Modulation of thyroid hormone action by environmental temperature

by

Stewart Austin Hammond
B.Sc., University of Victoria, 2011

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Biochemistry and Microbiology

Supervisory Committee

Modulation of thyroid hormone action by environmental temperature

by

Stewart Austin Hammond
B.Sc., University of Victoria, 2011

Dr Caren C. Helbing (Department of Biochemistry and Microbiology)
Supervisor

Dr Christopher J. Nelson (Department of Biochemistry and Microbiology)
Departmental Member

Dr Leigh Anne Swayne (Division of Medical Sciences)
Outside Member

Abstract

Supervisory Committee
Dr Caren C. Helbing (Department of Biochemistry and Microbiology)
Supervisor
Dr Christopher J. Nelson (Department of Biochemistry and Microbiology)
Departmental Member
Dr Leigh Anne Swayne (Division of Medical Sciences)
Outside Member

Thyroid hormone (TH) signaling is conserved across vertebrates, where it is important for normal growth and development, particularly in the perinatal period. TH has an additional critical role in amphibian metamorphosis as the sole signal that initiates the transition from a larval tadpole to juvenile frog. Premetamorphic tadpoles have a thyroid gland but are functionally athyroid, yet can be induced to undergo precocious metamorphosis by exogenous TH administration. This essential dependence upon TH makes amphibian metamorphosis an excellent model to study TH signaling.

Metamorphosis is sensitive to environmental stimuli such as temperature. Low temperature delays or slows metamorphosis, whereas high temperature advances or accelerates it. Whether a temperature is considered low or high varies by species and is related to its natural habitat. In temperate climes the North American bullfrog, *Rana catesbeiana*, does not undergo natural or precocious metamorphosis at low winter temperatures of 4-5°C. Tadpoles injected with TH at low temperature essentially clear it from their bodies after 60-80 days, but some manner of TH signaling has occurred such that they rapidly execute metamorphosis if returned to 20-25°C. This apparent molecular memory is poorly understood, but there is evidence that components of gene expression programs may be involved.

This thesis investigated the role of these factors in the molecular memory of TH formed at low temperature in the liver, brain, lung, back skin, and tail fin of *Rana catesbeiana*. The results suggested that TH receptor beta (*thrb*), CCAAT/enhancer binding protein 1 (*cebp1*), and Krüppel-like factor 9 (*klf9*) may contribute to the molecular memory to different extents in each tissue, and that TH-induced basic leucine zipper-containing protein (*thibz*) may have an important role in this process for every tissue examined. Assessment of additional genes was hampered by the limited genetic resources available for this species, so *de novo* high throughput RNA sequencing (RNA-seq) techniques were explored to alleviate this limitation. Trans-ABySS sequence assembly software produced a high quality *Rana catesbeiana* liver transcriptome that was annotated by BLAST alignment to established sequence databases and resulted in a more than ten-fold increase in *Rana catesbeiana* sequence information. This approach was supplemented with a software pipeline that was used to refine replicate *Rana catesbeiana* back skin assemblies, and by construction of a Bullfrog Annotation Resource for the Transcriptome (BART) that was used to quickly annotate more than 97% of the assembled back skin sequences.

In the future, the *Rana catesbeiana* transcriptome sequence resources can be leveraged to identify additional genes that may be involved in formation of the TH molecular memory, and chromatin immunoprecipitation could help characterize the factors and epigenetic marks in the promoter regions of these genes. Elucidation of the molecular memory mechanism provides a means to uncover key events in TH signaling.

Table of Contents

List of Tables

List of Figures

Acknowledgements

This work would not have been possible without the guidance and support of my supervisor, Dr. Caren Helbing; thank you for encouraging me to take the opportunity to expand my horizons and pursue training in bioinformatics. Thanks also to my committee members Dr. Chris Nelson and Dr. Leigh Anne Swayne for their excitement and support throughout this project.

I am indebted to Dr. İnanç Birol for welcoming me into his group at the Michael Smith Genome Sciences Centre, and to Tony Raymond, Ben Vandervalk, Greg Taylor, Bahar Behsaz, and Erdi Küçük for their instruction, advice, and productive collaboration.

Sincere thanks to Dr. Nik Veldhoen, whose profound expertise and uncompromising attention to detail improved my laboratory and informatics practices. I am also grateful for the support and camaraderie of past and present members of the Helbing Lab, in particular Amanda Carew, Pola Wojnarowicz, Stacey Maher, Mitchel Stevenson, Vicki Rehaume, and Taka-Aki Ichu.

Dedication

For Didem

List of Abbreviations

Use of capitalization and italics for gene transcripts and proteins follows the scheme given below, which is derived from http://www.xenbase.org/gene/static/geneNomenclature.jsp and http://www.informatics.jax.org/mgihome/nomen/gene.shtml.

