, such as birth, immigration, death, or emigration rate (Turchin 2003; Hixon and Johnson 2009). Specifically, compensatory density dependence occurs when population

growth rate decreases with increasing population density, and this negative feedback process acts to stabilize population size when a population exceeds carrying capacity (Rose et al. 2001). This mechanism can be documented with density-dependent fitness relationships, wherein per-capita fitness decreases with increasing population size. This phenomenon is commonly observed in salmonid populations when parent spawning density increases above carrying capacity, leading to decreased recruitment in the next generation (Chilcote 2003; Independent Scientific Advisory Board 2015).

Hatchery supplementation combined with environmentally-mediated reductions in carrying capacity may be responsible for increasing the magnitude of density-dependent effects. Here, we use 'supplementation' to refer to any program that aims to support wild population recovery by adding hatchery-raised individuals to the population with the expectation that these individuals will reproduce in the wild. These additional individuals are expected to provide a demographic boost to the population. However, when too many hatchery-produced individuals are added to a system, individual fitness can decrease for hatchery- and natural-origin fish in response, leading to a lower number of spawners (i.e., adults returning from the ocean to reproduce) than might have been achieved with less hatchery effort. Hatchery-origin fish also tend to have lower reproductive success than natural-origin fish, and increasing the proportion of hatchery-origin fish in the population can also reduce the number of recruits per spawner (Chilcote et al. 2011).

In addition to potentially negative demographic outcomes, hatchery supplementation can also alter population genetic parameters and decrease effective population size (N_e) in two main ways. First, initiation of captive breeding can often result in a bottleneck, leading to increased genetic drift and decreased genetic diversity in subsequent hatchery-supplemented generations (Ryman and Laikre 1991; Christie et al. 2012). Second, variance in reproductive success among captive-bred individuals can result in a high proportion of offspring being produced by a small number of parents. In this way, captive breeding practices can lead to a Ryman-Laikre effect, wherein N_e is reduced in supplemented populations compared to that of the original wild population if supplementation were not to have occurred (Ryman and Laikre 1991; Ryman et al. 1995). The intensity of this effect can increase with hatchery contributions to the population (Christie et al. 2012).

When a large proportion of a population near carrying capacity is hatchery-produced, density-dependent effects on fitness and Ryman-Laikre effects could act simultaneously to reduce

N_e. Although each of these effects has been identified in many supplemented populations, it is unclear how often both phenomena co-occur or how hatchery strategies should be modified to mitigate loss of genetic variation. We analyzed data from six studies of hatchery-supplemented salmonid populations to test for density-dependent effects on individual fitness. For a subset of these studies, we also used individual genotype information to test for the presence of Ryman-Laikre effects. Specifically, our objectives were to: (1) determine which populations were subject to these effects and (2) investigate whether stronger density-dependent effects on fitness also correspond to more pronounced Ryman-Laikre effects. Our results suggest that both density-dependent and Ryman-Laikre effects need to be taken into account when designing or modifying population supplementation strategies.

4.3 Methods

4.3.1 Data collection and population descriptions

We identified six studies encompassing six rivers and four species in which nearly all adult fish returning to spawn in a population were sampled for parentage analysis (i.e., the populations were fully pedigreed). In all studies, adults returning to spawn were sampled at a dam or weir that acts as a barrier to upstream migration. For five of the studies, offspring were sampled as returning adults and were assigned to parent pairs. In the remaining study, offspring were sampled as juveniles during the summer and were assigned back to parent pairs of individuals that had returned to spawn the previous fall (Milot et al. 2013).

Because all studies were conducted on hatchery-supplemented populations, parentage analysis also allowed offspring to be identified as either hatchery- or natural-origin fish. Hatchery-origin fish were produced by returning adults that had been collected from the population and brought into the captive hatchery environment to be spawned. Natural-origin fish were produced by adults spawning naturally in the wild, regardless of hatchery ancestry. For example, fish were classified as natural-origin even if their grandparents had been spawned in a hatchery (i.e., if their parents were hatchery-origin fish). For each study and each run year, we obtained the total number of hatchery- and natural-origin adults returning to spawn. We then used the total number of offspring assigned to those adults to calculate fitness (total number of offspring divided by total number of spawning adults). We also calculated the relative reproductive success (RRS) of

hatchery-origin fish compared to natural-origin fish by dividing fitness of hatchery-origin fish by fitness of natural-origin fish in each year.

Study 1: Wenatchee River Chinook salmon

The Wenatchee River (Washington, U.S.A.) population of spring-run Chinook salmon (*Oncorhynchus tshawytscha*) is listed as endangered under the U.S. Endangered Species Act and hatchery supplementation has been used to bolster population size since 1989 (Ford et al. 2013). This supplementation program has produced 50-80% of naturally spawning individuals in the population annually. Captive breeding is conducting using both natural- and hatchery-origin fish sampled at a weir in the Chiwawa River, a major tributary of the Wenatchee River (Williamson et al. 2010).

Study 2: Umpqua River coho salmon

The Umpqua River (Oregon, U.S.A.) coho salmon (*O. kisutch*) supplementation program began sustained stocking in 1980 and relied exclusively on local fish (30% natural-origin, 70% hatchery-origin) for captive breeding from 1983-2006 (Thériault et al. 2011). For each year of this study, hatchery-origin fish were released as unfed fry and smolts and we report data for fish released as smolts.

Study 3: Hood River steelhead trout

Hatchery supplementation of the Hood River steelhead (*O. mykiss*) population began in 1991 along with genetic sampling of nearly every returning individual (Araki et al. 2007a). This supplementation program marked a transition from stocking out-of-basin fish to reliance on within-river collections of individuals for captive breeding.

Study 4: Malbaie River Atlantic salmon

The Malbaie River (Québec, Canada) population of Atlantic salmon (*Salmo salar*) has been supported by hatchery supplementation since 1992 (Milot et al. 2013). Each year, approximately 20 returning individuals are used as broodstock, and fish used for captive breeding may have been

spawned in either the wild or the hatchery. This is the only study in our survey to only quantify fitness using juvenile offspring rather than offspring returning as adults.

Study 5: Little Sheep Creek steelhead trout

Little Sheep Creek (Oregon, U.S.A.) steelhead trout are listed as endangered under the Endangered Species Act, as part of the Snake River population designation (Berntson et al. 2011). Hatchery supplementation began in 1982 using wild returning adults. Beginning in 1985, natural- and hatchery-origin fish were used for captive breeding. For the six years studied, an average of 90% of the fish used for captive breeding were hatchery-produced due to the small number of fish returning annually, meaning that even natural-origin fish likely have high hatchery ancestry.

Study 6: Johnson Creek Chinook salmon

The Chinook salmon population in Johnson Creek (Idaho, U.S.A.) has been hatchery-supplemented since 1998. This program was initiated by the Nez Perce Tribe and exclusively uses natural-origin fish as captive broodstock in an effort to restore a nearly extirpated population (Hess et al. 2012; Janowitz-Koch et al. 2019). Before 1998, this system had not received input of any hatchery-reared individuals.

4.3.2 Testing for density-dependent effects on fitness

To test for the presence of density-dependent effects on fitness, we used linear regressions to quantify relationships between number of spawning adults (i.e., spawning density) and per capita lifetime reproductive success (hereafter, "fitness"). To meet assumptions of normality, both variables were $\log_{10}$-transformed prior to regression analysis. For Umpqua River coho salmon and Malbaie River Atlantic salmon (studies 2 and 4, respectively) only three years of data were available, and we did not formally test for density-dependent effects on fitness for these two populations given the small number of data points.

4.3.3 Estimating effective number of breeders (N_b)

We were able to obtain estimates of effective number of breeders (N_b) per broodyear in two populations. For Hood River steelhead trout, estimates of N_b were published for hatchery-

origin (N_h) and natural-origin (N_w) fish for each broodyear of the study (Christie et al. 2012). These estimates were calculated using offspring microsatellite genotypes at eight loci and a linkage disequilibrium-based approach implemented in LDNe v1.2 with the lowest allele frequency allowed set to 0.02 (Waples and Do 2008).

For Johnson Creek Chinook salmon, we used publicly available genotypes and phenotypic information to assign individuals to broodyear and hatchery vs. natural origin and to estimate N_b for each group (Janowitz-Koch et al. 2019). For the years 2000-2008, genotypes for 15 microsatellite loci were determined for all returning adults, whereas individuals were genotyped at 298 single nucleotide polymorphisms (SNPs) in 2003-2011 (Hess et al. 2012; Janowitz-Koch et al. 2019). We filtered the microsatellite data using the same criteria applied by Hess et al., removing individuals genotyped at <12 of 15 loci (2012). We also filtered the SNP data, removing SNP loci with ≥ 10% missing data and individuals with ≥10% missing genotypes. We analyzed and report results for microsatellite and SNP genotypes separately. To calculate N_b, we used the same approach and settings as were used for the Hood River steelhead data implemented in *NeEstimator* v2.1 (Do et al. 2014).

4.3.4 Testing for Ryman-Laikre effects

To test for Ryman-Laikre effects in the Johnson Creek Chinook population, we used the same approach applied by Christie et al. (2012) for Hood River steelhead. First, separate estimates of effective number of breeders for natural- (N_w) and hatchery-origin (N_h) fish obtained from *NeEstimator*. These values were then used to calculate combined effective number of breeders (N_T) as:

$$\frac{1}{N_T} = \frac{x^2}{N_h} + \frac{(1-x)^2}{N_w}$$

where x is the contribution of hatchery-origin fish to the next generation (Ryman and Laikre 1991). We calculated x as:

$$x = \frac{N_{hatchery} \times RRS}{N_{total}}$$

where $N_{hatchery}$ equals the number of hatchery-origin spawning adults, RRS is the reproductive success of hatchery-origin fish relative to natural-origin fish, and N_{total} equals the total number of potential parents (natural- and hatchery-origin combined). We obtained RRS for each year by

dividing the fitness of hatchery-origin fish by the fitness of natural-origin fish. We then averaged annual RRS values to obtain a single RRS estimate used to calculate year-specific values of x.

Next, we compared the combined effective number of breeders in the entire population (natural- and wild-origin fish combined; N_T) to the effective number of natural-origin breeders (N_w) by dividing N_T by N_w. When this ratio equals one, the reproductive contributions of hatchery-origin fish do not impact the effective number of breeders in the population (i.e., N_w is equal to N_T). However, when this ratio is less than one, the contribution of hatchery-origin fish to the populations decreases the effective population size below what it would be without supplementation and leads to a Ryman-Laikre effect (i.e., N_T is less than N_w). Finally, we tested the influence of x on N_T / N_w for each population and data type (i.e., microsatellites and SNPs for Johnson Creek Chinook) using linear regressions. We used the values provided by Christie et al. for x, N_T, and N_w to construct the linear model for Hood River steelhead (2012).

4.4 Results

4.4.1 Data collection and density-dependent effects

We identified six studies of four species in which nearly all returning adults in each population were sampled each year (range: 3-10 brood years per study; Table 4.1). From each study and for each year, we recorded the numbers of natural- and hatchery-origin fish returning to spawn. (Table C.1). With few exceptions, RRS was consistently <1 across populations, indicating lower reproductive success for hatchery-origin fish compared to natural-origin fish.

Table 4.1. Overview of characteristics for studies analyzed. Study species belong to two genera, *Oncorhynchus* (Pacific salmonids) and *Salmo* (Atlantic salmonids). 'Offspring life stage' refers to the stage at which parental fitness was evaluated (i.e., offspring either returning as adults or young-of-the-year juveniles). 'Years' refers to the number of brood years analyzed for each study.

