EVIDENCE FOR THERMAL ADAPTATION AMONG GEOGRAPHICALLY, GENETICALLY AND THERMALLY DISTINCT POPULATIONS OF THE AUSTRALIAN BARRAMUNDI, *LATES CALCARIFER* (BLOCH 1790): A MULTI-LEVEL APPROACH

Thesis submitted by
Richard C. Edmunds BSc (Hon) Towson
on 15 May 2009
(revision submitted 09 September 2009)
(revision accepted 14 September 2009)

For the degree of Doctor of Philosophy
in Marine Biology
within the School of Marine & Tropical Biology
James Cook University: Townsville
(conferred *in absentia* by University Council: 2 December 2009)

STATEMENT OF ACCESS

I, the undersigned, author of this work, understand that James Cook University will make this thesis available for use within the University Library and, via the Australian Digital Theses network, for use elsewhere.

I understand that, as an unpublished work, a thesis has significant protection under the Copyright Act and;

I do not wish to place any further restriction on access to this work.

Or

I wish this work to be embargoed until:

Or

I wish the following restrictions to be placed on this work:

09/09/09

Richard C. Edmunds Date

STATEMENT OF SOURCES

DECLARATION

I declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

09/09/09

Richard C. Edmunds Date

STATEMENT ON THE CONTRIBUTION OF OTHERS

Nature of Assistance	Category	Titles and Affiliations of Co-Contributors
Intellectual support	Proposal writing	Drs. Lynne van Herwerden (JCU), Dean Jerry (JCU) and Carolyn Smith-Keune (JCU)
	Data Analysis	Drs. Lynne van Herwerden (JCU), Christopher Fulton (ANU), Carolyn Smith-Keune (JCU) and Dean Jerry (JCU)
	Statistical support	Drs. Christopher Fulton (ANU), Carolyn Smith-Keune (JCU) and Craig Syms (JCU)
Editorial assistance	Financial support	Drs. Lynne van Herwerden and Dean Jerry (JCU R.A.P.)
	Aquarium research	Norton Perry (USA), Russell Graham (ANU) and Dr. Christopher Fulton (ANU)
	Stipend	Norton Perry (USA)
Research assistance	Project design	Drs. Lynne van Herwerden (JCU), Dean Jerry (JCU) and Carolyn Smith-Keune (JCU)
	Data collection	Paolo Momigliano (JCU) and Drs. Lynne van Herwerden (JCU), Christopher Fulton (ANU), Carolyn Smith-Keune (JCU) and Dean Jerry (JCU)

09/09/09

Richard C. Edmunds Date

ACKNOWLEDGMENTS

Firstly, I'd like to thank my supervisors (Drs. Lynne van Herwerden, Dean Jerry and Carolyn Smith-Keune) for their fantastic support and guidance throughout the my Ph.D. candidature. Without them I would not have come to the full realization or potential of my scientific passion.

Secondly, I'd like to thank my family (Barbara, Daryn, Phil and Lynda Edmunds) for their endless support, both emotionally and financially, throughout my candidature here at James Cook University in Townsville, QLD *Australia*. You all played a central role in my achieving this life goal. For this I am forever grateful.

I'd also like to thank Dr. Christopher Fulton, Mr. Russell Graham and Norton Perry for their fantastic assistance in my experimental manipulations at The Australian National University's aquarium in Canberra, ACT (February through July 2007). Moreover, I'd like to express my thanks to all the BoZo post-graduate students (namely, Suz Amaphone, Andrew "Smiles" Thornhill and Richard "Kiwi" Carter) for their providing me with the most fantastic Canberra experience possible. I wish you all the best and look forward to future collaborations.

Last, but in no way least, I'd like to thank all my friends here at JCU Townsville. If it wasn't for you (Jason B, Scott P., Mat M., Kasia M., Marios A, Paolo M., Colin S., Steve M., Loic T., Emma B., Adrian M. and Adrian L., Stewart P., Matt "Wiz Kid" W., Pete "Irish" C., etc) my Townsville experience would have not attained the magnitude of madness that it did. To the highest degree attainable by words, I thank you all for the amazing memories and experiences.

On a different note, I'd like to expression my gratitude to the approx. 600 juvenile *Lates calcarifer* (barramundi) that were sacrificed for the attainment of scientific data (presented in Chapter 5). This data has served to far greater lengths than the kilojoules of energy obtained from their mere ingestion.

“All that you touch and all that you see,

All that you taste and all that you feel,

All that you love and all that you hate,

All you distrust and all you save,

All that you give and all that you deal,

All that you buy, beg, borrow or steal,

All you create and all you destroy,

All that you do and all that you say,

All that you eat and everyone you meet,

All that you slight and everyone you fight,

All that is now and all that is gone,

All that's to come and everything under the sun

Is in tune…

But the sun is eclipsed by the moon.”

“Eclipse”
Written by Roger Waters
From the album <u>Dark Side of the Moon</u> by Pink Floyd (1973)
Recorded at Abbey Road Studios, London (June 1972 to January 1973)

ABSTRACT

The ability of tropical fish to adapt and/or acclimate to their native thermal habitat is of biological interest, especially in the face of climate change. Evidence for this can be sought at the phenotypic, genotypic, transcript abundance and/or deduced protein sequence level. In this thesis a multi-level approach is employed to examine the question of thermal adaptation in a tropical fish. I use the Australian barramundi (*Lates calcarifer*) as a target species because genetically distinct populations occur throughout their tropical distribution range. The phenotype of *L. calcarifer* from two such divergent populations was investigated for a response to two ecologically relevant stressors, namely temperature and swimming challenge.. Subsequently investigation of genotype, transcript abundance and deduced protein sequence of the candidate gene lactate dehydrogenase-B (*ldh-b*) in two discrete Australian *L. calcarifer* populations were carried out. This candidate gene was chosen because of its known metabolic (aerobic) function and previous association with thermal adaptation and swimming performance in the temperate fish *Fundulus heteroclitus*. This thesis presents novel evidence for local adaptation of tropical Australian *L. calcarifer* to their native thermal habitats using a comparative and integrative multi-level approach, as follows:

Chapter 2 investigated at the organismal level whether juveniles (*n* = approx. 300) from two genetically discrete Australian *L. calcarifer* populations (Darwin, NT: 12° S, 130° E and Bowen, QLD: 20° S, 148° E; herein referred to as northern and southern, respectively) possess differences in mortality rate, growth rate and exhaustive swimming capacity following 28 day acclimation to both cold- (20°C) and heat-stress (35°C) conditions. This chapter uses swimming challenge as an ecologically relevant aerobic stress, applied in the presence and absence of thermal stress, to establish physiological or metabolic differences between these discrete *L. calcarifer* populations. Among-population differences were consistent with local thermal adaptation, with southern juveniles from the

cooler climatic region exhibiting significantly higher performance than their northern counterparts under cold-stress (20°C) conditions ($F_{1, 12}$ = 18.023; p = 0.001). Conversely, northern populations from the warmer climatic region performed significantly better than their southern counterparts under heat-stress (35°C) conditions ($F_{1, 12}$ = 13.948; p = 0.003). Mortality rates differed significantly from expectation at the two temperature stress conditions, with cold-stress resulting in significantly reduced mortality (1.4 %) for the southern population ($p < 0.05$) and heat-stress resulting in significantly increased mortality for both populations (53.2 and 17.6 % respectively), but particularly for the southern population ($p < 0.02$). These phenotypic differences indicate that discrete *L. calcarifer* populations may be locally adapted to their native thermal regime across their tropical Australian distribution and provides the basis for subsequent investigations of genotype, transcript abundance and deduced amino acid sequences for the candidate gene lactate dehydrogenase-b (*ldh-b*).

Chapter 3 investigates the *ldh-b* genotype in *Lates calcarifer* and congeneric *Lates niloticus* and compares the nucleotide (exons and introns) and deduced amino acid sequence of the *ldh-b* locus among these two tropical fishes. *Ldh-b* was 5,004 and 3,527 bp in length in *L. calcarifer* and *L. niloticus*, respectively, with coding regions comprising 1,005 bp in both species. A high level of sequence homology existed between species for both coding and non-coding regions (> 97% homology), corresponding to a 98.5% amino acid sequence homology. All six known functional sites within the encoded protein sequence (LDH-B) were conserved between the two *Lates* species. This chapter also identifies putative regulatory motifs and elements embedded within non-coding (intron) regions including 10 simple sequence repeat (SSR) motifs and 30 putative microRNA elements (miRNAs).. Five single nucleotide polymorphisms (SNPs) were also identified within putative miRNA regions. This *ldh-b* characterization provided valuable sequence for the subsequent population genetics and transcript abundance studies.

Chapter 4 uses a population approach to assess the level of nucleotide (exons and introns) and deduced amino acid sequence variation in the *ldh-b* locus both among and

within Australian *L. calcarifer* populations from different thermal regimes (n = 8). We identified a high homology of nucleotide and amino acid sequences among discrete populations lending further support to a hypothesis for selective constraint acting on the *ldh-b* locus in this tropical fish (Chapter 6). Interestingly, we found that the southernmost population sampled (Gladstone, Queensland, 23°S 151°E) differed significantly ($p < 0.05$) from all the other six populations screened with F_{ST} values ranging from 0.12 to 0.30, making it the ideal target population for future research on this thermally sensitive species. This chapter also reveals the presence of several intronic SNPs which may impact several overlapping putative miRNA elements and may ultimately influence the expression of the *ldh-b* locus itself or other loci of the transcriptome.

Chapter 5 investigates transcript abundance by quantifying the magnitude of variation in hepatic *ldh-b* transcripts among the two sampled populations (Darwin, NT and Bowen, QLD) following 28 day acclimations to cold-stress (20°C), heat-stress (35°C) and native control temperature (25 or 30°C) treatments. *Ldh-b* transcript abundance was also quantified in individuals subjected to swimming stress in control temperatures and in the presence of cold- or heat-stress, following acclimation to these temperatures. This chapter uses *L. calcarifer*-specific qRT-PCR primers and Sybr GreenER fluorescence based assays to quantify *ldh-b* transcript abundance. Fish from the southern and northern populations exhibited a significant increase in hepatic *ldh-b* transcript abundance ($F_{5, 41} = 6.459$; $p < 0.001$ and $F_{5, 39} = 3.866$; $p = 0.006$ respectively) in response to swimming stress at their respective native culturing temperatures (25°C and 30°C respectively) (i.e. in the absence of thermal stress). Fish from both populations exhibited a significant increase in hepatic *ldh-b* transcript abundance following heat-stress (35°C) acclimation compared to controls. Fish from the southern, but not the northern, population possessed significantly higher abundance of hepatic *ldh-b* transcripts following cold-stress (20°C) acclimation compared to controls. These observations suggest that the southern population possesses a unique transcriptional response to that of the northern population and that southern fish may be evolutionarily accustomed to greater variance in seasonal temperatures.

Chapter 6 uses an analytical approach, in lieu of empirical protein data, to identify evidence of selective constraint acting on *ldh-b* in two tropical congeneric *Lates* (*L. calcarifer* and *L. niloticus*) and *Plectropomus (P. leopardus* and *P. laevis)* species compared to one temperate congeneric species pair (*Fundulus*). In regard to both coding nucleotide (*ldh-b*) and amino acid (LDH-B) sequences, pairwise and phylogenetic comparisons identify that the two tropical species pairs are more homologous to one another (94 – 96%) than either is to the temperate species pair (90 – 92%). Observed pairwise species differences between coding and deduced amino acid sequences do not occur at known functional residues. In summary, this analytical approach identifies that selective constraint appears to be acting on both coding nucleotide and deduced amino acid sequences of this candidate gene.

Chapter 7 synthesizes the combined results of the comparative and integrative multi-level approach used in this thesis and presents novel ideas for future studies which would expand on the biological insight generated by this thesis and could provided further evidence for local adaptation of Australian *L. calcarifer* to their native tropical habitats.

TABLE OF CONTENTS

Appendices

LIST OF TABLES

LIST OF FIGURES

Chapter 1

General Introduction

In this thesis I use a classic multi-level comparative approach to establish evidence for local adaptation in geographically, genetically and thermally distinct populations of the tropical catadromous Latidae, *Lates calcarifer* (barramundi). This is done empirically by evaluating the impact of temperature and swimming stress, both alone and concurrently, at the organismal (phenotype) and molecular (genotype and transcript abundance) levels. Deduced amino acid sequences are also investigated for evidence of selective constraint via an analytical approach. This chapter provides relevant background for the choice of ecological stressors, study species and targeted candidate gene and concludes with a succinct summary statement regarding the chapter-specific aims of this thesis.

Swimming and temperature: two model ecological stressors

Globally, temperature is a key determinant of species distribution ranges and these may change under current climate change conditions (Perry et al., 2005). Studies investigating the sensitivity of species to both temperature increases and decreases relative to their native thermal regimes are of particular importance. Multi-level studies are important to identify past adaptation among discrete populations. Studies are required that examine not only the phenotype but also the genotype, transcript abundance and protein sequences of relevant candidate genes for different populations experiencing a range of temperature regimes.

The entanglement of aquatic organisms with their respective thermal habitat is one of the most studied biological questions (Hochachka and Somero, 2002). Specifically, temperature is known to increase the pH of salt-water by 0.0114 units per

1°C at 1 atmosphere of pressure (Gieskes, 1969) and inversely affect the availability of dissolved oxygen, as per Henry's Law (Powers, 1980). The thermal regime inhabited by aquatic organisms also impacts internal biochemical schemes, the most noteworthy of which are the Michaelis–Menten constant (K_m), an indicator of catalytic efficiency (e.g. for the conversion of lactate to pyruvate), enzyme-substrate binding affinity and enzyme stability (Fields and Somero, 1997; Hochachka and Somero, 2002; Somero, 1996). Temperature driven effects also impact individuals directly in several ways: Firstly, the observed phenotype (e.g. swimming performance, growth and mortality) may differ with temperature. Second and thirdly, the underlying genotype and/or transcript abundance of candidate genes may differ between individuals and/or populations from different temperature environments due to local adaptation.

Such evidence of thermal adaptation has been observed in the regulatory machinery (via one fixed mutation), transcript abundance and enzyme variants of the candidate gene lactate dehydrogenase-B (*ldh-b*) in the temperate killifish *Fundulus heteroclitus* (Crawford et al., 1999b; Crawford and Powers, 1992; Pierce and Crawford, 1997; Place and Powers, 1984a, 1984b; Rees et al., 2001; Schulte et al., 2000; Schulte et al., 1997). Temperature has also been shown to impact the transcriptome of channel catfish (*Ictalurus punctatus*), common carp (*Cyprinus carpio*), zebrafish (*Danio rerio*), annual killifish (*Fundulus heteroclitus, F. grandis* and *Austrofundulus limnaeus*) and rainbow trout (*Oncorhynchus mykiss*) (reviewed by Douglas, 2006). Thermal and/or swimming stress induced a significant transcriptomic response (two-fold or greater increase or decrease in transcript abundances) in examined fish species (Douglas, 2006). In light of these previous transcriptomic findings, long-term exposure of fish to fixed cold- or heat-stress conditions are considered to be two unique and pertinent ecological stresses that differ from naturally encountered diurnal and seasonal fluctuations in temperature.

Ecologically relevant cold- and heat-stress conditions also impact on the swimming performance of both temperate and tropical fishes (reviewed by Fulton, in press; Plaut, 2001). The effects of temperature and swimming stress on fish are often explored with key physiological metrics. One useful metric employed in fish studies is critical swimming speed or U_{crit} (*sensu* (Brett, 1964). Swimming speed or ability plays a central role in accomplishing essential tasks such as acquiring food, avoiding predators, defending territories and finding mates (Batty and Domenici, 2000; Blake, 2004; Webb, 1994). Moreover, swimming activities are often one of the largest daily costs in the energy budget of fishes (Boisclair and Sirois, 1993; Boisclair and Tang, 1993; Feldmeth and Jenkins, 1973), such that even small changes in swimming performance can reduce the energy available for development and growth, repair, disease resistance and reproduction (reviewed by Fulton, in press; Plaut, 2001). As such, examining the swimming performance of fishes can provide key insights into the factors that affect the health and evolutionary fitness of individuals. Indeed, measures of swimming speed performance have uncovered important insights on how different temperature stresses influence the health and survival of fish (reviewed by Fulton, in press; Plaut, 2001). These insights include how temperature impedes or accelerates the rate of early development of larval fishes (Fisher *et al.*, 2000), the attainment of food (Aseada *et al.*, 2005), avoidance of predators (Langerhans *et al.*, 2004; Masuda, 2006), out-performance of competing males and improved reproductive success (Kodric-Brown and Nicoletto, 2005), migration to new habitats (Hinch and Rand, 1998) and/or competition for optimal locations within native habitats (Fulton *et al.*, 2005; Kingsolver and Huey, 2008). Swimming trials conducted on the temperate killifish *Fundulus heteroclitus* established significant differences in critical swimming speed between thermally distinct northern and southern populations endemic to the eastern seaboard of North America and distributed across a steep thermal gradient (1°C per 1° latitude) (reviewed by Powers et al., 1991; Powers and Schulte, 1998). However, these

performance differences were not observed when similar trials were recently repeated (Fangue et al., 2008). The reason for these contradictory results is unclear but may be due to unidentified genetic or other changes in the source populations or to differences in experimental conditions (Fangue et al., 2008). To summarize, critical swimming performance, as established in the presence or absence of thermal stress, is an optimal metric to establish if tropical fish from different thermal regions of their distribution range possess phenotypic differences. Such phenotypic differences may result from differences in the genotype, transcript abundance and/or deduced protein sequence of a candidate gene. This is examined here using a multi-level approach.

Adaptation by natural selection: a multi-level approach

Evolution by natural selection is the process by which beneficial mutations persist in populations inhabiting particular environments due to the increased fitness (reproductive success) of those individuals that carry and transmit them (Carroll, 2000; Cossins and Crawford, 2005; Darwin, 1859). Conversely, non-beneficial (neutral) mutations may or may not persist purely by chance. Evolution, therefore, appears to be driven by the perpetual and complex interactions present between individuals (within populations) and the stressors of their native environment (e.g. temperature and swimming stress) by way of fitness enhancing modifications to their phenome, genome, transcriptome and/or proteome, if not all four (Douglas, 2006; Hochachka and Somero, 2002; Plaut, 2001; Powers and Schulte, 1998; Somero, 2005). More specifically, chance mutations that are beneficial within the genome of an individual are honed by natural selection at the level of the phenotype. This observed phenotype is ultimately the manifestation of the particular protein construct of an individual (i.e. proteome), as modified by the environment over evolutionary time. The intermediate link between the proteome and the genome is the transcriptome or particular complement of transcripts

present at any given moment in any given tissue of an organism. Messenger RNA transcripts (mRNA) carry information from the genome to the cellular machinery, such as the cytoplasmic ribosomes, in a tissue-specific manner (Latchman, 1991). These actively generated, yet transient, transcripts encode for specific proteins which are translated as and when required by cytoplasmic ribosomes. This intermediate, and perhaps most critical, function of the transcriptome is regulated twofold. The transcription of a particular locus from genomic DNA (gDNA) to mRNA can be activated or repressed by point mutations, such as single nucleotide polymorphisms (SNPs), in associated regulatory regions, generally located immediately up- or down-stream of the coding nucleotide sequence (Carroll, 2005; Latchman, 1991). Recent studies have also identified regulatory motifs embedded within intervening introns (Bartel, 2004; Lambowitz and Belfort, 1993; Li et al., 2004; Mattick, 1994). Secondly, the silencing of a transcript can also occur prior to its cytosolic translation by way of mobile regulatory elements such as microRNAs that generally bind to 3' untranslated regions (UTR) (Ambros, 2004; Baek et al., 2008; Bartel, 2004; Selbach et al., 2008). This binding of a miRNA element to the 3' UTR blocks translation of mRNA to protein since the 3' UTR is the usual targeted binding site of ribosomal translational machinery (Latchman, 1991). This twofold transcriptomic regulatory system increases the precision with which natural selection can optimize the fitness of an individual within its native habitat, as point mutations in key regulatory regions of the genome can consequently impact the transcriptome, proteome and/or phenome.

Therefore, to understand how different phenotypes are shaped by natural selection, investigation across all three molecular levels is required. All four levels are now amenable to characterization following experimental manipulations which mimic ecologically relevant environmental stressors, such as temperature and swimming stress. In light of this we can now ask "Do alterations to the genotype, transcript abundance

and/or protein sequences allow species to exist in widely varying environmental conditions, such as the temperatures experienced in temperate or tropical regimes?" and if so, "Do these changes correspond to changes observed in the phenotype?" As such, studies investigating the impact of global climate change on biodiversity should utilize a comparative and integrative multi-level approach, as presented in this thesis, whenever possible.

Lates calcarifer as a model tropical study species

Lates calcarifer is a catadromous, protoandrous hermaphrodite denizen to rivers, estuaries and shallow marine environments of the Asian archipelago (13°N – 10°S), Papua New Guinea and tropical Australia (25°S – 12°S) (Figure 1.1). Within Australia, this species exhibits a high level of genetic population structure, possibly as a result of both the non-interconnecting nature of Australian coastal rivers (Figure 1.1) and the apparent non-migratory nature of this species (Davis, 1986, 1987; Griffin, 1987b; Russell and Garrett, 1985). Despite *L. calcarifer* being considered a eurythermal species, very little is actually known about the underlying biological mechanisms relating to how it tolerates a broad range of temperatures (15 to 40°C *sensu* (Katersky and Carter, 2007). Curiously, most studies on optimal growth temperatures to date have utilized fish originating from only a single genetic stock (Katersky and Carter, 2005, 2007). Regardless, *L. calcarifer* is the basis of a large and rapidly growing aquaculture industry in Australia and overseas (e.g. USA, Israel, Singapore, Thailand, India, the Netherlands, etc.). Within Australia genetically distinct stocks have been utilized as broodstock by geographically dispersed hatcheries (Rimmer, 1995) and may have accumulated thermal adaptations to their unique riverine habitats. Broodstock sourcing of this nature may be useful in the aquaculture of this tropical species, but to date the thermal tolerance differences of different stocks accessed by industry has not been

rigorously tested. As such, this thesis presents the first study which investigates phenotypic differences between genetically discrete *L. calcarifer* populations across all three molecular levels.

Given that *L. calcarifer* is commercially farmed, and thus successfully reared in captivity, it is ideal for use in scientific experimentation due to the ready availability of juveniles from discrete aquaculture facilities which source broodstock locally. This local sourcing of broodstock ensures that juveniles obtained from geographically distinct farms are in fact representative of local populations which may be adapted to their local thermal regime. Juveniles are ideal for experimental manipulation for four key reasons: 1) Their availability in large number and of similar sizes; 2) Their lower susceptibility to the impacts of cold- and/or heat-stress and thus they are more likely to survival under thermal extremes than adult fish (Meeuwig et al., 2004); 3) The adults are too large for aerobic challenges in our purpose-built swimming chamber, as described in (Fulton, 2007); and 4) Juveniles do not require an energy expenditure balance between growth and reproduction. In summary, the geographically, genetically and thermally distinct nature of Australian *L. calcarifer* populations, in conjunction with the local sourcing of broodstock by aquaculture facilities, makes this species an ideal target for a comparative and integrative multi-level approach regarding the local adaptation of tropical fish to their native thermal habitats.

Ldh-b: an ideal candidate gene for linking phenotype variation to genotype and/or transcriptional variation

Lactate dehydrogenase-B (*ldh-b*) is an ideal candidate gene to employ when investigating the impact of temperature and swimming stress on a tropical fish species such as *L. calcarifer* for several reasons. Firstly, the *ldh-b* locus encodes a glycolytic enzyme (LDH-B) which plays a critical role in maintaining aerobic metabolism by converting lactate, the major by-product of anaerobic glycolysis catalyzed by LDH-A,

to pyruvate via oxidation or directly to glucose via gluconeogenesis (Place and Powers, 1984a, 1984b). The lactate generated in the heart and skeletal muscles by LDH-A is converted back to pyruvate or glucose in the liver by LDH-B (Figure 1.2). This function of LDH-B allows desired and/or essential aerobic metabolic activity (e.g. swimming) to be sustained for extended lengths of time in fish (Powers et al., 1991; Powers and Schulte, 1998). Recent U_{crit} swimming trials on the rainbow trout *Oncorhynchus mykiss* highlight the importance of lactate conversion during such trials in that maximum swimming speed attained following repeated swimming was limited by the accumulation of serum lactate exclusively (Jain and Farrell, 2003).

In addition to this metabolic function LDH-B also affects oxygen binding affinity of hemoglobin (Hb) by altering intra-erythrocyte ATP concentrations of Hb in fish (Powers, 1980; Powers *et al.*, 1979). More specifically, in *Fundulus heteroclitus* the warm southern allozyme ($LDH\text{-}B^{a}$) reduces, while the cold northern allozyme ($LDH\text{-}B^{b}$) increases the abundance of ATP in hemoglobin, which inversely affects the $Hb\text{-}O_2$ binding affinity (Powers, 1980; Powers et al., 1979). This effect of LDH-B on $Hb\text{-}O_2$ binding affinity directly impacts delivery of Hb-bound oxygen to red muscle tissues and may therefore be an alternate mechanism by which LDH-B affects the swimming capacity of fish (DiMichele and Powers, 1982; Fangue et al., 2008; Powers et al., 1979). Swimming trials conducted on *F. heteroclitus* established significant differences in critical swimming speed (U_{crit}) between thermally distinct cold northern and warm southern populations (DiMichele and Powers, 1982). Following this observed difference in phenotype between genetically and thermally discrete populations the *ldh-b* locus was extensively investigated in the cold northern and warm southern *F. heteroclitus* populations (reviewed by Powers et al., 1991; Powers and Schulte, 1998).

The extensive investigation of *ldh-b* in *F. heteroclitus* revealed one non-fixed and two fixed single nucleotide polymorphisms (SNPs) in the coding nucleotide sequence, with all three mutations causing non-synonymous amino acid substitutions and thus unique LDH-B allozymes (reviewed by Powers et al., 1991; Powers and Schulte, 1998). In addition to these coding nucleotide and amino acid sequence differences, the level of *ldh-b* transcription also varied two-fold between cold northern and warm southern *F. heteroclitus* populations (Segal and Crawford, 1994b). This variation in *ldh-b* transcript abundance was linked directly to a one base pair mutation in the regulatory element identified within the 5' untranslated region (UTR) of the *ldh-b* locus in *F. heteroclitus* which resulted in a higher basal level of *ldh-b* transcripts in the cold northern population, even after acclimation to warmer water (Crawford et al., 1999a; Rees et al., 2001; Schulte et al., 2000; Schulte et al., 1997; Segal et al., 1996).

These reported differences in the *ldh-b* locus and LDH-B enzyme were concluded to be adaptive and likely generated by the unique selective pressure imposed by thermally distinct habitats of the *F. heteroclitus* populations examined (reviewed by Powers et al., 1991; Powers and Schulte, 1998). The *ldh-b* locus was also investigated in discrete populations of the fat-head minnow *Pimephales promelas* (Merritt, 1972), rainbow trout *Salmo gairdneri* (Kao and Farley, 1978) and crested blenny *Anoplarchus purpurescens* (Johnson, 1977); however, the extent of these studies and insights achieved are not as comprehensive as those conducted over two decades on *F. heteroclitus* (Powers, *et al*. 1979 – 2000). For additional information regarding the impact of thermal stress on *ldh-b* and/or LDH-B in *F. heteroclitus* (e.g. the impact of LDH-B isozymes on Hb-O_2 affinity and it's link to aerobic swimming performance; the proximal promoter and glucocorticoid responsive element (GRE) sequences and their link to *ldh-b* gene expression; mRNA and enzyme abundances in response to thermal and aerobic stresses, etc) see the literature review provided in Appendix 2.

In light of these previous investigations this thesis targets *ldh-b* as an ideal candidate gene for a comparative and integrative multi-level approach used to establish evidence for adaptation of Australian *L. calcarifer* populations to their native tropical habitats. To achieve this I compare the *L. calcarifer ldh-b* genotype with its congener *L. niloticus* as well as across discrete Australian populations and among other tropical (*Plectropomus*) and temperate (*Fundulus*) species (Chapters 3, 4 and 6 respectively). Noteworthy is that the full characterization of this locus generated the sequence information required for the design of species-specific quantitative real-time PCR (qRT-PCR) primers (Chapter 3). These qRT-PCR primers permitted the detection of population-specific transcriptomic differences (i.e. hepatic *ldh-b* transcript abundance) using Sybr GreenER fluorescence based assays (Chapter 5). The candidate gene approach utilized in this thesis also provides evidence of the usefulness of this approach when investigating the impact of ecological stressors on populations and/or species across multiple biological levels.

Thesis summary of the comparative and integrative multi-level approach:

Chapter 2 investigates the phenotype to establish whether discrete Australian populations of juvenile barramundi, *Lates calcarifer*, exhibit phenotypic differences in mortality, growth rates and swimming performance following 28 days acclimated to both cold-stress (20°C) and heat-stress (35°C) conditions. ***Chapter 3*** investigates the *ldh-b* genotype in a comparative manner to establish the level of variation inherent in this locus among the congeneric tropical fishes *Lates calcarifer* and *L. niloticus* in regard to nucleotide (exons and introns) and deduced amino acid sequences, as well as identifying putative regulatory motifs embedded within non-coding (intron) regions (e.g. SSRs and miRNAs). ***Chapter 4*** uses a population genetics approach to assess nucleotide (coding and non-coding), amino acid (deduced) and putative regulatory element (miRNA) sequence variation of the *ldh-b*

locus within and among thermally discrete *L. calcarifer* populations. ***Chapter 5*** investigates transcript abundance by quantifying the magnitude of variation in hepatic *ldh-b* transcripts in those juveniles subjected to ecological stressors, following acclimation, in Chapter 2. ***Chapter 6*** uses an analytical approach, in lieu of empirical protein data, to establish the presence of selective constraint on the *ldh-b* locus within two congeneric pairs of tropical fish as well as among these tropical fish and a congeneric temperate species pair. ***Chapter 7*** synthesizes the combined results of the organismal (phenotype) and molecular (genotype, transcript abundance and deduced protein sequences) level investigations to conclude that there is strong evidence for local adaptation of discrete tropical Australian *L. calcarifer* populations to their respective thermal habitats. This concluding chapter also presents directions for future research.

In addition to the primary chapters summarized above, I have included relevant peripheral work that I have undertaken, which provides additional depth and context to the main thesis. ***Appendix 1*** utilizes a comparative genomics approach to assess the level of variation in *ldh-b* among the tropical fish *Plectropomus leopardus* and *P. laevis* in regard to nucleotide (exons and introns) and deduced amino acid sequences, as well as the identification of putative regulatory motifs embedded within non-coding (intron) regions (e.g. SSRs and miRNAs). This information was required for the work presented in chapter 6. ***Appendix 2*** reviews the known response of *ldh-b* and LDH-B, as well as other loci and enzymes, to various ecological stressors (e.g. pH, temperature and swimming) in the killifish *F. heteroclitus* (Powers, *et al*. 1979 – 2000).

FIGURES

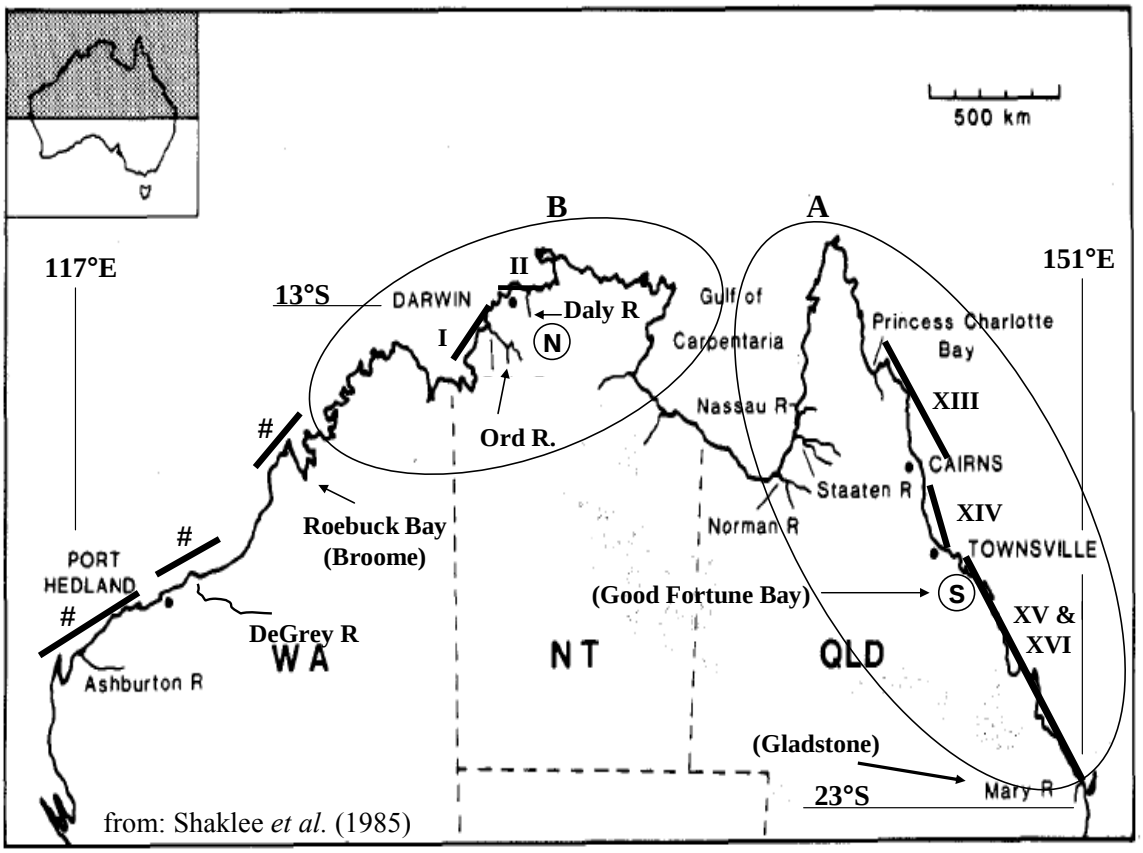


Fig. 1.1. Map of *L. calcarifer* distribution across tropical Australia (adapted with permission from Shaklee *et al*. 1985). Capital letters (A, B), hash marks (#) and roman numerals (I, II, XIII, XIV, XV and XVI) indicate genetic clades (Chenoweth et al., 1998), (Doupé et al., 1999) and (Keenan, 1994) established by mitochondrial markers and allozyme assays, respectively. Northern and southern experimental populations (Chapters 2 and 5) obtained from aquaculture facilities located in Darwin, NT and Bowen, QLD, and used in the work of this thesis are indicated by N and S respectively.

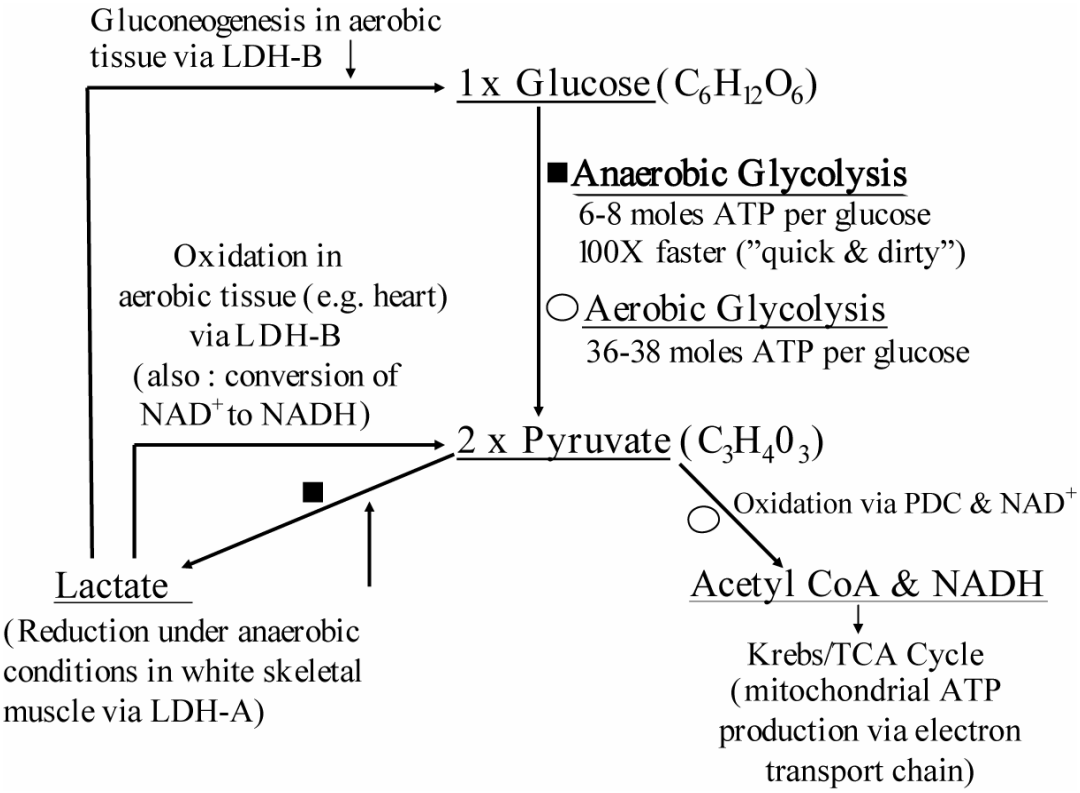


Fig. 1.2. Glycolysis summary highlighting the respective roles of LDH-B and LDH-A isozymes in aerobic and anaerobic pathways, respectively (Powers and Schulte, 1998).

Chapter 2

Effect of native and non-native thermal acclimation on the swimming performance of the tropical catadromous barramundi, *Lates calcarifer* (Bloch 1790)

2.1: ABSTRACT

We investigated evidence for thermal adaptation in the catadromous tropical species barramundi, *Lates calcarifer*, which is widely distributed across tropical Australia. Using the key physiological metric of critical swimming speed (U_{crit}), we assessed performance variation in discrete *L. calcarifer* populations after early growth and long-term acclimation to both native (25°C and 30°C) and non-native (20°C and 35°C) temperatures. Among-population differences were consistent with local thermal adaptation, with southern juveniles exhibiting significantly higher performance than their northern counterparts under cold-stress (20°C) ($F_{1,\ 12}$ = 18.023; p = 0.001). Conversely, northern populations performed significantly better than their southern counterparts under heat-stress (35°C) ($F_{1,\ 12}$ = 13.948; p = 0.003). Body, liver and heart mass were high in the 30°C and 25°C control treatments for the northern and southern populations, respectively ($F_{9,\ 576}$ = 28.936; p < 0.001). Mortality differed significantly between the two temperature stress conditions, with cold stress resulting in significantly reduced mortality for the southern population (p < 0.05) and hot stress resulting in significantly increased mortality for both, particularly the southern population (p < 0.02). We conclude that tropical *L. calcarifer* have adapted to their respective thermal environments, since mortality and performance are significantly altered at non-native temperatures, even after acclimation to those temperatures.

2.2: INTRODUCTION

Organisms can respond via changes in their morphology, physiology or survivorship to the temperature regime of their native habitat. Ectothermic, or cold-blooded, organisms are particularly responsive to their home-range temperatures due to the direct impact that temperature fluctuations (e.g. daily, seasonally or climatically) have on their internal temperatures and metabolic function. Temperature differences are also known to manifest at the population level in endothermic species, such as faster development, smaller size and shorter maturation times in warm water populations versus cold water populations (i.e. *Hotter is smaller* and *Hotter is better* rules (*sensu* (Kingsolver and Huey, 2008). These rules, or observed trends, are thought to be most pronounced in aquatic organisms due to the correlated decrease in available oxygen with increases in water temperature (Henry's Law), indicating that hotter conditions impose both thermal and hypoxic stresses (Powers, 1980).

Exploring temperature effects in ectothermic organisms is often achieved using key physiological metrics. One metric that has found great utility in the study of fishes is swimming performance, an attribute that plays a central role in accomplishing essential tasks such as acquiring food, avoiding predators and finding mates (Batty and Domenici, 2000; Blake, 2004; Webb, 1994). Swimming activities also constitute one of the largest daily costs in the energy budget of fishes (Boisclair and Sirois, 1993; Boisclair and Tang, 1993; Feldmeth and Jenkins, 1973), such that even small changes in swimming performance can have a major impact on the energy available for growth, repair and reproduction. As such, examining the swimming performance of fishes can provide key insights into the factors that affect the health of individuals and populations, as well as the evolutionary fitness of individuals (Huey and Stevenson, 1979). Indeed, measures of swimming speed performance, such as critical swimming speed performance (abbreviated

U_{crit} (*sensu* (Brett, 1964), have provided important insights into how different temperature stresses are linked to variations in their health and survival (reviewed by Fulton, in press; Plaut, 2001). Specific examples of these insights include how the environment has influenced the rate of early development of larval fishes (Fisher et al., 2000), the attainment of food (Aseada et al., 2005), avoidance of predators (Langerhans et al., 2004; Masuda, 2006), out-performance of competing males and improved reproductive success (Kodric-Brown and Nicoletto, 2005), migration to new habitats (Hinch and Rand, 1998) and/or competition for optimal locations within native habitats (Fulton et al., 2005; Kingsolver and Huey, 2008).

Establishing the critical swimming speed (U_{crit}) of a particular fish species has provided some valuable insights into the health and fitness requirements needed by individuals (both inter- and intra-specific) to occupy particular habitats (Carey and Franklin, 2009; Claireaux et al., 2006; DiMichele and Powers, 1982; Fulton, 2007; Fulton et al., 2005; Heap and Goldspink, 1986; Plaut, 2001). While a powerful metric, the accurate establishment of U_{crit} performance requires an appropriate acclimation period (e.g. ranging from 10 to 90 days) to any non-native temperature prior to swimming assessment to allow the adjustment of muscle fibers to either higher or lower water temperatures (Beitinger and Bennett, 2000; Beitinger, Bennett and McCauley, 2000; Claireaux *et al.*, 2006; Heap and Goldspink, 1986). In addition, variations in growth rate often occur between fish acclimated to native habitat temperatures and those acclimated to non-native temperatures (i.e. thermal stress treatments) (Beitinger and Bennett, 2000; Beitinger et al., 2000; Fangue and Bennett, 2003; Hokanson et al., 1977). Consequently, the measurement of swimming speed performance should be reported in a size-adjusted manner, which involves the minimization of size-effects via residualisation or a reference to body lengths (e.g. BL s^{-1}), as opposed to an absolute speed (centimeters per second, cm s^{-1}) that does not reflect such allometric effects (Claireaux et al., 2006; Fulton, 2007; Plaut, 2001).

In this study we used the critical swimming speed performance to determine the physiological fitness of juvenile barramundi, *Lates calcarifer* taken from populations native to two markedly different thermal environments (Central Queensland and Northern Territory, Australia). *L. calcarifer* is a catadromous hermaphrodite residing in rivers, estuaries and shallow marine environments at different stages of its lifecycle (Chenoweth et al., 1998; Keenan, 1994). It occurs throughout the Asian archipelago (13°N – 10°S) and the rivers of northern Australia (25°S – 12°S), with significant genetic partitioning (i.e. lack of gene flow) being found between each major river system in Australia (Chenoweth et al., 1998; Keenan, 1994). Due to such genetic partitioning, each barramundi population is likely to have accumulated unique adaptations to their native river. Therefore, population-specific responses to environmental stress (e.g. temperature, salinity, pH, flow rates, etc.) are expected in this widespread species. Furthermore, as this species is successfully reared in captivity, which is the basis of the aquaculture industry, barramundi is a model species for the investigation into such population-specific responses to temperature and/or swimming performance.

We tested the physiological response of two distinct populations of *L. calcarifer* at mean native water temperatures, as well as more extreme hot and cold variations that correspond with minimum winter and maximum summer water temperatures in the wild (20°C and 35°C, respectively). Using U_{crit} , morphological measurements and mortality rates from each population, we tested the following hypotheses: 1) *L. calcarifer* from southern and northern native populations exhibit significantly different U_{crit} performances in the two extreme thermal stress treatments (20°C and 35°C, respectively) used in this study; 2) *L. calcarifer* performance is optimal at the native acclimation temperature and lower at all other temperatures; and 3) acclimation to all non-native treatment temperatures imposes thermal stress on both *L. calcarifer* populations and manifests in increased mortality and decreased body and vital organ (liver and heart) masses.

2.3: METHODS

Study species

Juvenile *L. calcarifer* (approx. 90 days from hatching) individuals were obtained directly from two hatcheries within Australia, the Darwin Aquaculture Centre (hereafter referred to as northern) and Kuranda Fish Farm (hereafter referred to as southern) which are located in the Northern Territory and Northern Queensland and source their broodstock from rivers in the Darwin (12° 27.7' S, 130° 50.5' E) and Bowen (20° 0.7' S, 148° 14.8' E) regions, respectively (see Figure 1.1). The first batch of similar sized fingerlings was obtained from the northern hatchery on 21 March 2007, with an average total length (TL) of 60.4 ± 0.4 mm. Fingerlings were air-freighted in slightly brackish water (5 ppt) at an average temperature of 26.7 °C ± 1°C; zero mortality occurred during transportation. Fingerlings were immediately transferred to twelve 200 liter aquaria (distributed at an equal density of 1 individual per 11.8 L) housed in temperature-controlled rooms at the Australian National University (ANU) Aquarium Facility. Prior to the arrival of fingerlings all tanks were set to a temperature of 30.0 ± 1°C and a 14:10 light:dark regime. Each tank's filtration was maintained by air-powered, biologically-activated under-gravel filters. Fingerlings were fed twice daily with a set quantity (approximately 4 g) of 1.5mm pellets (Skretting Inc., Cambridge, Tasmania) per aquarium for the first 19 days, followed by 4 g of 2mm pellets thereafter.

The second batch of *L. calcarifer* was obtained from the southern hatchery on 30 May 2007, with an average TL of 61.9 ± 0.4 mm. Fingerlings were air-freighted in slightly brackish water as before, but at an average temperature of 22.0 ± 1°C; zero mortality occurred during transportation. As with the northern fingerlings, all tanks were set to a uniform 14:10 light:dark regime, although initial tank temperatures were set to 25.0 ± 1°C to match the lower initial temperature used in the hatchery and

southern brood stock's natural habitat. All southern fish were distributed among the same (unaltered) tanks at the same densities as for the northern fish and were, therefore, acclimated under the exact same conditions and feeding regime as the northern individuals (details above).

Temperature treatments

Each temperature treatment (35°C, 30°C, 25°C & 20°C) was applied to three separate tanks, with a total of 51 fish per treatment, divided into three replicate groups of 17 fish each (except 20°C treatment, $n = 52$). For the northern individuals, all tanks were held at 30.0 ± 1°C for 4 days to allow for general acclimation. Deviation from the general-acclimation temperature to the specific-acclimation temperature began on Day 5 at an increase (toward 35°C) or decrease (toward 25°C & 20°C) of approximately 1°C per day. The 30°C treatment was considered to be at the required population specific acclimation temperature on Day 1. Once at the required acclimation temperature all tanks of a given treatment were held at their respective temperatures for 28 days before swimming performance testing. The average temperature was held at the required acclimation temperature for 28 days. This was done across all three tanks per treatment as follows: 35 ± 1°C, 30 ± 1°C, 25 ± 1°C and 20 ± 1°C. Identical treatment conditions were applied for the southern individuals, except that they were initially acclimated to 25°C for the first 4 days before the 1°C d^{-1} change towards the specific-acclimation temperatures for each treatment (35°C, 30°C, 20°C). Control (25°C treatment) southern fish were considered to be at the population specific-acclimation temperature on Day 1 (Table 2.1).

Table 2.1 Overview of temperature treatments used for the two populations of *L. calcarifer*

Population	Control / Intermediate	Cold-stress	Heat-stress
Northern	30°C / 25°C	20°C	35°C
Southern	25°C / 30°C	20°C	35°C

Mortality

Individuals that displayed erratic behavior (e.g. loss of equilibrium) synonymous with viral infection (Noda-virus) were immediately removed, euthanized and excluded from the study. Otherwise, all other mortality during the experiments was assumed to occur as part of the thermal-induced physiological challenge of the different treatments. For all X^2 comparisons (Jonassen et al., 1999) the expected mortality rate for juvenile fish under thermal stress is considered to be 0.36% day^{-1} for the first 20 days and 0.18% day^{-1} thereafter (Hokanson et al., 1977).

Swimming speed performance

Juvenile *L. calcarifer* were acclimated at their specific temperature in still-water tanks for a minimum of 28 days and fasted 12-h prior to swimming performance trials designed to assess critical swimming speed (U_{crit}). Pre-U_{crit} fasting was conducted to ensure a post-absorptive state that maximizes metabolically available energy (Fulton, in press; Plaut, 2001) and minimizes variations in gut fullness between individuals as this is known to impact swimming performance (Billerbeck et al., 2001). All swimming performance trials were conducted on fishes from each of the four temperature treatments using a 194L recirculating flow tank, as described in (Fulton, 2007).

Each measure of U_{crit} involved placing two size-matched individuals, each randomly-selected from two of the three different holding tanks for each treatment temperature, into the working section of the flow tank (55.0 x 11.5 x 24.0 cm) and swimming them at an initial acclimation speed of approximately 1.5 total lengths per second (TLs^{-1}) for a period of 30mins. If no adverse behavior was observed in the initial acclimation period, the flow speed was then increased in increments of approximately 0.33 TLs^{-1} every 20 minutes until the trial end-point, which was taken as when the fish lost position in the working section and impinged on the downstream grid for more than

5 seconds. For each individual the U_{crit} value was calculated using the formula (following (Brett, 1964):

$$U_{crit} = U + [U_i (t \times t_i^{-1})]$$

where U = penultimate speed, t = time swum in final velocity increment, t_i = set time interval for each velocity increment (20min) and U_i = velocity increment (2.6cm s^{-1}, 3.0cm s^{-1}, 3.3cm s^{-1}, 3.6cm s^{-1}, 4.0cm s^{-1}or 4.3cm s^{-1} for fish with an estimated total length of 8, 9, 10, 11, 12 or 13cm, respectively).

Upon completion of the trial the test subjects were euthanized (via 5% clove-oil in 15% ethanol and salt water solution at 35 ppt, followed by ice-water slurry) and measured for total length, maximum body depth, maximum body width and mass. This allowed for the calculated adjustment of U_{crit} for solid blocking effects following (Fulton, 2007), with the average adjustment for solid blocking being a +0.49% increase in U_{crit} across the four treatments. This procedure was repeated for eight fish from each of the four temperature treatments for both populations (northern and southern), with the selection of test subjects taken randomly from the three aquaria housing fish for each temperature treatment.

An additional experiment was run on a unique group of southern fish (approx. 50) acclimated to the cold stress temperature of 20°C, but aerobically (swimming) challenged at all four treatment temperatures to examine the effect of temperature acclimation on U_{crit} performance. This involved subjecting individuals to U_{crit} trials (as per methodology described above) at a range of different temperatures above their acclimated temperature (20°C). Six individuals were trialed at 20°C, which allowed us to assess if U_{crit} performance had varied over the additional 5 days of acclimation from when the previous batch of 20°C individuals were tested as part of the previous procedures. U_{crit} trials were then conducted on the 20°C southern fish placed directly

into 25, 30 and 35°C water (n = 6 per temperature) in the flow tank and tested for their U_{crit} performance.

All swimming performance trials were conducted during daylight hours (5:30-19:30) and water temperatures in the flow tank were maintained at the same level as the treatment temperature for that specific fish (with the exception of the final experiment detailed in the paragraph above). Overall, a total of 204 juvenile *L. calcarifer* were tested over 137.7h of U_{crit} swimming trials. Following each individual being euthanized, all organs were removed, snap frozen (liquid nitrogen) and later weighed to nearest 0.01g and 0.001g (liver and heart, respectively).

Statistical analysis

Whenever a significant relationship between size and swimming speed was detected, all U_{crit} swimming speeds were converted to body lengths per second (BL s^{-1}) to minimize any allometric effects among treatments (Claireaux et al., 2006; Fulton, 2007; Plaut, 2001). Mortality rates were assessed for deviation from the expected via χ^2 tests, as per Jonassen *et al*. (1999). Two-way analysis of variance (ANOVA) was conducted on swimming data, with population and temperature considered fixed factors, to assess among population variation (Table 2.2a). Individual univariate ANOVA analysis was performed to establish Bonferroni corrected significant swimming performance differences between *L. calcarifer* populations within each temperature treatment (n = 4; alpha level = 0.0125; Table 2.2b). Within population differences in swimming performance were assessed via univariate ANOVA for Bonferroni corrected significance (alpha level = 0.025), with temperature as the fixed factor and subsequent Tukey's HSD post-hoc analysis (Table 2.2c). The unique 20°C acclimated southern treatment group was treated as a separate univariate ANOVA analysis and also utilized Tukey's HSD post-hoc analysis. A two-way multivariate analysis of variance

(MANOVA) was conducted on morphological data, with body, liver and heart mass as variables and population and temperature as fixed factors (Table 2.3). Two-way analysis of variance (ANOVA) was conducted on body, liver and heart mass separately, with population and temperature considered fixed factors, to assess among population variation at the Bonferroni corrected significance level (alpha = 0.017; Table 2.4, 2.5 and 2.6 respectively). Univariate ANOVA analyses were subsequently conducted on body, liver and heart mass as for the swimming analyses above (alpha = 0.0125; Table 2.4, 2.5, and 2.6 respectively). All analyses were performed using SPSS (Version 16.0).

2.4: RESULTS

Mortality

Mortality rates of 1.4% to 53.2% across the four treatments were recorded during the 28 day acclimation and subsequent swimming performance trials, constituting a total of 20 and 54 mortalities from the northern and southern populations, respectively (Figure 2.1). Observed mortality in the northern 20°C, 25°C and 30°C treatments did not differ significantly from the expected mortality (χ^2 test, df = 1, $p > 0.4$), whereas mortality in the northern 35°C treatment (9 fish) was significantly higher than expected (χ^2 test, df = 1, $p < 0.05$; Figure 2.1). Conversely, observed mortality of southern fish in the lower temperature treatments (20°C and 25°C) was significantly lower than expected (χ^2 test, df = 1, $p < 0.05$), whilst mortality of southern fish in the higher temperature treatments (30°C and 35°C) was significantly higher than expected (χ^2 test, df = 1, $p < 0.02$; Figure 2.1). Several mortalities from both northern and southern 35°C (n = 2 and 5 respectively) and 30°C (n = 1 and 2 respectively) treatments involved fish euthanized following the exhibition of erratic behavior synonymous with viral infection (e.g. Noda-

virus). The remaining mortalities cannot be directly linked to viral infection as these fish were deceased prior to observation.

Critical swimming performance (U_{crit}) and thermal stress (°C)

A significant relationship between critical swimming speed (U_{crit}) in terms of centimeters per second (U_{crit} cm s^{-1}) and total length (cm) was observed across all populations and treatment groups ($r^2 = 0.66$; $p < 0.001$; Figure 2.2a). However, when U_{crit} was adjusted to a scale of total lengths per second (TL s^{-1}) the significant relationship between total length and critical swimming speed was removed ($r^2 = 0.002$; $p > 0.5$; Figure 2.2b). Consequently, length-corrected U_{crit} values were used for all other analyses for each population and treatment level. When considering among-population differences in critical swimming performance (U_{crit}), a significant effect of temperature and the population x temperature interaction were observed; however the effect of population alone was not significant (Table 2.2a). Specifically, the 20°C (cold-stress) southern treatment group significantly outperformed (U_{crit} = 4.14 ± 0.14 BL s^{-1}) the 20°C (cold-stress) northern treatment group (U_{crit} = 3.27 ± 0.14 BL s^{-1}) (Table 2.2b; Figure 2.3c). Indeed, the 20°C (cold-stress) northern treatment group exhibited an average U_{crit} (3.27 ± 0.14 BL s^{-1}) that was the second lowest performance of all treatment groups assessed under all temperature treatments and populations examined (Figure 2.3c). The only treatment group to register a lower value was the unique 20°C acclimated southern treatment group when swum at 35°C (i.e. without acclimation to this higher temperature), with a U_{crit} performance of 2.19 ± 0.17 BL s^{-1} (Figure 2.4). Conversely, the 35°C (heat-stress) northern treatment group demonstrated a significantly higher average U_{crit} performance (4.57 ± 0.14 BL s^{-1}) than the 35°C (heat-stress) southern treatment group (3.79 ± 0.14 BL s^{-1}) (Table 2.2b; Figure 2.3c). By contrast, no significant differences in swimming performance were observed among

populations for the intermediate (25 and 30°C) treatment groups, with the average U_{crit} values for northern and southern treatment groups being 4.31 ± 0.14 and 4.37 ± 0.14 BL s^{-1} for 25°C and 4.34 ± 0.13 and 4.06 ± 0.13 BL s^{-1} for 30°C, respectively (Table 2.2b; Figure 2.3c). Significant within-population differences in swimming performance were also observed for both the northern and southern treatment groups (Table 2.2c and Figure 2.3). Within the northern population no significant differences in swimming speed performance were observed between the 25°C, 30°C and 35°C treatment groups (U_{crit} = 4.31 ± 0.15 BL s^{-1}, 4.34 ± 0.14 BL s^{-1} and 4.57 ± 0.15 BL s^{-1}, respectively); however, swimming performance was significantly lower in the 20°C treatment group compared to all other treatment groups (U_{crit} = 3.27 ± 0.15 BL s^{-1}) (Table 2.2c; Figure 2.3a). Similarly, no significant differences in swimming speed performance were observed between the 20°C, 25°C and 30°C southern treatment groups (U_{crit} = 4.14 ± 0.12 BL s^{-1}, 4.37 ± 0.12 BL s^{-1} and 4.06 ± 0.11 BL s^{-1}, respectively) yet a significant difference in swimming performance was observed between the 25°C and 35°C southern treatment groups (U_{crit} = 4.37 ± 0.12 BL s^{-1} and 3.79 ± 0.12 BL s^{-1}, respectively) (Table 2.2c; Figure 2.3b).

Table 2.2 Analyses of variance in swimming performance (BL s^{-1}) among & within populations

	Factor	Partial Eta Squared	*F*	df	*p*
a	Population	0.002	0.112	1, 50	0.739
	Temperature	0.329	8.158	3, 50	< 0.001
	Population x Temperature	0.437	12.961	3, 50	< 0.001
b	20°C	0.6	18.023	1, 12	0.001
	25°C	0.013	0.155	1, 12	0.701
	30°C	0.128	2.058	1, 14	0.173
	35°C	0.538	13.948	1, 12	0.003
c	Northern Treatment Group	0.638	14.699	3, 25	< 0.001
	Southern Treatment Group	0.319	3.895	3, 25	0.021

a: Two-way ANOVA with population and temperature as fixed factors; *b*: One-way ANOVA with population as fixed factor; *c*: One-way ANOVA with temperature as fixed factor.

The final and unique set of within-population swimming performance trials assessed southern fish which were acclimated to 20°C for an additional 7 days (35 days total) prior to U_{crit} assessment in each of the four treatment temperatures, without acclimation to increased temperatures (i.e. 25°C, 30°C and 35°C; see Methods). This unique 20°C acclimated southern treatment group also demonstrated significant within-population differences in swimming performance ($F_{3, 16} = 53.225$, $p < 0.001$; Figure 2.4). More specifically, this unique 20°C acclimated southern treatment group did not exhibit significantly different swimming speed performances between the 25°C and 30°C challenged treatment groups (4.62 ± 0.17 and 4.87 ± 0.17 BL s^{-1}, respectively; Figure 2.4). Only when swum at 35°C did this unique 20°C acclimated southern treatment group demonstrate a significantly lower U_{crit} performance (2.19 ± 0.17 BL s^{-1}; Figure 2.4). Moreover, this unique 20°C acclimated southern treatment group did not demonstrate a substantial difference in U_{crit} performance compared to that of the 20°C southern treatment group acclimated for the experimental standard of 28 days (4.15 ± 0.12 and 4.14 ± 0.14 BL s^{-1}, respectively) despite having 7 additional days for growth and acclimation. Furthermore, when this unique 20°C acclimated southern treatment group was swum at 25°C they did not out- or under-perform the 25°C southern treatment group (4.62 ± 0.12 and 4.37 ± 0.14 BL s^{-1}). When this unique 20°C acclimated southern treatment group was swum at 30°C, however, an unexpected out-performance of the 30°C southern treatment groups was observed (4.87 ± 0.23 BL s^{-1} versus 4.06 ± 0.13 BL s^{-1}).

Morphological responses to thermal stress

When considering the among-population differences across all morphometric variables (body, liver and heart mass), significant effects of population, temperature and the population x temperature interaction were observed (Table 2.3; Figure 2.5). However,

when considering the among-population differences in body mass, only temperature and the population x temperature interaction had significant effects (Table 2.4a; Figure 2.5c). More specifically, the average body mass of acclimated fish was significantly larger in the northern population for the 20°C (9.1 ± 0.32g) and 25°C (19.9 ± 0.70g) treatment groups than in the southern 20°C (7.3 ± 0.33g) and 25°C (11.1 ± 0.72g) treatment groups (Table 2.4b; Figure 2.5c). Conversely, for the intermediate (30°C) and the heat-stress (35°C) treatments the final average mass of the southern treatment group (22.2 ± 0.85g and 21.4 ± 1.21g, respectively) was significantly larger than the average mass of the northern treatment group (18.1 ± 0.69g and 16.1 ± 1.05g, respectively) (Table 2.4b; Figure 2.5c). As expected, the among-population variation in total length reflected those observed for body mass above in that population ($F_{1, 244} = 5.204$, $p = 0.023$), temperature ($F_{3, 244} = 221.395$, $p < 0.001$) and the population x temperature interaction ($F_{3, 244} = 48.688$, $p < 0.001$) were all significant (Figure 2.6c). Specifically, the average total length of acclimated fish was significantly larger in the northern population for the 20°C (8.3 ± 0.1 cm) and 25°C (11.6 ± 0.2 cm) treatment groups than in the southern 20°C (7.7 ± 0.1 cm) and 25°C (9.4 ± 0.2 cm) treatment groups ($F_{1, 64} = 16.746$, $p < 0.001$ and $F_{1, 64} = 107.953$, $p < 0.001$, respectively; Figure 2.6c). The average total length of acclimated fish was significantly larger in the southern population for the 30°C (12.1 ± 0.1 cm) and 35°C (11.6± 0.2 cm) treatment groups than in the northern 30°C (11.4 ± 0.1 cm) and 35°C (10.4 ± 0.2 cm) treatment groups ($F_{1, 62} = 13.011$, $p = 0.001$ and $F_{1, 54} = 14.205$, $p < 0.001$, respectively), a pattern which also mirrors the patterns of significance among-populations for body mass (Figure 2.6c).

Table 2.3 Two-way MANOVA of morphometric variables among *L. calcarifer* populations

Factor	Pillai's Trace	*F*	df	*p*
Population	0.848	352.296	3, 190	< 0.001
Temperature	0.934	28.936	9, 576	< 0.001
Population x Temperature	0.654	17.851	9, 576	< 0.001

Within-population differences in body mass were also observed to be significant in both the northern and southern treatment groups (Table 2.4c and Figure 2.5). The average body mass of 20°C northern treatment group (9.1 ± 0.73g) was significantly less than the average body mass of fish from all other northern treatment groups (19.9 ± 0.70g, 18.1 ± 0.73g and 16.1 ± 0.75g for 25, 30 and 35°C, respectively) (Figure 2.5a). The 25°C northern treatment group also exhibited an average body mass (19.9 ± 0.70g) significantly greater than that of fish from the hottest (35°C) northern treatment group (16.1 ± 0.75g); however, fish from the 30°C and 35°C northern treatment groups did not differ significantly from one another (Figure 2.5a). Within the southern population, the 20°C treatment group also exhibited an average body mass (7.3 ± 0.75g) significantly less than the average body mass observed for all other southern treatment groups (11.1 ± 0.71g, 22.2 ± 0.88g and 21.4 ± 0.84g for 25, 30 and 35°C, respectively) (Figure 2.5b). The 25°C southern treatment group exhibited an average body mass (11.1 ± 0.71g) significantly lower than the average for the 30°C and 35°C southern treatment groups (22.2 ± 0.88g and 21.4 ± 0.84g, respectively), which did not differ significantly from one another (Figure 2.5b). As observed in the among-population differences, the significant within-population differences observed for total length of the northern ($F_{3, 122}$ = 30.378, $p < 0.001$) and southern ($F_{3, 122}$ = 154.695, $p < 0.001$) treatment groups also reflected the patterns of significance seen for body mass above. Specifically, the average total length of the 20°C northern treatment group (9.0 ± 0.31g) was significantly less than the average body mass of fish from all other northern treatment groups (12.3 ± 0.31g, 11.1 ± 0.29g and 13.3 ± 0.31g for 25, 30 and 35°C, respectively) which did not differ significantly from one another (Figure 2.6a). The 20°C southern treatment group also exhibited an average total length (8.1 ± 0.31g) significantly less than the average total length observed for all other southern treatment groups (10.1 ± 0.31g, 13.2 ± 0.29g and 12.0 ± 0.31g for 25, 30 and 35°C, respectively) (Figure 2.6b).

Moreover, the 25°C southern treatment group exhibited an average total length (10.1 ± 0.31g) significantly lower than the average for the 30°C and 35°C southern treatment groups (13.2 ± 0.29g and 12.0 ± 0.31g, respectively), which did not differ significantly from one another (Figure 2.6b).

Table 2.4 Analyses of variance in body mass (g) among & within *L. calcarifer* populations

	Factor	Partial Eta Squared	*F*	df	*p*
a	Population	0.001	0.298	1, 217	0.586
	Temperature	0.576	98.207	3, 217	< 0.001
	Population x Temperature	0.342	37.562	3, 217	< 0.001
b	20°C	0.212	15.367	1, 57	< 0.001
	25°C	0.553	76.701	1, 62	< 0.001
	30°C	0.23	14.65	1, 49	< 0.001
	35°C	0.18	10.72	1, 49	0.002
c	Northern Treatment Group	0.519	43.174	3, 120	< 0.001
	Southern Treatment Group	0.728	86.616	3, 97	< 0.001

a: Two-way ANOVA with population and temperature as fixed factors; *b*: One-way ANOVA with population as fixed factor; *c*: One-way ANOVA with temperature as fixed factor.

When considering the among-population differences in liver mass a significant effect is observed for population, temperature and the population x temperature interaction (Table 2.5a). As observed for body mass, the average liver mass of acclimated fish was significantly larger in the northern population for the 20°C (0.299 ± 0.014g) and 25°C (0.367 ± 0.017g) treatment groups than in the southern 20°C (0.224 ± 0.014g) and 25°C (0.288 ± 0.017g) treatment groups (Table 2.5b; Figure 2.5f). Conversely, for the intermediate (30°C) and the heat-stress (35°C) treatments the final average mass of the southern treatment group (0.424 ± 0.017g and 0.541 ± 0.029g, respectively) was significantly larger than the average liver mass of the northern treatment group (0.264 ± 0.017g and 0.402 ± 0.030g, respectively) (Table 2.4b; Figure 2.5f**)**. Likewise, significant within-population differences in liver mass were also observed for the northern and southern populations (Table 2.5c; Figure 2.5). The

average liver masses of the 20°C, 25°C and 30°C northern treatment groups (0.299 ± 0.014g, 0.367 ± 0.017g and 0.264 ± 0.017g, respectively) do not differ significantly from each other (Figure 2.5d). The 20°C and 25°C southern treatment groups also do not differ from each other significantly in their average liver mass (0.224 ± 0.014g and 0.288 ± 0.017g, respectively); however, both treatments do differ significantly from the average liver mass of the 30°C (0.424 ± 0.017g) and 35°C (0.541 ± 0.029g) southern treatment groups, which also differ significantly from one another (Figure 2.5e). Moreover, the average liver mass of the 35°C southern treatment group (0.541 ± 0.029g) is significantly higher than all other southern treatment groups (Figure 2.5e).

Table 2.5 Analyses of variance in liver mass (g) among & within *L. calcarifer* populations

	Factor	Partial Eta Squared	*F*	df	*p*
a	Population	0.032	7.389	1, 225	0.007
	Temperature	0.334	37.696	3, 225	< 0.001
	Population x Temperature	0.255	25.689	3, 225	< 0.001
b	20°C	0.191	15.137	1, 64	< 0.001
	25°C	0.146	10.573	1, 62	0.002
	30°C	0.442	44.587	1, 61	< 0.001
	35°C	0.222	10.828	1, 38	0.002
c	Northern Treatment Group	0.177	8.009	3, 112	< 0.001
	Southern Treatment Group	0.677	78.956	3, 113	< 0.001

a: Two-way ANOVA with population and temperature as fixed factors; *b*: One-way ANOVA with population as fixed factor; *c*: One-way ANOVA with temperature as fixed factor.

Lastly, when considering the among-population differences in heart mass a significant effect is observed for population, temperature and the population x temperature interaction (Table 2.6a). Despite the non-significant differences in average heart mass observed between the 20°C northern and southern (0.041 ± 0.01g and 0.029 ± 0.01 respectively) and the 35°C northern and southern treatment groups (0.049 ± 0.03g and 0.047 ± 0.01g respectively), the average heart mass did differ significantly between the 25°C northern and southern (0.065 ± 0.01g and 0.042 ± 0.01g respectively) and the 30°C northern and southern (0.070 ± 0.01g and 0.044 ± 0.01g respectively)

treatment groups (Table 2.6b; Figure 2.5i). Significant within-population differences in heart mass were also observed for the northern and southern populations (Table 2.6c; Figure 2.5). Average heart mass of the 20°C (cold-stress) northern treatment group (0.041 ± 0.005g) was significantly less than the average heart mass of the 25°C and 30°C northern treatment groups (0.065 ± 0.005g and 0.070 ± 0.004g, respectively) while not significantly different from the average mass of the 35°C (heat-stress) northern treatment group (0.049 ± 0.030g) (Figure 2.5g). Moreover, average heart mass did not differ significantly between the 25°C and 30°C northern treatment groups (0.065 ± 0.005g and 0.070 ± 0.004g, respectively) or the 30°C and 35°C northern treatment groups (0.070 ± 0.004g and 0.049 ± 0.030g, respectively) (Figure 2.5g). Within southern population comparison revealed a non-significant difference between the average heart mass of the 20°C, 25°C and 30°C treatment groups (0.029 ± 0.01g, 0.042 ± 0.01g and 0.044 ± .01g, respectively) (Figure 2.5h). The 20°C (cold-stress) southern treatment group exhibited a significantly smaller average heart mass compared to that of the 35°C southern treatment group (0.047 ± 0.01g); however, the 25°C and 30°C southern treatment groups did not differ significantly from the 35°C southern treatment group (Figure 2.5h).

In terms of auto-correlation of variables, we found that liver and heart mass were generally correlated with body mass (i.e. larger liver and heart in fish with larger overall body mass) across all treatment temperatures for both the northern (Figure 2.7a, c respectively) and southern (Figure 2.7b, d respectively) populations. However, several major outliers did exist, with liver mass outliers either exclusively or near-exclusively from the 35°C treatments (illustrated as black circles) for northern and southern populations (Figure 2.7a, b respectively). Likewise, southern heart mass values were almost always seen as large deviations from the basic heart-length relationship (Figure 2.7d).

Table 2.6 Analyses of variance in heart mass (g) among & within *L. calcarifer* populations

	Factor	Partial Eta Squared	*F*	df	*p*
a	Population	0.088	22.8	1, 235	< 0.001
	Temperature	0.102	8.899	3, 225	< 0.001
	Population x Temperature	0.035	2.821	3, 225	0.04
b	20°C	0.049	3.237	1, 63	0.077
	25°C	0.173	12.358	1, 59	0.001
	30°C	0.229	17.778	1, 60	< 0.001
	35°C	0.001	0.05	1, 53	0.824
c	Northern Treatment Group	0.141	6.577	3, 120	< 0.001
	Southern Treatment Group	0.106	4.562	3, 115	0.005

a: Two-way ANOVA with population and temperature as fixed factors; *b*: One-way ANOVA with population as fixed factor; *c*: One-way ANOVA with temperature as fixed factor.