Class	Gene transcript	Protein
Mammalia	*Thrb*	Thrb
Amphibia	*thrb*	Thrb

BART	Bullfrog Annotation Resource for the Transcriptome
bcl6	B-cell CLL/ lymphoma 6
Bd	*Batrachochytrium dendrobatidis*, chytrid fungus
c3	Complement component 3
cfhr5	Complement factor H-related 5
cps1	Carbamyl phosphate synthetase 1
cat	Catalase
CCME	Canadian Council of the Ministers of the Environment
Cdk8	Cyclin-dependent kinase 8
cDNA	Complementary DNA
cebp1	CCAAT/enhancer binding protein 1
CEG	Core eukaryotic gene
C-fin	Cultured tail fin assay
ChIA-PET	Chromatin interaction analysis by paired-end tag sequencing

cirbp	Cold-inducible RNA-binding protein
CNS	Central nervous system
Co-A	Coactivator complex
Co-R	Corepressor complex
Crf	Corticotropin-releasing factor
Dio1	Type I iodothyronine deiodinase
Dio2	Type II iodothyronine deiodinase
Dio3	Type III iodothyronine deiodinase
EDC	Endocrine disruptive compound
eef1a	Eukaryotic translation elongation factor 1
EIF2	Eukaryotic initiation factor 2
EST	Expressed sequence tag
FPKM	Fragments per kilobase per million fragments mapped
GO	Gene ontology
Gs	Gosner developmental stage
hsp30	Heat shock protein 30 kD
IBF	Ibuprofen
IPA	Ingenuity pathway analysis
klf9	Krüppel-like factor 9
MAD	Median absolute deviation
mbl2	Mannose-binding lectin (protein C) 2
mRNA	Messenger RNA
N20	20^{th} percentile of contig length

N50	50th percentile of contig length, median contig length
N80	80th percentile of contig length
NaOH	Sodium hydroxide
NCBI	National Center for Biotechnology Information
Ncor	Nuclear receptor corepressor
ND	Not detected
NF	Nieuwkoop and Faber developmental stage
NR	National Center for Biotechnology Information Non-redundant database
nsun	Nucleolar protein 2/Sun domain RNA Methyltransferase
OECD	Organization for Economic Cooperation and Development
ORF	Open reading frame
otc	Ornithine transcarbamoylase
PCR	Polymerase chain reaction
Pkc	Protein kinase C
QC	Quality control
qPCR	Real-time quantitative polymerase chain reaction
RAM	Random-access memory
RIN	RNA integrity number
Rlk1	*Rana* larval keratin type I
ROS	Reactive oxygen species
RNA-seq	High-throughput RNA sequencing
RPKM	Reads per kilobase per million reads mapped
rpl8	Ribosomal protein L8

rrp8	Ribosomal RNA-processing 8
rps10	Ribosomal protein S10
rT_3	Reverse triiodothyronine, 3,3',5'-triiodothyronine
Rxr	Retinoid X receptor
Smrt	Silencing mediator of retinoic acid and thyroid hormone receptors
snrpa	U1 Small nuclear ribonucleoprotein A
sod	Superoxide dismutase
suv91	Suppressor of variegation 9 homologue 1
T_2	Diiodothyronine, 3,3'-diiodothyronine
T_3	3,5,3'-triiodothyronine
TCS	Triclosan
Tg	Thyroglobulin
TH	Thyroid hormone
THBP	Thyroid hormone binding protein
thibz	Thyroid hormone-induced basic leucine zipper-containing protein
Thr	Thyroid hormone receptor, either isoform
Thra	Thyroid hormone receptor alpha
Thrb	Thyroid hormone receptor beta
Thrb1	Thyroid hormone receptor beta - 1
Thrb2	Thyroid hormone receptor beta - 2
TK	Taylor and Kollros developmental stage
Tpo	Thyroid peroxidase
TRE	Thyroid response element

TRENCH	Transcriptome expression and characterization
Trh	Thyrotropin-releasing hormone
TSA	Transcriptome shotgun assembly
Tsh	Thyroid-stimulating hormone
Tyr	Tyrosine

Thesis Format and Manuscript Claims

This thesis is presented in a manuscript format. Chapter 1 provides background information and introduces the rationale of the thesis. Chapters 2, 3, and 4 are written in a manuscript style containing an Abstract, Introduction, Materials and Methods, Results, Discussion, and Conclusions. Chapter 5 synthesizes the major findings of the papers and suggests future experimental directions.

Chapter 2: **Hammond SA**, Veldhoen N, Helbing CC. 2015. Influence of temperature on thyroid hormone signaling and endocrine disruptor action in *Rana (Lithobates) catesbeiana* tadpoles. Gen Comp Endocrinol 219: 6-15. doi: 10.1016/j.ygcen.2014.12.001. S. Austin Hammond and Caren C. Helbing designed and performed the exposures. S. Austin Hammond performed the qPCR assays and data analysis. S. Austin Hammond and the co-authors prepared the manuscript.

Chapter 3: Birol, I, Behsaz, B, **Hammond, SA**, Kucuk E, Veldhoen, N, Helbing, CC. 2015. *De novo* transcriptome assemblies of *Rana* (*Lithobates*) *catesbeiana* and *Xenopus laevis* tadpole livers for comparative genomics without reference genomes. Public Library of Science ONE 10(6): e0130720. Caren C. Helbing and İnanç Birol conceived and designed the experiments, and Nik Veldhoen prepared the samples for sequencing. Bahar Behsaz performed the *de novo* sequence assemblies and open reading frame predictions. S. Austin Hammond performed the sequence annotation and qPCR analysis. Erdi Küçük performed the gene ontology and pathway analysis. S. Austin Hammond and the co-authors analyzed the data and prepared the manuscript.

Chapter 4: Behsaz, B, **Hammond, SA**, Kuçuk, E, Veldhoen, N, Helbing, CC, Birol, I. *De novo* assembly and synthesis of a shared reference transcriptome from replicate *Rana* (*Lithobates*) *catesbeiana* back skin samples. Bahar Behsaz wrote the software pipeline, performed the *de novo* sequence assembly, and made the open reading frame predictions. S. Austin Hammond prepared the samples for sequencing, created the Bullfrog Annotation Resource for the Transcriptome (BART), and annotated the sequences. Erdi Küçük performed the gene ontology analysis. S. Austin Hammond and the co-authors analyzed the data and prepared the manuscript.