Study	Common name	Species name	Offspring life stage	Years	Reference(s)
1	Chinook salmon	*O. tshawytscha*	adult	7	(Ford et al. 2013)
2	coho salmon	*O. kisutch*	adult	3	(Thériault et al. 2011)
3	steelhead trout	*O. mykiss*	adult	9	(Araki et al. 2007a, 2007b; Christie et al. 2012)
4	Atlantic salmon	*S. salar*	juvenile	3	(Milot et al. 2013)
5	steelhead trout	*O. mykiss*	adult	6	(Berntson et al. 2011)
6	Chinook salmon	*O. tshawytscha*	adult	10	(Hess et al. 2012; Janowitz-Koch et al. 2019)

For the four studies examined for density-dependent effects, we found negative relationships between number of spawning adults and per capita fitness for both hatchery- and natural-origin fish in all populations (Fig. 4.1; Table 4.2). Linear regression results were significant (α=0.05) for two of the four populations tested: Hood River steelhead and Johnson Creek Chinook, the two studies with the highest sample sizes (9 and 10 years, respectively).

Table 4.2. Linear regression results for number of spawning adults vs. per capita fitness. Both variables were $\log_{10}$-transformed prior to analysis to meet assumptions of normality. Note that Studies 2 and 4 were not analyzed due to small sample sizes (n<5). df = degrees of freedom for F-statistic.

Study	Origin	Slope	F	df	Adj. R^2	p
1	natural	-0.73	5.88	1,5	0.45	0.060
1	hatchery	-0.27	0.44	1,5	-0.10	0.54
3	natural	-1.29	28.84	1,7	0.78	**0.0010**
3	hatchery	-1.69	7.36	1,7	0.44	**0.030**
5	natural	-1.06	4.02	1,4	0.38	0.12
5	hatchery	-0.81	3.79	1,4	0.36	0.12
6	natural	-1.97	37.01	1,8	0.80	**0.0003**
6	hatchery	-1.90	23.80	1,8	0.72	**0.0012**

4.4.2 Estimating effective number of breeders (N_b)

For Johnson Creek Chinook salmon, the filtered microsatellite data included 4,564 individuals genotyped at ≥12 of 15 loci, representing 9 brood years (2000-2008). The filtered SNP data included 6,216 individuals from 9 brood years (2003-2011) genotyped at ≥90% of 263 loci. For both data types and for Hood River steelhead, *NeEstimator* results are provided in Table 4.3. Across all estimates of N_b, 95% jackknife confidence intervals did not overlap for hatchery- and natural-origin estimates, with the exception of brood year 2003 for Johnson Creek Chinook using SNP genotypes. Natural-origin N_b (N_w) was consistently greater than hatchery-origin N_b (N_h). Within the Johnson Creek data set, both microsatellite and SNP data were available for brood years 2003-2008 and N_b estimates for the two genotyping methods were similar across years (Fig. 4.2).

4.4.3 Ryman-Laikre effects

For Johnson Creek Chinook salmon, we determined the average RRS values across years (2000-2011) to be 0.71 and we used this value to calculate year-specific values for x (Table 4.3).

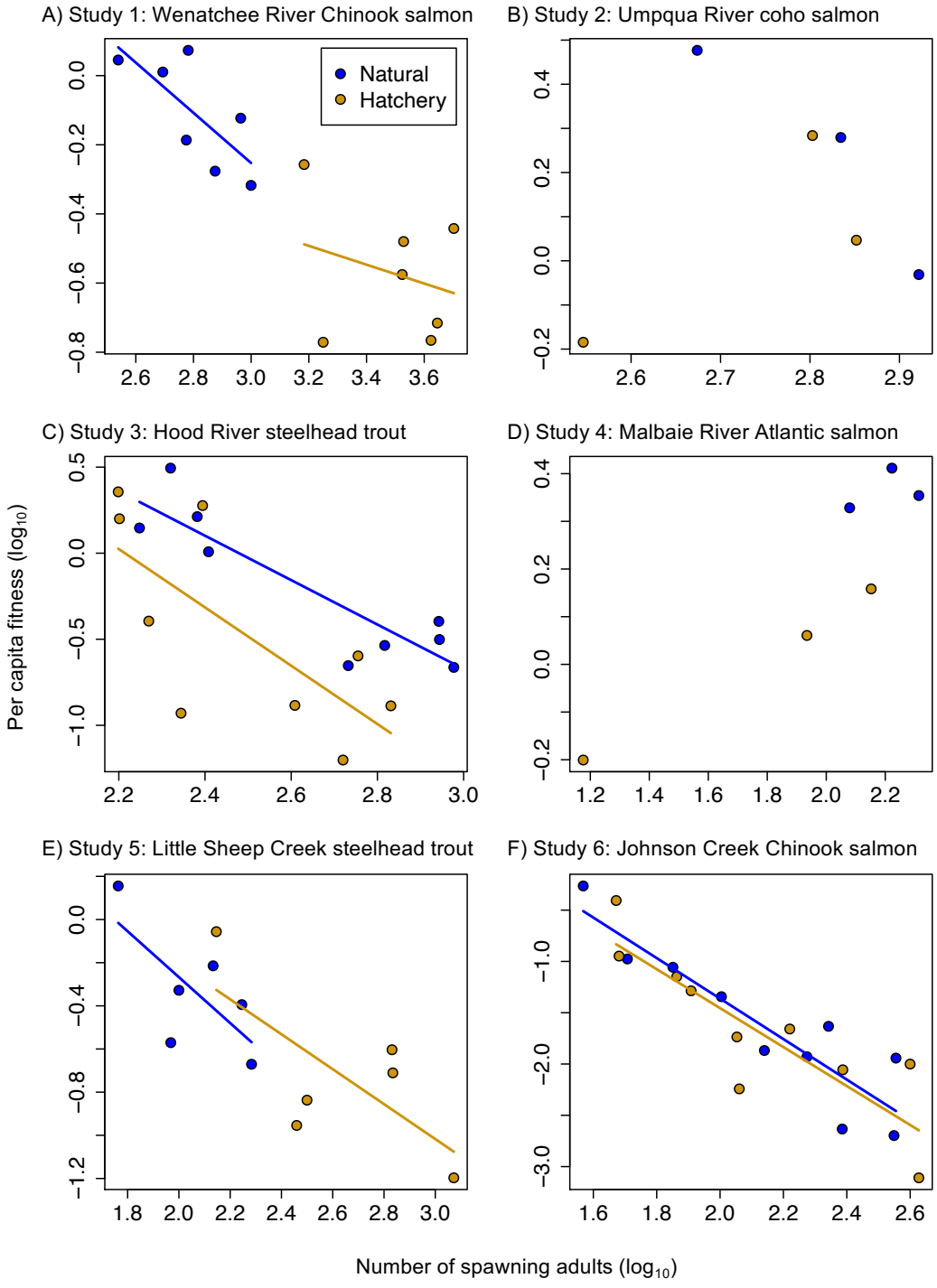


Figure 4.1. Relationship between number of spawning adults and per capita fitness for six populations (A-F). Corresponding references for each study are provided in Table 4.1. Linear regression trendlines are plotted for all studies except Umpqua River coho salmon and Malbaie River Atlantic salmon (studies 2 and 4, respectively), for which sample sizes were small. Both variables were $\log_{10}$-transformed to meet the assumption of normality prior to linear regression.

Table 4.3. *NeEstimator* results and calculated values for x and N_T. Marker type can be either microsatellite (M) or SNP (S).

Study	Marker type	Brood year	Natural			Hatchery			x	N_T
			N_w	Jackknife CI		N_h	Jackknife CI			
3	M	1993	222.4	196.9	251.6	23.6	21.7	25.7	0.37	100.9
3	M	1994	285.5	232.9	362.2	30.3	27.8	32.9	0.49	120.1
3	M	1995	180.1	157.3	208.2	22.2	20.0	24.6	0.34	141.7
3	M	1996	250.3	230.5	271.8	21.2	19.2	23.3	0.10	264.6
3	M	1997	426.6	372.7	491.6	30.7	28.1	33.6	0.19	382.1
3	M	1998	517.2	452.2	600.1	29.1	27.2	31.1	0.35	202.0
3	M	1999	588.9	500.1	710.4	38.7	35.3	42.2	0.42	187.9
3	M	2000	663.6	559.6	808.9	21.6	20.2	23.1	0.31	158.6
3	M	2001	577.3	492.0	694.4	27.8	25.3	30.5	0.46	116.8
3	M	2002	650.4	497.0	922.7	33.2	29.3	37.6	0.26	345.0
3	M	2003	866.1	604.9	1468.9	27.7	24.7	30.9	0.32	256.8
6	S	2003	223.5	146.3	437.1	47.5	22.7	399.2	0.22	267.8
6	S	2004	118.6	102.7	138.1	20.8	17.4	24.6	0.31	116.0
6	S	2005	96.0	83.5	111.1	33.5	30.9	36.2	0.25	129.5
6	S	2006	70.0	63.4	77.2	26.2	24.5	28.0	0.30	95.8
6	S	2007	109.6	96.3	125.4	27.1	23.2	31.4	0.42	103.6
6	S	2008	160.2	145.1	177.3	27.0	23.7	30.6	0.38	127.8
6	S	2009	187.5	159.3	223.9	25.9	21.7	31.0	0.49	93.3
6	S	2010	200.2	180.1	222.7	25.1	23.2	27.2	0.38	132.2
6	S	2011	189.6	154.0	239.2	22.4	21.1	23.9	0.31	147.0
6	M	2000	47.7	41.0	55.9	21.4	18.4	24.7	0.06	53.4
6	M	2001	335.1	203.3	795.5	36.6	28.8	47.4	0.36	221.9
6	M	2002	287.1	211.5	429.9	30.6	22.2	43.8	0.43	141.7
6	M	2003	176.1	137.9	236.3	22.8	18.5	28.0	0.22	180.3
6	M	2004	124.9	106.6	147.8	25.5	23.0	28.1	0.31	132.0
6	M	2005	84.4	70.9	101.7	32.1	29.0	35.4	0.25	116.2
6	M	2006	73.7	61.8	87.8	24.3	22.0	26.8	0.30	96.5
6	M	2007	98.8	84.1	117.0	28.0	23.8	32.7	0.42	102.4
6	M	2008	193.5	171.4	219.7	25.7	22.7	29.1	0.38	130.0

Values of N_T/N_w were similar between SNP and microsatellite data within brood years, with neither measure appearing to consistently over- nor underestimate this ratio across values of x (Fig. 4.3).We found a negative relationship between x and N_T/N_w for all three data sets (Fig. 4.4), and a significant association between these two variables for Hood River steelhead and Johnson Creek Chinook SNP data (Table 4.4). For Hood River steelhead and Johnson Creek Chinook SNP data, most N_T/N_w values fell below 1, indicating that hatchery-origin fish contribute to decreases in N_T for most brood years. For Johnson Creek Chinook microsatellite data, 3 of 9 N_T/N_w estimates were <1.

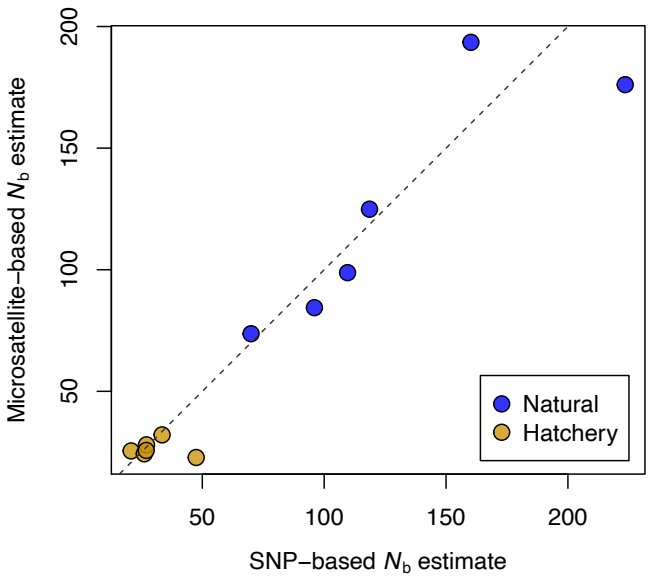


Figure 4.2. Relationship between estimates for effective number of breeders (N_b) using SNP and microsatellite genotypes. Dashed line represents a 1:1 relationship.