FIGURES

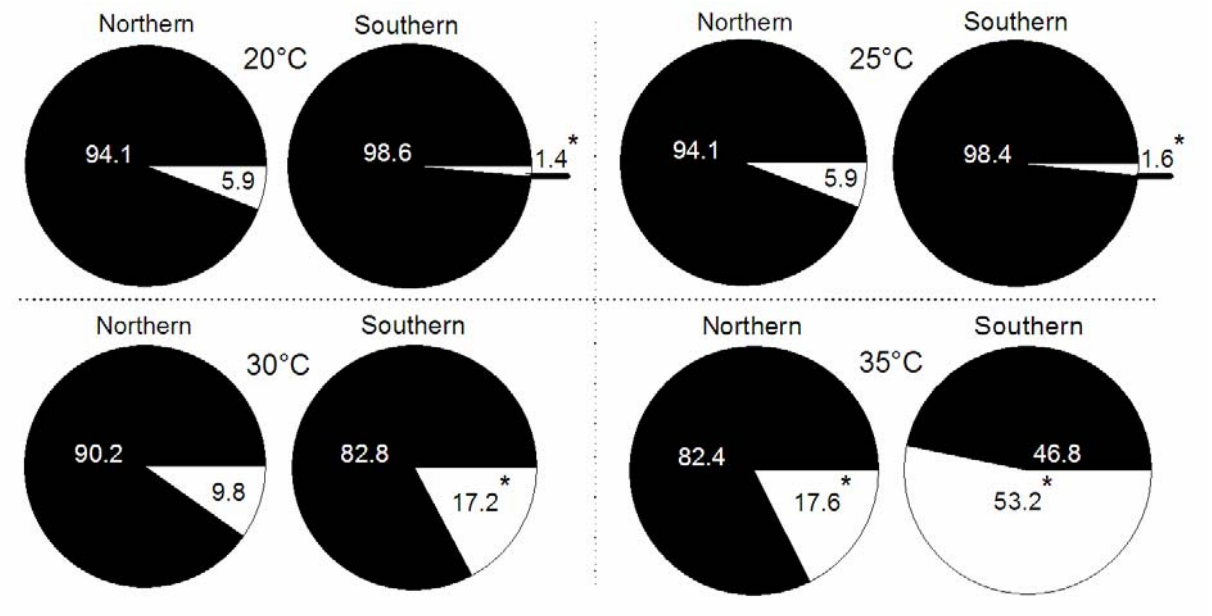


Figure 2.1. Mortality Rates in northern and southern *L. calcarifer* population temperature treatments (20, 25, 30 and 35°C). Survival and mortality rates, proportional to each treatment size, are displayed as black and grey fills, respectively. Numbers displayed within each segment are percentage of survivors (white text) and deaths (black text). Significant deviation from expected mortality as per methods (χ^2 test; d.f. = 1; $\acute{\alpha} = 0.05$) is indicated by *.

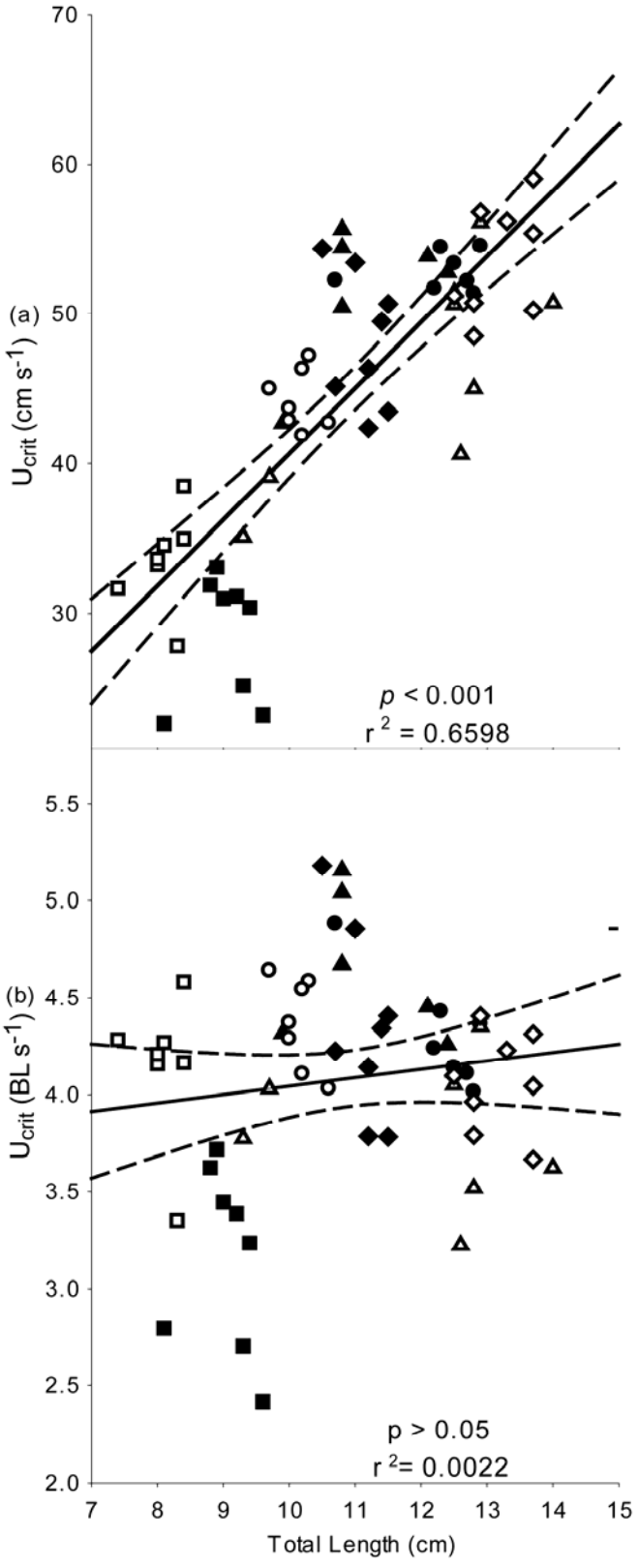


Figure 2.2. Interaction between both *L. calcarifer* experimental populations total length (TL) in centimeters (cm) and critical swimming performance (U_{crit}) reported in centimeters per second (cm s^{-1}; panel a) and body lengths per second (BL s^{-1}; panel b) at acclimated control and temperature stress treatments (see Methods). Symbols for each treatment are as follows: triangle (▲), diamond (♦), circle (●) and square (■) represent 35, 30, 25 and 20°C respectively. Filled and unfilled symbols represent the northern and southern populations respectively.

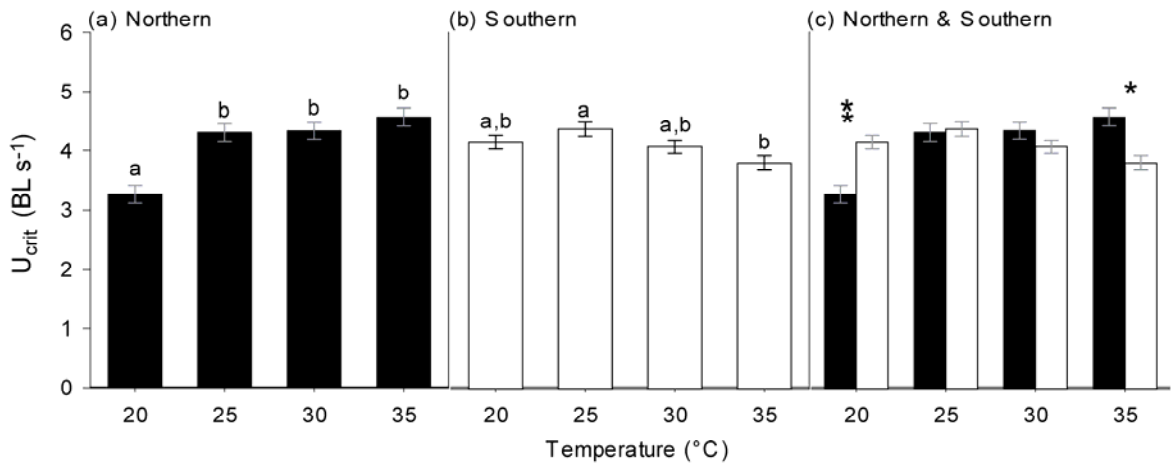


Figure 2.3. Critical swimming performance (U_{crit}) comparison between *L. calcarifer* treatment groups exposed to various thermal treatments. Panels (a) and (b) depict within population responses in swimming performance at different temperatures whereas panel (c) depicts among population responses in swimming performance at specific temperatures. Homogenous groupings ($p < 0.025$) of intra-population treatment responses are provided by letters in (a) and (b). Among-population univariate ANOVA significance of $p = 0.03$ and $p = 0.001$ are indicated by * and **, respectively (Table 2b).

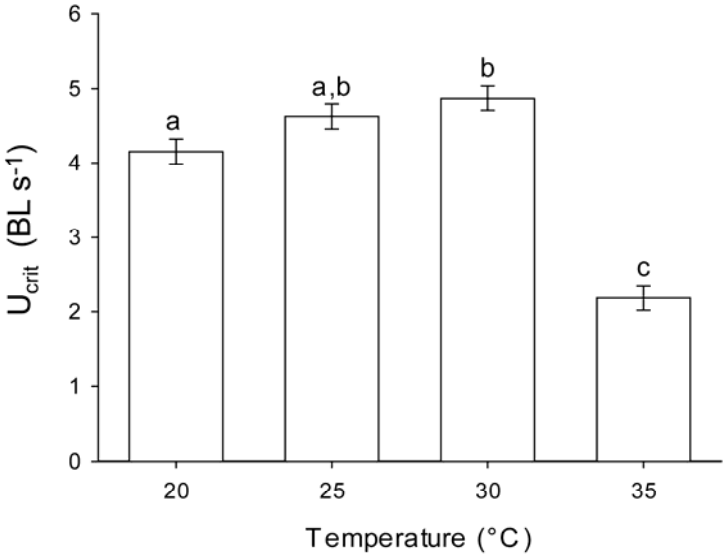


Figure 2.4. Critical swimming performance (U_{crit}) comparison within the unique 20°C southern treatment group challenged at various temperatures (see Methods). Homogenous groupings ($p < 0.05$) of intra-population treatment responses, as per Tukey HSD post-hoc analysis, are provided by letters.

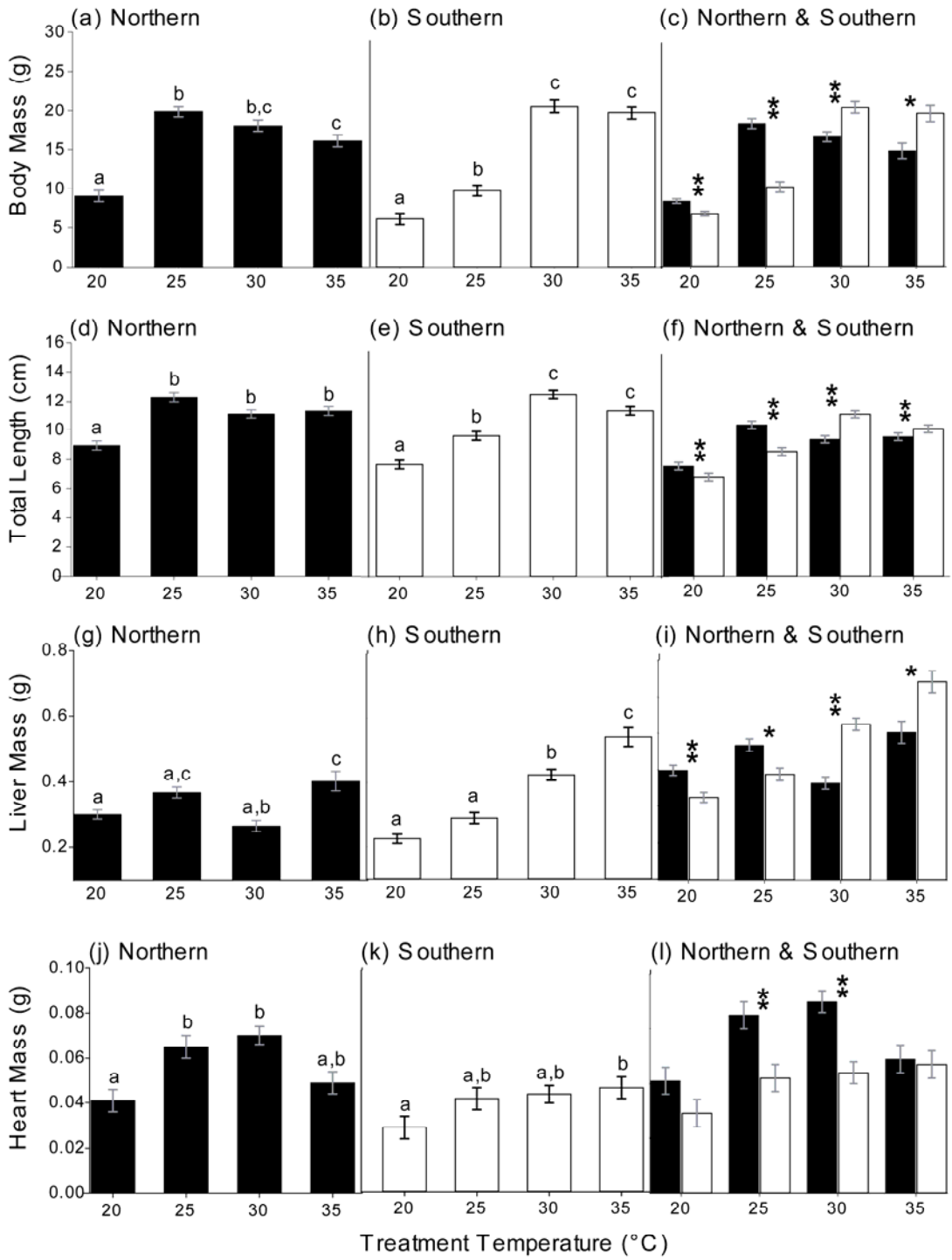


Figure 2.5. Variations in growth parameters for all acclimated fish from each treatment temperature (n = 4) and *L. calcarifer* population (n = 2). Panels (a) and (b), (d) and (e), (g) and (h), (j) and (k) depict response of body mass (a, b) total length (d, e), liver mass (g, h) and heart mass (j, k), respectively, to the four temperature treatments within each northern and southern treatment group (as shown). Panels (c), (f), (i) and (l) depict between population responses in body mass (c), total length (f), liver mass (i) and heart mass (l) to the four temperature treatments (as shown). Significant homogenous groupings ($p < 0.017$) of intra-population treatment responses are provided by letter groupings (Tables 4, 5 and 6 for body, liver and heart mass respectively). Among-population univariate ANOVA significance levels of $p < 0.01$ and $p < 0.001$ are indicated by * and ** respectively (Tables 4, 5 and 6 for body, liver and heart mass respectively).

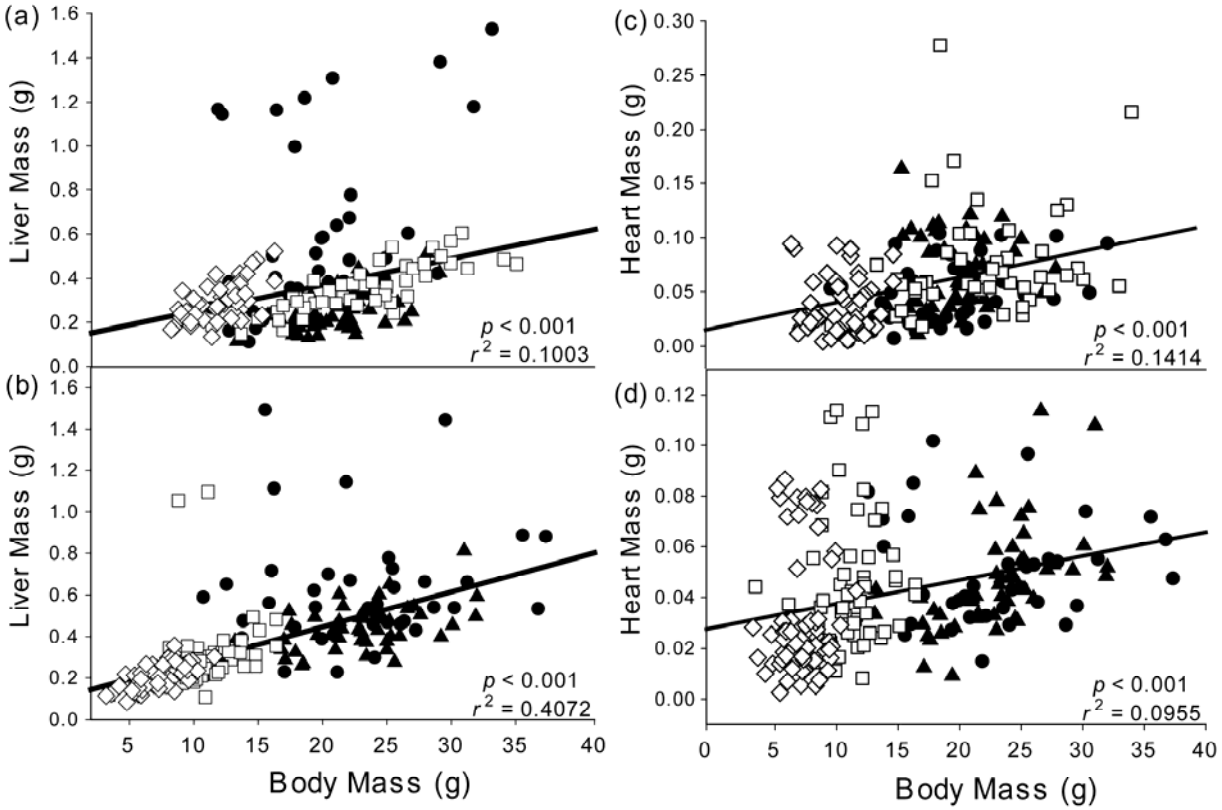


Figure 2.6. Interaction between both the body mass (g), liver mass (a, b) and heart mass (c, d) of both the northern and southern treatment groups at acclimated control and temperature stress treatments (see Methods). Symbols for each treatment are as follows: circle (●), 35°C; triangle (▲), 30°C; square (□), 25°C; and diamond (◊), 20°C.

2.5: DISCUSSION

Based on our swimming performance and morphometric analyses, tropical *Lates calcarifer* (barramundi) populations appear to be adapted to each of their native thermal environments. Lowest levels of mortality and highest physiological performance across the range of temperatures examined appeared to be optimum at the temperature that most closely matched the native thermal environment of each population. Conversely, growth rates appeared to be optimum in intermediate and non-native temperatures for both populations. Specifically, southern population individuals, who experience colder mean annual temperatures, performed significantly better and exhibited lesser growth at colder temperatures than northern population individuals, who performed best and exhibited lesser growth at warmer temperatures which more closely reflect their native tropical climate. Of equal note, however, was that both of these populations displayed relatively high levels of performance across a broad range of temperatures, suggesting the capacity for wide thermal tolerance in this species. This fact may go a long way towards explaining the extensive range of this species, and indicates their resilience to wide changes in their thermal environment.

Mortality rates

Our observed mortalities suggest a strong population-specific acclimatory tolerance to thermal stress in each population. The significantly higher than expected survival of southern fish in their cold stress (20°C) and control (25°C) treatments suggests that this population is better adapted to colder water compared to that of northern fish, which did not exhibit a mortality rate significantly different from the expected mortality at these temperatures as the majority of the observed mortality appeared to be driven by an inability to acclimate to non-native temperature treatments. Conversely, southern fish

exhibited a significantly higher mortality rate than expected in the 30°C and 35°C treatments, which suggests a decrease in tolerance for warmer waters. Likewise, northern fish did not exhibit significantly different mortality to that expected at 30°C. However, northern fish, like the southern fish exhibited a significantly higher mortality rate than expected in the 35°C treatment. Moreover, the mortality rate of southern fish was 4 ½ times higher in the 35°C treatment than that of the northern fish under the heat-stress treatment (53.2 and 17.6% respectively). This suggests a strong population-specific acclimatory tolerance to extreme heat stress in the northern population (see (Kingsolver and Huey, 2008). Furthermore, any such acclimatory advantage would likely be driven by the unique genotypes inherent within these two geographically, thermally and genetically distinct populations (Chenoweth et al., 1998; Keenan, 1994).

Complex physiological effects are likely, with the stress imposed on 35°C treated fish being not only thermal in nature. All temperature-linked water properties, such as decreasing oxygen saturation with increasing water temperature (Henry's Law), would likely have affected fish at this temperature stress (Powers, 1980). Decrease in oxygen saturation at 35°C is likely to be causing major physiological stress in southern fish, who normally would not be subjected to prolonged high temperatures in their native habitats (Beitinger and Bennett, 2000; Beitinger *et al.*, 2000; Claireaux *et al.*, 2006; Fitzgibbon, Strawbridge and Seymour, 2007; Herbert and Steffensen, 2005; Powers, 1980; Powers, Greaney and Place, 1979; Virani and Rees, 2000). Notably, the impact of combined oxygen and temperature stress is likely to also be significant at 30°C, a temperature at which the southern population was seen to readily perform at a high level. Such a response would require both an adequate and specific molecular response for survival (Crawford, Pierce and Segal, 1999b; Crawford and Powers, 1992; Fangue, Hofmeister and Schulte, 2006; Podrabsky and Somero, 2004; Schulte *et al.*, 2000; Somero, 1996). Our *L. calcarifer* populations, may therefore, possess unique

gene variants (i.e. allozymes) of critical metabolic and/or oxygen binding proteins such as lactate dehydrogenase-B and/or have different expression profiles of such proteins (i.e. transcript abundance) (Crawford et al., 1999b; Crawford and Powers, 1992; Fangue et al., 2006; Hochachka and Somero, 2002; Kassahn et al., 2007; Oleksiak et al., 2002; Place and Powers, 1984a, 1984b; Podrabsky and Somero, 2004; Powers, 1980; Powers et al., 1979; Schulte et al., 2000; Somero, 1996). The population-specific acclimatory response to such temperature stress is likely to occur at both levels, thereby likely altering the proteome and/or transcriptome in such discrete tropical populations.

Swimming performance and morphological response

Acclimation temperature was also found to have a profound effect on the swimming speed performance of barramundi from each population. As for their survivorship, the direction and magnitude of this effect appeared to center around the native temperature of the population relative to the experimental temperature being examined, i.e. individuals at a temperature nearest to their native thermal climate performed the best. The significantly faster U_{crit} of southern fish in 20°C and northern fish in 35°C treatments was anticipated (see review by Powers and Schulte, 1998). The fact that seasonal temperature ranges differ by 5-10°C between these two locations, as well as being geographically and genetically distinct (Chenoweth et al., 1998; Keenan, 1994), provides evidence for the potential adaptation of the more southern population to colder waters and the more northern population to warmer waters. Moreover, the native rivers of northern fish (Daly River) and southern fish (Burdekin River) have currents which vary with seasonal rains, from approx. 10 to 185 cm s^{-1} (Warfe, D., personal communication), and approx. 30 to 550 cm s^{-1} (Pocock, G., personal communication), respectively, which could explain (in part) the relatively high upper U_{crit} speed observed for both populations (approx. 55 to 60 cm s^{-1}) in that swimming performance tends to

mirror home-range flow rates (Nelson et al., 2003). However, since juvenile barramundi migrate up-stream to tidal creeks around April (i.e. after the wet season) they are not likely to encounter the peak wet-season flow rates as juveniles (Griffin, 1987a; Russell and Garrett, 1985). Regardless, this fluctuation in river currents imposes an inescapable necessity for the northern fish to achieve adequate critical swimming speeds, while in warmed (hypoxic) waters, if they are to persist and survive the wet seasons of their tropical Australian distribution ranges (Fitzgibbon *et al.*, 2007; Nilsson *et al.*, 2009; Virani and Rees, 2000; Herbert and Steffensen, 2005). Therefore, it is not entirely unexpected that northern fish possess a relatively high U_{crit} compared to southern fish when assessed in 35°C following acclimation (Claireaux et al., 2006; Fulton, 2007; Fulton et al., 2005; Heap and Goldspink, 1986; Plaut, 2001). The potential for such a population-specific adaptation to warm northern water, as well as an equally plausible adaptation to cold southern water is supposed by the significantly faster U_{crit} of northern and southern fish under heat-stress (35°C) and cold-stress (20°C) conditions, respectively. Furthermore, the observed difference in acclimatory ability may explain the observed swimming performance differences where both populations were acclimated to treatment-specific temperatures for an identical 28-day period, but southern fish exhibited a nearly five-fold higher mortality rate at 35°C than did northern fish. This substantially higher southern mortality rate at 35°C suggests that these fish were more thermally stressed (upper limit) than northern fish, even prior to swimming (aerobic metabolic) challenge (assessed by U_{crit} performance) in 35°C water.

Such variations in critical swimming performance among each of our barramundi populations are likely due to population-specific variables (e.g. genotype or physiology) that are closely linked to metabolic processes strongly affected by the thermal environment. Potential variables which explain this performance difference in the two extreme temperature stresses (20°C & 35°C) could be, but are not limited to, the

sustainability of aerobic metabolic activity (i.e. to overcome accumulating muscle lactate) and/or the ability to extract sufficient amounts of oxygen from the flowing water column during a critical swimming speed trial (Claireaux et al., 2006; Fangue and Bennett, 2003; Jain and Farrell, 2003). Decreased levels of oxygen saturation in higher water temperatures, an inverse correlation more pronounced in salt-water than freshwater, impacts the critical swimming speed (U_{crit}) of juvenile fish due to limited oxygen consumption limiting the active metabolic rate, as was demonstrated in the coastal mulloway (*Argyrosomus japonicus*) (Fitzgibbon et al., 2007). Similar reductions in aerobic scope were observed under increasingly hypoxic conditions (31, 32 and 33°C as compared to 29°C) in two marine cardinalfishes, *Ostorhinchus cyanosoma* and *O. doederleini*, suggesting a reduction in survivorship in these cardinalfishes following minor increases (2°C) in sea-surface temperatures (SST) compared to the three damselfishes, *Dascyllus anuarus*, *Chromis atripectoralis* and *Acanthochromis polyacanthus* (Nilsson et al., 2009). This species-specific reduction in aerobic scope in response to heightened SST postulates the impact of predicted increases in sea-surface temperatures, due to global climate change over the next century, on less tolerant species and/or population sustainability in that they may be out-competed for habitats and/or food by other species and/or populations possessing greater resilience under hypoxic conditions (Nilsson et al., 2009). The impact of a similar hypoxic stress on the aerobic scope of juvenile *L. calcarifer* is unknown; however, our data suggests that northern and southern populations may possess unique capabilities to deal with sustaining aerobic metabolism under hypoxic conditions (i.e. heat-stress of 35°C). While the impact of temperature on oxygen availability at extreme temperatures is a uniform physio-chemical relationship, our evidence suggests that the biological response of barramundi to these thermal/oxygen extremes varies considerably according to their population history. Individuals from a northern population had

seemingly little difficulty with extracting enough oxygen to maintain a swimming performance considerably higher than the southern population individuals. Population-specific abilities to tolerate low oxygen saturation could be linked to differences in the surface area of gill filaments between these thermally distinct populations, as such a mechanism was observed between populations of the African cichlid *Pseudocrenilabrus multicolor* endemic to normoxic and hypoxic lakes (Chapman et al., 2000).

Interestingly, the 35°C southern treatment group, aerobically challenged at 35°C following 28-day acclimation, had a significantly faster U_{crit} than the unique 20°C acclimated southern treatment group aerobically (swimming) challenged at 35°C, which suggests that an acclimation period is in fact beneficial when assessing U_{crit} of fish under heat-stress conditions. On the other hand, when the unique 20°C acclimated southern treatment group was aerobically challenged in 25°C and 30°C no significant decrease in U_{crit} performance was observed, consistent with these temperatures being within the range of naturally experienced temperatures for the more southern population. Of particular interest is the observation that the unique 20°C acclimated southern treatment group, when subjected to aerobic (swimming) challenge at 30°C, outperformed all other metabolically challenged (swum) fish. This suggests that temperature-specific acclimation is not required for all temperature treatments, if these treatments are not thermal stresses. Furthermore, the unexpectedly high U_{crit} performance of this unique 20°C acclimated southern treatment group when aerobically challenged in 30°C goes against the standing belief that acclimation is necessary to allow for muscle fiber adjustment to any non-native temperatures and therefore providing a more realistic U_{crit} assessment (Beitinger and Bennett, 2000; Beitinger *et al.*, 2000; Claireaux *et al.*, 2006; Heap and Goldspink, 1986). With no parallel data available for northern fish, no conclusions can be drawn as to the existence of a similar trend for both juvenile barramundi populations. However, the observed U_{crit} difference

between normal southern treatment groups and this unique 20°C acclimated southern treatment group is substantial enough to be noteworthy and perhaps even warrant further investigation via a genotype-by-environment (GxE) experiment.

Remarkably homogenous performances were found at intermediate temperatures across the two populations, suggesting a wide thermal tolerance in these discrete *L. calcarifer* populations. The fact that no significant U_{crit} swimming performance differences exist either within or between populations at the intermediate/respective control temperature treatments (25°C and 30°C for southern and northern, respectively), supports the experimental assumption that these intermediate temperatures do not impose thermal stress on either study population. The significant differences in body and organ masses and total body length between populations in the intermediate/control treatments, however, suggests the existence of optimal growth rates for each population in the other population's control treatment temperature (i.e. northern and southern fish grew heavier and longer at 25°C and 30°C, respectively). These unexpected growth rates in each population's non-control temperature may provide insight to a triggered physiological mechanism operating between geographically isolated and genetically distinct *L. calcarifer* populations (Chenoweth et al., 1998; Keenan, 1994). We speculate that tropical northern populations, which are presumably accustomed to dealing with a warmer climate and annual high temperature extremes, may commit energy reserves to cope with such extremes and, therefore, when exposed to a less stressful temperature (25°C), more energy may be available to commit to growth. The opposite may be true for southern fish, as their relatively cold water habitat imposes no premise for such a "heat-stress precaution" mechanism and, therefore, southern fish may respond to the warmer temperature treatment (30°C) by committing extra energy to growth at that temperature. Alternatively, the spike in southern fish growth at 30°C could be an artifact due to the decrease in tank stocking density as a result of the

significantly greater than expected mortality rate at this temperature. Regardless, these observed variations in phenotypes between discrete *L. calcarifer* populations potentially provide novel evidence of countergradient variation across the range of this tropical catadromous species (i.e. phenotypic variation (V_p) is greater than expected) (Conover and Schultz, 1995).

2.6: CONCLUSIONS

Our findings indicate that tropical *L. calcarifer* display markedly different thermal optima at which levels of growth, survivorship and performance are maximized. The fact that these optima closely match the mean annual temperate of the native waterways of each population suggests an intrinsic adaptation of each population to their native thermal regime. Despite having such thermal optima, we found a surprisingly wide thermal tolerance of these species. These results raise several intriguing questions. Most pertinent of these is what underlying genetic, physiological or other biological mechanism(s) explain the observed U_{crit} performance differences? To explore this question, candidate genes are required. However, as *L. calcarifer* is not a model species, only limited genetic information is available. An ideal candidate gene for such a preliminary assessment of a genetically-linked metabolic mechanism is lactate dehydrogenase-B (*ldh-b*), a gene previously associated with population-specific temperature differences in swimming performance of the temperate estuarine killifish *Fundulus heteroclitus* (DiMichele and Powers, 1982); see also reviews by Powers et al., 1991; Powers and Schulte, 1998). Moreover, these swimming performance differences were observed when geographically and thermally distinct acclimated *F. heteroclitus* populations were assessed under similar thermal stress treatments as *L. calcarifer* populations examined in this study. Additional evidence for the validity of *ldh-b* as a

candidate gene in barramundi is the existence of two *Fundulus ldh-b* alleles which are alternatively fixed in northern and southern *Fundulus* populations (reviewed by Powers et al., 1991; Powers and Schulte, 1998) and the differential expression profiles of this gene in thermally distinct populations of *F. heteroclitus* (Crawford and Powers, 1992; Fangue et al., 2006; Oleksiak et al., 2002; Schulte et al., 2000; Segal and Crawford, 1994b; Segal et al., 1996). The characterization of this candidate gene in barramundi is the subject of Chapter 3.

Chapter 3

Comparative characterization of a temperature responsive gene (lactate dehydrogenase-B, *ldh-b*) in two congeneric tropical fish, *Lates calcarifer* and *Lates niloticus*

3.1: ABSTRACT

The characterization of candidate loci is a critical step in obtaining insight into adaptation and acclimation of organisms. In this study of two non-model tropical congeneric perciformes (*Lates calcarifer* and *L. niloticus*), we characterized both coding and non-coding regions of lactate dehydrogenase-B (*ldh-b*), which is adaptive in the temperate fish, *Fundulus heteroclitus*. *Ldh-b* was 5,004 and 3,527 bp in length in *L. calcarifer* and *L. niloticus*, respectively, with coding regions comprising 1,005 bp in both species. A high level of sequence homology existed between species for both coding and non-coding regions of *ldh-b* (> 97% homology), corresponding to a 98.5% amino acid sequence homology. All six known functional sites within the encoded protein sequence (LDH-B) were conserved between the two *Lates* species. Ten simple sequence repeat (SSR) motifs (mono-, di-, tri- and tetranucleotide) and thirty putative microRNA elements (miRNAs) were identified within introns 1, 2, 5 and 6 of both *Lates* species. Five single nucleotide polymorphisms (SNPs) were also identified within miRNA containing intron regions. Such SNPs are implicated in several complex human behaviors and/or diseases (as demonstrated by extensive genome-wide association studies). This characterization serves as a platform to further examine how non-model species may respond to changes in their native temperatures, which are expected to increase by up to 6°C during this century.

3.2: INTRODUCTION

The lactate dehydrogenase-B enzyme (LDH-B) plays a critical role in maintaining aerobic metabolism by converting lactate, the major by-product of anaerobic glycolysis, to pyruvate via oxidation in the presence of its coenzyme nicotinamide adenine dinucleotide, (NADH) (reviewed by Powers et al., 1991; Powers and Schulte, 1998). LDH-B can also convert lactate directly to glucose via gluconeogenesis. This conversion of accumulating lactate from aerobic tissues (e.g. heart, skeletal muscle) occurs in the liver and allows desired aerobic metabolic activity (e.g. swimming) to be sustained for extended lengths of time (reviewed by Powers et al., 1991; Powers and Schulte, 1998). In addition to these metabolic functions, the LDH-B enzyme affects the oxygen binding affinity of hemoglobin (Hb) by altering intra-erythrocyte ATP concentrations of Hb in fish (Powers, 1980; Powers et al., 1979). This effect of LDH-B on Hb-O_2 binding affinity directly impacts delivery of Hb-bound oxygen to red muscle tissues and may therefore be an alternate mechanism by which LDH-B affects sustainability of aerobic performance such as swimming performance in fish (DiMichele and Powers, 1982; Fangue et al., 2008; Powers and Schulte, 1998).

Oxygen available to aquatic organisms is inversely correlated with water temperature (Henry's Law) leading to the potential for variability in the ability of Hb to uptake and transport oxygen under differing thermal regimes. In natural populations of aquatic organisms the genes involved in aerobic metabolism and oxygen transport, such as *ldh-b* and hemoglobin may therefore be subjected to strong selective pressure (Crawford et al., 1999b; Crawford et al., 1999a; Powers, 1980; Powers et al., 1979; Rees et al., 2001; Schulte et al., 2000; Segal et al., 1996). In fact, within thermally distinct populations of the temperate estuarine killifish, *Fundulus heteroclitus*, Hb-O_2 affinity varies together with intra-erythrocyte ATP concentrations dependent on which

LDH-B isozyme (LDH-B^a or LDH-B^b) is fixed (reviewed by Powers et al., 1991; Powers and Schulte, 1998). In addition to the extensive characterization and investigation of *ldh-b* in *F. heteroclitus* (reviewed by Powers et al., 1991; Powers and Schulte, 1998), the translated protein of this candidate gene has also been characterized and investigated in other temperate fishes like rainbow trout, *Salmo gairdneri* (Kao and Farley, 1978; Klar et al., 1979), crested blenny, *Anoplarchus purpurescens* (Johnson, 1977). However, this gene has not been fully characterized in any tropical perciform to date and there have been no investigations into the role this gene may have in thermal acclimation or the capacity to cope with thermal stress in tropical fishes.

As a first step in understanding the role *ldh-b* may have in thermal adaptation or acclimation of tropical fish species, we characterized this gene in two tropical congenerics, the Australian barramundi, *Lates calcarifer* and the African nile perch, *L. niloticus*. *L. calcarifer* is a catadromous, protoandrous hermaphrodite, denizen to rivers, estuaries and shallow marine environments throughout northern Australia (25°S – 12°S) and the south-east Asian archipelago (13°N – 10°S) (Chenoweth et al., 1998; Doupé et al., 1999; Keenan, 1994), while *L. niloticus* originates from east African rivers and lakes (7°S - 27°N) (Froese and Pauly, 2008). *Ldh-b* sequences were characterized in these fishes to allow for comparisons of coding sequences between these congeneric tropical perciformes in that fish are ideal candidates for studies investigating adaptations or enhanced acclimatory ability to native thermal regimes (Cossins and Crawford, 2005).

In addition to the traditional characterization of coding regions (exons) we also characterize, for the first time in a non-model fish, the non-coding regions (introns) of the *ldh-b* locus in these perciform species to establish if regulatory motifs and/or elements (simple sequence repeats (SSRs) or microRNAs (miRNAs)), which are known to be embedded in or encoded by non-coding regions, are present. Previous studies have demonstrated the presence of such elements within introns of other genes, where

they are implicated in regulation or silencing of transcription (Ambros, 2004; Baek et al., 2008; Bartel, 2004; Doench and Sharp, 2004; Lai, 2003; Li et al., 2004; Meloni et al., 1998; Reisman et al., 1988; Wittekindt et al., 2000).

3.3: METHODS

Complementary DNA (cDNA) synthesis from hepatic messenger RNA (mRNA)

As no *ldh-b* sequence information was available for either species total liver RNA was first required for reverse transcription to obtain a *Lates ldh-b* cDNA sequence for primer design. This cDNA sequence was necessary to enable the design of *ldh-b* specific primers targeting the non-coding (intron) sequences of the *ldh-b* locus that can only be obtained from genomic DNA. Liver RNA was targeted for this initial cDNA sequencing as this tissue exclusively expresses the LDH-B protein as opposed to the alternative isozymes, LDH-A or C, which are expressed in other, non-hepatic tissues (Crawford and Powers, 1992; Quattro et al., 1993; Segal and Crawford, 1994b). Total RNA was extracted from snap-frozen livers from four Darwin Harbour, Northern Territory, *L. calcarifer* individuals, using Trizol and following manufacturer's instructions (Invitrogen Australia Pty, Mount Waverley Victoria). Extracted RNA was treated with Turbo DNA-free (Ambion, Austin, TX, USA) to remove all traces of contaminating genomic DNA. Messenger RNA (mRNA) was reverse transcribed to generate cDNA immediately via IM-Prom II Reverse Transcriptase with Oligo dT_{20} and random primers (Promega, Madison, WI USA), as per manufacturer's instructions.

Amplification of the *ldh-b* locus off hepatic cDNA

Amplification and sequencing of *ldh-b* from the *L. calcarifer* hepatic cDNA was accomplished with general fish primers designed by aligning *ldh-b* coding sequences

from a taxonomically diverse range of fishes from the National Center for Biotechnology Information (NCBI) sequence database (GenBank) (Zhang and Madden, 1997). The following sequences were aligned and primers were designed based on conserved regions - Gadiformes *Trachyrincus murrayi* [AJ609235], *Merlangius merlangus* [AJ609234], *Gadus morhua* [AJ609233] *Coryphaenoides armatus* [AJ609232]; Cypriniformes *Danio rerio* and *Cyprinus carpio* [AY644476]; Cyprinodontiformes, *F. heteroclitus* [L43525], *F. heteroclitus* (D. Crawford personal communication); Squaliformes *Squalus acanthias* [AF059035]. While several primers were initially designed those that produced a specific product from *L. calcarifer* hepatic cDNA was the forward primer designed off Gadiformes and Cypriniformes (ATGGCCTGTGCCGTCAGC) with the reverse primer designed of Cypriniformes (TCTTTCAGGTCTTTCTGGAT) which annealed in exon 2 and exon 7, respectively. Further sequence upstream (exon 1 to 5' end of exon 2) was obtained using a previously published primer for *F. heteroclitus* (Bernardi et al. 1993) and primers *L. calcarifer*-Intron2-R1 or *L. niloticus*-Intron2-R1 to give the full coding sequence (Table 3.1).

All PCR reactions were conducted in the following manner: Amplification reactions (20 µL) contained the following final concentrations: 1X Buffer [2.5 mM Tris pH 8.7, 5 mM KCl, 5 mM $(NH_4)_2SO_4$, containing 1.5 mM $MgCl_2$] (Qiagen, Doncaster, VIC, Australia) or 1x Buffer [2.5 mM Tris pH 8.7, 5 mM KCl, 5 mM $(NH_4)_2SO_4$ not containing 1.5 mM $MgCl_2$] (Bioline Pty Ltd., Alexandria NSW, Australia) (unless more was required as per Table 3.1)], 250 µM each dNTP, 250 nM each primer (Table 3.1), 10 ng gDNA template and 0.75 to 1.5 units of Taq Polymerase (Qiagen, Doncaster, Victoria and Bioline Pty Ltd., Alexandria New South Wales, Australia). Thermal cycling was conducted on a MJR DNA Engine thermal cycler (Bio-Rad Laboratories Pty., Ltd., Gladesville, New South Wales) as follows: initial denaturation at 94°C for 3 min followed by 35 cycles of 94°C denaturation for 30 sec, annealing at primer specific

T_a for 30 sec (Table 3.1), 72°C extension for 30 to 90 sec depending on target fragment size with larger fragments (> 1,000 bp) requiring longer (> 60 sec) extension times (Table 3.1) and a final 72°C extension for 10 min. Melting temperature (T_m) was calculated via (A/T x 2 + G/C x 4) method, with the annealing temperature (T_a) set at 5°C less than T_m, for all primer-pairs (Table 3.1).

Table 3.1. PCR and sequencing primers used to obtain *ldh-b* sequences in *Lates* species. Segments 1 to 5 refer to regions depicted in Figure 1. All primers anneal to gDNA (see Methods). Primers used for forward and reverse sequencing reactions only (nested primers) are indicated by (SeqF) and (SeqR), respectively. Final $MgCl_2$ concentrations in brackets refer to *L. niloticus* amplification requirements, which differed from that required for *L. calcarifer*.

Lates spp. *ldh-b* Primers	5' to 3' Sequence	$MgCl_2$	T_a	Amplicon (bp)
Segment 1*				
L. calcarifer-Seg1-R1	GATTTAGACATGTCGTTCCTCAG	1.5 - 2.5	61°C	2000
L. calcarifer-Seg1-R2	ATAATGACACCATCAATGTTCACG	1.5 - 2.5	61°C	1000
L. calcarifer Intron1 (Seq-R)	ATGGATGAATGTCTCAATCAG	1.5 - 2.5	53°C	500
L. calcarifer Intron2-R1	TCGGATAACAGAAGCACTCAC	1.5 - 2.5	55°C	1500
L. niloticus Intron2-R1	TAATCACTCATGGCCTCGG	1.5 - 2.5	53°C	1300
L. niloticus Intron1-R1	AACTGGAAACTAATCTAGGCC	1.5 - 2.5	55°C	450
L. niloticus Intron1 (SeqR)	TCAGGTTAGCACTGCTGC	1.5 - 2.5	51°C	350
Segment 2				
L. calcarifer-F1	TGATGGAGGATCGTCTGAAAGG	1.5 - 2.5	61°C	800
L. calcarifer-R1	TCGGCCATCAGGTAACGGAAG	1.5 - 2.5	61°C	800
Segment 3				
L. calcarifer-F2'	GTTGATGTGCTGACCTACGTC	1.5-2.5 [3.5-4.5]	59°C	1500
L. calcarifer-R2'	AGCCCTTCAGCTTGATCACC	1.5-2.5 [3.5-4.5]	57°C	1500
Segment 4				
L. calcarifer-F3i	ACAGAGCTTCCACTGTATCAC	1.5 - 2.5	57°C	850
L. calcarifer-R3i	GCAAAAGGTTCCTAGGCATGTA	1.5 - 2.5	59°C	850
L. calcarifer-F3	AGAAGCTGAACCCTGAGATCG	[3.5-4.5]	59°C	800
L. calcarifer-R3	TTCTGGATGCCCCACAGTGTG	[3.5-4.5]	61°C	800
Segment 5**				
L. calcarifer-F3i'	TGGTTGCTAGGATAAAGAATGTG	1.5 - 2.5	59°C	700
L. calcarifer-F3i (SeqF)	AGTTGTAAATAATTCAGGCATC	1.5 - 2.5	53°C	500
L. niloticus Intron6-F1	ATGTGGATAGCCTAGCTTAGC	1.5 - 2.5	55°C	400

* and **: Published forward and reverse *F. heteroclitus ldh-b* primers (Bernardi et al., 1993) used in conjunction with designed primers to amplify terminal (5' and 3') segments in both *Lates* species, respectively.

Verification of *ldh-b* sequences amplified off hepatic cDNA

Amplification of cDNA with the fish specific *ldh-b* primers (F3 and R1: Table 3.1) generated a strong single-band product of approximately 800 bp for all four fish examined. Subsequent amplification of genomic DNA with the published forward (*F. heteroclitus* - F) and designed reverse (*L. calcarifer* - Intron2 - R1 or *L. niloticus* - Intron2 - R1) primers resulted in a strong single-band product of approximately 1,500 bp and 1,300 bp (*L. calcarifer* and L. *niloticus*, respectively) containing the missing exon 1 and 5'end of exon 2 fragment. These products were precipitated with 120 µL isopropanol (70%) for 15 minutes followed by a 500 µL wash with 70% isopropanol prior to drying, re-suspension in water (10 µL) and subsequent sequencing (Macrogen, Inc., South Korea). Sequences were edited in BioEdit (Hall, 1999) and a contig made to give the full length coding sequence (1,005 bp) and produce a consensus cDNA sequence. To check that the correct *ldh* gene homologue had been obtained the sequence was used in a BLAST search of GenBank and also directly aligned using ClustalW in Mega 3.1 (Kumar et al., 2004) to those *ldh-b* sequences previously utilized for general fish primer design (see above). The obtained sequence shared 92% homology with *F. heteroclitus* LDH-B amino acid sequence, a level of homology that is well above that reported for *F. heteroclitus* LDH-C *v.* LDH-B and LDH-C *v.* LDH-A (78% and 70% homology, respectively) (Quattro et al., 1993). As a further check that the correct *ldh* gene homologue was obtained all nucleotide (and deduced amino acids) sequences were aligned with those of *Danio rerio ldh-b* [AF067202] and *ldh-a* [NM_131246] genes. These two gene homologues differ by 600 bp in length and 34.1% of the nucleotide sequences, respectively. The *L. calcarifer ldh-b* sequence obtained from hepatic cDNA most closely matched that of the *D. rerio ldh-b* gene homologue. Once the initial cDNA sequence was obtained from *L. calcarifer* the full characterization of the *ldh-b* locus from both *Lates* species (*L. calcarifer* and *L.*

niloticus) was accomplished via primer walking along genomic DNA from representative individuals of each species.

Study species and genomic DNA (gDNA) extraction

Genomic DNA was extracted from *L. calcarifer* samples collected from four locations within tropical Australia. Samples from Gladstone, Queensland (23°S, 151°E) and Darwin, Northern Territory (12°S, 130°E) were obtained directly from fish farms while Archer River (Cape York, Queensland: 13°S, 142°E) and Tully River (Tully, Queensland: 17°S, 145°E) samples originated from wild caught fish. *L. niloticus* was purchased as two imported frozen fillets at a local supermarket in Townsville, Queensland and therefore the exact geographical origin of the *L. niloticus* samples examined is unknown, but assumed to be from one of the African Rift Valley lakes (4°N – 14°S) where a large export fishery of this species exists. Fin-clips/muscle tissue were taken from all fish and preserved in ethanol (90%). DNA extractions were performed via proteinase-K digestion (20 mg/mL) in CTAB buffer at 60°C for 1 hr and DNA was subsequently cleaned with a salt and chloroform:isoamyl alcohol (24:1) procedure (Sambrook et al., 1989). All extractions resulted in high molecular weight gDNA, as visualized on a 0.8% agarose gel, with quantities ranging from 20 – 100 ng/µL.

Amplification of the *ldh-b* locus off genomic DNA

Full length *ldh-b* gene sequences (including intron sequences) were obtained from genomic DNA extracts using the primers and primer specific PCR conditions outlined in Table 3.1 and Figure 3.1. In some cases different primers were required for amplification of *L. niloticus* and *L. calcarifer* introns (e.g. intron 5 & 6 in Figure 3.1). See Figure 3.1 for primer binding locations within the *ldh-b* locus. Thermocycling

parameters were the same as that used for cDNA amplification with the exception of the annealing temperature and $MgCl_2$ concentration which varied as listed in Table 3.1. All PCR products obtained from both *L. calcarifer* and *L. niloticus* genomic DNA were verified as *ldh-b* in two ways: 1) searching the NCBI database (GenBank) via nucleotide (blastn) and protein (blastp) basic local alignment and search tool (b.l.a.s.t.) with resulting matches specific to LDH-B exclusively and 2) direct alignment with *F. heteroclitus ldh-b* nucleotide and LDH-B protein sequence. After confirmation of product identity full length gene sequences for each individual (n = 4 *L. calcarifer* and n = 2 *L. niloticus*) were assembled in BioEdit and the consensus nucleotide and deduced amino acid sequences from each species were aligned using ClustalW in Mega 3.1. Consensus gene sequences for each species were submitted to GenBank under accession numbers [**FJ439509**] and [**FJ439510**] for *L. calcarifer* and *L. niloticus* respectively. Additionally, *L. calcarifer ldh-b* coding sequence isolated from hepatic mRNA was submitted to GenBank under accession number [**FJ439507**]. Deduced LDH-B amino acid sequences were manually assessed for the presence of variation within the NH_2-terminal arm (residues 1-20), coenzyme binding domain (residues 21-95 and 118-163), substrate binding domain (residues 164-333) and loop helix ά D region (residues 96-117) (Li et al., 1983).

Assessment of non-coding (intron) sequences for micro RNA (miRNA) and simple sequence repeat (SSR) motifs

Several recent studies have demonstrated intron sequences may contain simple sequence repeat (SSR) motifs of functional importance as they potentially bind regulatory machinery (e.g. promoters and/or enhancers) and may affect gene expression levels (Li et al., 2004; Yue et al., 2001). Intron sequences from *ldh-b* of both *Lates* species were manually assessed for the presence of SSRs. In addition, short microRNA (miRNA) elements (21-23 bp) may be encoded for by introns and these are believed to be spliced

out of pre-messenger RNA (mRNA) subsequently targeting regions within the 3'UTR of actively expressed mRNAs regulating translation from mRNA transcripts to functional proteins (Ambros, 2004; Baek et al., 2008; Bartel, 2004; Chen et al., 2004; Doench and Sharp, 2004; Enright et al., 2003). Intron sequences from *Lates* spp. were therefore assessed for putative miRNA elements with the software package miRanda (Enright et al., 2003). All presented miRNA elements are located within conserved intron regions of *Lates* spp. and matched known miRNA motifs from *Danio rerio* (zebrafish), *Takifugu rubripes* (Japanese pufferfish) and *Tetraodon nigroviridis* (spotted green pufferfish). Several of the identified elements also matched known *Xenopus tropicalis* (African frog) miRNA motifs. A nucleotide match score (score), affinity to bind measure (energy [kcal mol^{-1}]), statistical assessment of match quality (z-score) and homology of motif to query intron sequence (percentage) were all calculated by miRanda for each matching miRNA element (Enright et al., 2003). Threshold values were set to 100, -19 kcal mol^{-1} and 5.0 for score, energy and z-score, respectively, to avoid false-positive miRNA identification (Bonnet et al., 2004; Brennecke et al., 2005; Enright et al., 2003). Overly-stringent thresholds were avoided to maximize the likelihood of identifying putative miRNA elements which have recently been documented at high abundance within eukaryotic genomes (i.e. average of 100 binding sites per miRNA elements genome-wide) (Brennecke et al., 2005).

3.4: RESULTS & DISCUSSION

Descriptive characterization of the *ldh-b* locus exon and deduced amino acid sequences

Variation in the primary sequence and level of gene expression of the *ldh-b* locus has been linked to differences in aerobic performance (reviewed by Powers et al., 1991; Powers and Schulte, 1998) and natural selection (Crawford et al., 1999b; Pierce and

Crawford, 1997) in the temperate fish *F. heteroclitus*. The present study is the first to characterize the full length gene sequence of this important metabolic gene in two tropical perciform fish, namely *L. calcarifer* and *L. niloticus*. This locus consisted of 5,004 and 3,527 bp, of which 1,005 bp was coding, in these two species, respectively (Figure 3.1). Consistent with other fish species the *ldh-b* locus of both *L. calcarifer* and *L. niloticus* contained seven exons and six introns, with all seven *ldh-b* exons being conserved in size between the characterized *Lates* species (Figure 3.1). Additional coding nucleotide sequence comparison revealed 29 base differences between *L. calcarifer* and *L. niloticus* (2.9% divergence); however, the majority of these occurred in the third codon positions which lead to silent (i.e. synonymous) amino acid substitutions. This high level of conservation was expected between *Lates* species being that this locus is known to be under functional constraint due to it role in maintaining aerobic metabolism (reviewed by Powers and Schulte, 1998); (Crawford et al., 1999b; Pierce and Crawford, 1997). Moreover, the encoded enzyme (LDH-B) has also been shown to impact on hemoglobin-oxygen binding affinity in fish, with allozyme variants altering this critical interaction (Powers, 1980; Powers et al., 1979).

Each residue of the post-translation modified protein sequences has a role in the specific functioning of the molecule, from its internal stability to its external functional interactions. Five amino acid differences (1.5%) are present between the LDH-B of both *Lates* spp. These variable amino-acids involve the substitution of threonine (T) for methionine (M), leucine (L) for alanine (A), valine (V) for isoleucine (I), lysine (K) for asparagine (N) and valine (V) for isoleucine at residues 34, 35, 126, 127 and 147, respectively (Figure 3.2). All of these variable amino acid residues reside within the coenzyme (NADH) binding domain (residues 21-95 and 118-163), of LDH-B, which is otherwise conserved between *Lates* spp. (Figure 3.2). This is unanticipated in such a conserved LDH-B functional domain (Li et al., 1983; Quattro et al., 1993) as this

functional domain is also known to be conserved for the LDH-A_4 isoform across fish taxa (Fields and Somero, 1997; Johns and Somero, 2004; Somero, 2003, 2005). Interestingly the amino acid variation found to explain the differences in LDH-A catalytic efficiency across temperate fish taxa resides exclusively in the loop region (Fields and Somero, 1997, 1998; Somero, 1995), a region which is completely conserved between the tropical *Lates* congeners (Figure 3.2). Noteworthy are the two shifts occurring at residues 34 and 127 as these invoke changes in polarity (polar – non-polar) and acidity (neutral polar – basic polar) between *Lates* species, respectively. The effect of these amino acid shifts on the catalytic efficiency (k_{cat}) of LDH-B between *L. calcarifer* and *L. niloticus* is unknown and warrants further investigation. Conversely, the NH_2-terminal arm (residues 1-20), substrate binding domain (residues 164-333) and loop helix ά D region (residues 96-117) (Li et al., 1983), as well as the LDH-B substrate binding (residues 100, 107, 139, 170, 249) and proton acceptor (residue 194) sites found within (The UniProt, 2008), are conserved between both *Lates* species, as expected (Li et al., 1983; Quattro et al., 1993).

Comparison between *Lates* species and *F. heteroclitus* full length LDH-B amino acid sequences revealed a relatively extensive divergence, with 30 amino acids (9.0% divergence) observed between these phylogenetically distant species (data not shown). The two residues demonstrated to have fixed differences between thermally and geographically distinct populations of *F. heteroclitus* were serine (S) *v*. alanine (A) and alanine (A) *v*. aspartic acid (D) at residues 185 and 311 for cold northern and warm southern populations, respectively (reviewed by Powers et al., 1991; Powers and Schulte, 1998). *L. calcarifer* and *L. niloticus* both possess serine (S) and aspartic acid (D) at residues 185 and 311, respectively (Figure 3.2). The deduced LDH-B amino acid sequence of both *Lates* species examined are therefore similar to the cold northern *F. heteroclitus* population at residue 185 and to the warm southern *F. heteroclitus*

population at residue 311. These residues are located on the internal and external surfaces of the folded LDH-B protein, respectively (reviewed by Powers et al., 1991; Powers and Schulte, 1998). The former change (residue 185) has been hypothesized to be associated with a variation in thermal stability due this residue being located at a hairpin turn in the center of the folded protein; whereas the latter change (residue 311) has been hypothesized to be associated with a variation in substrate binding affinity due to this residue being located on the external surface of the conformed protein (reviewed by Powers et al., 1991; Powers and Schulte, 1998). In light of these previous hypothetical explanations, future studies should strive to compare the structure and function (e.g. enzymatic activity, affect on Hb-O_2 binding affinity) of LDH-B in the *Lates* species in parallel to those studies conducted on within and among thermally discrete *F. heteroclitus* populations (Place and Powers, 1984a, 1984b; Powers et al., 1979).

Descriptive characterization of the *ldh-b* locus introns and identification of embedded putative regulatory motifs and/or elements

Nucleotide sequence comparisons of the homologous introns between the two *Lates* spp. characterized reveals a high level of homology (97.7%), a level surprisingly similar to that observed between exons of the *Lates* species (97.1% exon homology). The fact that sequence homology between non-coding introns is similar to that of coding exons suggests similar selective constraint (i.e. functionality) may be driving the high levels of homology observed across these historically non-characterized intronic regions (Hare and Palumbi, 2003; Mattick, 1994). Interestingly, *L. niloticus* introns are consistently smaller than *L. calcarifer* introns with the exception of intron 2, which is 20 bp longer in *L. niloticus* relative to *L. calcarifer* (Figure 3.1). Moreover, one insertion-deletion event (indel) occurs between the two *Lates* species in all introns, with indels ranging in size across introns 1 to 6 (157, 20, 6, 19, 571 and 735 bp, respectively) (Figure 3.1).

Whether these indels are historical insertion or deletion events and whether putative regulatory elements embedded within indel regions have an impact on the transcriptome or functional proteome of either Lates species is unknown. Therefore, future investigations into the impact these indels may or may not have on *L. niloticus ldh-b* transcript abundance (i.e. expression), as compared to the variation observed among thermally discrete *L. calcarifer* population (Chapter 4), is warranted.

Numerous simple sequence repeat (SSR) sequences were detected within *ldh-b* introns of both *Lates* species (Table 3.2). Intronic SSRs are of interest because they may regulate gene transcription, lead to abnormal splicing and disrupt export of mRNA to the cytoplasm (Li et al., 2004; Mattick, 1994; Yue et al., 2001). Four mononucleotide SSRs (T, C and A), ranging from 5 to 11 repeats, are present within intron 1 of both *Lates* spp. and these were conserved in size between species (Table 3.2). In addition to these, two mononucleotide repeats (T_{10} and $C_{7\text{-}11}$) are also present exclusively in *L. calcarifer* introns 1 and 6, respectively (Table 3.2). One dinucleotide repeat (AC) is present at the same location within intron 6 of both *L. calcarifer* and *L. niloticus*, but it differs in repeat number between the species (AC_5 and AC_8, respectively) (Table 3.2). In addition there is one trinucleotide repeat (CAA) present within intron 1 of *L. niloticus* and *L. calcarifer* which varies between 3 and 4 repeats respectively (Table 3.2). Another trinucleotide repeat (TCC_4) is present within a region in intron 2 and is conserved between both *Lates* species (Table 3.2). Noteworthy is that a similar (TCC_4) SSR was identified in the 5' flanking untranslated region (UTR) of the *ldh-b* locus in *F. heteroclitus* and, more importantly, that variation in repeat number of this SSR in the 5' UTR of *ldh-b* impacted the level of *ldh-b* transcription (i.e. gene expression) observed in thermally discrete *F. heteroclitus* populations (Crawford and Powers, 1992; Crawford et al., 1999a; Schulte et al., 2000; Schulte et al., 1997; Segal et al., 1996). Moreover, this region within the 5' proximal promoter of *F. heteroclitus*, in

addition to the 6fp and Sp1 binding sites also identified, were concluded to be under functional constraint by way of a phylogenetic analysis on the nucleotide sequences of these 5' UTR regulatory motifs which clearly differentiated the cold northern from the warm southern population (Crawford et al., 1999b; Pierce and Crawford, 1997). Lastly, a tetranucleotide repeat ($TGTA_4$) is observed in a region of intron 6 exclusive to *L. calcarifer* (Table 3.2). The variation and potential functional role, if any, of these SSRs on *ldh-b* gene expression itself and/or on the greater transcriptome functionality within and among thermally discrete *L. calcarifer* and *L. niloticus* populations is currently unknown; however, further investigation into such potential implications on the transcriptome is warranted.

Table 3.2. Simple sequence repeat (SSR) motifs present within intron sequences of *Lates* species

SSR Motifs	Location	# of Repeats: *L. calcarifer*	# of Repeats: *L. niloticus*
Mononucleotide			
T^	Intron 1	10	-
C	Intron 1	6	5
C	Intron 1	8	5
T	Intron 5	8	5
A	Intron 6	7	8
C^	Intron 6	7-11	-
Dinucleotide			
AC	Intron 6	5	8
Trinucleotide			
CAA	Intron 1	4	3
TCC*	Intron 2	4	4
Tetranucleotide			
$TGTA^{\wedge}$	Intron 6	4	-

^: Indicates repeats exclusive to *L. calcarifer*. *: Indicates a simple sequence repeat motif identified within the 5' proximal promoter of *ldh-b* in *F. heteroclitus*, of which the repeat number impacts *ldh-b* expression (Crawford et al., 1999b; Segal et al., 1996).

Numerous potential microRNA elements (miRNAs) were also identified within *Lates* species intron sequences (Table 3.3). The encoding of thirty putative miRNA elements were identified within conserved intron regions and these had a score, energy (kcal mol^{-1}), z-score and homology ranging from 102 to 140, -26.96 to -19.01, 5.12 to 14.94 and 61.76 to 87.5%, respectively (see Methods). Four of the putative miRNA elements identified (Table 3.3: *miR-let7b*, *miR-124, miR-181* and *miR-223*) have previously been associated with specific regulatory functions. Of these, *miR-let7b* and *miR-223* knockout in mouse HeLa cells (Selbach et al., 2008) and neutrophils (Baek et al., 2008) resulted in reduced expression of approx. 2,700 and approx. 3,800 proteins, respectively. Two additional sub-families (*d* and *e*) from the highly investigated *let-7* miRNA element family were identified within *Lates* spp. *ldh-b* introns, which provides additional data consistent with the widespread dispersal and high abundance of targets and/or functions of miRNA elements encoded by introns throughout the genome (Baek et al., 2008; Bartel, 2004; Brennecke et al., 2005; Selbach et al., 2008) (Table 3.3). Noteworthy is that the *dre-let-7 (-b, -c, -d)* miRNA elements have also been recently identified within the 5' UTR of the *L. calcarifer* myostatin gene (*mstn*) (De Santis et al., 2008).

Table 3. Putative microRNA (miRNA) motifs identified within *ldh-b* intron sequences of *Lates* spp. Score: nucleotide match score; Energy: affinity to bind measure; Z-Score: statistical assessment of match quality; Homology: motif homology to query intron sequence

miRNA Family	Sub-families	Score	Energy (kcal mol^{-1})	Z-Score	Homology (%)
let-7	b, d, e	104 to 128	-22.65 to -19.17	5.12 to 11.04	65.52 to 86.36
miR-10	b, c, d	102 to 127	-24.50 to -19.01	5.75 to 10.02	65.38 to 76.00
miR-15	a	109 to 128	-21.88 to -20.64	7.29 to 12.20	66.67 to 82.61
miR-23	a, b	113 to 128	-20.19 to -19.53	10.02 to 12.92	75.00 to 80.00
miR-24	-	105 to 112	-19.81 to -19.29	6.72 to 8.75	75.00 to 80.00
miR-25	-	113 to 122	-25.79 to -20.77	8.52 to 10.96	72.41 to 85.71
miR-27	c, e	106 to 130	-24.44 to -19.36	6.91 to 13.77	76.92 to 83.33
miR-101	a	109 to 110	-20.02 to -19.23	7.18 to 7.91	68.97 to 73.91
miR-103	-	111 to 114	-22.68 to -19.15	6.84 to 8.19	73.08 to 87.50
miR-107	-	104 to 110	-21.16 to -20.59	5.56 to 7.31	67.86 to 70.37
miR-122	-	109 to 130	-24.44 to -19.67	6.35 to 12.72	64.71 to 87.50
miR-124	-	126 to 133	-21.18 to -19.68	11.54 to 13.73	71.43 to 72.41
miR-125	a, b	102 to 118	-21.84 to -19.47	6.10 to 10.16	64.00 to 76.92
miR-138	-	106 to 117	-20.40 to -19.19	7.11 to 9.45	65.62 to 81.82
miR-140	-	104 to 115	-23.55 to -19.32	6.17 to 9.55	70.00 to 79.17
miR-148	-	109 to 115	-23.74 to -19.82	7.21 to 9.3	86.36
miR-152	-	102 to 140	-25.95 to -19.30	5.75 to 14.94	69.57 to 84.00
miR-181	a, b	103 to 112	-21.00 to -19.61	5.25 to 8.89	61.76 to 76.00
miR-183	-	118 to 127	-26.75 to -19.73	8.69 to 11.3	68.75 to 77.78
miR-184	-	101 to 128	-24.99 to -20.76	5.73 to 12.28	70.37 to 80.00
miR-187	-	103 to 122	-22.26 to -19.02	6.93 to 13.45	73.08 to 81.82
miR-196	a, b	107 to 127	-22.22 to -19.98	6.51 to 11.52	73.08 to 83.33
miR-202	-	107 to 126	-24.76 to -19.41	7.42 to 11.74	68.97 to 82.61
miR-210	-	102 to 127	-26.96 to -19.62	6.08 to 12.54	68.00 to 82.61
miR-214	-	102 to 114	-21.45 to -19.68	6.78 to 9.93	70.37 to 77.27
miR-216	b	111 to 112	-19.16 to -19.15	7.98 to 8.68	72.41 to 77.78
miR-217	-	112 to 132	-26.17 to -19.47	7.67 to 12.26	70.37 to 84.00
miR-221	-	105 to 133	-21.67 to -19.42	5.59 to 12.58	70.00 to 81.48
miR-222	-	105 to 127	-25.88 to -19.36	5.13 to 10.61	70.00 to 76.92
miR-338	-	121	-19.65	10.32 to 11.04	75.00

Five intronic single nucleotide polymorphisms (SNPs) also were identified in intronic regions of the *ldh-b* locus of individuals representing eight discrete *L. calcarifer* populations in a pilot screening (Edmunds, R.C. PhD Thesis, 2009). Interestingly, these SNPs are present at sites where multiple miRNA elements overlap (Edmunds, R.C. PhD Thesis, 2009). The fact that a single SNP can impact multiple putative miRNA elements concurrently may provide insight in regard to recent findings of genome-wide association studies, which show a relationship between such SNPs and variation in complex human behaviour (e.g. schizophrenia and bipolar disorder) and/or susceptibility to complex diseases (e.g. type 2 diabetes and Crohn's disease) (Duerr et

al., 2006; Hirschhorn and Daly, 2005; Scott et al., 2007; Sladek et al., 2007; Wang et al., 2005). Future studies could utilize cross-species hybridized microarray methodology (Kassahn et al., 2007) as a novel mechanism for the identification of new candidate genes (e.g. concurrently up- and/or down-regulating) for the investigation as to the impact that such intron embedded regulatory motifs and/or elements (e.g. SSRs and/or miRNAs) may have on the active expression (i.e. transcript abundance) of those chosen and subsequently characterized candidate loci.

Evidence for selective constraint on non-coding (intron) sequences

The putative encoding of known functional motifs and/or elements (SSRs and miRNAs) by intronic regions of loci, in addition to the even distribution of such motifs and/or elements throughout longer introns (Haddrill et al., 2005), provides evidence that functional constraint is arguably acting on these historically less characterized non-coding regions. Indeed, such functional constraint is more likely to occur for introns of loci whose coding nucleotide or amino acid sequences are known to be under selection (Ambros, 2004; Baek et al., 2008; Bartel, 2004; Hare and Palumbi, 2003; Lai, 2003; Li et al., 2004; Mattick, 1994; Reisman et al., 1988; Wittekindt et al., 2000; Yue et al., 2001). More specifically, longer introns (> 87 bp) have been found to exhibit less divergence than shorter introns (< 87 bp) because of either an increased likelihood of embedded functional motifs being present in longer introns or the potential impact any mutations may have on the secondary structure of precursor messenger RNA (pre-mRNA) (Haddrill et al., 2005). A recent pairwise and cross-taxa comparison of intron sequences between three mammalian species (human, whale and seal) revealed sequence homologies 14% and 12% higher, respectively, than those expected from a neutral model of evolution based on expected rates of substitution for non-coding DNA, further suggesting functional constraint acting upon non-coding intronic sequences

(Hare and Palumbi, 2003). Moreover, the existence of numerous intron motifs and/or elements is known to be essential for the functioning of complex multi-cellular organisms, as they permit a two-fold regulatory system in eukaryotic organisms: one for the transcriptome (Ambros, 2004; Bartel, 2004; Nakaya et al., 2007) and one for the proteome (Baek et al., 2008; Chen et al., 2004; Selbach et al., 2008). Regardless, further research is required to determine if these motifs and/or elements (SSRs and/or miRNAs) have similar target sites and/or impacts on the transcriptome (i.e. gene expression) and/or the proteome (i.e. gene silencing) within and among *Lates* species and/or fish in general.

FIGURES

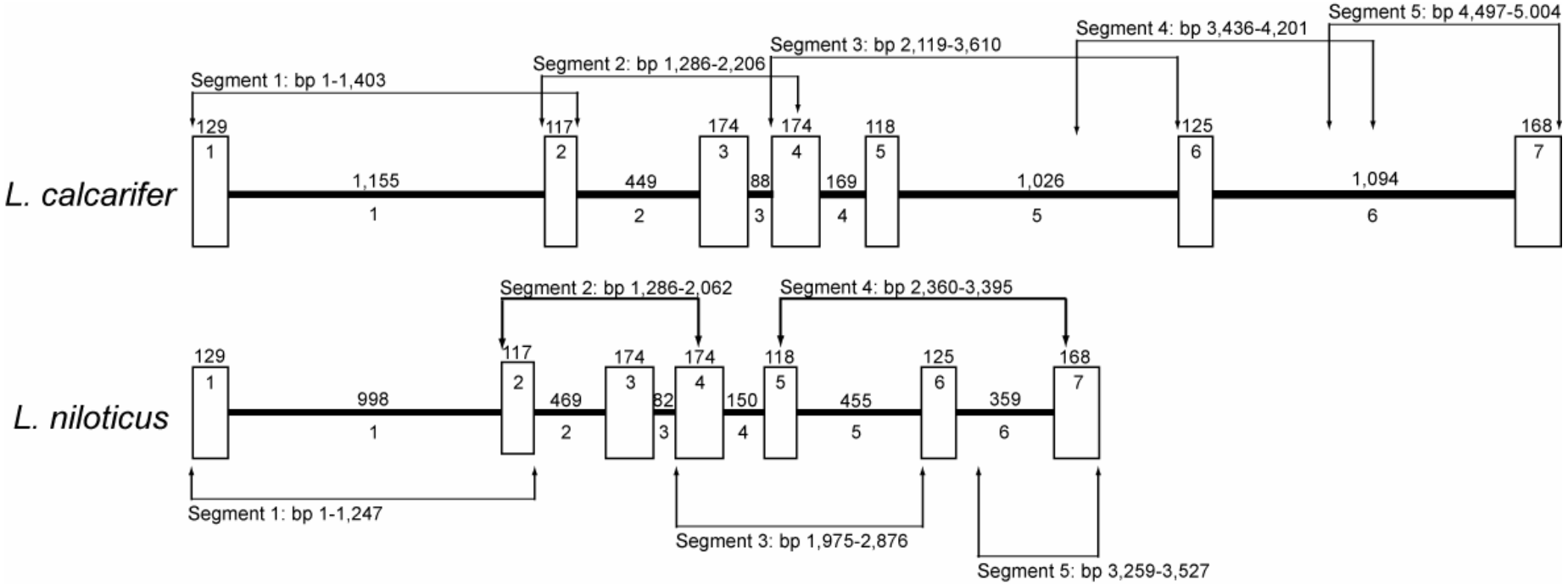


Figure 3.1. Comparative *ldh-b* gene map of congeneric *Lates* spp. Exon (white boxes) and intron (black bars) sizes in number of base pairs (bp) are given above their respective graphic representation. Sequential numbering within white boxes and below black bars is for exons and introns, respectively. Arrowhead lines indicate region amplified by specific primer pairs (as per Table 3.1) along with the size of each segment in number of base pairs.