1 Introduction

1.1 Thyroid hormone (TH)

1.1.1 TH importance in vertebrates

Hormones are molecules synthesized in specialized metazoan organs that are transmitted throughout the body to convey signals to target tissues (Gardner and Shoback 2007). TH plays an important role in normal metabolism, growth, and development throughout vertebra (Figure 1.1). Through modulation of anabolic and catabolic pathways it is involved in homeostasis, and also affects cardiac function and cationic transport in cardiomyoctes (Cordeiro et al. 2013; Davis et al. 1996; Klein et al. 2007). Although TH may act nongenomically its major mode of action is by regulation of gene expression, by which it plays a role in cell proliferation, differentiation, and apoptosis (Bassett et al. 2003; Davis et al. 2008; Sirakov et al. 2013). TH is of particular importance to development during gestation and the mammalian perinatal period due to its involvement in differentiation of neuronal cell lineages and myelination, skin keratinization, globin switching (Grimaldi et al. 2013; Pascual et al. 2013; Préau et al. 2015); disruption of TH action in the human fetus can cause severe and irreversible developmental defects (Cao et al. 1994).

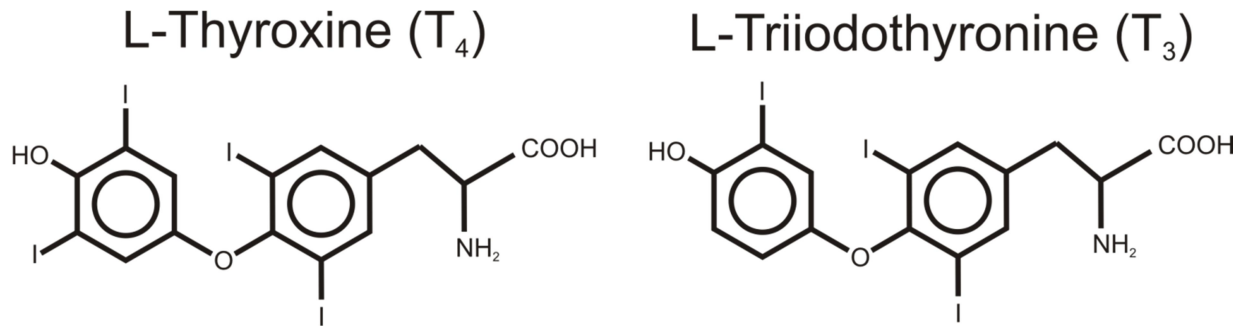

Figure 1.1 Structure of the two major endogenous THs. Adapted from Yen (2001).

TH also plays important roles in nonmammalian vertebrate development. In flatfish such as halibut and sole, the bilaterally-symmetric pelagic larva metamorphoses into a benthic juvenile

with both eyes on the same pigmented side of the body (Galay-Burgos et al. 2008; Isorna et al. 2009). In diadromous species, like salmon and some trout that have both freshwater and saltwater life phases, TH is involved in the transition from the juvenile freshwater parr to the saltwater smolt (Björnsson et al. 2011; Larsen et al. 2011). Molting in birds requires careful thermoregulation as new feathers replace old ones, and TH is responsible for increasing the basal metabolic rate in response to the heat loss experienced in the interim (Vezina et al. 2009). Plasma TH levels also increase in birds during the laying season, and the steep decline in circulating TH may signal the end of that period (Lien et al. 1993).

Amphibians also rely on TH to regulate vital processes, but it is of particular importance as the critical signal that initiates metamorphosis (Galton 1992; Shi 2000). The thyroid gland is present but nonfunctional in amphibian larvae, hence TH levels in the blood and tissues are negligible (Leloup et al. 1977; Regard et al. 1978). As the larva matures and the thyroid gland begins to produce TH, drastic changes to body plan and physiology occur to allow the transition to the juvenile phenotype (Gilbert et al. 1981; Gosner 1960; Nieuwkoop et al. 1956; Taylor et al. 1946)

1.1.2 TH synthesis, regulation, and metabolism

The critical role played by TH in vertebrate development necessitates stringent control of its activities. This control of TH is achieved through modulation of its synthesis and secretion, sequestration of circulating and intracellular TH by protein binding, and modification of TH to an active or inactive form (Shi 2000).

TH production and release is under the control of the hypothalamic – pituitary – thyroid axis (Figure 1.2). Environmental stimuli are registered by the central nervous system (CNS), which then communicates with the hypothalamus. In vertebrates, the hypothalamus stimulates the pituitary with thyrotropin-releasing hormone (Trh), whereas amphibians instead use

corticotropin-releasing factor (Crf) for this stage of the pathway (Denver 2013). Once stimulated by Trh or Crf, the pituitary releases thyroid-stimulating hormone (Tsh), which then promotes the thyroid gland to produce TH in the follicular cells. These outer cells of the gland draw iodine from the bloodstream *via* sodium-iodide symporters and then transport it to the follicular lumen through a chloride-iodide antiporter (Figure 1.3) (Bizhanova et al. 2009; Levy et al. 1997).