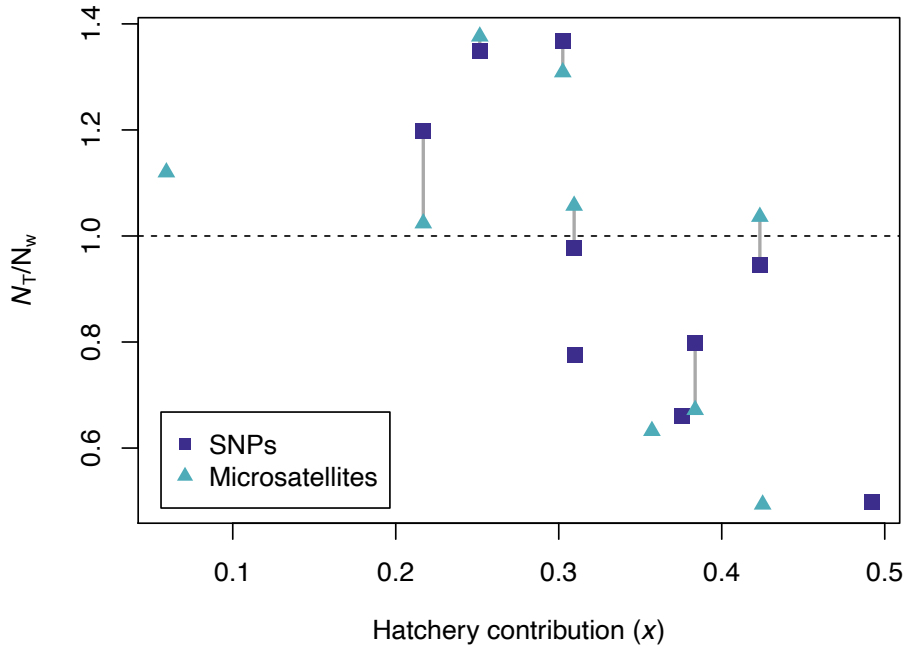


Figure 4.3. For Johnson Creek Chinook salmon, microsatellite and SNP estimates of N_T/N_W for the same brood years are similar. Estimates linked by grey lines correspond to the same brood year. Neither approach consistently results in a greater estimate of N_T/N_W, indicating that neither approach reliably under- nor overestimates N_b.

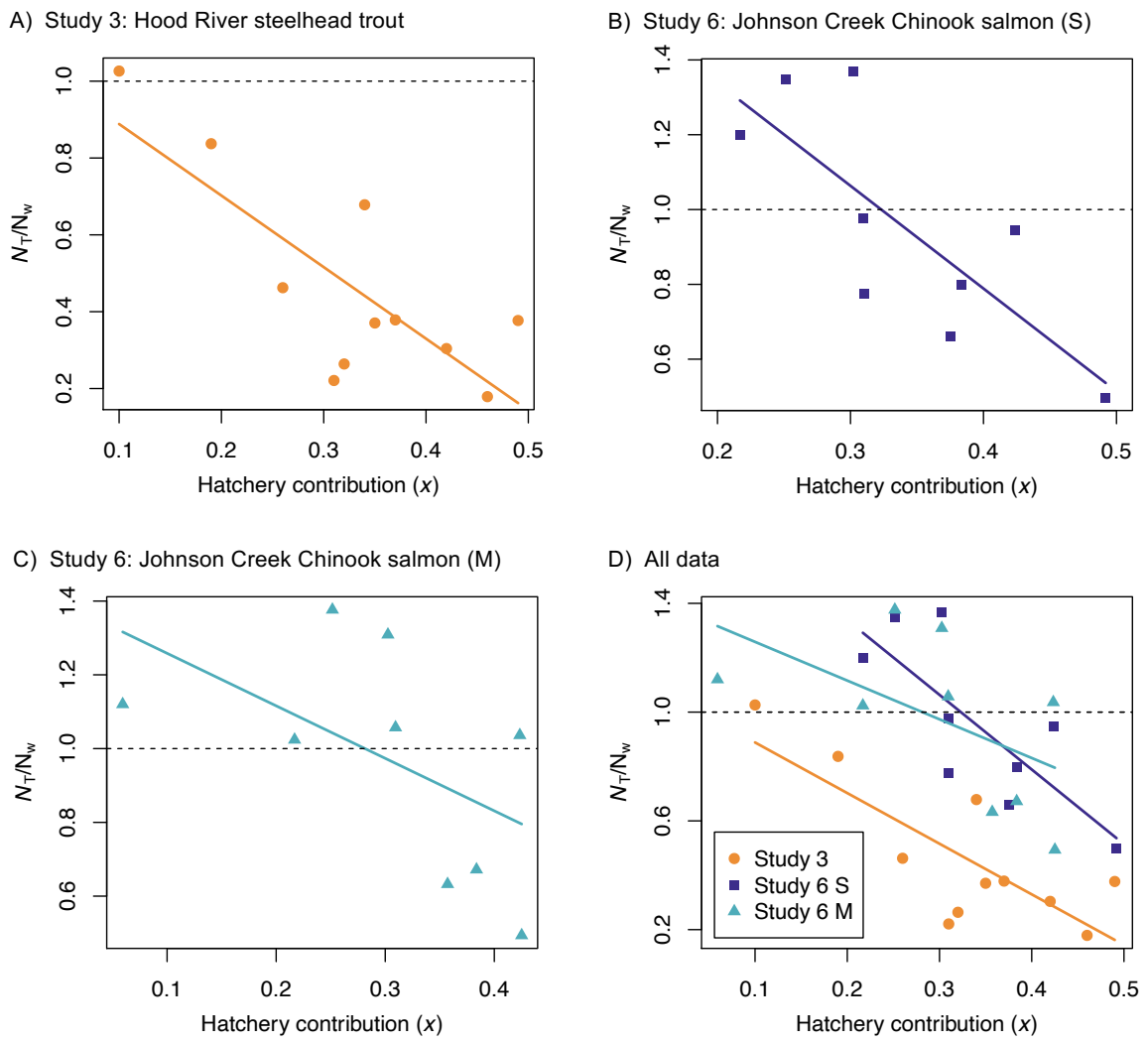


Figure 4.4. N_T/N_w decreases with increasing hatchery contributions to the next generation (x) across Hood River steelhead trout (A,B,D) and Johnson Creek Chinook salmon (C,D). Results for two genetic marker types are displayed for Johnson Creek Chinook: SNPs (S) and microsatellites (M) in panels (B) and (C), respectively. The dashed line represents the value of N_T/N_w below which Ryman-Laikre effects are observed. The majority of N_T/N_w estimates are <1, with stronger Ryman-Laikre effects at higher values of x (D).

Table 4.4. Linear regression results for x vs. N_T/N_w. Marker type can be either microsatellite (M) or SNP (S). df = degrees of freedom for F-statistic.

Study	Marker type	Slope	F	df	Adj. R^2	p
3	M	-1.862	14.65	1,9	0.577	**0.0040**
6	S	-2.746	10.80	1,7	0.551	**0.0134**
6	M	-1.423	2.915	1,7	0.193	0.1315

4.5 Discussion

The negative relationship we identified between number of spawning adults (i.e., density) and fitness in four of the studies suggests the presence of density-dependent effects on fitness in these populations (Fig. 4.1). Although this relationship was only significant for two of the four studies, acquiring additional years of data for the two studies with non-significant relationships may support more robust conclusions (Table 4.2). Two studies—Umpqua River coho salmon and Malbaie River Atlantic salmon—only reported data from three years each. For Malbaie River Atlantic salmon, fitness appears to increase with increasing spawner density but this is the only study in which fitness was only measured as number of juvenile offspring rather than number of returning adult offspring (Milot et al. 2013). This population may be subject to density-dependent effects on fitness but only at later life stages (e.g., as smolts migrate to the ocean or as adults return to spawn). The Umpqua River coho data suggest density-dependent effects on fitness, but for both studies, small sample size precludes definitive assessment. However, for the majority of populations we examined, our results suggest that fitness decreases with increasing spawner density and that hatchery stocking of a greater number of individuals in these populations would not result in a demographic boost in the subsequent generation.

We also detected Ryman-Laikre effects in each of the populations with available genotype data, with N_T/N_w values <1 observed for multiple years within each population. We found that increasing hatchery contributions to the population corresponded to stronger decreases in effective population size (Fig. 4.4). Our sample size is insufficient to formally test for a relationship between the strength of density-dependent effects on fitness (i.e., the slopes of linear regressions in Fig. 4.1 and Table 4.2) and degree of Ryman-Laikre effects observed (i.e., the slopes of linear regressions in Fig. 4.4 and Table 4.4). However, the Hood River steelhead and Johnson Creek Chinook populations exhibit both Ryman-Laikre effects and density-dependent fitness, suggesting the potential for interactive effects between these two phenomena. Incorporating genotype data for additional populations will allow for quantitative assessment of whether stronger compensatory dynamics increase the magnitude of Ryman-Laikre effects.

From a management perspective, the presence of density-dependent effects on fitness means that stocking hatchery-origin fish may only be beneficial up to a point. As more and more fish are introduced into the system, the per capita decrease in fitness (of which survival is a component) means that fewer individuals (proportionately) will survive to adulthood. In extreme

cases, stocking 1,000,000 smolts may result in the same number of returning adult individuals as stocking 10,000 smolts. The stronger the relationship (i.e., the steeper the slope) between spawner density and individual fitness, the greater the trade-off between stocking numbers and returns. Increasing the proportion of hatchery-origin adults that spawn in the population could lead to the intensification of Ryman-Laikre effects, further reducing N_e. Collection of genotypic and demographic data from additional populations and years for populations with existing data will allow for quantification of relationships between density-dependent and Ryman-Laikre effects. Furthermore, these analyses will allow for the parameterization of agent-based model designed to test the effects of varying hatchery strategies on individual fitness and N_e in hatchery-supplemented populations.

As natural populations of terrestrial and aquatic taxa continue to decline, captive breeding may emerge as a viable conservation option in some contexts. Our results demonstrate that adding too many captive-origin parents to a population may not only decrease the number of offspring in subsequent generations due to density-dependent effects on fitness, but may also lead to loss of genetic diversity via Ryman-Laikre effects. Particularly for populations faced with future loss of habitat and consequent decreases in carrying capacity, captive breeding and supplementation strategies must account for these potential effects to mitigate loss of genetic diversity.

4.6 References

Araki, H., Ardren, W.R., Olsen, E., Cooper, B., and Blouin, M.S. 2007a. Reproductive success of captive-bred steelhead trout in the wild: evaluation of three hatchery programs in the Hood River. Conservation Biology 21(1): 181–190. doi:10.1111/j.1523-1739.2006.00564.x.

Araki, H., Cooper, B., and Blouin, M.S. 2007b. Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. Science 318(5847): 100–103. doi:10.1126/science.1145621.

Berntson, E.A., Carmichael, R.W., Flesher, M.W., Ward, E.J., and Moran, P. 2011. Diminished reproductive success of steelhead from a hatchery supplementation program (Little Sheep Creek, Imnaha Basin, Oregon). Transactions of the American Fisheries Society 140(3): 685–698. doi:10.1080/00028487.2011.584489.

Butchart, R.B. 2010. Global biodiversity: indicators of recent declines. Science 328(5982): 1164–1168. doi:10.1126/science.1186777.