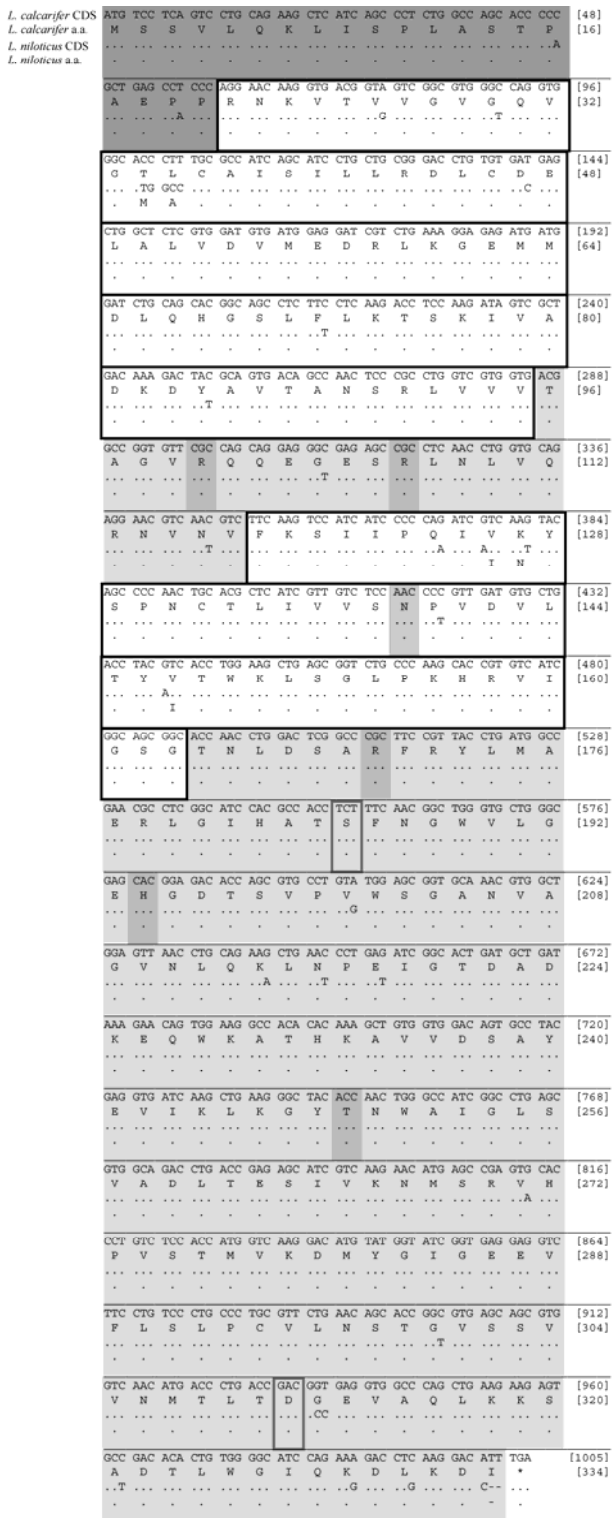


Figure 3.2. Coding nucleotide sequence (CDS) for *L. calcarifer* and *L. niloticus* are given on lines 1 and 3, respectively. Deduced amino acid sequence (a.a.) for *L. calcarifer* and *L. niloticus* are given on lines 2 and 4, respectively. Bracketed numbers at ends of line 1 and 2 refer to nucleotide position and amino acid residue, respectively. NH_2-terminal arm (residues 1-20) is dark-gray box outlined. Coenzyme binding domains (residues 21-95 and 118-163) are boldface box outlined. Substrate binding domain (residues 164-333) and loop helix ά D region (residues 96-117) are light-gray box outlined. Other known functional sites (residues 100, 107, 139, 170, 194 and 249) within domains are medium-gray box outlined. Residues (185 and 311) of known fixed difference between *F. heteroclitus* populations are boldface box outlined (see Results).

3.5: CONCLUSIONS

The *ldh-b* locus was found to be highly conserved between two tropical perciformes, *L. calcarifer* and *L. niloticus,* with just 2.9% divergence of coding regions and five amino acid differences between deduced LDH-B protein sequences. Variation within the nucleotide binding region of the translated protein sequences may possibly confer a variation in the specific functioning (i.e. binding affinity) of the LDH-B enzyme between the geographically isolated *L. calcarifer* and *L. niloticus* and warrants further exploration. Non-coding (intron) sequences of the *ldh-b* locus in both *Lates* species were as conserved as the coding regions of this gene, despite comprising 72.5 to 80% of the entire gene sequence. Ten SSR motifs and thirty putative miRNA elements were observed within the introns of the *ldh-b* locus in both *Lates* species. These putative regulatory elements and/or motifs warrant further investigation for their potentially functional importance in the regulation and/or expression of *ldh-b* or other constituent loci contributing to the transcriptome and proteome. A pilot population genetics screening of the newly characterized *ldh-b* locus among and within discrete *L. calcarifer* populations is required to determine if there is any evidence suggestive of thermal adaptation in *ldh-b* among populations inhabiting different thermal environments in the wild. This is the subject of the following chapter (Chapter 4).

Chapter 4

Pilot population screening of the *ldh-b* locus within and among discrete populations of the Australian barramundi, *Lates calcarifer* (Bloch 1790)

4.1: ABSTRACT

Assessing genomic variation across discrete populations of fish can provide essential insight regarding their genetic variation. This chapter presents a pilot population genetics screening of the lactate dehydrogenase-b (*ldh-b*) locus across geographically, genetically and thermally distinct populations of the Australian barramundi, *Lates calcarifer*. We found that the southernmost population (Gladstone, Queensland) differed significantly ($p < 0.05$) from all other screened populations ($n = 7$) with F_{ST} values ranging from 0.12 to 0.30, making it the ideal target population for future research on this thermally sensitive species. Two parsimoniously informative single nucleotide polymorphisms (SNPs) are identified within the coding (exon) sequence of the *ldh-b* locus across discrete populations; however both are silent mutations (i.e. render synonymous amino acid substitutions). Within the previously characterized non-coding (intron) regions, five SNPs are identified with two occurring in regions which contain multiple overlapping putative microRNA (miRNA) elements. The conservation of these intronic miRNA elements warrants further investigation regarding their functional importance in the regulation and/or expression of the *ldh-b* locus as a constituent of the transcriptome and the proteome. In conclusion, the pilot population genetics screening presented in this chapter identifies a minimal degree of nucleotide and amino acid sequence variation in the *ldh-b* locus within and among geographically, genetically and thermally discrete Australian *L. calcarifer* populations.

4.2: INTRODUCTION

Genetic variation within coding regions (exons) is expected to be reduced when compared to the observed level of variation within non-coding regions (introns) since coding regions are known to be under functional constraint. However introns, despite being considered non-functional and therefore less constrained, are now arguably as functionally constrained as exons (Ambros, 2004; Baek et al., 2008; Bartel, 2004; Bonnet et al., 2004; Brennecke et al., 2005; Chen et al., 2004; Enright et al., 2003; Hare and Palumbi, 2003; Hobert, 2004; Lai, 2003; Li et al., 2004; Mattick, 1994; Meloni et al., 1998; Roger and Doolittle, 1995; Wittekindt et al., 2000). In line with this view on intron conservation due to functional constraint, the lactate dehydrogenase-B locus (*ldh-b*) exhibited a near equal level of intron and exon conservation (97.7% and 97.1% respectively) between Australian *Lates calcarifer* (barramundi) and its African congeneric *L. niloticus* (nile perch) (Chapter 3). To our knowledge this finding is unique, at least for fish, even though this locus has been extensively investigated across thermally distinct populations of the temperate North American estuarine killifish *Fundulus heteroclitus* (see reviews by (Powers et al., 1991; Powers and Schulte, 1998) and the Puget Sound crested blenny *Anoplarchus purpurescens* (Johnson, 1971, 1977). Neither the killifish nor the crested blenny investigations considered the level of intronic conservation as these intervening regions have yet to be characterized in these species. In light of this, as well as the highly conserved intron and exon regions of *ldh-b* between *Lates* spp., the aim of this chapter was to conduct a pilot population screening of the coding and non-coding regions accounting for 72.3% of this locus (from the 3' end of exon 2 to the 5' end of exon 7) across geographically, genetically and thermally discrete populations of *Lates calcarifer*, the first such screening to be conducted for full-length *ldh-b* on any fish species.

Low levels of variation are expected for the *ldh-b* locus across barramundi populations because an allozyme screening conducted by Keenan (1994) disregarded LDH-B based on a lack of detectable allozyme differences between the 6,000 individuals assayed, which represented fifty locations spanning across Australia from Gladstone (southern Queensland) up the east coast, around the tip of Cape York, across the Gulf of Carpentaria and westward to Darwin (Northern Territory). However, this 1994 study did not screen populations from the Western Australian clade as this was established by Doupé *et al*. in 1999 using mitochondrial control region 1 (CO1) based phylogenetics (Chapter 4). The lack of detectable LDH-B allozyme difference between these discrete populations from similar thermal environments does not eliminate the possibility that unique and advantageous genetic differentiation exists between the geographically, genetically and thermally distinct Australian populations of this tropical perciform species. As such, the pilot population screening presented in this chapter assessed not only the coding regions (exons) but also the non-coding regions (introns) of the *ldh-b* locus. Moreover, screening *ldh-b* introns allowed for the identification of regulatory motifs and/or elements, such as simple sequence repeats (SSRs) and/or microRNAs (miRNAs), which are known to be embedded in or encoded by non-coding regions and have been demonstrated to reside within introns of other genes, where they have been shown to affect either the transcriptome (transcribed RNA present) or the proteome (translated protein) (Ambros, 2004; Baek et al., 2008; Bartel, 2004; Bonnet et al., 2004; Brennecke et al., 2005; Doench and Sharp, 2004; Enright et al., 2003; Hobert, 2004; Lai, 2003; Lambowitz and Belfort, 1993; Li et al., 2004; Meloni et al., 1998; Nakaya et al., 2007; Rehmsmeier et al., 2004; Reisman et al., 1988; Wittekindt et al., 2000; Yue et al., 2001).

Lates calcarifer resides in rivers, estuaries and shallow marine environments throughout northern Australia (25°S – 12°S) and the south-east Asian archipelago

(13°N – 10°S) (Keenan, 2000) at different stages of its lifecycle (Katersky and Carter, 2005, 2007). Significant genetic partitioning (i.e. lack of gene flow) has previously been established between each major river system in Australia by way of allozyme assays (Keenan, 1994; Shaklee and Salini, 1985) and microsatellite screenings (Chenoweth et al., 1998; Doupé et al., 1999). Genetic partitioning is linked to the non-migratory nature of barramundi associated with major Australian river ways (Chenoweth et al., 1998; Doupé et al., 1999; Katersky and Carter, 2005, 2007; Keenan, 1994, 2000; Salini and Shaklee, 1988). Therefore, each discrete *L. calcarifer* population is likely to have accumulated unique adaptations to the environmental parameters (e.g. thermal regime) of their native river. In this chapter we determine if there is any evidence that sequence variation within either coding or non-coding regions of *ldh-b* is partitioned into the genetic lineages identified most recently within eastern and northern Australia by Chenoweth (1998) and within Western Australia by Doupé (1999), both of which utilized putatively neutral microsatellite markers. Moreover, we utilize a population genetics approach to detect genotypic differences across coding and non-coding *ldh-b* sequences, both within and among *L. calcarifer* populations to indicate particular populations which possess a reduction in genetic variation and alternative variants of putative functional motifs, which are of particular interest if genetic lineages correspond with environmental variables of the populations studied, such as temperature.

4.3: METHODS

Genomic DNA was extracted from *L. calcarifer* samples collected from eight locations within tropical Australia. Samples from Gladstone, Queensland (23°S, 151°E; n = 5) and Good Fortune Bay (Bowen, Queensland: 20°S, 148°E; n = 5) were obtained directly

from fish farms while Archer River (Cape York, Queensland: 13°S, 142°E; *n* = 5), Daly River (southwest of Darwin, Northern Territory: 13°S, 130°E; *n* = 3), Ord River (north of Halls Creek, Northern Territory: 17°S, 127°E; *n* = 5), Roebuck Bay (Broome, Western Australia: 17°S, 122°E; *n* = 4), DeGrey River (north of Port Headland, Western Australia: 20°S, 119°E; *n* = 4) and the Ashburton River (south of Paraburdoo, Western Australia: 23°S, 117°E; *n* = 5) samples originated from wild caught fish (Figure 4.1). Fin-clips from all fish were preserved in ethanol (80%). Genomic DNA (gDNA) was extracted via proteinase-K digestion (20 mg/mL) in CTAB buffer at 60°C for 1 hr and subsequently cleaned with a salt and chloroform:isoamyl alcohol (24:1) procedure (Sambrook, 2002). All extractions resulted in high molecular weight gDNA, as visualized on a 0.8% agarose gel, with quantities ranging from 20 – 100 ng/µL, as was reported for such extractions in Chapter 3.

Partial *ldh-b* gene sequences (3' end of exon 2 through 5' end of exon 7 including intervening intron sequences) were obtained from PCR amplifications off gDNA using the primers and primer specific PCR conditions outlined in Chapter 3 (for specific primer binding locations within the *ldh-b* locus see Figure 3.1**)**. Thermocycling parameters were the same as those utilized in Chapter 3 with the exception of the annealing temperature and $MgCl_2$ concentration which varied as listed in the *Lates* spp. primer table (see Table 3.1). All PCR products obtained from *L. calcarifer* gDNA were verified as *ldh-b* by direct alignment to the full length *Lates* spp. (*n* = 4 *L. calcarifer* and *n* = 2 *L. niloticus*) *ldh-b* sequences (Chapter 3). Partial *ldh-b* nucleotide sequence data was aligned for all individuals (*n* = 36) in BioEdit (Hall, 1999) and assessed therein for parsimoniously informative single nucleotide polymorphisms (SNPs) as well as for the conservation of those simple sequence repeats (SSRs) and micro RNA elements (miRNAs) identified in Chapter 3.

A maximum parsimony phylogenetic analysis was conducted on the partial *ldh-b* sequences (n = 36) and out-group rooted with one *L. niloticus ldh-b* sequence (Chapter 3), which was truncated to match the partial length of *ldh-b* utilized in this screening, was conducted in Mega 3.1 (Kumar et al., 2004). Furthermore, this phylogenetic MP analysis utilized all parsimony informative sites within the partial *ldh-b* sequence (n = 7/3,637). Among population pairwise F_{st} comparisons and subsequent analyses of molecular variance (AMOVA) were employed using Arlequin 2.0.1.1 (Schneider et al., 2000) on samples grouped by: a) thermal regime (Ashburton, DeGrey, Gladstone and Good Fortune Bay vs. Roebuck Bay, Ord, Daly and Archer Rivers); and b) established genetic clades (Queensland: Archer River, Good Fortune Bay and Gladstone; Northern Territory: Ord and Daly Rivers; Western Australia: Roebuck Bay, DeGrey and Ashburton Rivers) to identify if populations are genetically distinct and if so, at what level are they partitioned.

4.4: RESULTS

Descriptive characterization of obtained *ldh-b* locus sequences

This chapter presents the first population screening of the *ldh-b* locus across discrete populations of the tropical estuarine fish *Lates calcarifer*. The partial gene sequences utilized in this pilot population screening included base pairs 1,369 to 4,981 of the 5,004 bp *ldh-b* locus (Chapter 3). Moreover, this partial sequence covered 769 bp and approx. 255 amino acids of the 1,005 bp and 334 aa which comprise the coding nucleotide and amino acid sequences, respectively (Chapter 3). As was observed for both *Lates* spp. in Chapter 3, as well as with other fish species, the *ldh-b* locus from all *L. calcarifer* populations contained seven exons and six introns. Of these, six exons and five introns were screened and found to be conserved in size across all *L. calcarifer* populations, as

expected. Exon nucleotide sequence comparison revealed two parsimoniously informative SNPs at base pairs 1,890 and 2,256 (exons 3 and 4, respectively) (Figure 4.2). However, these SNPs were "silent" since no amino acid substitutions resulted amongst the variants. Otherwise all observed mutations were singletons (n = 22) occurring across all six screened exons. Interestingly, exon 5 was 100% conserved across all populations while exons 3, 4 and 6 only contained 3, 3 and 2 singletons, respectively. The majority of observed singletons coming from the terminal portions of exon 2 (bp 1,369 to 1,390) and exon 7 (bp 4,960 to 4,981) may be due to errors in the sequencing reads at respective proximal ends of these gDNA amplified *ldh-b* segments.

As was observed between congeneric *Lates* spp. in Chapter 3, similar levels of sequence homology are also observed for non-coding (intron) regions across discrete *L. calcarifer* populations. Introns 2, 3, 4, 5 and 6 (449, 88, 169, 1,023 and 1,103 bp respectively) contain 11, 0, 0, 32 and 18 singletons, respectively. Moreover, the only cross-population SNPs present within the non-coding region of *ldh-b* are in intron 5 (base pair 2, 907, 2, 943 and 2,960) and intron 6 (base pair 4, 804 and 4, 815) (Figure 4.3). However, the presence of these intronic SNPs spans across populations from the previously established genetic clades (Chenoweth et al., 1998; Doupé et al., 1999) thus rendering them non-informative (Figure 4.3).

When partial *ldh-b* sequences were assessed for pairwise nucleotide differences and nucleotide diversity within populations, the southern Queensland population (Gladstone) is of particular interest due to its exceptionally low values for both indices (0.9 ± 0.7 and 0.02 ± 0.02 respectively; Table 4.1). Moreover, when F_{ST} values are compared among populations, this population differs significantly ($p < 0.05$) from all others, with pairwise F_{ST} values ranging from 0.12 to 0.30 (Table 4.2). This genetic partitioning between Gladstone and other sampled sites is also reflected in the "among populations within groups" (F_{SC}) and "among populations" (F_{CT}) components of the

hierarchical analysis. However, when *ldh-b* sequences are grouped together by thermal regime (cool and warm tropical) or by established genetic clade (Queensland, Northern Territory and Western Australia) (see Methods) no significant partitioning among the defined groups is observed in analyses of molecular variance (AMOVAs; Table 4.3a, b respectively). Additionally, the MP consensus tree, which encompasses 833 total trees and utilizes a cut-off value of 50%, reflects this lack of significant genetic structure between discrete populations, even when grouped by thermal regimes and/or established clades (Figure 4.4). However, the smaller clade almost exclusively contains individuals from Western Australia and Northern Territory populations (Ashburton, DeGrey and Daly Rivers), a structure likely due to these populations sharing ancestry with *L. calcarifer* from more equatorial Indo-Australian Archipelago and Southeast Asia populations, as gene flow between these populations and eastern Australian populations has been evolutionarily inhibited by the Torres Straight land-bridge between Cape York and Papua New Guinea (Chenoweth et al., 1998; Salini and Shaklee, 1988; Zhu et al., 2006). Furthermore, these more equatorial populations are arguably a unique sub-species to the Australian and Papua New Guinea *L. calcarifer*, a recent finding based on DNA bar-coding loci (Ward et al., 2008).

Table 4.1. Pairwise differences and nucleotide diversity within *L. calcarifer* populations

Population	*n*	# of Pairwise Differences	Nucleotide diversity (%)
Ashburton River, WA	5	12.3 ± 6.7	0.3 ± 0.2
DeGrey River, WA	4	5.0 ± 3.1	0.1 ± 0.1
Roebuck Bay, WA	4	3.2 ± 2.1	0.09 ± 0.07
Ord River, NT	5	1.8 ± 1.2	0.05 ± 0.04
Daly River, NT	3	6.7 ± 4.3	0.2 ± 0.1
Archer River, N. QLD	5	5.2 ± 3.0	0.1 ± 0.1
Good Fortune Bay, C, QLD*	5	4.2 ± 2.5	0.1 ± 0.08
Gladstone, S. QLD**	5	0.9 ± 0.7	0.02 ± 0.02

Individuals obtained from central and southern Queensland aquaculture facilities are indicated by * and ** respectively.

Table 4.2. Pairwise F_{ST} differences among populations

	Ash	DeGy	Roe	Ord	Daly	Arch	GFB	Glad
Ash	0							
DeGy	0.02045	0						
Roe	0.06737	-0.02083	0					
Ord	0.12935*	0.13698	0.2333*	0				
Daly	-0.04563	-0.01493	0.12878	0.24171*	0			
Arch	-0.04067	0.04659	0.07801	0.25439*	-0.01715	0		
GFB	0.10326	0.12058*	0.17322*	0.18956*	0.1508	0.16964	0	
Glad	0.15829*	0.11537*	0.17654*	0.3*	0.29155*	0.24757*	0.17188*	0

Asterisks (*) indicate a significantly differing pairwise F_{st} distances ($p < 0.05$).

Table 4.3. Analyses of Molecular Variance (AMOVAs)

	Source of variation	d.f.	Sum of Squares	Variance component	*F*-statistic (*p* value)	Percentage of variation
a	AG	1	2.833	-0.06278 (V_a)	F_{CT}: -0.02333 (0.85)	-2.33
	APWG	6	23.506	0.33567 (V_b)	F_{SC}: 0.12188 (0.004)	12.47
	WP	28	67.717	2.41845 (V_c)	F_{ST}: 0.10139 (0.002)	89.86
b	AG	2	6.531	-0.06094 (V_a)	F_{CT}: -0.02255 (0.73)	-2.25
	APWG	5	19.808	0.34497 (V_b)	F_{SC}: 0.12483 (0.001)	12.76
	WP	28	67.717	2.41845 (V_c)	F_{ST}: 0.10510 (0.003)	89.49

a: Populations grouped by cool and warm tropical regimes; *b*: populations grouped by previously established QLD, NT and WA genetic clades (see Methods).

Among population conservation of simple sequence repeat (SSR) motifs and microRNA (miRNA) elements

Non-coding (intron) regions of the *ldh-b* locus were assessed by eye for variation in the simple sequence repeat (SSR) motifs and microRNA (miRNA) elements identified in the previous chapter (Tables 3.2 and 3.3 respectively). However, due to the incomplete sequencing of this locus in individuals representing each distinct *L. calcarifer* population (72.3%), only those motifs and elements identified in introns 2 to 6 are assessed for inherent variation (Table 4.4).

The ten SSRs identified within *ldh-b* introns of *L. calcarifer* vary in repeat length (4 to 11) and encompass mono-, di-, tri- and tetranucleotide motif structures (Table 3.2). Of the six mononucleotide SSRs only three reside is assessable introns, with two (T_8 and A_7 in introns 5 and 6 respectively) exhibiting complete conservation across discrete populations. The third mononucleotide SSR, occurring in intron 6, does

vary in repeat length from 7 to 11 with the majority being C_7 across individuals representing the major genetic clades previously established for Australian barramundi by way of both allozyme and microsatellite variation (Chenoweth et al., 1998; Doupé et al., 1999; Keenan, 1994; Shaklee and Salini, 1985). Individuals from the Ashburton River in southern Western Australia show the greatest level of within-population repeat length variation with two C_7, one C_9 and two C_{11} motifs observed, followed by individuals from the Archer River in far north Queensland (Cape York) exhibiting three C_7, one C_9 and one C_{11} motifs. In addition, one C_9 and one C_{10} motif are observed in a single screened individual from Good Fortune Bay (northern Queensland) and the DeGrey River (central Western Australia), respectively. Otherwise, all individuals across all populations possess a C_7 repeat in intron 6 as well as complete conservation across all discrete populations for the dinucleotide (AC_5), trinucleotide (TCC_4) and tetranucleotide ($TGTA_4$) motifs identified in introns 6, 2 and 6 respectively (Table 3.2).

Of the 57 miRNA elements identified within introns 2 to 6, the highest abundances (19 and 27) reside within introns 5 and 6, respectively, which is expected due to these being the longest *ldh-b* introns in *L. calcarifer* (Table 4.4). Introns 2 and 4, being the third and fourth largest, also contain miRNA motifs (9 and 2 respectively; Table 4.4); however, no miRNA elements are identified within the shortest and most conserved (100%) of the *ldh-b* locus non-coding regions, intron 3. In regard to mutations, 30 singletons are observed across 16 of the 57 identified miRNA elements (Table 4.4) whereas one SNP in intron 5 (bp 2, 906) affects three discrete miRNA elements (miR-27c, miR-27e and miR-103) due to the overlapping nature of miRNA element sequences (Figure 4.3). This overlapping nature is further evident in three miRNAs (miR-10d-5, miR125a-2 and miR125b-2) that all identify in the same region of intron 5 (bp 3812 – 3833) as do two additional miRNAs (miR-148 and miR-152-2) which also overlap in another region of intron 5 (bp 3586 – 3606) (Table 4.4).

Table 4.4. Variation in miRNA element sequences across individuals from thermally discrete populations

miRNA Family and/or Element	miRNA datbase	Intron	Location	Size (bp)	Score	Energy (kcal mol^{-1})	Z-Score	Homology	Variation (SNPs/Singletons)
let-7 Family									
let-7b-1	dre, fru, tni	6	bp 4064-4089	28	128	-21.41	11.04	71.4%	-
let-7b-2	dre, fru, tni	6	bp 4737-4760	22	116	-22.23	8.12	86.4%	-
let-7b-3	dre, fru, tni	5	bp 3111-3138	28	116	-20.3	8.12	75.0%	-
let-7b-4	dre, fru, tni	6	bp 3737-3764	29	104	-19.82	5.2	65.5%	-
let-7b-5	xtr	6	bp 3758-3780	25	114	-19.17	7.89	80.0%	-
let-7d-1	dre, fru, tni	6	bp 4061-4082	24	119	-20.01	8.79	75.0%	-
let-7d-2	dre, fru, tni	5	bp 3223-3247	25	118	-20.2	8.55	80.0%	- / 1 (*a*)
miRNA-10 Family									
miR-10b-1	dre, fru, tni	5	bp 3505-3533	29	119	-19.01	9.14	75.9%	- / 5 (*b* x 4, *c* x 1)
miR-10b-2	dre, fru, tni	6	bp 3813-3832	24	115	-21.15	8.24	70.8%	-
miR-10c-1	dre, fru, tni	6	bp 3810-3833	24	118	-22.64	9.85	75.0%	-
miR-10d-1	dre	6	bp 3813-3833	24	123	-23.63	9.76	75.0%	-
miR-10d-2	dre	5	bp 3508-3533	28	120	-21.27	9.09	75.0%	- / 5 (*b* x 4, *c* x 1)
miR-10d-3	dre	6	bp 3787-3809	26	117	-19.5	8.43	73.1%	- / 1 (*d*)
miR-10d-4	fru, tni	6	bp 3785-3809	28	127	-20.67	9.22	75.0%	- / 1 (*d*)
miR-10d-5	fru, tni	6	bp 3812-3833	26	126	-24.5	9	73.1%	-
miR-10d-6	fru, tni	5	bp 3507-3533	29	125	-21.27	8.78	75.9%	- / 5 (*b* x 4, *c* x 1)
miRNA-15 Family									
miR-15a-1	dre	6	bp 4177-4198	23	120	-20.64	9.65	82.6%	- / 1 (*b*)
miR-15a-2	dre	6	bp 3980-4008	30	110	-21.88	7.29	66.7%	-
miR-24	dre, fru, tni	5	bp 3373-3396	24	105	-19.29	6.77	75.0%	-
miR-25	dre, fru, tni, xtr	6	bp 4143-4168	26	122	-25.79	10.96	80.8%	- / 1 (*c*)
miRNA-27 Family									
miR-27c	xtr	5	bp 2904-2927	21	113	-19.36	9.63	81.0%	1 (*b, e, f*) / 1 (*e*)
miR-27e	dre, fru, tni	5	bp 2902-2927	24	126	-20.99	11.65	83.3%	1 (*b, e, f*) / 1 (*e*)
miRNA-101 Family									
miR-101a-1	dre, fru, tni, xtr	6	bp 3921-3948	29	110	-19.23	7.64	69.0%	-
miR-101a-2	dre	6	bp 3847-3870	23	109	-19.55	7.39	73.9%	-
miR-103	dre, fru, tni, xtr	5	bp 2901-2927	26	111	-19.15	7.08	73.1%	1 (*b, e, f*) / 1 (*e*)
miR-107-1	dre, fru, tni, xtr	4	bp 2685-2712	28	110	-21.16	7	67.9%	-
miR-107-2	dre, fru, tni, xtr	5	bp 3260-3286	27	104	-20.59	5.56	70.4%	-
miR-124	dre, fru, tni, xtr	6	bp 3885-3914	28	126	-19.68	11.54	71.4%	-
miRNA-125 Family									
miR-125a-1	dre, fru, tni, xtr	6	bp 3951-3976	25	115	-21.28	9.02	76.0%	-
miR-125a-2	dre, fru, tni, xtr	6	bp 3812-3833	25	108	-21.03	7.53	64.0%	-
miR-125b-1	dre, fru, tni, xtr	6	bp 3950-3976	26	118	-20.77	9.79	76.9%	-
miR-125b-2	dre, fru, tni, xtr	6	bp 3812-3833	24	116	-21.84	9.34	70.8%	-
miR-148	dre, fru, tni	5	bp 3586-3606	22	109	-19.82	7.44	86.4%	-
miR-152-1	dre	2	bp 1581-1605	25	140	-25.95	14.94	84.0%	-
miR-152-2	fru,tni	5	bp 3586-3606	23	113	-19.82	8.27	82.6%	-
miRNA-181 Family									
miR-181a-1	dre, fru, tni, xtr	2	bp 1576-1600	25	108	-21	6.6	76.0%	- / 1 (*c*)
miR-181a-2	dre, fru, tni, xtr	5	bp 2675-2709	34	103	-20.49	5.4	61.8%	-
miR-181b-1	dre, fru, tni, xtr	5	bp 2678-2705	27	112	-20.59	8.3	74.1%	-
miR-181b-2	dre, fru, tni, xtr	2	bp 1536-1561	24	103	-19.61	6.11	75.0%	-
miR-183	dre, fru, tni, xtr	2	bp 1639-1669	32	119	-26.75	9.33	68.8%	- / 1 (*c*)
miRNA-196 Family									
miR-196a	dre, fru, tni	2	bp 1524-1546	25	127	-20.37	11.52	80.0%	-
miR-196b	dre, fru, tni	5	bp 3547-3571	26	117	-22.22	8.88	80.8%	-
miR-210-1	dre, fru, tni, xtr	2	bp 1410-1433	23	126	-24.74	12.06	82.6%	- / 3 (*b*)
miR-210-2	dre, fru, tni, xtr	2	bp 1576-1604	28	108	-26.96	7.56	71.4%	- / 1 (*c*)
miR-214-1	dre, fru, tni, xtr	2	bp 1631-1654	26	112	-21.45	9.39	73.1%	-
miR-216b	dre, fru, tni	4	bp 2357-2385	29	111	-19.15	7.98	72.4%	-
miR-217-1	dre, fru, tni	6	bp 3882-3906	25	128	-26.17	12.26	84.0%	-
miR-217-2	dre, fru, tni	5	bp 2769-2793	24	122	-19.47	10.79	79.2%	-
miR-217-3	xtr	6	bp 3877-3906	30	132	-26.17	10.82	76.7%	-
miR-217-4	xtr	5	bp 2785-2807	25	124	-22.57	9	80.0%	-
miR-221-1	dre, fru, tni, xtr	6	bp 4144-4171	27	133	-21.67	12.05	81.5%	- / 1 (*c*)
miR-221-2	dre, fru, tni, xtr	6	bp 3843-3869	25	126	-21.5	10.43	80.0%	-
miR-221-3	dre, fru, tni, xtr	5	bp 2858-2888	30	121	-19.42	9.28	70.0%	-
miR-221-4	dre, fru, tni, xtr	5	bp 3297-3323	27	119	-20.37	8.82	74.1%	-
miR-221-5	dre, fru, tni, xtr	2	bp 1687-1714	29	105	-19.52	5.59	72.4%	-
miR-222-1	dre, fru, tni, xtr	6	bp 3844-3869	26	127	-21.66	10.57	76.9%	-
miR-222-2	dre, fru, tni, xtr	6	bp 3983-4012	30	105	-23.19	5.41	70.0%	-

Identified variation is indicated as either a SNP or a singleton, as shown for the number of samples from discrete locations are as follows, *a*: Archer River, QLD (*n* = 5); *b*: Ashburton River, WA (*n* = 5); *c:* DeGrey River, WA (*n* = 4); *d*: Roebuck Bay, WA (*n* = 4); *e*: Daly River, NT (*n* = 3); *f*: Good Fortune Bay, QLD (*n* = 5). MicroRNA database abbreviations are as follows, dre: *Danio rerio*; fru: *Takifugu rubripes*; tni: *Tetraodon nigroviridis*; xtr: *Xenopus tropicalis*.

FIGURES

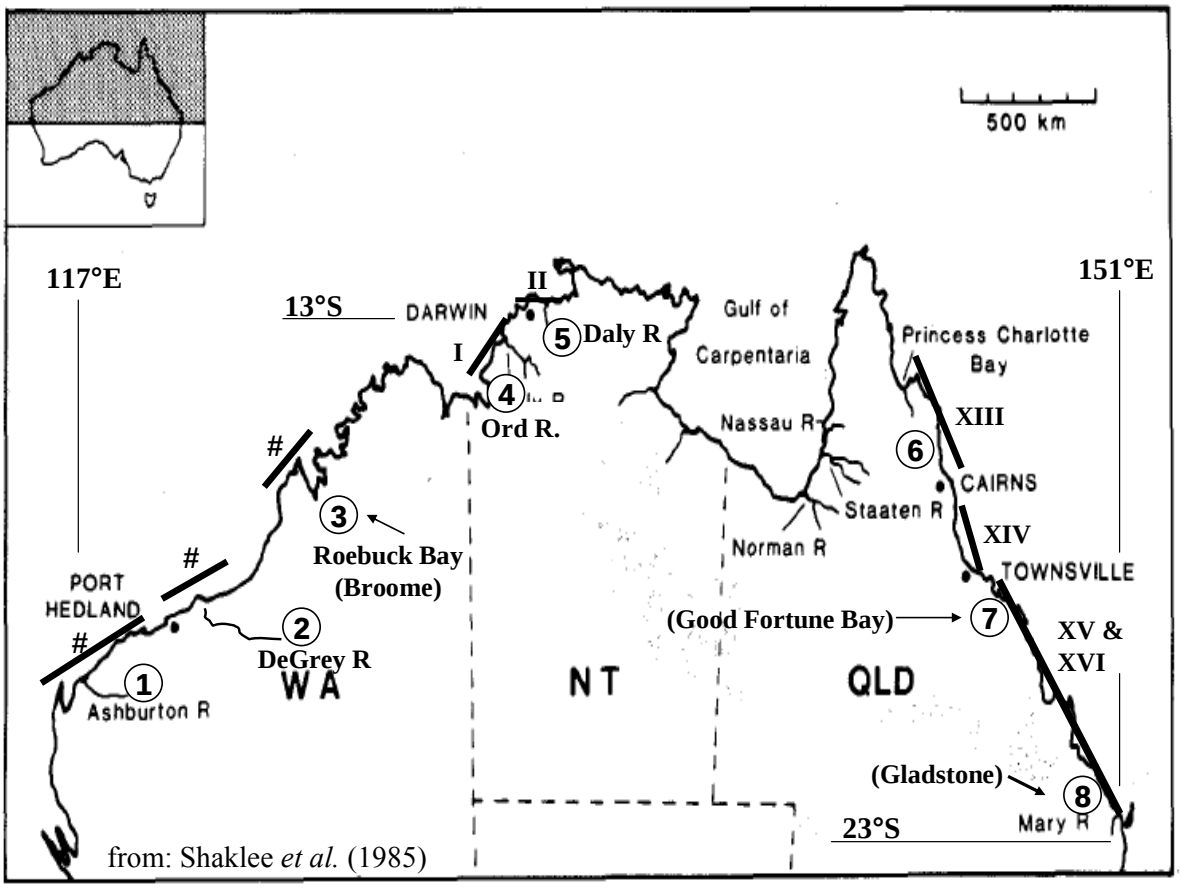


Figure 4.1. Map of *L. calcarifer* sampling locations across genetically distinct populations from tropical Australia (adapted with permission from Shaklee *et al*. 1985). Roman numerals and hash marks indicate unique genetic clades established by (Chenoweth et al., 1998; Keenan, 1994) and (Doupé et al., 1999) respectively.

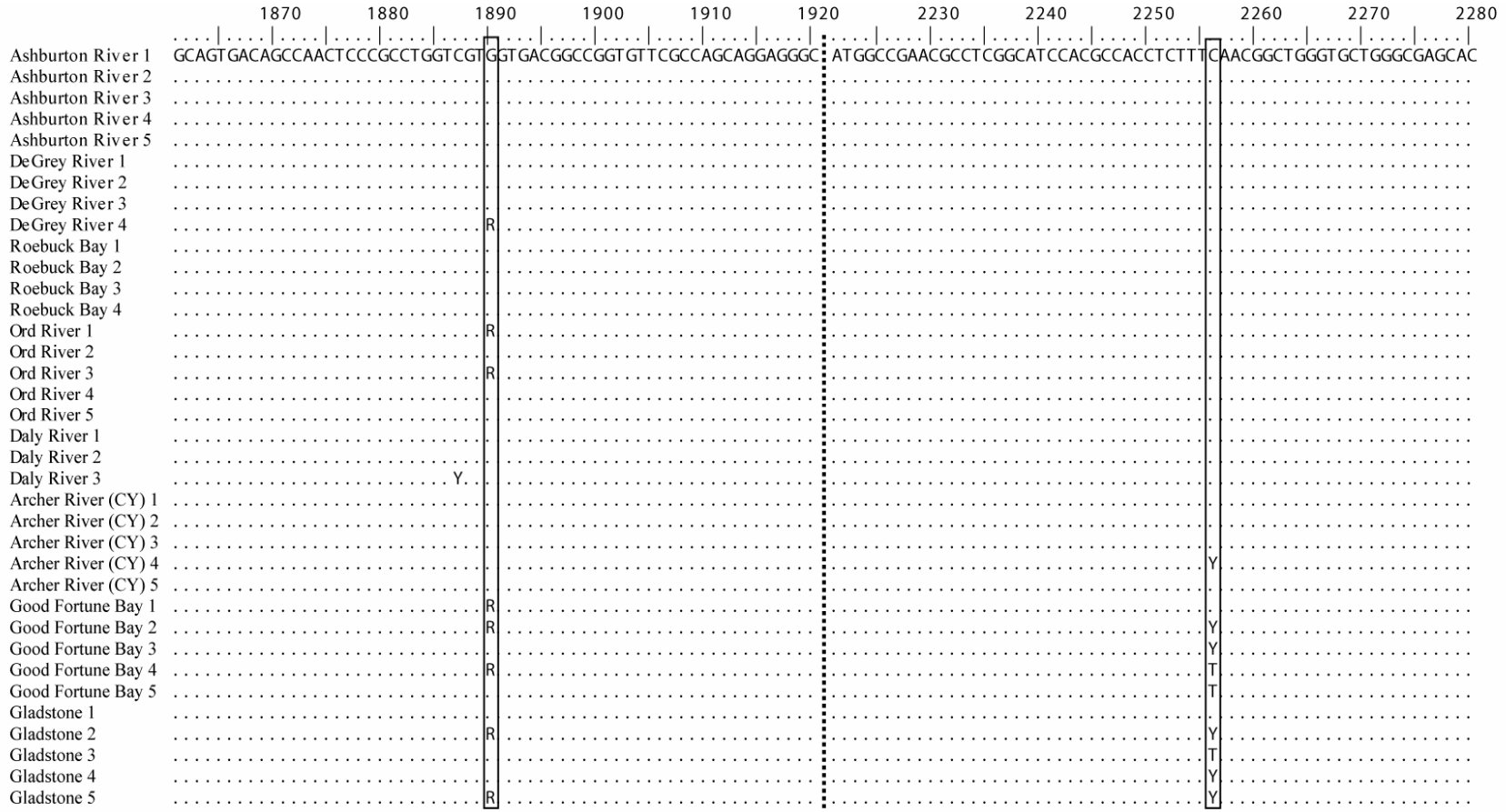


Figure 4.2. Two exon single nucleotide polymorphisms (SNPs) present across discrete *L. calcarifer* populations. R and Y refer to individuals heterozygous for a purine (A or G) and pyrimidine (C or T) at bp 1,890 and 2,256 respectively. A break in the continuity of the sequence at bp 1,921 is indicated by the vertical dashed line.

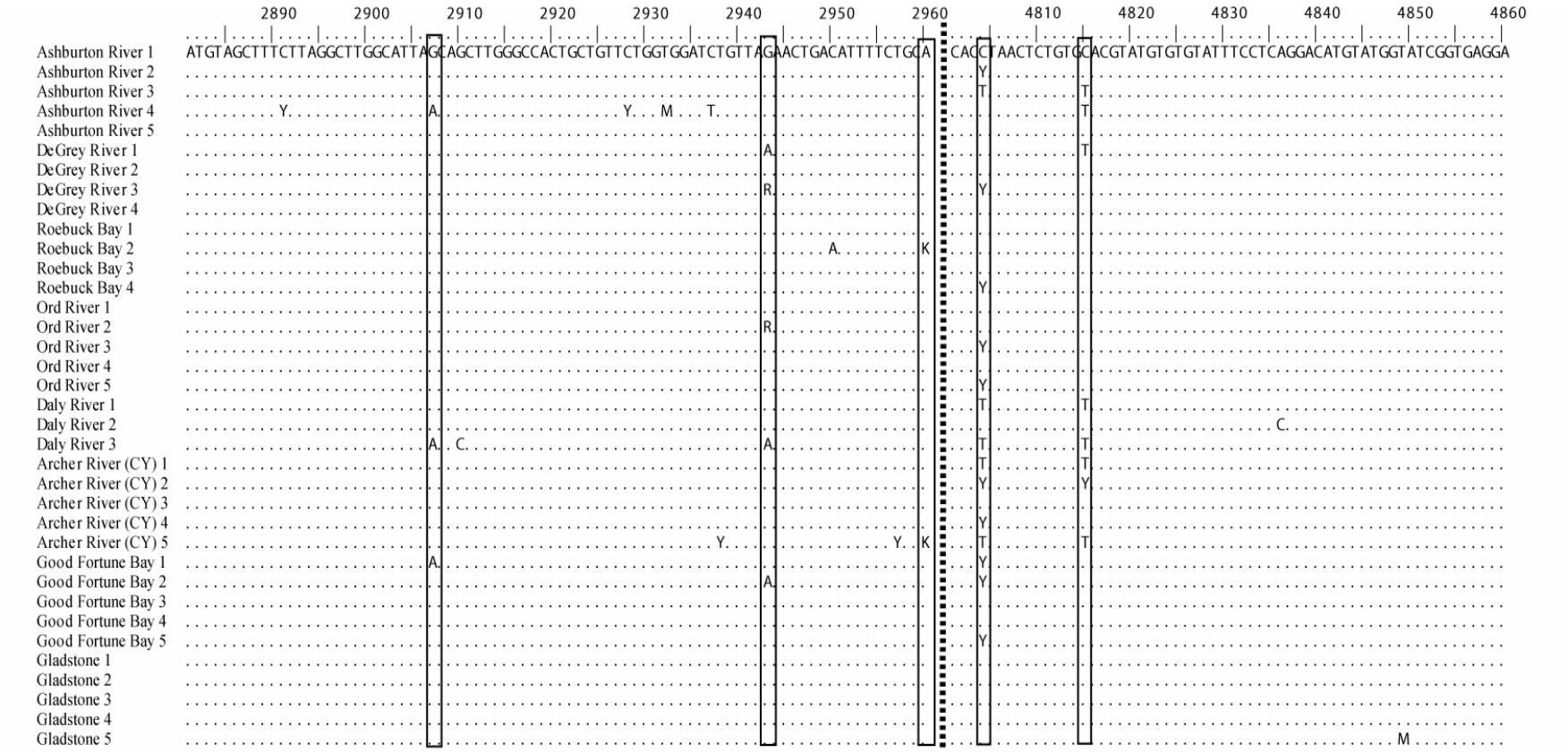


Figure 4.3. Five intron single nucleotide polymorphisms (SNPs) present across discrete *L. calcarifer* populations. R, K and Y refer to individuals heterozygous for a purine (A or G), a keto (G or T) and a pyrimidine (C or T) at bp 2, 943, 2,960 and 4, 804 respectively. A break in the continuity of the sequence at bp 2,963 is indicated by the vertical dashed line.

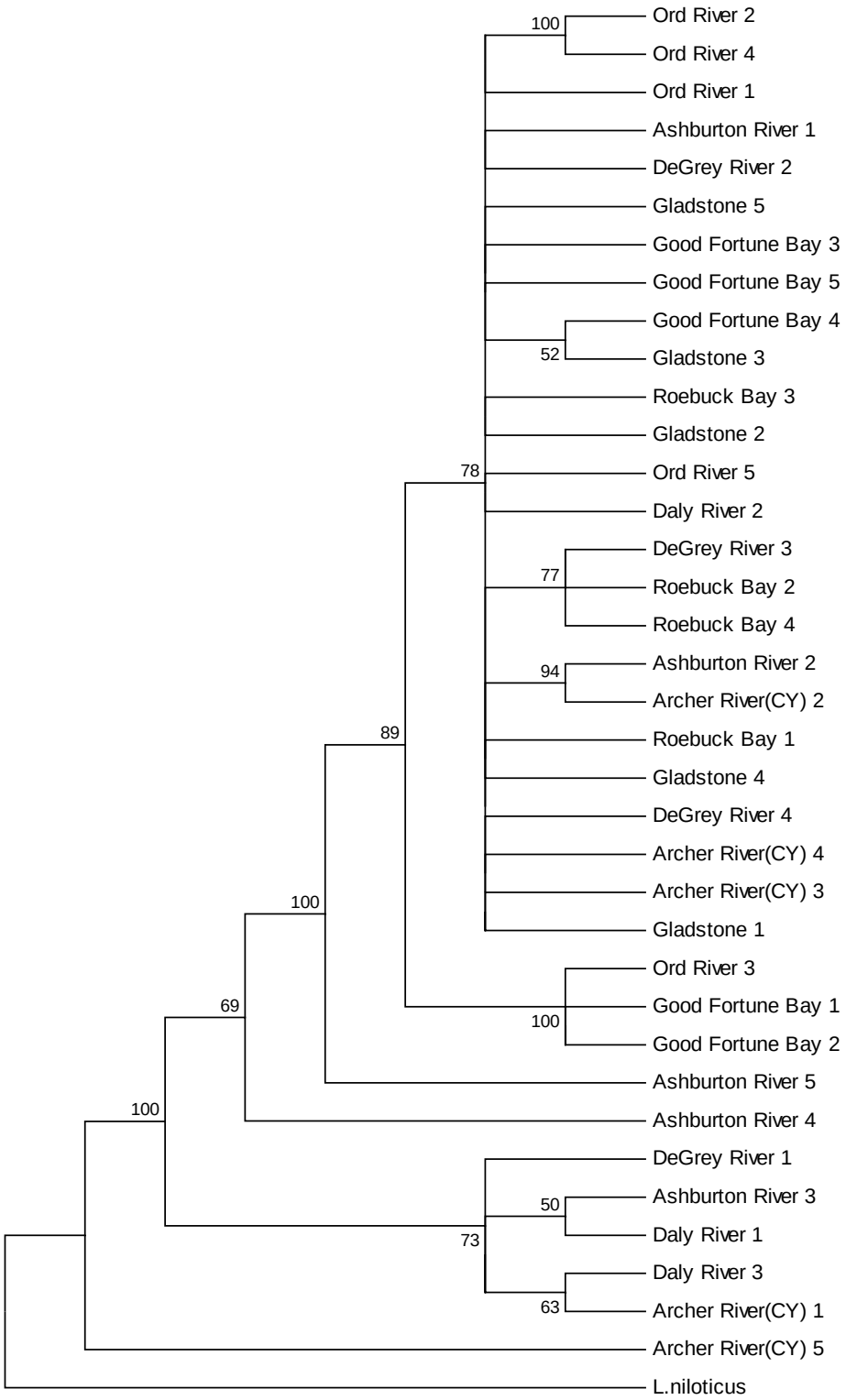


Figure 4.4. Consensus maximum parsimony tree of partial *ldh-b* sequences (coding and no-coding regions). Tree is rooted to *L. niloticus ldh-b* sequence (Chapter 3).

4.5: DISCUSSION

Variation in the coding nucleotide sequence of the *ldh-b* locus has been linked to differences in swimming performance and natural selection in the temperate fish *F. heteroclitus* (reviewed by Powers *et al.* 1998; (Crawford et al., 1999b); however, previous investigations into *ldh-b* have not assessed non-coding regions for potential functionality. This chapter presents the first among and within population genetic screening of *ldh-b* sequences in individuals representing eight populations which span the geographically, genetically and thermally distinct range of the tropical estuarine perciform *L. calcarifer*. The minimal coding region differences present across screened individuals and/or populations leads to no amino acid differences, which was expected based on the finding of previous allozyme screenings for variation across a similar range of *L. calcarifer* populations (Keenan, 1994; Shaklee and Salini, 1985). Moreover, this lack of variation in the coding and amino acid sequences across *L. calcarifer* populations is reflective of the minimal variation seen between *L. calcarifer* and its African congeneric *L. niloticus* (Chapter 3). Interestingly, a recent DNA bar coding experiment by Ward et al. (2008) called for a new species of Asia sea bass, which raises additional questions regarding the conservation of *ldh-b* and/or LDH-B. Specifically, high phylogenetic differentiation was established between Australia/Singapore (n = 10) and Myanmar/French Polynesia (n = 6) *L. calcarifer* via neighbor-joining analyses of the cytochrome oxidase I (*cox I*) and cytochrome b (*cyt b*) mitochondrial sequences (n = 21; 9.5% and 11.3% respectively), with Tanzanian *L. niloticus* (n = 5) as the out-group, (Ward et al., 2008). The level of variation in *ldh-b* and/or LDH-B between Australia/Singapore and Myanmar/French Polynesia sea bass is unknown. Therefore, comparing *ldh-b* and/or LDH-B sequences of Australian and Myanmar/French Polynesia sea bass would provide valuable insight regarding the response of *ldh-b* (e.g. up-regulation of transcripts) or LDH-B (e.g. enhanced Hb-O_2 binding efficiency due to

a unique allozyme) to extreme equatorial climate conditions (i.e. higher temperature and lower oxygen availability).

The high level of conservation observed for non-coding (intron) regions across discrete *L. calcarifer* populations, as was also observed between *Lates spp*. congenerics (Chapter 3), is potentially indicative of selective constraint on these regions. More specifically, the abundance of simple sequence repeats (SSRs) and microRNA elements (miRNAs) identified within *ldh-b* introns, in addition to their inherent low levels of variation, suggests that selection itself may be the driving force behind intron homology being greater than exon homology both among *Lates spp*. (Chapter 3) and between *L. calcarifer* populations (Hare and Palumbi, 2003; Lambowitz and Belfort, 1993; Mattick, 1994). The described impact that these identified SSR motifs and miRNA elements have on gene expression further highlights the potential for selective constraint to be maintaining their functional integrity, regardless of their location within introns (Ambros, 2004; Baek et al., 2008; Bartel, 2004; Brennecke et al., 2005; Enright et al., 2003; Hobert, 2004; Lai, 2003; Li et al., 2004; Meloni et al., 1998; Nakaya et al., 2007; Reisman et al., 1988; Wittekindt et al., 2000; Yue et al., 2001). As pointed out in Chapter 3, longer introns (> 87 bp) have been found to exhibit less divergence than shorter introns (< 87 bp) because of either an increased likelihood of embedded functional motifs being present in longer introns or the potential impact any mutations may have on the secondary structure of precursor messenger RNA (pre-mRNA) (Haddrill et al. 2005). This trend is upheld in regard to the identified *ldh-b* miRNAs, in that their abundance was higher in longer introns across all screened *L. calcarifer* populations. Moreover, non-coding intron sequences of *ldh-b* across all discrete *L. calcarifer* populations were more conserved than the *ldh-b* coding regions, despite comprising approx. 80% of the entire gene sequence. The identification of these SSR motifs and/or miRNA elements warrants further investigation regarding their functional

importance in the regulation and/or expression of the *ldh-b* locus as a constituent of the transcriptome and the proteome.

The *ldh-b* locus was found to be highly conserved across all screened *L. calcarifer* populations with only two SNPs occurring in the coding region, both of which were silent. Despite this lack of difference in the amino acid sequence, variation between thermally discrete *L. calcarifer* populations could still reside in transcript (mRNA) stability (Segal and Crawford, 1994b), transcript abundance (i.e. expression) (Crawford and Powers, 1992; Oleksiak et al., 2002; Schulte et al., 2000; Schulte et al., 1997; Segal et al., 1996) or in a population-specific efficiency in ribosomal translation (i.e. *ldh-b* : LDH-B not 1:1) (Latchman, 2002; Powers and Schulte, 1998). Therefore, further research is needed to assess this metabolically linked locus on such levels in an effort to establish if discrete tropical *L. calcarifer* populations are adapted to their respective habitats. The next chapter investigates the variation in *ldh-b* expression (i.e. transcript abundance) between discrete barramundi populations following long-term acclimation to cold-stress (20°C), heat-stress (35°C) and a control temperature (25°C and 30°C) for the southern and northern populations, respectively. It also examines the response of *ldh-b* to swimming (aerobic) challenge under all three treatment types (Chapter 5).

Interestingly, pairwise F_{ST} comparisons of *ldh b* sequences from each sampled location revealed the southernmost population (Gladstone, S. QLD) to be the only one which differed significantly from all other locations, whilst the Ord River sample differed significantly from six of the seven sampled sites. Further investigation into why the southernmost east coast population (Gladstone, QLD) and the northern Ord River, NT populations (see Methods) are most different from all other samples in regard to *ldh-b* sequences is warranted. Interestingly, these populations also have the lowest *ldh-b* genetic diversity indicating that they may have undergone either a selective sweep or a

bottleneck, but more data is required to evaluate this further. To achieve this, additional samples are required for screening of the established coding and non-coding SNPs within the *ldh-b* locus. Experimental acclimation to thermal stress (20 and 35°C) conditions and subsequent swimming (aerobic) challenges on fish from the southernmost (Gladstone, QLD: 23°S, 151°E) and northern (Ord River, NT: 17°S, 127°E) would provide key insight regarding whether or not greater selective constraint is acting upon the *ldh-b* locus of these genetically unique populations. The southernmost (Gladstone, QLD) population was originally planned to be the study population for the comparative transcriptomic analyses (Chapter 5), but this was prevented due to a Noda-virus outbreak at the Gladstone Waterboard Facility just prior to juveniles being sourced for experimental manipulation (see Chapter 2). Future studies should aim to investigate the swimming performance of Gladstone fish following cold-stress acclimation. Should such phenomic investigations find the U_{crit} of Gladstone fish to be significantly faster than that established for northern (Darwin, NT) and southern (Bowen, QLD) populations (Chapter 2), a transcriptomic investigation, such as that conducted in Chapter 5, should follow. The establishment of phenomic and transcriptomic difference unique to those observed for the northern and southern populations considered in this thesis (Chapters 2 and 5) would serve as key evidence toward the local adaptation of discrete Australian *L. calcarifer* populations to their respective thermal habitats.

4.6: CONCLUSIONS

The *ldh-b* locus appears to be highly conserved among discrete Australian *L. calcarifer* populations, as was expected based on previous population genetics studies which utilized neutral markers (i.e. allozymes and microsatellites). However, two coding region and five non-coding region SNPs are identified within *L. calcarifer ldh-b*, with

the non-coding SNPs occurring at locations embedded within miRNA elements. Single nucleotide polymorphisms such as these have been implicated in increasing or decreasing susceptibility to various human diseases (e.g. type 2 diabetes, Crohn's disease, obesity, schizophrenia, etc) by way of recent genome-wide association studies (Duerr et al., 2006; Hirschhorn and Daly, 2005; Scott et al., 2007; Sladek et al., 2007; Wang et al., 2005). Therefore, the single SNPs which impact several microRNA elements identified herein may be of relevance in the context of their potential impact on other loci in the metabolic cascade and/or on overall transcriptome functionality by way of altering *ldh-b* gene expression itself or altering the expression of proximal and/or distal loci (Ambros, 2004; Baek et al., 2008; Bartel, 2004; Enright et al., 2003; Hobert, 2004; Lai, 2003; Reisman et al., 1988; Selbach et al., 2008). In light of these differences in both coding and non-coding regions of *ldh-b,* among and within *L. calcarifer* populations, there is a need to extend the screening of *ldh-b* for such variants arising to determine if there are population specific variants that could be advantageous in different environments (e.g. temperature). Such miRNA variants are known to modify mRNA folding and/or SNPs are known to alter the binding efficiency or target-site recognition of miRNA elements (Ambros, 2004; Baek et al., 2008; Bartel, 2004; Brennecke et al., 2005; Enright et al., 2003; Lai, 2003; Li et al., 2004; Meloni et al., 1998; Yue et al., 2001). Both of these are known to ultimately alter the transcriptome and/or proteome.

Lastly, the partial characterization of the *ldh-b* locus in individuals representing discrete *L. calcarifer* populations, as well as across established genetic clades, presented in this chapter allows for a wide range of future studies on intra- and/or inter-species variation in the expression profile of *ldh-b,* an important metabolic enzyme (LDH-B). The subject of the next chapter (Chapter 5) is the examination of the differential response in the expression of this gene, to various environmental stressors (e.g.

temperature or swimming) in acclimated *L. calcarifer* sourced from geographically, genetically and thermally distinct locations. Lastly, the partial characterization of the *ldh-b* locus in individuals representing discrete *L. calcarifer* populations, as well as across established genetic clades, presented in this chapter allows for a wide range of future studies on intra- and/or inter-species variation in the expression profile of *ldh-b,* an important metabolic enzyme (LDH-B). The next chapter investigates variation in *ldh-b* transcript abundance following long-term (28 day) acclimation to cold-stress (20°C), heat-stress (35°C) and a control temperature (25°C and 30°C) conditions for juveniles from a southern (Bowen, QLD: 20°S, 148°E) and northern (Darwin, NT: 12°S, 130°E) population, respectively. Chapter 5 also examines the response of hepatic *ldh-b* transcript abundance to swimming challenge both in the presence and absence of thermal stress.

Chapter 5

Different transcriptional responses to separate and combined stresses in discrete populations of Australian barramundi, *Lates calcarifer*: the "double whammy" effect

5.1: ABSTRACT

Environmentally induced thermal stress impacts both the phenome and transcriptome of fish, affecting overall growth, performance and metabolism as well as altering transcription levels of associated genes. The candidate gene lactate dehydrogenase-B (*ldh-b*) was targeted due to its essential role in maintaining aerobic metabolism by catalyzing the conversion of lactate to pyruvate via oxidation. We demonstrate that juvenile *Lates calcarifer* from two distinct Australian populations vary in their magnitude of *ldh-b* transcript abundance following 28 day acclimations to ecologically relevant cold-stress (20°C) and heat-stress (35°C) conditions. *Ldh-b* transcript abundance was also quantified following exposure to swimming (aerobic) challenge both in the presence and absence of these thermal stresses. Fish from southern and northern populations exhibited a significant increase in hepatic *ldh-b* transcript abundance ($F_{5, 41} = 6.459$; $p < 0.001$ and $F_{5, 39} = 3.866$; $p = 0.006$ respectively) in response to swimming challenge at their respective native culturing temperatures (25°C and 30°C respectively). Fish from both populations exhibited a significant increase in hepatic *ldh-b* transcripts following 28 day acclimation to heat-stress (35°C) condition. However, only fish from the southern population possessed significantly more hepatic *ldh-b* transcripts following 28 day acclimation to cold-stress (20°C) conditions. This evidence suggests that southern populations of barramundi possesses a unique transcriptional response compared to a distinct northern population, and that the southern population may be evolutionarily accustomed to a greater variance in seasonal temperatures. Moreover, *ldh-b* transcriptional mechanisms (e.g. promoter) are likely to be under enhanced selective pressure should temperature increase over the next century, as predicted, due to climate-change.

5.2: INTRODUCTION

Environmental temperature is a pivotal factor in many biological processes. Poikilothermic organisms, such as fish, are dependent on environmental temperature for homeostatic maintenance and are highly susceptible to changes in their thermal environment. A range of biological activities in fish are known to be influenced by thermal stress, or extended exposure to temperatures above and/or below those normally experienced in native environments. For example, increases in mortality rate, as well as reduced growth and reproductive output have been demonstrated in hard corals subjected to extreme heat/cold conditions (Harriott and Banks, 2002; Jokiel and Coles, 1990). Similarly, cold- and heat-stress conditions have been found to impact the swimming performance of both temperate and tropical fishes (reviewed by Fulton, in press; Plaut, 2001).

Cold- and heat-stress conditions are known to invoke significant changes in the transcriptome of fish in ways that are linked to their ability to cope with thermal change (reviewed by Basu et al., 2002; Douglas, 2006). More specifically, the up-regulation of loci known to reduce mRNA secondary structure (*DDX21* and *CIRBP*), as well as high-mobility group loci known to assist in maintaining chromosome conformation (*HMG1*, *HMGT1*, *HMG4*, *NHPX*), were observed across various tissue types in the common carp *Cyprinus carpio* following 12 day exposure to 10°C compared to 30°C control (Gracey et al., 2004). Increased HMG1 expression was also observed to be highly negatively correlated with temperature in unchallenged annual killifish (*Austrofundulus limnaeus*) following exposure to ecologically relevant diurnal fluctuations (20 - 37°C) and was concluded to be a putative and ideal "global regulator of transcription in response to temperature" (Podrabsky and Somero, 2004). Loci involved in enhancing the capacity for ATP synthesis (ADP/ATP translocases, phosphate carrier protein and

mitochondrial electron transport chain), as well as lipid metabolism in the liver, were also found to be up-regulated in response to 12 day cold-stress (10°C) in the common carp (Gracey et al., 2004).

Conversely, the transcriptional response of heat shock proteins can also vary with extent of heat-stress exposure (reviewed by Basu et al., 2002). For example, northern and southern *F. heteroclitus* populations exhibit population-specific expression patterns for heat-shock proteins (*hsc70, hsp70-1*and *hsp70-2*) following heat-shock treatments of 33°C for 2 hours (Fangue et al., 2006). Moreover, differential responses across populations indicated a heightened susceptibility to heat-stress in cold water northern *F. heteroclitus* populations based on the significantly higher expression of *hsp70-2* and an increased tolerance to heat-stress in warm water southern *F. heteroclitus* populations (Fangue et al., 2006). Moreover, a locus involved in glucose metabolism (glycogen synthase) was up-regulated in the liver of *Austrofundulus limnaeus* following heat-stress alone (37°C) (Podrabsky and Somero, 2004). Other loci involved in the gluconeogenesis pathway were found to be up-regulated following short-term (4 hours) exposure to ramped (11 to 30°C) heat-stress conditions that mirror the experience of porcelain crab, *Petrolisthes cinctipes exposed* in summer low tides (Teranishi and Stillman, 2006).

In addition to these examples, other studies have reported on the transcriptomic response to varying lengths of thermal (cold- and/or heat)-stress conditions in an assortment of model temperate fish, including channel catfish (*Ictalurus punctatus*), zebrafish (*Danio rerio*), annual killifish (*Fundulus heteroclitus, F. grandis*) and rainbow trout (*Oncorhynchus mykiss*) (reviewed by Basu et al., 2002; Douglas, 2006). However, in some instances, transcriptomic variation is as high within (i.e. between individuals) as it is among fish populations (Oleksiak et al., 2002). In light of these identified phenomic and transcriptomic responses, we consider extended (28 day) exposure to

ecologically relevant cold (20°C) or heat-stress (35°C) conditions to be equally powerful and unique forms of thermal stress.

The candidate gene lactate dehydrogenase-B (*ldh-b*) was selected because it encodes the metabolic enzyme LDH-B, which plays an essential role in maintaining aerobic metabolism by catalyzing the conversion of lactate to pyruvate via oxidation (Powers and Schulte, 1998). Lactate is the major by-product of anaerobic glycolysis which results from burst swimming (Plaut, 2001). The conversion of lactate to glucose via gluconeogenesis is also catalyzed by LDH-B (Place and Powers, 1984a, 1984b). Removal of accumulating lactate from aerobic tissues (e.g. heart, skeletal muscle) via these LDH-B catalyzed conversions occurs in the liver and allows the sustained aerobic metabolic activity required for competition for mates, active predation or escape from predators (reviewed by Powers *et al.*, 1991; Powers and Schulte, 1998). The oxygen binding affinity of hemoglobin (Hb) is also impacted by LDH-B, which alters intra-erythrocyte ATP concentrations of Hb in fish and thus directly impacts delivery of Hb-bound oxygen to red muscle tissues (Powers, 1980; Powers *et al.*, 1979). Hb-O_2 interplay may be an alternate mechanism by which LDH-B affects sustainability of aerobic performance and/or processes (e.g. swimming). The significant fluctuation in transcript abundance of *ldh-b* in response to both thermal and swimming stress in the temperate killifish *Fundulus heteroclitus* (reviewed by Powers *et al.*, 1991; Powers and Schulte, 1998) makes *ldh-b* an ideal candidate gene for the stress response study undertaken here.

In this study we use the candidate gene lactate dehydrogenase-B (*ldh-b*) to quantify and compare transcriptomic responses in barramundi exposed to 28 day acclimation in cold-stress, heat-stress and native thermal environments. We examine this in two geographically, genetically and thermally discrete populations of the non-model tropical Australian barramundi, *Lates calcarifer*. Additionally, for both

barramundi populations, transcriptomic variation in *ldh-b* gene expression is quantified and compared for swimming (aerobic) stress in the presence and absence of thermal stress. This was done in order to determine if *ldh-b* expression is responsive to swimming stress alone and in combination with both thermal stresses for either or both barramundi populations.

Lates calcarifer is a catadromous, protoandrous hermaphrodite, native to rivers, estuaries and shallow marine environments throughout the northern half of Australia (25°S – 12°S) and the south-eastern Asian archipelago (13°N – 10°S). High genetic partitioning is present between Australian populations of *L. calcarifer* due to the non-interconnecting nature of the Australian river-ways and the non-migratory aspect of barramundi's larval stage (Russell and Garrett, 1985). Two major genetic clades have been established between the eastern (Queensland) and northern (Northern Territory) coasts (Chenoweth et al., 1998), as well as a unique third genetic clade encompassing the west coast (Western Australia) (Doupé et al., 1999). While *L. calcarifer* is often described as eurythermal, very little is known about how this species tolerates the broad range of temperatures (18 to 34°C) encountered across its natural geographic range (Katersky and Carter, 2005, 2007). To date, most studies on optimal growth temperatures have utilized only one genetic stock/cohort (Katersky and Carter, 2005, 2007) and those studies which did utilize different "stock" have been confined to the Eastern Australia clade (Keenan, 1994). Given that *L. calcarifer* is the basis of a large rapidly growing aquaculture industry in Australia and overseas (e.g. USA, Israel, Singapore, Thailand, India, the Netherlands, etc), there is the potential to select genetically distinct stocks for broodstock so as to match the thermal environments of geographically dispersed hatcheries (Rimmer, 1995).

The aim of this study was to test five hypotheses on the response of *ldh-b* gene expression in *L. calcarifer* from two genetically and thermally distinct populations, one

population from the Eastern Australian clade and one population from a Northern Australian clade (Chenoweth et al., 1998; Keenan, 1994). Specifically, the null hypotheses were that: 1) Aerobic challenge alone, under native culture temperatures (25°C and 30°C) impacts both *L. calcarifer* populations similarly; 2) Acclimation to the cold-stress (20°C) condition impacts southern (sub-tropical) fish differently to northern (tropical) fish; 3) Aerobic challenge combined with acclimation to cold-stress (20°C) conditions impacts southern fish differently to northern fish; 4) Acclimation to the heat-stress (35°C) condition impacts southern fish differently to northern fish and 5) Aerobic challenge combined with acclimation to the heat-stress (35°C) condition impacts southern fish differently to northern fish. These hypotheses were tested in juvenile *L. calcarifer* following a 28 day (long-term) acclimation by comparing the fold-difference in *ldh-b* transcript abundance relative to a non-stress control condition (25°C and 30°C for southern and northern respectively).

5.3: METHODS

Study populations and tissue sampling

Juvenile *L. calcarifer* were obtained from two geographically isolated hatcheries within Australia: Darwin Aquaculture Centre (n = 205; standard length = 60.47 ± 0.73mm) and Gladstone Area Water Board Fish Hatchery (n = 311; standard length = 61.93 ± 0.56mm) (Figure 1). These aquaculture facilities are located in Darwin, Northern Territory (12° 27.7’ S, 130° 50.5’ E) and Bowen, central Queensland (20° 0.7’ S, 148° 14.8’ E) and are hereafter referred to as northern and southern, respectively. Fish attained from these two distinct aquaculture facilities were first generation offspring from locally sourced broodstock. First generation *L. calcarifer* were sourced from two genetically and thermally distinct aquaculture reared populations rather than natural

stocks because: 1) Their availability in large number and of similar sizes; 2) Their lower susceptibility to the impacts of cold- and/or heat-stress and thus they are more likely to survival under thermal extremes than adult fish (Meeuwig et al., 2004); 3) The adults are too large for aerobic challenges in our purpose-built swimming chamber, as described in (Fulton, 2007); and 4) Juveniles do not require an energy expenditure balance between growth and reproduction. Juvenile barramundi were acclimated for 28 days to the culturing temperature used by the southern or northern aquaculture facility from which they were obtained, being 25ºC and 30°C, respectively, to remove environmental-induced differences that may have existed in the transcriptome prior to swimming challenge and tissue sampling.

Fish from northern and southern *L. calcarifer* populations were acclimated to 20, 25, 30 and 35°C for 28 days and subsequently swimming challenged by way of an incremental increase in water velocity (cm s^{-1}) in a purpose built swimming chamber until the point of exhaustion (Chapter 2). Liver tissue samples were collected from fish both unchallenged acclimation and aerobic challenge following acclimation (n = 16 for all four treatment temperatures for both populations). Liver samples were immediately snap-frozen in liquid nitrogen (LNO_2). Following isolation of total RNA from all liver samples, variation in *ldh-b* expression was subsequently assessed by an optimized quantitative real-time PCR (qRT-PCR) assay.

Isolation of total RNA

Total RNA was extracted with UltraSpec RNA Isolation reagent (Fisher Biotec Pty Ltd, Sublaco Western Australia) by homogenizing whole liver samples with 1ml of reagent per 0.1g of tissue (as per manufacturer's instructions). Isolated total RNA was cleaned once by precipitation with an equal volume (1mL) of 100% isopropanol, to remove trace amounts of remaining UltraSpec reagent, and once by precipitation with [2.5M]

lithium chloride (Amresco, Ohio USA), to remove all carbohydrates, lipids and proteins (including LDH-B). Total RNA was then treated with Turbo DNAse Free (Ambion, Austin, TX, USA) to remove any remaining trace DNA contamination. DNA-free total RNA was precipitated overnight in equal parts [0.17M] ammonium acetate and absolute ethanol [50%] to assure no carryover of DNAse Inactivation agent from the Turbo-DNAse Free Kit. Purified total RNA was quantified (ng/µl) and purity verified (260:280 and 260:230nm absorbance ratios) using a NanoDrop spectrophotometer (Invitrogen Australia Pty. Ltd., Mount Waverley Victoria, Australia). Quantity and purity values ranged from 400-1,000ng/µL and 1.9-2.1 (260:230nm), respectively.

Synthesis of complementary DNA (cDNA) via reverse transcription (RT)

First-strand complimentary DNA (cDNA) was synthesized with Super-Script III reverse transcriptase (SybrGreenER 2-Step qRT-PCR Kit; Invitrogen Australia Pty. Ltd., Mount Waverley Victoria, Australia). The supplied reverse transcriptase master mix (Step 1) contained both random hexamers and Oligo d_T 20 for complementary-strand synthesis and was used according to manufacturer's instructions. This method of reverse transcription was chosen over primer-specific cDNA synthesis due to the added advantage of permitting future screenings of the "global" cDNA library to quantify any mRNAs known to express in fish liver. All cDNA synthesis reactions utilized a normalized amount (1µg) of total RNA, as recommended by (Bustin et al., 2005; Bustin and Nolan, 2004). Following reverse transcription the synthesized cDNA was cleaned via Microspin S-400 HR columns (GE Healthcare Bio-Sciences Pty. Ltd. New South Wales, Australia) to remove any remaining RT reagents, as these were found to decrease qRT-PCR efficiency.