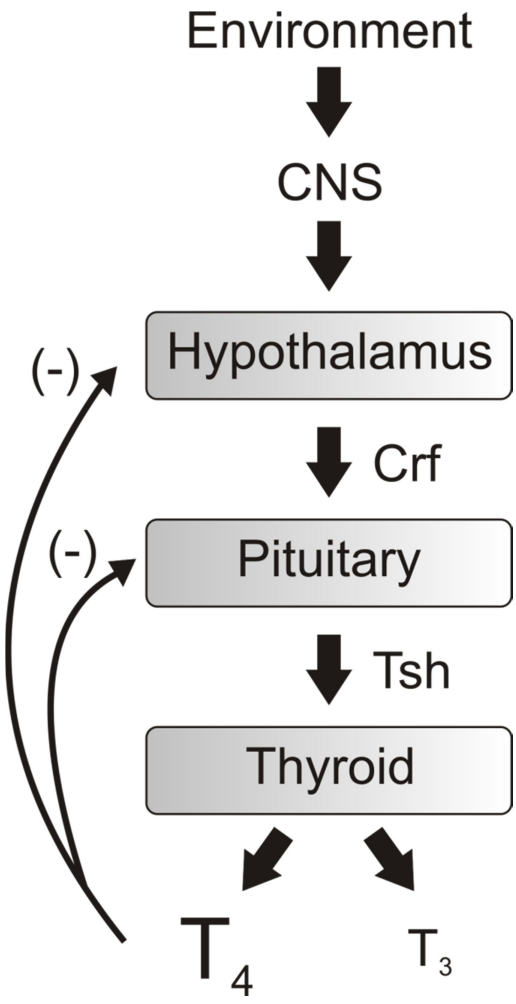

Figure 1.2 The hypothalamic – pituitary – thyroid axis. Environmental cues are registered by the CNS, which signals the hypothalamus to release Crf (Trh in mammals). Crf triggers the pituitary to produce Tsh, which stimulates thyrocytes in the thyroid gland to produce the THs T_4 and T_3, although the latter to a lesser degree. Release of these THs exerts negative feedback upon the hypothalamic – pituitary –thyroid axis to moderate the amount of TH. Adapted from Nussey and Whitehead (2001).

The follicular cells also produce and transport to the lumen the glycoprotein thyroglobulin (Tg), which is rich in tyrosine residues that are the substrate for thyroid peroxidase (Tpo) (Taurog et al. 1996). This enzyme catalyzes one to two iodinations of tyrosine residues on Tg to form monoiodotyrosine or diiodotyrosine, respectively, and then conjugates two diiodotyrosine residues to form the thyroxine (T_4) or one diiodotyrosine and one monoiodotyrosine to form 3-3'-5-triiodothyronine (T_3) (Taurog 1996). These THs are ultimately freed from the modified

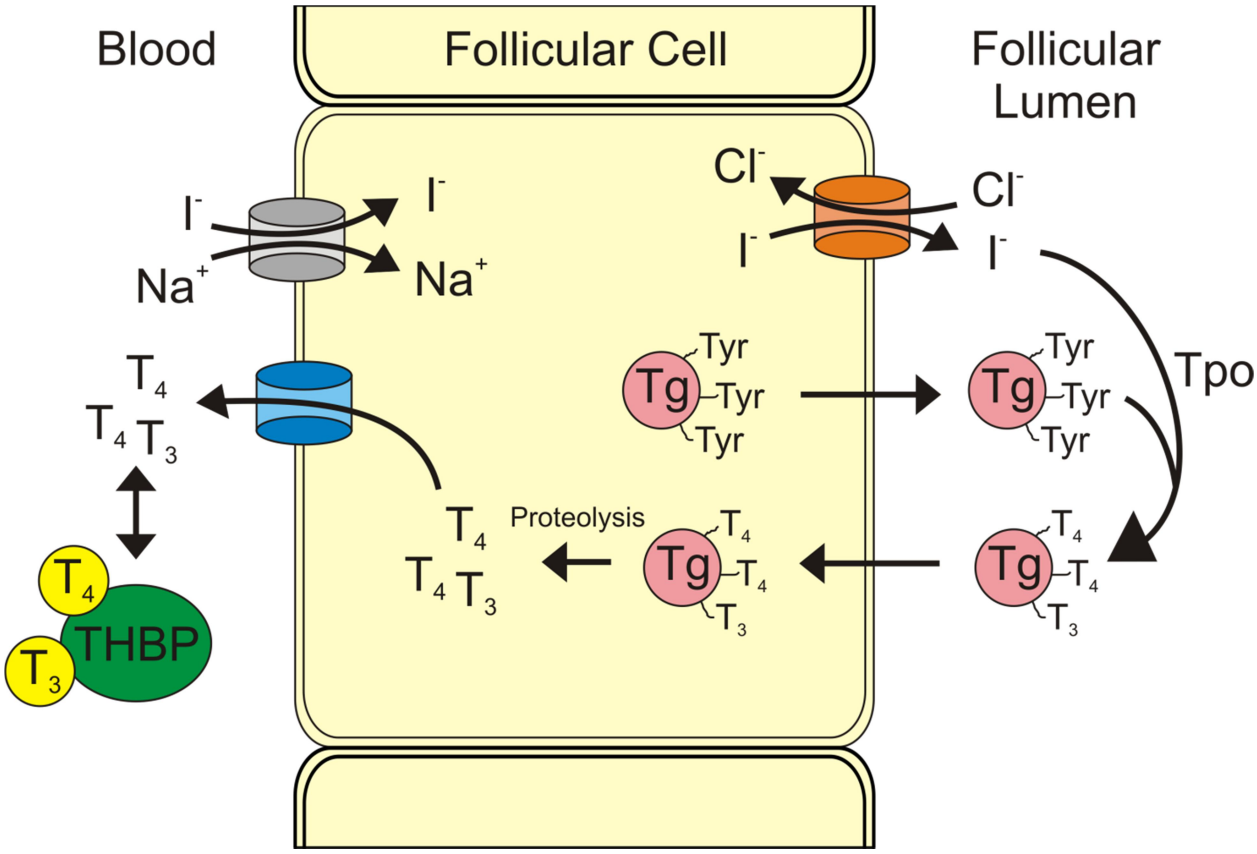

Figure 1.3 Production of TH in the thyroid gland. Iodide enters the follicular cells of the thyroid gland through a sodium-iodide symporter. It and Tg are then exported to the follicular lumen, where tyrosine residues (Tyr) on Tg are iodinated by Tpo. Conjugation of iodinated Tyr by Tpo results in Tg-bound T_4 and T_3, which re-enters the follicular cell where proteolysis releases the THs. These are then exported to the bloodstream, where the majority of the THs are bound by TH binding proteins (THBPs) for transport throughout the body. Adapted from Nussey and Whitehead (2001).