Chase, J.M., Blowes, S.A., Knight, T.M., Gerstner, K., and May, F. 2020. Ecosystem decay exacerbates biodiversity loss with habitat loss. Nature 584(7820): 238–243. doi:10.1038/s41586-020-2531-2.

Chilcote, M.W. 2003. Relationship between natural productivity and the frequency of wild fish in mixed spawning populations of wild and hatchery steelhead (*Oncorhynchus mykiss*). Can. J. Fish. Aquat. Sci. 60(9): 1057–1067. doi:10.1139/f03-092.

Chilcote, M.W., Goodson, K.W., and Falcy, M.R. 2011. Reduced recruitment performance in natural populations of anadromous salmonids associated with hatchery-reared fish. Can. J. Fish. Aquat. Sci. 68(3): 511–522. doi:10.1139/F10-168.

Christie, M.R., Marine, M.L., French, R.A., Waples, R.S., and Blouin, M.S. 2012. Effective size of a wild salmonid population is greatly reduced by hatchery supplementation. Heredity 109(4): 254–260. doi:10.1038/hdy.2012.39.

Do, C., Waples, R.S., Peel, D., Macbeth, G.M., Tillett, B.J., and Ovenden, J.R. 2014. NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. Mol Ecol Resour 14(1): 209–214. doi:10.1111/1755-0998.12157.

Ford, M.J., Howard, S., Murdoch, A.R., and Hughes, M.S. 2013. Monitoring the reproductive success of naturally spawning hatchery and natural spring Chinook salmon in the Wenatchee River. Bonneville Power Administration.

Hess, M.A., Rabe, C.D., Vogel, J.L., Stephenson, J.J., Nelson, D.D., and Narum, S.R. 2012. Supportive breeding boosts natural population abundance with minimal negative impacts on fitness of a wild population of Chinook salmon. Mol Ecol 21(21): 5236–5250. doi:10.1111/mec.12046.

Hixon, M.A., and Johnson, D.W. 2009. Density dependence and independence. In Encyclopedia of Life Sciences. Edited by John Wiley & Sons, Ltd. John Wiley & Sons, Ltd, Chichester, UK. p. a0021219. doi:10.1002/9780470015902.a0021219.

Independent Scientific Advisory Board. 2015. Density dependence and its implications for fish management and restoration programs in the Columbia River Basin. Northwest Power and Conservation Council, Columbia River Basin Indian Tribes, and National Marine Fisheries Service, Portland, OR. Available from https://www.nwcouncil.org/sites/default/files/isab2015-1_0.pdf [accessed 22 October 2020].

Janowitz-Koch, I., Rabe, C., Kinzer, R., Nelson, D., Hess, M.A., and Narum, S.R. 2019. Long-term evaluation of fitness and demographic effects of a Chinook salmon supplementation program. Evol Appl 12(3): 456–469. doi:10.1111/eva.12725.

Milot, E., Perrier, C., Papillon, L., Dodson, J.J., and Bernatchez, L. 2013. Reduced fitness of Atlantic salmon released in the wild after one generation of captive breeding. Evol Appl 6(3): 472–485. doi:10.1111/eva.12028.

Rose, K.A., Cowan, J.H., Winemiller, K.O., Myers, R.A., and Hilborn, R. 2001. Compensatory density dependence in fish populations: importance, controversy, understanding and prognosis. Fish and Fisheries 2(4): 293–327. doi:10.1046/j.1467-2960.2001.00056.x.

Ryman, N., Jorde, P.E., and Laikre, L. 1995. Supportive breeding and variance effective population size. Conservation Biology 9(6): 1619–1628. doi:10.1046/j.1523-1739.1995.09061619.x.

Ryman, N., and Laikre, L. 1991. Effects of supportive breeding on the genetically effective population size. Conservation Biology 5(3): 6.

Thériault, V., Moyer, G.R., Jackson, L.S., Blouin, M.S., and Banks, M.A. 2011. Reduced reproductive success of hatchery coho salmon in the wild: insights into most likely mechanisms. Molecular Ecology 20(9): 1860–1869. doi:10.1111/j.1365-294X.2011.05058.x.

Turchin, P. 2003. Connecting mathematical theory to empirical dynamics. In Complex population dynamics: a theoretical/empirical synthesis. Princeton University Press.

Waples, R.S., and Do, C. 2008. LDNE: a program for estimating effective population size from data on linkage disequilibrium. Molecular Ecology Resources 8(4): 753–756. doi:10.1111/j.1755-0998.2007.02061.x.

Williamson, K.S., Murdoch, A.R., Pearsons, T.N., Ward, E.J., and Ford, M.J. 2010. Factors influencing the relative fitness of hatchery and wild spring Chinook salmon (Oncorhynchus tshawytscha) in the Wenatchee River, Washington, USA. Can. J. Fish. Aquat. Sci. 67(11): 1840–1851. doi:10.1139/F10-099.

APPENDIX A. CHAPTER 2 SUPPLEMENTAL MATERIALS

INTRODUCTION

Coordinated reintroduction efforts for Lake Champlain Atlantic salmon began in 1972 through stocking hatchery reared salmon, but self-sustaining wild populations have not been reestablished. A major limiting factor to the reintroduction effort was the illegal introduction of the non-native alewife to Lake Champlain in 2003 (Marsden & Hauser, 2009). Alewife currently comprise a sufficient proportion of Atlantic salmon diets to cause thiamine-related mortality in eggs collected for hatchery rearing (Marsden & Langdon, 2012), and the Atlantic salmon population remains entirely supported by hatchery supplementation. In the hatchery, fertilized salmon eggs must be treated with thiamine to prevent high rates of mortality due to thiamine deficiency.

METHODS

Experimental crosses

We transported gametes at 4 °C to the White River National Fish Hatchery (Bethel, Vermont, USA), where we systematically combined milt and eggs to generate 35 families (32 males paired to 35 females; in 2016, low availability of males required that one male be crossed with three females and that an additional male be crossed with two females; half-sib families were used in dose-response curve calculations, but not for RNA-seq or Cox proportional hazards regressions. To prevent pathogen introduction to the hatchery, we treated all fertilized eggs with a 0.005% iodophor solution for 30 minutes. We next rinsed fertilized eggs with fresh water to prepare them for treatment with thiamine (treated) or water (untreated).

DISCUSSION

Genes hypothesized to be differentially expressed *a priori*

One thiamine derivative, thiamine diphosphate (TDP), is required as a cofactor for several enzymes involved in energy production. Conversion of TDP to thiamine triphosphate is catalyzed by adenylate kinase (Fig. A.2). Under low thiamine conditions, downregulation of two paralogous genes encoding adenylate kinase may decrease the amount of TDP converted to thiamine triphosphate, resulting in conservation of intracellular TDP pools. Reduced folate carrier transports folate and phosphorylated thiamine derivatives across the cell membrane, and decreased

expression of this carrier is associated with increased accumulation of intracellular TDP (Zhao et al., 2001). Downregulation of reduced folate carrier in untreated individuals may also conserve intracellular TDP by reducing the amount of TDP that is exported from the cell. Maintenance of cellular TDP concentration is required for proper cellular metabolism (Bettendorff, 2013), and adenylate kinase and reduced folate carrier may comprise an important cellular response to thiamine-poor conditions.

The downregulation of thiamine transporter 2 in untreated samples is unexpected, given extensive documentation of increased thiamine transporter 2 mRNA transcription and promoter activity under thiamine deficiency conditions in a variety of study organisms and tissue types (Nabokina, Subramanian, Valle, & Said, 2013; Reidling & Said, 2005). Our observation may conflict with these studies because expression of thiamine transporter 2 is dependent on developmental stage, with expression generally decreasing with maturation (Reidling, Nabokina, Balamurugan, & Said, 2006). Downregulation of thiamine transporter 2 under thiamine deficiency has been described in human cell line culture (Nabokina et al., 2013) and mature mice (Reidling & Said, 2005), but the combined effects of thiamine deficiency and developmental stage on transporter expression have not been evaluated. Juvenile individuals may respond to thiamine deficiency differently than mature individuals at the level of gene expression.

Popeye domain-containing protein 2: additively expressed and putatively adaptive

Popeye domain-containing protein 2 (*popdc2*), an additively and adaptively expressed gene, plays a crucial role in cardiac and skeletal muscle development, as evidenced by abnormal muscle fiber morphology, pericardial edema, and irregular heart rate observed in *popdc2*-knockdown zebrafish (*Danio rerio*) (Kirchmaier et al., 2012). These effects of *popdc2*-knockdown are mirrored in thiamine deficient juvenile Atlantic salmon (Essa et al., 2011; Fisher, Spitsbergen, Iamonte, Little, & Delonay, 1995; Sechi & Serra, 2007). Genes exhibiting patterns similar to *popdc2* may represent differences in how thiamine is used or allocated across different genetic backgrounds, regardless of thiamine status (i.e., treated or untreated) (Fig. 4A,B).

Example family x treatment genes

Expression of optineurin (*optn*) falls in the first family x treatment category, with thiamine treatment decreasing expression across families with the greatest decrease occurring for families

with low survival (Fig. 4C). Accumulation of optineurin has been documented in nervous system tissues affected by such neurodegenerative diseases as amyotrophic lateral sclerosis (ALS), Parkinson's disease, Creutzfeldt-Jakob disease, and glaucoma (Osawa et al., 2011). Another gene following a similar pattern, gamma-crystallin M2, is differentially upregulated in response to thiamine treatment (Fig. 4D). Gamma-crystallin proteins are the primary lens proteins in the vertebrate eye, and high concentrations of these proteins are required for proper lens clarity (H. Zhao, Brown, Magone, & Schuck, 2011). Multiple copies of gamma-crystallin M2 (n = 3) and M3 (n = 6) are differentially downregulated in families with low survival. Downregulation of this gene has previously been described in Atlantic salmon with diet-induced cataracts (Trösße et al., 2009). In our study, downregulation of this gene in families with low survival may lead to gamma-crystallin protein concentrations insufficient for normal lens development, and may help explain diminished foraging efficiency, predator avoidance, and visual acuity in juvenile salmonids that are thiamine deficient (Carvalho et al., 2009; Fitzsimons et al.,e 2009).

Example putatively adaptive genes

For genes with expression levels positively associated with mortality risk (n = 812), related overrepresented GO terms are broadly indicative of physiological stress. Genes upregulated in individuals at greater risk of mortality may serve to mitigate unfavorable cellular conditions. For example, one such putatively adaptive gene belongs to the glutathione peroxidase family, a group of enzymes found in all domains of life with antioxidant functions (Toppo, Vanin, Bosello, & Tosatto, 2008). Antioxidant defense mechanisms serve to prevent oxidative stress, which results from reactive oxygen species production and can cause tissue damage (Betteridge, 2000). In our study, expression of the glutathione peroxidase 2 gene increased with decreasing survival in untreated individuals, indicating the need for a stronger response to oxidative stress in families with lower survival rates (Fig. 6A). This result is consistent with previous findings that thiamine deficiency increases glutathione peroxidase expression in Baltic Sea Atlantic salmon (Lundström et al., 1999) and that inactivation of one glutathione peroxidase gene leads to oxidative stress in mice (Esposito et al., 2000). Oxidative stress often accompanies thiamine deficiency (Vuori & Nikinmaa, 2007), and due to the capacity of thiamine to act as an antioxidant, oxidative stress may exacerbate thiamine deficiency, and vice versa (Bettendorff, 2013). This positive feedback loop is observed in many human neurodegenerative disorders, including Alzheimer's disease, Parkinson's

disease, and Huntington's disease (Gibson & Zhang, 2002), and may be occurring in families with low rates of survival in our study.