Quantitative real-time PCR (qRT-PCR) primer design

Primers designed for qRT-PCR assays conformed to the following stringencies: 1) amplicon length was between 50-150 bp; 2) binding sites spanned exon/exon boundaries; 3) synthesized at a final concentration of 200nm; 4) a GC content of 30 - 80 %; 5) T_m between 58°C and 60°C; 6) last 5 bp at 3'end contained no more than 2 G's or C's; 7) minimal chance of dimerization with one another (kcal mol^{-1}). In addition to these criteria the physical location of qRT-PCR primers was skewed toward the 3'end of the *ldh-b* locus coding sequence in an effort to prevent a potential loss of binding efficiency due to an incomplete representation of the 5'end following reverse transcription with a combination of Oligo-dT_{20} and random hexamers (Bustin and Nolan, 2004). Following optimization of three different primer-pairs, the third set (F3: CTGGAGTTAACCTGCAGAAG and R3: CAGCTTGATCACCTCGTAG) yielded the best PCR efficiency (105%) and were therefore utilized in the chemistry of all subsequent qRT-PCR assays.

Quantitative real-time PCR (qRT-PCR) quality control

PCR efficiency was determined by generation of standard curves from a five-fold Log_{10} dilution series (neat, 10^0, 10^{-1}, 10^{-2}, 10^{-3} and 10^{-4}). Triplicate reactions were run for all dilution factors to evaluate consistency between amplifications. A background calibration dye (ROX) was added to all master-mixes, as per manufacturer's instructions (SuperScript-III 2-Step qRT-PCR kit; Invitrogen). The coefficient of variation (% CV) across triplicate reactions for all dilution factors in the serial dilution was less than 5% indicating low technical variation. All assayed cDNA samples were diluted to 10^{-2} before qRT-PCR assessment, as this dilution factor exhibited the lowest coefficient of variation across optimization runs. The coefficients of variation (i.e. CV) across triplicate reactions for all cDNA samples were less than 1%, indicating low technical

variation. A calibrator sample (unchallenged southern individual) was also included in triplicate within each qRT-PCR run to establish the level of variation inherent from run-to-run as well as between replicated RT reactions. These sources of technical variation produced CV's of less than 4%, further indicating a low level of technical variance within the accepted margin of statistical error ($p < 0.05$).

Quantitative real-time PCR (qRT-PCR) assays

Following first-strand synthesis, a standardized volume and quantity of cDNA, 2µL and 1µg respectively, were utilized as template in all qRT-PCR reactions. All reactions utilized primers designed off *ldh-b* coding region sequence (GenBank Accession number **FJ439507)**, consisting of one forward (CTGGAGTTAACCTGCAGAAG) and one reverse (CAGCTTGATCACCTCGTAG) primer (Chapter 3). All qPCR reactions (15µl) utilized the Invitrogen "Two-Step qRT-PCR Kit Universal". Reactions contained 200nM each *ldh-b* gene-specific primer and 1µg of cDNA template added to 1X SYBR GreenER qPCR SuperMix Universal with supplied 500nM ROX reference dye (Invitrogen). Thermal cycling, SYBR GreenER fluorescence detection and melt curve analysis were conducted on a MJR DNA Engine with Chromo4 detector and Opticon detection software (Bio-Rad Laboratories Pty., Ltd., Gladesville, New South Wales). Optimal thermal cycling conditions were as follows: 50°C for 2min (UDG incubation) and 95°C for 10min ("hot-start" *Taq* activation) followed by 40 cycles of 95°C denaturation for 15sec and 60°C annealing/elongation for 30sec (modified from manufacturer's instructions). The threshold cycle (C_t) is established internally by Opticon software (Bio-Rad Laboratories) as the cycle at which SYBR GreenER fluorescence crossed the user-defined threshold fluorescence (0.1). The increase in C_t cycle number with decrease in qRT-PCR reaction template per Log_{10} dilution factor resulted in a standard curve with a slope of - 3.21, translating to PCR efficiencies ($E = 10^{1/slope}$) of 105%, which is within the recommended window of qRT-PCR efficiency

(95 – 105%). Lastly, melt curve analysis was performed to confirm the presence of a single target product (113 bp), as well as the absence of primer dimer, by progressing incrementally (in steps of 0.2°C) across the range 70 – 90°C. The presence of a single melt curve indicated that there was no primer-dimer formation or non-specific PCR amplification.

Establishment of fold-differences in *ldh-b* transcript abundance

Quantification of *ldh-b* expression between samples within each population was established by differences in the threshold cycle (C_t) relative to that sample's respective control C_t. In this way, *ldh-b* transcript abundance was normalized to total RNA as recommended (Bustin, 2000; Bustin *et al.*, 2005; Bustin and Nolan, 2004). The lack of established housekeeping genes for both the experimental stress (temperature) and the study species (*Lates calcarifer*) dictated our use of this internal normalization method.

Unchallenged fish from both the northern and southern control (native) treatment temperatures (30°C and 25°C, respectively) were used as the baseline against which all C_t comparisons were made in establishing the fold-differences in *ldh-b* expression. All reported data refer to fold-differences in *ldh-b* expression as compared to these standards. PCR efficiency of 105% was used in conjunction with the number of cycles differing between the C_t of a treatment sample and that of the population-specific base-line C_t to establish the fold difference in abundance of starting template. Calculations are based on the equation:

$$\Delta = E^{(\text{control Ct} - \text{sample Ct})}$$

where Δ is the fold-change in expression level, E is the qRT-PCR efficiency, control C_t is the average critical cycle of the population-specific control samples and sample C_t is the average critical cycle across triplicate reactions for individual treatment

samples. All fold differences in *ldh-b* expression are given as the mean ± s.e. for $n = 8$ individuals per treatment group.

Statistical analysis

All data are presented as (mean ± S.E.). Data points were assessed for homogeneity of variance by way of Levine's test ($p > 0.05$) and SPSS (version 16.0) generated Q-Q plots. Following the justification of inbuilt assumptions, the Type III sum of squares based one-way analysis of variance (ANOVA) test was conducted with temperature treatment and aerobic state (i.e. non-swum or swum) as fixed factors. Subsequent Tukey's HSD post-hoc analysis was used to establish the levels of significance in *ldh-b* gene expression between temperature treatments within each *L. calcarifer* population (Table 5.1). Direct comparisons of *ldh-b* expression levels between populations were not possible for individual treatments as rearing and swimming trials were unable to be conducted simultaneously due to space limitations and availability of juvenile fish. Subsequent regression analyses (Figure 2) were performed to establish the relationship between fold-difference in hepatic *ldh-b* transcript abundance and the relative critical swimming performance (U_{crit} in BL s^{-1}) of juvenile *L. calcarifer* exposed to ecologically relevant stressors (Chapter 2). Data points from southern and northern fish more than two standard deviations from the mean of the 20°C and 35°C treatments ($n = 3$ and 1, 1 and 3 respectively) as well as the population specific control treatments (25°C and 30°C; $n = 2$ and 3 respectively) were considered outliers and removed from regression analyses. All statistical analyses were performed with SPSS 16.0.

5.4: RESULTS

Long term (28 day) acclimation to the cold-stress (20°C) and heat-stress (35°C) conditions had a pronounced effect on the mortality rate (survivorship) of *L. calcarifer* juveniles from both experimental populations examined here (Chapter 2). Specifically, a significantly lower than expected mortality rate was observed for the southern (1.4%), but not the northern (5.9%) population under cold-stress (20°C) conditions; whereas both populations exhibited a higher than expected mortality rate (53.2% and 17.6% respectively) under heat-stress (35°C) conditions (Figure 2). Under native temperature (control) conditions (25°C and 30°C), mortality rates were not significantly lower than expected for the southern (1.6%) and northern (9.8%) populations, respectively (Figure 2).

Population-specific baseline (unchallenged control) expression level

Population specific baseline *ldh-b* expression levels were based on culturing temperatures used by *L. calcarifer* aquaculture facilities located in the southern and northern geographic ranges (25ºC and 30°C, respectively). The average *ldh-b* expression observed in juveniles acclimated for 28 days to their respective native culture temperature and not exposed to an aerobic swimming (unchallenged) were used as the baseline level of expression for comparisons within each of the southern and northern populations (n = 8; 1.0 ± 0.37 and 1.0 ± 0.29, respectively; Figure 2). All subsequent temperature treatments involved an identical 28 day period of acclimation and method of tissue collection, with or without an aerobic swimming challenge.

3.2. Fold-differences in hepatic *ldh-b* transcript abundance following aerobic (swimming) challenge at native (control) temperatures

Southern *L. calcarifer* subjected to aerobic challenge at a native temperature (25°C) exhibited a significant, 3.69 ± 0.37 - fold increase in hepatic *ldh-b* transcript abundance ($p < 0.001$) compared to their baseline, unchallenged expression (1.0 ± 0.37; Table 1; Figure 2a). Likewise, northern fish subjected to aerobic challenged under their native culture temperature (30°C) exhibited a significant, 2.50 ± 0.31 - fold increase in hepatic *ldh-b* transcript abundance ($p = 0.025$) compared to baseline expression (1.0 ± 0.29; Table 1, Figure 2b).

Table 5.1 Analyses of variance in *ldh-b* expression within populations

Population	Partial Eta Squared	*F*	df	*p*
Southern	0.441	6.459	5, 41	< 0.001
Northern	0.331	3.866	5, 39	0.006

Fold-differences in hepatic *ldh-b* transcript abundance in response to 28 day cold-stress (20°C) without and with concurrent swimming (aerobic) challenge

Cold-stressed (not swimming challenged) southern *L. calcarifer* exhibited a significant, 2.87 ± 0.40 - fold increase in hepatic *ldh-b* transcript abundance ($p = 0.03$) compared to baseline expression (Table 5.1, Figure 5.2a). However, when southern fish were subjected to both cold-stress and swimming challenge (double stress) there was no significant change in hepatic *ldh-b* transcript abundance ($p = 0.19$) compared to baseline expression (Table 5.1, Figure 5.2a). Moreover, southern fish from the double stress treatment exhibited a level of hepatic *ldh-b* transcript abundance no different from their respective unchallenged (non-swum) counterparts (0.26–fold decrease, $p = 0.96$; Table 5.1, Figure 5.2a).

Cold-stressed (not swimming challenged) northern *L. calcarifer* had similar levels of hepatic *ldh-b* transcript abundance to that of the native culture temperature (unchallenged baseline) ($p = 0.78$) (Table 5.1; Figure 5.2b). Similarly, when northern

fish were subjected to both swimming challenge and cold-stress (double stress) they exhibited no significant difference in hepatic *ldh-b* transcript abundance (1.78 ± 0.29 fold increase; $p = 0.58$) from baseline expression (Table 5.1; Figure 5.2b). Moreover, as was the case for the southern population, northern fish from the concurrent swimming challenged cold-stress treatment exhibited a level of hepatic *ldh-b* transcript abundance no different from their respective unchallenged cold-stress counterparts ($p = 1$) (Table 5.1; Figure 5.2b).

Fold-differences in hepatic *ldh-b* transcript abundance in response to 28 day heat-stress (35°C) without and with concurrent swimming (aerobic) challenge

Heat-stressed, but not swimming challenged southern *L. calcarifer* exhibited a significant, 3.00 ± 0.37 fold increase in hepatic *ldh-b* transcript abundance ($p = 0.01$) compared to baseline expression (Table 5.1; Figure 5.2a). Southern fish from the concurrent swimming challenged and heat-stressed treatment paralleled the response observed for southern fish from the concurrently swimming challenged cold-stress (20°C) treatment in that there was no significant change in hepatic *ldh-b* transcript abundance ($p = 0.97$) compared to baseline expression (Table 5.1; Figure 5.2a). However, southern fish from the swimming challenged heat-stress treatment exhibited a biologically significant reduction ($p = 0.09$) in hepatic *ldh-b* transcript abundance compared to the unchallenged but heat-stressed counterparts (Table 5.1, Figure 5.2b).

Heat-stressed, but not swimming challenged northern *L. calcarifer* exhibited a significant, 2.62 ± 0.31 fold increase in hepatic *ldh-b* transcript abundance ($p = 0.01$) compared to baseline expression (Table 5.1, Figure 5.2b). Northern fish from the concurrent swimming challenged heat-stress treatment paralleled the response observed for northern fish from the challenged cold-stress (20°C) treatment in that there was no significant change in hepatic *ldh-b* transcript abundance ($p = 0.97$) compared to baseline expression (Table 5.1, Figure 5.2b). However, as for the southern population, northern

fish from this double-stress treatment (i.e. double whammy effect) exhibited a biologically significant reduction (p = 0.10) in hepatic *ldh-b* transcript abundance compared to their unchallenged heat-stress treatment counterparts (Table 5.1, Figure 5.2b).

Autocorrelation between hepatic *ldh-b* transcript abundance and critical swimming performance following 28 day acclimation

Regression analyses were used to establish the presence of underlying correlations between fold-differences in *ldh-b* transcript abundance and relative critical swimming performance (U_{crit} in BL s^{-1}) (Figure 5.3). Fold-differences in hepatic *ldh-b* transcript abundance observed in swimming challenged cold-stress northern and southern fish are negatively correlated with reported U_{crit} swimming speeds (n = 7 and 5 respectively; Figure 5.3a) (Chapter 2). Fold-differences in hepatic *ldh-b* transcript abundance observed in swimming challenged heat-stressed northern fish are also negatively correlated with these reported U_{crit} swimming speeds (n = 6), whereas no such correlation is observed for swimming challenged heat-stressed southern fish (n = 7; Figure 5.3b) (see Chapter 2). Interestingly, fold-differences in hepatic *ldh-b* transcript abundance observed for northern and southern fish from the swimming challenged native temperature treatments (30°C and 25°C, respectively) are not correlated with their reported U_{crit} swimming speeds (n = 5 and 6 respectively; Figure 5.3c**)** (Chapter 2). These relationships between fold-differences in *ldh-b* transcript abundance and relative U_{crit} swimming speed appear to be interconnected with the complex response of metabolic pathways to cold-stress and heat-stress conditions, as discussed below.

FIGURES

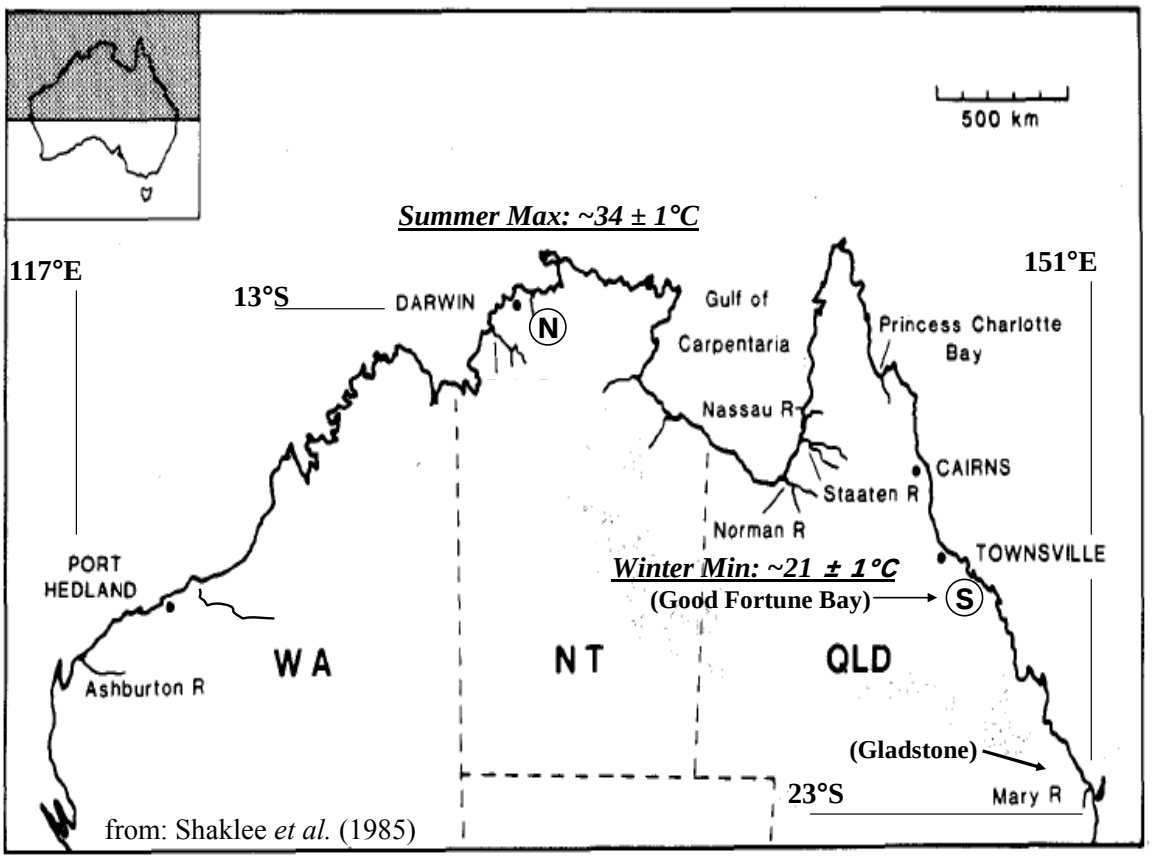


Figure 5.1. Map indicating the source location of thermally discrete and experimentally manipulated (Chapter 2) northern and southern *L. calcarifer* populations, represented by N and S respectively (adapted with permission from Shaklee *et al*. 1985). Summer max and winter min temperatures based on sea-surface temperatures recorded on 31 January and 31 July, 2006 respectively.

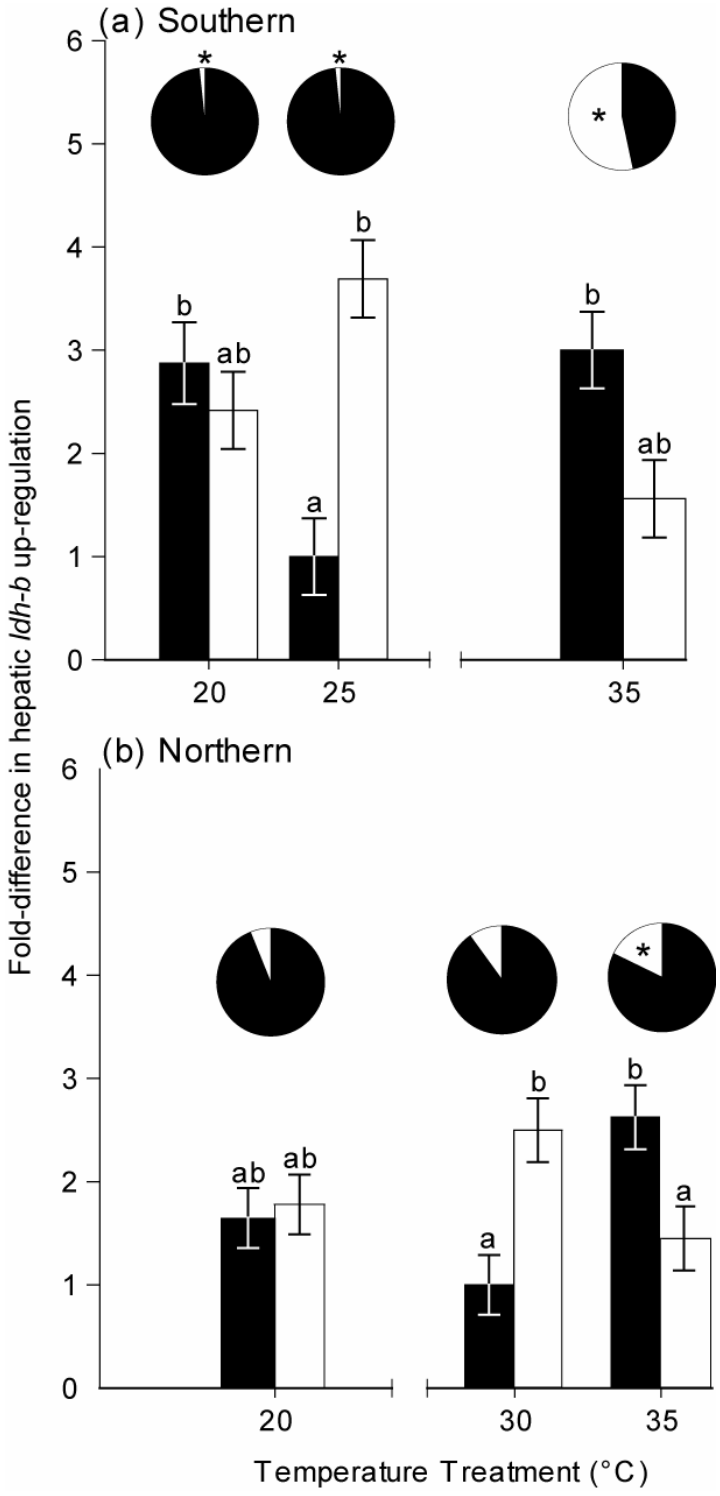
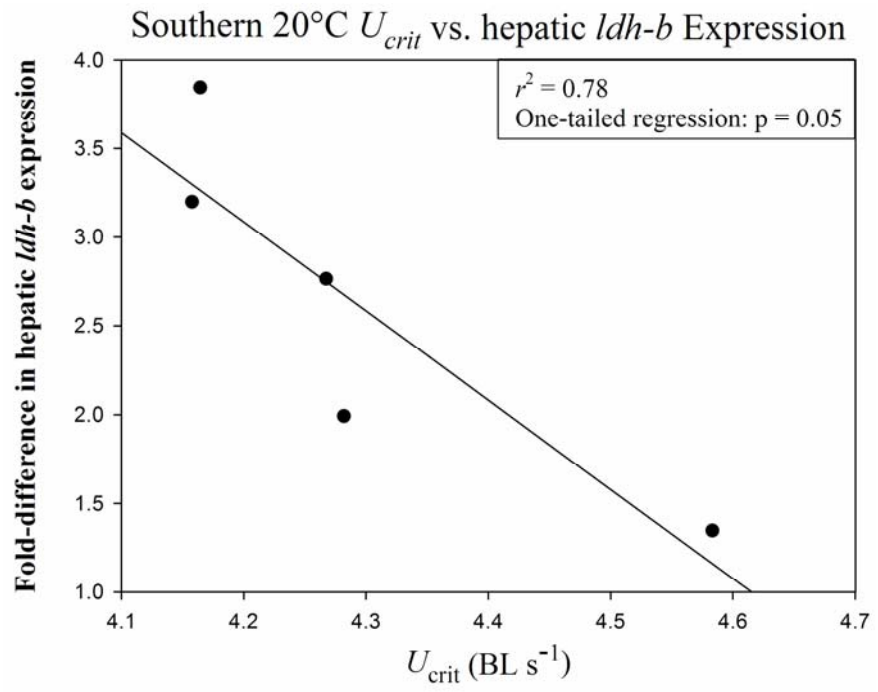


Figure 5.2. Fold-difference in hepatic *ldh-b* template abundance as detected by qRT-PCR assays. Inset pie charts refer to previously observed mortality rates under each experimental temperature treatment (Chapter 2). Data presented as mean ± S.E. Histogram: panels (a) and (b) refer to the southern and northern populations respectively, while black and grey fills represent unchallenged and challenged treatments, respectively. Inset pie charts: black and grey proportions represent % survival and % mortality, respectively. Significant homogenous groupings ($p < 0.05$) of intra-population expression levels as compared to their respective baseline are provided by letter groupings (a, b, ab), whereas asterisks indicate when mortality (%) differed significantly from expectation (X^2, $p < 0.05$) (Chapter 2).



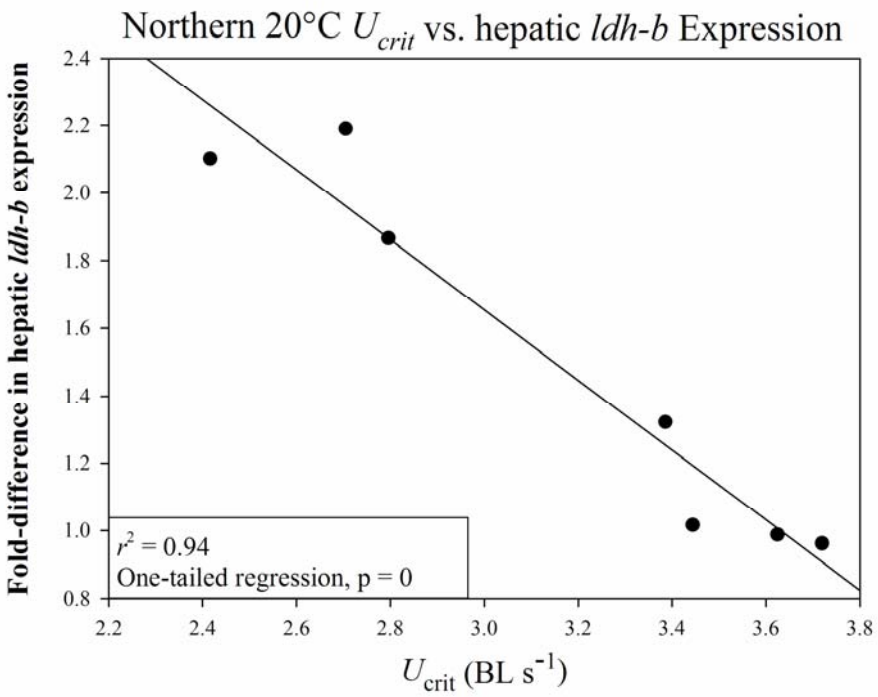


Figure 5.3a. Simple regression analysis comparing *ldh-b* template abundance (fold-difference) vs. relative U_{crit} swimming speed (BL s^{-1}) for cold-stress northern and southern *L. calcarifer* juveniles. Outliers more than two standard deviations from the mean were removed (see Methods).

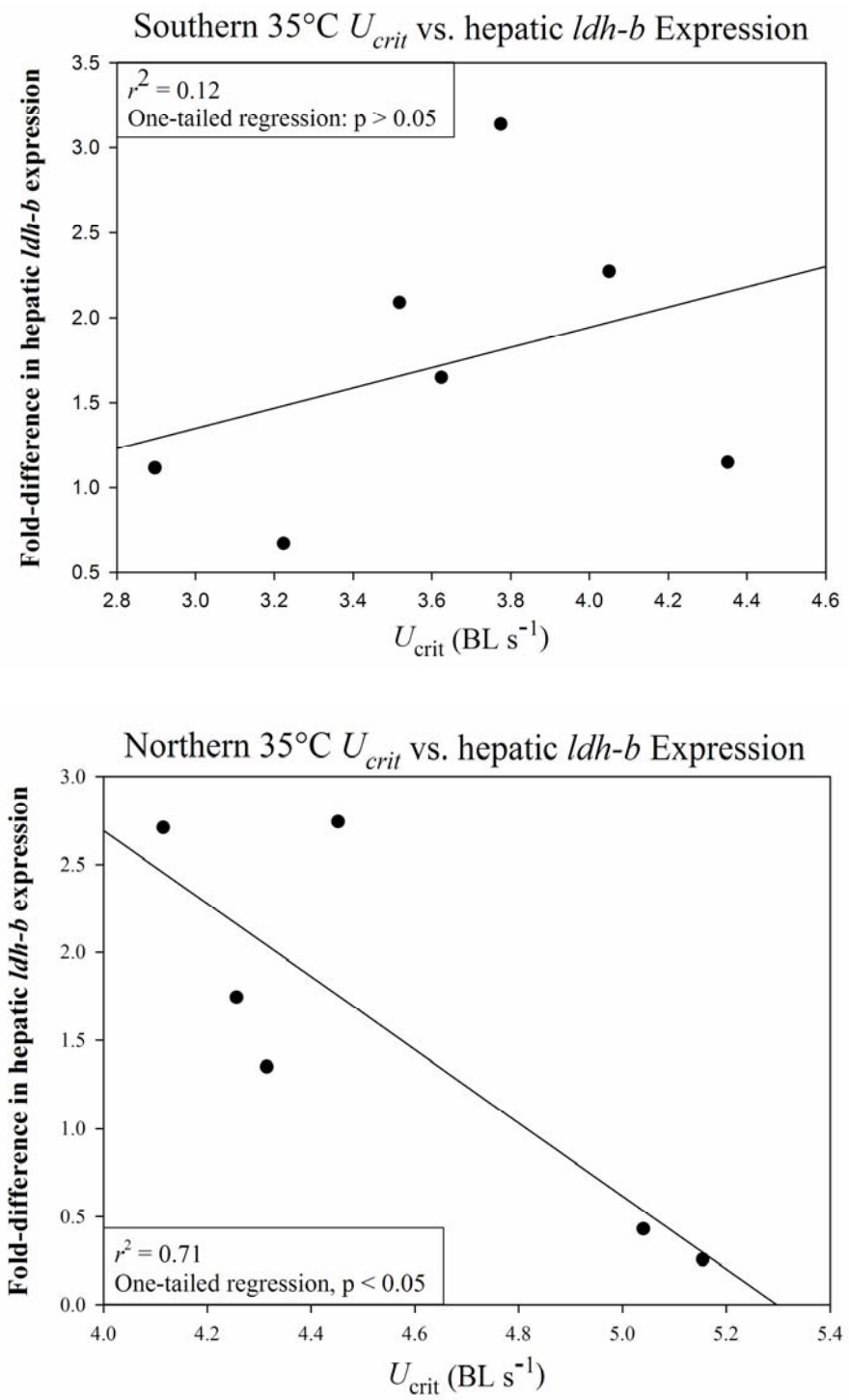


Figure 5.3b. Simple regression analysis comparing *ldh-b* template abundance (fold-difference) vs. relative U_{crit} swimming speed (BL s^{-1}) for heat-stress northern and southern *L. calcarifer* juveniles. Outliers more than two standard deviations from the mean were removed (see Methods).

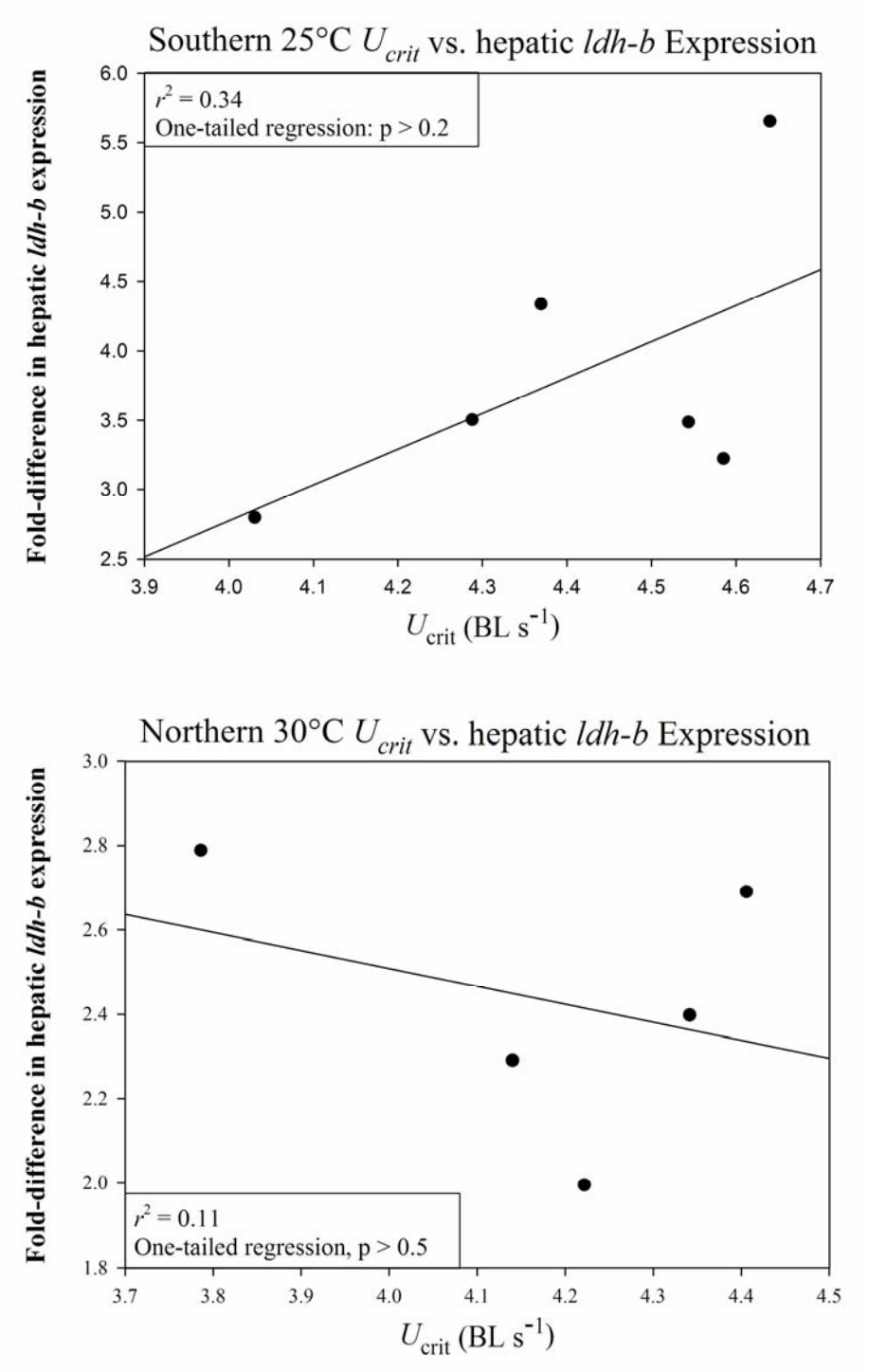


Figure 5.3c. Simple regression analysis comparing *ldh-b* template abundance (fold-difference) vs. relative U_{crit} swimming speed (BL s^{-1}) for non-stressed (control) northern and southern *L. calcarifer* juveniles. Outliers more than two standard deviations from the mean were removed (see Methods).

5.5: DISCUSSION

On the phenomic level, different *L. calcarifer* populations with varying levels of mortality (survivorship) under cold-stress (20°C) and heat-stress (35°C) also differed in their critical swimming performance under these thermal stress conditions (Chapter 2). In light of these phenomic differences we tested hypotheses that cold-stress (20°C), heat-stress (35°C) and non-thermal stress conditions (25°C and 30°C) elicit population-specific responses in their levels of hepatic *ldh-b* transcript abundance (transcriptomic variation) within discrete Australian *L. calcarifer* populations. Additional hypotheses were tested regarding the impact of swimming challenge in the presence and absence of thermal stress, for both northern and southern populations.

We showed that *ldh-b* transcript levels were significantly elevated (> 2-fold) in response to cold-stress (20°C) and heat-stress (35°C) after extended acclimation. Furthermore, we showed that swimming challenge in the absence, but not the presence, of thermal stress also significantly altered hepatic *ldh-b* transcript abundance, as follows: 1) hepatic *ldh-b* gene transcription is significantly elevated in response to swimming stress alone in both the southern and northern populations at their control temperatures (25°C and 30°C respectively); 2) *ldh-b* gene transcription is significantly elevated in response to cold-stress (20°C) alone in the southern population, but not in the northern population; 3) *ldh-b* gene transcription is not significantly elevated following concurrently applied cold- and swimming stress in either the southern or northern populations; 2) *ldh-b* gene transcription is significantly elevated in response to heat-stress (35°C) alone in both populations; 5) *ldh-b* gene transcription is not significantly elevated, but rather collapses to a biologically significant degree, following concurrently applied heat- and swimming stress in both populations.

***Ldh-b* response to swimming (aerobic) challenge in the absence of thermal stress**

Exposure to swimming challenge induced an increase in hepatic *ldh-b* transcript abundance in both the northern and southern *L. calcarifer* populations compared to baseline-transcription levels (at 30°C and 25°C, respectively). This observed increase in hepatic *ldh-b* transcript abundance in both populations is presumably in direct response to the accumulation of lactate in skeletal muscle under the increased aerobic metabolism requirement following the subjection of fish to swimming trials (approx. 200min) designed to impose complete aerobic exhaustion upon challenged fish (Chapter 2). For example, increased levels of serum lactate were shown to be the sole limiting physiological factor in swimming challenged rainbow trout (*Oncorhynchus mykiss*) (Jain and Farrell, 2003). Serum lactate levels have also been reported as a key physiological stress in several other fish species, namely the model North American killifish, *Fundulus heteroclitus* (DiMichele and Powers, 1982; Fangue *et al.*, 2008) and the Atlantic cod, *Gadus morhua* (Herbert and Steffensen, 2005). This physiological increase in serum lactate levels, in concert with the physiological demand for lactate detoxification, is the most parsimonious explanation for the observed fold-increase in hepatic *ldh-b* transcript abundance in both *L. calcarifer* populations following a similar swimming challenge in the absence of thermal stress.

***Ldh-b* response to single, cold-stress**

Cold-stress (20°C) resulted in a significant, almost three fold, increase in hepatic *ldh-b* transcript abundance in *L. calcarifer* from the southern but not the northern population. Our results suggest an additional functional requirement of the hepatic *ldh-b* transcript in responding to cold-stress conditions, such as the maintenance of basal metabolic rates in the southern, but not the northern, *L. calcarifer* populations examined. One potential explanation for this population-specific response to the cold-stress condition would be

the presence of allelic variation in the nucleotide sequences of up- and/or down-stream (5' or 3') untranslated regions of the *ldh-b* locus. Such an allelic variation (1 bp) was identified in the glucocorticoid responsive element (GRE) of the 5' proximal promoter of the *ldh-b* locus in the temperate *F. heteroclitus* and was, furthermore, demonstrated to be the source of the two-fold difference in hepatic *ldh-b* transcript abundance observed in thermally discrete northern and southern populations (Crawford et al., 1999a; Rees et al., 2001; Schulte et al., 2000; Schulte et al., 1997; Segal et al., 1996). Moreover, the significant increase in hepatic *ldh-b* transcript abundance observed in southern juveniles following cold-stress (20°C) exposure is in line with a negative correlation between LDH-B protein abundance and native environmental temperature previously observed across populations of the model temperate *F. heteroclitus* (reviewed by Powers et al., 1991; Powers and Schulte, 1998). In temperate *F. heteroclitus* the observed increase in LDH-B protein abundance with decreasing environmental temperature was attributed to the required doubling of LDH-B per 10°C reduction in temperature in order to maintain kinetic rates (i.e. $K_m = V_{max} / 2$) (reviewed by Powers et al., 1991; Powers and Schulte, 1998). Additional examinations of LDH-B protein abundance and comparisons of metabolic rates in the northern and southern Australian *L calcarifer* populations examined here are required to fully explore the effect of the observed population difference in *ldh-b* transcript abundance under different temperature stresses.

The significant increase in hepatic *ldh-b* transcript abundance in the southern, but not northern, population could be indicative of a local adaptation of *L. calcarifer* to their native thermal habitat. More specifically, southern *L. calcarifer* are evolutionarily accustomed to low winter minimum temperatures and thus exhibit a significant response to the cold-stress (20°C) condition; whereas the northern population is not evolutionarily accustomed to such low winter minimum temperature and thus does not

exhibit a significant increase in *ldh-b* transcript abundance in response to the cold-stress condition. Moreover, the observed significantly lower than expected mortality rate and significantly faster critical swimming speed observed in the southern *L. calcarifer* population under cold-stress (20°C) conditions (Chapter 2) supports the notion of a local adaptation to the native thermal habitats of discrete *L. calcarifer* populations. Evidence toward such a local adaptation to native thermal habitats is especially important in light of the increased winter mortality rates observed for *L. calcarifer* in aquaculture grow-out ponds over the winter months (e.g. Townsville QLD winter, 2007) (Jerry, Dean pers. comm.). Further studies are needed to establish a more comprehensive understanding of how the hepatic transcriptome of *L. calcarifer* responds to extended cold-stress (single whammy), as well as to concurrent cold-stress and swimming challenge (double whammy), since no other study has investigated such transcriptome responses in this non-model tropical species.

***Ldh-b* response to single, heat-stress**

Acclimation to our experimental heat-stress condition (35°C) instilled a significant increase in hepatic *ldh-b* transcript abundance in *L. calcarifer* from both populations. The significant increase in hepatic *ldh-b* transcript abundance in response to heat-stress (35°C) alone demonstrates that this locus is responsive to heat-stress, but not cold-stress, in both *L. calcarifer* populations. However, additional studies encompassing multiple time points of exposure/acclimation to heat-stress are needed to confirm the nature of this observed increase in hepatic *ldh-b* transcript abundance (i.e. durable *v.* transient). For example, an increase in blood glucose concentrations is a common indicator of stress levels in fish and, taken in conjunction with the significant increase in serum lactate levels reported under hypoxic (heat-stress) conditions (Virani and Rees, 2000), is suggestive that heat-stress response requires not only additional glucose, due

to the lesser efficiency of anaerobic metabolism, but also requires more glycolytic enzyme (e.g. *ldh-b*) to detoxify accumulating levels of lactate and preserve homeostasis. Therefore, the increase in hepatic *ldh-b* transcript abundance in response to heat-stress acclimation in both *L. calcarifer* populations is in agreement with these trends and could be driven by its role in sustaining an elevated basal aerobic metabolism, gluconeogenesis, or both, under reduced oxygen conditions linked to increased water temperatures, as observed between thermally discrete *F. heteroclitus* populations (reviewed by Powers et al., 1991; Powers and Schulte, 1998). Moreover, this increase in hepatic *ldh-b* transcript abundance does not appear to convey an increased heat-stress tolerance, as both populations suffered significantly greater than expected mortality rates (Chapter 2); however, and perhaps more interestingly, the less responsive northern population did not suffer mortality to the extent suffered by the southern population following the extended heat-stress (35°C) exposure applied in this study (Chapter 2). The heat shock proteins (*hsc70, hsp70-1*and *hsp70-2*) identified to vary in transcript abundance between thermally discrete temperate *F. heteroclitus* populations (Fangue et al., 2006) could serve as ideal candidate genes for future investigations into the heat-stress responses and thermal tolerance of tropical *L. calcarifer* populations.

In contrast to the documented increase in hepatic *ldh-b* transcript abundance observed following 28 day acclimation of *L. calcarifer* to 35ºC temperatures, a down-regulation of loci involved in lipid metabolism, detoxification, oxygen transport and oxidative phosphorolation, which are believed to reduce the production of reactive oxygen species (ROS), were also observed in the porcelain crab *Petrolisthes cinctipes* following 4 hours of ramping heat-stress (11 to 30°C) exposure (Teranishi and Stillman, 2006). Moreover, the reduction in oxidative phosphorylation and concurrent increase in lactate production via anaerobic glycolysis was observed in nucleated red blood cells of rainbow trout, *Oncorhynchus mykiss* under heat-stress (25°C), as compared to control

(10°C), conditions (Currie et al., 1999). This contradicting evidence suggests that the increase in hepatic *ldh-b* transcript abundance observed in both *L. calcarifer* populations following 28 day heat-stress (35°C) acclimation isn't driven by the role this locus plays in maintaining efficient aerobic metabolism but is instead likely driven by its role in gluconeogenesis. Increasing hepatic *ldh-b* transcript levels for this purpose would effectively detoxify accumulating lactate by catalyzing its conversion to glucose to sustain the less efficient anaerobic metabolism encountered under the hypoxic conditions associated with heat-stress (Podrabsky and Somero, 2004; Powers and Schulte, 1998; Somero, 2005).

The observed increase in hepatic *ldh-b* transcript abundance in these discrete *L. calcarifer* populations may also be linked to the role that the enzyme (LDH-B) plays in moderating hemoglobin-oxygen binding affinity (i.e. "stickiness" of red blood cells), as was demonstrated in the temperate killifish *Fundulus heteroclitus* (Powers, 1980; Powers et al., 1979). However, this specific function of translated LDH-B in *L. calcarifer,* or in any other tropical fish, has yet to be established. Moreover, LDH-B moderated differences in hemoglobin-oxygen affinity were linked to different isozyme variants (i.e. LDH-B^a vs. LDH-B^b) in geographically and thermally discrete *F. heteroclitus* populations (Powers, 1980; Powers et al., 1979); therefore, the impact of increasing hepatic *ldh-b* transcript abundance on Hb-O_2 affinity in *L. calcarifer*, where no such allelic variation has been demonstrated to date (Chapter 4), is unknown.

Ldh-b response to swimming (aerobic) challenge in the presence of cold-stress: the double whammy effect

Neither population exhibited an additional increase in hepatic *ldh-b* transcript abundance following concurrent swimming challenge and cold-stress exposure. This lack of a significant elevation in hepatic *ldh-b* transcript abundance in both populations following subjection to concurrent swimming challenge and cold-stress (20°C) suggests

that in combination these two stresses, surprisingly, have a dampening effect on the rate of gene transcription for this locus, regardless of population. In other words, hepatic *ldh-b* transcription is unable to increase in swimming challenged cold-stressed fish, in contrast to the response when each stress is experienced independently of the other stress.

This response of hepatic *ldh-b* transcript levels to concurrent stresses in thermally discrete and tropical *L. calcarifer* populations mirrors the response observed in the warm southern population *F. heteroclitus* exposed to similar concurrent stresses, in that the doubling of *ldh-b* transcript abundance occurred in response to cold-stress alone but did not increase further in response to a similar and concurrently imposed swimming challenge (reviewed by Powers and Schulte, 1998). Moreover, recent studies indicate a trend of up-regulation for loci involved in transcriptional or translational pathways as they compensate for the depressing effect of cold-stress on transcriptome machinery (reviewed by Douglas, 2006). Our observed increase in hepatic *ldh-b* transcript abundance in the southern population is in line with the general up-regulation observed for loci involved in overcoming the depressing (slowing) effects of cold-stress exposure in an effort to maintain homeostasis (e.g. chromatin structure, translational efficiency and basal metabolic rates) (Gracey et al., 2004; Podrabsky and Somero, 2004). However, the up-regulation of critical homeostasis maintaining loci could in turn impact on the transcription (i.e. transcript abundance) of non-critical loci under the combination of cold- and swimming stress (i.e. double whammy effect). Such a transcriptomic response could explain the lack of significant fold-increases in hepatic *ldh-b* transcript abundance in *L. calcarifer* from both thermally discrete populations following concurrent subjection to cold- and swimming stress (double whammy effect).

Juveniles from the southern population, in which a significant increase (> 3-fold) in hepatic *ldh-b* transcript abundance was observed following 28 day exposure to

the cold-stress (20°C) condition, may have a genetically linked enhancement in their ability to manage the lactate levels which accumulate during swimming challenge in that fish from this southern population were able to sustain aerobic metabolic activity for longer than fish from the northern population when challenged under the cold-stress condition (Chapter 2). However, the role of hepatic *ldh-b* transcript or LDH-B protein in the maintenance of aerobic metabolism under concurrent cold- and swimming stress is unknown for *L. calcarifer* as well as tropical fish in general.

Ldh-b response to swimming (aerobic) challenge in the presence of heat-stress: the double whammy effect

Both populations exhibited a biologically significant decline in *ldh-b* transcript abundance following concurrent exposure to swimming challenge and heat-stress (35°C). This reduction in *ldh-b* transcript abundance following swimming challenge is plausibly caused by the lack of transcriptional machinery available for the continued expression of *ldh-b* in light of the essential synthesis of proteins required for the maintenance of homeostasis (e.g. heat-shock proteins and other molecular chaperones) following such double-stress exposure (i.e. the double whammy effect). For example, only two loci involved in protein synthesis (40S ribosomal proteins S17 and S5) are significantly up-regulated in response to unchallenged heat-stress in both the African lake fish *Austrofundulus limnaeus* at 37°C (Podrabsky and Somero, 2004) and the coral reef fish *Pomacentrus moluccensis* at 34°C (Kassahn et al., 2007). These up-regulated ribosomal loci are likely reserved for the synthesis of critical proteins only, as all other assessed loci involved in transcription or translational processes are significantly down-regulated in response to unchallenged heat-stress exposure (Kassahn et al., 2007; Podrabsky and Somero, 2004). The observed decrease in *ldh-b* transcript abundance following concurrent swimming challenge and heat-stress exposure suggests, as suggested above for concurrently applied cold- and swimming stress, that this locus is

not critical in the maintenance of homeostasis under such extreme circumstances. The exact mechanism behind this double whammy effect is unknown; however, the observed transcriptomic response to these pertinent and concurrently applied stressors (heat and swimming) may be indicative of the southern populations' upper thermal tolerance limit.

Autocorrelation between hepatic *ldh-b* transcript abundance and critical swimming performance

The significant negative correlations observed between critical swimming speed and fold-differences in hepatic *ldh-b* transcript levels are confounding and raise several questions: Is the abundance of hepatic *ldh-b* transcript 1:1 with the abundance of translated LDH-B protein following thermal acclimation in tropical *L. calcarifer*? If so, is the ability of LDH-B to convert lactate to pyruvate 1:1? The task of establishing the translational efficiency of *ldh-b* mRNA to LDH-B protein in Australian *L. calcarifer* should be the aim of future studies and can be achieved by standard protein quantification using barramundi-specific LDH-B antibodies (via Western blot). Following this, autocorrelation analyses can be preformed to verify if the relationship between *ldh-b* mRNA and LDH-B protein is 1:1 for each discrete *L. calcarifer* population following acclimation to 20, 25, 30 and 35°C.

This method was utilized to investigate the relationship between *ldh-b* mRNA and LDH-B protein in thermally discrete *F. heteroclitus* populations (Segal and Crawford, 1994a). These were found to be highly correlated ($r^2 = 0.81$, $p < 0.01$) for both cold northern and warm southern *F. heteroclitus* populations acclimated to 20°C (Segal and Crawford, 1994a). This correlation did not hold true, however, following the acclimation of *F. heteroclitus* populations to 10°C ($r^2 = 0.22$, $p > 0.25$ and $r^2 = 0.57$, $p = 0.086$ respectively) (Segal and Crawford, 1994a). Such investigations also provided key insight regarding differences in the regulatory processes of *ldh-b* in discrete populations

of temperate *F. heteroclitus* (Segal and Crawford, 1994a) and are thus ideal to employ for similar investigations in tropical Australian *L. calcarifer*. More specifically, if *ldh-b* mRNA abundance decreases while LDH-B protein abundance remains consistent in *L. calcarifer* populations following acclimation to 20, 25, 30 and 35°C this would indicate a difference in the transcriptional regulation of *ldh-b* in response to temperature acclimation. Moreover, the maintenance of LDH-B protein abundance despite reductions in the abundance of *ldh-b* mRNA would also suggest the presence of a translational and/or post-translational mechanism (Crawford and Powers, 1992; Segal and Crawford, 1994a). Unfortunately, the hepatic tissue used for quantification of *ldh-b* transcript abundance quantification in the present study cannot be subsequently used for protein quantification as LDH-B was not retained during the RNA extraction protocol, for technical reasons (see Methods). However, this correlation could be established by quantifying both *ldh-b* transcript and LDH-B protein abundances in the collected (snap-frozen) cardiac tissue.

5.6 CONCLUSIONS

This chapter assesses the level of variation in hepatic *ldh-b* transcript abundance (i.e. transcriptomic variation) and demonstrates a significant difference in the abundance of hepatic *ldh-b* transcripts between thermally distinct *L. calcarifer* populations (Figure 5.1). Specifically, these populations demonstrate significant differences in the levels of their hepatic *ldh-b* transcript abundance following 28 day acclimation to both cold-stress (20°C) and heat-stress (35°C) conditions as well as in response to swimming challenge both in the presence and absence of thermal stress. In light of these results it appears as if the southern population possesses a unique transcriptomic response to cold-stress due to it being evolutionarily accustomed to such conditions through

minimum winter temperatures. The lesser mortality and higher critical swimming performance previously observed for the northern population following acclimation to the heat-stress (35°C) condition (Chapter 2) suggests that this population is more evolutionarily accustomed to a warmer thermal regime than the southern cool water population. It does not appear, however, as if hepatic *ldh-b* transcript abundance can explain the significant difference in U_{crit} swimming speed observed between these populations when under cold-stress and heat-stress conditions (Chapter 2). In conclusion, investigating differences in *ldh-b* transcript abundance, as compared to the *ldh-b* genotype (Chapters 3 and 4), provides unique insights regarding observed phenotypic differences (Chapter 2). The final molecular component required to complete the multi-level approach of this thesis is analytically addressed in the following data chapter by comparing *ldh-b* coding nucleotide and deduced LDH-B amino acid sequences.

Chapter 6

Evidence for selection on the lactate dehydrogenase-B (*ldh-b*) locus within and among genera of tropical fish, *Lates* and *Plectropomus*: an analytical approach

---

6.1: ABSTRACT

Identifying selective constraint on thermally sensitive loci in non-model fish species can provide novel insight regarding their potential adaptation to local thermal regimes. This chapter presents analytical evidence, in lieu of empirical proteomic data, toward the identification of selective constraint acting upon the lactate dehydrogenase-B (*ldh-b*) locus in the non-model tropical congeneric *Lates* (*L. calcarifer* and *L. niloticus*) and *Plectropomus* (*P. leopardus* and *P. laevis*) species. Analyses utilize multiple pairwise comparisons within and among the non-model tropical species and the model temperate *Fundulus* (*F. heteroclitus* and *F. parvipinnis*) species. Pairwise and phylogenetic comparisons of both coding nucleotide and amino acid sequences indicate that tropical *Lates* and *Plectropomus* species are more homologous to one another (94 – 96%) than either is to the temperate *Fundulus* species (90 – 92%). The congeneric tropical *Plectropomus* species exhibit a higher homology with the warm southern than the cold northern *F. heteroclitus* population, suggesting that increased water temperature is an environmental stressor for this metabolic locus/enzyme. Curiously, valine and arginine are in significantly higher and lower abundances than expected ($\chi^2_{0.001,\,18} = 42.31$; $p < 0.001$), respectively. Evidence of such selective constraint provides additional evidence for local adaptation of tropical fishes to their respective thermal habitats.

6.2: INTRODUCTION

Environmental variables, such as temperature and oxygen availability, are selective forces which actively shape the evolution of species that are adapted to specific environments. More specifically, evolutionary adaptations manifest as alterations (mutations) in the nucleotide composition of relevant parts of the genome and are more commonly reported for coding (exon) (Calladine *et al.*, 2004; Frank-Kamenetskii, 1997; Kao and Farley, 1978; Kirpichnikov *et al.*, 1990; Merritt, 1972; Oleksiak *et al.*, 2002; Powers *et al.*, 1991; Powers and Schulte, 1998; Quattro *et al.*, 1993; Steffensen, 2002; Tsuji *et al.*, 1994; Xia, 1998) than for non-coding (intron) regions (Hare and Palumbi, 2003; Lambowitz and Belfort, 1993; Mattick, 1994; Meloni et al., 1998; Roger and Doolittle, 1995; Wittekindt et al., 2000). However, we now know that non-coding regions are also implicated in adaptive change, due to selection for gene expression levels to be up- or down regulated (Ambros, 2004; Baek et al., 2008; Bartel, 2004; Bonnet et al., 2004; Brennecke et al., 2005; Lai, 2003; Li et al., 2004; Mattick, 1994). In this chapter we focus on evidence for selection in exons (coding regions) of *ldh-b* alone, since this has been considered for the non-coding introns previously (Chapters 3 and 4; Appendix 1). Genetic substitutions (transitions or transversions) occurring within exons lead to either synonymous or nonsynonymous (replacement) amino acid substitutions (Xia, 2000). Nonsynonymous substitutions tend to involve the replacement of amino acids with those of similar properties (e.g. polarity) (Grantham, 1974; Xia, 1998). However, being that amino acids can be characterized based on a wide array of traits (n = 134 *sensu* Sneath, 1966), this study focuses on those values given to each amino acid by Grantham (1974) (polarity, volume and side-chain composition) entitled Grantham's distances (D_G) herein, to characterize the nature and impact of observed nonsynonymous substitutions.

To determine if evolutionary adaptation is occurring between geographically and thermally distinct species pairs of both temperate and tropical fishes, this study utilized a candidate gene approach. Lactate dehydrogenase-B (*ldh-b*) is an ideal candidate gene to identify evidence of thermal adaptation due to its response to both thermal and swimming stresses in the model temperate killifish *Fundulus heteroclitus* (see reviews by (Powers *et al.*, 1991; Powers and Schulte, 1998) as well as in the non-model but tropical *L. calcarifer* (Chapter 5). More specifically, the enzyme (LDH-B) plays a critical role in maintaining aerobic metabolism by converting lactate, the major by-product of anaerobic glycolysis, to pyruvate via oxidation or to glucose via gluconeogenesis (Place and Powers, 1984a, 1984b). This conversion of accumulating lactate from aerobic tissues (e.g. heart and skeletal muscle) occurs in the liver and allows aerobic metabolic activity to be sustained for extended lengths of time (DiMichele and Powers, 1982; Powers et al., 1991; Powers and Schulte, 1998; Segal and Crawford, 1994b). In addition to these metabolic functions, LDH-B affects the oxygen binding affinity of hemoglobin (Hb) by altering intraerythrocyte ATP concentrations of Hb in fish (Powers, 1980; Powers *et al.*, 1979). The increase or decrease in hemoglobin's ability to bind and transport oxygen is affected by the particular LDH-B variant (i.e. LDH-B^{a}, ^{b}) selected for, as demonstrated in *F. heteroclitus* by Powers *et al*. (1979). This effect of LDH-B on Hb-O_2 binding affinity directly impacts delivery of Hb-bound oxygen to red muscle tissues and may therefore be an alternate mechanism by which LDH-B affects sustainability of aerobic performance (e.g. swimming performance) in fish (reviewed by Powers et al., 1991; Powers and Schulte, 1998; Segal and Crawford, 1994b).

Thermally distinct populations of the temperate estuarine killifish *F. heteroclitus* were studied extensively by Powers *et al*. (1979 - 2000). These studies provided conclusive evidence that the *ldh-b* locus is under selection in populations native to the

two thermal and geographical extremes of this species' distribution range (reviewed by Powers et al., 1991; Powers and Schulte, 1998). More specifically, *F. heteroclitus* populations are separated into "cold" northern (Nova Scotia, Canada), "warm" southern (Florida, USA) and "intermediate" (New Jersey through Georgia, USA) water temperature regimes across the steep thermal gradient (1°C per 1° latitude) present along the eastern seaboard of the United States (reviewed by Powers *et al.*, 1991; Powers and Schulte, 1998). In summary, one non-fixed and two fixed SNPs were reported between thermally distinct populations for the coding nucleotide (exon) sequence of the *ldh-b* locus, with all three cause nonsynonymous amino acid substitutions (Bernardi *et al.*, 1993; Place and Powers, 1984a, 1984b). These reported differences for the *ldh-b* locus exons and protein were concluded to be adaptive and likely generated by the unique selective pressure imposed by the thermally distinct regions inhabited by cold northern and warm southern populations of *F. heteroclitus* (i.e. depressing effects of northern cold water or hypoxic effects of warm southern waters) (reviewed by Powers et al., 1991; Powers and Schulte, 1998). This chapter utilizes a comparative genomics approach to identify evidence regarding whether or not selective constraint appears to be acting upon the *ldh-b* locus in the congeneric pairs of the non-model tropical *Lates* and *Plectropomus* species, as *ldh-b* is known to be under selection in *F. heteroclitus* (Crawford et al., 1999b; Pierce and Crawford, 1997).

Temperate and tropical regions span relatively narrow geographical distances, yet impose quite different selective forces upon aquatic organisms inhabiting either. Temperate zones span 23.5°N to 66.6°N and 23.5°S to 66.5°S, whereas the tropical zone extends between 23.5°S and 23.5°N. These distinct geographical regions impose unique selective pressures, which have impacted on (adaptive) genes in both tropical fish (e.g. catadromous (Chenoweth et al., 1998; De Santis et al., 2008; Katersky and Carter, 2005, 2007; Keenan, 1994) and coral reef (Kassahn et al., 2007) species) and

temperate fish (e.g. North American (Crawford et al., 1999b; Oleksiak et al., 2002; Rees et al., 2001) and African (Podrabsky and Somero, 2004) killifish). However, adaptation of the *ldh-b* locus in tropical fish living at temperatures greater than 20°C has not been reported to date. In light of the recent characterization of the ldh-b locus in tropical fish species, representing two congeneric *Lates* (*Lates calcarifer* and *L. niloticus*) and *Plectropomus (P. leopardus* and *P. laevis*) species pairs (Chapter 3 and Appendix 1 respectively), we can, for the first time, utilize a comparative genomics approach to assess this thermally sensitive locus for evidence of selective constraint both among and within tropical fishes. More specifically, multiple pairwise comparisons between these tropical congeneric species pairs and the model temperate congeneric *Fundulus* (*F. heteroclitus* and *F. parvipinnis*) species pair are utilized.

The first pair of congeneric species pair, barramundi (*L. calcarifer*) and nile perch (*L. niloticus*) are catadromous and freshwater perciformes, respectively, with allopatric tropical distribution ranges. *L. calcarifer* occurs throughout the Asian archipelago (13°N – 10°S) and the central to northern rivers of Australia (25°S – 12°S) (Chenoweth et al., 1998; Doupé et al., 1999; Katersky and Carter, 2005, 2007; Keenan, 1994; Salini and Shaklee, 1988; Shaklee and Salini, 1985), while *L. niloticus* occurs throughout the rivers and lakes of East Africa (7°S – 27°N) (Ribbink, 1987). The natural ranges of *L. calcarifer* and *L. niloticus* encompass similar tropical thermal environments with water temperatures ranging from approx. 18 to 34°C at the extremes.

The second pair of congeneric species, the coral trouts *Plectropomus leopardus* and *P. laevis,* are sympatric tropical perciform fishes that inhabit coral reefs and exhibit a natural distribution range from 28°S (Abrolhos Islands) to 23°N (Taiwan) (Condie and Dunn, 2006; Hwang and Wong, 2005; Lough, 1998; van Herwerden et al., 2002). The habitat ranges of both *Plectropomus* species span temperatures of 19.5°C to > 30°C (Lough, 1998) and are of similar body size, making them ideal marine counter-parts to

the tropical catadromous and freshwater *L. calcarifer* and *L. niloticus*, respectively. Moreover, these congeneric coral trout species are sympatric with some level of ecological partitioning (*P. leopardus* is most abundant on reefs intermediate to inshore and off shore habitats, but does occur in all three, while *P. laevis* is most common in offshore habitats, but is occasionally seen on reefs closer to shore) (van Herwerden et al., 2002). This chapter presents a novel comparative genomics assessment between these non-model tropical congeneric species and the model temperate *F. heteroclitus*, as a means to identify evidence toward the presence of selective constraint acting upon the coding nucleotide and deduced amino acid sequences of the *ldh-b* locus as this locus is known to be under selection in *F. heteroclitus* (Crawford et al., 1999b; Pierce and Crawford, 1997).

6.3: METHODS

Sequence editing, alignment, verification, analysis and gene map construction were completed with the use of several programs. The programs and the respective uses are as follows: BioEdit for sequence editing (Hall, 1999); ClustalW for sequence alignment and verification (Higgins *et al.*, 1994; Lopez and Lloyd, 1997); NCBI Blast for sequence verification (Zhang and Madden, 1997; Zhang et al., 2000); *MEGA 3.1* (Kumar *et al.*, 2004) for comparisons of *ldh-b* and LDH-B sequences among fish species; *DAMBE (v.5.20)* (Xia and Xie, 2001) for analysis of natural selection and molecular evolution among and within species; Adobe Illustrator (CS3) and Microsoft Excel for figure and table construction respectively. Parsimoniously informative single nucleotide polymorphisms (SNPs) present within exon sequences were also assessed both among and within congeneric species of the genera *Lates*, *Plectropomus* and *Fundulus* via *MEGA 3.1* (Kumar *et al.*, 2004) and are presented in a pairwise fashion

(Table 6.1). Isoelectric point differences were assessed via multiple t-tests (df = 1; ά = 0.05) in SPSS 16.

More specifically, *D.A.M.B.E. (v.5.20)* (Xia and Xie, 2001) was used to assess *ldh-b* nucleotide and amino acid sequences for evidence of selection in the following ways: a) abundance of particular amino acids and available codon number; b) observed vs. expected transition (s) and transversion (v) values rendering single-step nonsynonymous codon substitutions (SSNCSs) at all codon positions, as well as corresponding Grantham's Distances (D_G) based on selected sequences (observed) and standard genetic code (expected); c) variation in amino acid sequences and d) assessment of isoelectric point (pI) between considered species (Xia, 2000; Xia and Xie, 2001). Moreover, when assessing variation across amino acid sequences, Grantham's distance was utilized exclusively. The distance between amino acids was internally established by *D.A.M.B.E.* based on Grantham (1974), which assessed three characteristics: a) polarity; b) volume; c) composition of side-chain (Grantham, 1974; Xia, 1998, 2000). The distances output by *D.A.M.B.E.* are presented in a pairwise fashion and divided into two groups, $D_G > 100$ and $D_G < 100$ as this relates to highly dissimilar and similar amino acids, respectively (Grantham, 1974) (Table 6.1c). All pairwise comparisons with *F. parvipinnis* are across slightly non-homologous sequence lengths, on the order of 37bp and 11 residues, as these are lacking from the coding nucleotide and amino acid sequences, respectively, due to the available GenBank sequence for this species being incomplete [**L23780**].

6.4: RESULTS

Ldh-b locus structure

In the six species examined all seven *ldh-b* exons, encoding LDH-B, were found to be conserved in size; whereas intron sizes varied from 88 to 2560bp in length (Figure 6.1**)**. Size comparisons of homologous introns between congenerics were generally similar (e.g. *Lates spp*., *Plectropomus spp*. and *Fundulus spp*. intron 2 of 88, 92 and 104bp, respectively); however, major size differences were observed in several introns (e.g. introns 1 and 5 of *L. calcarifer* and *Plectropomus spp*. differ by 1,308 bp and 1,537 bp, respectively) (Figure 6.1). The only comparison of indels that could be made between *Fundulus spp*. is the comparison of intron 2 between *F. heteroclitus* and *F. grandis*, as this non-coding sequence, obtained by Reese, B.B. and Schulte, P.M. (unpublished), is the only such intron sequence available for *Fundulus* on GenBank (AF322874, AF322873, AF322872, AF322871). This comparison reveals the presence of three insertion-deletion (indel) events (2, 7 and 8bp) in an otherwise highly conserved sequence (98% homologous). All such indels observed are confined to intron sequences (Figure 6.1**)**. More specifically, three indels (187, 567 and 745bp) exist between congeneric *Lates* species in introns 2, 5 and 6, respectively (Figure 6.1**)**. Congeneric *Plectropomus* species also possess a variety of indels across introns 1, 4, 5 and 6; however, these indels are an order of magnitude smaller that those present between *Lates* species (Figure 6.1).

Comparative statistics: *Ldh-b* exon (nucleotide) sequences

Assessment of *ldh-b* exon sequence (1,005bp) across all six species presents no significant deviation from the expected equal frequency (0.25) of each nucleotide ($\chi^2_{0.05, 3} = 7.82$; $p > 0.05$). In contrast, a significant deviation does appear when considering

dinucleotide abundance (AA, CC, TT, GG) at codon positions 1 - 2, 2 - 3 and 3 - 1 ($\chi^2_{0.001,\ 47}$ = 82.72; $p < 0.001$; data not shown). Pairwise comparison demonstrates that maximum homology (99.4%) exists within *F. heteroclitus* populations ("warm" southern vs. "cold" northern), as expected (Table 6.1). Additional pairwise comparisons on the species level reveals the second highest sequence homology to exist between *ldh-b* of congeneric pairs (Table 6.1). Moreover, *Lates* species are more homologous to *Plectropomus* species (89.8 - 91.9%) than *Lates* species are to *Fundulus* species (86.8 - 88.3%) (Table 6.1). A subsequent comparison among the four tropical species reveals that both *Lates* species are more homologous to *P. laevis* (91.3 - 91.9%) than they are to *P. leopardus* (89.8 - 90.4%) (Table 6.1**)**. When such pairwise comparisons incorporate the temperate species as well, *P. leopardus* and *P. laevis* both exhibit an increased homology with the warm southern *F. heteroclitus* population (88.2 and 89.2%, respectively) than to the cold northern population (88.0 and 89.0%, respectively); however, *L. calcarifer* and *L. niloticus* exhibit an equal level of homology to both southern and northern *F. heteroclitus* populations (88.3 and 87.5%, respectively) (Table 6.1). Interestingly, *F. parvipinnis* exhibits a slightly higher homology to *L. calcarifer* and *P. laevis* (87.7 and 87.2%, respectively) than to *L. niloticus* and *P. leopardus* (86.8 and 86.5%, respectively; Table 6.1); however, this trend does not hold true for amino acid sequence comparisons (Table 6.1; see below**)**.

Table 6.1. Pairwise comparison of variation between coding nucleotide and deduced amino acid sequences.

%V	*L.cal*	*L.nilo*	*P.leo*	*P.lae*	*F.het*(N)	*F.het*(S)	*F.parv*
L.cal		2.90	9.57	8.08	11.74	11.74	12.29
L.nilo	1.50		10.17	8.67	12.46	12.46	13.22
P.leo	4.80	6.01		0.90	11.96	11.76	13.53
P.lae	3.90	5.11	0		10.97	10.77	12.81
F.het(N)	7.49	8.71	7.21	6.31		0.60	5.37
F.het(S)	7.19	8.41	6.91	6.01	0.90		5.17
F.parv	8.67	9.60	8.36	7.43	2.79	2.48	

Numbers above and below the grayed diagonal refer to the variation (%) in coding nucleotide and deduced amino acid sequences, respectively.

Comparative statistics: amino acid sequences

Assessment of the entire LDH-B amino acid sequence (334 amino acids) across all six considered species presents no significant variation in the abundance of each individual amino acid ($\chi^2_{0.05, 19} = 30.14$; $p > 0.05$). Amino acid abundance based on the relative codon abundance of each residue does, however, present a significant deviation from the expected amino acid abundance based on the number of encoding codons ($\chi^2_{0.001, 18} = 42.31$; $p < 0.001$; Table 6.2). Most noteworthy of these unexpectedly abundant amino acids are the respective increase and decrease of Valine and Arginine, which have 4 and 6 encoding codons, respectively (Table 6.2). Pairwise comparisons of amino acid sequences revealed that maximum homology (99.1%) exists between *F. heteroclitus* populations (i.e. warm southern vs. cold northern), as was observed for exon sequences (see above). Species-level pairwise comparisons revealed that the second highest amino acid homology exists between congeneric species pairs (Table 6.1), as was also observed for exon sequences (see above). Moreover, within tropical species comparison revealed higher homology between *L. calcarifer* and *P. laevis* (96.1%) than between *L. niloticus* and *P. leopardus* (94.0%) as well as that all four tropical species are more homologous to warm southern (91.6 – 94.0%) than to cold northern populations (91.3 – 93.7%) of *F. heteroclitus* (Table 6.1).

Table 6.2. Summary of amino acid abundance by type across all considered species (pooled).

a	AA	Number	Percent	Cod/AA	Pred	Resid	*b*	r^2	Slope	df	*p*
	Ala	154	6.63	4	142.9	11.1		0.393	28.189	18	0.0008
	Arg	87	3.75	6	199.3	-112.3					
	Asn	111	4.78	2	86.6	24.4					
	Asp	129	5.55	2	86.6	42.4					
	Cys	31	1.33	2	86.6	-55.6					
	Gln	75	3.23	2	86.6	-11.6					
	Glu	101	4.35	2	86.6	14.4					
	Gly	172	7.4	4	142.9	29.1					
	His	42	1.81	2	86.6	-44.6					
	Ile	113	4.86	3	114.7	-1.7					
	Leu	246	10.59	6	199.3	46.7					
	Lys	149	6.41	2	86.6	62.4					
	Met	68	2.93	1	58.4	9.6					
	Phe	41	1.76	2	86.6	-45.6					
	Pro	81	3.49	4	142.9	-61.9					
	Ser	201	8.65	6	199.3	1.7					
	Thr	133	5.73	4	142.9	-9.9					
	Trp	42	1.81	1	58.4	-16.4					
	Tyr	47	2.02	2	86.6	-39.6					
	Val	300	12.91	4	142.9	157.1					

Section *a*: Column headings "Cod/AA", "Pred" and "Resid" refer to the number of codons encoding each amino acid, the predicted abundances based on number of available codons (per D.A.M.B.E. (Xia, 2000) and the residual difference (Number – Predicted) respectively. Section *b*: Regression analysis statistics generated by D.A.M.B.E.