thyroglobulin through proteolysis and enter the bloodstream through a process that may involve monocarboxylate transporters (Brix et al. 2011). The majority of secreted THs are bound by transthyretin, thyroxine-binding globulin, and albumin, but a small fraction travels throughout the body unbound (Oppenheimer 1968; Refetoff et al. 1970). Regulation of TH production and release is achieved through a negative feedback loop, where high levels of TH downregulate release of Tsh from the pituitary (Wang et al. 2009).

In both humans and amphibians the majority of TH released from the thyroid gland is T_4 (Bianco et al. 2002; Galton 1983). After entering the cell, THs are bound by cytoplasmic TH-binding proteins. Interestingly, these proteins often have two unrelated functions: pyruvate kinase subtype m2 participates in energy metabolism as a tetramer, but is able to bind TH with high affinity as a monomer (Shi et al. 1994). T_4 can be converted to T_3 after entering the cell by intracellular type I and type II 5'-deiodinases (Dio1 and Dio2, respectively). Additional

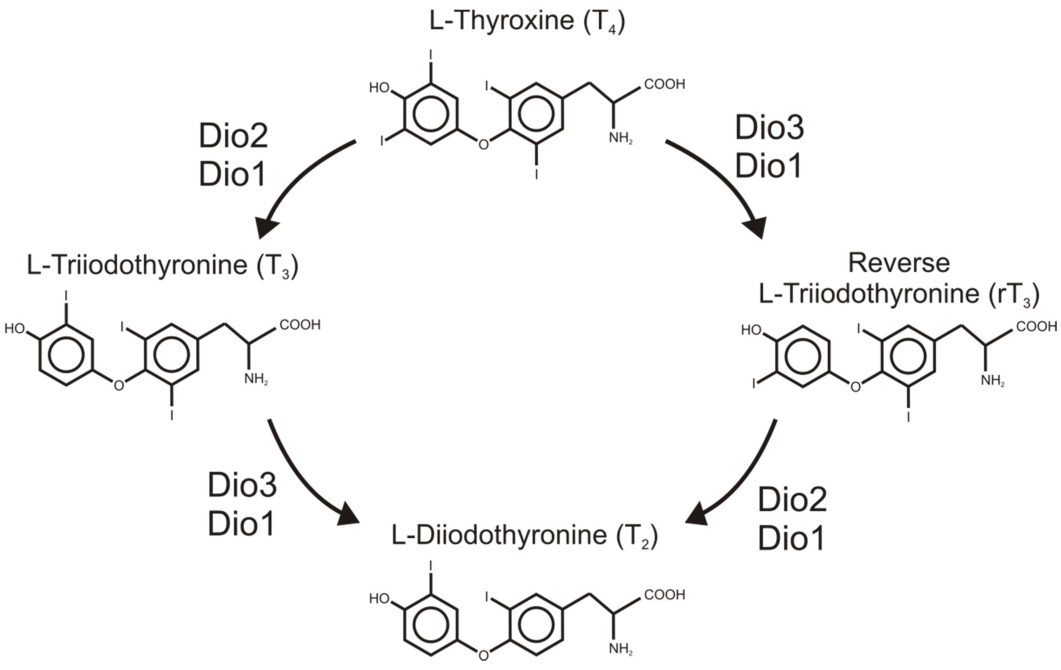

Figure 1.4 Interconversion of TH by Dios. Adapted from Yen (2001).

deiodinations can be performed to convert T_4 into reverse T_3 (rT_3), or to produce T_2 from T_3 or rT_3 (Figure 1.4). These forms of TH were considered to be non-functional intermediates destined for complete deiodination and excretion, but T_2 has been recently identified as an alternate ligand for a particular Thr isoform in tilapia (Bianco et al. 2002; Mendoza et al. 2013). While T_4 has long been considered as primarily a transport form of the hormone due to the five-fold lower biological effect and affinity for the Thrs relative to T_3 (Galton 1986; Schneider et al. 1993; White et al. 1981; Zhang et al. 2006), evidence of its distinct effects on gene expression and biological activity in Dio2-poor amphibian tissues and knockout mice have prompted re-evaluation of this dogma (Galton et al. 2014; Galton et al. 2009; Helbing et al. 2007; Helbing et al. 2007; Maher et al. 2015).