Putatively adaptive genes negatively associated with mortality risk were associated with GO terms related to development and growth (e.g., embryo development, neurogenesis, and kidney development). Upregulation of genes in this category likely underlie maintenance of normal cellular processes in families exhibiting higher rates of survival under thiamine-poor conditions. For example, ATP-sensitive inward rectifier potassium channel 12 (*kcnj12*) belongs to a family of potassium channels with important roles in heart rate regulation (Anumonwo & Lopatin, 2010; Vornanen, 2017) (Fig. 6B). In addition to *kcnj12*, two other members of this family (2 copies of *kcnj1* and 2 copies of *kcnj15*) are also differentially expressed across families, with expression upregulated in families exhibiting higher survival.

LITERATURE CITED

Anumonwo, J. M. B., & Lopatin, A. N. (2010). Cardiac strong inward rectifier potassium channels. Journal of Molecular and Cellular Cardiology, 48(1), 45–54. doi: 10.1016/j.yjmcc.2009.08.013

Bettendorff, L. (2013). Thiamine. In Handbook of Vitamins, Fifth Edition (pp. 267–324). CRC Press.

Betteridge, D. J. (2000). What is oxidative stress? Metabolism, 49(2), 3–8.

Carvalho, P. S. M., Tillitt, D. E., Zajicek, J. L., Claunch, R. A., Honeyfield, D. C., Fitzsimons, J. D., & Brown, S. B. (2009). Thiamine deficiency effects on the vision and foraging ability of lake trout fry. Journal of Aquatic Animal Health, 21(4), 315–325.

Esposito, L. A., Kokoszka, J. E., Waymire, K. G., Cottrell, B., MacGregor, G. R., & Wallace, D. C. (2000). Mitochondrial oxidative stress in mice lacking the glutathione peroxidase-1 gene. Free Radical Biology and Medicine, 28(5), 754–766. doi: 10.1016/S0891-5849(00)00161-1

Essa, E., Velez, M. R., Smith, S., Giri, S., Raman, S. V., & Gumina, R. J. (2011). Cardiovascular magnetic resonance in wet beriberi. Journal of Cardiovascular Magnetic Resonance, 13(1), 41. doi: 10.1186/1532-429X-13-41

Fisher, J. P., Spitsbergen, J. M., Iamonte, T., Little, E. E., & Delonay, A. (1995). Pathological and behavioral manifestations of the "Cayuga syndrome," a thiamine deficiency in larval landlocked Atlantic salmon. Journal of Aquatic Animal Health, 7(4), 269–283.

Fitzsimons, J. D., Brown, S. B., Williston, B., Williston, G., Brown, L. R., Moore, K., … Tillitt, D. E. (2009). Influence of thiamine deficiency on lake trout larval growth, foraging, and predator avoidance. Journal of Aquatic Animal Health, 21(4), 302–314.

Gibson, G. E., & Zhang, H. (2002). Interactions of oxidative stress with thiamine homeostasis promote neurodegeneration. Neurochemistry International, 40(6), 493–504.

Kirchmaier, B. C., Poon, K. L., Schwerte, T., Huisken, J., Winkler, C., Jungblut, B., … Brand, T. (2012). The Popeye domain containing 2 (*popdc2*) gene in zebrafish is required for heart and skeletal muscle development. Developmental Biology, 363(2), 438–450. doi: 10.1016/j.ydbio.2012.01.015

Lundström, J., Carney, B., Amcoff, P., Pettersson, A., Börjeson, H., Förlin, L., & Norrgren, L. (1999). Antioxidant systems, detoxifying enzymes and thiamine levels in Baltic salmon (*Salmo salar*) that develop M74. Ambio, 28, 24–29.

Marsden, J. E., & Hauser, M. (2009). Exotic species in Lake Champlain. Journal of Great Lakes Research, 35(2), 250–265.

Marsden, J. E., & Langdon, R. W. (2012). The history and future of Lake Champlain's fishes and fisheries. Journal of Great Lakes Research, 38, 19–34. doi: 10.1016/j.jglr.2011.09.007

Nabokina, S. M., Subramanian, V. S., Valle, J. E., & Said, H. M. (2013). Adaptive regulation of human intestinal thiamine uptake by extracellular substrate level: A role for THTR-2 transcriptional regulation. American Journal of Physiology-Gastrointestinal and Liver Physiology, 305(8), G593–G599. doi: 10.1152/ajpgi.00237.2013

Osawa, T., Mizuno, Y., Fujita, Y., Takatama, M., Nakazato, Y., & Okamoto, K. (2011). Optineurin in neurodegenerative diseases. Neuropathology, 31(6), 569–574. doi: 10.1111/j.1440-1789.2011.01199.x

Reidling, J. C., Nabokina, S. M., Balamurugan, K., & Said, H. M. (2006). Developmental maturation of intestinal and renal thiamin uptake: Studies in wild-type and transgenic mice carrying human THTR-1 and 2 promoters. Journal of Cellular Physiology, 206(2), 371–377. doi: 10.1002/jcp.20492

Reidling, J. C., & Said, H. M. (2005). Adaptive regulation of intestinal thiamin uptake: Molecular mechanism using wild-type and transgenic mice carrying hTHTR-1 and -2 promoters. American Journal of Physiology-Gastrointestinal and Liver Physiology, 288(6), G1127–G1134. doi: 10.1152/ajpgi.00539.2004

Sechi, G., & Serra, A. (2007). Wernicke's encephalopathy: New clinical settings and recent advances in diagnosis and management. The Lancet Neurology, 6(5), 442–455.

Toppo, S., Vanin, S., Bosello, V., & Tosatto, S. C. (2008). Evolutionary and structural insights into the multifaceted glutathione peroxidase (Gpx) superfamily. Antioxidants & Redox Signaling, 10, 1501–1514.

Tröße, C., Waabø, R., Breck, O., Stavrum, A.-K., Petersen, K., & Olsvik, P. (2009). Genome-wide transcription analysis of histidine-related cataract in Atlantic salmon (*Salmo salar* L). Molecular Vision, 15, 1332–1350.

Vornanen, M. (2017). Electrical excitability of the fish heart and its autonomic regulation. In Fish Physiology (Vol. 36, pp. 99–153). doi: 10.1016/bs.fp.2017.04.002

Vuori, K. A. M., & Nikinmaa, M. (2007). M74 syndrome in Baltic salmon and the possible role of oxidative stresses in its development: Present knowledge and perspectives for future studies. AMBIO: A Journal of the Human Environment, 36(2), 168–172.

Zhao, H., Brown, P. H., Magone, M. T., & Schuck, P. (2011). The molecular refractive function of lens γ-crystallins. Journal of Molecular Biology, 411(3), 680–699. doi: 10.1016/j.jmb.2011.06.007

Zhao, R., Gao, F., Wang, Y., Diaz, G. A., Gelb, B. D., & Goldman, I. D. (2001). Impact of the reduced folate carrier on the accumulation of active thiamin metabolites in murine leukemia cells. Journal of Biological Chemistry, 276(2), 1114–1118. doi: 10.1074/jbc.M007919200

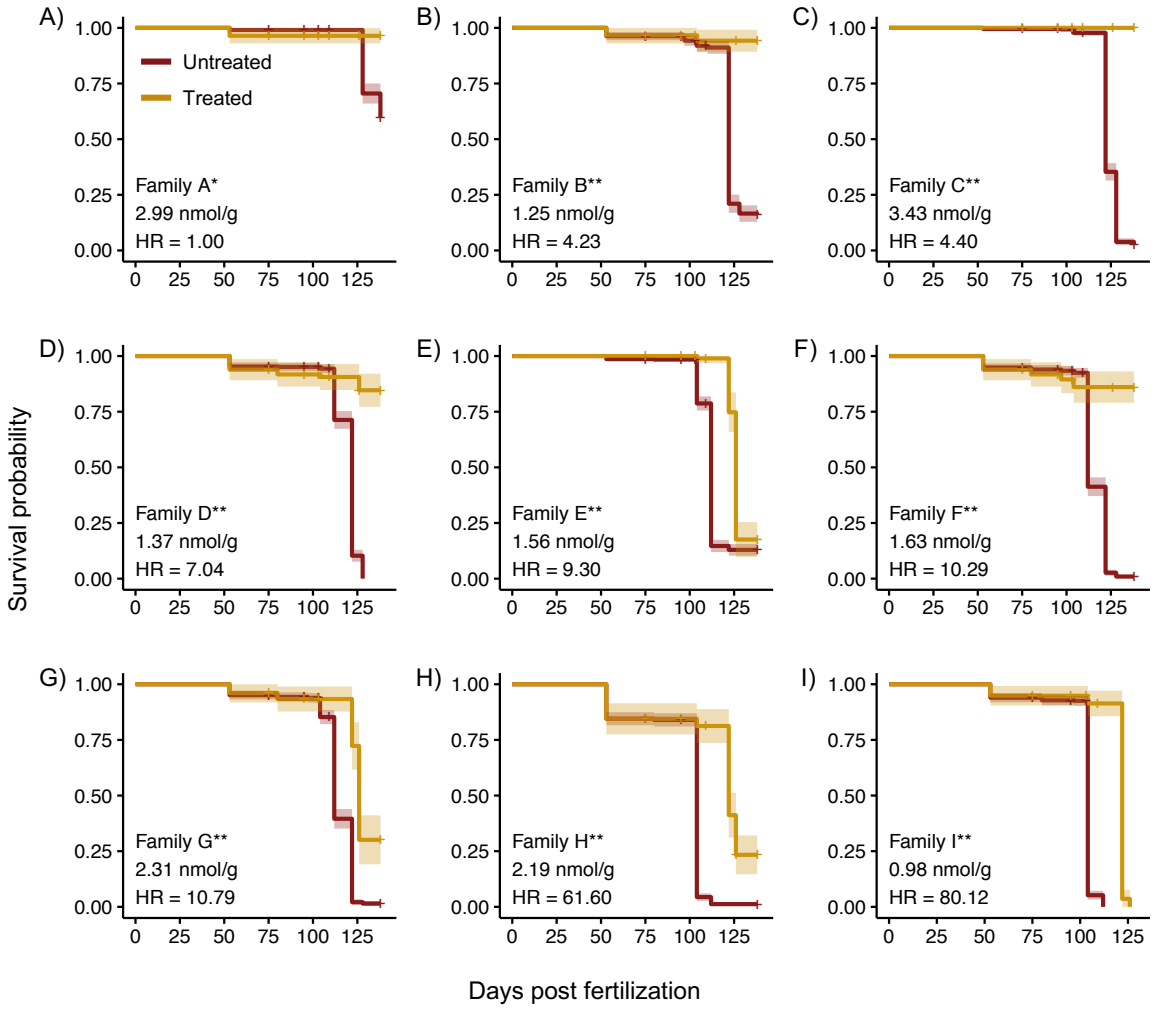


Figure A.1. Kaplan-Meier survival distributions for treated and untreated groups within each family A-I. Significant differences between treated and untreated survival distributions within each family denoted by * ($p < 0.01$) and ** ($p < 0.0001$). Total unfertilized egg thiamine concentration for each family is provided, as well as hazard ratios (HR) calculated using a Cox proportional hazards regression. All HR values were calculated using the Family A untreated group as the reference group. Hatch marks on survival distributions indicate censored individuals (i.e., samples removed for RNA-seq sampling or disease testing). The eye-up stage was reached at 50 days post-fertilization (dpf) and sampling for RNA-seq was conducted at 95 dpf.