Of additional interest are the significant similarities and differences observed between species when assessed on the standard deviation (S.D.) of their respective isoelectric points (pI) (Table 6.3). The observed pI for *L. calcarifer* (7.47) is 1.65 S.D. lower than its congeneric *L. niloticus* (7.85), a similar degree of difference as observed for the pI value of the southern warm (7.43) *F. heteroclitus* population which is 1.48 S.D. lower than its cold northern counterpart (7.77). Conversely, the pI values of *L. calcarifer* and the southern warm *F. heteroclitus* population are exceedingly similar in that they only differ by 0.04 S.D. (Table 6.3). The pI values of *P. leopardus* (7.86), *P. laevis* (7.85) and *L. niloticus* (7.85) are all also exceedingly similar to one another (± 0.01 S.D.) as well as to the cold northern *F. heteroclitus* population (± 0.09 S.D.; Table

6.3). Lastly, and perhaps most interesting, the pI of *F. parvipinnis* (8.08) appears to be the most divergent, and basic, with an isoelectric point 1 S.D. higher than *P. leopardus*, *P. laevis* and *L. niloticus*, 1.5 S.D. higher than the cold northern *F. heteroclitus* population and 2.8 S.D. higher than both the warm southern *F. heteroclitus* population and *L. calcarifer* (Table 6.3).

Table 6.3. Amino acids utilized for isoelectric point (pI) determination.

Species	Arg	Lys	His	Tyr	Cys	Glu	Asp	pI
L. calcarifer	12	22	6	7	4	15	19	7.47
L. niloticus	12	22	6	7	4	13	20	7.85
P. leopardus	13	20	6	7	4	15	17	7.86
P. laevis	13	21	6	7	4	15	18	7.85
F. heteroclitus (N^{th})	12	22	6	6	5	14	19	7.77
F. heteroclitus (S^{th})	12	22	6	6	5	15	19	7.43
F. parvipinnis	13	20	6	7	5	14	17	8.08
Mean	12	21	6	7	4	14	18	7.76
Standard Deviation	1	1	0	0	1	1	1	0.23

Amino acid abbreviation and their relative pK values utilized by D.A.M.B.E. to compute isoelectric point (pI) are as follows: Arginine (Arg: 12.50), Lysine (Lys: 10.79), Histidine (His: 6.50), Tyrosine (Tyr: 10.95), Cysteine (Cys: 8.30), Glutamine (Glu: 4.25), Asparagine (Asp: 3.91). Additional computation value is put on NH_2 (8.56) and COOH (3.56) groups (data not shown).

Comparative statistics: nonsynonymous amino acid substitutions

A total of 393 nonsynonymous amino acid substitutions are present across all pairwise comparisons (Table 6.4). The clear majority of these are between amino acids with Grantham's distances (D_G) < 100 (Table 6.4), which indicates the presence of selective constraint (Grantham, 1974; Xia, 2000). Those nonsynonymous amino acid substitutions remaining (i.e. D_G > 100) follow the trends present for both exon and amino acid sequence pairwise comparisons (see above) in that the fewest substitutions occur among congenerics and between tropical congenerics (*Lates* spp. *v*. *Plectropomus* spp.), followed by the tropical versus temperate comparison (*Lates* spp. or *Plectropomus* spp. *v*. *Fundulus* spp.) (Table 6.4).

Table 6.4. Pairwise comparison of nonsynonymous amino acid substitutions.

D_G	*L.cal*	*L.nilo*	*P.leo*	*P.lae*	*F.het*(N)	*F.het*(S)	*F.parv*
L.cal		1	1	1	3	2	4
L.nilo	4		2	2	4	2	6
P.leo	15	18		0	2	1	4
P.lae	12	15	0		4	3	6
F.het(N)	22	25	22	17		1	4
F.het(S)	22	26	22	17	2		3
F.parv	24	24	23	18	5	5	

Numbers above and below the grayed diagonal refer to nonsynonymous amino acid substitutions (n = 393) which are either dissimilar (D_G > 100) or similar (D_G < 100) in nature (polarity, volume and side-chain composition, *sensu* (Grantham, 1974), respectively (see Methods).

Of these 393 nonsynonymous substitutions 262 utilize single-step nonsynonymous codon substitutions (SSNCSs), the majority of which (78.6%) are due to transversions (Table 6.5). A significant deviation from neutral chance expectation exists, suggesting that selective pressure or constraint is plausibly driving the total abundance of observed transitions (s) and transversions (v) ($\chi^2_{0.001,\ 1}$ = 10.83; p < 0.001). Moreover, the observed s and v occurrences at all three codon positions, which is presumably driven by s/v occurrences in the 1st and 2nd codon positions, being that 3rd codon position SSNCSs are rare and were not separated from the other codon positions in this test, also differed significantly from neutral chance expectation ($\chi^2_{0.001,\ 5}$ = 20.52; p < 0.001). A significant deviation from chance also exists for each codon position's mean D_G, which again includes the rare 3rd codon position, providing further evidence toward the presence of selective constraint upon *ldh-b* ($\chi^2_{0.001,\ 2}$ = 13.82; p < 0.001). A similar trend is observed when only tropical species (*Lates spp*. and *Plectropomus spp*.) are considered (Table 6.5**)** in that total s and v abundances also deviate significantly from chance expectations ($\chi^2_{0.01,\ 5}$ = 15.09; p < 0.01). Furthermore, the D_G for all codon positions also deviate significantly from neutral expectations ($\chi^2_{0.001,\ 2}$ = 13.82; p < 0.001). The proportion of substitutions at 1st, 2nd or 3rd codon positions leading to SSNCSs also present a significant deviation from the

chance expectation of neutrality ($\chi^2_{0.01, 2} = 9.21$; $p < 0.01$; Table 6.5). The observed proportion of such codon substitutions leading to SSNCSs differ more significantly from expected proportions when only *Lates* species are considered ($\chi^2_{0.001, 2} = 13.82$; $p < 0.001$) (data not shown). This trend, however, does not hold true for comparisons between congeneric *Plectropomus* species or *Fundulus* species (data not shown).

Table 6.5. Single-step nonsynonymous codon substitutions (SSNCSs)

a		1st Codon Position	2nd Codon Position	3rd Codon Position	Sub	Prop	*c*		1st Codon Position	2nd Codon Position	3rd Codon Position	Sub
	s	39	17	0	56	0.214		*s*	35.11	39.038	1.31	75.46
	Mean G	13.08	102.8	40.3				Mean D	54.46	79.36	10	66.63
	v	93	84	29	206	0.786		*v*	74.41	79.386	33.012	186.5
	Mean G	58.46	56.31	60.93	57.93			Sum	109.5	118.42	34.322	262
	Sum	132	101	29	262			Mean D	67.41	104.6	70.33	83.73
	Prop	0.504	0.385	0.111				Sum	0.418	0.451	0.131	1
	Mean G	45.05	64.13	60.93				Mean D	63.26	96.28	68.19	
b	*s*	3	3	0	6	0.12	*d*	*s*	6.65	7.45	0.25	14.35
	Mean G	38	76	57				Mean D	54.45	79.15	10	66.49
	v	16	16	12	44	0.88		*v*	14.15	15.1	6.4	35.65
	Mean G	32.06	50.5	64.25	47.55			Sum	20.8	22.55	6.65	50
	Sum	19	19	12	50			Mean D	67.04	104.77	70.26	83.59
	Prop	0.38	0.38	0.24	6.9			Sum	0.416	0.45	0.133	1
	Mean G	33	54.53	64.25				Mean D	63.01	96.32	68.05	

Transitions and transversions are indicated by *s* and *v*, respectively. Panels *a* and *b*: observed SSNCSs occurrence by codon position and type (*s* or *v*) considering all species together (i.e. pooled) and only tropical species (*Lates* and *Plectropomus*), respectively. Panels *c* and *d*: expected SSNCSs occurrence by codon position and type (*s*/*v*) considering all species together (i.e. pooled) and only tropical species (*Lates* and *Plectropomus*), respectively. Mean Grantham's Distances are indicated by row headings "Mean G" and "Mean D" in (*a, b*) and (*c, d*) respectively (see **Methods**). The subtotal and proportion of SSNCSs occurring across all three codon positions, due to transitions and transversions, are indicated by row headings "Sub" and "Prop", respectively.

Comparative statistics: non-coding (intron) sequences

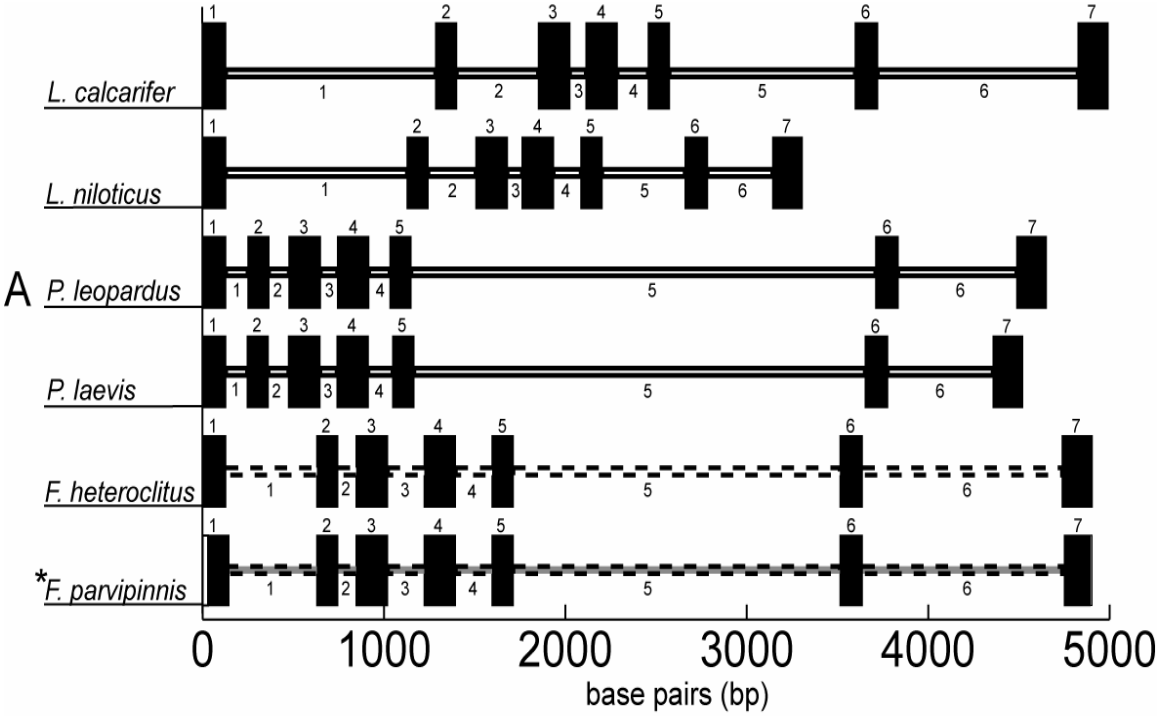
Comparison of intron sequences reveals that only 2.3% and 3.6% variation exists within *Lates spp.* and *Plectropomus spp*., respectively (Figure 6.1). Within *Plectropomus* species there is an increase in variation from 2.6% to 3.6% when considering exon and intron sequences, respectively (Figure 6.1). However, the decrease exon and intron sequence variation within *Lates* species from 4.8% to 2.3%, respectively, is unanticipated (Figure 6.1). For full *ldh-b* non-coding region characterization for *Lates spp*. and *Plectropomus spp*. see Chapter 3 and Appendix 1, respectively. No *ldh-b* intron sequence is available for *F. parvipinnis* on GenBank; thus, the gene characterization (Figure 6.1) assumes *F. parvipinnis* to possess homologous intronic structure with *F. heteroclitus* as both exon and amino acid sequence comparisons demonstrate a maximum homology to exist within, rather than among, these temperate congenerics (Table 6.1).

Phylogenetic analysis of LDH-B amino acid sequences

Phylogenetic analysis of amino acid sequences partitioned tropical genera (*Lates* and *Plectropomus*) and temperate genera (*Fundulus*) into two distinct clades (Figure 6.2). Within the tropical clade *Lates* species are clearly partitioned from *Plectropomus* species, as is also true for *Fundulus* species within the temperate clade (Figure 6.2). Within the *Fundulus*, *Lates* and *Plectropomus* species pairs 10, 5 and 0 acid differences were observed, respectively (Figure 6.2).

FIGURES

A



B

Species	Intron					
	1	2	3	4	5	6
L. calcarifer	1155	449	88	169	1023	1103
L. niloticus	998	262	80	150	456	358
P. leopardus	120	110	92	115	2560	652
P. laevis	117	110	92	134	2486	580
F. heteroclitus	500	100	200	200	1800	1100
**F. parvipinnis*	500	100	200	200	1800	1100

Figure 6.1. *Ldh-b* comparative gene maps. **A.** Top numbers refer to exons, bottom numbers refer to introns. Solid, clear-dashed and grayed horizontal lines refer to non-coding (intron) sequences. *Ldh-b* intron sizes are based on obtained nucleotide sequences for *Lates* and *Plectropomus* species while the intron sizes of *F. heteroclitus ldh-b* are approximated (*sensu* Powers, *et al.* 1998). *F. parvipinnis ldh-b* intron sizes are assumed to be homologous to those approximated for *F. heteroclitus*, as indicated by * (**see Results Id**). White areas at the extreme 5' and 3' ends of exons 1 and 7 indicate regions of unknown *ldh-b* coding sequence for *F. parvipinnis*. **B.** Total number of base pairs present within each intron (numbered) for all considered species (n = 6). Intron sizes assumed to be homologous between *F. heteroclitus* and *F. parvipinnis* are represented by *.

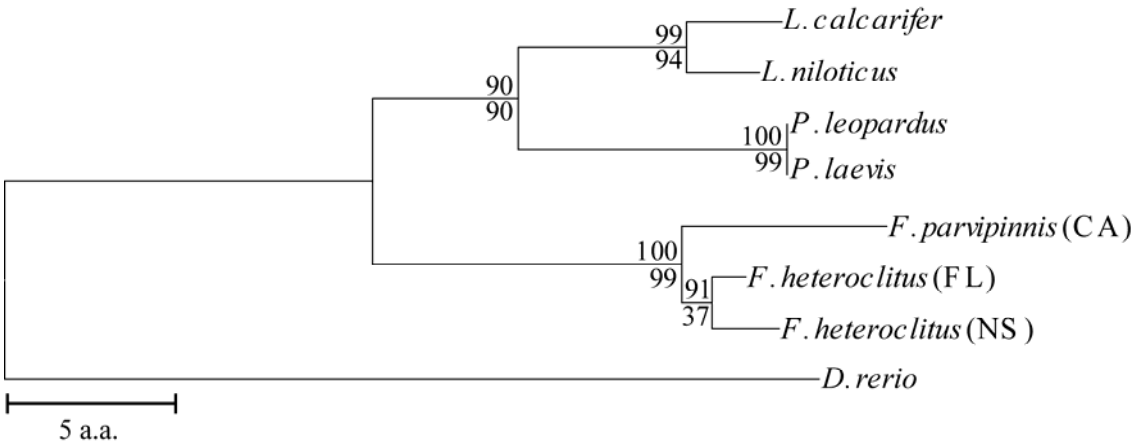


Figure 6.2. Neighbor joining distances based on translated amino acid sequences with complete deletion of all missing residues. Both neighbor joining (NJ) and maximum parsimony (MP; 500 replicates) trees ran 2,000 bootstrap replicates with support values for the consensus trees given above and below phylogram lines, respectively (Kumar *et al.*, 2004). Zebrafish (*Danio rerio*), a distant tropical taxa, is utilized as the rooted out-group due to this being the only other tropical species for which LDH-B sequence is available on GenBank (AF067202).

6.5: DISCUSSION

This chapter presents a comparative assessment of the *ldh-b* locus in tropical and temperate fish to look for evidence of selective constraint on the evolution of this critical metabolic enzyme, which has been shown to be under thermal selection in the model killifish *Fundulus heteroclitus* (Crawford et al., 1999b; Pierce and Crawford, 1997). More specifically, the characterized lactate dehydrogenase–B locus of the *Lates* and *Plectropomus* species pairs (Chapter 3 and Appendix 1, respectively) are compared to the *ldh-b* locus of the temperate congeneric *Fundulus* species pair. This comparative genomics assessment of the *ldh-b* locus for evidence toward selective constraint focuses on coding nucleotide and deduced amino acid sequences, as these are the more conserved regions between phylogenetically distinct taxa as well as being the regions more so constrained by selective forces (Bernardi et al., 1993; Quattro et al., 1993).

Increased exon sequence homology of the *ldh-b* locus across tropical congeneric *Lates* and *Plectropomus* species pair is consistent with the hypothesis that increasing water temperature, and linked decrease in oxygen saturation (Henry's Law), are likely the strongest selective forces acting upon the *ldh-b* locus and encoded LDH-B protein in fish (Powers, 1980; Powers et al., 1979). Furthermore, the high levels of *ldh-b* locus sequence homology among and within congeneric tropical species are also observed when comparing intron sequences within congeneric species pairs; however, intron comparison among congenerics is prevented by a lack of sequence alignment. These high levels of homology in the *ldh-b* locus intron sequences within congeneric species, as well as the presence of known functional motifs within introns of *Lates* and *Plectropomus* species pairs (Chapter 3 and Appendix 1, respectively) suggests that non-coding regions are under a level of selective constraint similar to that of the more historically characterized and considered coding and amino acid sequences (Baek et al.,

2008; Bartel, 2004; Hare and Palumbi, 2003; Mattick, 1994; Selbach et al., 2008). The mechanisms behind exactly how intronic sequences respond to such selective constraint are relatively unknown; therefore, future studies should aim to investigate this in both model and non-model species (see reviews by (Baek et al., 2008; Bartel, 2004; Mattick, 1994; Selbach et al., 2008). For example, microRNA elements (miRNAs) are encoded by small 21 - 24 bp regions within introns and have been shown to be involved the regulation of gene expression by way of binding the 3' untranslated region (UTR) of expressed mRNA transcripts, causing a subsequent "silencing" (i.e. down-regulation) of miRNA-bound transcripts throughout the transcriptome (Ambros, 2004; Baek et al., 2008; Bartel, 2004; Bonnet et al., 2004; Brennecke et al., 2005; Chen et al., 2004; Doench and Sharp, 2004; Mattick, 1994; Selbach et al., 2008). Indeed, the utilization of fish for such studies would be ideal in that a substantial amount of genomic and transcriptomic information is known for an array of potential study species (Cossins and Crawford, 2005).

The observed lower abundance of amino acid substitutions with Grantham's distances (D_G) greater than 100 is explained most parsimoniously by the presence of a strong selective force (e.g. purifying, disruptive or balancing selection) acting upon the translated amino acid sequence of LDH-B (Grantham, 1974; Xia, 2000; Xia and Xie, 2001). Functional necessity restricts nonsynonymous substitutions to similar amino acids ($D_G < 100$) since substitutions between divergent amino acids ($D_G > 100$) would manifest functionally non-equivalent proteins, as described for oxytocin (Xia, 1998). The translated LDH-B amino acid sequence, therefore, appears to be under a level of constraint which presumably restricts evolutionary modifications of this essential glycolytic protein to synonymous amino acid shifts in any randomly occurring, yet selectively advantageous, mutations which optimize the proteome (e.g. catalytic efficiency) or phenome (e.g. swimming performance) in response to the environmental

stresses present in the native habitats of the *Lates* and *Plectropomus* species pairs (e.g. temperature, availability of oxygen). Therefore, the identified differences in the coding nucleotide and amino acid sequences of the *ldh-b* locus within and among these congeneric species pairs are plausibly due to independent evolutionary trajectories following the diversification from a common ancestral *Lates* and *Plectropomus* species, an event which presumably predates the separation of Australia from Africa (approx. 150 m.y. ago *sensu* (Veevers and McElhinny, 1976). These independent trajectories may reflect evidence for selective constraint on the *ldh-b* variants which have persisted through to current populations of these two congeneric species pairs.

The observed reduction in arginine (Arg) abundance in the amino acid sequences of LDH-B of the tropical fishes examined herein, despite the fact that this amino acid is encoded for by six codons, is in line with a previous observation of this unexplained trend (Xia, 1998, 2000). However, the consensus increase in valine (Val) abundance in the LDH-B amino acid sequences inherent across all congeneric species pairs examined is, to our knowledge, an unprecedented anomaly. Valine, being a non-polar amino acid, may be more beneficial than a polar amino acid being that proteins tend to evolve toward possessing the lowest polarity (i.e. highest abundance of non-polar residues) in an effort to the decrease their susceptibility to undesirable allosteric interactions and/or mutations (Xia, 1998, 2000). Alternatively, valine may appear in higher than expected abundance due to its being one of ten "primitive" amino acids (Ala, Asp, Glu, Gly, Ile, Leu, Pro, Ser, Thr and Val) and, therefore, may have accumulated within the amino acid sequence of this ancient and metabolically critical enzyme (Miller, 1986; Xia, 1998).

In addition to nucleotide and amino acid sequence comparisons, those based on isoelectric point (pI) reveal several noteworthy homologies, as pI is a protein characteristic directly linked to habitat pH (see Methods and Results: 3.2.2).

Plectropomus species and *L. niloticus* can avoid unfavorable water temperatures experienced at the surface by moving into deeper water since they live on coral reefs and in large African lakes, respectively, whilst *L. calcarifer* and both *F. heteroclitus* populations live in shallow estuarine habitats (Bernardi et al., 1993; Chenoweth et al., 1998). This habitat difference may explain why *Plectropomus* species possess a pI more homologous to *L. niloticus* and the cold northern *F. heteroclitus* population than do their tropical counterpart *L. calcarifer.* In further support of this, *L. calcarifer* possess a pI more homologous to the warm southern *F. heteroclitus* population than to either *Plectropomus* species or *L. niloticus*, possibly because both these species are forced to deal with warmer annual water temperatures (i.e. behaviorally inescapable) than are either *Plectropomus* species or *L. niloticus* due to their relatively shallow estuarine habitats (Bernardi et al., 1993; Chenoweth et al., 1998). Further isoelectric point comparisons suggest that *L. niloticus* evolved in non-tropical waters due to its high pI homology with the exclusively tropical marine *Plectropomus* species pair and the cold northern *F. heteroclitus* population; whereas in all other comparisons *L. niloticus* follow the tropical species trend of being more homologous to the warm southern *F. heteroclitus* population. The high level of pI divergence observed for *F. parvipinnis*, native to the west coast of North America, suggests an independent evolutionary pathway from its congener *F. heteroclitus*, which is native to the eastern seaboard of North America (reviewed by Powers et al., 1991; Powers and Schulte, 1998). Moreover, the exceedingly basic pI values observed for *F. parvipinnis* may indicate that this species has adapted to a niche with environmental parameters (e.g. pH, salinity) distinctly different from those of all other considered species (Somero, 1996, 2005; Xia, 1998).

6.6: CONCLUSIONS

All organisms achieve the most optimal performance within their respective native habitats (i.e. homeostasis), regardless of whether they inhabit tropical or temperate habitats, by way of selection-driven evolutionary adaptation. Therefore, and in light of the fact that selection has been shown to be acting upon the *ldh-b* locus in the temperate killifish *Fundulus heteroclitus* (Crawford et al., 1999b; Pierce and Crawford, 1997), a similar selective constraint appears to be acting upon the *ldh-b* locus in the tropical *Lates* and tropical *Plectropomus* species. More specifically, the *ldh-b* locus appears to be under selection in both non-model species pairs considered here (*Lates* and *Plectropomus*) based on the high homologies observed for the coding nucleotide and deduced amino acid sequences of these congenerics. Moreover, the observed functional constraint (measured by $D_G < 100$) present between the majority of nonsynonymous amino acid substitutions supports the conclusion that selective constraint appears to be acting upon both the *ldh-b* locus and translated LDH-B protein in the non-model tropical congeneric *Lates* and *Plectropomus* species pairs.

In light of these findings, future studies should aim to characterize the 5' and 3' untranslated regions (UTR) of the *ldh-b* locus in both congeneric *Lates* and *Plectropomus* species pairs being that these regions contain functional elements (e.g. promoters, enhancers) that are likely to be under similar, if not greater, selective constraint than the coding nucleotide and deduced amino acid sequences. For example, the proximal promoter and glucocorticoid response element (GRE) sequences of the *ldh-b* locus 5' UTR in the model temperate *F. heteroclitus* were shown to be under greater selective constraint than other UTR sequences up-stream of the *ldh-b* locus start codon (Crawford et al., 1999b; Crawford and Powers, 1992; Crawford et al., 1999a; Pierce and Crawford, 1997; Rees et al., 2001; Schulte et al., 2000; Schulte et al., 1997;

Segal et al., 1996). Therefore, in addition to the characterization of ldh-b UTR sequences in *Lates* and *Plectropomus*, future studies should aim to quantify hepatic *ldh-b* gene expression in *L. niloticus*, *P. leopardus* and *P. laevis* as this would allow for a comparative transcriptomic assessment within and among these tropical species. Such transcriptomic comparison within and among phylogenetically distinct species, as was done on the expression of glycolytic enzymes in temperate fish (Crawford et al., 1999b; Pierce and Crawford, 1997), would provide further evidence regarding the specific type of selective constraint acting upon the *ldh-b* locus and/or transcript in non-model tropical species.

Taken together, this comparative (analytical) based evidence toward selective constraint acting upon the *ldh-b* coding and deduced LDH-B amino acid sequences provides insight regarding the local adaptations of the tropical *Lates* and *Plectropomus* species. Furthermore, the comparison of these sequences further validates the use of *ldh-b* as a candidate gene for future studies wishing to adopt such a multi-level approach to the question of thermal adaptation in fish as presented here.

Chapter 7

General Discussion

Lates calcarifer from two geographically, genetically and thermally discrete Australian populations exhibited evidence consistent with adaptation to their local thermal environments. Specifically, following acclimation, the tropical northern population (Darwin, Northern Territory) had lower mortality, higher growth (body mass) and better physiological performance (U_{crit}) under heat-stress (35°C) conditions compared to a population originating from a more southern (and cooler water) population (Bowen, central Queensland) (Chapter 2). When a cold-stress (20°C) was applied, inverse results were observed whereby the southern population exhibited lower mortality, higher growth (body mass) and better physiological performance (U_{crit}) (Chapter 2). The fact that seasonal temperature extremes differ between these two locations by 5 - 10°C, as well as their geographic and genetic distinction (Chenoweth et al., 1998; Keenan, 1994), makes these populations candidates for thermal adaptation. Additionally, river flow rates in the native northern and southern environments differ by approximately threefold (Chapter 2), so the swimming abilities of these populations should reflect this difference in flow rates, if populations are adapted to their native environments. If adaptation is evident, the phenome should differ accordingly between these evolutionarily discrete *L. calcarifer* populations. Such adaptive change may originate from genome, transcriptome and/or proteome differences (Cossins and Crawford, 2005; Douglas, 2006; Hochachka and Somero, 2002; Powers and Schulte, 1998; Somero, 2005).

A multi-level approach was used to examine all four levels (phenotypic and molecular) in this study, albeit analytically rather than empirically at the proteomic level, in order to identify at which levels these adaptations are evident. Firstly,

phenotypic differences were examined and confirmed for swimming performance and different growth parameters (Chapter 2). Secondly, a candidate gene approach, targeting *ldh-b* variation at the level of the genotype was examined, but no population specific differences were identified in *ldh-b* genotypes (Chapter 3). These two species exhibited only 2.9% divergence within coding regions (exons), rendering five amino acid differences (i.e. nonsynonymous substitutions) between the deduced amino acid sequences of LDH-B (Chapter 3). This level of conservation was further illustrated by very low levels of sequence variation within and among Australian *L. calcarifer* populations observed following a population genetics pilot screening of the *ldh-b* locus (Chapter 4). More specifically, there were only two parsimoniously informative single nucleotide polymorphisms (SNPs) in the coding nucleotide sequences among the 36 individuals screened (n = 3 to 5 from each of eight discrete populations), representing the two major genetic clades (east coast and Northern Territory) established by mitochondrial control region and ATPase 6 and 8 sequences (Chenoweth et al., 1998). Since both coding region SNPs were synonymous, unique LDH-B allozymes between these discrete populations are unlikely, an expected finding based on the lack of detectable LDH-B allozymes reported by Keenan (1994). These high levels of conservation among discrete *L. calcarifer* populations (Chapter 4), as was also observed between the two congeneric *Lates* spp. (Chapter 3), are potentially indicative of selective constraint on the coding region of *ldh-b* in this species,

Thirdly, despite the absence of *ldh-b* genotype and deduced LDH-B amino acid sequence differences within and among discrete Australian populations (Chapter 4), this thesis presents compelling evidence for stabilizing selection acting on the deduced LDH-B protein sequence of *L. calcarifer* (and other tropical species) (Chapter 6 and Appendix 1). Such evidence for selection was obtained by assessing the level of variation in the *ldh-b* locus when compared between discrete congeneric *Lates* (Chapter

3), *Plectropomus* (Appendix 1) and the model temperate *Fundulus* species (reviewed by Powers and Schulte, 1998) (Chapter 6). In light of such comparisons, substantial evidence toward selective constraint acting upon the *ldh-b* locus in these tropical species is presented, notably a locus which is known to be under thermal selection in cold northern and warm southern *F. heteroclitus* sub-populations (Crawford et al., 1999b; Pierce and Crawford, 1997) (Chapter 6). The selective constraint identified indirectly (via temperature stress) in Chapter 6, which is believed to be acting upon the coding nucleotide (*ldh-b*) and/or deduced amino acid (LDH-B) sequences among tropical *Lates* and *Plectropomus* species, is perhaps the driving force behind the minimal allelic variation and the presence of a single allozyme in Australian *L. calcarifer* populations (Chapter 4). Such population genetic data is also available for the *ldh-b* locus in the *Plectropomus* species and indicates that there are likewise unlikely to be allozyme variants in this tropical coral reef fish species (Hillersøy, G. Honors Thesis, 2008).

Having found no clear inter-population differences in the *ldh-b* genotype among thermally distinct Australian *L. calcarifer* populations in Chapter 4, this population genetics screening of the *ldh-b* locus does not explain the phenotypic differences observed in Chapter 2. Therefore, as a fourth approach I examined the transcriptome for evidence of population-specific differences. Chapter 5 demonstrates that the expression profile of hepatic *ldh-b* differs in magnitude between northern and southern *L. calcarifer* following acclimation and exposure to various ecologically relevant stresses (both 20°C cold-stress and 35°C heat-stress alone, as well as swimming stress both in the presence and absence of thermal stress; see Chapter 2). More specifically, hepatic *ldh-b* transcript abundance increases in response to thermal or swimming stress alone, but not when both stressors are concurrently applied, as follows: 1) *ldh-b* transcript abundance increases significantly in response to swimming stress alone under native culturing temperatures (25°C and 30°C) in the southern and northern populations,

respectively (Chapter 5); 2) *ldh-b* transcript abundance increases significantly in response to cold-stress alone, after acclimation in the southern population exclusively; 3) *ldh-b* transcript abundance increases significantly in response to heat-stress alone, after acclimation in both populations; 4) *ldh-b* transcript abundance does not increase additionally following concurrently applied cold- and swimming stress in either population; 5) *ldh-b* transcript abundance does not increase additionally, but rather collapses, following concurrently applied heat- and swimming stress in both populations. In light of these results it appears as if the southern population possesses a unique transcriptomic response to cold-stress due possibly due to this populations evolutionary exposure to cooler winter water temperatures (Chapter 5). Similarly, the lower mortality and higher critical swimming performance observed for the northern population following acclimation to the heat-stress (35°C) condition (Chapter 2) suggests that this population is more evolutionarily accustomed to a warmer thermal regime than the southern cool water population (Chapter 5). An observed increase in hepatic *ldh-b* transcript abundance in both populations under the increased aerobic metabolism requirement following the subjection of fish to swimming trials (approx. 200min) in the absence of thermal stress, which were designed to impose complete aerobic exhaustion upon challenged fish (Chapter 2) is presumably in direct response to the accumulation of lactate in skeletal muscle. The impact of *ldh-b* transcript abundance on the observed differences in swimming performance (a phenotypic difference) (Chapter 2) is unclear due to the unknown correlation between abundances of ldh-b transcripts (mRNA) and LDH-B protein following exposure to this metabolic (aerobic) stress (see below). However, these findings do illustrate the importance of investigating observed differences in transcript levels as such responses can be pronounced and insightful despite minimal (or no) variation observed in the genotype or deduced amino acid sequence of the candidate gene (Chapters 3, 4 and 6). To disregard

the degree of variation in hepatic *ldh-b* transcript abundance due to a lack of variation in the *ldh-b* genotype is to disregard key evidence for possible acclimatory capability, thermal tolerance or thermal adaptation that may exist among discrete populations of tropical Australian *L. calcarifer* (Chapter 5).

Indeed, adaptations rendering a variation in the transcriptome can be equally, if not more powerful, than mutations which alter the sequence of the expressed protein itself (Carroll, 2005; Douglas, 2006; Li et al., 2004; Selbach et al., 2008). The differences I observed in hepatic *ldh-b* transcript abundance following thermal and/or swimming (aerobic) stress in northern and southern *L. calcarifer* populations (Chapter 5), despite the lack of observed differences in *ldh-b* genotype (Chapters 3 and 4), suggests that genotypic variation may exist in either the upstream and/or downstream untranslated regions (UTRs) of this locus, but these were not characterized in this study. Such a difference was found when a two-fold higher basal expression of hepatic *ldh-b* was linked to a single base pair mutation in the stress-responsive glucocorticoid responsive element (GRE) within the TATA-less 5' proximal promoter (UTR) of *ldh-b* in thermally discrete *F. heteroclitus* populations (Crawford et al., 1999a; Rees et al., 2001; Schulte et al., 2000; Schulte et al., 1997; Segal et al., 1996). An eight-fold higher expression of warm acclimation protein (WAP), which is not known to have unique allozyme variants, was also observed in the warm adapted southern population of *F. heteroclitus* (Picard and Schulte, 2004). Since a similar 2.5 to 3.5-fold increase in hepatic *ldh-b* transcript abundance was observed in tropical Australian *L. calcarifer* populations following 28 day cold- or heat-stress acclimation (Chapter 5), future research should aim to characterize the up- and down-stream regulatory regions of the *ldh-b* locus in this species (see Future Research section below). The characterization of these UTR regions may identify inherent transcriptional regulatory elements and establish if alleles unique to thermally discrete populations are in fact present. If so,

additional studies will be required to establish the direct link between population-specific UTR mutation and levels of hepatic *ldh-b* transcript abundance in Australian *L. calcarifer*. Preliminary attempts were made by the author to obtain the sequence of these regions by inverse PCR (I-PCR) (Hartl and Ochman, 1994) but, unfortunately, the full characterization of *L. calcarifer* UTR sequences was not possible as part of this thesis due to time and resource limitations.

Intervening non-coding regions (introns) were successfully characterized for both *L. calcarifer* and *L. niloticus* (Chapter 3) and subsequently assessed within and among discrete Australian *L. calcarifer* populations (Chapter 4). The historically less characterized non-coding (intron) sequences of the *ldh-b* locus in both *Lates* species are, interestingly, as conserved (97.7%) as the coding regions (97.1%) of this locus despite comprising 72.5 to 80% of the entire gene sequence (Chapter 3). Several simple sequence repeat (SSR) motifs and putative microRNA (miRNA) elements were identified within the introns of the *ldh-b* locus in both congeneric *Lates* species as well as within the introns of congeneric *Plectropomus* species (Chapter 3 and Appendix 1 respectively). SSR motifs and miRNA elements within introns 2 to 6 are also conserved across screened tropical Australian *L. calcarifer* populations (Chapter 4). These SSR motifs and/or miRNA elements are worthy of further investigation for their potential functional importance in the regulation and/or expression of *ldh-b*, or other candidate genes involved in the metabolic (glycolysis) pathway, as such motifs have been directly linked to the increased expression or silencing of other loci (Ambros, 2004; Baek et al., 2008; Brennecke et al., 2005; Li et al., 2004).

To date, the specific mechanisms underlying transcriptomic regulation and their impact on the manifested phenome of individuals, or species, is unclear. However, regulating mechanisms, such as feedback loops in molecular cascades, do appear to play a part in fine-tuning the transcriptome, and thus the proteome, in response to

environmental stimulus (Carroll, 2005; Li et al., 2004; Selbach et al., 2008). Therefore, relatively minor modifications at the level of the transcriptome are likely to be manifested in the phenome of individuals or species (Crawford et al., 1999b; Crawford and Powers, 1992; Gracey et al., 2004; Powers and Schulte, 1998; Rees et al., 2001; Somero, 2005). For example, the observed swimming performance (phenotypic) difference in northern *v.* southern Australian *L. calcarifer* after acclimation to the cold- or heat-stress condition could be due to differences in the transcriptomic response of barramundi from two evolutionarily discrete locales to such thermal stress. More specifically, the change in hepatic *ldh-b* transcript abundance is markedly different in magnitude within these two discrete populations following exposure to thermal and swimming stress alone as well as when these stressors are concurrently applied (Chapter 5).

These population-specific levels of hepatic *ldh-b* transcript abundance, in conjunction with the lack of detectable allozymes in the deduced LDH-B amino acid sequence of wild-caught individuals from the same evolutionarily discrete populations (Chapter 4), suggests that stabilizing selection manifests differences (mutations) in the regulatory machinery located in the 5' and/or 3' UTR of selection-sensitive loci, such as *ldh-b*. In light of this it appears as if the coding region of *ldh-b* might be under stabilizing selection due to functional constraints on the primary (to tertiary) sequence of the protein while the UTR regions may be under directional selection which favors optimal transcriptional elements producing varying concentrations of transcript (and perhaps protein) suited to the natural environment of this tropical species. Such directional selection has been shown for the GRE element in the 5'UTR of *ldh-b* in *F. heteroclitus*, wherein point mutations have been shown to alter the abundance of hepatic *ldh-b* transcripts and enzyme in southern fish acclimated to native (20°C) and non-native (10°C) conditions (Crawford et al., 1999b; Crawford and Powers, 1992;

Crawford et al., 1999a; Pierce and Crawford, 1997; Rees et al., 2001; Schulte et al., 2000; Schulte et al., 1997). Therefore, the up- and down-stream UTR regions of the *ldh-b* locus should be characterized in these Australian *L. calcarifer* populations to identify if a similar point mutation exists and, if so, establish whether or not such a mutation impacts on the basal or stressed level of hepatic *ldh-b* transcript abundance.

It does not immediately appear as if the fold-differences in hepatic *ldh-b* transcript abundance can be directly linked to the variation in critical swimming performance (U_{crit}) in northern *v*. southern *L. calcarifer* populations (Chapter 5). In light of this, further transcriptomic and/or proteomic investigations should examine if there is a correlation between *ldh-b* transcript abundance, LDH-B protein abundance and swimming performance (phenotype) (see Future Research section below for details). In regard to the variation in *ldh-b* transcript abundance (Chapter 5), the observed magnitude of response in southern (~3.5 fold) *v.* northern (~2.5 fold) fish suggests the possible presence of a population-specific difference in the 5' UTR (e.g. proximal promoter) of this locus (see above). Such a mutation could allow these thermally distinct populations to respond accordingly to the seasonal, and/or daily, thermal flux of their respective native habitats. However, the equal translation of *ldh-b* mRNA into LDH-B protein (1:1) needs to be verified first, in order to confirm that an increase or decrease in *ldh-b* transcript does in fact convey an equivalent increase or decrease in active LDH-B protein and thus a more likely impact on the observed differences in swimming performance (phenotype).

Future research

Several hypotheses are raised by the findings of this thesis. In particular: i) Is LDH-B protein abundance correlated with *ldh-b* mRNA abundance in *L. calcarifer* subjected to thermal acclimation and/or swimming stress?; ii) Do the kinetic properties of LDH-B

differ between thermally discrete Australian *L. calcarifer* populations despite the homology of deduced amino acid sequences?; iii) Do the 5' or 3' UTR sequences of *ldh-b* in discrete Australian *L. calcarifer* populations differ and do these regions contain regulatory motifs?; iv) What broad-scale transcriptomic differences exist between these *L. calcarifer* populations following exposure to ecological stressors?; and v) Does the use of multiple candidate genes enhance our molecular-level understanding of organismal differences?. These hypotheses are testable along the following lines:

i) Is LDH-B protein abundance correlated with *ldh-b* mRNA abundance in *L. calcarifer* subjected to thermal acclimation and/or swimming stress?:

The immediate task of establishing the translational efficiency of *ldh-b* mRNA to LDH-B protein in Australian *L. calcarifer* can be achieved by standard protein quantification using *L. calcarifer* LDH-B specific antibodies (via Western blot). Following this, an autocorrelation analysis can be preformed to establish if the relationship between *ldh-b* mRNA and LDH-B protein is positive (i.e. 1:1) for each discrete *L. calcarifer* population. This method was utilized to investigate the relationship between *ldh-b* mRNA and LDH-B protein in thermally discrete *F. heteroclitus* populations (Segal and Crawford, 1994a). These were found to be highly correlated ($r^2 = 0.81$, $p < 0.01$) for both cold northern and warm southern *F. heteroclitus* populations acclimated to 20°C (Segal and Crawford, 1994a). This correlation did not hold true, however, following acclimation to 10°C ($r^2 = 0.22$, $p > 0.25$ and $r^2 = 0.57$, $p = 0.086$ respectively) (Segal and Crawford, 1994a). Being that *L. calcarifer* represents a novel tropical system, the correlation between *ldh-b* mRNA and LDH-B protein at 20, 25, 30 and 35°C cannot be assumed and therefore require direct investigation.

Such investigation can also provide key insight regarding differences in the regulatory processes of *ldh-b* in discrete populations of tropical Australian *L. calcarifer*, as it did for discrete populations of temperate *F. heteroclitus* (Segal and Crawford,

1994a). More specifically, if *ldh-b* mRNA abundance decreases while LDH-B protein abundance remains consistent in thermally discrete *L. calcarifer* populations following acclimation to 20, 25, 30 and 35°C this would indicate a difference in the transcriptional regulation of *ldh-b* in response to temperature acclimation. Moreover, the maintenance of LDH-B protein abundance despite reductions in the abundance of *ldh-b* mRNA would also suggest the presence of a translational and/or post-translational mechanism (Crawford and Powers, 1992; Segal and Crawford, 1994a). Unfortunately, the hepatic tissue used for quantification of *ldh-b* transcript abundance quantification in the present study (Chapter 5) cannot be subsequently used for protein quantification as LDH-B was not retained during the RNA extraction protocol, for technical reasons (see Chapter 3 methods). However, this correlation could be established by quantifying both *ldh-b* transcript and LDH-B protein abundances in snap-frozen cardiac tissue collected from the original experimental animals used in this thesis.

ii) Do the kinetic properties of LDH-B differ between thermally discrete Australian *L. calcarifer* populations despite the homology of deduced amino acid sequences?:

In addition to the quantification of LDH-B protein levels and examination of the relationship to *ldh-b* transcript abundance, the activity of this metabolic enzyme should also be assessed in Australian *L. calcarifer* populations. Key parameters of interest for future studies of the thermal sensitivity of LDH-B from locally adapted barramundi populations are kinetic properties such as the Michaelis–Menten constant (K_m) and catalytic efficiency (k_{cat}). Such differences in K_m and k_{cat} were detected for LDH-A (the anaerobic isozyme of LDH-B, see Figure 1.2) in two congeneric species pairs of goby fishes living at different temperatures (*Gillichthys* and *Coryphopterus* spp.) despite minimal (n = 1) or no difference in deduced LDH-A amino acid sequences (Fields and Somero, 1997). These differences in kinetic properties could arise due to population-specific translational and/or post-translational modifications designed to optimize

functionality under native environmental conditions, as suggested for LDH-A in the above goby species (Fields and Somero, 1997). Since LDH-B and LDH-A are alternative components of the same metabolic pathway and both are expressed in cardiac tissue exclusively, these isozymes are ideal targets for future studies aiming to link transcription levels with protein abundance and metabolic-based phenotypic characteristics such as swimming performance.

iii) Do the 5' or 3' UTR sequences of *ldh-b* in discrete Australian *L. calcarifer* populations differ and do these regions contain regulatory motifs?

Considering the possible existence of different promoter region variants in *ldh-b* from northern and southern *L. calcarifer* populations, this can easily be achieved by the characterization of both up- and down-stream UTR regions (1,000 and 500 bp respectively). Since up-stream regulatory elements are known to be under heightened selective constraint (Crawford et al., 1999b; Rees et al., 2001; Schulte et al., 2000; Schulte et al., 1997), the characterization of these regions may uncover additional evidence for local adaptation of Australian *L. calcarifer* to their native tropical habitats. In addition, this additional characterization would also identify putative SSR and miRNA targeted binding sites in the 5' and 3' UTR respectively. This identification of the 5' and 3' UTR targeted binding sites is required for establishing whether or not SSR motifs and/or miRNA elements embedded within other loci in the metabolic pathway (e.g. *ldh-a*) impact on *ldh-b* expression.

iv) What broad-scale transcriptomic differences exist between these *L. calcarifer* populations following exposure to ecological stressors?:

The convoluted web of transcriptomics is not easily resolved, but other thermally-sensitive loci have been identified in fish. These include *ldh-a* (Somero, 1996), *HMG-1* (Gracey et al., 2004) and *hsp-90* (Fangue et al., 2006). The recent validation of cross-

hybridization of *Pomacentrus moluccensis* cDNA onto a *Danio rerio* microarray chip (Kassahn et al., 2007) permits a broad-scale transcriptomic approach to understanding multi-gene responses to thermal stresses in non-model fish (i.e. those fish for which no oligonucleotide microarray chip exists). This cross-hybridization microarray approach could also be utilized for the identification of additional candidate genes in *L. calcarifer* that respond to thermal and swimming stress alone or in combination. The hepatic cDNA generated for two thermally discrete *L. calcarifer* populations (n = 48 each; see Chapter 5) would be ideal for such a study. The number of loci known to up- or down-regulate in response to ecological stressors such as temperature, hypoxia or swimming (see (Douglas, 2006) provides evidence that such additional candidate genes should be identifiable in tropical Australian *L. calcarifer* by cross hybridizing the hepatic cDNA generated in Chapter 5 onto a *Danio rerio* microarray chip (*sensu* (Kassahn et al., 2007).

v) Does the use of multiple candidate genes enhance our molecular-level understanding of organismal differences?

The broad-scale microarray approach (discussed above) would be secondary to the more simplistic and holistic approach of using additional candidate genes known to function within the same metabolic pathway and, ideally, express within the same tissue as *ldh-b*. Based on these requirements, the *ldh-a* locus would serve as an ideal secondary candidate gene for such future studies.

Linking the abundance of one transcript to a particular phenotypic trait is clearly difficult, if not impossible. However, through the use of multiple candidate genes, e.g. *ldh-b* and *ldh-a,* both of which serve complementary functions in the same metabolic (glycolysis) pathway (see Figure 1.2), it may be possible to identify their combined role at the level of both transcript and protein abundance as it relates to the observed phenotype. Additionally, both LDH-B and LDH-A isozymes have been extensively

characterized by previous studies and shown to be adaptive to different thermal environments (Powers et al., 1979 – 2000 and Somero, et al., 1969 – 2005, respectively). The full sequence characterization of additional candidate genes such as *ldh-a* would be required for *L. calcarifer* and/or other target species if the role of SSR motifs and/or miRNA elements embedded within the intervening introns of these loci is to be examined (as done for *ldh-b* in Chapter 3 and Appendix 1). Once SSR motifs and/or miRNA elements are identified within the introns of identified candidate genes (e.g. *ldh-a*), in conjunction with the UTRs of the same candidate genes, it will be possible to determine if these genes have the capacity to regulate each others expression and/or translation levels. In other words, SSR motifs and miRNA elements embedded within *ldh-b* introns may have targeted binding sites in the 5' or 3' UTR of the *ldh-a* locus (and vice versa) which cause the activation of promoter elements (i.e. concurrent expression) or prevention of ribosomal binding (i.e. linked repression) to these metabolic loci, respectively (Figure 7.1). The role of such intron embedded SSR motifs and miRNA elements are known to impact the transcriptional and/or translational efficiency of other loci (Chen et al., 2004; Selbach et al., 2008; Wittekindt et al., 2000; Yue et al., 2001).

In summary, such hypothesis testing studies could provide valuable scientific insight regarding the transcriptomic and linked proteomic response of tropical fish to changes in their native thermal habitats. Such insight is especially valuable in the face of the temperature increases estimated to occur during this century as a result of global climate change (IPCC 4th Assessment, 2007).

Thesis Conclusions

This thesis presents novel multi-level (organismal and molecular) evidence regarding the response of the Australian *Lates calcarifer* to two ecologically relevant stressors, temperature and aerobic swimming challenge. These findings suggest that *L. calcarifer* is a useful model species for future investigations regarding how the phenome, genome, transcriptome and/or proteome of a tropical fish responds to changes in its thermal habitat (i.e. do differences in the observed phenotype of individuals provide population-specific advantages in an altered thermal habitat? And, if so, are such phenotypic differences observed in the genotype, transcript and/or protein abundance of candidate genes within and among discrete populations?). The enhanced understanding of thermal tolerance in this tropical species could also benefit the Australian aquaculture industry by including this trait as a breeding objective in selection programs. In conclusion, the multi-level approach used in this thesis presents phenotypic (swimming, growth and mortality), hepatic *ldh-b* transcript abundance and deduced LDH-B amino acid sequence evidence which strongly suggests that geographically and genetically distinct Australian *L. calcarifer* populations are adapted to their native tropical habitats. Moreover, this multi-level approach can be applied to a range of biological questions concerned with understanding the processes that underlie phenotypic differences between species and/or populations in response to ecological stressors. More importantly, this approach can be applied to any target species (tropical or temperate) and can use any relevant target candidate genes selected on the basis of the specific ecological stressors under examination. This thesis presents compelling evidence that future studies, through the use of such a multi-level approach, can achieve a more comprehensive understanding of adaptation than can be obtained by focusing on discrete phenomic, genomic, transcriptomic or proteomic approaches.

FIGURE

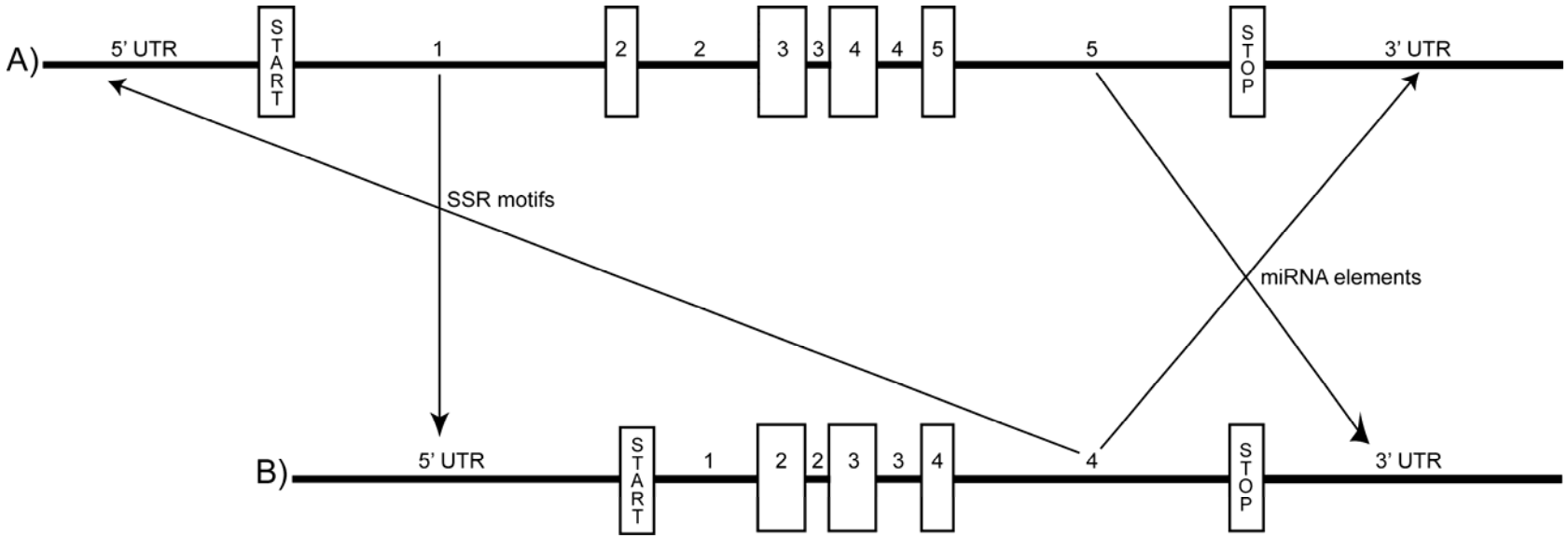


Figure 7.1. Schematic of putative *ldh-b* intronic SSR motifs and miRNA elements binding sites (A) within up- and down-stream untranslated regions of other loci (B).

References

Ambros, V. (2004) The functions of animal microRNAs. Nature, 431: 350-356.

Aseada, T., Vu, T. K. & Manatunge, J. (2005) Effects of flow velocity on feeding behaviour and microhabitat selection of the Stone Moroko *Pseudorasbora parva*: A trade-off between feeding and swimming costs. Trans Am Fish Soc, 134: 537-547.

Baek, D., Villen, J., Shin, C., Camargo, F. D., Gygi, S. P. & Bartel, D. P. (2008) The impact of microRNAs on protein output. Nature, 455: 64-73.

Bartel, D. P. (2004) MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. Cell, 116: 281-298.

Basu, N., Todgham, A. E., Ackerman, P. A., Bibeau, M. R., Nakano, K., Schulte, P. M. & Iwama, G. K. (2002) Heat shock protein genes and their functional significance in fish. Gene, 295: 173–183.

Batty, R. S. & Domenici, P. (2000) Fish swimming behavior. IN: Domenici, P. & Blake, R. W. (Ed.^(Eds.), Biomechanics in animal behaviour, BIOS, Oxford, pp. 237–257.

Beitinger, T. L. & Bennett, W. A. (2000) Quantification of the role of acclimation temperature in temperature tolerance of fishes. Environ Biol Fish, 58: 277–288.

Beitinger, T. L., Bennett, W. A. & Mccauley, R. W. (2000) Temperature tolerances of North American freshwater fishes exposed to dynamic changes in temperature. Environ Biol Fish, 58: 237–275.

Bernardi, G., Sordino, P. & Powers, D. A. (1993) Concordant mitochondrial and nuclear DNA phylogenies for populations of the teleost fish *Fundulus heteroclitus*. Proc Natl Acad Sci USA, 90: 9271-9274.

Billerbeck, J. M., Lankford, T. E. J. & Conover, D. O. (2001) Evolution of intrinsic growth and energy acquisition rates. I. Trade-offs with swimming performance in *Menidia menidia*. Evolution, 55: 1863–1872.

Blake, R. W. (2004) Fish functional design and swimming performance. J Fish Biol, 65: 1193-1222.

Boisclair, D. & Sirois, P. (1993) Testing assumptions of fish bioenergetics models by direct estimation of growth, consumption, and activity rates. Trans Am Fish Soc, 122: 784-796.

Boisclair, D. & Tang, M. (1993) Empirical analysis of the swimming pattern on the net energetic cost of swimming in fishes. J Fish Biol, 42: 169-183.

Bonnet, E., Wuyts, J., Rouzé, P. & Van De Peer, Y. (2004) Evidence that microRNA precursors, unlike other non-coding RNAs, have lower folding free energies than random sequences. Bioinformatics, 20: 2911-2818.

Brennecke, J., Stark, A., Russell, R. B. & Cohen, S. M. (2005) Principles of microRNA–target recognition. PLoS Biol, 3: 0404-0418.

Brett, J. R. (1964) The respiratory metabolism and swimming performance of young sockeye salmon. J Fish Res Board Can, 21: 1183-1226.

Bustin, S. A. (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. Journal of Molecular Endocrinology, 25: 169–193.

Bustin, S. A., Benes, V., Nolan, T. & Pfaffl, M. W. (2005) Quantitative real-time RT-PCR – a perspective. Journal of Molecular Endocrinology, 34: 597–601.

Bustin, S. A. & Nolan, T. (2004) Pitfalls of Quantitative Real-Time Reverse-Transcription Polymerase Chain Reaction. Journal of Biomolecular Techniques, 15: 155-166.

Calladine, C. R., Drew, H. R., Luisi, B. F. & Travers, A. A. (2004) Understanding DNA: the molecule and how it works (Third Edition). Third (Ed.^(Eds.), Elsevier Academic Press San Diego, pp. 334.

Carey, G. R. & Franklin, C. E. (2009) Effect of incubation and rearing temperature on locomotor ability in barramundi, *Lates calcarifer* Bloch, 1790. Marine and Freshwater Research, 60: 203-210.

Carroll, S. B. (2000) Endless Forms: The Evolution of Gene Regulation and Morphological Diversity. Cell, 101: 577-580.
Carroll, S. B. (2005) Evolution at Two Levels: On Genes and Form. PLoS Biology, 3: e245.
Chapman, L. L., Galis, F. & Shinn, J. (2000) Phenotypic plasticity and the possible role of genetic assimilation: Hypoxia-induced trade-offs in the morphological traits of an African cichlid. Ecology Letters, 3: 387-393.
Chen, C.-Z., Li, L., Lodish, H. F. & Bartel, D. P. (2004) MicroRNAs Modulate Hematopoietic Lineage Differentiation. Science, 303: 83-86.
Chenoweth, S. F., Hughes, J. M., Keenan, C. P. & Lavery, S. (1998) When Oceans Meet: A Teleost Shows Secondary Intergradation at an Indian-Pacific Interface. Proc R Soc Lond B Biol Sci, 265: 415-420.
Claireaux, G., Couturier, C. & Groison, A. L. (2006) Effect of temperature on maximum swimming speed and cost of transport in juvenile European sea bass (*Dicentrarchus labrax*). J Exp Biol, 209: 3420-3428.
Condie, S. A. & Dunn, J. R. (2006) Seasonal characteristics of the surface mixed layer in the Australasian region: implications for primary production regimes and biogeography. . Mar Freshwater Res 57: 569-590.
Conover, D. O. & Schultz, E. T. (1995) Phenotypic similarity and the evolutionary significance of countergradient variation. TREE, 10: 248-252.
Cossins, A. R. & Crawford, D. L. (2005) Opinion: Fish as models for environmental genomics. Nat Rev Genet, 6: 324.333.
Crawford, D. L., Pierce, V. A. & Segal, J. A. (1999b) Evolutionary Physiology of Closely Related Taxa: Analyses of Enzyme Expression. Am Zool, 39: 389-400.
Crawford, D. L. & Powers, D. A. (1992) Evolutionary Adaptation to Different Thermal Environments via Transcriptional Regulation. Mol Biol Evol, 9: 806-813.
Crawford, D. L., Segal, J. A. & Barnett, J. L. (1999a) Evolutionary Analysis of TATA-less Proximal Promoter Function. Mol Biol Evol, 16: 194-207.
Currie, S., Tufts, B. L. & Moyes, C. D. (1999) Influence of bioenergetic stress on heat shock protein gene expression in nucleated red blood cells of fish. Am J Physiol Regulatory Integrative Comp Physiol, 276: 990-996.
Darwin, C. R. (1859) The origin of species by means of natural selection. (Ed.^(Eds.), London, pp. 454.
Davis, T. L. O. (1986) Migration patterns in barramundi, *Lates calcarifer* (Bloch), in Van Diemen Gulf, Australia, with estimates of fishing mortality in specific areas. Fisheries Research, 4: 243-258.
Davis, T. L. O. (1987) Biology of Wild Stock *Lates calcarifer* in Northern Australia. IN: Copeland, J. W. & Grey, D. L. (Ed.^(Eds.), The Management of Wild and Cultured Sea Bass/Barramundi *Lates calcarifer,* ACIAR Proceedings, pp. 22-29.
De Santis, C., Evans, B., Smith-Keune, C. & Jerry, D. (2008) Molecular characterization, tissue expression and sequence variability of the barramundi (*Lates calcarifer*) myostatin gene. BMC Genomics, 9: 82.
Dimichele, L. & Powers, D. A. (1982) Physiological Basis for Swimming Endurance Differences between LDH-B Genotypes of *Fundulus heteroclitus*. Science, 216: 1014-1016.
Doench, J. G. & Sharp, P. A. (2004) Specificity of microRNA target selection in translational repression. Genes & Dev, 18: 504-511.
Douglas, S., E. (2006) Microarray Studies of Gene Expression in Fish. OMICS A Journal of Integrative Biology, 10: 474-489.
Doupé, R. G., Horwitz, P. & Lymbery, A. J. (1999) Mitochondrial genealogy of Western Australian barramundi: applications of inbreeding coefficients and coalescent analysis for separating temporal population processes. J Fish Biol, 54: 1197-1209.
Duerr, R. H., Taylor, K. D., Brant, S. R., Rioux, J. D., Silverberg, M. S., Daly, M. J., Steinhart, A. H., Abraham, C., Regueiro, M., Griffiths, A., Dassopoulos, T., Bitton, A., Yang, H., Targan, S., Datta, L. W., Kistner, E. O., Schumm, L. P., Lee, A. T., Gregersen, P. K., Barmada, M. M., Rotter, J. I., Nicolae, D. L. & Cho, J. H. (2006)

A Genome-Wide Association Study Identifies *IL23R* as an Inflammatory Bowel Disease Gene. Science, 314: 1461-1463.
Enright, A. J., John, B., Gaul, U., Tuschl, T., Sander, C. & Marks, D. S. (2003) MicroRNA targets in *Drosophila*. Genome Biol, 5: R1-14.
Fangue, N. A. & Bennett, W. A. (2003) Thermal Tolerance Responses of Laboratory-Acclimated and Seasonally Acclimatized Atlantic Stingray, *Dasyatis sabina*. Copeia, 2003: 315–325.
Fangue, N. A., Hofmeister, M. & Schulte, P. M. (2006) Intraspecific variation in thermal tolerance and heat shock protein gene expression in common killifish, *Fundulus heteroclitus*. J Exp Biol, 209: 2859-2872.
Fangue, N. A., Mandic, M., Richards, J. G. & Schulte, P. M. (2008) Swimming Performance and Energetics as a Function of Temperature in Killifish *Fundulus heteroclitus*. Physiol Biochem Zool, 81: 389–401.
Feldmeth, C. R. & Jenkins, T. M. (1973) An estimate of energy expenditure by rainbow trout (*Salmo gairdneri*) in a small mountain stream. J Fish Res Board Can, 30: 1755-1759.
Fields, P. A. & Somero, G. N. (1997) Amino Acid Sequence Differences Cannot Fully Explain Interspecific Variation In Thermal Sensitivities Of Gobiid Fish A_4-Lactate Dehydrogenases (A_4-LDHs). The Journal of Experimental Biology, 200: 1839-1850.
Fields, P. A. & Somero, G. N. (1998) Hot spots in cold adaptation: Localized increases in conformational flexability in lactate dehydrogenase A_4 orthologs of Antarctic notothenioid fishes. Proc Natl Acad Sci USA, 95: 11476-11481.
Fisher, R., Bellwood, D. R. & Job, S. D. (2000) Development of swimming abilities in reef fish larvae. Mar Ecol Prog Ser, 202: 163-173.
Fitzgibbon, Q. P., Strawbridge, A. & Seymour, R. S. (2007) Metabolic scope, swimming performance and the effects of hypoxia in the mulloway, *Argyrosomus japonicus* (Pisces: Sciaenidae). Aquaculture, 270: 358–368.
Frank-Kamenetskii, M. D. (1997) Unraveling DNA: The Most Important Molecule of Life. Second (Ed.^(Eds.), Perseus Books, Reading, Massachusetts, pp. 214.
Froese, R. & Pauly, D. (2008) FishBase. World Wide Web electronic publication.
Fulton, C. J. (2007) Swimming speed performance in coral reef fishes: field validations reveal distinct functional groups. Coral Reefs, 26: 217-228.
Fulton, C. J. (in press) The role of swimming in reef fish ecology. IN: Domenici, P. & Kapoor, B. G. (Ed.^(Eds.), Fish swimming: an etho-ecological perspective., Chapter 12, Science Publishers, Enfield.
Fulton, C. J., Bellwood, D. R. & Wainwright, P. C. (2005) Wave energy and swimming performance shape coral reef fish assemblages. Proc R Soc Lond B Biol Sci, 272: 827–832.
Gieskes, J. M. (1969) Effect of Temperature on the pH of Seawater. Limnology and Oceanography, 14: 679-685.
Gracey, A. Y., Fraser, E. J., Li, W., Fang, Y., Taylor, R. R., Rogers, J., Brass, A. & Cossins, A. R. (2004) Coping with cold: An integrative, multitissue analysis of the transcriptome of a poikilothermic vertebrate. Proc Nati Acad Sci USA, 101: 16970–16975.
Grantham, R. (1974) Amino Acid Difference Formula to Help Explain Protein Evolution. Science, 185: 3.
Griffin, R. K. (1987a) Life history, distribution, and seasonal migration of barramundi in the Daly River, Northern Territory, Australia. IN: Dadswell, M. J., Klauda, R. J., Moffitt, C. M. & Saunders, R. L. (Ed.^(Eds.), Common strategies of anadromous and catadromous fishes, American Fisheries Society, Symposium 1, Bethesda, Maryland.
Griffin, R. K. (1987b) Life history, distribution, and seasonal migration of barramundi in the Daly River, Northern Territory, Australia. IN: M. J. Dadswell, R. J. Klauda, C.

M. Moffitt & Saunders, R. L. (Ed.^(Eds.), Common strategies of anadromous and catadromous fishes, American Fisheries Society, Bethesda, Maryland.
Haddrill, P. R., Charlesworth, B., Halligan, D. L. & Andolfatto, P. (2005) Patterns of intron sequence evolution in *Drosophila* are dependent upon length and GC content. Genome Biol, 6: R67.
Hall, T. A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser, 41: 95-98.
Hare, M. P. & Palumbi, S. R. (2003) High Intron Sequence Conservation Across Three Mammalian Orders Suggests Functional Constraints. Mol Biol Evol, 20: 969-978.
Harriott, V. J. & Banks, S. A. (2002) Latitudinal variation in coral communities in eastern Australia: a qualitative biophysical model of factors regulating coral reefs. Coral Reefs, 21: 83-94.
Hartl, D. L. & Ochman, H. (1994) Inverse Polymerase Chain Reaction. Methods in Molecular Biology, 31: Protocols for Gene Analysis: 9.
Heap, S. P. & Goldspink, G. (1986) Alterations to the swimming performance of carp, *Cyprinus carpio*, as a result of temperature acclimation. J Fish Biol, 29: 747-753.
Herbert, N. A. & Steffensen, J. F. (2005) The response of Atlantic cod, *Gadus morhua*, to progressive hypoxia: fish swimming speed and physiological stress. Mar Biol, 147: 1403–1412.
Higgins, D. G., Thompson, J. D. & Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl Acids Res, 22: 4673-4680.
Hinch, S. G. & Rand, P. S. (1998) Swim speeds and energy use of upriver-migrating sockeye salmon (*Oncorhynchus nerka*) — role of local environment and fish characteristics. Can J Fish Aquat Sci, 55: 1821-1831.
Hirschhorn, J. N. & Daly, M. J. (2005) Genome-wide association studies for common diseases and complex traits. Nature Reviews: Genetics, 6: 95-108.
Hobert, O. (2004) Common logic of transcription factor and microRNA action. TRENDS in Biochemical Sciences, 29: 7.
Hochachka, P. W. & Somero, G. N. (2002) Biochemical Adaptation: Mechanism and Process in Physiological Evolution. (Ed.^(Eds.), Oxford University Press, Oxford, pp. 466.
Hokanson, K. E. F., Kleiner, C. F. & Thorslund, T. W. (1977) Effect of Constant Temperatures and Diel Temperature Fluctuations on Specific Growth and Mortality Rates and Yield of Juvenile Rainbow Trout, *Salmo gairdneri*. J Fish Res Board Can, 34: 639-648.
Huey, R. B. & Stevenson, R. D. (1979) Integrating thermal physiology and ecology of ectotherms: a discussion of approaches. Am Zool, 19: 357-366.
Hwang, J. & Wong, C. K. (2005) The China coastal current as a driving for transporting Calanus sinicus (Copepoda: Calanoida) from its population senters to waters off Taiwan and Hong Kong during the winter northeast monsoon period. J Plankton Res, 27: 205-210.
Jain, K. E. & Farrell, A. P. (2003) Influence of seasonal temperature on the repeat swimming performance of rainbow trout *Oncorhynchus mykiss*. J Exp Biol, 206: 3569-3579.
Johns, G. C. & Somero, G. N. (2004) Evolutionary Convergence in Adaptation of Proteins to Temperature: A_4-Lactate Dehydrogenases of Pacific Damselfishes (*Chromis* spp.) Mol Biol Evol, 21: 314-320.
Johnson, M. S. (1971) Adaptive lactate dehydrogenase variation in the crested blenny, *Anoplarchus*. Heredity, 27: 205-226.
Johnson, M. S. (1977) Association of allozymes and temperature in the crested blenny *Anoplarchus purpurescens*. Mar Biol, 41: 147-152.
Jokiel, P. L. & Coles, S. L. (1990) Response of Hawaiian and other Indo-Pacific reef corals to elevated temperature. Coral Reefs, 8: 155-162.

Jonassen, T. M., Imsland, A. K. & Stefansson, S. O. (1999) The interaction of temperature and fish size on growth of juvenile halibut. J Fish Biol, 54: 556–572.
Kao, Y.-H. J. & Farley, T. M. (1978) Purification and properties of allelic lactate dehydrogenase isozymes at the B^2 locus in rainbow trout, *Salmo gairdneri*. Comp Biochem & Physio Pt B, 61: 507-512.
Kassahn, K. S., Caley, M. J., Ward, A. C., Connolly, A. R., Stone, G. & Crozier, R. H. (2007) Heterologous microarray experiments used to identify the early gene response to heat stress in a coral reef fish. Mol Ecol, 16: 1749-1763.
Katersky, R. S. & Carter, C. G. (2005) Growth efficiency of juvenile barramundi, *Lates calcarifer*, at high temperatures. Aquaculture, 250: 775– 780.
Katersky, R. S. & Carter, C. G. (2007) High growth efficiency occurs over a wide temperature range for juvenile barramundi *Lates calcarifer* fed a balanced diet. Aquaculture, 272: 444–450.
Keenan, C. P. (1994) Recent Evolution of Population Structure in Australian barramundi, *Lates calcarifer* (Bloch): An Example of Isolation by Distance in One Dimension. Aust J Mar Freshwat Res, 45: 1123-1148.
Keenan, C. P. (2000) Should we allow human-induced migration of the Indo-West Pacific fish, barramundi *Lates calcarifer* (Bloch) within Australia? Aquaculture Research, 31: 121-131.
Kingsolver, J. G. & Huey, R. B. (2008) Size, temperature, and fitness: three rules. Evol Ecol Res, 10: 251–268.
Kirpichnikov, V. S., Muske, G. A., Scholl-Engberts, A. D., Chernov, V. M. & Borchsenius, S. N. (1990) Genetic structure and allele frequency dynamics in the sockeye salmon population of Lake Dalneye, Kamchatka. Aquaculture, 84: 13-25.
Klar, G. T., Stalnaker, G. B. & Farley, T. M. (1979) Comparative blood lactate response to low oxygen concentrations in rainbow trout, *Salmo gairdneri*, LDH-B2 phenotypes. Comp Biochem & Physio Pt A, 63: 237-240.
Kodric-Brown, A. & Nicoletto, P. F. (2005) Courtship behavior, swimming performance, and microhabitat use of Trinidadian guppies. Environ Biol Fish, 73: 299-307.
Kumar, S., Tamura, K. & Nei, M. (2004) MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Briefings in Bioinformatics, 5: 150-163.
Lai, E. C. (2003) microRNAs: Runts of the Genome Assert Themselves. Curr Biol, 13: R925-R936
Lambowitz, A. M. & Belfort, M. (1993) Introns as mobile genetic elements. Annu Rev Biochem, 62: 587-622.
Langerhans, R. B., Layman, C. A., Shokrollahi, A. M. & Dewitt, T. J. (2004) Predator-driven phenotypic diversification in *Gambusia affinis*. Evolution, 58: 2305-2318.
Latchman, D. (2002) Gene Regulation: A eukaryotic perspective. 4th Edition (Ed.^(Eds.), Nelson Thornes Ltd., Cheltenham, pp. 323.
Latchman, D. S. (1991) Eukaryotic Transcription Factors. (Ed.^(Eds.), Academic Press Limited, London, pp. 270.
Li, S. S., Fitch, W. M., Pan, Y.-C. P. & Sharief, F. S. (1983) Evolutionary Relationships of Vertebrate Lactate Dehydrogenase Isozymes A_4 (Muscle), B_4, (Heart), and C_4, (Testis). J Biol Chem, 258: 7029-7032.
Li, Y.-C., Korol, A. B., Fahima, T. & Nevo, E. (2004) Microsatellites Within Genes: Structure, Function, and Evolution. Mol Biol Evol, 21: 991-1008.
Lopez, R. & Lloyd, A. (1997) ClustalW WWW Service at the European Bioinformatics Institute (http://www.ebi.ac.uk/clustalw): On-line sequence alignment.
Lough, J. M. (1998) Coastal climate of northwest Australia and comparisons with the Great Barrier Reef: 1960 to 1992. Coral Reefs 17: 351-367.
Masuda, R. (2006) Ontogeny of anti-predator behavior in hatchery-reared jack mackerel *Trachurus japonicus* larvae and juveniles: patchiness formation, swimming capability, and interaction with jellyfish. Fish Sci, 72: 1225-1235.
Mattick, J. S. (1994) Introns: evolution and function. Curr Opin Genet Dev, 4: 823-831.