1.1.3 Regulation of transcription by TH

The Thrs are class II nuclear hormone receptors that bind DNA both in the presence and absence of TH as a heterodimer with the retinoid X receptor (Rxr) (Mangelsdorf et al. 1995). The two major isoforms, Thra and Thrb, bind to thyroid response elements (TREs) composed of two half sites, separated by a small number of bases, located in gene regulatory sequences (Paquette et al. 2014). While most often found near the core promoter of the target gene, these TREs can be as far away as half a megabase upstream as identified through chromatin interaction analysis by paired-end tag sequencing (ChIA-PET) techniques (Aranda et al. 2001; Buisine et al. 2015). In the unliganded state, the Thr-Rxr heterodimer is complexed with factors that silence expression of the target gene, including nuclear corepressor (Ncor) and silencing mediator of retinoic acid and Thrs (Smrt) (Figure 1.5) (Astapova et al. 2008). These factors are part of a multiprotein complex that mediates histone deacetylation around the promoter and thereby represses expression of the target gene (Guenther et al. 2001; Heinzel et al. 1997; Huang et al.

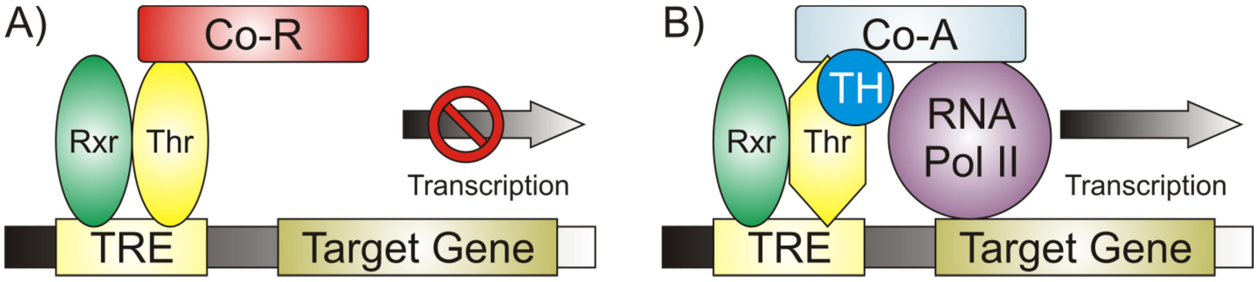

Figure 1.5 Regulation of gene expression by TH. A: In the unliganded state, the TRE-bound Thr-Rxr heterodimer recruits co-repressors (Co-Rs) to the gene promoter and transcription of the target gene is repressed. B: Upon binding TH, the Thr undergoes a conformational change and exchanges the Co-Rs for co-activators (Co-As) and transcription is able to proceed. Adapted from Yen (2001).

2000). Binding of TH results in the dissociation of silencing factors from the complex and recruitment of the multisubunit Mediator complex, which then interacts with RNA polymerase II and the basal transcriptional machinery (Fondell 2013). Recruited histone acetyltransferases promote an open chromatin state permissive to transcription, leading to expression of TH-responsive genes (Oetting et al. 2007).

Activation of gene expression by TH binding to Thr is the mechanism that is best understood. However, some genes are repressed by the TH/Thr complex. Control of endogenous TH is achieved through negative feedback, where increasing TH levels cause a decrease in Tsh and Trh expression (Sugrue et al. 2010; Vella et al. 2009). Tsh is a two subunit glycoprotein whose receptor specificity is determined by the *Tshb*-encoded β subunit (Steinfelder et al. 1997; Wolff et al. 1974). The existence of one or more negative TREs in the *Tshb* promoter has been proposed, consisting of an inverted repeat that may support Thr homodomerization and a lone half-site that may bind Thr as a monomer (Bodenner et al. 1991; Carr et al. 1994; Wondisford et al. 1989). Chromatin immunoprecipitation assays have supported Thr binding in both regions and, in conjunction with knockdown experiments, identified preference for Thrb (Chiamolera et

al. 2012). The makeup of the repressive protein complexes interacting with the Thrs upon TH binding remains unclear, but interestingly the coactivation domain of the Thr appears to be necessary (Ortiga-Carvalho et al. 2005).

Originally identified as *c-erb-A*, the cellular counterpart to the viral oncogene *v-erb-A*, the highly homologous Thra and Thrb proteins are encoded by two genes in most vertebrates: *Thra* and *Thrb* (Sap et al. 1986; Shi 2000; Thormeyer et al. 1999). They share a similar overall structure with a C-terminal ligand binding domain and highly-conserved DNA binding domain closer to the N-terminus. Each gene can produce multiple isoforms due to alternative promoter usage and splicing (Chassande et al. 1997; Williams 2000). In humans and mice, only one of the four *Thra* products is an actual nuclear receptor; the majority of the others have repressive roles due to their ability to bind DNA but not TH (Koenig et al. 1989). An example is c-erb-Aα2, a protein generated through alternative splicing of the *Thra* mRNA (Oetting et al. 2007). This protein can bind DNA, but not TH due to its extended C-terminus, thereby competing with true Thrs for binding to TREs in upstream promoter regions (Oetting et al. 2007). Interestingly, in rats c-erb-Aα2 has inverted properties: it can bind DNA with nanomolar affinity, but cannot bind TH (Lazar et al. 1990). In humans, Thra is actually downregulated by T_3 in the heart, kidneys and pituitary gland (Lazar et al. 1990). Two proteins are produced from the *Thrb* gene, and aside from the length of their amino termini, Thrb1 and Thrb2 are identical (Shi 2000). While these Thrb isoforms are not uniformly expressed between tissues, no isoform-specific effects with respect to differential regulation of gene expression have been reported outside of some fish species, where T_2 is an alternate ligand for Thrb1 (Mendoza et al. 2013; Oetting et al. 2007).