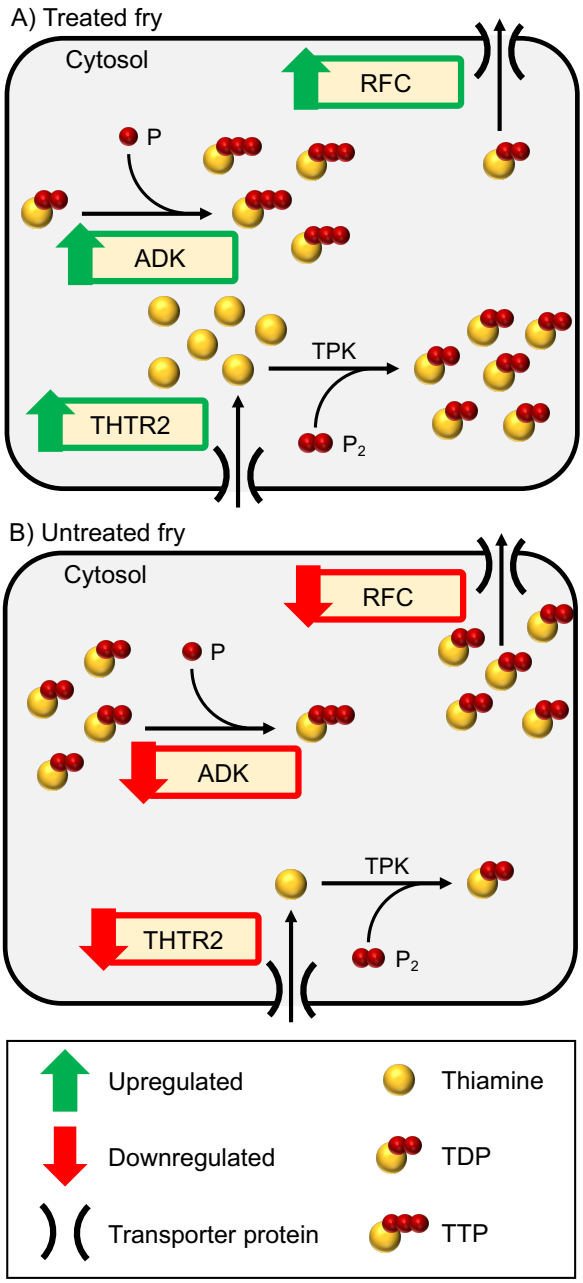


Figure A.2. Schematic describing how concentrations of thiamine diphosphate (TDP) may be impacted by thiamine status (treated vs. untreated) at the cellular level through differential expression of genes hypothesized *a priori* to be affected by thiamine deficiency. Direction of gene regulation in treated vs. untreated individuals is indicated by arrows next to gene names. THTR transports thiamine across the cell membrane, ADK interconverts forms of thiamine, and RFC exports thiamine derivates from the cell. ADK = adenylate kinase; P = phosphate group; THTR2 = thiamine transporter 2; TPK = thiamine pyrophosphokinase; TTP = thiamine triphosphate; RFC = reduced folate carrier.

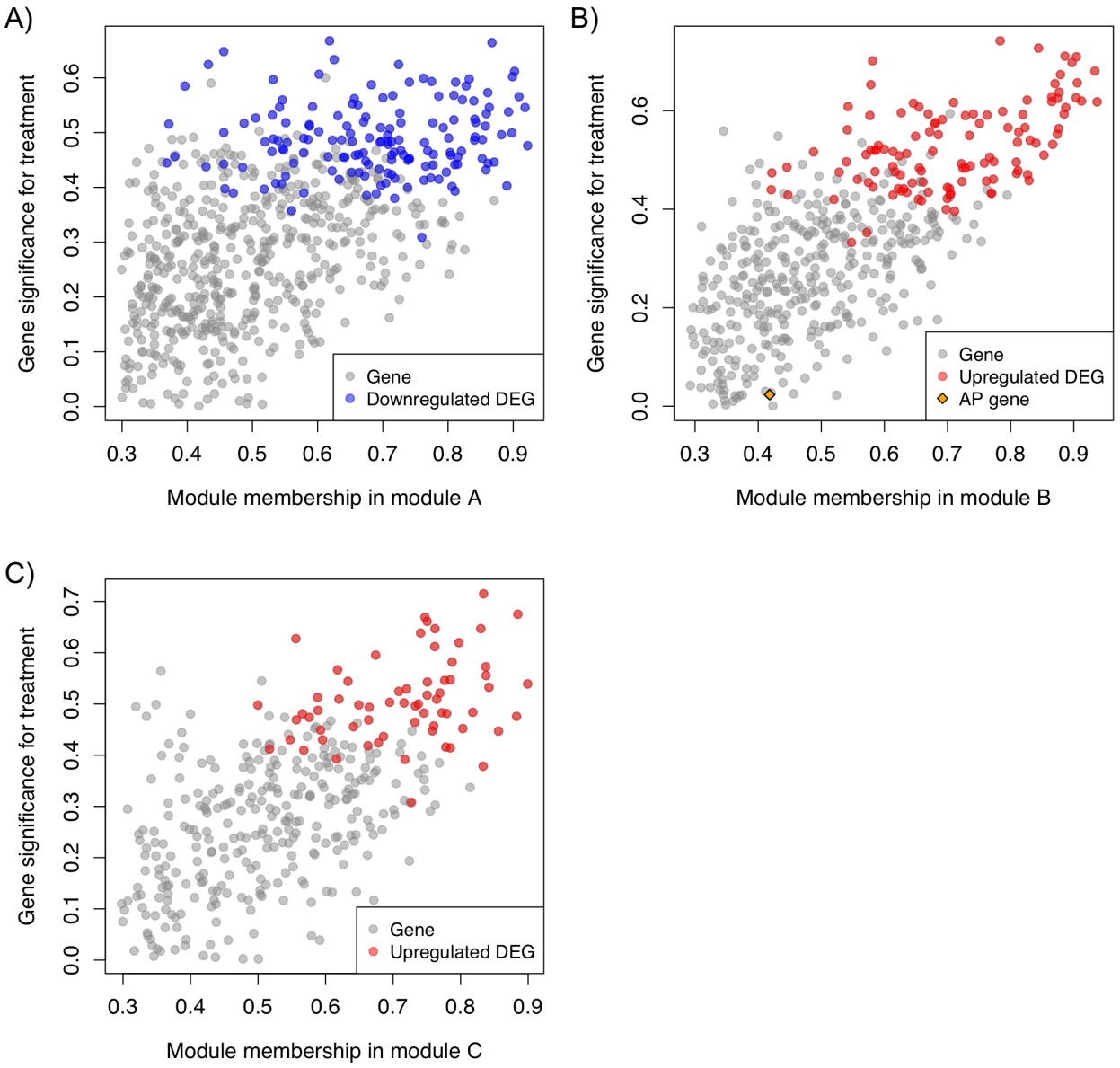


Figure A.3. Module membership vs. gene significance for each of the modules (A-C) that were significantly correlated with treatment status after Bonferroni correction. Module membership is a value calculated during *WGCNA* module construction that indicates how many connections (i.e., tight correlations in terms of expression) a gene has to other genes within that module. Genes that are significantly differentially expressed are highlighted in blue (downregulated) or red (upregulated). Genes hypothesized to be differentially expressed *a priori* (AP) are indicated by an orange diamond. DEG = differentially expressed gene.

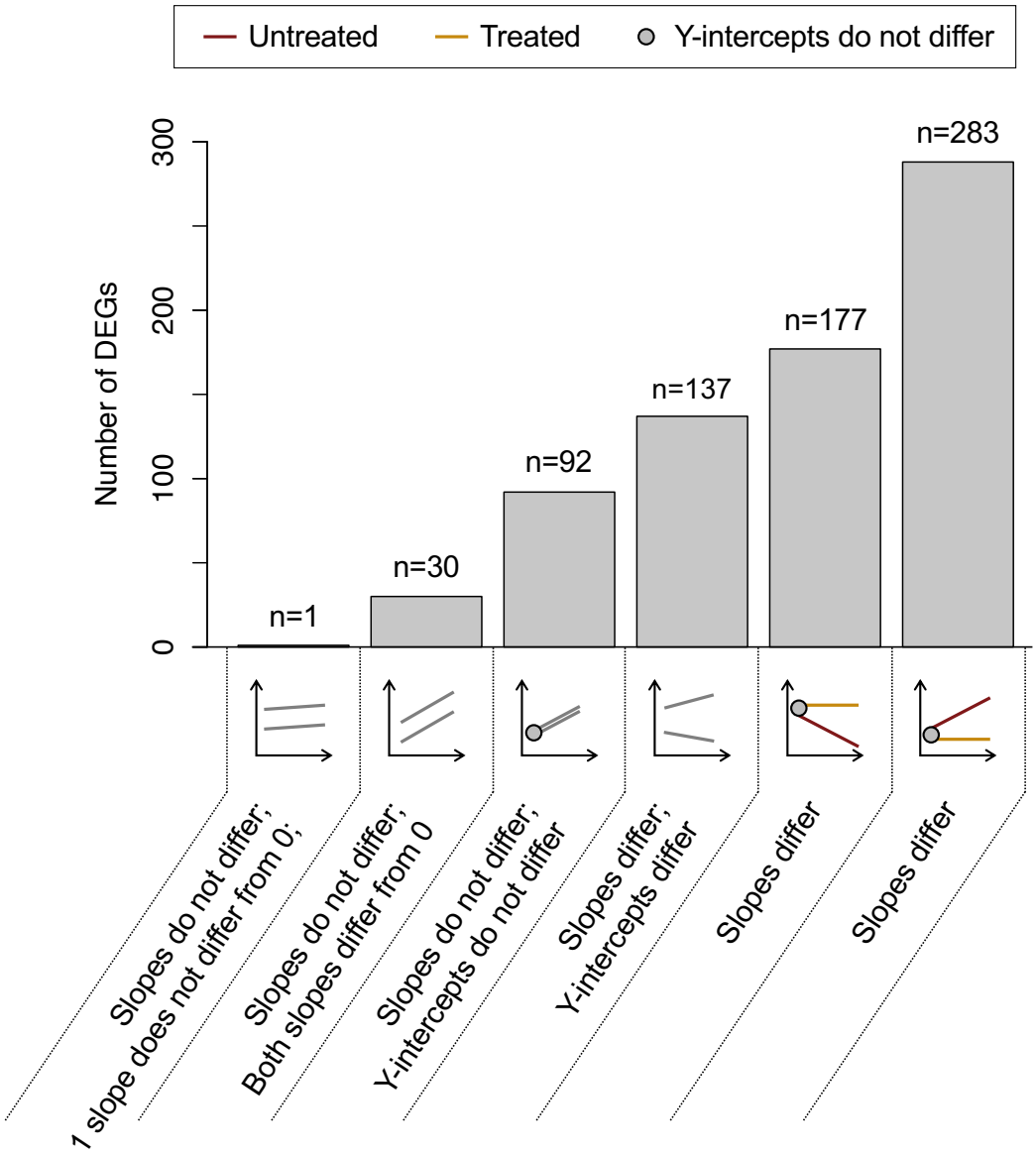


Figure A4. Six categories into which all family x treatment differentially expressed genes (DEGs) can be classified (n = 720). Grey lines can represent either treated or untreated regression lines. Numbers of DEGs in each category are provided, as well as an example plots and statistical description for each category. For example plots below the *x*-axis of the barplot, the *x*-axis is log(hazard ratio) and the *y*-axis is fragments per million mapped (FPM). These example plots are representative of the criteria used to define each category. Empirical examples of these types of plots may be found in Fig. 4A,B and Fig. 6A-C in the main text. Although all 720 genes represented here are genotype x environment DEGs, our stringent filtering of regression results identified 92 genes as having slopes and *y*-intercepts that do not differ despite being identified as differentially expressed with *DESeq2*.