Meeuwig, M. H., Dunham, J. B., Hayes, J. P. & Vinyard, G. L. (2004) Effects of constant and cyclical thermal regimes on growth and feeding of juvenile cutthroat trout of variable sizes. Ecology of Freshwater Fish, 13: 208–216.

Meloni, R., Albanèse, V., Ravassard, P., Treilhou, F. & Mallet, J. (1998) A tetranucleotide polymorphic microsatellite, located in the first intron of the tyrosine hydroxylase gene, acts as a transcription regulatory element *in vitro*. Hum Mol Genet, 7: 423-428.

Merritt, R. B. (1972) Geographical distribution and enzymatic properties of lactate dehydrogenase allozymes in the fat-head minnow, *Pimephales promelas*. American Naturalist, 106: 173-184.

Miller, S. L. (1986) Current status of the prebiotic synthesis of small molecules. Molecular evolution of life, Cambridge University Press, Cambridge: 5-11.

Nakaya, H. I., Amaral, P. P., Louro, R., Lopes, A., Fachel, A. A., Moreira, Y. B., El-Jundi, T. A., Da Silva, A. M., Reis, E. M. & Verjovski-Almeida, S. (2007) Genome mapping and expression analyses of human intronic noncoding RNAs reveal tissue-specific patterns and enrichment in genes related to regulation of transcription. Genome Biol, 8: R43.

Nelson, J. A., Gotwalt, P. S. & Snodgrass, J. W. (2003) Swimming performance of blacknose dace (*Rhinichthys atratulus*) mirrors home-stream current velocity. Can J Fish Aquat Sci, 60: 301-308.

Nilsson, G. E., Crawley, N., Lunde, I. G. & Munday, P. L. (2009) Elevated temperature reduces the respiratory scope of coral reef fishes. Global Change Biol, 1: 1-8.

Oleksiak, M. F., Churchill, G. A. & Crawford, D. L. (2002) Variation in gene expression within and among natural populations. Nat Genet, 32: 261-266.

Perry, A., L., Low, P. J., Ellis, J. R. & Reynolds, J. D. (2005) Climate Change and Distribution Shifts in Marine Fishes. Science, 308: 1912-1915.

Picard, D. J. & Schulte, P. M. (2004) Variation in gene expression in response to stress in two populations of *Fundulus heteroclitus*. Comp Biochem Physiol Physiol, 137: 205-216.

Pierce, V. A. & Crawford, D. L. (1997) Phylogenetic Analysis of Glycolytic Enzyme Expression. Science, 276: 256-259.

Place, A. R. & Powers, D. A. (1984a) Purification and Characterization of the Lactate Dehydrogenase (LDH-B_4) Allozymes of *Fundulus heteroclitus*. J Biol Chem, 259: 1299-1308.

Place, A. R. & Powers, D. A. (1984b) Kinetic Characterization of the Lactate Dehydrogenase (LDH-B_4) Allozymes of *Fundulus heteroclitus*. J Biol Chem, 259: 1309-1318.

Plaut, I. (2001) Critical swimming speed: its ecological relevance. Comp Biochem Physiol Physiol, 131: 41-50.

Podrabsky, J. E. & Somero, G. N. (2004) Changes in gene expression associated with acclimation to constant temperatures and fluctuating daily temperatures in an annual killifish *Austrofundulus limnaeus*. J Exp Biol, 207: 2237-2254.

Powers, D. A. (1980) Molecular Ecology of Teleost Fish Hemoglobins: Strategies for Adapting to Changing Environments. Am Zool, 20: 139-163.

Powers, D. A., Greaney, G. S. & Place, A. R. (1979) Physiological correlation between lactate dehydrogenase genotype and haemoglobin function in killifish. Nature, 277: 240-241.

Powers, D. A., Lauerman, T., Crawford, D. & Dimichele, L. (1991) Genetic Mechanisms for Adapting to a Changing Environment. Annu Rev Genet, 25: 629-659.

Powers, D. A. & Schulte, P. M. (1998) Evolutionary Adaptations of Gene Structure and Expression in Natural Populations in Relation to a Changing Environment: A Multidisciplinary Approach to Address the Million-Year Saga of a Small Fish. J Exp Zool Comp Exp Biol, 282: 71-94.

Quattro, J. M., Woods, H. A. & Powers, D. A. (1993) Sequence analysis of teleost retina-specific lactate dehydrogenase C: Evolutionary implications for the vertebrate lactate dehydrogenase gene family. Proc Natl Acad Sci USA, 90: 242-246.

Rees, B. B., Bowman, J. A. L. & Schulte, P. M. (2001) Structure and Sequence Conservation of a Putative Hypoxia Response Element in the Lactate Dehydrogenase-B Gene of *Fundulus*. Biol Bull, 200: 247-251.

Rehmsmeier, M., Steffen, P. & Höchsmann, M. (2004) Fast and effective prediction of microRNA/target duplexes. RNA, 10: 12.

Reisman, D., Greenberg, M. & Rotter, V. (1988) Human p53 oncogene contains one promoter upstream of exon 1 and a second, stronger promoter within intron 1. Proc Nati Acad Sci USA, 85: 5146-5150.

Ribbink, A. J. (1987) African lakes and their fishes: conservation and suggestions. Environ Biol Fish, 19: 3-26.

Rimmer, M. (1995) Barramundi farming: an introduction. (Ed.^(Eds.), The State of Queensland, Department of Primary Industries, Brisbane.

Roger, A. J. & Doolittle, W. F. (1995) Why Introns-in-pieces? Nature, 364: 2.

Russell, D. J. & Garrett, R. N. (1985) Early life history of barramundi, *Lates calcarifer* (Bloch), in northeastern Queensland. Aust J Mar Freshwat Res, 36: 191-201.

Salini, J. P. & Shaklee, J. B. (1988) Genetic structure of barramundi (*Lates calcarifer*) stocks from Northern Australia. Aust J Mar Freshwat Res, 39: 317-329.

Sambrook, J. (2002) Molecular cloning: A Laboratory Manual Third (Ed.^(Eds.), CSHL Press, Woodbury, NY USA, pp. 600.

Sambrook, J., Fritsch, E. & Maniatus, T. (1989) Molecular cloning: a laboratory manual. 2nd (Ed.^(Eds.), Cold Spring Harbour Laboratory Press, New York, USA.

Schneider, S., Roessli, D. & Excoffier, L. (2000) Arlequin: A software for population genetics analysis. Ver. 2.000. Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva.

Schulte, P. M., Glemet, H. C., Fiebig, A. A. & Powers, D. A. (2000) Adaptive variation in lactate dehydrogenase-B gene expression: Role of a stress-responsive regulatory element. Proc Nati Acad Sci USA, 97: 6597-6602.

Schulte, P. M., Gomez-Chiarri, M. & Powers, D. A. (1997) Structural and Functional Differences in the Promoter and 5' Flanking Region of *Ldh-B* Within and Between Populations of the Teleost *Fundulus heteroclitus*. Genetics, 145: 759-769.

Scott, L. J., Mohlke, K. L., Bonnycastle, L. L., Willer, C. J., Li, Y., Duren, W. L., Erdos, M. R., Stringham, H. M., Chines, P. S., Jackson, A. U., Prokunina-Olsson, L., Ding, C.-J., Swift, A. J., Narisu, N., Hu, T., Pruim, R., Xiao, R., Li, X.-Y., Conneely, K. N., Riebow, N. L., Sprau, A. G., Tong, M., White, P. P., Hetrick, K. N., Barnhart, M. W., Bark, C. W., Goldstein, J. L., Watkins, L., Xiang, F., Saramies, J., Buchanan, T. A., Watanabe, R. M., Valle, T. T., Kinnunen, L., Abecasis, G. R., Pugh, E. W., Doheny, K. F., Bergman, R. N., Tuomilehto, J., Collins, F. S. & Boehnke, M. (2007) A Genome-Wide Association Study of Type 2 Diabetes in Finns Detects Multiple Susceptibility Variants. Science, 316: 1341-1345.

Segal, J. A. & Crawford, D. L. (1994a) LDH-B enzyme expression: the mechanisms of altered gene expression in acclimation and evolutionary adaptation. Am J Physiol Regulatory Integrative Comp Physiol, 267: 1150-1153.

Segal, J. A. & Crawford, D. L. (1994b) LDH-B enzyme expression: the mechanisms of altered gene expression in acclimation and evolutionary adaptation. Am Physio Soc: 1150-1153.

Segal, J. A., Schulte, P. M., Powers, D. A. & Crawford, D. L. (1996) Descriptive and Functional Characterization of Variation in the *Fundulus heteroclitus Ldh-B* Proximal Promoter. J Exp Zool Comp Exp Biol, 275: 355-364.

Selbach, M., Schwanhausser, B., Thierfelder, N., Fang, Z., Khanin, R. & Rajewsky, N. (2008) Widespread changes in protein synthesis induced by microRNAs. Nature, 455: 58-63.

Shaklee, J. B. & Salini, J. P. (1985) Genetic variation and population subdivision in Australian barramundi *Lates calcarifer* (Bloch). Aust J Mar Freshwat Res, 36: 203-218. http://www.publish.csiro.au/nid/126/paper/MF9850203.htm.

Sladek, R., Rocheleau, G., Rung, J., Dina, C., Shen, L., Serre, D., Boutin, P., Vincent, D., Belisle, A., Hadjadj, S., Balkau, B., Heude, B., Charpentier, G., Hudson, T. J., Montpetit, A., Pshezhetsky, A. V., Prentki, M., Posner, B. I., Balding, D. J., Meyre, D., Polychronakos, C. & Froguel, P. (2007) A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature, 445: 881-885.

Somero, G. N. (1995) Proteins and temperature. A Rev Physiol, 57: 43-68.

Somero, G. N. (1996) Temperature and Proteins: Little Things Can Mean a Lot. News Physiol Sci, 11: 72-77.

Somero, G. N. (2003) Protein adaptations to temperature and pressure: complementary roles of adaptive changes in amino acid sequence and internal milieu. Comp Biochem Physiol B Biochem Mol Biol, 136: 577–591.

Somero, G. N. (2005) Linking biogeography to physiology: Evolutionary and acclimatory adjustments of thermal limits. Frontiers in Zoology, 2: 9.

Steffensen, J. F. (2002) Metabolic cold adaptation of polar fish based on measurements of aerobic oxygen consumption: fact or artefact? Artefact! Comp Biochem Physiol Physiol, 132: 789-795.

Teranishi, K. S. & Stillman, J. H. (2006) A cDNA microarray analysis of the response to heat stress in hepatopancreas tissue of the porcelain crab *Petrolisthes cinctipes*. Comparative Biochemistry and Physiology, Part D, 2: 53–62.

The Uniprot, C. (2008) The Universal Protein Resource (UniProt). Nucl Acids Res, 36: D190-195.

Tsuji, S., Qureshi, M. A., Hou, E. W., Fitch, W. M. & Li, S. S.-L. (1994) Evolutionary relationships of lactate dehydrogenases (LDHs) from mammals, birds, an amphibian, fish, barley, and bacteria: LDH cDNA sequences from *Xenopus*, pig, and rat. Proc Natl Acad Sci USA, 91: 9392-9396.

Van Herwerden, L., Davies, C. R. & Choat, J. H. (2002) Phylogenetic and evolutionary perspectives of the Indo-Pacific grouper Plectropomus species on the Great Barrier Reef, Australia. J Fish Biol, 60: 1591-1596.

Veevers, J. J. & Mceihinny, M. W. (1976) The Separation of Australia from Other Continents. Earth-Science Reviews, 12: 139-159.

Virani, N. A. & Rees, B. B. (2000) Oxygen consumption, blood lactate and inter-individual variation in the gulf killifish, *Fundulus grandis*, during hypoxia and recovery. Comp Biochem Physiol Physiol, 126: 397–405.

Wang, W. Y. S., Barratt, B. J., Clayton, D. G. & Todd, J. A. (2005) Genome-wide association studies: theoretical and practical concerns. Nat Rev Genet, 6: 109-118.

Ward, R. D., Holmes, B. H. & Yearsley, G. K. (2008) DNA barcoding reveals a likely second species of Asian sea bass (barramundi) (*Lates calcarifer*). J Fish Biol, 72: 458–463.

Webb, P. W. (1994) The biology of fish swimming. IN: Maddock, L., Bone, Q. & Rayner, J. M. V. (Ed.^(Eds.), Mechanics and physiology of animal swimming, Cambridge University Press, Cambridge, pp. 45-62.

Wittekindt, N. E., Hortnagel, K., Geltinger, C. & Polack, A. (2000) Activation of the *c-myc* promoter P1 by immunoglobulin κ gene enhancers in Burkitt lymphoma: functional characterization of the intron enhancer motifs κB, E box 1 and E box 2, and of the 3' enhancer motif PU. Nucleic Acids Res, 28: 800-808.

Xia, X. (1998) What Amino Acid Properties Affect Protein Evolution? J Mol Evol, 47: 557-564.

Xia, X. (2000) Data Analysis in Molecular Biology and Evolution (D.A.M.B.E.). (Ed.^(Eds.), Kluwer Academic Publishers, Dordrecht, pp. 277.

Xia, X. & Xie, Z. (2001) DAMBE: Software Package for Data Analysis in Molecular Biology and Evolution Journal of Heredity, 92: 371-373.

Yue, G., Li, Y. & Orban, L. (2001) Characterization of Microsatellites in the IGF-2 and GH Genes of Asian Seabass (*Lates calcarifer*). Mar Biotechnol, 3: 1-3.

Zhang, J. & Madden, T. L. (1997) PowerBLAST: A new network BLAST application for interactive or automated sequence analysis and annotation. Genome Res, 7: 649-656.

Zhang, Z., Schwartz, S., Wagner, L. & Miller, W. (2000) A greedy algorithm for aligning DNA sequences. J Comput Biol, 7: 203-214.

Zhu, Z. Y., Lin, G., Lo, L. C., Xu, Y. X., Feng, F., Chou, R. & Yue, G. H. (2006) Genetic analyses of Asian seabass stocks using novel polymorphic microsatellites. Aquaculture, 256: 167–173.

Appendix 1

Classic approach revitalizes genomics: complete characterization of a candidate gene for thermal adaptation in two coral reef fishes

ABSTRACT

Lactate dehydrogenase-B (*ldh-b*) encodes a metabolic enzyme (LDH-B) which plays an important role in maintaining aerobic performance and in thermal acclimation and/or adaptation of fish. As the first step in understanding the effect this enzyme has on the ability of tropical coral reef fishes to cope with thermal stress, we characterized both coding and non-coding regions of *ldh-b* in two congeneric perciformes, *Plectropomus leopardus* and *Plectropomus laevis. Ldh-b* was 4,666 and 4,539 bp in length in *P. leopardus* and *P. laevis*, respectively, with coding regions comprising 1,005 bp in both species. We report a high level of sequence homology between the coding regions of *ldh-b* in these two species, with 98.1% identity of nucleotides corresponding to 100% amino-acid identity between the deduced protein sequences. Comparison between non-coding (intron) regions of both species revealed the presence of several indels, despite the high level of homology observed (95.9% identity of intron nucleotides). Potential regulatory motifs and elements, including twenty-six simple sequence repeat motifs (mono-, di-, tri- and tetranucleotide) and twenty-three putative microRNA elements are identified within the introns of both species, further supporting recent demonstrations that such short motifs and elements exhibit widespread positioning throughout non-coding regions of the genome. This characterization of *ldh-b* in these two coral reef fishes allows for a wide range of future studies (e.g. analytical comparisons of *ldh-b* and LDH-B among different fish genera from different thermal environments and habitats).

1. INTRODUCTION

The candidate gene approach is particularly useful in studies aiming to understand adaptive phenotypes at the molecular level, particularly in the genomics era, as it comprehensively links phenotypic variation to a particular constituent of the genome, transcriptome and proteome [1-3]. This approach has been utilized in studies of temperate fishes (e.g. lactate dehydrogenase-B (*ldh-b*) in killifish [4, 5] as well as ectodysplasin (*EDA*) and pantophysin (*Pan I*) in stickleback and Atlantic cod, respectively [6]). However, the use of such a targeted approach in non-model tropical fish is lacking (but see [7] for use of the multi-level candidate gene approach in the Australian *Lates calcarifer*). As such, this study is the first to present a comparative characterization of the *ldh-b* locus in two congeneric coral reef fish, *Plectropomus leopardus* and *P. laevis*, from the ecologically and commercially important group Epinephelinae.

Lactate dehydrogenase-B (*ldh-b*) encodes a metabolic enzyme (LDH-B), which plays an important role in maintaining aerobic performance of fish by converting the major by-product of anaerobic glycolysis (lactate) to either pyruvate or glucose via oxidation or gluconeogenesis, respectively (see reviews by [4, 5]). Accumulating lactate in aerobic tissues (e.g. heart, skeletal muscle) is transported via the blood to the liver, where this conversion occurs. An enhanced ability to convert accumulating lactate back to pyruvate may therefore allow aerobic metabolic activity to be sustained for an extended period of time, as was proposed for the population-specific differences in swimming performance of thermally discrete populations of the temperate killifish *Fundulus heteroclitus* subjected to ecologically relevant thermal stresses [8]. In support of this conclusion is the recent demonstration that accumulating serum lactate is the primary biochemical factor responsible for the diminished repeat swimming performance of rainbow trout *Oncorhynchus mykiss* acclimated to ecologically relevant thermal stresses

[9]. The *ldh-b* locus is also implicated in thermal acclimation and/or adaptation of fish to thermal extremes at the peripheries of their distribution ranges, with different *ldh-b* alleles (*ldh-b* ^{aa} *v. ldh-b* ^{bb}) being fixed within thermally distinct populations of the temperate estuarine killifish *Fundulus heteroclitus* (see reviews by [4, 5]). These two *ldh-b* alleles encode for two LDH-B allozymes (LDH-B$^{a\ or\ b}$), which differ by only two amino acids [10, 11]. Additionally, and in opposition to the neutralist theory, these population-specific allozymes have been shown to impact the flow of carbon (glucose) through the metabolic pathway by replacing one allozyme with the other in *F. heteroclitus* embryos [12]. A 1 bp mutation in a glucocorticoid responsive element (GRE) located in the 5' flanking region of *ldh-b* is fixed in thermally distinct *F. heteroclitus* populations, where it impacts the population-specific expression of *ldh-b*, which is believed to be adaptive [13-18]. The adaptation and/or acclimation of *ldh-b* in coral reef fishes has not been reported to date despite the fact that *ldh-b* is an ideal candidate gene to utilize for the investigation of adaptation and/or acclimation to water temperatures varying between 18 and 35^{o}C in tropical oceans (but see [19] for the tropical estuarine and freshwater fish *Lates calcarifer* and *L. niloticus*, respectively).

Plectropomus leopardus and *P. laevis* (coral trout) are commercially and ecologically important Epinepheline perciform fish that inhabit Indo-Pacific coral reefs [20]. These habitats are largely confined to depths of less than 50 meters and have experienced substantial damage and mortality associated with temperature stress in recent times [21]. *P. leopardus* and *P. laevis* exhibit a natural distribution range encompassing tropical and subtropical Indo-Pacific waters from 29°S (Abrolhos Islands, on the west Australian coast) to southern Japan in the north and eastwards to the Caroline Islands and Fiji [22]. The habitat ranges of both *Plectropomus* species span minimum temperatures of 19.5°C to maximum temperatures of more than 30°C [23]. Moreover, *P. leopardus* exhibits genetic partitioning at regional scales [20] making it an ideal

marine counter-part to the tropical estuarine species *Lates calcarifer* [19] and allowing for a comprehensive comparison of *ldh-b* among diverse congeneric fish taxa [7].

In addition to the traditional characterization of coding regions (exons) we also characterize the non-coding regions (introns) of the *ldh-b* locus in these coral reef perciformes. Such characterization allows for the identification of putative regulatory motifs (simple sequence repeat, SSR) and/or elements (microRNA, miRNA) encoded therein. Previous studies have identified such embedded functional motifs and elements within non-coding regions of other genes, which have been implicated in the regulatory pathways (e.g. transcriptome and/or proteome) of both plants [24] and animals [25-34].

2. METHODS

2.1. Complementary DNA (cDNA) synthesis from hepatic messenger RNA (mRNA)

Messenger RNA (mRNA) was extracted from a snap-frozen liver of two west Australian (WA) *P. leopardus* individuals, one from the Abrolhos Islands (28.3°S, 113.6°E) and one from Scott Reef (14.0°S, 121.8°E), WA via Trizol, following manufacturer's instructions (Invitrogen Australia Pty, Mount Waverley VIC). Extracted RNA was treated with Turbo DNA-free (Applied Biosystems/Ambion, Austin, TX, USA). Messenger RNA (mRNA) was reverse transcribed to generate cDNA immediately via IM-Prom II Reverse Transcriptase with Oligo dT and random primers (Promega, Madison, WI USA), which exclusively bound mRNA, as per manufacturer's instructions.

2.2. Amplification of ldh-b off hepatic derived cDNA and sequence verification

Amplification and sequencing of *ldh-b* from the *P. leopardus* hepatic cDNA was accomplished with general fish primers designed by aligning *ldh-b* coding sequences

from a taxonomically diverse range of fishes from the National Center for Biotechnology Information (NCBI) sequence database (GenBank). *Ldh-b* sequences from the following species were aligned in order to design primers based on conserved coding regions - Gadiformes *Trachyrincus murrayi* [**AJ609235**], *Merlangius merlangus* [**AJ609234**], *Gadus morhua* [**AJ609233**] *Coryphaenoides armatus* [**AJ609232**]; Cypriniformes *Danio rerio* and *Cyprinus carpio* [**AY644476**]; Cyprinodontiformes, *F. heteroclitus* [**L43525**], *F. heteroclitus* (D. Crawford personal communication); Squaliformes *Squalus acanthias* [**AF059035**]. While several primers were designed, those that produced a specific product from *P. leopardus* hepatic cDNA were F3 (ATGGCCTGTGCCGTCAGC) with R1 (TCTTTCAGGTCTTTCTGGAT) and these were designed to anneal in exon 2 and exon 7, respectively. Further sequence upstream (Exon 1 to end of Exon 2) was obtained using a previously published primer for *F. heteroclitus* [35] and the primer *Plectro*-Seg1-R (CAGAGCAGCTGTGGTGCGTA) to give the full *ldh-b* coding sequence for two *P. leopardus* individuals (Table 1).

All PCR reactions were conducted in the following manner: Amplification reactions (20 µL) contained the following final concentrations: 1X Buffer [2.5 mM Tris pH 8.7, 5 mM KCl, 5 mM $(NH_4)_2SO_4$ containing 1.5 mM $MgCl_2$] (Qiagen, Doncaster, VIC, Australia) or 1x Buffer [2.5 mM Tris pH 8.7, 5 mM KCl, 5 mM $(NH_4)_2SO_4$ not containing 1.5 mM $MgCl_2$] (Bioline Pty Ltd., Alexandria NSW, Australia) at 1.5mM $MgCl_2$ unless more was required (as per Table 1), 250 µM each dNTP, 250 nM each primer (Table 1), 10 ng gDNA template and 0.75 to 1.5 units of Taq Polymerase (Qiagen, Doncaster, Victoria and Bioline Pty Ltd., Alexandria New South Whales). Thermal cycling was conducted on a MJR DNA Engine thermal cycler (Bio-Rad Laboratories Pty., Ltd., Gladesville, New South Whales) as follows: initial denaturation at 94°C for 3 minutes followed by 35 cycles of 94°C denaturation for 30 seconds, annealing at primer specific T_a for 30 seconds (Table 1), 72°C extension for 30 to 90

seconds depending on target fragment size with larger fragments (> 1,000bp) requiring longer (> 60s) extension times (Table 1) and a final 72°C extension for 10 minutes. Melting temperature (T_m) was calculated according to the A/T x 2 + G/C x 4 method with the annealing temperature (T_a) considered to be 5°C less than the calculated T_m.

Amplification of cDNA with *Plectropomus*-specific *ldh-b* primers (F3 and R1) generated a strong single-band product of approximately 800 bp for both fish examined. Subsequent amplification of *Plectropomus* spp. genomic DNA with the published forward (*F. heteroclitus* - F) and designed reverse (*Plectro* - Seg1 - R) primers resulted in a strong single-band product of approximately 500 bp, which contained the missing exon 1 and 5'end of exon 2. All PCR products were precipitated with 120 µL isopropanol (70%) for 15 minutes followed by a 500 µL wash with 70% isopropanol prior to drying, re-suspension in water (10 µL) and subsequent sequencing (Macrogen, Inc., South Korea). Sequences were edited in BioEdit [36] and a contig made to give the full length coding sequence (1,005 bp) and produce a consensus cDNA sequence. To check that the correct *ldh* gene homologue had been obtained the sequence was used in a BLAST search of GenBank and also directly aligned using ClustalW in Mega 3.1 [37] to those *ldh-b* sequences previously utilized for general fish primer design (see above). The obtained sequence shared 92.5% homology with *F. heteroclitus* LDH-B amino acid sequence, a level of homology that is well above that reported for *F. heteroclitus* LDH-C vs. LDH-B and LDH-C vs. LDH-A (78% and 70% homology, respectively) [38]. As a further check that the correct *ldh* gene homologue was obtained, all nucleotide (and deduced amino acids) sequences were aligned with those of *Danio rerio ldh-b* [**AF067202**] and *ldh-a* [**NM_131246**] loci, whereby the two gene homologues differ by 600bp in length and 34.1% of the nucleotide sequence, respectively. The *P. leopardus ldh-b* sequence obtained from hepatic cDNA most closely matched that of the *D. rerio ldh-b* homologue. Once the initial cDNA sequence

was obtained from *P. leopardus* the full characterization of the *ldh-b* locus from both *Plectropomus* species (*P. leopardus* and *P. laevis*) was accomplished via primer walking along genomic DNA from representative individuals of each species.

2.3. Study species and genomic DNA (gDNA) extraction

Genomic DNA (gDNA) was extracted from samples collected by line fishing from two locations along the Western Australian Coast; the Abrolhos Islands (28.3°S, 113.6°E) and Scott Reef (14.0°S, 121.8°E), and by spear fishing from North Herald Cay, Coral Sea (16.9°S, 149.2°E). As these tissue collections were made during an independent and previous study, the gDNA obtained from these samples and utilized in this study required no animal ethics approval.

Liver and white muscle samples were immediately placed in vials on liquid nitrogen and stored at -80˚C and fin-clips were preserved in 80% ethanol. DNA extractions were performed via proteinase-K digestion (20 mg/mL) in CTAB buffer at 60°C for 1 hour and DNA was subsequently cleaned with a salt and chloroform:isoamyl alcohol (24:1) procedure [39]. All extractions resulted in high molecular weight gDNA, as visualized on a 0.8% agarose gel, with quantities ranging from 20 – 100 ng/μL.

2.4. Amplification of the full ldh-b locus off genomic DNA

Ldh-b sequences (including intron sequences) were obtained from genomic DNA extracts from thirty *P. leopardus* and two *P. laevis* individuals using the primers and primer specific PCR conditions outlined in Table 1 and Figure 1. However, the complete *ldh-b* sequences were only obtained for four of the 30 *P. leopardus* samples and both *P. laevis* individuals, which are therefore the focus of this study. In some cases different primers were required for amplification of *P. laevis* and *P. leopardus* introns (e.g. intron 5, Fig. 1). See Figure 1 for primer binding locations within the *ldh-b* locus.

Thermocycling parameters were the same as that used for cDNA amplification with the exception of the annealing temperature and magnesium chloride concentration which varied as listed in Table 1. All PCR products obtained from both *P. leopardus* and *P. laevis* genomic DNA were verified as *ldh-b* in two ways: 1) searching the NCBI database (GenBank) via nucleotide (blastn; [40, 41]) and protein (blastp; [42]) basic local alignment and search tool (b.l.a.s.t.) with resulting matches specific to *ldh-b* exclusively and 2) direct alignment with *F. heteroclitus ldh-b* nucleotide and LDH-B protein sequence. After confirmation of product identity full length *ldh-b* sequences for each individual (n = 4 *P. leopardus* and n = 2 *P. laevis*) were assembled in BioEdit and the consensus nucleotide and deduced amino acid sequences from each species were aligned using ClustalW in Mega 3.1. Consensus gene sequences for each species were submitted to GenBank under accession numbers **FJ439511** and **FJ439512** for *P. leopardus* and *P. laevis,* respectively. In addition, *ldh-b* coding sequence obtained from hepatic cDNA of *P. leopardus* was also submitted to GenBank under accession number **FJ439508**. Deduced LDH-B amino acid sequences were manually assessed for the presence of variation within the NH_2-terminal arm (residues 1-20), coenzyme binding domain (residues 21-95 and 118-163), substrate binding domain (residues 164-333) and loop helix ά D region (residues 96-117) [43].

2.5. Assessment of non-coding (intron) sequences for micro RNA (miRNA) and simple sequence repeat (SSR) motifs

Several recent studies have demonstrated intron sequences may contain simple sequence repeat (SSR) motifs of possible functional importance (e.g. when associated with promoters and/or enhancers) which potentially bind regulatory machinery and may affect gene expression levels [29, 34]. Intron sequences from *ldh-b* of both *Plectropomus* species were manually assessed for the presence of SSR motifs. In addition, short microRNA (miRNA) elements (21-23bp) may occur within introns

which when spliced out of pre-messenger RNA (mRNA) may target regions within the 3'UTR of actively expressed mRNAs, thereby regulating translation from mRNA transcripts to functional proteins [25, 44-48]. Intron sequences from both *Plectropomus* species in this study were therefore assessed for putative miRNA elements using the software package miRanda [47]. All identified miRNA elements matched known miRNAs from *Danio rerio* (zebrafish), *Takifugu rubripes* (Japanese pufferfish) and *Tetraodon nigroviridis* (spotted green pufferfish), with several also matching known *Xenopus tropicalis* (African frog) miRNAs (miRBase: [49-51]). A nucleotide match score (score), affinity to bind measure (energy [kcal mol^{-1}]) and statistical assessment of match quality (z-score) were all calculated by miRanda for each matching miRNA motif. Threshold values were set to 100, -19 kcal mol^{-1} and 5.0 for score, energy and z-score, respectively, to avoid false-positive miRNA identification [26, 47, 52] in accordance with miRNA detection in *Lates spp*. [19]. Overly-stringent thresholds were avoided to maximize the likelihood of identifying putative miRNA elements which have recently been documented at high abundance within eukaryotic genomes (an average of 100 binding sites per miRNA motif genome-wide) [26].

3. RESULTS AND DISCUSSION

3.1. Characterization of *ldh-b* exons and comparison of deduced LDH-B protein sequences of *Plectropomus* spp. and *F. heteroclitus* populations at residues 185 and 311

Variation in the primary sequence and level of *ldh-b* expression has been linked to differences in aerobic performance and natural selection in the temperate fish *F. heteroclitus* (reviewed by [4, 5]). The present study is the first to characterize the full length *ldh-b* sequence of this important metabolic locus in two congeneric coral reef fish, namely *P. leopardus* and *P. laevis*. The entire gene was 4,870 and 4,536 bp for *P.*

leopardus and *P. laevis* respectively, 1,005 bp of which were coding sites, as reported for other fish *ldh-b* coding regions (see Methods). Also consistent with other fish species, the *ldh-b* locus of both *P. leopardus* and *P. laevis* contained seven exons and six introns. All seven *ldh-b* exons were conserved in size between the two *Plectropomus* species examined. In contrast intron sizes varied between the two species and intron lengths ranged in size from 92 to 2,560 bp (Fig. 1). An inter-specific comparison of exon sequences revealed 9 base differences between *P. leopardus* and *P. laevis* (0.9% divergence), all of which are third codon position sites and therefore do not convey any amino acid substitutions. The exceptionally low divergence on both the coding nucleotide and amino acid level between *Plectropomus* spp. could be a reflection of their close phylogenetic relationship, which is supported by the fact that these two species occasionally hybridize with one another naturally [53]. However, such high levels of homology between these *Plectropomus* species could be suggestive of selective forces (e.g. purifying selection) acting upon *ldh-b* in these coral reef fish that occupy very similar thermal environments.

The two LDH-B variants fixed between thermally and geographically distinct populations of *F. heteroclitus* involved the amino acids serine (S) *v*. alanine (A) and alanine (A) *v*. aspartic acid (D) at residues 185 and 311 in northern and southern populations, respectively [35]. *P. leopardus* and *P. laevis* both possess serine (S) and aspartic acid (D) at residues 185 and 311, respectively, (Fig. 2). LDH-B of both *Plectropomus* spp. examined is therefore similar to the northern (cold water) *F. heteroclitus* population at residue 185 and to the southern (warm water) *F. heteroclitus* population at residue 311, as were both *Lates* spp. [19]. These residues are located on the internal and external surfaces of the folded LDH-B protein, respectively [10]. The former has been hypothesized to be associated with a variation in thermal stability due to its location at a potentially critical hairpin turn at the center of the folded molecule

while the latter has been associated with charge differences of the conformed protein (see reviews by [4, 5]). When considering full-length LDH-B amino acid sequences, comparison between *Plectropomus* spp. and *F. heteroclitus* revealed extensive divergence (25 out of 335 amino acids, or 7.5%), as was expected when comparing between phylogenetically distinct taxa occupying tropical and temperate environments, respectively.

Each residue of the protein sequence plays a role in the specific functioning of the molecule, from its internal stability to its external functional interactions [43, 54]. No variation was detected between *Plectropomus* spp. in the NH_2-terminal arm (residues 1-20), coenzyme (NADH) binding domain (residues 21-95 and 118-163), substrate (lactate) binding domain (residues 164-333) or in the loop helix ά D region (residues 96-117) (Fig. 2). Noteworthy is that differences were observed between the LDH-B coenzyme binding domain of two *Lates* spp**.** [19]. The LDH-B substrate binding and proton acceptor sites located at residues 100, 107, 139, 170, 249 and 194, respectively (The UniProt, '08) were also conserved between *Plectropomus* spp. (Fig. 2), a finding consistent with LDH-B comparison between two *Lates* spp. [19]. When considering the thermally distinct northern and southern populations of temperate *F. heteroclitus* as well as its congener *F. parvipinnis*, these six functional sites are also conserved. Moreover, this conservation holds even when these six functional sites are compared across all three congeneric species pairs.

3.2. Characterization of *ldh-b* introns and identification of embedded putative regulatory motifs (SSRs) and elements (miRNAs)

Sequence comparisons of the homologous introns from the two *Plectropomus* spp. reveals a high level of homology (95.9%), similar to that between exons of these two species (99.1% homology across all coding regions). Interestingly, *P. laevis* introns are

consistently smaller than *P. leopardus* introns with the exception of intron 4 which is 19 bp longer in *P. laevis* (Fig. 1). A total of 15 indels occur when non-coding *ldh-b* introns are compared between *P. leopardus* and *P. laevis* (Fig. 1). More specifically and in the order of presence within the respective introns: intron 1 contains one indel of 3 bp; intron 4 contains one indel of 18 bp; intron 5 contains nine indels of 8, 9, 12, 26, 5, 41, 19, 18 and 19 bp; and intron 6 contains four indels of 169, 4, 36 and 54 bp (Fig. 1). No potential regulatory elements/motifs (microRNA elements and simple sequence repeat motifs) are identified within *ldh-b* intronic indels (see below), which is consistent with these elements having a functional role [25, 44, 48, 55, 56].

Numerous SSR motifs were detected within the *ldh-b* introns of both *Plectropomus* species examined (Table 2). Intronic SSR motifs are of interest because they may regulate gene transcription, lead to abnormal splicing and/or disrupt export of mRNA to the cytoplasm [29, 30, 34, 57]. One mononucleotide SSR motif (A_5) is present within intron 3 of both *Plectropomus* spp. and is conserved in size between species (Table 2). Two T_5 mononucleotide repeat motifs are present within intron 4 of both *Plectropomus* spp.; however, an additional T_7 mononucleotide repeat motif is present within *P. laevis* intron 4 (Table 2). Eight mononucleotide repeat motifs (T, A and G), ranging from 5 to 13 repeats, are present in intron 5 of both *Plectropomus* spp.; additionally, two mononucleotide repeat motifs (T_{11} and A_8) occur exclusively in intron 5 of *P. laevis* (Table 2). One mononucleotide repeat motif (T_9 and T_{10}) exists within intron 6 of *P. leopardus* and *P. laevis*, respectively; however, *P. laevis* contains an additional T_7 repeat motif while *P. leopardus* contains additional A_6 and A_8 repeat motifs in intron 6 (Table 2). Two perfect dinucleotide repeat motifs (TA and AC) and one imperfect repeat motif ($TG_4AG_1TG_1$) are present at the same locations within intron 5 of both *Plectropomus* spp., with the two perfect repeat motifs differing in repeat number between the species (Table 2). In addition, two perfect dinucleotide repeat

motifs (AC and AT) are present at the same locations within intron 6 of both species with the number of repeats differing in number between species (Table 2). Two trinucleotide repeat motifs (TCC and TAC) occur within introns 1 and 5 of both species and these vary in repeat number between species (Table 2). The TCC simple sequence repeat motif present within intron 1 of both *P. leopardus* and *P. laevis* are of particular interest as a similar TCC repeat motif is present in the 5' flanking untranslated region (UTR) of *F. heteroclitus ldh-b* and, moreover, that variation in repeat number of this motif in the 5' UTR of *ldh-b* significantly affects the levels of *ldh-b* transcription in *F. heteroclitus* [18, 58]. Furthermore, this TCC repeat motif is located at the 3' end of intron 1 in both *Plectropomus* spp. is curiously similar to the TCC_4 repeat motif identified at the 3' end of intron 2 in both *Lates* spp. investigated to date [19]. This observation is further supported as relevant, by the recent demonstration that intronic sequences proximal to exon boundaries exhibit increased conservation and suggests selective constraint [59]. Lastly, one tetranucleotide repeat motif ($AAAT_3$) is present within intron 6 of *P. leopardus,* but is absent from *P. laevis ldh-b* introns (Table 2). The variation and potential functional roles, if any, of these SSR motifs present within *Plectropomus* spp. *ldh-b* introns is currently unknown. However, further investigation is warranted to determine if length variants of these SSR motifs are associated with variation in levels of *ldh-b* transcript present in the transcriptome and/or LDH-B enzyme present in the proteome, granted all other variables are controlled for.

Numerous potential microRNA elements (miRNAs) were also identified within *Plectropomus* spp. *ldh-b* introns (Table 3). Twenty-three putative miRNA elements were identified within conserved intron regions and these had a score, energy (kcal mol^{-1}), z-score and homology ranging from 112 to 140, -28.87 to -19.01, 5.01 to 6.96 and 62.1 to 90.0%, respectively, to assess *ldh-b* for miRNAs under strict parameters so as to avoid false-positive identifications (≤ 5%) [26, 47, 60] (see Methods and Table 3).

Two of the putative miRNA elements identified (Table 3: *miR-let7b, miR-92*) have previously been associated with a specific regulatory function: knockout trials of *miR-let7b* in mouse HeLa cells resulted in reduced expression of ~2,700 proteins [25, 48] and *miR-92* was enriched two- to four-fold in *Drosophila melanogaster* larval development genes [47]. In addition, two more sub-families from the highly investigated *let-7* miRNA motif family (let-7*f* and -*g*) were identified within *Plectropomus* spp. *ldh-b* introns (Table 3), which provides additional evidence for the genome-wide occurrence of miRNA targets in the 3'UTR of expressed proteins (Baek et al. 2008; Bartel 2004; Brennecke et al. 2005; Selbach et al. 2008). Noteworthy is the existence of the *let-7b, miR-15a, miR-122, miR-138, miR-148, miR-152, mirR-184, miR-210, miR-214, miR-216, miR-221 and miR-222* putative miRNA elements in coral trout *ldh-b* introns as these have also been recently identified within the *ldh-b* introns of the tropical fishes *Lates calcarifer* (barramundi) and *L. niloticus* (nile perch) [19] (Table 3).

3.7. Evidence for selective constraint on *ldh-b* non-coding (intron) sequences

The presence of known functional motifs and/or elements (e.g. SSRs and miRNAs) within non-coding introns of several genes, in addition to their even distribution across longer introns [59] provides additional evidence toward the possibility of selective constraints on non-coding nucleotide sequences (Ambros 2004; Baek et al. 2008; Bartel 2004; Hare and Palumbi 2003; Li et al. 2004; Mattick 1994; Selbach et al. 2008). More specifically, longer introns (> 87 bp) exhibit less divergence than shorter introns (< 87 bp) [59]. This could potentially be driven by two processes: a) an increased likelihood of embedded functional motifs and/or elements in longer introns, or b) the impact that nucleotide mutations may have on precursor messenger RNA (pre-mRNA) secondary structure [59]. Additional evidence for functional constraint acting on intronic sequences was evident from pairwise and cross-taxa comparisons of intron sequences of

three mammalian species (human, whale and seal) [61]. This revealed homologies 12% to 14% higher than expected from a neutral model of evolution based on expected rates of substitution for non-coding DNA. The existence of such intron motifs and/or elements is now known to be essential for the functioning of complex multi-cellular organisms as they permit a two-fold regulatory system in eukaryotic organisms, one for the transcriptome [26, 44, 62, 63] and one for the proteome (Baek et al. 2008; Chen et al. 2004; Selbach et al. 2008). Further investigations are required, however, to determine if these motifs and/or elements (SSRs, miRNAs) have similar target sites or impacts on the transcriptome (i.e. gene expression) and/or the proteome (i.e. gene silencing) within *Plectropomus* spp. and fish species in general.

TABLES AND FIGURES

Table 1. PCR and sequencing primers used to obtain *ldh-b* sequences in *Plectropomus* species. Segments 1 to 6 refer to regions depicted in Figure 1. All primers anneal to gDNA (see Methods). Primers used for forward and reverse sequencing reactions are indicated by (F) and (R), respectively.

Plectropomus spp. *Ldh-b* Primers	5' to 3' Sequence	$MgCl_2$	T_a	T_m	Extension Time (s)	Fragment Size (bp)
Segment 1*						
Plectro - Seg1-R	CAGAGCAGCTGTGGTGCGTA	3.0	53°C	64°C	30	450
Segment 2						
CTF1b (F)	ATGTGATGGAGGACCGTCTCAAG	1.5	60°C	65°C	60	400
CTR1 (R)	AGATAGCGGAAGCGGGCAGAG	-	-	68°C	-	-
Segment 3						
CTF2b (F)	TTGTTGGAACACCCTCTCTGC	1.5	59°C	64°C	60	1000
CTR2c (R)	TGTCCACTACAGCCTTGTGCG	-	-	66°C	-	-
Segment 4.1						
CTF3c (F)	AGAAGCTGAACCCTGAGATCG	3.5	59°C	64°C	90	1500
CTR3i' (R)	ACAGACCAGTTAAATAACGGCAC	-	-	66°C	-	-
Segment 4.2						
CTF3ib (F)	TCAGGCTGCCTCCTGTTATGG	1.5	59°C	66°C	90	1500
CTR3ic (R)	TGTCCCAAATTCTGAGAAACATC	-	-	64°C	-	-
CTR3ic3 (R)^	AATCATATCTGCTCTCCAATGTC	-	-	64°C	-	-
Segment 4.3						
CTF3i'' (F)	TAGGCTTTCAGGATGTTTCTCAC	4.5	59°C	66°C	90	1500
CTR3c (R)	ACACGGCTCATGTTCTTGACG	-	-	64°C	-	-
Segment 5						
CTF4 (F)	TATGAGGTGATCAAGCTGAAGG	1.5	59°C	64°C	90	1500
CTR3 (R)	TGGATGCCCCACAGCGTGTC	-	-	66°C	-	-
Segment 6**						
Plectro - I6 - F	GAAAGGTGTCTGAATGCAGC	3.0	55°C	60°C	30	400

^: Primer utilized exclusively for Segment 4.2 amplification from *P. laevis* gDNA.

* and **: Published forward and reverse *F. heteroclitus ldh-b* primers (Bernardi, et al., 1993) used in conjunction with designed primers to amplify terminal (5' and 3') segments in both *Plectropomus* species, respectively.

Table 2. Simple sequence repeat (SSR) motifs present within non-coding intron sequences of *Plectropomus* species.

SSR Motif	Location	# of Repeats in *P. leopardus*	# of Repeats in *P. laevis*
Mononucleotide			
A	Intron 3	5	5
T	Intron 4	5	5
	Intron 4	5	5
	Intron 4	-	7
T	Intron 5	7	7
	Intron 5	8	9
	Intron 5	8	13
	Intron 5	6	5
	Intron 5	-	11
	Intron 5	10	9
A	Intron 5	5	6
	Intron 5	-	8
G	Intron 5	5	6
	Intron 5	5	5
T	Intron 6	9	10
	Intron 6	-	7
A	Intron 6	6	-
	Intron 6	8	-
Dinucleotide			
$TG_4AG_1TG_1$	Intron 5	1	1
TA	Intron 5	3	4
AC	Intron 5	4	3
AC	Intron 6	6	8
AT	Intron 6	4	15
Trinucleotide			
TCC^	Intron 1	6	5
TAC	Intron 5	9	5
Tetranucleotide			
AAAT	Intron 6	3	-

^: A simple sequence repeat motif identified within intron 2 of *Lates* spp. (Edmunds, et al., 2009) and the 5' proximal promoter of *F. heteroclitus ldh-b*, of which the repeat number impacts *ldh-b* expression (Crawford, et al., 1999b; Segal, et al., 1996).

Table 3. Putative microRNA (miRNA) motifs identified within *ldh-b* intron sequences of *Plectropomus* spp. Score: nucleotide match score; Energy: affinity to bind measure; Z-Score: statistical assessment of match quality; Homology: motif homology to query intron sequence

miRNA Family	Sub-Familes	Score	Energy (kcal/mol)	Z-Score	Homology (%)
let-7	b*, f, g	129 to 132	-19.81 to -22.22	5.21 to 5.52	70.4 to 74.1
miR-15	a*, b, c	117 to 131	-19.09 to -22.49	5.01 to 6.04	62.1 to 76.9
miR-20	-, a, b	127 to 133	-19.03 to -20.45	5.0 to 5.56	65.2 to 72.0
miR-29	a, b, c	120 to 128	-19.23 to -20.98	5.43 to 6.16	75.0 to 78.3
miR-30	b, c	126 to 133	-19.78 to -22.64	5.63 to 6.8	78.3 to 90.0
miR-92	-, a, b	114 to 123	-19.02 to -28.87	5.15 to 6.95	72.0 to 79.2
miR-122*	-	126 to 131	-19.55	5.16 to 5.29	74.1 to 75.0
miR-138*	-	128	-25.34	5.89 to 6.06	80.0
miR-148	-*, b	126 to 132	-19.01 to -20.67	5.29 to 6.49	76.0
miR-152*	-	124 to 132	-22.07 to -22.35	5.25 to 6.18	80.0 to 83.3
miR-184*	-	118 to 130	-19.41 to -23.01	5.01 to 6.44	65.2 to 72.0
miR-205	-, a, b	124	-20.77 to -21.63	5.88 to 6.67	70.4
miR-210*	-	125 to 133	-20.16 to -25.19	5.69 to 6.82	69.2 to 72.0
miR-214*	-	115 to 119	-20.52 to -24.59	5.15 to 5.83	72.0 to 79,0
miR-216*	b	124 to 132	-20.68 to -23.38	5.33 to 6.45	76.0 to 80.0
miR-221*	-	128	-19.47	5.32 to 5.56	71.4
miR-222*	-	124 to 136	-20.83 to -21.63	5.15 to 6.45	71.4 to 72.4
miR-301^	b	126	-19.32	5.01	83.3
miR-456^	-	124	-21.78	5.38	67.9
miR-457^	a	136	-21.86	6.57 to 6.66	78.3
miR-489	-	131	-19.73	5.67 to 5.82	71.4
miR-723^	-	137 to 140	-21.27 to -21.86	5.83 to 6.15	75.0 to 78.6
miR-738^	-	112	-24.71	5.03	63.0

*: Identical putative miRNA elements also identified within *Lates* spp. (see Edmunds, et al., 2009). ^: MicroRNA elements matching only *Danio rerio* miRNAs (miRBase, 2009).

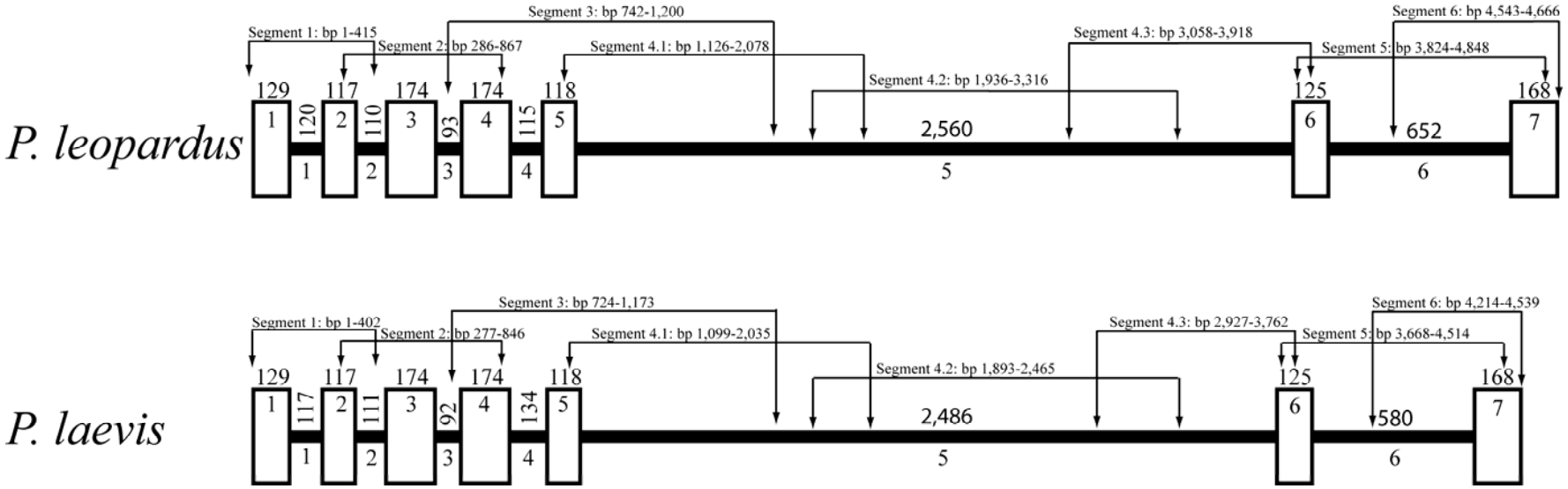


Figure 1. Comparative *ldh-b* gene map of congeneric *Plectropomus* spp. Exon (white boxes) and intron (black bars) sizes in number of base pairs (bp) are given above their respective graphic representation. Sequential numbering within white boxes and below black bars is for exons and introns, respectively. Arrowheads lines indicate region amplified by specific primer pairs (as per Table 1) along with the size of each segment in number of base pairs.

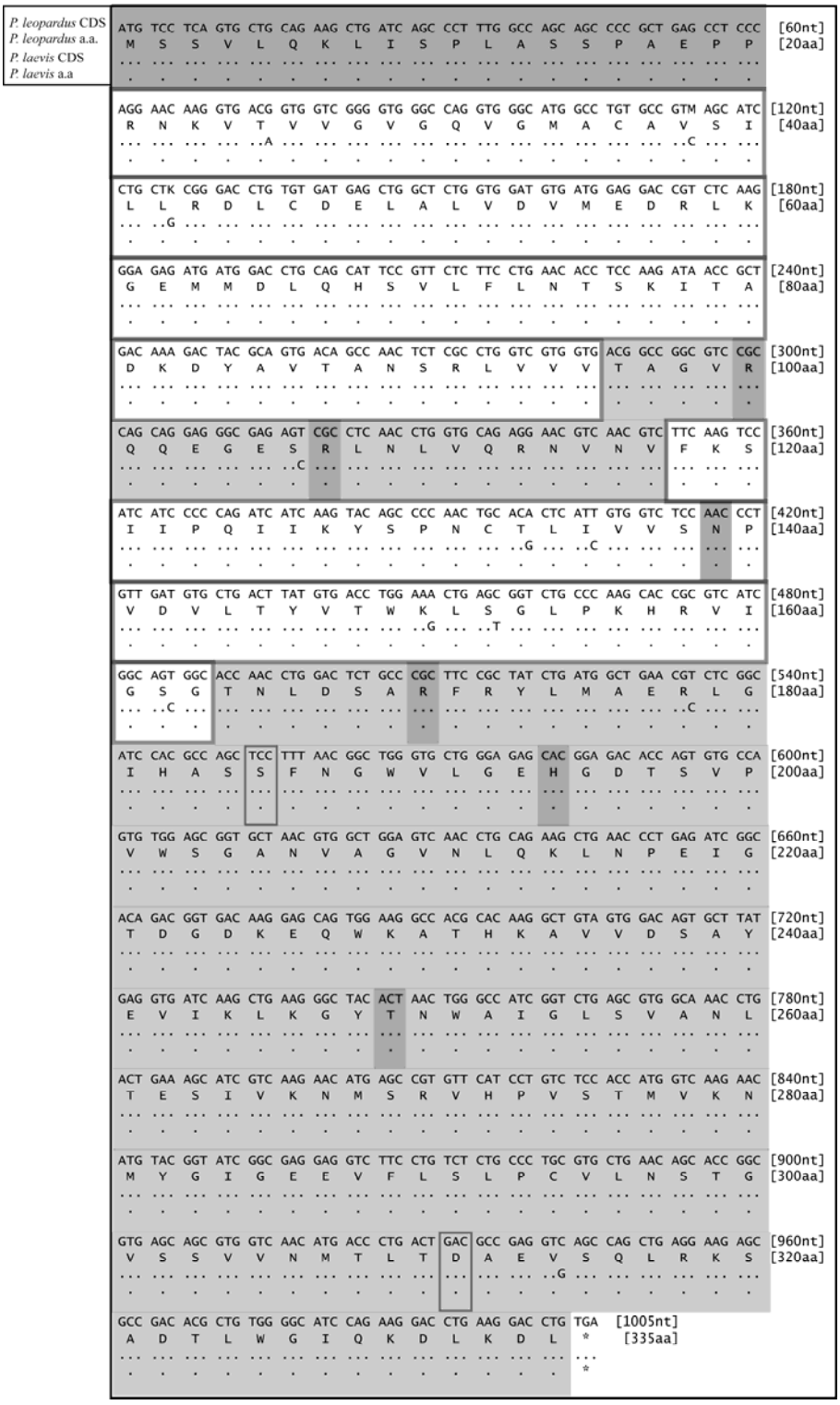


Figure A1.2. Coding nucleotide sequence (CDS) for *P. leopardus* and *P. laevis* are given on lines 1 and 3, respectively. Deduced amino acid sequence (a.a.) for *P. leopardus* and *P. laevis* are given on lines 2 and 4, respectively. Bracketed numbers at ends of line 1 and 2 refer to nucleotide position and amino acid residue, respectively. NH_2-terminal arm (residues 1-20) is dark-gray box outlined. Coenzyme binding domains (residues 21-95 and 118-163) are boldface box outlined. Substrate binding domain (residues 164-333) and loop helix ά D region (residues 96-117) are light-gray box outlined. Other known functional sites (residues 100, 107, 139, 170, 194 and 249) within domains are medium-gray box outlined. Residues (185 and 311) of known fixed difference between *F. heteroclitus* populations are boldface box outlined (see Results).

4. Conclusions

The *ldh-b* locus was found to be highly conserved between two tropical coral reef perciformes, *P. leopardus* and *P. laevis,* with just 0.9% divergence of coding regions rendering no amino acid differences between deduced LDH-B protein sequences. Non-coding intron sequences of *ldh-b* in both *Plectropomus* species were as conserved as the coding regions of this gene, despite comprising approximately 80% of the entire gene sequence. Moreover, several putative miRNA elements and SSR motifs were observed within non-coding regions of *ldh-b* in both species. These are worthy of investigation for potential functional importance in the regulation and/or expression of *ldh-b* as a constituent of the transcriptome and/or proteome. The characterization of the *ldh-b* locus presented here allows for a range of future studies such as: a) intra and inter-species variation in *ldh-b* expression of these and other Epinepheline species, as it pertains to the thermal tolerance of these important coral reef fishes; b) the importance of characterizing introns and identifying the putative regulatory motifs and/or elements within them, as they pertain to genomic, transcriptomic and/or proteomic regulatory networks; and c) an examination for evidence of adaptation at the level of the protein across fishes occupying both temperate and tropical environments in habitats of estuarine, saline and freshwater conditions.

Acknowledgements

Thanks to Dianne Rowe and Carolyn Smith-Keune for help in the Molecular Evolution and Ecology Lab; Stephen Newman (Dept of Fisheries, Western Australia) and Will Robbins for tissue samples from the study species; James Cook University Research Advancement Program for Fin-fish Genetics (RAP), AIMS @ JCU initiative and Garry Russ for funding to generate the data as well as the Department for Environment, Water Heritage and Arts for funding to collect the Coral Sea coral trout samples.

REFERENCES

1. Cossins, A.R. and D.L. Crawford. Opinion: Fish as models for environmental genomics. Nat Rev Genet 2005; **6**(April): 324.333.
2. De Santis, C. and D. Jerry. Candidate growth genes in finfish – where should we be looking? Aquaculture 2007; **272**: 22-38.
3. Nielsen, E.E., J. Hemmer-Hansen, P.F. Larsen, et al. Population genomics of marine fishes: identifying adaptive variation in space and time. Mol Ecol 2009; **18**: 3128–3150.
4. Powers, D.A. and P.M. Schulte. Evolutionary Adaptations of Gene Structure and Expression in Natural Populations in Relation to a Changing Environment: A Multidisciplinary Approach to Address the Million-Year Saga of a Small Fish. J Exp Zool Comp Exp Biol 1998; **282**: 71-94.
5. Schulte, P.M. Environmental adaptations as windows on molecular evolution. Comp Biochem Physiol B Biochem Mol Biol 2001; **128**: 597-611.
6. Guinand, B., C. Lemaire and F. Bonhomme. How to detect polymorphisms undergoing selection in marine fishes? A review of methods and case studies, including flatfishes. Journal of Sea Research 2004; **51**: 167–182.
7. Edmunds, R.C. Evidence for thermal adaptation among geographically, genetically and thermally distinct populations of the Australian barramundi, *Lates calcarifer* (Bloch 1790): a multi-level approach. Ph.D. Thesis, 2009; School of Marine and Tropical Biology, James Cook University: Townsville.
8. DiMichele, L. and D.A. Powers. Physiological Basis for Swimming Endurance Differences between LDH-B Genotypes of *Fundulus heteroclitus*. Science 1982; **216**(4549): 1014-1016.
9. Jain, K.E. and A.P. Farrell. Influence of seasonal temperature on the repeat swimming performance of rainbow trout *Oncorhynchus mykiss*. J Exp Biol 2003; **206**: 3569-3579.
10. Place, A.R. and D.A. Powers. Purification and Characterization of the Lactate Dehydrogenase (LDH-B_4) Allozymes of *Fundulus heteroclitus*. J Biol Chem 1984a; **259**: 1299-1308.
11. Place, A.R. and D.A. Powers. Kinetic Characterization of the Lactate Dehydrogenase (LDH-B_4) Allozymes of *Fundulus heteroclitus*. J Biol Chem 1984b; **259**: 1309-1318.
12. DiMichele, L., K.T. Paynter and D.A. Powers. Evidence of lactate dehydrogenase-B allozyme effects in the teleost, *Fundulus heteroclitus*. Science 1991; **253**: 898-900.
13. Crawford, D.L., V.A. Pierce and J.A. Segal. Evolutionary Physiology of Closely Related Taxa: Analyses of Enzyme Expression. Am Zool 1999b; **39**(2): 389-400.
14. Crawford, D.L. and D.A. Powers. Evolutionary Adaptation to Different Thermal Environments via Transcriptional Regulation. Mol Biol Evol 1992; **9**(5): 806-813.
15. Crawford, D.L., J.A. Segal and J.L. Barnett. Evolutionary Analysis of TATA-less Proximal Promoter Function. Mol Biol Evol 1999a; **16**(2): 194-207.
16. Picard, D.J. and P.M. Schulte. Variation in gene expression in response to stress in two populations of *Fundulus heteroclitus*. Comp Biochem Physiol Physiol 2004; **137**(2004): 205-216.
17. Schulte, P.M., H.C. Glemet, A.A. Fiebig, et al. Adaptive variation in lactate dehydrogenase-B gene expression: Role of a stress-responsive regulatory element. Proc Nati Acad Sci USA 2000; **97**(12): 6597-6602.

18. Schulte, P.M., M. Gomez-Chiarri and D.A. Powers. Structural and Functional Differences in the Promoter and 5' Flanking Region of *Ldh-B* Within and Between Populations of the Teleost *Fundulus heteroclitus*. Genetics 1997; **145**: 759-769.
19. Edmunds, R.C., L. van Herwerden, C. Smith-Keune, et al. Comparative characterization of a temperature responsive gene (lactate dehydrogenase-B, *ldh-b*) in two congeneric tropical fish, *Lates calcarifer* and *Lates niloticus*. Int J Biol Sci 2009; **5**(6): 558-569.
20. van Herwerden, L., J.H. Choat, S.J. Newman, et al. Complex patterns of population structure and recruitment of *Plectropomus leopardus* (Pisces: Epinephelidae) in the Indo-West Pacific: implications for Wsheries management. Mar Biol 2009; **156**: 1595–1607.
21. Hughes, T.P., A.H. Baird, D.R. Bellwood, et al. Climate Change, Human Impacts, and the Resilience of Coral Reefs. Science 2003; **301**: 929-933.
22. Heemstra, P.C. and J.E. Randall. FAO species catalogue, vol 16. Groupers of the world (family Serranidae, subfamily Epinephelinae). An annotated and illustrated catalogue of the grouper, rockcod, hind, coral grouper and lyretail species known to date. FAO Fish Synop 1993; **125**(16): 382.
23. Lough, J.M. Coastal climate of northwest Australia and comparisons with the Great Barrier Reef: 1960 to 1992. Coral Reefs 1998; **17**: 351-367.
24. Puzey, J.R. and E.M. Kramer. Identification of conserved *Aquilegia coerulea* microRNAs and their targets. Gene 2009; **doi:10.1016/j.gene.2009.08.005**
25. Baek, D., J. Villen, C. Shin, et al. The impact of microRNAs on protein output. Nature 2008; **455**(4 September): 64-73.
26. Brennecke, J., A. Stark, R.B. Russell, et al. Principles of microRNA–target recognition. PLoS Biol 2005; **3**(3): 0404-0418.
27. Lagos-Quintana, M., R. Rauhut, A. Yalcin, et al. Identification of tissue-specific microRNAs from mouse. Curr Biol 2002; **12**: 735-739.
28. Lai, E.C., P. Tomancak, R.W. Williams, et al. Computational identification of Drosophila microRNA genes. Genome Biol 2003; **4**(R42): 1-20.
29. Li, Y.-C., A.B. Korol, T. Fahima, et al. Microsatellites Within Genes: Structure, Function, and Evolution. Mol Biol Evol 2004; **21**: 991-1008.
30. Mattick, J.S. Introns: evolution and function. Curr Opin Genet Dev 1994; **4**: 823-831.
31. Meloni, R., V. Albanèse, P. Ravassard, et al. A tetranucleotide polymorphic microsatellite, located in the first intron of the tyrosine hydroxylase gene, acts as a transcription regulatory element *in vitro*. Hum Mol Genet 1998; **7**: 423-428.
32. Reisman, D., M. Greenberg and V. Rotter. Human p53 oncogene contains one promoter upstream of exon 1 and a second, stronger promoter within intron 1. Proc Nati Acad Sci USA 1988; **85**: 5146-5150.
33. Wittekindt, N.E., K. Hortnagel, C. Geltinger, et al. Activation of the *c-myc* promoter P1 by immunoglobulin κ gene enhancers in Burkitt lymphoma: functional characterization of the intron enhancer motifs κB, E box 1 and E box 2, and of the 3' enhancer motif PU. Nucleic Acids Res 2000; **28**(3): 800-808.
34. Yue, G., Y. Li and L. Orban. Characterization of Microsatellites in the IGF-2 and GH Genes of Asian Seabass (*Lates calcarifer*). Mar Biotechnol 2001; **3**: 1-3.
35. Bernardi, G., P. Sordino and D.A. Powers. Concordant mitochondrial and nuclear DNA phylogenies for populations of the teleost fish *Fundulus heteroclitus*. Proc Natl Acad Sci USA 1993; **90**: 9271-9274.

36. Hall, T.A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 1999; **41**: 95-98.
37. Kumar, S., K. Tamura and M. Nei. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Briefings in Bioinformatics 2004; **5**: 150-163.
38. Quattro, J.M., H.A. Woods and D.A. Powers. Sequence analysis of teleost retina-specific lactate dehydrogenase C: Evolutionary implications for the vertebrate lactate dehydrogenase gene family. Proc Natl Acad Sci USA 1993; **90**: 242-246.
39. Sambrook, J., Molecular cloning: A Laboratory Manual Woodbury, NY USA: CSHL Press; 2002.
40. Altschul, S.F., W. Gish, W. Miller, et al. Basic local alignment search tool. J Mol Biol 1990; **215**(3): 403-10.
41. Zhang, Z., S. Schwartz, L. Wagner, et al. A greedy algorithm for aligning DNA sequences. J Comput Biol 2000; **7**: 203-14.
42. Gish, W. and D.J. States. Identification of protein coding regions by database similarity search. Nat Genet 1993; **3**: 266-272.
43. Li, S.S., W.M. Fitch, Y.-C.P. Pan, et al. Evolutionary Relationships of Vertebrate Lactate Dehydrogenase Isozymes A_4 (Muscle), B_4, (Heart), and C_4, (Testis). J Biol Chem 1983; **258**(11): 7029-7032.
44. Bartel, D.P. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. Cell 2004; **116**: 281-298.
45. Chen, C.-Z., L. Li, H.F. Lodish, et al. MicroRNAs Modulate Hematopoietic Lineage Differentiation. Science 2004; **303**: 83-86.
46. Doench, J.G. and P.A. Sharp. Specificity of microRNA target selection in translational repression. Genes & Dev 2004; **18**: 504-511.
47. Enright, A.J., B. John, U. Gaul, et al. MicroRNA targets in *Drosophila*. Genome Biol 2003; **5**: R1-14.
48. Selbach, M., B. Schwanhausser, N. Thierfelder, et al. Widespread changes in protein synthesis induced by microRNAs. Nature 2008; **455**: 58-63.
49. Griffiths-Jones, S. The microRNA Registry. Nucleic Acids Res 2004; **32**(Database issue): D109 - D111.
50. Griffiths-Jones, S., R.J. Grocock, S. van Dongen, et al. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res 2006; **34**(Database issue): D140–D144.
51. Griffiths-Jones, S., H.K. Sain, S. van Dongen, et al. miRBase: tools for microRNA genomics. Nucleic Acids Res 2007; **36**(Database issue): D154–D158.
52. De Santis, C., B. Evans, C. Smith-Keune, et al. Molecular characterization, tissue expression and sequence variability of the barramundi (*Lates calcarifer*) myostatin gene. BMC Genomics 2008; **9**(1): 82.
53. van Herwerden, L., C.R. Davies and J.H. Choat. Phylogenetic and evolutionary perspectives of the Indo-Pacific grouper Plectropomus species on the Great Barrier Reef, Australia. J Fish Biol 2002; **60**: 1591-1596.
54. Somero, G.N. Protein adaptations to temperature and pressure: complementary roles of adaptive changes in amino acid sequence and internal milieu. Comp Biochem Physiol B Biochem Mol Biol 2003; **136**: 577–591.
55. Lai, E.C. MicroRNAs are complementary to 3' UTR motifs that mediate negative post-transcriptional regulation. Nat Genet 2002; **30**: 363-364.
56. Lai, E.C. microRNAs: Runts of the Genome Assert Themselves. Curr Biol 2003; **13**: R925-R936

57. Lambowitz, A.M. and M. Belfort. Introns as mobile genetic elements. Annu Rev Biochem 1993; **62**: 587-622.
58. Segal, J.A., P.M. Schulte, D.A. Powers, et al. Descriptive and Functional Characterization of Variation in the *Fundulus heteroclitus Ldh-B* Proximal Promoter. J Exp Zool Comp Exp Biol 1996; **275**: 355-364.
59. Haddrill, P.R., B. Charlesworth, D.L. Halligan, et al. Patterns of intron sequence evolution in *Drosophila* are dependent upon length and GC content. Genome Biol 2005; **6**(8): R67.
60. Bonnet, E., J. Wuyts, P. Rouzé, et al. Evidence that microRNA precursors, unlike other non-coding RNAs, have lower folding free energies than random sequences. Bioinformatics 2004; **20**(17): 2911-2818.
61. Hare, M.P. and S.R. Palumbi. High Intron Sequence Conservation Across Three Mammalian Orders Suggests Functional Constraints. Mol Biol Evol 2003; **20**: 969-978.
62. Ambros, V. The functions of animal microRNAs. Nature 2004; **431**: 350-356.
63. Nakaya, H.I., P.P. Amaral, R. Louro, et al. Genome mapping and expression analyses of human intronic noncoding RNAs reveal tissue-specific patterns and enrichment in genes related to regulation of transcription. Genome Biol 2007; **8**: R43.

Appendix 2

Evolutionary (conventional) genetic, downstream biochemical & epigenetic factors for the interaction of organisms with environmental stressors

---

1. GENERAL INTRODUCTION

Evolution by natural selection is the process by which beneficial phenotypic mutations persist due to their contribution to the organism's reproductive fitness or longevity; whereas functional, but non-beneficial, phenotypes do not survive [1]. Considering this mechanism of evolution to be true, then evolution is grounded in an organism's perpetual and complex interaction with its specific environmental parameters (e.g. temperature, pH). Its specific environment will, therefore, exude precise pressures on the organism which will ultimately generate an optimized phenotype. However, this "fine-tuned" phenotype (biochemically regulated) requires alterations to the genotype (genetically and epigenetically regulated). These genetic and epigenetic adaptations have aroused great interest in the scientific community; especially among scientists concerned with the effect of global climate change on biodiversity (e.g. how minor alterations to temperature will affect those species adapted for life at only very specific temperatures). One of the most heavily targeted model species for such studies on the effects of environmental temperature on both biochemical and genetic induced alternations is the temperate killifish *Fundulus heteroclitus* [2-8].