1.2 TH-mediated metamorphosis of amphibians

Much like humans and other animals, amphibians undergo physical changes during their development. However, unlike most vertebrates, amphibia have free-living larval and juvenile/adult stages in vastly different ecological niches. Amphibian metamorphosis can be divided into three main phases: premetamorphosis, prometamorphosis and metamorphic climax (Shi 2000; Yaoita et al. 1990). Progression within and between these phases is staged according to morphological criteria, including hind limb digitation and snout definition. Development of the African clawed frog, *Xenopus laevis*, is staged using the system developed by Nieuwkoop and Faber (NF) (Nieuwkoop et al. 1956), while the Taylor and Kollros (TK) or Gosner (Gs) systems are often used for the North American bullfrog (*Rana catesbeiana*) (Gosner 1960; Taylor et al. 1946). During premetamorphosis (TK stages I – IX), the thyroid gland is present but not active and there is no circulating TH (Figure 1.6). Entry into prometamorphosis (TK X – XIX) coincides with maturation of the thyroid gland and rising levels of circulating TH (Figure 1.5) (Regard et al. 1978). The most overt morphological change to occur during this phase is the rapid growth of the hindlimbs from the small buds that were formed in the preceding phase. The ultimate phase of metamorphosis, metamorphic climax (TK XX – XXIV), encompasses the most dramatic developmental changes (Shi 2000). The resorption of the tail and gills concludes, as do development of the limbs and lungs. Internally, the process of nitrogen excretion completes the switch from an ammoniotelic to a ureotelic system in the liver (Brown et al. 1959). The tadpole's glandular sheath regresses and is replaced by a larger, permanent stomach and the intestine shortens to suit the switch from the microphagous or herbivorous diet of the tadpole to the predatory carnivorism of the adult (Hourdry et al. 1996). Extensive alterations to the nervous system also accompany the new lifestyle and adult means of locomotion (Sillar et al. 2008). TH levels reach their peak at approximately 10 nM during TK XXIII, before dropping to a

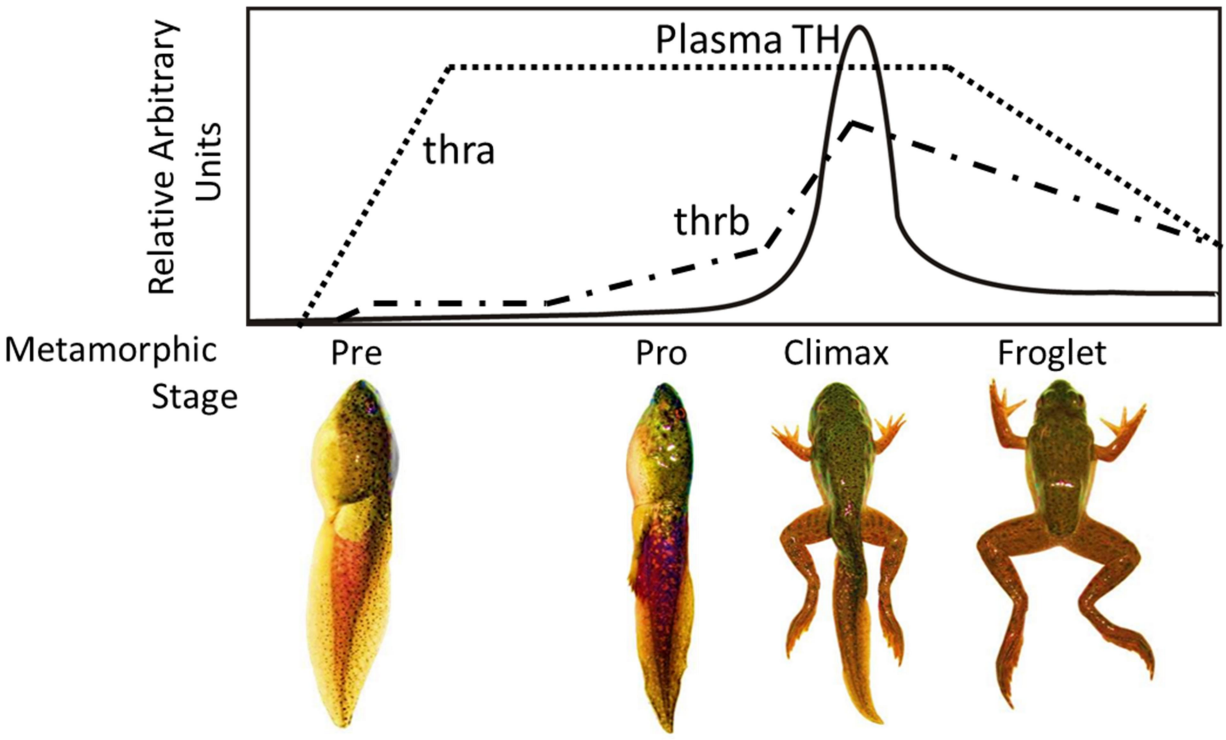

Figure 1.6 Plasma TH levels in relation to Thr expression and morphological change during *Rana catesbeiana* metamorphosis. The premetamorphic tadpole is functionally athyroid, but thra expression during this phase bestows competence to undergo precocious metamorphosis if exposed to exogenous TH. As the level of circulating TH increases during prometamorphosis, so does expression of thrb. This is accompanied by morphological changes, which become more pronounced during metamorphic climax. Post-climax, the levels of TH and the Thrs fall to a suprabasal state as the juvenile body plan is achieved. Representative bullfrog images are shown for each stage category. Adapted from Das et al (2010).

suprabasal level of about 2 nM (Figure 1.6) (Regard et al. 1978). Metamorphosis is considered to have concluded when the tail has been fully resorbed (TK XXV). Over the course of these changes, the tadpole's mass has remained approximately constant; the remainder of physical changes consist of growth of the froglet to a full-sized adult.