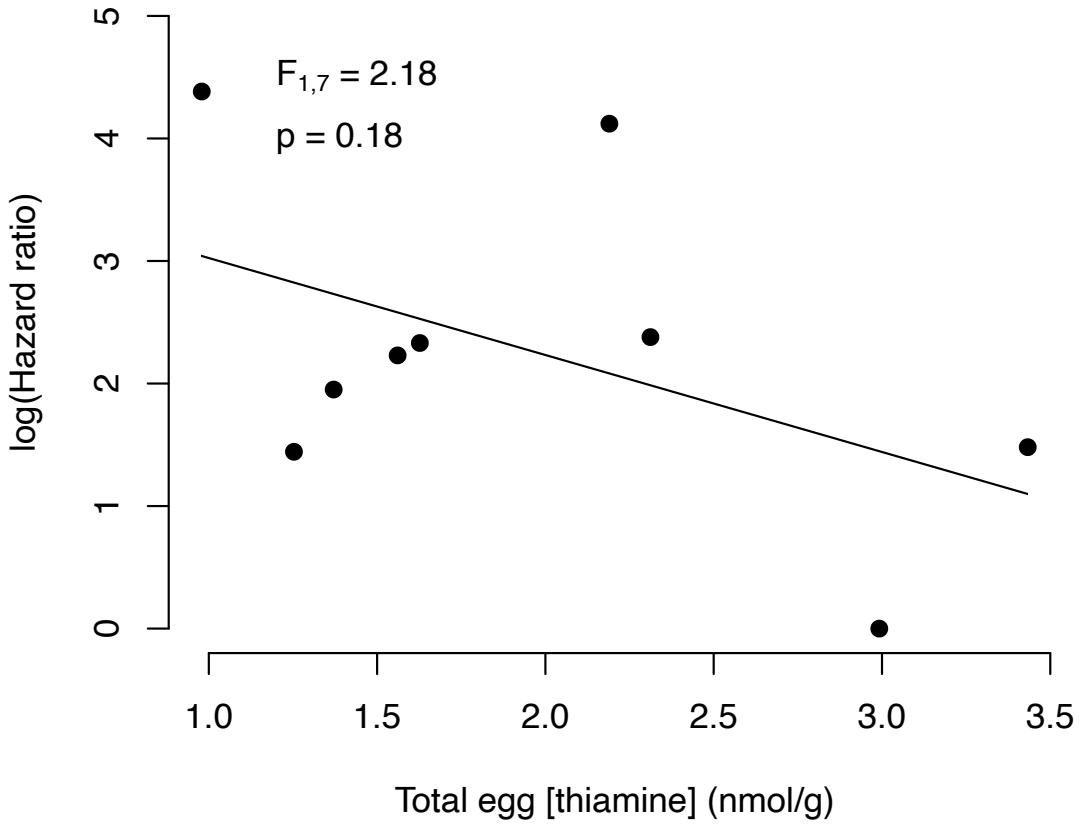


Figure A.5. Linear regression results demonstrating no relationship between total egg thiamine concentration (nmol/g) and hazard ratio ($\log_e$). This pattern indicates that individual survival is not simply a product of egg thiamine allocation.

Table A.1. Gene ontology (GO) terms associated with numbered terminal nodes of GO hierarchy trees in Figure 2.3. Terms marked with * are related to neurological function and development (Module A), metabolism (Module B), and cardiovascular system development and function (Module C). The central node in all networks represents the biological process level of the GO hierarchy.

Module	Number	GO term
A	1	amino acid transmembrane transport
A	2	arginine transport
A	3	endosomal transport
A	4	protein localization
A	5	phosphorylation
A	6	cAMP catabolic process
A	7	cGMP catabolic process
A	8	glycerol ether metabolic process
A	9	response to endogenous stimulus
A	10	positive regulation of Rac protein signal transduction
A	11	positive regulation of necrotic cell death
A	12	positive regulation of ruffle assembly
A	13	negative regulation of receptor-mediated endocytosis
A	14	regulation of cellular senescence
A	15	negative regulation of leukocyte chemotaxis
A	16	positive regulation of angiogenesis
A	17	positive regulation of miRNA metabolic process
A	18	activation of MAPKK activity
A	19	positive regulation of gene expression
A	20	establishment or maintenance of cell polarity
A	21	lateral inhibition
A	22	actin filament organization
A	23	neuromuscular junction development*
A	24	protein heterooligomerization
A	25	cell redox homeostasis
A	26	neuropeptide signaling pathway*
A	27	Ras protein signal transduction
A	28	cell death
A	29	cellular senescence
A	30	tube development
B	31	regulation of metabolic process*
C	32	midbrain development
C	33	thyroid gland development
C	34	embryonic hemopoiesis*
C	35	post-embryonic hemopoiesis*
C	36	nucleate erythrocyte development*
C	37	megakaryocyte development*
C	38	embryonic heart tube development*
C	39	endocardium formation*
C	40	cell morphogenesis
C	41	epidermis morphogenesis
C	42	erythrocyte maturation*
C	43	blood vessel maturation*
C	44	stem cell population maintenance
C	45	cellular response to cycloheximide
C	46	negative regulation of response to cytokine stimulus
C	47	positive regulation of cell division
C	48	positive regulation of protein complex assembly
C	49	regulation of rhodopsin mediated signaling pathway
C	50	negative regulation of phosphatidylinositol 3-kinase signaling
C	51	negative regulation of protein kinase B signaling

Table A.1 continued

C	52	negative regulation of cysteine-type endopeptidase activity involved in apoptotic process
C	53	regulation of stem cell population maintenance
C	54	positive regulation of endothelial cell differentiation
C	55	regulation of mast cell differentiation
C	56	positive regulation of hemoglobin biosynthetic process*
C	57	regulation of phosphatidylinositol 3-kinase activity
C	58	negative regulation of interleukin-6 production
C	59	negative regulation of heterotypic cell-cell adhesion
C	60	epidermal cell differentiation
C	61	type I pneumocyte differentiation
C	62	hemangioblast cell differentiation*
C	63	platelet formation*
C	64	mesodermal cell fate determination

APPENDIX B. CHAPTER 3 SUPPLEMENTAL MATERIALS

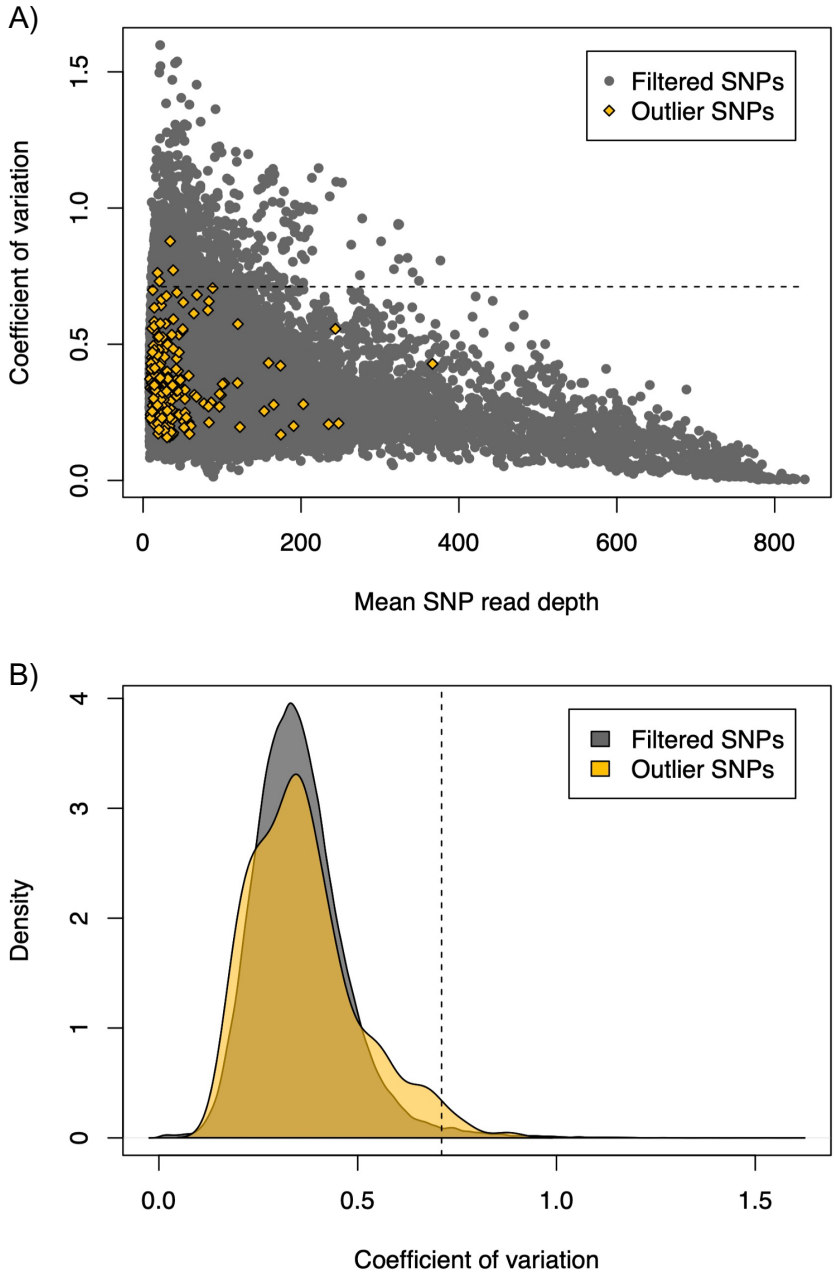


Figure B.1. A) Relationship between mean read depth and coefficient of variation (CV) in read depth across 10 pools of RNA sequenced for the Sebago Lake and LaHave River populations. Horizontal dashed line indicates mean CV + 3 standard deviations. Four outlier SNPs exhibited high CV and were not considered in downstream analyses. B) Kernel density plots of CV values for all filtered Sebago Lake-LaHave River SNPs (n=135,839) and for outlier SNPs (n=228).

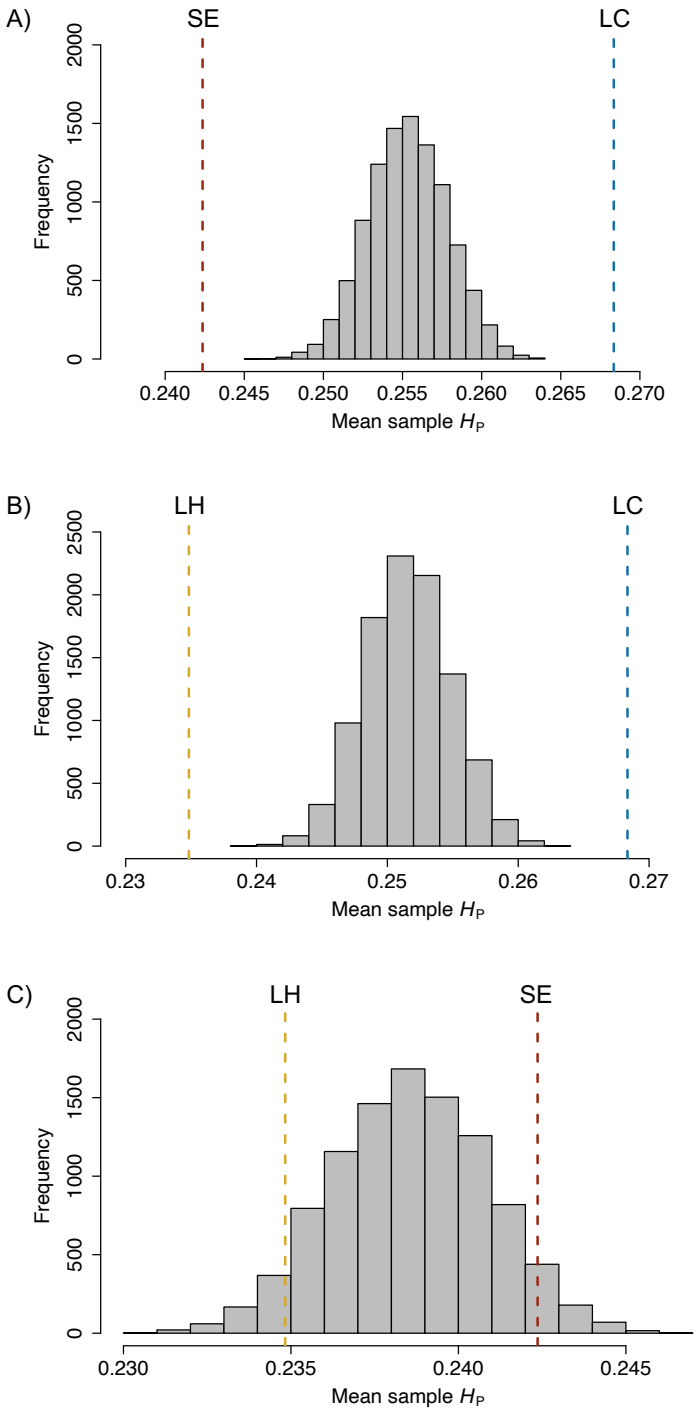


Figure B.2. Results of randomization test demonstrating that H_P differs significantly between Lake Champlain (LC) and both Sebago Lake (SE) and LaHave River (LH) (p<0.0001). For each population comparison, chromosome-specific H_P values were pooled for both populations, 29 values were drawn randomly, and the mean value was calculated; this process was repeated 10,000 times. Distributions are for mean sample H_P per comparison: (A) Lake Champlain vs Sebago Lake, (B) Lake Champlain vs. LaHave River, and (C) Sebago Lake vs. LaHave River. Vertical dashed lines represent genome-wide average H_P for each population.