Genetic adaptation was first proposed to be linked directly with alterations in the regulation of genetic expression (i.e. nucleotide sequence variation [protein coding or promoter sequence] and/or chromatin methylation/acetylation patterns) in 1969 [9]. The study of variation in genetic expression via the investigation of functional sequence differences (e.g. length of microsatellites, nucleotide composition of protein coding, promoter, or upstream flanking regulatory regions) is one of the most predominant in the field of conventional (evolutionary) genetics [2-4, 6-8, 10-18]. Such studies utilize conventional methods (e.g. PCR/Sequencing, cDNA generation from mRNA, qRT-PCR, Microarrays, *in vivo* Footprinting, microsatellite

analysis) to determine the differences in specific genetic composition (i.e. nucleotide sequence variation) both within and between individuals and/or populations of model organisms such as *F. heteroclitus* and *Austrofundulus limnaeus* [12]. Such conventional studies also aim to unravel the complex mechanisms involved in the regulation of gene expression (i.e. transcription and translation), which is currently believed to be most heavily regulated at the level of transcription initiation [19]. The analysis of both mRNA and protein produced by differing DNA sequences of a candidate gene (between populations) or by differing tissue types (within individuals) are readily utilized [7]. These differences, regardless of whether they are within or between individuals/populations, are believed to have been induced by the specific environmental pressures being imposed these model organisms (e.g. *Fundulus heteroclitus* [[4]], *Austrofundulus limnaeus* [[12]], *E. coli* [[20]]) by their specific habitat (e.g. temperature & pH). The specific effects of environmental pressures on genotypes will be one of the main topics discussed in detail in "Part A" of this review.

The effects of environmental pressures, such as temperature and/or pH, have an equally important impact on downstream biochemical components (i.e. proteins and enzymes) as they do on the specific nucleotide sequences of the genes by which they are encoded. These biochemical components, of which some are encoded for by stress-sensitive genes (e.g. LDH, GAPDH, PK), are responsible for the manifested phenotypic fitness optimization (e.g. aerobic performance) of individuals in response to the pressures of their native habitat [5, 21-48]. Therefore, it is important to consider the effects of environmental stress not only on the specific genetic composition (i.e. nucleotide sequence), but also on the biochemical components (i.e. proteins) which regulate the physiological processes (e.g. aerobic metabolism) that directly determine an individual's optimized phenotype (i.e. fitness). The impact of such environmental stresses on the specific downstream biochemical functions will also be discussed in detail in "Part A" of this review.

The conventional methods of investigating genetic variation (e.g. PCR, mRNA to cDNA synthesis, nucleotide sequencing) and biochemical variation (e.g. kinetic rate constants, maximum velocity, allozyme analysis, mass spectronomy) are no longer the only methods used by the scientific community when looking at the effects of environmental pressure on an

organism's specific genetic adaptation. There is, in fact, an entirely new area of research which aims to illustrate this same environmental interaction by way of epigenetic factors (i.e. methylation and/or acetylation patterns of candidate genes) [49-62]. These epigenetic studies, unlike those conventional genetic and biochemical ones mentioned above, focus primarily on the methylation/deacetylation or demethylation/acetylation patterns of the candidate genes as the method to analyze and/or determine how candidate genes regulate their expression (e.g. increased availability of the chromatin to transcriptional machinery in unmethylated or acetylated genes). It is the potential regulation of these previously studied candidate genes (e.g. LDH, GAPDH, and PK) via the environmentally driven patterns of methylation and/or acetylation which, in conjunction with the known specific nucleotide patterns of these candidate genes, is believed to ultimately determine an organism's fitness under its specific environmental parameters. The specifics of environmentally induced methylation/acetylation patters in rats will be discussed in detail below in "Part B" of this review.

Conventional genetic techniques, biochemical variation, and epigenetic techniques provide insight as to exactly how environmental pressures generate organisms with phenotypes optimized for performance within their native or evolutionary habitat. Only by considering what all three fields have demonstrated to be true on model organisms such as *F. heteroclitus* (biochemical and genetic effects) and rats (epigenetic) can the scientific community begin to formulate a full understanding of the potential impact that global climate change will have on biodiversity.

Part A: Conventional "evolutionary" genetic & downstream biochemical factors for the interaction of organisms with specific environmental stresses (temperature & pH)

A.1. Utilization of Micro-organisms (e.g. E. coli)

The utilization of micro-organisms, such as *E. coli*, in conducting studies that look at genetic adaptation over evolutionary time are highly beneficial for many reasons: a) easily propagated, b) fast reproduction and therefore short generation times which can be easily observed, c) can induce suspended animation, d) many micro-organisms reproduce asexually, d) asexual

reproduction allows for fitness markers to be utilized, e) easily manipulated environment and genetic composition of parental generations, f) extensive amounts of data and molecular tools which can be utilized only for micro-organisms [20].

According to a recent study, the fitness gained in the first 5,000 generations was ten times greater than that which was obtained by the same individual after 15-20,000 generations [63]. Furthermore, it appears as if there is no end to genetic adaptation of an organism to a specific environment, which is evident by the fact that less beneficial mutations take exponentially longer to obtain (i.e. if a 10% beneficial mutation were to be obtained in 250 generations, then a 0.1% beneficial mutation would take 25,000 generations to obtain) than more beneficial mutations [64]. This observation, coupled with the effects of random genetic drift or the effects of chance on fitness and reproduction, make it near impossible for an organism to ever obtain a genotype that is 100% beneficial to its specific environment [20]. Furthermore, it was demonstrated by Leroi *et al.* (1994) that the acclimation of *E. coli* to a specific temperature does not always confer an advantage to those individuals acclimated. The effects of acclimation did, however, increase the survival rate of *E. coli* when exposed to lethal temperatures, but this is believed to be coupled with the induction of a stress response [35]. This observation provides evidence for the complex, and yet to be fully understood interaction between an organism (e.g. *E. coli*) and the environment in which it lives. Therefore, the utilization of micro-organisms, such as *E. coli*, have provided much insight as to the extensive timeframe on which evolutionary adaptation operates (i.e. thousands of generations), as well as the seemingly never ending ability for organisms to further adapt to their ever-changing habitat.

A.2. Water temperature as a major environmental pressure as seen via the effects on biochemical function

The water temperature in which an aquatic organism lives is thought to be one of the most pertinent stresses imposed by the environment [2-4, 6-8, 10-16, 21-23, 25, 27-30, 32, 35, 40-45, 47, 48, 65-68]. Temperature has many effects on both the specific composition of the water itself (pH [[26, 39]], oxygen saturation), and the internal biochemical scheme of the individual

(kinetic rate constants [[1, 3, 6, 7, 28, 45]], enzyme-substrate affinity [[28, 45]], and enzyme stability [[6, 7, 24, 40, 41, 45, 47]]).

The effect of water temperature on pH levels was demonstrated to be +0.0114 pH units per 1°C increase in temperature under 1 atmosphere (atm) of pressure [26]. Despite this fluctuation in water pH, the intra-mitochondrial pH was shown to be held relatively constant (0.41U) during fluctuating temperatures (10-30°C) in the red muscle of *Cyprinus carpio* [39]. This ionic stability indicates a level of protection on the intra-mitochondrial pH that may prevent effects of non-optimal pH on protein heat denaturation [69]. Interestingly, heat denaturation is believed to play a role in determining the tissue-specific expression of LDH allozymes (LDH-A, LDH-B, & LDH-C) [69]. Overall, it is evident that pH is directly linked with water temperature, that temperature-specific pH is maintained at optimum within the mitochondria, and that temperature-sensitive-pH plays a role in the distribution of metabolic enzymes into specific tissues based on optimal functioning.

Kinetic rate constant (K_m), or the amount of substrate which brings the reaction velocity to half of the maximum (V_{max} / 2), is highly influenced by the organism's environmental temperature, as has been demonstrated for LDH-B and other proteins/enzymes in *F. heteroclitus* [1, 3, 6, 7, 28, 45]. This thermal effect, by which twice as much enzyme is required to maintain the same kinetic rate per 10°C decrease in temperature directly supports the fact that kinetic rate reaches a minimum, and therefore an optimal reaction velocity, at the organism's habitat temperature [6, 28, 45]. In addition to the optimal reaction velocity being obtained at the habitat temperature of the model organism *F. heteroclitus*, sharp increases in K_m have been observed at the lower extremes of such organisms' thermal range, which has unknown but potentially important implications on maintaining biochemical function [45]. This direct correlation between optimal kinetic rate and habitat temperature shows a clear link between biochemical function (e.g. aerobic glycolysis) and both evolutionary adaptation and short-term acclimation of an organism to temperature [28, 46].

Water temperature has also been shown to directly affect enzyme-substrate affinity [6, 7, 28, 45, 46]. The enzyme-substrate affinity is measured by the reciprocal of the kinetic rate constant of substrate (K_m) [45, 46]. This affinity of enzymes for their corresponding substrate is

essential in governing the role of intracellular (i.e. low substrate levels) catalysis [45, 46]. In addition to the naturally occurring substrate affinity, most activators (i.e. positive modulators) act to increase enzyme-substrate affinity via decreasing K_m and, therefore, optimizing the velocity of that particular intracellular reaction (e.g. glycolysis) at any given time [45, 46]. The effect of water temperature on enzyme-substrate affinity is an effective example of the ability of such an environmental pressure to impact the specific biochemical functioning, and therefore overall fitness, of an organism.

The effects of water temperature are also seen on the lipids of the plasma membrane [70]. Similar effects on lipids were observed when investigating the impact of habitat temperature on the plasma membrane of rainbow trout [71]. The ratio of cholesterol to phospholipids in the plasma membrane was elevated in liver cultures that were warm-acclimated (20°C), but not in those that were cold-acclimated (5°C) [71]. These alterations in the composition of the plasma membrane, due to shifts in temperature, could in fact play a key role in the movement of critical chemical signals (e.g. stress hormones) both in and out of the cell when exposed to non-habitat temperatures. Such alterations in permeability, due to shifts in temperature, could greatly affect intra-cellular function and efficiency as alterations to cytosolic composition could very well interfere with ribosomal function (i.e. mRNA translation) or perhaps even specific gene expression (e.g. activation of upstream GRE or HIF sites). These aforementioned effects of temperature on biochemical systems show a clear link between environmental stresses and the optimization and/or refinement of an organism's molecular construct (e.g. enzyme-substrate affinity optimization). These interactions are all clear biochemical examples, obtained using conventional analysis techniques (e.g. allozyme analysis, western blot, protein gel electrophoresis), which demonstrate that an organism is "fine-tuned" to operating within the specific habitat in which it has prevailed over evolutionary time.

There are obviously many environmental factors which impose such refining pressures, however the effects of temperature are some of the clearest and most well documented [2-4, 6-8, 10-16, 21-23, 25, 27-30, 32, 35, 40-48, 65-68]. Therefore, temperature must be regarded as one of the most direct and powerful environmental stresses acting upon all organisms living within

one specific thermal regime, and thus is the reason for such extensive investigation into its impact on the genetic variation of model organisms such as *F. heteroclitus* and *A. limnaeus*.

A.3. Environmental (water) temperature and its evolutionary effect on the generation of genetically and functionally distinct lactate dehydrogenase-B (LDH-B) isozymes in populations of *Fundulus heteroclitus* from thermally distinct regions

Lactate dehydrogenase (LDH) is a critical metabolic enzyme with three allozymes: A, B, and C [6, 7]. The loci encoding for these different allozymes exhibit highly specific tissue expression [1, 2, 4-7, 14, 40, 41]. LDH-A is found predominantly in white skeletal muscle due to its role in the reduction of pyruvate to lactate following glycolysis, and may have a role in DNA replication [6, 7]. LDH-B is found predominately in heart ventricles [red muscle] and liver tissues due to its role in the oxidation of lactate back to pyruvate following glycolysis, as well as its role in gluconeogenesis (conversion of lactate back to glucose) occurring in the liver [6, 7]. In addition, LDH-B is believed to be involved with some other function because it shares an epitope with centrosomal proteins [7]. LDH-C is found in various tissues depending on the specific organism (e.g. liver of cod, eyes of salmon, mature testes of mammals and coulmid birds) [72]. Furthermore, it is believed that the LDH locus of mammals and fish was derived independently [67], and thus far there is no evidence of similarity between LDH-C in mammals and fishes [6]. Lactose dehydrogenase is highly involved in the metabolism of an organism and is therefore a likely candidate for being under direct pressure from the environmental temperature.

Lactate dehydrogenase-B (LDH-B) has thus far been the most extensively studied due to its direct role in aerobic metabolism, there therefore directly linked to the phenotypic fitness exhibited via swimming performance differences between individuals from varying thermal regimes (Maine, USA vs. Florida, USA) [1-6]. LDH-B has isozymes (-a and –b) which are fixed in populations of *F. heteroclitus* which reside in thermal habitats that vary by ~12°C annually [1, 2, 4-6, 14-16, 40, 41]. LDH-Ba isozyme is found predominately in southern or warm waters (Florida, USA) while LDH-Bb isozyme is found predominately in northern or cold waters (Nova Scotia, Canada and Maine, USA) [1, 2, 5-7, 14, 16, 40, 41]. The heterozygote isozyme

LDH-Ba/b is found predominately in the intermediate regions on the Atlantic Seaboard of the United States (New Jersey and Georgia, USA) [6, 7]. Interestingly, it has been demonstrated that LDH-B heterozygous individuals actually have an increased level of fitness due to adopting the most beneficial characteristics from both LDH-Ba and LDH-Bb [6, 32]. For example, those embryos expressing heterozygous LDH-B alleles of Sockeye salmon had an increased survival rate over the winter months [73]. The isozymes of LDH-B and their function have been well established and constitute a direct link between genetic expression and fitness within a specific thermal habitat.

Isozymes of LDH-B do not simply exhibit fixed differences between thermally varying populations of *F. heteroclitus*; they also differ significantly in the following ways: a) catalytic efficiency at low temperatures, b) degree of inhibition by substrate, c) stability at high temperatures, d) protein structure, e) effect on hemoglobin-oxygen affinity, f) downstream effect on swimming performance [1, 5-7, 14, 16, 40, 41]. Therefore, the fixed isozyme difference, which are encoded for my the fixed allelic differences, which have been demonstrated between northern and southern populations of F. heteroclitus clearly demonstrates that these two populations, from thermally distinct habitats, have not only genetic but functional biochemical differences. Thus, these differences are the product of the genetic and, therefore, downstream biochemical adaptation of *F. heteroclitus* to its specific thermal habitat.

At low temperatures (10°C), LDH-B^b (northern isozyme) has a greater catalytic efficiency at low substrate concentrations (V*max*/*Km*) than does LDH-B^a (southern isozyme) [5, 41]. This situation is reversed at higher temperatures (40°C) [5, 41]. In addition, there is an even more dramatic effect seen at 10°C when pH was held at 6.5 [5, 41]. Moreover, the fact that the pH of a cell decreases as the cell's temperature increases [7] further exemplifies the effects of pH mentioned in the previous section. The LDH-B^a isozyme is less inhibited by substrate during both lactate to pyruvate oxidation and pyruvate to lactate reduction than is LDH-B^b isozyme [6, 41]. This level of inhibition increases for both isozymes as temperature decreases [6, 41]. This variation could be a mechanism by which the organism deals with accumulating lactate levels during stressed swimming or embryonic development [6, 41].

In regard to thermal stability, LDH-B^b (cold water) is more stable than LDH-B^a (warm water) [40] but both allelic isozymes are completely stable up to 50°C, which is above the lethal limit of *F. heteroclitus* [6]. Due to the equal stability of LDH-B up to a temperature (50°C) which exceeds the upper thermal tolerance of *F. heteroclitus* (40°C), other loci were investigated to determine whether or not variations in thermal tolerance existed. Four loci (*Idh-B, H6pdh, Gpi-B, Aat-A*) were demonstrated to exhibit variation in stability within the thermal tolerance range of *F. heteroclitus* (reviewed by [6, 7]). Hence, in addition to LDH-B, detectable differences in thermal stability exist for the products of the aforementioned thermally-sensitive loci which do remain stable beyond the lethal temperature of *F. heteroclitus*.

The main difference in structure between the northern and southern allelic isozymes of LDH-B in *F. heteroclitus* (LDH-B^a and LDH-B^b) may very well be due to the two amino acid differences between them [6, 7]. These differences occur at residues 185 and 311 [2, 40, 41] and appear to be linked to the geographic location and/or thermal habitat of the individuals which express the fixed allelic difference. The northern population of *F. heteroclitus* expresses Serine (Ser) at residue 185 while southern populations express Alanine (Ala) (Powers and Schulte 1998). At residue 311, northern population express Alanine (Ala) while southern populations express Aspartic Acid (Asp) [7]. More specifically, residue 311 is located on the exterior of the molecule and the shift from Ala to Asp most likely results in a charge difference, but has no known effect on structural stability [7]. "Residues 182 and 185 form a hairpin loop that extends deep into the crevice of the other subunit. This structure may well be stabilized by hydrogen bonding between the hydroxyl hydrogen of Ser 185 and the imidazolium nitrogen of His 182" (Powers and Schulte 1998). Because this change directly affects the internal hydrogen bonding of the molecule it is more likely that this shift at residue 185 between *F. heteroclitus* populations results in the increased heat denaturation (i.e. thermal sensitivity) of the LDH-Ba isozyme (southern); moreover, this effect is most likely due to the alteration of the molecule's internal structural stability (Powers and Schulte 1998). This alteration, perhaps, could be due to different posttranslational modifications (e.g. phosphorolation) of the enzyme by the Golgi apparatus (Powers and Schulte 1998).

The direct impact of LDH-B^a *v.* LDH-B^b on hemoglobin-oxygen affinity and ATP concentrations [5, 7] goes hand-in-hand with the demonstrated differences in swimming performance between populations [2, 6, 7]. More specifically, adenine tri-phosphate (ATP) is an allosteric effector of fish hemoglobin that alters with both intra- and extra-cellular pH [5-7]. "LDH-B^b phenotype fish were able to sustain swimming speed 20% higher than that of LDH-B^a fish [at 10°C]. Blood oxygen affinity, serum lactate, liver lactate, and muscle lactate levels also differed between phenotypes when exercised to fatigue at 10°C" [6]. This increased swimming performance by the LDH-B^b phenotype at 10°C is due to this phenotype being optimized for performance in cold waters by way of its different hemoglobin-oxygen affinity (i.e. decreased saturation) and ATP concentration. This allows for the increased aerobic swimming ability of LDH-B^b individuals at 10°C when compared to that of LDH-B^a phenotypic individuals which are genetically adapted for optimal performance in warmer water. Furthermore, the connectedness of LDH-B and hemoglobin-oxygen affinity is especially high during the early stages of fish embryonic development [32]. Therefore, the effect of temperature on the specific allelic expression of LDH-B of an individual affects a lot more than just this one metabolic enzyme. It affects both hemoglobin-oxygen affinity and ATP concentrations, which ultimately determine an individual's swimming performance and/or fitness at a specific temperature.

The role of environmental temperature on generating the fixed allelic differences between populations of *F. heteroclitus*, living in thermal habitats varying by ~12°C, is a clear indication that temperature is one environmental parameter which has the ability to alter an organism's genome over evolutionary time. This alteration of the genome will also be evident in the varying abundance of LDH-B enzyme between populations, as well as thru mRNA levels and differences in upstream regulatory regions of the LDH-B gene loci. All of these differences are necessary to confer an increased fitness to one genotype over another, as demonstrated phenotypically via differential swimming performance between populations [2, 6, 7]. It appears evident that natural selection has acted upon *Ldh-B* over an evolutionary timeframe to generate genetically distinct populations of *F. heteroclitus*. Such genetic distinction most likely exists within and between populations of many other species of fish, amphibians, reptiles, and mammals. However, research into the specific genetic variability between populations of much

species is only now beginning. The variation in LDH-B nucleotide sequences, intron structure, mRNA expression, and protein abundance following aerobic exhaustion at both native and non-native water temperatures for populations of the tropical eurythermal *Lates calcarifer* (barramundi) is currently underway.

A.4. Differences in abundance of both LDH-B enzyme and *ldh-b* mRNA between thermally discrete *F. heteroclitus* populations*:*

There is a two-fold difference in the level of LDH-B enzyme abundance in liver tissue between northern populations and southern populations of *F. heteroclitus* [6, 7, 15, 28, 46, 47]. This degree of difference in enzyme concentration is expected in populations living in thermal habitats that differ by more than 10°C [1]. The difference in mRNA levels between populations is the same, two fold higher in northern population, when observed at 20°C [6, 15]. This difference in mRNA abundance is, however, lost when observed at 10°C [6, 15]. This loss of difference in mRNA abundance is likely due to one or more of the following possibilities: a) variations in the proximal promoter [3, 16], b) presence of GREs in the upstream 5' flanking region [11, 14, 16, 18], or c) both.

The two fold difference in both enzyme concentration and mRNA abundance is due to a two fold increase in the level of *ldh-b* locus transcription [2]. It is not, interestingly, due to an increase in the stability of the *ldh-b* mRNA molecule [2] This increased level of transcription could possibly be linked to the presence of a glucocorticoid receptor element (GRE) in the upstream 5' flanking region of the *ldh-b* locus [14]. GRE elements are known to increase levels of transcription in the downstream loci when in the presence of stress hormones [13]. Such hormones can potentially be released by the pituitary gland when exposed to thermal stress due to the pituitary gland's role as the "master gland" [74].

The increased levels of expression (i.e. transcription and translation), which generates the increased levels of both *ldh-b* mRNA and LDH-B enzyme, were also observed within the heart during aerobic metabolism [42]. These increased levels of heart metabolism in northern populations are tied to the differences in oxygen consumption rates [42] which are directly

linked to the specific allelic phenotype's affect on ATP concentration and, therefore, hemoglobin-oxygen affinity [5]. Furthermore, the increased expression of PK, GAPDH, and LDH explained 87% of the difference in oxygen consumption between northern and southern populations of *F. heteroclitus* [42]. These three genes must, therefore, affect metabolic rate as well as oxygen availability for metabolic consumption in highly aerobic tissues such as cardiac and red muscle [42].

The increased levels of both LDH-B enzyme and *ldh-b* mRNA levels in the northern *F. heteroclitus* populations' liver tissue when acclimated to 20°C are directly due to the increased levels of transcription of the *ldh-b* locus [2]. This increase in level of transcription must be linked to genetic differences in the transcriptional regulatory region of the *ldh-b* locus in liver and cardiac tissue of *F. heteroclitus* [2, 4, 6-8, 14, 16, 42]. Any alteration to the nucleotide sequence of these regulatory regions (i.e. proximal promoter and 5' flanking region) must have been imposed via the specific thermal stress individuals experienced being either from the northern (cold water) or the southern (warm water) population of *F. heteroclitus*.

A.5. Environmentally induced differences in the proximal promoter region and the 5' flanking region of *ldh-b* in thermally varying populations of *F. heteroclitus*

A promoter, by definition, is "a DNA segment with which RNA polymerase binds to initiate messenger RNA (mRNA) synthesis" [75]. The promoter sequence itself, however, is not transcribed [76]. Promoter regions of a gene are composed of three distinct parts: the core promoter, the proximal promoter, and the distal promoter [77]. The core promoter contains the transcription start site (-35bp), binding site for RNA polymerases (I, II, or III), and binding sites for general transcriptional machinery [77]. The proximal promoter is generally located 250bp upstream of the transcription start site and is the binding site for specific transcription factors (e.g. SAT proteins) [16, 78]. These factors are necessary in establishing the pre-initiation complex, which is critical for RNA polymerase binding and therefore fundamental in the initiation of DNA expression (i.e. transcription and translation) [16, 78].

The proximal promoter of *ldh-b* in *F. heteroclitus* has been found to be strikingly similar to the human epidermal growth factor receptor gene, a gene of which the evolutionary

and functional significance is unknown [7]. The proximal promoter of *ldh-b* in *F. heteroclitus* contains a Sp1-like motif [3, 16]; however, a perfect Sp1 motif is seen only in the southern populations [16]. More interestingly is that *Ldh-B*'s proximal promoter for *F. heteroclitus* lacks a TATA-box [3, 7, 16]. This lack of a TATA-box in the proximal promoter region of *Ldh-B* substantially alters the mechanism by which transcription initiation is undertaken. The most noteworthy impacts are the elevated role of the Sp1 site during transcription initiation [3, 18] and the dependency of transcription on the still mysterious TATA-binding protein (TBP) [18]. In addition to these alterations in TATA-less promoters, TCC-repeats occur instead of the typical TATA-box and the initiator (Inr) motifs do not bind TFIID (i.e. transcription factor) proteins [3, 16]. "*F. heteroclitus Ldh-B* promoter completely lacks a TATA-box consensus sequence, and transcription is instead mediated by interactions of basal transcription factors with the initiator motifs [Sp1, Inr, TFIID]" [7]. Interestingly, the lack of a TATA-box is considered typical of housekeeper genes [14]. Also, it is believed that indirect protein-protein interaction plays a large role in Inr-mediated transcription of TATA-less proximal promoters [18].

In the study conducted by Segal *et al*. (1996) it was observed that a five-fold decrease in transcriptional output followed the removal of the proximal promoter's Sp1 site [16]. The removal of the TCC-repeat, which replaced the TATA-box, resulted in an additional five-fold decrease in transcription of *ldh-b* [16]. Removal of the Inr, and thus the removal of all three functionally important promoter regions, terminated all promoter activity (i.e. halted transcriptional output of the gene entirely) [16]. Being that the removal of the three core components of the *ldh-b* promoter (i.e. Sp1, Inr, TFIID [TCC-repeat]) in *F. heteroclitus* had such a drastic effect on transcription rate it can, therefore, be concluded that any alteration to the nucleotide sequence of any one of these TATA-less promoter components will directly effect the transcriptional output of *Ldh-B* [16]. Furthermore, according to Crawford *et al*. (1999a), slight alterations to functional proximal promoter sequence would have substantial effects on transcriptional output of a tissue-specific enzyme. This is the proposed mechanism by which fluctuation in enzyme concentration can and do remain plastic within a specific tissue, as well as throughout an entire organism [1]. "Sequence variability between northern and southern *ldh-b*

proximal promoters [in *F. heteroclitus*] is correlated with differential *in vivo* protein binding in two functionally important regions: Sp1 site region and putative TFIID binding site [TCC-repeat]" [16]. It would seem, conclusively, that evolutionary mechanisms have a foothold with which they can impose genetic adaptation via a mere few nucleotide changes, rendering the genotype of one individual more fit under specific habitat pressures (e.g. water temperature, pH) than another slightly different genotypic individual.

Minor alterations in the promoter region of the *ldh-b* locus have been observed in *F. heteroclitus* originating from thermal environments which vary by ~12°C annually [4, 7, 16]. There is an average pair-wise difference between populations of 5.5%, whereas the pair-wise difference within populations was seen to be less (Maine [Northern]: 1.5% vs. Florida [Southern]: 2.9%) [16]. These findings support those of Oleksiak *et al.* (2002) in that there is more variation between populations than within populations, but, more importantly, that the intraspecies variation is substantial enough that it should not be overlooked [4]. Furthermore, the cytosine (C) to thymine (T) transition within the TCC-repeat at -27, which is speculated to remove the TCC-repeat's protection, was observed only in the northern *F. heteroclitus* population; whereas the T in the southern population's TCC-repeat retains its protection [16]. Minor alterations such as these hold the power to drastically alter specific gene expression, either in specific tissues or across the entire organism.

In addition to the variation found within the proximal promoter of *ldh-b* [16], variation between populations of *F. heteroclitus* has been demonstrated in the further upstream 5' flanking region [14]. One of the most noteworthy findings is the presence of a glucocorticoid receptor element (GRE), located approximately 400-500bp upstream of the transcription start site, which may play a key role in altering *Ldh-B* transcription rate and/or degree [7, 13, 14]. The GRE present in the southern *F. heteroclitus* population was found to be highly similar to the mammary tumor virus glucocorticoid responsive element (MTV-GRE) [7, 13, 14]. The MTV-GRE is a negative regulatory motif that is activated under stressful conditions [13]. The upstream MTV-GRE is not present in the northern *F. heteroclitus* population due to a single nucleotide difference between populations [7, 14]. This difference in MTV-GRE presence between northern and southern populations of *F. heteroclitus* could potentially be the cause of

the observed difference in *ldh-b* transcription rates (i.e. higher transcription rates in northern population due to their lack of a MTV-GRE site) [7, 13, 14]. However, this difference in MTV-GRE presence does not explain the loss of mRNA difference when the northern *F. heteroclitus* population was acclimated to 20°C [1]. The true role of the MTV-GRE in southern *F. heteroclitus* populations is still to be fully understood, however the knowledge of its presence is useful for other scientists aiming to answer similar questions regarding the environmental interaction for other fish species.

Another source of variation in the 5' flanking region was the length of microsatellites regions (TCC and GC repeats) [7, 14]. This variation in repeat length generates a potential for recombination in the promoter sequence [7, 14]. There are different numbers of Sp1 sites between populations (1, 2, or 3), around which two different sets of TCC-repeats were observed [7, 14]. Also, an ATC microsatellite was observed 200bp upstream from the promoter region [7, 14]. These variations between the specific microsatellite length and location is likely to play a substantial role in distinguishing between the specific gene expression patterns observed between populations of *F. heteroclitus* from different thermal habitats due to environmental pressures (e.g. temperature). There was no similarity between *F. heteroclitus ldh-b* introns and those known and available for the mouse genome [7]. However variation was observed in the sixth intron in several individuals which suggests a possible polymorphism within this intron exclusively [7]. The role and differences between intron sequences across populations and their potential role in altering gene expression is a relatively untouched area in conventional genetic studies thus far; however, much can be learned from these specific intron sequences and their possible functions. One such potential function would be the regulatory ability of the small molecules produced from the removed intron sequences on specific mRNA cytosolic translation [79-86]. Whether these small regulatory molecules act upon the gene from which their intron forbearer came or some other distant gene is still yet to be concluded [79-86]. Either way, such a role would establish much understanding as to the evolution of introns in the first instance. Intermediate populations of *F. heteroclitus*, which possess both northern and southern LDH-B alleles (LDH-B^{ab}), are more similar to whichever genotype (LDH-B^{a} or LDH-B^{b}) is most beneficial for a specific trait in a specific habitat [41]. When looking on the broader scale, it

was observed that more heterozygous individuals shared similar expression patterns with the southern genotype (LDH-B^a) than with the northern genotype (LDH-B^b) [41]. This goes hand in hand with the observed selective advantage of those individuals heterozygous for LDH-B [6, 32]. These findings support the idea that those individuals heterozygous for LDH-B have an overall increased fitness due to their "well rounded" and versatile genotype and should, therefore, be less susceptible to the pressures of climate change.

Lastly, all nucleotide polymorphisms between northern and southern *F. heteroclitus* were observed to be in functional gene regions, including those fixed differences [1, 3]. To clarify, an allele becomes "fixed" when it is expressed exclusively in one population while another allele is expressed exclusively in the other population [1]. This is critical because only fixed allelic differences generate differences in transcription [3]. In regards to *F. heteroclitus* all eight fixed nucleotide polymorphisms are unique and exclusive to the northern population [1, 3]. These fixed differences, which occur only in functional regions of the *ldh-b* gene, are what clearly differentiates the northern *F. heteroclitus* population from both southern *F. heteroclitus* and *F. grandis* populations phylogenetically [1, 3]. "These fixed differences could explain two-fold difference in transcription rate, and therefore enzyme concentration, found between the two population in both stressed and unstressed conditions due to genetic adaptation to specific thermal environment because of temperature's direct effect on enzyme function" [1].

Much of the genetic variation between populations of *F. heteroclitus* originating from thermally varying habitats is found to be not only within the protein coding sequence of *ldh-b*, but also in the proximal promoter and 5' flanking regions as well [1, 3, 4, 6, 7, 14, 16, 32, 40, 87, 88]. These observed differences in gene sequence, mRNA production, and enzyme concentration can all be linked to the specific thermal habitat from which the individual in question originated. Thus providing clear scientific evidence for the interaction between *F. heteroclitus* and the specific environmental conditions (e.g. temperature) under it has prevailed throughout evolutionary time.

A.6. Using microarrays & qRT-PCR to determine widespread gene expression differences

Both microarrays and quantitative real-time PCR (qRT-PCR) are coming under heavier and more widespread use in the scientific community as of late. Both genetic analysis techniques provide researchers a unique opportunity to directly comparing differences in the levels of expression of large numbers of candidate genes. The findings of Podrabsky and Somero (2004), Oleksiak *et al.* (2002), and Hirayama *et al.* (2005) are highly informative to those questioning the genome on a functional level and attained such insight via the use of microarrays and qRT-PCR (see also reviews by [89-92]).

Temperature, as discussed above, is known to have a direct impact on genetic expression. However, there are some differences that are only notable when individuals are subjected to constant vs. cycling temperatures [12]. Using microarray technology, 540 cDNA genes out of 4992 from *Austrofundulus limnaeus* differed by at least two fold [12]. These 540 cDNA samples were broken into eleven specific cellular function categories [12]; Figure 5 therein). Two of the most interesting fluctuations occurred in the genes encoding for heat-shock proteins & chaperones and glycolysis/gluconeogenesis & blood glucose homeostasis [12]. Interestingly, it was found that heat-shock proteins increased by at least two fold under constant thermal stress, while metabolic proteins (e.g. LDH, PK, GAPDH) increased by at least two fold under cycling thermal stress [12]. These two examples illustrate the complex interactions (i.e. cycling vs. constant temperature) between an organism and its environment and how different variations of the same stress render different alterations in specific gene expression as seen via the utilization of micro-array technology.

In a study by Oleksiak *et al.* (2002) it was observed that the variation in expression for 37 out of 907 genes (4%) was significantly different between individuals within the same population of *F. heteroclitus*. Under less stringent statistical methods 161 genes out of the 907 (18%) screened showed significant variation in expression within populations of *F. heteroclitus* [4]. When comparing the variation in expression of the same 907 genes but between populations of *F. heteroclitus* from both warm and cold environments, plus the out-group *F. grandis* from the Gulf of Mexico (warm), only 15 genes (1.65%) exhibited significant variation

in expression [4]. These 15 genes, which were shown to exhibit differential expression between populations of *F. heteroclitus,* established a clear phylogenetic grouping of southern *F. heteroclitus* with *F. grandis* and northern *F. heteroclitus* by itself (Oleksiak, Churchill et al. 2002; Figure 4a within). Further separation was established when microarrays were used on the 27 genes which exhibited significant levels of variation in northern *F. heteroclitus* population vs. both southern *F. heteroclitus* and *F. grandis* grouped together (Oleksiak, Churchill et al. 2002; Figure 6 within). This study's use of microarray technology is another clear example of the extent to which new techniques are being utilized to look at widespread genomic expression in relation to specific environmental pressures (e.g. temperature).

In a study on temperature-sensitive genes in genetically distinct *Oryzias latipes* from Nth. Japan, Sth. Japan, E. Korea, and a related tropical species *O. celebensis* by Hirayama *et al*. (2005) there was significantly different genetic expression based on the specific thermal stress. 13 out of 102 temperature-sensitive genes analyzed via qRT-PCR demonstrated significantly differential expression (i.e. at least two-fold more or less expression) at 25°C vs. 33°C [27]. Those temperature-sensitive genes found to vary in expression based on 25°C vs. 33°C were: molecular chaperones (HSP47, gp96), protein translation (PABPC)/degradation (COP9), stabilization (spermidine synthase)/decay (PABPC1) of nucleic acids, mitochondrial components (SDHD, ND1, α-F-ATPase, TIM10), heme metabolism (ALAD), glycolysis (LDH-B), intracellular trafficking (Rab-1c), immune response (gp96,IxBα) [27]. "Our data demonstrate that the temperature range in which culture cells grow is different among *medaka* populations and that these variations are associated with certain gene expression patterns" [27]. The establishment of these temperature-sensitive genes varying in expression by at least two-fold in fishes other than *F. heteroclitus* establishes that even though the temperate killifish from the Atlantic Coast of the USA has been a highly model species in functional genomic studies thus far, they are clearly not the only species to exhibit sensitivity and adaptation to specific environmental conditions (e.g. temperature).

The above mentioned studies are clear examples of how new technology (e.g. qRT-PCR and microarrays) are being implemented to observe alterations in genetic expression across a much broader ranges of genes, as well as across a much broader range of species. The

following are some of the other fishes demonstrated to exhibit differential expression of *ldh-b* due to environment stress (e.g. temperature): rainbow trout (*Oncorhynchus mykiss*) [21, 31, 33, 43, 93]; brook trout (*Salvelinus fontinalis*) [94]; lake trout (*Salvelinus namaycush*) [94]; sockeye salmon (*Oncorhynchus nerka*) [73]; fathead minnow (*Pimephales promelas*) [37]; crested blenny (*Hypleurochilus germinatus*) [29, 30].

These various studies show the degree to which the environment's interaction with *ldh-b* and LDH-B, specifically in fishes, has been studied. Moreover, these recent additions to the body of scientific knowledge provide a slightly more refined concept of just how deep the "organism-environment-interaction-iceberg" goes.

A.7. Other effects of specific thermal environment and hypoxia on various physiological, cellular, and genetic components of fishes

Lactate dehydrogenase-A is know to be expressed exclusively in white skeletal muscle as it is predominately involved in the reduction of pyruvate to lactate during anaerobic glycolysis [6, 7]. The effects of temperature on LDH-A_4 from white skeletal muscle of *Gillichthys mirabilis* and *Coryphopterus nicholsi* were analyzed [25]. Five nucleotide differences and one amino acid shift clearly distinguished *Gillichthys mirabilis* (9-30°C annual temperature range) from *Coryphopterus nicholsi* (10-18°C annual temperature range) [25]. The one amino acid shift occurred at residue 78 and involved a shift from glycine (*C. nicholsi*) to alanine (*G. mirabilis*) [25]. This one amino acid shift very well may account for the difference in thermal stability between the two LDH-A_4 orthologues [25]. In addition, *G. mirabilis* was compared to a closely related species, *G. seta*, which revealed a difference of 4 nucleotides; however, all substitutions were synonymous and therefore there was no alteration in the translated amino acid sequence [25]. This lack of effect on the amino acid sequence could be evident of different protein conformational variants (conformers) being responsible for the differences in kinetic rate and thermal stability [25]. These differences could be due to a lack of variation in primary structure, as is evident by identical nucleotide sequences (i.e. folding) [25]. Furthermore, this potential explanation is supported by the identical weights of the conformers as identified by mass spectronomy [25]. If the difference between LDH-A_4 conformers in *G. mirabilis* vs. *G. seta* is

indeed due to differential folding, then the particular folding of a protein and/or enzyme could be a fundamental mechanism by which thermal adaptation is manifested. "…Our results suggest that differences in amino acid sequence are not always necessary for adapting proteins to temperature. If the kinetics and stabilities of enzymes can be modified through the production of stable alternative conformers during the maturation of a newly synthesized protein, then another mechanism for the adaptation of enzymes, in addition to changes in their amino acid sequence, may exist" [25]. In addition, it has been demonstrated that cold temperatures render a significant effect on the flexibility of LDH-A, which ultimately causes alterations in K_m and K_{cat}, and could very well go hand-in-hand with the differential folding between conformers [25, 48]. The difference in LDH-A_4 between *G. mirabilis* vs. *C. nicholsi* and *G. mirabilis* vs. *G. seta* provides evidence toward the susceptibility of LDH-A_4 to thermal stress over evolutionary time, as well as evidence toward the mechanism of folding which may be yet another dimension at which molecules manifest there specific adaptation to the parameters of their evolutionary habitat.

Another way in which temperature is known to impact fish, in addition to genetic expression and protein/enzyme folding, is via the alteration it imposes on the blood flow to the brain [74]. Most notably is this change in blood flow's effect on the pituitary gland, or the "master endocrine gland" [74]. This impact of temperature on the endocrine signaling of the pituitary glad could easily have downstream effects on genetic expression via the GRE sites, which are known to be associated with upstream flanking regions of *Ldh-B* and other TATA-less promoters [3, 6, 7, 11, 13, 14, 18]. Endocrine regulation is also affected by lunar cycle and photoperiod via light inhibition of pineal melatonin synthesis, as was demonstrated via photoperiod regulation of the reproductive cycle in salmon [65]. According to this study, photoperiod may be more pertinent than the effects of temperature on fish behavior [65]. These studies provide evidence of yet another level at which individual organisms are susceptible and reacting to the environmental pressures (e.g. temperature, photoperiod) imposed on them by their native habitats.

Also, in addition to temperature and photoperiod, hypoxia has been shown to have an effect on genetic expression [8, 67]. Hypoxia can have a direct effect on genetic expression via

the hypoxia inducible factor (HIF-1) [8]. HIF-1 has been linked with the increased expression of LDH-B in the liver of *F. heteroclitus* when exposed to hypoxic conditions, most likely because of the critical role LDH-B plays in the oxidation of lactate to pyruvate during aerobic metabolism [6, 8, 67]. In addition, as was demonstrated in Adriatic sturgeon (*Acipenser naccarii*), the role of diet can effect oxygen consumption and therefore alters an individual's ability to maintain fitness under hypoxic conditions [36]. The degree of oxygen consumption during hypoxia was also shown to be dependent on the individuals rate of oxygen consumption during normoxia [36]. However, the rate of oxygen consumption by *A. naccarii* during both hypoxia and normoxia was dependent on dietary fatty acid content [36]. "…Specific components of a diet can influence an animal's response to environmental change" [36].

Temperature has direct effects on the folding and flexibility [25] of molecules, which appears to be evidence of yet another level of complexity through which an organism's specific evolutionary adaptation manifests itself. Temperature can also directly effect endocrine signaling via the pituitary gland [74] and therefore can have an indirect effect on those genes known to possess a GRE in their 5' flanking region [7, 13]. Even the amount of fatty acids in an organisms diet can have an effect on their ability to withstand hypoxic conditions, which will most certainly have an effect on that individuals fitness via optimized genetic functioning [36]. It appears as if the degree to which organisms are "fine-tuned" to their specific habitat is limitless, with some individuals being shaped into generalists while others are specialized to the point that any change in habitat (e.g. shift in temperature) will generate an environment in which they cannot survive [68]. "Through evolutionary processes, each species has eked out a particular thermal niche, within which it may function well, but outside of which it may fail to survive" [95]. Therefore, our ever-expanding knowledge of exactly how in-tune organisms' are to their local habitat is necessary before realistic attempts can be made to foresee the effects that global climate change will have on the biosphere.

A.8. How many molecular changes are necessary for such adaptational changes?

Exactly how much molecular variation is necessary to establish adaptational changes in organisms? For once the answer appears to be less complicated: very few [88]. The variation in

8 of the 11 glycolytic enzymes shown to be differentially expressed in *F. heteroclitus* [7] can be explained by either inherent random variation or phylogenetic distance between those species studied [88]. Based on this, it is essential to factor in any known evolutionary distance between the model organisms targeted for such studies [88]. However, the increased expression of 3 of the 11 glycolytic enzymes (glyceraldehydes 3-phosphate dehydrogenase [GAPDH], pyruvate kinase [PK], and lactate dehydrogenase [LDH]) in colder environments could not be explained by either random variation or phylogenetic distance [88]. "The most likely, and the most parsimonious, explanation for the variation in these three enzymes is that they have evolved by natural selection" [88]. The inability to explain the differential expression of the above mentioned glycolytic enzymes provides solid evidence that as little as three enzymatic differences between individuals and/or populations is grounds enough to put forth the argument that these differences are a direct result of an organism's genetic adaptation to its unique environmental pressures (e.g. temperature, pH, photoperiod) over evolutionary time.

A.9. Conclusion

To understand the specific ongoing interactions between an organism and its environment it is necessary to first investigate and understand the effects of such environmental pressure (e.g. temperature, pH) on molecular (i.e. biochemical) characteristics [5, 7, 25, 28-30, 32, 40, 41, 47, 48, 93], as well as the effects of the same environmental pressures on the gene sequences themselves which encode for all fundamental molecular components (e.g. proteins and enzymes) [2-4, 6-9, 11-16, 18, 87].

There are many environmental pressures that exist across the endless array of habitats that exist on Earth; however, not all such pressure imposed alterations can be easily investigated scientifically (e.g. magnetic fields, lunar gravitational forces, and celestial forces). Because of this seemingly endless environmentally specific pressure that is forever being imposed on all the organisms living within it, it is exceedingly necessary to target those pressures which can render scientifically credible evidence towards their impact on specific organisms. "…Given the clear evidence for functional differentiation in the *Ldh-B* regulatory sequences, and the extensive data

on physiological and developmental differences linked to the *Ldh-B* locus, it is plausible that selection has acted on the regulation of the *Ldh-B* gene in *F. heteroclitus*" [7].

One such ideal environmental pressure that can be analyzed in regards to genetic alterations, biochemical alterations, or both, is temperature [1-7, 12, 14, 15, 21-23, 27-30, 32, 35, 38, 41, 42, 44, 47, 48, 66, 68, 74]. Through the investigation of this environmental pressure (e.g. temperature) on model organisms' (e.g. *F. heteroclitus*) genetic and biochemical constructs, the scientific community is gaining an ever-clearer view of the degree to which all organisms are adapted to their specific habitat. With this heightened understand of the perpetual interactions between the environment and the organisms living within it, perhaps we can further comprehend the ever-evolving and ever-adapting nature of life on this planet.

Part B: Epigenetic factors for the interaction of organisms with specific environmental stresses

B.1. Introduction

The specific mechanisms by which epigenetics regulate gene expression have received little scientific scrutiny. DNA methylation, histone deacetylation/acetylation, and histone phosphorolation are thus far the most understood regulators of epigenetic information [49-51, 54, 56, 58-60, 62, 76]. The term "epigenetics", first used by Conrad Waddington in 1940, was described as the necessity to look at the "causal interactions between genes and their products which bring phenotypes into being" [96]. Recent studies provide substantial evidence that there is in fact a "need to study events 'over' or beyond the gene", as is imbedded in the Aristotelian root of the word "epi", meaning "upon" or "over" [52]. This up and coming field of scientific research has already begun to answer a lot of questions which have thus far remained unanswerable.

B.2. DNA methylation and histone acetylation/deacetylation

DNA methylation, according to Calladine, Drew *et al*.. 2004, has five major effects on biological activity. Firstly, DNA methylation is directly involved in the formation of specific tissues in the embryo, via *de novo* methylation by the methylases *Dnmt* 3a and 3b [76]. Second, DNA methylation is linked to the imprinting of genes in a newly fertilized egg cell; this imprinting requires the binding of methylases *Dnmt* 3a and 3L [76]. Thirdly, transcriptional repression of one of the two X chromosomes in females has been directly linked to complete DNA methylation as triggered by *Xist* RNA [76]. DNA methylation, via methylases *Dnmt* 1, 3a, & 3b, can block transcriptional factors (i.e. RNA polymerase) from binding promoter sequences and thus regulate specific nucleosome expression [76]. Lastly, DNA methylation is believed to have originally derived from the need to suppress foreign invading DNA, such as trypanosomes that are injected into an organism's bloodstream by tsetse flies [76]. Mechanistic repression of translation by DNA methylation is thought to be active on 99% of all genomic CpG sites [76].

DNA compacted around the histone octamer complex in the nucleosome is 10,000 times smaller than unpacked gDNA [76]. Repression of translation is therefore achieved by increasing nucleosome compaction, which decreases the accessibility of RNA polymerase to promoter sequences. Mechanistically, methylation is achieved by replacing a hydrogen atom with a methyl group (CH3) on a cytosine [97]. Methylated cytosine, whether located in a promoter sequence or not, will not only decrease the accessibility of RNA polymerase but will also recruit histone deacetylase 1 [76] .

This biochemical reaction is catalyzed by the methylase proteins *Dnmt*1, 3a, 3b, and 3L [76]. Methylase *Dnmt* 3L has no effect on its own; however, it has been linked to the imprinting of female pro-nuclei when coupled with *Dnmt* 3a [76]. Cytosine is converted, reversibly, into 5-methylcytosine by the addition of a methyl group to the 5^{th} position carbon [76]. Adenine can also be methylated into 6-nitro-methyladenine by the addition of a methyl group to the 6-nitrogen position; however the function of methylated adenine is unknown [76]. The addition of a methyl group generally occurs at CpG sites, or when a cytosine nucleotide is directly linked to a guanine nucleotide via a phosphate molecule [97]. A nucleosome is the combination of DNA with the histone octamer complex [98]. The histone octamer complex consists of two

copies of each non-linker histone (i.e. H2A, H2B, H3, & H4) [76]. The linker histone (H1) is found between each chromatin nucleosome repeat [76]. A nucleosome repeats approximately every 160-220 base pairs (1.7 turns of DNA) [76].

Histone deacetylation, or the removal of an acetyl (CH_3CO) group via histone de-acetylate 1 (HDAC-1), has been shown to be coupled with DNA methylation due to HDAC-1's affinity for *Dnmt* 1, 3a, 3b, & 3L [76]. The removal of an acetyl group from lysine 9 (i.e. one of the histone composing amino acids) further increases that histone's positive charge [76]. This increase in positive charge will cause the deacetylated histone H_3 to bind more tightly to the DNA of that nucleosome. Such an increased binding efficiency would further inhibit the ability of RNA polymerase to bind the promoter sequence [76]. Chromatin structure associated with transcriptional repression is referred to as "closed", or heterochromatin [59]

In contrast, the acetylation, or the addition of an acetyl group, to lysine 9 residues of histone H3 and H4 via histone acetylase neutralizes these histones' positive charge [76]. This loss of positive charge decreases the binding efficiency of histones H_3 and H_4 to the DNA of that nucleosome. This decrease in binding efficiency dissociates the histone octamer complex from DNA, which increases DNA's availability to RNA polymerase binding, and thus increases overall levels of gene expression [76]. Chromatins associated with increased levels of expression are referred to as "open", or euchromatins [59]. This inverse relationship between DNA methylation and histone acetylation was also seen in the differential glucocorticoid receptor promoter sequence methylation of rat neurons after receiving either positive or negative maternal care during the first week of life [60]. In addition to the effects of DNA methylation and histone H_3 deacetylation on the decreased levels of overall translation of a particular nucleosome, the effect of histone H_1 acetylation would also be significant on chromatin compaction. Histone H_1 plays a critical role in "stabilizing the folding of nucleosomes into this [toroidal configuration] compact form: in the absence of histone H_1, the nucleosomes are not 'frozen' by the fixative into some sort of spiral structure, but instead lie like 'beads-on-a-string'…" [76]. Therefore, the acetylation of histone H_1 would neutralize its positive charge and would decrease its binding efficiency to DNA, causing a lack in necessary compaction and unknown possible downstream or long-term effects. It would seem as if this is why the

acetylation of histone H_1 is uncommon [76]. In repressing the transcriptional ability of cells, DNA methylation has also been shown to be involved in the passing of epigenetic information from one generation to the next [99]. Methylation occurs in the gametes of the parents (i.e. spermatocytes and oocytes) and so it is these methylation patterns that are first imposed on the embryo [58]. However, spermatocytes undergo a much greater demethylation than do oocytes [57]. Spermatocytes are exposed to high amounts of demethylase in the oocyte's cytoplasm, causing substantial demethylation in sperm while having no such effect on the oocyte itself [57].

Moreover, the specific patterns of methylation and demethylation that occur during embryonic development are still unconfirmed. However, the following was seen in rats by Weaver *et al*. 2004:

> "Just before birth (embryonic day 20; E_{20}) the entire region [exon 1(7) of the glucocorticoid receptor promoter sequence] was unmethylated in both groups [those experiencing positive maternal behavior in one group (high-LG-ABN) and those experiencing negative maternal behavior in the other (low-LG-ABN)]. Strikingly, one day after birth (Postnatal day 1; P1) the exon 1[7] GR promoter was *de novo* methylated in both groups. The 5' and 3' CpG sites of the exon 1[7] GR NGFI-A [nerve growth factor-inducible protein A] response element in the offspring of both high- and low-LG-ABN mothers, which exhibit differential methylation later in life, were *de novo* methylated to the same extent. These data show that both the basal state of methylation and the first wave of *de novo* methylation after birth occur similarly in both groups. Because similar analyses are not documented for other genes, it remains unknown whether changes in methylation are common around birth or whether they are unique to this GR promoter." [60]

The methylation patterns of an individual that are established during embryonic development are passed on to all daughter cells via mitosis throughout the life of the individual [53]. The role of a cell during development (e.g. stem cell) is known to be drastically different than the functional role of a specific cell type in a mature eukaryotic organism (e.g. insulin production by pancreatic acini cells). This means that some mechanism is driving the development of specific tissue types following a specific stage in embryonic development. This

tissue specificity is thought to be achieved by the appearance of *Dnmt* 3a, 3b, and 1o (variant of *Dnmt* 1) at the 4-to-8 cell stage of development [76].

The noticeable difference between cellular function during embryonic development and tissue-specific cellular function in mature organisms highlights the fact that each individual cell physically contains the entire genome in its nucleus but cannot actively express any aspect other than those fragments essential to its tissue-specific function [53]. Only these specific CpG islands (i.e. 500-1000bp fragments of unmethylated DNA) that exhibit tissue-specific expression (e.g. insulin production by acini cells) can do so because of the constant presence of transcriptional machinery [59]. This present machinery is, in turn, what "protects" the necessary CpG islands from methylases [76]. Therefore, there must be an active mechanism by which the unexpressed majority of genomic information is repressed. The most probable mechanism suggested to date are the use of methylases *Dnmt*1, 3a, 3b, 3L and histone de-acetylase 1 in maintaining the repression of all non-expressed genes in the genome.

B.3. Potential role of epigenetics in specific biological systems

Epigenetics, or "the third component" [51], refers to the information stored in the small chemical changes that occur (e.g. methylation) in addition to the information inherent in the base paring of adenine (A) with thymine (T) and guanine (G) with cytosine (C) [76]. It is believed that the majority, if not all, of epigenetic information that is passed from one generation to the next does so via specific methylation patterns [49, 60, 99]. However, the role of histone acetylation has also been found to play a role in epigenetic regulation of gene expression [49, 60].

Studies by Weaver *et al.* (2004) and Fish *et al.* 2004 on the effects of maternal behavior on rats, found both DNA methylation and histone acetylation played major roles in hypothalamic-pituitary-adrenal (HPA) stress response. Those rats nurtured by a positive mother (e.g. licking and grooming [LG] or arched back nursing [ABN]) in the first week following birth, irrespective of which mother they were born to, had decreased DNA methylation and increased histone acetylation of glucocorticoid receptor (GR) promoter sequence at exon 1[7] [49, 60]. This difference in methylation patterns was observed in both glial cells and neurons

[49, 60]. Conversely, increased levels of methylation were observed in those rats experiencing negative maternal care (e.g. lack of LG and ABN) during the first week following birth [49, 60]. This biochemical change (e.g. DNA methylation or histone acetylation) in receptor chemistry significantly affected the rat's response to hypothalamic-pituitary-adrenal (HPA) stress due to either increasing (histone acetylation) or decreasing (DNA methylation and histone deacetylation) the binding efficiency of nerve growth factor-inducible protein A (NGFI-A) to exon 1[7] of the GR promoter sequence [49, 60]. Furthermore, the decrease in HPA stress response observed in those rats receiving positive maternal behavior were also found to have histone acetylation of exon 1[7] of the GR promoter [49, 60]. This histone acetylation caused a decrease in HPA stress response due to an increased GR promoter availability and expression; therefore, an increased sensitivity to the negative feedback control of corticotrophin-releasing factor (CRF) on HPA stress was generated [49, 60]. Moreover, methylated DNA actually prevented the acetylation of neighboring histones via attracting HDAC-1 (histone deacetylase) and methylation binding proteins (*Dnmt1*) [49, 60]. However, such effects were found to be reversible in ectopically methylated genes via the use of HDAC inhibitor trichostatin A (TSA) *in vitro* [60]. Thus, this effect of maternal behavior on either DNA methylation/histone deacetylation or histone acetylation is a clear example of how a chemical shift on promoter sequence (e.g. methylation of exon 1[7]) can render differential responses between individuals.

A second example of differential responses observed between individuals exposed to varying environmental pressures (e.g. temperature) would be the levels of high mobility group b1 (HMGB1) proteins expressed in *Austrofundulus limnaeus* when subjected to both steady-state and diurnal thermal cycling [12]. HMGB1 proteins have wide-range effects on the transcriptional activity of associated transcriptomes, one such effect being the generation of open chromatin structure when HMGB1 is present thereby providing cellular machinery access to the transcriptome [17]. HGMB1 proteins are believed to be under extremely tight control based on changes in mRNA levels [12]. Increase in the abundance of HMGB1 proteins was observed at low temperatures, but not at high temperatures (Podrabsky and Somero 2004; figure 4e). The inverse cycling difference (i.e. increased expression at low temperatures [20°C] and decreased expression at high temperatures [37°C]) of HMGB1 protein expression, which was

observed during thermal cycling (20°C to 37°C) of *Austrofundulus limnaeus,* may be tied to the difference in overall thermal tolerance limits of individuals depending on the specific implications HMGB1 has on fitness. In addition to the fluctuation in HMGB1 expression, "Arrhenius break temperature" (ABT) of heart function (i.e. a point from which individuals cannot recover) differs between populations of porcelain crabs (genus *Petrolisthes*) which naturally occur at either high temperatures (31°C) or low temperatures (16°C) [100, 101]. These alterations in expression, which appear to be directly linked to the fluctuation of naturally occurring thermal stress (e.g. high or low temperatures), provide strong scientific insight as to exactly what degree genetic and biochemical systems are operating in conjunction with habitat-specific stresses (e.g. temperature).

Individuals from the intertidal population of *P. cinctipes* have a maximum habitat temperature of 31°C and were observed to undergo ABT of heart function at 31.5°C [100, 101]. Conversely, individuals from the subtidal population of *P. eriomerus* have a maximum habitat temperature of 16°C and were observed to undergo ABT of heart function at 26.5°C [100, 101]. This clear difference in *P. cinctipes's* ability to cope with increased thermal stress may be linked to their lack of open chromatin, due to a lack of HMGB1 proteins at high temperatures, and therefore potentially tied to their lack in ability to sufficiently adjust the expression of their transcriptome accordingly. Moreover, the ability of *P. eriomerus* to sustain heart function under significant thermal stress (+10°C) is possibly linked to their increased production of HMGB1 protein due to HMGB1's speculated necessity in maintaining unstable, or open, chromatin [17] in cold environmental conditions, and therefore provides them with the ability to cope with increased levels of thermal stress [100, 101]. Overall, this difference in levels of HMGB1 protein between closely related species from different thermal environments is another clear link between thermal stress and specific gene expression, and is likely to be tightly regulated by epigenetic factors. Similar heart function limitations may be found in Barramundi, *Lates calcarifer*, when subjected to swimming trials under varying regimes (+/- 15°C of naturally occurring seasonal maximum/minimum) of thermal stress.

In addition to the role of DNA methylation/histone acetylation on response to HPA stress and HGMB1 protein production, epigenetics is considered to be a predominant

mechanism of imprinting [76]. The exact mechanism by which imprinting occurs is still unclear however it may involve the impact of *Dnmt3a/3L* complex on the oocyte pro-nuclei [76]. This would mean that external environmental factors have the ability to promote DNA methylation and/or histone acetylation in the specific nucleosomes of the correct chromosomes, in such a manner as to invoke a lifelong response by that individual toward environmental stimuli (e.g. water temperature, maternal care, pH). It is speculated, for example, that methylation imprints the knowledge of which individual bird is "mother" onto a newborn chick. This chick's ability to recognize its mother is essential for survival, so positive selection for such a mechanism is present and strong.

So how is this external environmental queue translated into an act of *de novo* methylation? Could this external stimulus trigger an increase in the expression of methylases? Moreover, how can such an external queue effectively impose an up-regulation of DNA methylase or histone acetylase expression and then accurately methylate/acetylate the appropriate chromatin within the correct cell types? The internal biochemical cascade which ensues after the chick sees (i.e. visual stimulation) and/or hears (i.e. auditory stimulation) and/or smells (i.e. olfactory stimulation) the first individual it comes into contact with, is perhaps one and the same as the cascade invoked by the maternal behavior of rats in their offspring, as per Weaver *et al*. (2004). This biochemical cascade, which ultimately ends with the appropriate methylation of the appropriate nucleosomes of the precise chromosome within the specific cells which possess the ability to connect that chick to its mother (or birthplace), is one aspect of genetics (i.e. epigenetic mechanisms) that is just now coming under scientific scrutiny.

Nonetheless, if methylation is, in fact, the major mechanism by which individuals imprint, then many other questions may be answered. For instance, such information may help to understand precisely how coral-dwelling Gobies' identify their host coral [102]. Could it be that coral hosts invoke an act of methylation/demethylation on the developing Gobie embryo, following its attachment to a branch of the host coral? This knowledge of the imprinting mechanism could also answer questions about other organisms (e.g. salamanders, frogs, turtles, etc.) that appear to undergo imprinting on the location in which they were born because they return to their birthplace to breed themselves [103-105]. This initial methylation/demethylation

of the embryo, which is invoked by either environmental pressures or inherited from the parental gametes, is a necessary mechanism when constructing complex eukaryotic organisms. In many ways, methylation-based imprinting could be the mechanism behind ants being born with the instilled knowledge of their specific job. This mechanism would serve to benefit overall population success due to not having to train/specify which individual does what after maturity (e.g. worker ants begin working immediately upon maturing). One question which remains perplexing: is the first task of the newly formed embryo to express the sequence coding for methylases (*Dnmt* 3a, 3b, 3L) and/or histone acetylase so that appropriate genes may be activated at the appropriate time and in the appropriate tissue as to render an individual optimized for functioning within a specific habitat?

B.4. Other implications of environmentally induced methylation or acetylation

An investigation of DNA methylation and histone acetylation's ability to imprint lifelong messages onto an individual will prove of great interest, particularly because it may provide evidence of the interaction between environmental stimuli and the biochemical construct of an individual (e.g. differential methylation/acetylation in glial cells and neurons of rats' brains due to different experiences of maternal behavior). Does the environment's ability to impose an act of methylation only exist during development or immediately following birth? If not, how often are living organism's genomes affected in this epigenetic context (i.e. DNA methylation or histone acetylation) by the pressures exhibited on it from the surrounding environment? Furthermore, how significant a change in the environment is necessary to invoke an alteration in the epigenetic makeup of a cell? In regard to living organisms, could day-to-day shifts in the environment (e.g. lunar cycle, light/dark cycle, seasonality, and harmonics/vibration) cause an increase or decrease in the amount of specific gene expression via the stimulation or repression of methylases and/or acetylase production? Moreover, could this altered gene expression affect protein production in such a way as to ultimately affect perceptions and behavior due to subtle changes in brain chemistry? Is it then theoretically possible that brains, like all the individual cells that constitute a eukaryotic organism, contain a massive amount of communal information that we simply cannot access due to the way we have been "methylated" since birth? Could

social pressure via stress response imprint individuals to the extent that an individual genome becomes epigenetically modified in such a way as to mold the individual with its social role (e.g. social insects, reef fishes)?

These questions may have seemed ridiculous up until now. But with mounting evidence of the constant interaction between the environment and specific gene expression as regulated by epigenetic modifications within individuals, it seems imperative to consider that all living beings are affected by the specific habitat in which they live. Therefore, epigenetics could, indeed, play a large role in the interactions between Nature and Nurture. It is no longer appropriate to consider Nature independently of Nurture. The aforementioned epigenetic studies provide clear evidence that Nature and Nurture are not mutually exclusive.

B.5. Conclusion

Recent studies undertaken on methylation and the repressing/silencing effects it has on genes [52-62, 99, 106] brings to light yet another level of complexity inherent in genetics, or the "Third Component" deemed epigenetics [107]. It is clear now that, even though every cell physically contains the entirety of genomic information, not all information can be accessed from any cell-type without steps being taken to remove any such silencing/repression (e.g. methylation and/or deacetylation). This need to remove any such silencing and/or confounding effects (e.g. DNA methylation and/or presence of introns) becomes especially clear when looking at the difference in functional protein expression between individuals and/or populations due to environmental factors (i.e. temperature, pH, salinity, etc). This further understanding of the mechanisms by which an organism's specific environment interacts with its biochemical construct (e.g. invoking methylation/deacetylation of specific DNA sequences or acetylation of specific histones to obtain a desired and beneficial result) has the power to answer many questions. The number of questions generated by this epigenetic understanding will most certainly be many times greater than the number answered by it.

2. GENERAL CONCLUSIONS

The ways in which the environment interacts with the organisms living within it can be investigated on several levels. Some of the most predominant areas of research aiming to answer such questions are in effects of environmental stress (e.g. temperature, pH, hypoxia) on biochemical components (e.g. proteins and enzymes) [5, 24, 25, 28, 40, 41, 45-47, 69, 93], genetic components (e.g. coding sequence, upstream and/or downstream regulatory sequence) [1-4, 6-8, 11-18, 42, 66, 67, 72, 87, 100, 101], and epigenetic components (e.g. (de)methylation, (de)acetylation) [49-62]. The insight provided to the scientific community by such studies is beginning to provide a clear image of precisely how integrated organisms are with their specific habitat (i.e. suite of environmental pressures). In other words, these studies provide critical insight as to how adapted certain organisms have become, over evolutionary time, to function with optimally (i.e. fitness) under the specific environmental stressors imposed by their native habitat. With further scientific inquiry into these convoluted and complex interaction between organisms and their specific native environment we can strive to obtain a comprehensive understanding as to the effect that climate shift will have on global biodiversity.

REFERENCES

1. Crawford, D.L., V.A. Pierce, and J.A. Segal, *Evolutionary Physiology of Closely Related Taxa: Analyses of Enzyme Expression.* American Zoologist, 1999b. **39**(2): p. 389-400.

2. Crawford, D.L. and D.A. Powers, *Evolutionary Adaptation to Different Thermal Environments via Transcriptional Regulation.* Molecular Biology and Evolution, 1992. **9**(5): p. 806-813.

3. Crawford, D.L., J.A. Segal, and J.L. Barnett, *Evolutionary Analysis of TATA-less Proximal Promoter Function.* Mol. Biol. Evol., 1999a. **16**(2): p. 194-207.

4. Oleksiak, M.F., G.A. Churchill, and D.L. Crawford, *Variation in gene expression within and among natural populations.* Nature Genetics, 2002. **32**(October): p. 261-266.

5. Powers, D.A., G.S. Greaney, and A.R. Place, *Physiological correlation between lactate dehydrogenase genotype and haemoglobin function in killifish.* Nature, 1979. **277**: p. 240-241.

6. Powers, D.A., et al., *Genetic Mechanisms for Adapting to a Changing Environment.* Annual Review Genetics, 1991. **25**: p. 629-659.

7. Powers, D.A. and P.M. Schulte, *Evolutionary Adaptations of Gene Structure and Expression in Natural Populations in Relation to a Changing Environment: A Multidisciplinary Approach to Address the Million-Year Saga of a Small Fish.* Journal of Experimental Zoology Part A: Comparative Experimental Biology, 1998. **282**: p. 71-94.

8. Rees, B.B., J.A.L. Bowman, and P.M. Schulte, *Structure and Sequence Conservation of a Putative Hypoxia Response Element in the Lactate Dehydrogenase-B Gene of Fundulus.* Biol. Bull., 2001. **200**(June): p. 247-251.

9. Britten, R.J. and E. Davidson, Science, 1969. **165**: p. 349-357.

10. Ohno, S., et al., *Genetic Control of Lactate Dehydrogenase Formation in the Hagfish Eptatretus.* Science, 1967. **156**: p. 96-98.

11. O'Shea-Greenfield, A. and S.T. Smale, *Roles of TATA and Initiator Elements in Determining the Start Site Location and Direction of RNA Polymerase II Transcription.* Journal of Biological Chemistry, 1992. **267**(January 15): p. 1391-1402.

12. Podrabsky, J.E. and G.N. Somero, *Changes in gene expression associated with acclimation to constant temperatures and fluctuating daily temperatures in the annual killifish Austrofundulus limnaeus.* Journal of Experimental Biology, 2004. **207**: p. 2237-2254.

13. Schulte, P.M., et al., *Adaptive variation in lactate dehydrogenase-B gene expression: Role of a stress-responsive regulatory element.* Proceedings of the National Academy of Science, USA, 2000. **97**(12): p. 6597-6602.

14. Schulte, P.M., M. Gomez-Chiarri, and D.A. Powers, *Structural and Functional Differences in the Promoter and 5′ Flanking Region of Ldh-B Within and Between Populations of the Teleost Fundulus heteroclitus.* Genetics, 1997. **145**: p. 759-769.

15. Segal, J.A. and D.L. Crawford, *LDH-B enzyme expression: the mechanisms of altered gene expression in acclimation and evolutionary adaptation.* American Physiological Society, 1994(July): p. 1150-1153.

16. Segal, J.A., et al., *Descriptive and Functional Characterization of Variation in the Fundulus heteroclitus Ldh-B Proximal Promoter.* Journal of Experimental Zoology Part A: Comparative Experimental Biology, 1996. **275**: p. 355-364.

17. Thomas, J.O. and A.A. Travers, *HMG1 and 2 and related 'architectural' DNA- binding proteins.* Trends Biochem. Sci., 2001. **26**: p. 167-174.

18. Zenzie-Gregory, B., et al., *Mechanism of Initiator-Mediated Transcription: Evidence for a Functional Interaction between the TATA-Binding Protein and DNA in the Absence of a Specific Recognition Sequence.* Molecular and Cellular Biology, 1993. **13**(7): p. 3841-3849.

19. Latchman, D., *Gene Regulation: A eukaryotic perspective*. 4th Edition ed. 2002, Cheltenham: Nelson Thornes Ltd. 323.

20. Elena, S.F. and R.E. Lenski, *Evolution Experiments with Microorganisms: The Dynamics and Genetic Bases of Adaptation.* Nature Reviews: Genetics, 2003. **4**: p. 457-469.

21. Allendorf, F.W., K.L. Knudsen, and R.F. Leary, *Adaptive significance of differences in the tissue-specific expression of a phosphoglucomutase gene in rainbow trout.* Proceedings ojthe National Academy of'Sciences, USA, 1983. **80**: p. 1397-1400.

22. Feder, M.E., A.F. Bennett, and R.B. Huey, *Evolutionary Physiology.* Annu. Rev. Ecol. Syst., 2000. **31**: p. 315-341.

23. Feldmeth, C.R. and E.A. Stone, *An Increased Scope for Thermal Tolerance upon Acclimating Pupfish (Cyprinodon) to Cycling Temperatures.* Journal of Comparative Physiology, 1974. **89**: p. 39-44.

24. Fields, P.A. and G.N. Somero, *Amino Acid Sequence Differences Cannot Fully Explain Interspecific Variation In Thermal Sensitivities Of Gobiid Fish A_4-Lactate Dehydrogenases (A_4-LDHs).* The Journal of Experimental Biology, 1997. **200**: p. 1839-1850.

25. Fields, P.A. and G.N. Somero, *Hot spots in cold adaptation: Localized increases in conformational flexability in lactate dehydrogenase A_4 orthologs of Antarctic notothenioid fishes.* Proc. Natl. Acad. Sci. USA, 1998. **95**(September): p. 11476-11481.

26. Gieskes, J.M., *Effect of Temperature on the pH of Seawater.* Limnology and Oceanography, 1969. **14**(5): p. 679-685.

27. Hirayama, M., H. Mitani, and S. Watabe, *Temperature-dependent growth rates and gene expression patterns of various medaka Oryzias latipes cell lines derived from different populations.* Journal of Comparative Physiology B, 2006. **176**: p. 311-320.

28. Hochachka, P.W. and G.N. Somero, *The Adaptation of Enzymes of Temperature.* Comp. Biochem. Physiol., 1968. **27**: p. 659-668.

29. Johnson, M.S., *Adaptive lactate dehydrogenase variation in the crested blenny, Anoplarchus.* Heredity, 1971. **27**: p. 205-226.

30. Johnson, M.S., *Association of allozymes and temperature in the crested blenny Anoplarchus purpurescens.* Marine Biology, 1977. **41**: p. 147-152.

31. Kao, Y.-H.J. and T.M. Farley, *Purification and properties of allelic lactate dehydrogenase isozymes at the B^2 locus in rainbow trout, Salmo gairdneri.* Comp. Biochem. & Physio. Pt B, 1978. **61**: p. 507-512.

32. Kirpichnikov, V.S., *Adaptive nature of intrapopulational biochemical polymorphisms in fish.* Journal of Fish Biology, 1992. **40**: p. 1-16.

33. Klar, G.T., G.B. Stalnaker, and T.M. Farley, *Comparative blood lactate response to low oxygen concentrations in rainbow trout, Salmo gairdneri, LDH-B2 phenotypes.* Comp. Biochem. & Physio. Pt A, 1979. **63**: p. 237-240.

34. Krasnov, A., et al., *Gene expression in the brain and kidney of rainbow trout in response to handling stress.* BMC Genomics, 2005. **6**(3): p. 11.

35. Leroi, A.M., A.F. Bennett, and R.E. Lenski, *Temperature acclimation and competitive fitness: An experimental test of the beneficial acclimation assumption.* Proc. Natl. Acad. Sci. USA, 1994. **91**(March): p. 1917-1921.