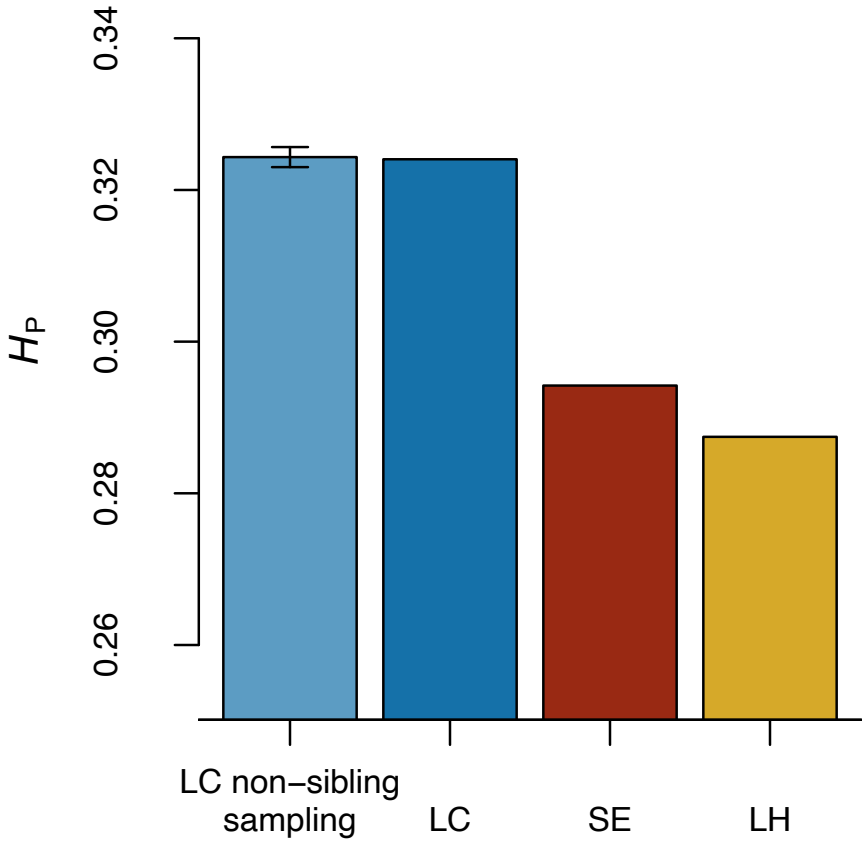


Figure B.3. Pooled heterozygosity (H_P) estimated for (i) 1,000 random, non-sibling samples (n=9; 1 sample per family) drawn from Lake Champlain individuals (mean +/- 1 standard deviation), (ii) Lake Champlain, (iii) Sebago Lake, and (iv) LaHave River.

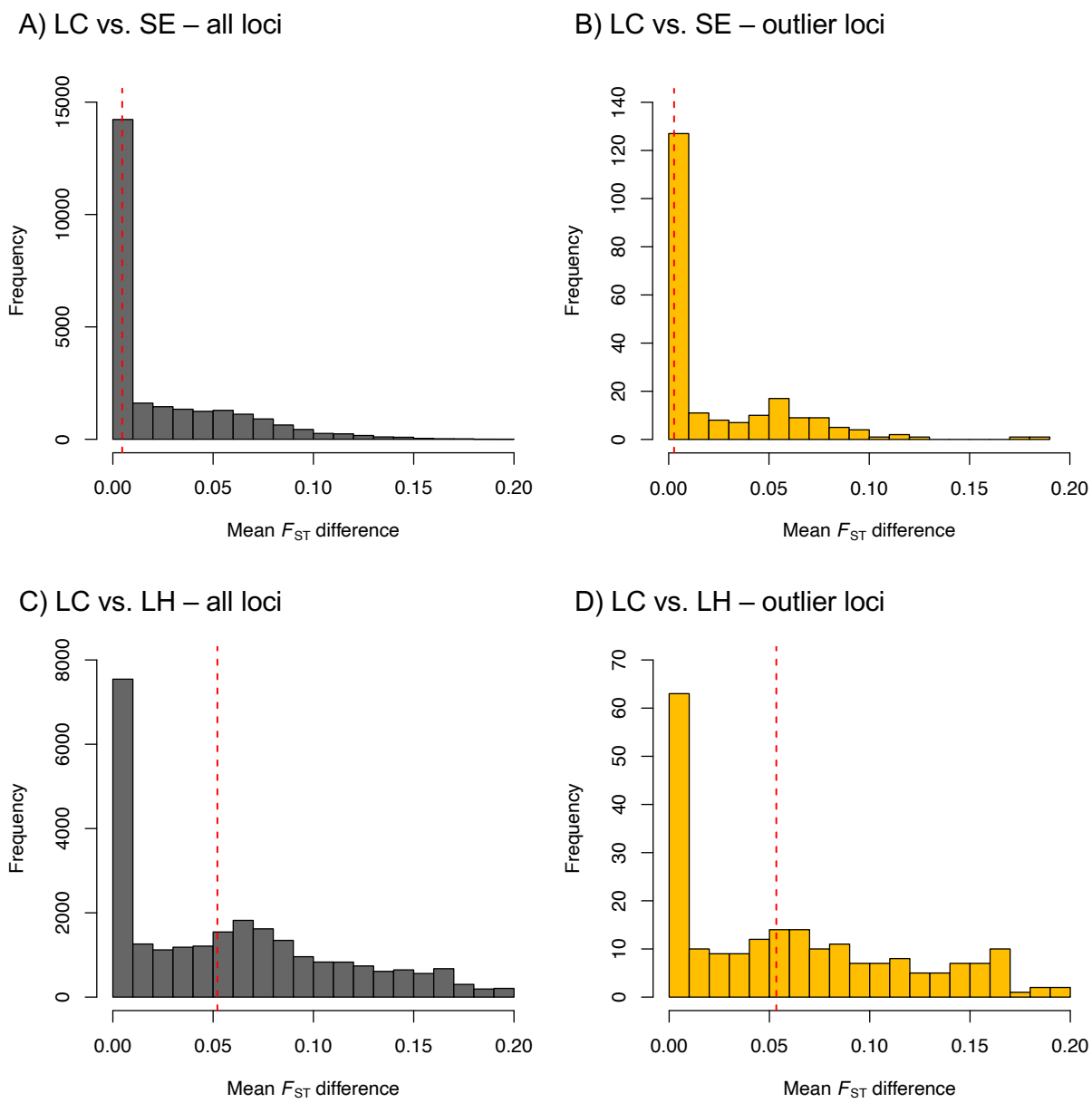


Figure B.4. Distribution of mean differences between pooled F_{ST} estimated for nine unrelated Lake Champlain individuals and for all 36 individuals. Pooled F_{ST} was estimated at each locus using 10,000 random draws of nine unrelated individuals, the absolute difference between each estimate and the 36-individual estimate was recorded, and finally, the mean difference was calculated for each locus. This process was conducted for both the Lake Champlain-Sebago Lake (A,B) and the Lake Champlain-LaHave River comparison (C,D), with results presented for all loci (A,C) and for 228 outlier loci (B,D). Vertical dashed line indicates median value for each distribution.

APPENDIX C. CHAPTER 4 SUPPLEMENTAL MATERIALS

Table C.1. Data collected from each of the six studies analyzed. Offspring stage refers to the life stage at which parental fitness was evaluated (i.e., offspring either returning as adults or young-of-the-year juveniles). Origin-specific numbers of potential parents are provided in the 'Natural and 'Hatchery' columns. Fitness values for each origin are provided in the '$w_{natural}$' and '$w_{hatchery}$' columns.

Study	Parent run year	Offspring stage	RRS	Natural	Hatchery	$w_{natural}$	$w_{hatchery}$
1	2004	adult	0.22	920	1,777	0.753	0.169
1	2005	adult	0.41	596	3,340	0.651	0.266
1	2006	adult	0.54	494	1,526	1.024	0.553
1	2007	adult	0.30	346	3,378	1.110	0.331
1	2008	adult	0.31	605	5,043	1.183	0.361
1	2009	adult	0.33	750	4,204	0.529	0.172
1	2010	adult	0.40	999	4,421	0.481	0.192
2	2004	adult	0.70	835	352	0.931	0.654
2	2005	adult	0.59	683	711	1.903	1.114
2	2006	adult	0.64	472	635	2.998	1.922
3	1995	adult	0.73	209	158	3.120	2.272
3	1996	adult	1.16	241	248	1.631	1.891
3	1997	adult	1.13	177	159	1.401	1.585
3	1998	adult	0.40	256	186	1.020	0.403
3	1999	adult	0.29	876	221	0.402	0.118
3	2000	adult	0.80	879	569	0.315	0.253
3	2001	adult	0.60	949	678	0.217	0.130
3	2002	adult	0.45	656	406	0.291	0.131
3	2003	adult	0.28	540	525	0.222	0.063
4	2002	juvenile	0.30	120	15	2.130	0.630
4	2003	juvenile	0.45	167	86	2.580	1.150
4	2004	juvenile	0.64	206	142	2.260	1.440
5	2000	adult	0.61	58	140	1.431	0.879
5	2001	adult	0.53	100	678	0.470	0.249
5	2002	adult	0.30	192	1,181	0.213	0.064
5	2003	adult	0.54	93	316	0.269	0.146
5	2004	adult	0.32	136	683	0.610	0.195
5	2005	adult	0.28	176	288	0.403	0.111
6	2002	adult	0.50	243	424	0.002	0.001
6	2003	adult	3.00	354	115	0.002	0.006
6	2004	adult	1.16	101	81	0.045	0.052
6	2005	adult	1.07	51	48	0.106	0.113

Table C.1 continued

6	2006	adult	0.72	37	47	0.544	0.392
6	2007	adult	0.81	71	73	0.088	0.071
6	2008	adult	0.96	220	166	0.023	0.022
6	2009	adult	0.64	138	244	0.014	0.009
6	2010	adult	0.91	359	398	0.011	0.010
6	2011	adult	1.50	188	113	0.012	0.018