36. McKenzie, D.J., et al., *Effects of diet on spontaneous locomotor activity and oxygen consumption in Adriatic sturgeon (Acipenser naccarii).* Fish Physiology and Biochemistry, 1995. **14**(5): p. 341-355.

37. Merritt, R.B., *Geographical distribution and enzymatic properties of lactate dehydrogenase allozymes in the fat-head minnow, Pimephales promelas.* American Naturalist, 1972. **106**: p. 173-184.

38. Mitton, J.B. and R.K. Koehn, *Genetic Organization and Adaptive Response of Allozymes to Ecological Variables in Fundulus heteroclitus.* Genetics, 1974. **79**(January): p. 97-111.

39. Moyes, C.D., L.T. Buck, and P.W. Hochachka, *Temperature effects on pH of mitochondria isolated from carp red muscle.* Am J Physiol Regulatory Integrative Comp Physiol, 1988. **254**: p. 611-615.

40. Place, A.R. and D.A. Powers, *Purification and Characterization of the Lactate Dehydrogenase (LDH-B_4) Allozymes of Fundulus heteroclitus.* Journal of Biological Chemistry, 1984a. **259**: p. 1299-1308.

41. Place, A.R. and D.A. Powers, *Kinetic Characterization of the Lactate Dehydrogenase (LDH-B_4) Allozymes of Fundulus heteroclitus.* Journal of Biological Chemistry, 1984b. **259**: p. 1309-1318.

42. Podrabsky, J.E., et al., *Intraspecific variation in aerobic metabolism and glycolytic enzyme expression in heart ventricles.* Am J Physiol Regulatory Integrative Comp. Physiol, 2000. **279**: p. 2344-2348.

43. Rolle, N.N., *Comparative investigation of fish physiological and biochemical polymorphisms in connection with rational using of thermal waters.*, in *Institute of Evolutionary Physiology and Biochemistry*. 1981: Leningrad.

44. Sakamoto, K., M. Nakajima, and N. Taniguchi, *Thermal tolerance traits of juveniles of the ayu Plecoglossus altivelis and Japanese flounder Paralichthys olivaceus evaluated by their caudal fin cells.* Aquaculture, 2005. **246**: p. 93-99.

45. Somero, G.N., *Enzymic Mechanisms of Temperature Compensation: Immediate and Evolutionary Effects of Temperature on Enzymes of Aquatic Poikilotherms.* American Naturalist, 1969. **103**(933): p. 517-530.

46. Somero, G.N., *Temperature Adaptation of Enzymes: Biological Optimization Through Structure-Function Compromises.* Annu. Rev. Ecol. Syst., 1978. **9**: p. 1-29.

47. Somero, G.N., *Temperature and Proteins: Little Things Can Mean a Lot.* News In Physiological Science, 1996. **11**(April): p. 72-77.

48. Steffensen, J.F., *Metabolic cold adaptation of polar fish based on measurements of aerobic oxygen consumption: fact or artefact? Artefact!* Comparative Biochemistry and Physiology Part A, 2002. **132**: p. 789-795.

49. Fish, E.W., et al., *Epigenetic Programming of Stress Responses through Variations in Maternal Care.* Ann. N.Y. Acad. Sci., 2004. **1036**(2004): p. 167-180.

50. Francis, D.D., et al., *Maternal Care, Gene Expression, and the Development of Individual Differences in Stress Reactivity.* Ann. N.Y. Acad. Sci., 1999. **896**(December): p. 66-84.

51. Gartner, K., *A third component causing random variability beside environment and genotype. A reason for the limited success of a 30 year long effort to standardize laboratory animals?* Lab Anim., 1990. **24**(1): p. 71-77.

52. Jablonka, E. and M.J. Lamb, *The Changing Concept of Epigenetics.* Ann. N.Y. Acad. Sci., 2002. **981**(2002): p. 82-96.

53. Jaenisch, R. and A. Bird, *Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals.* Nature Genetics Supplement 2003. **33**: p. 245-254.

54. Kuzmin, I. and L. Geil, *DNA Methylation and Chromatin Modifications in Cancer and Development.* Int Arch Biosci, 2001: p. 1047-1056.

55. Lorincz, M.C., et al., *Dynamic Analysis of Proviral Induction and De Novo Methylation: Implications for a Histone Deacetylase-Independent, Methylation Density-Dependent Mechanism of Transcriptional Repression.* Molecular and Cellular Biology, 2000. **20**(3): p. 842-850.

56. Morange, M., *The Relations between Genetics and Epigenetics: A Historical Point of View.* Ann. N.Y. Acad. Sci., 2002. **981**(2002): p. 50-60.

57. Oswald, J., et al., *Active demethylation of the paternal genome in the mouse zygote.* Current Biology, 2000. **10**: p. 475-478.

58. Sapienza, C., *Parental imprinting of genes.* Sci. Am., 1990. **263**: p. 52-60.

59. Sutherland, J.E. and M. Costa, *Epigenetics and the Environment.* Ann. N.Y. Acad. Sci., 2003. **983**(2003): p. 151-160.

60. Weaver, I.C.G., et al., *Epigenetic programming by maternal behavior.* Nature Neuroscience, 2004. **7**(8): p. 847-854.

61. Yanagimachi, R., *Male Gamete Contributions to the Embryo.* Ann. N.Y. Acad. Sci., 2005. **1061**(2005): p. 203-207.

62. Zhang, T.-Y., et al., *Maternal Programming of Individual Differences in Defensive Responses in the Rat.* Ann. N.Y. Acad. Sci., 2004. **1032**(2004): p. 85-103.

63. Cooper, V.S. and R.E. Lenski, *The population genetics of ecological specialization in evolving Escherichia coli populations.* Nature, 2000. **407**: p. 736-739.

64. Lenski, R.E., et al., *Long-term experimental evolution in Escherichia coli. I. Adaptation and divergence during 2,000 generations.* American Naturalist, 1991. **138**: p. 1315-1341.

65. Pankhurst, N.W. and M.J.R. Porter, *Cold and dark or warm and light: variations on the theme of environmental control of reproduction.* Fish Physiology and Biochemistry, 2003. **28**: p. 385-389.

66. Picard, D.J. and P.M. Schulte, *Variation in gene expression in response to stress in two populations of Fundulus heteroclitus.* Comparative Biochemistry and Physiology Part A, 2003. **137**(2004): p. 205-216.

67. Quattro, J.M., H.A. Woods, and D.A. Powers, *Sequence analysis of teleost retina-specific lactate dehydrogenase C: Evolutionary implications for the vertebrate lactate dehydrogenase gene family.* Proc. Natl. Acad. Sci. USA, 1993. **90**: p. 242-246.

68. Weinstein, R.B. and G.N. Somero, *Effects of Temperature on mitochondrial function in the Antarctic fish Trematomus bernacchii.* Journal of Comparative Physiology B, 1998. **168**: p. 190-196.

69. Vesell, E.S. and K.L. Yielding, *Effects of pH, Ionic Strength, and Metabolic Intermediates on the Rates of Heat Inactivation of Lactate Dehydrogenase Isozymes.* PNAS, 1966. **56**(4): p. 1317-1324.

70. Hazel, J.R., *THERMAL ADAPTATION IN BIOLOGICAL MEMBRANES: Is Homeoviscous Adaptation the Explanation?* Annu. Rev. Physiol., 1995. **57**: p. 19-42.

71. Zehmer, J.K. and J.R. Hazel, *Plasma membrane rafts of rainbow trout are subject to thermal acclimation.* The Journal of Experimental Biology, 2003. **206**: p. 1657-1667.

72. Tsuji, S., et al., *Evolutionary relationships of lactate dehydrogenases (LDHs) from mammals, birds, an amphibian, fish, barley, and bacteria: LDH cDNA sequences from Xenopus, pig, and rat.* Proc. Natl. Acad. Sci. USA, 1994. **91**(Evolution): p. 9392-9396.

73. Kirpichnikov, V.S., et al., *Genetic structure and allele frequency dynamics in the sockeye salmon population of Lake Dalneye, Kamchatka.* Aquaculture, 1990. **84**: p. 13-25.

74. van den Burg, E.H., et al., *Brain Responses to Ambient Temperature Fluctuations in Fish: Reduction of Blood Volume and Initiation of a Whole-Body Stress Response.* Journal of Neurophysiology, 2005. **93**: p. 2849-2855.

75. Frank-Kamenetskii, M.D., *Unraveling DNA: The Most Important Molecule of Life*. Second ed. 1997, Reading, Massachusetts: Perseus Books. 214.

76. Calladine, C.R., et al., *Understanding DNA: the molecule and how it works (Third Edition)*. Third ed. 2004, San Diego: Elsevier Academic Press 334.

77. Wikipedia. *Transcription Factors*. 2006 [cited 2006 1/8/2006]; Available from: http://en.wikipedia.org/wiki/Transcription_factor.

78. Wikipedia. *Promoter Elements*. 2006 [cited 2006 1/8/2006]; Available from: http://en.wikipedia.org/wiki/Promoter#Promoter_elements.

79. Ambros, V., *The functions of animal microRNAs.* Nature, 2004. **431**: p. 350-356.

80. Baek, D., et al., *The impact of microRNAs on protein output.* Nature, 2008. **455**(4 September): p. 64-73.

81. Bartel, D.P., *MicroRNAs: Genomics, Biogenesis, Mechanism, and Function.* Cell, 2004. **116**: p. 281-298.

82. Brennecke, J., et al., *Principles of microRNA–target recognition.* PLoS Biol, 2005. **3**(3): p. 0404-0418.

83. Lai, E.C., *microRNAs: Runts of the Genome Assert Themselves.* Current Biology, 2003. **13**: p. 12.

84. Li, Y.-C., et al., *Microsatellites Within Genes: Structure, Function, and Evolution.* Mol. Bio. & Evo., 2004. **21**: p. 991-1008.

85. Meloni, R., et al., *A tetranucleotide polymorphic microsatellite, located in the first intron of the tyrosine hydroxylase gene, acts as a transcription regulatory element in vitro.* Human Molecular Genetics, 1998. **7**: p. 423-428.

86. Selbach, M., et al., *Widespread changes in protein synthesis induced by microRNAs.* Nature, 2008. **455**: p. 58-63.

87. Bernardi, G., P. Sordino, and D.A. Powers, *Concordant mitochondrial and nuclear DNA phylogenies for populations of the teleost fish Fundulus heteroclitus.* Proc. Natl. Acad. Sci. USA, 1993. **90**: p. 9271-9274.

88. Pierce, V.A. and D.L. Crawford, *Phylogenetic Analysis of Glycolytic Enzyme Expression.* Science, 1997. **276**(11 April): p. 256-259.

89. Bustin, S.A., *Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays.* Journal of Molecular Endocrinology, 2000. **25**: p. 169–193.

90. Bustin, S.A., et al., *Quantitative real-time RT-PCR – a perspective.* Journal of Molecular Endocrinology, 2005. **34**: p. 597–601.

91. Bustin, S.A. and T. Nolan, *Pitfalls of Quantitative Real-Time Reverse-Transcription Polymerase Chain Reaction.* Journal of Biomolecular Techniques, 2004. **15**: p. 155-166.

92. Douglas, S., E., *Microarray Studies of Gene Expression in Fish.* OMICS A Journal of Integrative Biology, 2006. **10**(4): p. 474-489.

93. Tsuyuki, H. and S.N. Williscroft, *The pH activity relations of the LDH homotetramers from trout liver and their physiological significance.* Journal of the Fisheries Research Board of Canada, 1973. **30**: p. 1023-1026.

94. Wuntch, T. and E. Goldberg, *A comparative physico-chemical characterization of lactate dehydrogenase isozymes in brook trout, lake trout and their hybrid splake trout.* Journal of Experimental Zoology, 1970. **174**: p. 233-252.

95. Hochachka, P.W. and G.N. Somero, *Biochemical Adaptation: Mechanism and Process in Physiological Evolution*. 2002, Oxford: Oxford University Press. 466.

96. Waddington, C.H., *The basic ideas of biology*. 1-32 ed. Towards a Theoretical Biology. Vol. 1: Prolegomena. 1968, Edinburgh: Edinburgh University Press.

97. Wikipedia. *Methylation*. 2005 [cited 2006 20/6/2006]; Available from: www.en.wikipedia.org/wiki/methylation.

98. Wikipedia. *Nucleosome*. 2005 [cited 2006 20/6/2006]; Available from: www.en.wikipedia.org/wiki/nucleosome.

99. Wong, A.H.C., I.I. Gottesman, and A. Petronis, *Phenotypic differences in genetically identical organisms: the epigenetic perspective.* Human Molecular Genetics, 2005. **14**(Review 1): p. R11-R18.

100. Stillman, J.H. and G.N. Somero, *Adaptation to temperature stress and aerial exposure in congeneric species of intertidal porcelain crabs (genus Petrolisthes): Correlation of physiology, biochemistry and morphology with vertical distribution.* Journal of Experimental Biology, 1996. **199**: p. 1845-1855.

101. Stillman, J.H. and G.N. Somero, *A comparative analysis of the upper thermal tolerance limits of eastern Pacific porcelain crabs, genus Petrolisthes : Influences of latitude, vertical zonation, acclimation, and phylogeny.* Physiological Biochemical Zoology 2000. **73**: p. 200-208.

102. Munday, P.L., *Does habitat availability determine geographical-scale abundances of coral-dwelling fishes?* Coral Reefs, 2002. **21**: p. 105-116.

103. Diego-Rasilla, E.J. and R.M. Luengo, *Celestial orientation in the marbled newt (Triturus marmoratus).* J. of Ethol., 2002. **20**: p. 137-141.

104. Limpus, C.J., et al., *Migration of Green (Chelonia mydas) and Loggerhead (Caretta caretta) Turtles to and from Eastern Australian Rookeries.* Wildl. Res., 1992. **19**: p. 347-358.

105. Sinsch, U., *Orientation and navigation in Amphibia.* Marine and Freshwater Behaviour and Physiology, 2006. **39**(1): p. 65-71.

106. Monk, M., M. Boubelik, and S. Lehnert, *Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development.* Development, 1987. **99**: p. 371-382.

107. Waddington, C.H., *Organisers and Genes*. 1940, Cambridge: Cambridge University Press.

Appendix 3

Permission from CSIRO Publishing for reproduction of Figure 1 from Shaklee *et al.* 1985

Dear Richard
Thank you for your email.

We grant you permission for this Figure 1 from the article cited below to be included in your PhD thesis, provided you fully acknowledge the journal, the authors and **CSIRO** PUBLISHING as the original source, as well as cite the link to our website - http://www.publish.csiro.au/nid/126/paper/MF9850203.htm.

This permission is granted for this sole purpose only.

With best wishes,

Carla Flores
Rights & Permissions, **CSIRO** PUBLISHING
150 Oxford Street (PO Box 1139)
Collingwood VIC 3066

Fax (03) 9662 7595

International Journal of Biological Sciences
2009; 5(6):558-569

Research Paper

Comparative characterization of a temperature responsive gene (lactate dehydrogenase-B, *ldh-b*) in two congeneric tropical fish, *Lates calcarifer* and *Lates niloticus*

Richard C. Edmunds ✉, Lynne van Herwerden, Carolyn Smith-Keune, Dean R. Jerry

Molecular Evolution and Ecology Laboratory, School of Marine & Tropical Biology, James Cook University, Townsville QLD 4811, Australia

✉ Correspondence to: Richard C. Edmunds, School of Marine & Tropical Biology, James Cook University, Townsville QLD 4811 *Australi*a. Tel: +61 7 4781 5735; Fax: +61 7 4725 1570; E-mail: richard.c.edmunds@gmail.com

Received: 2009.05.19; Accepted: 2009.08.26; Published: 2009.09.01

Abstract

The characterization of candidate loci is a critical step in obtaining insight into adaptation and acclimation of organisms. In this study of two non-model tropical (to sub-tropical) congeneric perciformes (*Lates calcarifer* and *Lates niloticus*) we characterized both coding and non-coding regions of lactate dehydrogenase-B (*ldh-b*), a locus which exhibits temperature-adaptive differences among temperate and sub-tropical populations of the North American killifish *Fundulus heteroclitus*. *Ldh-b* was 5,004 and 3,527 bp in length in *L. calcarifer* and *L. niloticus*, respectively, with coding regions comprising 1,005 bp in both species. A high level of sequence homology existed between species for both coding and non-coding regions of *ldh-b* (> 97% homology), corresponding to a 98.5% amino acid sequence homology. All six known functional sites within the encoded protein sequence (LDH-B) were conserved between the two *Lates* species. Ten simple sequence repeat (SSR) motifs (mono-, di-, tri- and tetranucleotide) and thirty putative microRNA elements (miRNAs) were identified within introns 1, 2, 5 and 6 of both *Lates* species. Five single nucleotide polymorphisms (SNPs) were also identified within miRNA containing intron regions. Such SNPs are implicated in several complex human conditions and/or diseases (as demonstrated by extensive genome-wide association studies). This novel characterization serves as a platform to further examine how non-model species may respond to changes in their native temperatures, which are expected to increase by up to 6°C over the next century.

Key words: Barramundi; Nile perch; exons; introns; microRNAs; SSRs

Introduction

The lactate dehydrogenase-B enzyme (LDH-B) plays a critical role in maintaining aerobic metabolism by converting lactate, the major by-product of anaerobic glycolysis, to pyruvate via oxidation in the presence of its coenzyme nicotinamide adenine dinucleotide, (NADH) (reviewed by [1, 2]). LDH-B can also convert lactate directly to glucose via gluconeogenesis. This conversion of accumulating lactate from aerobic tissues (e.g. heart, skeletal muscle) occurs in the liver and allows desired aerobic metabolic activity to be sustained for extended lengths of time (reviewed by [1, 2]). In addition to these metabolic functions, the LDH-B enzyme affects the oxygen binding affinity of hemoglobin (Hb) by altering intra-erythrocyte ATP concentrations of Hb in fish [3, 4]. This effect of LDH-B on Hb-O_2 binding affinity directly impacts delivery of Hb-bound oxygen to red muscle tissues and may therefore be an alternate mechanism by

which LDH-B affects sustainability of aerobic performance such as swimming performance in fish [2, 5, 6].

The dissolved oxygen which is available to aquatic organisms is inversely correlated with water temperature (Henry's Law) and thus leads to the potential for variability in the ability of Hb to uptake and transport oxygen under differing thermal regimes. In natural populations of aquatic organisms the genes involved in aerobic metabolism and oxygen transport, such as *ldh-b* and hemoglobin may therefore be subjected to strong selective pressure [3, 4, 7-11]. In fact, within thermally distinct populations of the temperate estuarine killifish *Fundulus heteroclitus* Hb-O_2 affinity varies together with intra-erythrocyte ATP concentrations dependent on which LDH-B isozyme (LDH-B^a or LDH-B^b) is fixed [3, 4]. Moreover, a significant difference in transcript abundance of *ldh-b* was observed in response to both thermal and aerobic stress in discrete North American *F. heteroclitus* populations, even after acclimation to a common temperature. This transcript response was linked to a one base pair mutation in the glucocorticoid responsive element (GRE) identified within the *ldh-b* 5′ proximal promoter ([7, 10]; see also reviews by [1, 2]). In addition to the extensive characterization and investigation of *ldh-b* in *F. heteroclitus*, the translated protein of this candidate gene has also been characterized and investigated in other temperate fishes like rainbow trout, *Salmo gairdneri* [12, 13] and crested blenny, *Anoplarchus purpurescens* [14]. However, this gene has not been fully characterized in any tropical perciform to date and there have been no investigations into the role this gene may have in thermal acclimation or the capacity to cope with thermal stress in tropical fishes. Therefore, the *ldh-b* locus appears to be an ideal candidate gene for the investigation of thermal adaptation and/or acclimation to native thermal regimes [15] in non-model tropical fish..

As a first step in understanding the role *ldh-b* may have in thermal adaptation or acclimation of tropical fish species, we characterized this gene in two tropical congenerics, the Australian barramundi (*Lates calcarifer*) and the African nile perch (*Lates niloticus*). *Lates calcarifer* is a catadromous, protoandrous hermaphrodite, native to rivers, estuaries and shallow marine environments throughout northern Australia (25°S – 12°S) and the south-east Asian archipelago (13°N – 10°S) [16-18], while *Lates niloticus* originates from east African rivers and lakes (7°S - 27°N) [19]. These species were targeted in this study for four reasons. Firstly, they occupy a range of different thermal environments; secondly, they are commercially valuable; thirdly, they are well suited for experimental manipulation and, fourthly, they are congeners which allows for comparisons of nucleotide sequences between these congeneric tropical perciformes. In addition to the traditional characterization of coding regions (exons) we also characterize, for the first time in non-model fish, the non-coding regions (introns) of the *ldh-b* locus in these perciform species to establish if regulatory motifs and/or elements (simple sequence repeats (SSRs) or microRNAs (miRNAs)), which are known to be embedded in or encoded by non-coding regions, are present. Previous studies on a diverse range of taxa, including fish, have demonstrated the presence of such elements within introns of other genes, where they are implicated in regulation or silencing of transcription [20-28].

Materials and methods

Complementary DNA (cDNA) synthesis from hepatic messenger RNA (mRNA)

As no *ldh-b* sequence information was available for either species, total liver RNA was first required for reverse transcription to obtain a *Lates ldh-b* cDNA sequence for primer design. This cDNA sequence was necessary to enable the design of *ldh-b* specific primers targeting the non-coding (intron) sequences of the *ldh-b* locus that can only be obtained from genomic DNA. Liver RNA was targeted for this initial cDNA sequencing as this tissue exclusively expresses the LDH-B protein as opposed to the alternative isozymes, LDH-A or C, which are expressed in other, non-hepatic tissues [29-31]. Total RNA was extracted from snap-frozen livers from four Darwin Harbour, Northern Territory, *L. calcarifer* individuals using Trizol as per manufacturer's instructions (Invitrogen Australia Pty, Mount Waverley Victoria). Extracted hepatic RNA was treated with Turbo DNA-free (Ambion, Austin, TX, USA) to remove any remaining trace DNA contamination. To ensure that trace amounts of genomic DNA, if present, do not contribute to the synthesis of cDNA, intron spanning primers were designed (detailed below). RNA purity was verified by NanoDrop Spectrophotometer (Invitrogen Australia Pty, Mount Waverley Victoria) analyses of 260/230 and 260/280nm absorbance. Messenger RNA (mRNA) was reverse transcribed to generate cDNA immediately via IM-Prom II Reverse Transcriptase with Oligo dT_{20} and random primers (Promega, Madison, WI USA), as per manufacturer's instructions.

Amplification of *ldh-b* from hepatic cDNA

Amplification and sequencing of *ldh-b* from the *L. calcarifer* hepatic cDNA was accomplished with general fish primers designed by aligning *ldh-b* coding

sequences from a taxonomically diverse range of fishes from the National Center for Biotechnology Information (NCBI) **s**equence database (GenBank) [32]. The following sequences were aligned and primers were designed based on conserved regions - Gadiformes *Trachyrincus murrayi* [AJ609235], *Merlangius merlangus* [AJ609234], *Gadus morhua* [AJ609233] *Coryphaenoides armatus* [AJ609232]; Cypriniformes *Danio rerio* and *Cyprinus carpio* [AY644476]; Cyprinodontiformes, *F. heteroclitus* [L43525], *F. heteroclitus* (D. Crawford personal communication); Squaliformes *Squalus acanthias* [AF059035]. Following the design of several primers the most specific *ldh-b* product was attained using the forward primer designed from Gadiformes and Cypriniformes (ATGGCCTGTGCCGTCAGC) in conjunction with the reverse primer designed from Cypriniformes (TCTTTCAGGTCTTTCTGGAT), which anneal in exons 2 and 7 respectively. The remaining upstream sequence (exon 1 to 5′ end of exon 2) was subsequently obtained using a previously published *ldh-b* forward primer for *F. heteroclitus* [33] and the reverse primers *L. calcarifer*-Intron2-R1 or *L. niloticus*-Intron2-R1 were designed from intron sequences subsequent to the initial cDNA amplification (Table 1). These intron sequences were obtained from genomic DNA amplification with *Lates*-specific *ldh-b* primers designed in exons 2 to 7.

Table 1. PCR and sequencing primers used to obtain *ldh-b* sequences in *Lates* species. Segments 1 to 5 refer to regions depicted in Figure 1. All primers anneal to gDNA (see Methods). Primers used for forward and reverse sequencing reactions only (nested primers) are indicated by (SeqF) and (SeqR), respectively. Final $MgCl_2$ concentrations in brackets refer to *L. niloticus* amplification requirements, which differed from that required for *L. calcarifer*.

Lates spp. ldh-b Primers	5' to 3' Sequence	$MgCl_2$	T_a	Amplicon (bp)
Segment 1*				
L. calcarifer-Seg1-R1	GATTTAGACATGTCGTTCCTCAG	1.5 - 2.5	61°C	2000
L. calcarifer-Seg1-R2	ATAATGACACCATCAATGTTCACG	1.5 - 2.5	61°C	1000
L. calcarifer-Intron1 (SeqR)	ATGGATGAATGTCTCAATCAG	1.5 - 2.5	53°C	500
L. calcarifer-Intron2-R1	TCGGATAACAGAAGCACTCAC	1.5 - 2.5	55°C	1500
L. niloticus-Intron2-R1	TAATCACTCATGGCCTCGG	1.5 - 2.5	53°C	1300
L. niloticus-Intron1-R1	AACTGGAAACTAATCTAGGCC	1.5 - 2.5	55°C	450
L. niloticus-Intron1 (SeqR)	TCAGGTTAGCACTGCTGC	1.5 - 2.5	51°C	350
Segment 2				
L. calcarifer-F1	TGATGGAGGATCGTCTGAAAGG	1.5 - 2.5	61°C	800
L. calcarifer-R1	TCGGCCATCAGGTAACGGAAG	1.5 - 2.5	61°C	800
Segment 3				
L. calcarifer-F2'	GTTGATGTGCTGACCTACGTC	1.5 - 2.5 [3.5 - 4.5]	59°C	1500
L. calcarifer-R2'	AGCCCTTCAGCTTGATCACC	1.5 - 2.5 [3.5 - 4.5]	57°C	1500
Segment 4				
L. calcarifer-F3i	ACAGAGCTTCCACTGTATCAC	1.5 - 2.5	57°C	850
L. calcarifer-R3i	GCAAAAGGTTCCTAGGCATGTA	1.5 - 2.5	59°C	850
L. calcarifer-F3	AGAAGCTGAACCCTGAGATCG	1.5 - 2.5 [3.5 - 4.5]	59°C	800
L. calcarifer-R3	TTCTGGATGCCCCACAGTGTG	1.5 - 2.5 [3.5 - 4.5]	61°C	800
Segment 5**				
L. calcarifer-F3i'	TGGTTGCTAGGATAAAGAATGTG	1.5 - 2.5	59°C	700
L. calcarifer-F3i (SeqF)	AGTTGTAAATAATTCAGGCATC	1.5 - 2.5	53°C	500
L. niloticus-Intron6-F1	ATGTGGATAGCCTAGCTTAGC	1.5 - 2.5	55°C	400

* and **: Published forward and reverse *F. heteroclitus ldh-b* primers [33] used in conjunction with designed primers to amplify terminal (5′ and 3′) segments in both *Lates* species, respectively.

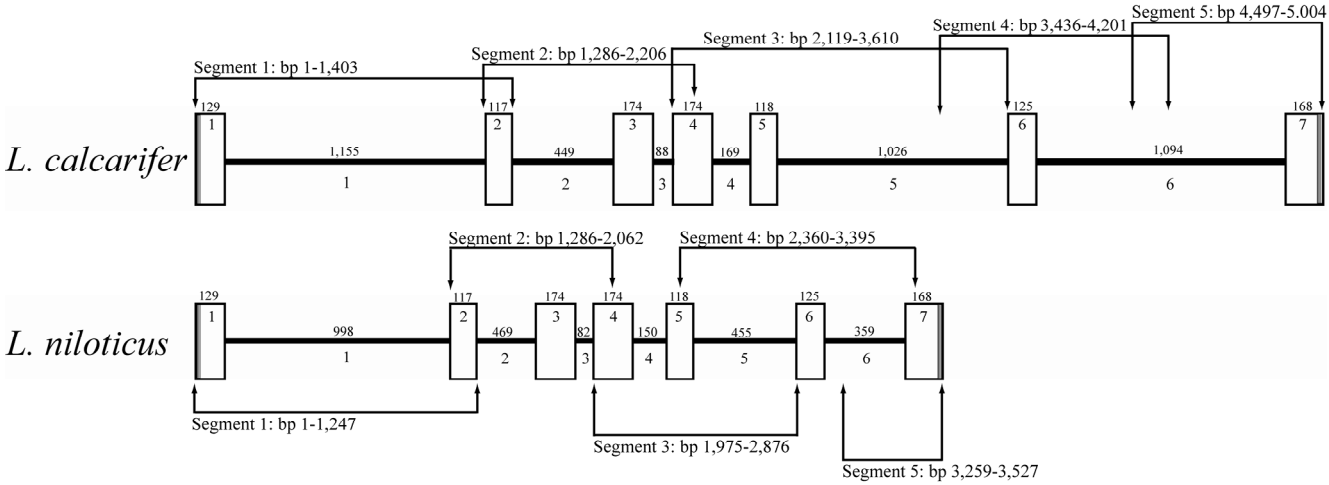


Fig. 1. Comparative *ldh-b* gene map of congeneric *Lates* species. Exon (white boxes) and intron (black bars) sizes in number of base pairs (bp) are given above their respective graphic representation. Sequential numbering within white boxes and below black bars is for exons and introns, respectively. Arrowhead lines indicate region amplified by specific primer pairs (as per Table 1) along with the size of each segment in number of base pairs.

All PCR reactions were conducted in the following manner: Amplification reactions (20 µL) contained the following final concentrations: 1X Buffer [2.5 mM Tris pH 8.7, 5 mM KCl, 5 mM $(NH_4)_2SO_4$, containing 1.5 mM $MgCl_2$] (Qiagen, Doncaster, Victoria Australia) or 1x Buffer [2.5 mM Tris pH 8.7, 5 mM KCl, 5 mM $(NH_4)_2SO_4$ not containing 1.5 mM $MgCl_2$] (Bioline Pty Ltd., Alexandria, New South Wales Australia) (unless more was required as per Table 1)], 250 µM each dNTP, 250 nM each primer (Table 1), 10 ng gDNA template and 0.75 to 1.5 units of *Taq* Polymerase (Qiagen and Bioline Pty Ltd.). Thermal cycling was conducted on a MJR DNA Engine thermal cycler (Bio-Rad Laboratories Pty., Ltd., Gladesville, New South Wales) as follows: initial denaturation at 94°C for 3 min followed by 35 cycles of 94°C denaturation for 30 sec, annealing at primer specific T_a for 30 sec (Table 1), 72°C extension for 30 to 90 sec depending on target fragment size with larger fragments (> 1,000 bp) requiring longer (> 60 sec) extension times (Table 1) and a final 72°C extension for 10 min. Melting temperature (T_m) was calculated via (A/T x 2 + G/C x 4) method, with the annealing temperature (T_a) set at 5°C less than T_m, for all primer-pairs (Table 1).

Verification of *ldh-b* sequences amplified from hepatic cDNA

Ldh-b amplification from hepatic RNA derived cDNA with *Lates*-specific primers (see above) generated a strong single-band product of approximately 800 bp for all *L. calcarifer* examined (n = 4). Subsequent amplification of genomic DNA with the published forward (*F. heteroclitus*-F, [33]) and designed reverse (*L. calcarifer*-Intron2-R1 and *L. niloticus*-Intron2-R1) primers resulted in a strong single-band product of approximately 1,500 bp and 1,300 bp (*L. calcarifer* and L*. niloticus*, respectively) containing the missing exon 1 and 5′ end of exon 2 fragment. These products were precipitated with 120 µL isopropanol (70%) for 15 minutes followed by a 500 µL wash with 70% isopropanol prior to drying, re-suspension in water (10 µL) and subsequent sequencing (Macrogen, Inc., South Korea). Sequences were edited in BioEdit [34] and a contig made to give the full length coding sequence (1,005 bp) and produce a consensus cDNA sequence. To check that the correct *ldh* gene homologue had been obtained the sequence was used in a BLAST search of GenBank and also directly aligned using ClustalW in Mega 3.1 [35] to those *ldh-b* sequences previously utilized for general fish primer design (see above). The obtained sequence shared 92% homology with *F. heteroclitus* LDH-B amino acid sequence, a level of homology that is well above that reported for *F. heteroclitus* LDH-C *v.* LDH-B and LDH-C *v.* LDH-A (78% and 70% homology, respectively) [30]). As a further check that the correct *ldh* gene homologue was obtained all nucleotide (and deduced amino acids) sequences were aligned with those of *Danio rerio ldh-b* [AF067202] and *ldh-a* [NM_131246] genes. These two gene homologues differ by 600 bp in length and 34.1% of the nucleotide sequences, respectively. The *L. calcarifer ldh-b* sequence obtained from hepatic cDNA most closely matched that of the *D. rerio ldh-b* gene homologue. Once the initial cDNA sequence was obtained from *L. calcarifer* the full characterization of the *ldh-b* locus from both *Lates* species (*L. calcarifer* and *L.*

niloticus) was accomplished via primer walking along genomic DNA from representative individuals of each species.

Study species and genomic DNA (gDNA) extraction

Genomic DNA was extracted from *L. calcarifer* samples collected from four locations within tropical Australia. Samples from Gladstone, Queensland (23°S, 151°E) and Darwin, Northern Territory (12°S, 130°E) were obtained directly from fish farms while Archer River (Cape York, Queensland: 13°S, 142°E) and Tully River (Tully, Queensland: 17°S, 145°E) samples originated from wild caught fish. *L. niloticus* was purchased as two imported frozen fillets at a local supermarket in Townsville, Queensland and therefore the exact geographical origin of the *L. niloticus* samples examined is unknown but assumed to be from one of the African Rift Valley lakes (4°N – 14°S) where a large export fishery of this species exists. *L. niloticus* fillets may or may not have been snap-frozen; regardless, extracted genomic DNA appeared to be of equal quality to the DNA from *L. calcarifer* samples. Fin-clips/muscle tissue were taken from all fish and preserved in ethanol (90%). DNA extractions were performed via proteinase-K digestion (20 mg/mL) in CTAB buffer at 60°C for 1 hr and DNA was subsequently cleaned with a salt and chloroform:isoamyl alcohol (24:1) procedure [36]. All extractions resulted in high molecular weight gDNA, as visualized on a 0.8% agarose gel, with quantities ranging from 20 – 100 ng/μL.

Amplification of the *ldh-b* locus from genomic DNA

Full length *ldh-b* gene sequences (including intron sequences) were obtained from genomic DNA extracts using the primers and primer specific PCR conditions outlined in Table 1 and Figure 1. In some cases different primers were required for amplification of *L. niloticus* and *L. calcarifer* introns (e.g. intron 5 & 6 in Figure 1). See Figure 1 for primer binding locations within the *ldh-b* locus. Thermocycling parameters were the same as that used for cDNA amplification with the exception of the annealing temperature and $MgCl_2$ concentration which varied as listed in Table 1. All PCR products obtained from both *L. calcarifer* and *L. niloticus* genomic DNA were verified as *ldh-b* in two ways: 1) searching the NCBI database (GenBank) via nucleotide (blastn) and protein (blastp) basic local alignment and search tool (b.l.a.s.t.) with resulting matches specific to LDH-B exclusively and 2) direct alignment with *F. heteroclitus ldh-b* nucleotide and LDH-B protein sequence. After confirmation of product identity full length gene sequences for each individual (*L. calcarifer: n* = 4 and *L. niloticus: n* = 2) were assembled in BioEdit and the consensus nucleotide and deduced amino acid sequences from each species were aligned using ClustalW in Mega 3.1. The high level of conservation across individuals from both *Lates* species allowed for the compilation of *L. calcarifer* and *L. niloticus* consensus sequences. Consensus gene sequences for each species were submitted to GenBank under accession numbers [**FJ439509**] and [**FJ439510**] for *L. calcarifer* and *L. niloticus* respectively. Additionally, *L. calcarifer ldh-b* coding sequence isolated from hepatic mRNA was submitted to GenBank under accession number [**FJ439507**]. Deduced LDH-B amino acid sequences were manually assessed for the presence of variation within the NH_2-terminal arm (residues 1-20), coenzyme binding domain (residues 21-95 and 118-163), substrate binding domain (residues 164-333) and loop helix ά D region (residues 96-117) [37].

Assessment of non-coding (intron) sequences for micro RNA (miRNA) and simple sequence repeat (SSR) motifs

Several recent studies have demonstrated intron sequences may contain simple sequence repeat (SSR) motifs of functional importance as they potentially bind regulatory machinery (e.g. promoters and/or enhancers) and may affect gene expression levels [25, 38]. Intron sequences from *ldh-b* of both *Lates* species were manually assessed for the presence of simple sequence reseat (SSR) motifs. In addition, short microRNA (miRNA) elements (21-23 bp) may be encoded by introns and these are believed to be spliced out of pre-messenger RNA (mRNA) subsequently targeting regions within the 3'UTR of actively expressed mRNAs regulating translation from mRNA transcripts to functional proteins [20-23, 39, 40]. Intron sequences from *Lates* spp. were therefore assessed for putative miRNA elements with the software package miRanda [40]. All presented miRNA elements are located within conserved intron regions of *Lates* spp. and matched known miRNA motifs from *Danio rerio* (zebrafish), *Takifugu rubripes* (Japanese pufferfish) and *Tetraodon nigroviridis* (spotted green pufferfish) (miRBase: [41-43]). Several of the identified elements also matched known *Xenopus tropicalis* (African frog) miRNA motifs (miRBase: [41-43]). A nucleotide match score (score), affinity to bind measure (energy [kcal mol^{-1}]), statistical assessment of match quality (z-score) and homology of motif to query intron sequence (percentage) were all calculated by miRanda for each matching miRNA element [40]. Threshold values were set to 100, -19 kcal mol^{-1} and 5.0 for score, energy and z-score, respectively, to assess *ldh-b* for

miRNAs under strict parameters so as to avoid false-positive identifications (≤ 5%) [40, 44, 45]. However, overly-stringent thresholds were avoided to maximize the likelihood of identifying putative miRNA as such elements have recently been documented in high abundance within eukaryotic genomes at an average of 100 binding sites per miRNA element, genome-wide [45].

Results and Discussion

Descriptive characterization of *ldh-b* coding nucleotide (exon) and deduced amino acid sequences

Variation in the primary sequence and level of gene expression of the *ldh-b* locus has been linked to differences in aerobic performance (reviewed by [1, 2]) and natural selection [7, 46] in the temperate (to sub-tropical) fish *F. heteroclitus*. The present study is the first to characterize the full length gene sequence of this important metabolic gene in two tropical perciform fish, namely *L. calcarifer* and *L. niloticus*. This locus consisted of 5,004 and 3,527 bp, of which 1,005 bp was coding, in these two species, respectively (Figure 1). Consistent with other fish species the *ldh-b* locus of both *L. calcarifer* and *L. niloticus* contained seven exons and six introns, with all seven *ldh-b* exons being conserved in size between the characterized *Lates* species (Figure 1). Additional coding nucleotide sequence comparison revealed 29 base differences between *L. calcarifer* and *L. niloticus* (2.9% divergence); however, the majority of these occurred in the third codon positions which lead to silent (i.e. synonymous) amino acid substitutions. This high level of conservation was expected between *Lates* species being that this locus is known to be under functional constraint due to it role in maintaining aerobic metabolism ([7, 46]; see also review by [2]). Moreover, the encoded enzyme (LDH-B) has also been shown to impact on hemoglobin-oxygen binding affinity in fish, with allozyme variants altering this critical interaction [3, 4].

Each residue of the post-translation modified protein sequences has a role in the specific functioning of the molecule, from its internal stability to its external functional interactions. Five amino acid differences (1.5%) are present between the LDH-B of both *Lates spp*. These variable amino-acids involve the substitution of threonine (T) for methionine (M), leucine (L) for alanine (A), valine (V) for isoleucine (I), lysine (K) for asparagine (N) and valine (V) for isoleucine (I) at residues 34, 35, 126, 127 and 147, respectively (Figure 2). All of these variable amino acid residues reside within the coenzyme (NADH) binding domain (residues 21-95 and 118-163), of LDH-B, which is otherwise conserved between *Lates* spp. (Figure 2). This is unanticipated in such a conserved functional domain [30, 37]. Noteworthy are the two shifts occurring at residues 34 and 127 as these invoke changes in polarity (polar – non-polar) and acidity (neutral polar – basic polar) between *Lates* species, respectively (Figure 2). The effect of these amino acid shifts on the catalytic efficiency (k_{cat}) of LDH-B between *L. calcarifer* and *L. niloticus* is unknown and warrants further investigation. Conversely, the NH_2-terminal arm (residues 1-20), substrate binding domain (residues 164-333) and loop helix ά D region (residues 96-117) [37], as well as the LDH-B substrate binding (residues 100, 107, 139, 170, 249) and proton acceptor (residue 194) sites found within [47], are conserved between both *Lates* species (Figure 2), as expected [30, 37].

Comparison between *Lates* species and *F. heteroclitus* full length LDH-B amino acid sequences revealed a relatively extensive divergence, with 30 amino acids (9.0% divergence) observed between these phylogenetically distant species (data not shown). The two residues demonstrated to have fixed differences between thermally and geographically distinct populations of *F. heteroclitus* were serine (S) *v.* alanine (A) and alanine (A) *v.* aspartic acid (D) at residues 185 and 311 for cold northern and warm southern populations, respectively (see review by [1, 2]). *L. calcarifer* and *L. niloticus* both possess serine (S) and aspartic acid (D) at residues 185 and 311, respectively (Figure 2). The deduced LDH-B amino acid sequence of both *Lates* species examined are therefore similar to the cold northern *F. heteroclitus* population at residue 185 and to the warm southern *F. heteroclitus* population at residue 311. These residues are located on the internal and external surfaces of the folded LDH-B protein, respectively ([48, 49]; see also reviews by [1, 2]). The former change (residue 185) has been hypothesized to be associated with a variation in thermal stability due to this residue being located at a hairpin turn in the center of the folded protein; whereas the latter change (residue 311) has been hypothesized to be associated with a variation in substrate binding affinity due to this residue being located on the external surface of the conformed protein ([48, 49]; see also reviews by [1, 2]). In light of these previous hypothetical explanations, future studies should strive to compare the structure and function (e.g. enzymatic activity, effect on Hb-O_2 binding affinity) of LDH-B in the *Lates* species in parallel to those studies conducted on within and among thermally discrete *F. heteroclitus* populations [4, 48, 49] and cod from two different temperatures [50].

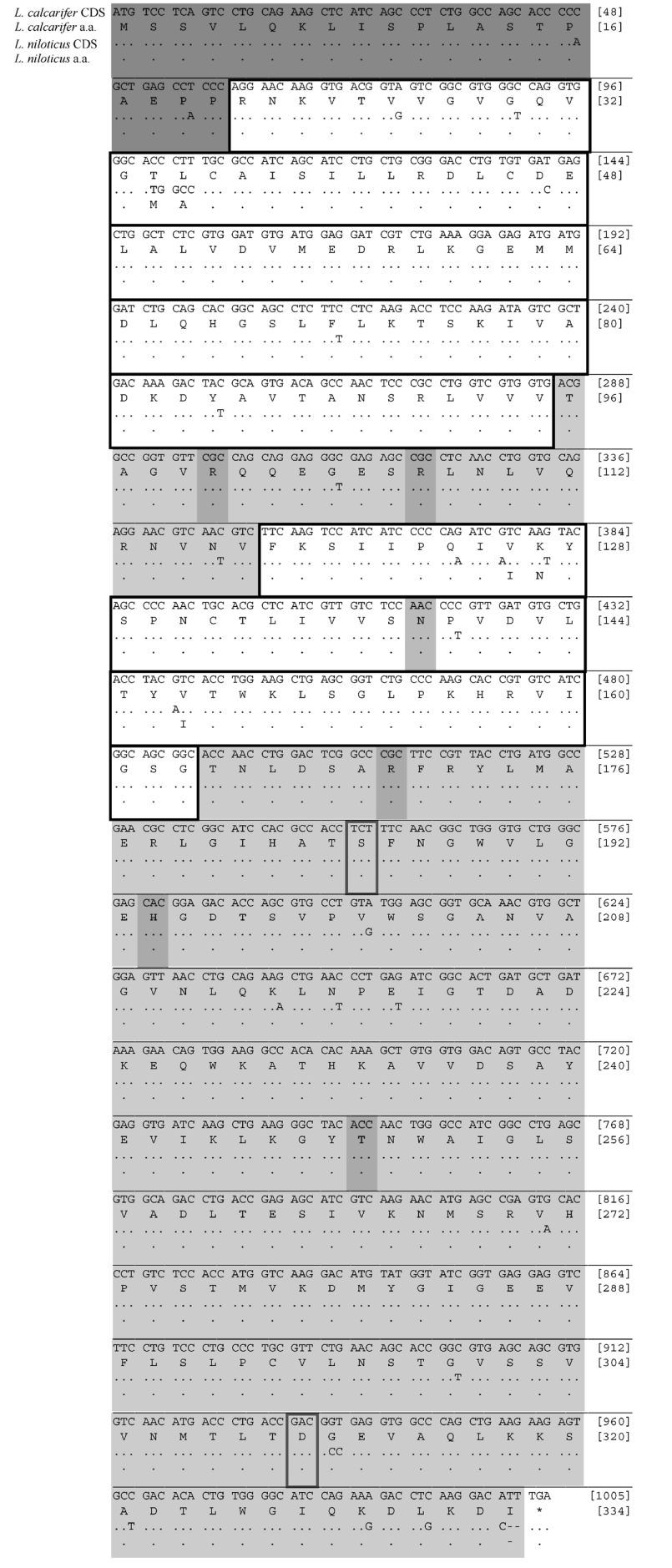


Fig. 2. Coding nucleotide sequence (CDS) for *L. calcarifer* and *L. niloticus* are given on lines 1 and 3, respectively. Deduced amino acid sequence (a.a.) for *L. calcarifer* and *L. niloticus* are given on lines 2 and 4, respectively. Bracketed numbers at ends of line 1 and 2 refer to nucleotide position and amino acid residue, respectively. NH_2-terminal arm (residues 1-20) is dark-gray box outlined. Coenzyme binding domains (residues 21-95 and 118-163) are boldface box outlined. Substrate binding domain (residues 164-333) and loop helix ά D region (residues 96-117) are light-gray box outlined. Other known functional sites (residues 100, 107, 139, 170, 194 and 249) within domains are medium-gray box outlined. Residues (185 and 311) of known fixed difference between *F. heteroclitus* populations are boldface box outlined (see Results).

Descriptive characterization of *ldh-b* introns and identification of embedded putative regulatory motifs and/or elements

Nucleotide sequence comparisons of the homologous introns between the two *Lates spp*. characterized reveals a high level of homology (97.7%), a level surprisingly similar to that observed between exons of the *Lates* species (97.1% exon homology). The fact that sequence homology between non-coding introns is similar to that of coding exons suggests similar selective constraint (i.e. functionality) may be driving the high levels of homology observed across these historically non-characterized intronic regions [51, 52]. Interestingly, *L. niloticus* introns are consistently smaller than *L. calcarifer* introns with the exception of intron 2, which is 20 bp longer in *L. niloticus* relative to *L. calcarifer* (Figure 1). Moreover, one insertion-deletion event (indel) occurs between the two *Lates* species in all introns, with indels ranging in size across introns 1 to 6 (157, 20, 6, 19, 571 and 735 bp, respectively) (Figure 1). Whether these indels are historical insertion or deletion events and whether putative regulatory elements embedded within indel regions have an impact on the transcriptome or functional proteome of either *Lates* species is unknown. Therefore, fu-

ture investigations into the impact these indels may or may not have on *L. niloticus ldh-b* gene expression, as compared to the variation observed among thermally discrete *L. calcarifer* populations [53], is warranted.

Numerous simple sequence repeat (SSR) sequences were detected within *ldh-b* introns of both *Lates* species (Table 2). Intronic SSRs are of interest because they may regulate gene transcription, lead to abnormal splicing and disrupt export of mRNA to the cytoplasm [25, 38, 52, 54]. Four mononucleotide SSRs (T, C and A), ranging from 5 to 11 repeats, are present within intron 1 of both *Lates spp*. and these were conserved in size between species (Table 2). In addition to these, two mononucleotide repeats (T_{10} and $C_{7\text{-}11}$) are also present exclusively in *L. calcarifer* introns 1 and 6, respectively (Table 2). One dinucleotide repeat (AC) is present at the same location within intron 6 of both *L. calcarifer* and *L. niloticus*, but it differs in repeat number between the species (AC_5 and AC_8, respectively) (Table 2). In addition there is one trinucleotide repeat (CAA) present within intron 1 of *L. niloticus* and *L. calcarifer* which varies between 3 and 4 repeats, respectively (Table 2). Another trinucleotide repeat (TCC_4) is present within a region in intron 2 and is conserved between both *Lates* species (Table 2). Noteworthy is that a similar (TCC_4) SSR was identified in the 5′ flanking untranslated region (UTR) of the *ldh-b* locus in *F. heteroclitus* and, more importantly, that variation in repeat number of this SSR in the 5′ UTR of *ldh-b* impacted the level of *ldh-b* transcription (i.e. gene expression) observed in thermally discrete *F. heteroclitus* populations [8, 10, 11, 29, 55]. Moreover, this region within the 5′ proximal promoter of *F. heteroclitus*, in addition to the 6fp and Sp1 binding sites also identified, were concluded to be under functional constraint by way of a phylogenetic analysis on the nucleotide sequences of these 5′ UTR regulatory motifs which clearly differentiated the cold northern from the warm southern population [7, 46]. Lastly, a tetranucleotide repeat ($TGTA_4$) is observed in a region of intron 6 exclusive to *L. calcarifer* (Table 2). The variation and potential functional role, if any, of these SSRs on *ldh-b* gene expression itself and/or on the greater transcriptome functionality within and among thermally discrete *L. calcarifer* and *L. niloticus* populations is currently unknown; however, further investigation into such potential implications on the transcriptome is warranted.

Numerous potential microRNA elements (miRNAs) were also identified within *Lates* species intron sequences (Table 3). The encoding of thirty putative miRNA elements was identified within conserved intron regions and these had a score, energy (kcal mol^{-1}), z-score and homology ranging from 102 to 140, -26.96 to -19.01, 5.12 to 14.94 and 61.76 to 87.5%, respectively (see Methods). Four of the putative miRNA elements identified (Table 3: *miR-let7b*, *miR-124, miR-181* and *miR-223*) have previously been associated with specific regulatory functions [21, 39, 56-59]. Of these, *miR-let7b* and *miR-223* knockout in mouse HeLa cells [59] and neutrophils [21] resulted in reduced expression of approx. 2,700 and approx. 3,800 proteins, respectively. Two additional sub-families (*d* and *e*) from the highly investigated *let-7* miRNA element family were identified within *Lates* spp. *ldh-b* introns (see Table 3), which provides additional data consistent with the widespread dispersal and high abundance of targets and/or functions of miRNA elements encoded by introns throughout the genome [21, 22, 45, 59]. Noteworthy is that the *dre-let-7 (-b, -c, -d)* miRNA elements have also been recently identified within the 5′ UTR of the *L. calcarifer* myostatin gene (*mstn*) [60].

Five intronic single nucleotide polymorphisms (SNPs) also were identified in intronic regions of the *ldh-b* locus of individuals representing eight discrete *L. calcarifer* populations in a pilot population genetics screening [53]. Interestingly, these SNPs are present at sites where multiple miRNA elements overlap [53]. The fact that a single SNP can impact multiple putative miRNA elements concurrently may provide insight in regard to recent findings of genome-wide association studies, which show a relationship between such SNPs and variation in complex human behaviour (e.g. schizophrenia and bipolar disorder) and/or susceptibility to complex diseases (e.g. type 2 diabetes and Crohn's disease) [61-65].

Table 2. Simple sequence repeat (SSR) motifs present within non-coding intron sequences of *Lates* species.

SSR Motifs	Location	# of Repeats: *L. calcarifer*	# of Repeats: *L. niloticus*
Mononucleotide			
T^	Intron 1	10	-
C	Intron 1	6	5
C	Intron 1	8	5
T	Intron 5	8	5
A	Intron 6	7	8
C^	Intron 6	7 - 11	-
Dinucleotide			
AC	Intron 6	5	8
Trinucleotide			
CAA	Intron 1	4	3
TCC*	Intron 2	4	4
Tetranucleotide			
TGTA^	Intron 6	4	-

^: Indicates repeats exclusive to *L. calcarifer*. *: Indicates a simple sequence repeat motif identified within the 5′ proximal promoter of *ldh-b* in *F. heteroclitus*, of which the repeat number impacts *ldh-b* expression [7, 11].

Table 3. Putative microRNA (miRNA) motifs identified within *ldh-b* intron sequences of *Lates* spp. Score: nucleotide match score; Energy: affinity to bind measure; Z-Score: statistical assessment of match quality; Homology: motif homology to query intron sequence.

miRNA Family	Sub-families	Score	Energy (kcal mol^{-1})	Z-Score	Homology (%)
let-7	b, d, e	104 to 128	-22.65 to -19.17	5.12 to 11.04	65.52 to 86.36
miR-10	b, c, d	102 to 127	-24.50 to -19.01	5.75 to 10.02	65.38 to 76.00
miR-15	a	109 to 128	-21.88 to -20.64	7.29 to 12.20	66.67 to 82.61
miR-23	a, b	113 to 128	-20.19 to -19.53	10.02 to 12.92	75.00 to 80.00
miR-24	-	105 to 112	-19.81 to -19.29	6.72 to 8.75	75.00 to 80.00
miR-25	-	113 to 122	-25.79 to -20.77	8.52 to 10.96	72.41 to 85.71
miR-27	c, e	106 to 130	-24.44 to -19.36	6.91 to 13.77	76.92 to 83.33
miR-101	a	109 to 110	-20.02 to -19.23	7.18 to 7.91	68.97 to 73.91
miR-103	-	111 to 114	-22.68 to -19.15	6.84 to 8.19	73.08 to 87.50
miR-107	-	104 to 110	-21.16 to -20.59	5.56 to 7.31	67.86 to 70.37
miR-122	-	109 to 130	-24.44 to -19.67	6.35 to 12.72	64.71 to 87.50
miR-124	-	126 to 133	-21.18 to -19.68	11.54 to 13.73	71.43 to 72.41
miR-125	a, b	102 to 118	-21.84 to -19.47	6.10 to 10.16	64.00 to 76.92
miR-138	-	106 to 117	-20.40 to -19.19	7.11 to 9.45	65.62 to 81.82
miR-140	-	104 to 115	-23.55 to -19.32	6.17 to 9.55	70.00 to 79.17
miR-148	-	109 to 115	-23.74 to -19.82	7.21 to 9.3	86.36
miR-152	-	102 to 140	-25.95 to -19.30	5.75 to 14.94	69.57 to 84.00
miR-181	a, b	103 to 112	-21.00 to -19.61	5.25 to 8.89	61.76 to 76.00
miR-183	-	118 to 127	-26.75 to -19.73	8.69 to 11.3	68.75 to 77.78
miR-184	-	101 to 128	-24.99 to -20.76	5.73 to 12.28	70.37 to 80.00
miR-187	-	103 to 122	-22.26 to -19.02	6.93 to 13.45	73.08 to 81.82
miR-196	a, b	107 to 127	-22.22 to -19.98	6.51 to 11.52	73.08 to 83.33
miR-202	-	107 to 126	-24.76 to -19.41	7.42 to 11.74	68.97 to 82.61
miR-210	-	102 to 127	-26.96 to -19.62	6.08 to 12.54	68.00 to 82.61
miR-214	-	102 to 114	-21.45 to -19.68	6.78 to 9.93	70.37 to 77.27
miR-216	b	111 to 112	-19.16 to -19.15	7.98 to 8.68	72.41 to 77.78
miR-217	-	112 to 132	-26.17 to -19.47	7.67 to 12.26	70.37 to 84.00
miR-221	-	105 to 133	-21.67 to -19.42	5.59 to 12.58	70.00 to 81.48
miR-222	-	105 to 127	-25.88 to -19.36	5.13 to 10.61	70.00 to 76.92
miR-338	-	121	-19.65	10.32 to 11.04	75.00

Evidence for selective constraint on non-coding (intron) sequences

The putative encoding of known functional motifs and/or elements (SSRs and miRNAs) by intronic regions of loci, in addition to the even distribution of such motifs and/or elements throughout longer introns [66], provides evidence that functional constraint is arguably acting on these historically less characterized non-coding regions. Indeed, such functional constraint is more likely to occur for introns of loci whose coding nucleotide or amino acid sequences are known to be under selection [22, 51, 52]. More specifically, longer introns (> 87 bp) have been found to exhibit less divergence than shorter introns (< 87 bp) because of either an increased likelihood of embedded functional motifs being present in longer introns or the potential impact any mutations may have on the secondary structure of precursor messenger RNA (pre-mRNA) [66]. A recent pairwise and cross-taxa comparison of intron sequences between three mammalian species (human, whale and seal) revealed sequence homologies 14% and 12% higher, respectively, than those expected from a neutral model of evolution based on expected rates of substitution for non-coding DNA, further suggesting functional constraint acting upon non-coding intronic sequences [51]. Moreover, the existence of numerous intron motifs and/or elements is known to be essential for the functioning of complex multi-cellular organisms, as they permit a two-fold regulatory system in eukaryotic organisms: one for the transcriptome [20, 22, 25, 26, 28, 67] and one for the proteome [21, 39, 59]. Regardless, further research is required to determine if these motifs and/or elements (SSRs and/or miRNAs) have similar target sites and/or impacts on the transcriptome (i.e. gene expression) and/or the proteome (i.e. gene silencing) within and among *Lates* species and/or fish in general.

Conclusions

The *ldh-b* locus was found to be highly conserved between two tropical perciformes, *L. calcarifer* and *L. niloticus,* with just 2.9% divergence of coding regions and five amino acid differences between deduced LDH-B protein sequences. Variation within the co-enzyme binding domains (residues 21-95 and 118-163) of the deduced protein sequences may possibly confer a variation in the specific catalytic efficiency of LDH-B in geographically isolated *L. calcarifer* and *L. niloticus* and warrants further exploration. Non-coding (intron) sequences of the *ldh-b* locus in both *Lates* species were as conserved as the coding regions of this gene, despite comprising 72.5 to 80% of the entire gene sequence. Ten SSR motifs and thirty putative miRNA elements were observed within the introns of the *ldh-b* locus in both *Lates* species. These putative regulatory elements and/or motifs warrant further investigation for their potentially functional importance in the regulation and/or expression of *ldh-b* or other constituent loci contributing to the transcriptome and proteome. The characterization of the *ldh-b* locus has spawned additional studies, one of which confirmed that *ldh-b* expression differs significantly between *L. calcarifer* populations from different temperature environments, suggesting that *ldh-b* is adaptive in this species [53], as was shown for the temperate (to sub-tropical) killifish *Fundulus heteroclitus* [7, 46]. These findings, in conjunction with a pilot population genetics screening of the *ldh-b* locus among and within discrete *L. calcarifer* populations [53], are suggestive of thermal adaptation occurring within this tropical estuarine species. Moreover, future studies on other non-model tropical species are also permitted by the *ldh-b* characterization presented herein. Indeed, future studies of this adaptive locus in other non-model tropical species will provide insight into the response of thermally sensitive species (living at the edge of their thermal maxima) to changes in their native temperatures, which are expected to increase by up to 6°C over the next century.

Acknowledgements

Thanks to Paolo Momigliano and Dianne Rowe for help in the Molecular Evolution and Ecology lab, Christian De Santis for *L. niloticus* gDNA samples, Manue Botte for RNA precipitation protocols, the Kuranda Fish Farm and Darwin Aquaculture Center for *L. calcarifer* fingerlings and the James Cook University Research Advancement Program for Fin-fish Genetics (RAP) for funding.

Conflict of Interest

The authors have declared that no conflict of interest exists.

References

1. Powers D.A, Lauerman T., Crawford D., et al. Genetic Mechanisms for Adapting to a Changing Environment. Annu Rev Genet 1991; 25: 629-659.
2. Powers D.A and Schulte P.M. Evolutionary Adaptations of Gene Structure and Expression in Natural Populations in Relation to a Changing Environment: A Multidisciplinary Approach to Address the Million-Year Saga of a Small Fish. J Exp Zool Comp Exp Biol 1998; 282: 71-94.
3. Powers D.A. Molecular Ecology of Teleost Fish Hemoglobins: Strategies for Adapting to Changing Environments. Am Zool 1980; 20: 139-163.
4. Powers D.A, Greaney G.S and Place A.R. Physiological correlation between lactate dehydrogenase genotype and haemoglobin function in killifish. Nature 1979; 277: 240-241.
5. DiMichele L and Powers D.A. Physiological Basis for Swimming Endurance Differences between LDH-B Genotypes of Fundulus heteroclitus. Science 1982; 216: 1014-1016.
6. Fangue N.A, Mandic M., Richards J.G., et al. Swimming Performance and Energetics as a Function of Temperature in Killifish Fundulus heteroclitus. Physiol Biochem Zool 2008; 81: 389–401.
7. Crawford D.L, Pierce V.A and Segal J.A. Evolutionary Physiology of Closely Related Taxa: Analyses of Enzyme Expression. Am Zool 1999; 39: 389-400.
8. Crawford D.L, Segal J.A and Barnett J.L. Evolutionary Analysis of TATA-less Proximal Promoter Function. Mol Biol Evol 1999; 16: 194-207.
9. Rees B.B, Bowman J.A.L and Schulte P.M. Structure and Sequence Conservation of a Putative Hypoxia Response Element in the Lactate Dehydrogenase-B Gene of Fundulus. Biol Bull 2001; 200: 247-251.
10. Schulte P.M, Glemet H.C., Fiebig A.A., et al. Adaptive variation in lactate dehydrogenase-B gene expression: Role of a stress-responsive regulatory element. Proc Nati Acad Sci USA 2000; 97: 6597-6602.
11. Segal J.A, Schulte P.M., Powers D.A., et al. Descriptive and Functional Characterization of Variation in the Fundulus heteroclitus Ldh-B Proximal Promoter. J Exp Zool Comp Exp Biol 1996; 275: 355-364.
12. Kao Y.-H.J and Farley T.M. Purification and properties of allelic lactate dehydrogenase isozymes at the B2 locus in rainbow trout, Salmo gairdneri. Comp Biochem & Physio Pt B 1978; 61: 507-512.
13. Klar G.T, Stalnaker G.B and Farley T.M. Comparative blood lactate response to low oxygen concentrations in rainbow trout, Salmo gairdneri, LDH-B2 phenotypes. Comp Biochem & Physio Pt A 1979; 63: 237-240.
14. Johnson M.S. Association of allozymes and temperature in the crested blenny Anoplarchus purpurescens. Mar Biol 1977; 41: 147-152.
15. Cossins A.R and Crawford D.L. Opinion: Fish as models for environmental genomics. Nat Rev Genet 2005; 6: 324-333.
16. Chenoweth S.F, Hughes J.M., Keenan C.P., et al. When Oceans Meet: A Teleost Shows Secondary Intergradation at an Indian-Pacific Interface. Proc R Soc Lond B Biol Sci 1998; 265: 415-420.
17. Doupé R.G, Horwitz P and Lymbery A.J. Mitochondrial genealogy of Western Australian barramundi: applications of inbreeding coefficients and coalescent analysis for separating temporal population processes. J Fish Biol 1999; 54: 1197-1209.

18. Keenan C.P. Recent Evolution of Population Structure in Australian barramundi, Lates calcarifer (Bloch): An Example of Isolation by Distance in One Dimension. Aust J Mar Freshwat Res 1994; 45: 1123-1148.
19. Ribbink A.J. African lakes and their fishes: conservation and suggestions. Environ Biol Fish 1987; 19: 3-26.
20. Ambros V. The functions of animal microRNAs. Nature 2004; 431: 350-356.
21. Baek D, Villen J., Shin C., et al. The impact of microRNAs on protein output. Nature 2008; 455: 64-73.
22. Bartel D.P. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. Cell 2004; 116: 281-298.
23. Doench J.G and Sharp P.A. Specificity of microRNA target selection in translational repression. Genes & Dev 2004; 18: 504-511.
24. Lai E.C. microRNAs: Runts of the Genome Assert Themselves. Curr Biol 2003; 13: R925-R936
25. Li Y.-C, Korol A.B., Fahima T., et al. Microsatellites Within Genes: Structure, Function, and Evolution. Mol Biol Evol 2004; 21: 991-1008.
26. Meloni R, Albanèse V., Ravassard P., et al. A tetranucleotide polymorphic microsatellite, located in the first intron of the tyrosine hydroxylase gene, acts as a transcription regulatory element in vitro. Hum Mol Genet 1998; 7: 423-428.
27. Reisman D, Greenberg M and Rotter V. Human p53 oncogene contains one promoter upstream of exon 1 and a second, stronger promoter within intron 1. Proc Nati Acad Sci USA 1988; 85: 5146-5150.
28. Wittekindt N.E, Hortnagel K., Geltinger C., et al. Activation of the c-myc promoter P1 by immunoglobulin  gene enhancers in Burkitt lymphoma: functional characterization of the intron enhancer motifs B, E box 1 and E box 2, and of the 3' enhancer motif PU. Nucleic Acids Res 2000; 28: 800-808.
29. Crawford D.L and Powers D.A. Evolutionary Adaptation to Different Thermal Environments via Transcriptional Regulation. Mol Biol Evol 1992; 9: 806-813.
30. Quattro J.M, Woods H.A and Powers D.A. Sequence analysis of teleost retina-specific lactate dehydrogenase C: Evolutionary implications for the vertebrate lactate dehydrogenase gene family. Proc Natl Acad Sci USA 1993; 90: 242-246.
31. Segal J.A and Crawford D.L. LDH-B enzyme expression: the mechanisms of altered gene expression in acclimation and evolutionary adaptation. Am J Physiol Regulatory Integrative Comp Physiol 1994; 267: 1150-1153.
32. Zhang J and Madden T.L. PowerBLAST: A new network BLAST application for interactive or automated sequence analysis and annotation. Genome Res 1997; 7: 649-656.
33. Bernardi G, Sordino P and Powers D.A. Concordant mitochondrial and nuclear DNA phylogenies for populations of the teleost fish Fundulus heteroclitus. Proc Natl Acad Sci USA 1993; 90: 9271-9274.
34. Hall T.A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 1999; 41: 95-98.
35. Kumar S, Tamura K and Nei M. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Briefings in Bioinformatics 2004; 5: 150-163.
36. Sambrook J, Fritsch E and Maniatus T. Molecular cloning: a laboratory manual. New York, USA: Cold Spring Harbour Laboratory Press; 1989.
37. Li S.S, Fitch W.M., Y.- Pan C.P., et al. Evolutionary Relationships of Vertebrate Lactate Dehydrogenase Isozymes A4 (Muscle), B4, (Heart), and C4, (Testis). J Biol Chem 1983; 258: 7029-7032.
38. Yue G, Li Y and Orban L. Characterization of Microsatellites in the IGF-2 and GH Genes of Asian Seabass (Lates calcarifer). Mar Biotechnol 2001; 3: 1-3.
39. Chen C.-Z, Li L., Lodish H.F., et al. MicroRNAs Modulate Hematopoietic Lineage Differentiation. Science 2004; 303: 83-86.
40. Enright A.J, John B., Gaul U., et al. MicroRNA targets in Drosophila. Genome Biol 2003; 5: R1-14.
41. Griffiths-Jones S. The microRNA Registry. Nucleic Acids Res 2004; 32: D109 - D111.
42. Griffiths-Jones S, Grocock R.J., van Dongen S., et al. miRBase: micro sequences RNA, targets and gene nomenclature. Nucleic Acids Res 2006; 34: D140–D144.
43. Griffiths-Jones S, Sain H.K., van Dongen S., et al. miRBase: tools for microRNA genomics. Nucleic Acids Res 2007; 36: D154–D158.
44. Bonnet E, Wuyts J., Rouzé P., et al. Evidence that micro precursors RNA, unlike other non-coding RNAs, have lower folding free energies than random sequences. Bioinformatics 2004; 20: 2911-2818.
45. Brennecke J, Stark A., Russell R.B., et al. Principles of microRNA–target recognition. PLoS Biol 2005; 3: 0404-0418.
46. Pierce V.A and Crawford D.L. Phylogenetic Analysis of Glycolytic Enzyme Expression. Science 1997; 276: 256-259.
47. The UniProt C. The Universal Protein Resource (UniProt). Nucl Acids Res 2008; 36: D190-195.
48. Place A.R and Powers D.A. Purification and Characterization of the Lactate Dehydrogenase (LDH-B4) Allozymes of Fundulus heteroclitus. J Biol Chem 1984; 259: 1299-1308.
49. Place A.R and Powers D.A. Kinetic Characterization of the Lactate Dehydrogenase (LDH-B4) Allozymes of Fundulus heteroclitus. J Biol Chem 1984; 259: 1309-1318.
50. Zakhartsev M, Johansen T., Pörtner H.O., et al. Effects of temperature acclimation on lactate dehydrogenase of cod (Gadus morhua): genetic, kinetic and thermodynamic aspects. J Exp Biol 2004; 207: 95-112.
51. Hare M.P and Palumbi S.R. High Intron Sequence Conservation Across Three Mammalian Orders Suggests Functional Constraints. Mol Biol Evol 2003; 20: 969-978.
52. Mattick J.S. Introns: evolution and function. Curr Opin Genet Dev 1994; 4: 823-831.
53. Edmunds R.C. Evidence for thermal adaptation among geographically, genetically and thermally distinct populations of the Australian barramundi, Lates calcarifer (Bloch 1790): a multi-level approach. PhD Thesis. Townsville: School of Marine and Tropical Biology, James Cook University. 2009.
54. Silva M.C and Edwards S.V. Structure and Evolution of a New Avian MHC Class II B Gene in a Sub-Antarctic Seabird, the Thin-Billed Prion (Procellariiformes: Pachyptila belcheri). J Mol Evol 2009; 68: 279-291.
55. Schulte P.M, Gomez-Chiarri M and Powers D.A. Structural and Functional Differences in the Promoter and 5' Flanking Region of Ldh- Within B and Between Populations of the Teleost Fundulus heteroclitus. Genetics 1997; 145: 759-769.
56. Lagos-Quintana M, Rauhut R., Yalcin A., et al. Identification of tissue-specific microRNAs from mouse. Curr Biol 2002; 12: 735-739.
57. Lai E.C. MicroRNAs are complementary to 3' UTR motifs that mediate negative post-transcriptional regulation. Nat Genet 2002; 30: 363-364.
58. Lai E.C, Tomancak P., Williams R.W., et al. Computational identification of Drosophila microRNA genes. Genome Biol 2003; 4: 1-20.
59. Selbach M, Schwanhausser B., Thierfelder N., et al. Widespread changes in protein synthesis induced by microRNAs. Nature 2008; 455: 58-63.
60. De Santis C, Evans B., Smith-Keune C., et al. Molecular characterization, tissue expression and sequence variability of the barramundi (Lates calcarifer) myostatin gene. BMC Genomics 2008; 9: 82.

61. Duerr R.H, Taylor K.D., Brant S.R., et al. A Genome-Wide Association Study Identifies IL23R as an Inflammatory Bowel Disease Gene. Science 2006; 314: 1461-1463.
62. Hirschhorn J.N and Daly M.J. Genome-wide association studies for common diseases and complex traits. Nature Reviews: Genetics 2005; 6: 95-108.
63. Scott L.J, Mohlke K.L., Bonnycastle L.L., et al. A Genome-Wide Association Study of Type 2 Diabetes in Finns Detects Multiple Susceptibility Variants. Science 2007; 316: 1341-1345.
64. Sladek R, Rocheleau G., Rung J., et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 2007; 445: 881-885.
65. Wang W.Y.S, Barratt B.J., Clayton D.G., et al. Genome-wide association studies: theoretical and practical concerns. Nat Rev Genet 2005; 6: 109-118.
66. Haddrill P.R, Charlesworth B., Halligan D.L., et al. Patterns of intron sequence evolution in Drosophila are dependent upon length and content GC. Genome Biol 2005; 6: R67.
67. Nakaya H.I, Amaral P.P., Louro R., et al. Genome mapping and expression analyses of human intronic noncoding RNAs reveal tissue-specific patterns and enrichment in genes related to regulation of transcription. Genome Biol 2007; 8: R43.

ADMINISTRATIVE DOCUMENTATION HAS BEEN REMOVED