These varied and dichotomous changes to the tadpole physiology are initiated solely by TH (Shi 2000). The foremost lines of evidence supporting this role for TH are that tadpoles that have undergone thyroidectomies or have had synthesis of TH chemically blocked will not undergo metamorphosis, and conversely that metamorphosis may be precociously-induced by treatment

with TH (Gudernatsch 1912; Regard et al. 1978). This induction is not only possible with the whole animal, but can be induced in many of the individual tissues in *ex vivo* culture as well. (Hammond et al. 2013; Hinther et al. 2010; Veldhoen et al. 2005)

The major role of TH in initiating these diverse metamorphic changes is accomplished by alteration of tissue-specific gene expression networks. For example, TH induces expression of Thrb in the liver, which binds to TREs in the promoter of CCAAT/enhancer binding protein 1 (*cebp1*) and increases expression of that gene (Chen et al. 1997). Cebp1 in turn activates expression of ornithine transcarbamylase (Otc) and increases that of carbamoyl phosphate synthetase 1 (Cps1), which are integral members of the urea cycle that is characteristic of the adult physiology (Helbing et al. 1992). The tissue-intrinsic quality of the metamorphic gene expression program allows fine study of the direct tissue effects of TH, as well as disruption of normal TH action by endocrine disruptive chemicals (EDCs) (Hammond et al. 2013; Hinther et al. 2010; Ji et al. 2007).

Investigations into the makeup and kinetics of these TH-induced metamorphic gene expression networks have revealed that they may be roughly separated into an initiation phase and an execution phase based upon the sensitivity of certain genes to chemical interference of RNA transcription or protein translation or phosphorylation (Buckbinder et al. 1992; Ji et al. 2007; Kanamori et al. 1993; Skirrow et al. 2008). Thrb and Cebp1 were identified as components of this early, inhibitor-resistant phase of expression: by virtue of the TRE-bound Thr complexes in their gene promoters, their expression can be activated directly by TH (Menéndez-Hurtado et al. 2000; Sakurai et al. 1992). “Direct response” genes such as these are integral to control of the TH-induced gene expression program (Shi 2000). However, the chemical methods used to obtain this valuable information have serious consequences for the tadpoles or cultured cells in these

experiments, as within eight hours the protein synthesis inhibitors make them noticeably sick or cause a general decline in transcription (Kanamori et al. 1993; Wang et al. 1993).

The molecular mechanisms of amphibian metamorphosis have been predominantly studied using *Xenopus laevis* and *Xenopus tropicalis*, likely because their amenability to captive breeding ensures a ready supply of research specimens (Parker et al. 1947). However, *Xenopus* larvae are typically much smaller than those of *Rana catesbeiana*, with the consequence that each individual animal yields a smaller quantity of tissue for analysis. Indeed, *Rana catesbeiana* tadpoles are large enough that techniques such as the cultured tail fin (C-fin) assay are possible, where multiple tissue biopsies are collected from an individual animal and cultured *ex vivo* in a variety of hormone or chemical conditions (Hammond et al. 2013; Hinther et al. 2011; Hinther et al. 2010; Hinther et al. 2012). This assay design, in effect, means that each animal is simultaneously and independently exposed to every condition in the experiment, and the result of these different conditions can be evaluated within each individual animal using powerful repeated-measures statistics. As a model for human perinatal development, including the transition from the aquatic environment of the womb to the outside world, *Xenopus* species are again inferior to *Rana catesbeiana* because they remain aquatic and never depend on their lungs to obtain oxygen (Shi 2000).

1.3 Impacts of low environmental temperature on amphibian development

The central importance of TH in anuran development does not preclude involvement of other biological or environmental factors that serve to modulate the timing of progression through metamorphosis (Dodd et al. 1976). In temperate latitudes many amphibians overwinter as adults, but *Rana catesbeiana* typically remains as a tadpole until its second year of life (Cecil et al. 1979; Dent 1968). Thus, the postembryonic developmental program of this particular anuran

species has evolved to include environmental temperature as an influencing modulator in the life stage transition. Investigation into the relationship between low environmental temperature and amphibian development began in the late nineteenth and early twentieth centuries, when it was noted that low temperature delayed metamorphosis and reduced the efficacy of TH treatment (Atlas 1935; Hertwig 1897; Huxley 1929). More detailed analyses revealed attenuated induction of urea cycle enzymes in *Rana catesbeiana* larvae acclimated to 15°C, and that larvae acclimated to 5°C prior to T_3-treatment will not undergo metamorphosis at all (Frieden et al. 1965; Paik et al. 1960). The apparent lack of response to T_3 at the cold temperature is not simply due to a slowed metabolism, nor from an inability of the hormone to enter the cell and bind to the Thrs (Murata et al. 2005). However, once transferred from this apparent nonpermissive temperature to 20 - 25°C, the barrier to full urea cycle activity is lifted and the tadpoles rapidly undergo precocious metamorphosis (Ashley et al. 1968; Frieden et al. 1965; Paik et al. 1960; Yamamoto et al. 1966). This ability to resume metamorphosis without delay is maintained even after an extended period of time at the nonpermissive cold temperature where TH has been effectively cleared from the tadpole body (Yamamoto et al. 1966). It is evident that TH acts upon the larval cells at low temperature and commits them to undergo metamorphosis, but progression from the initiation phase to the execution phase of the program is blocked. This block can be lifted by returning the animal to a permissive temperature, whereupon the initial TH signal is remembered and metamorphosis proceeds; the precise nature of this molecular memory of TH remains to be elucidated, and may yield insight into regulation of gene expression by TH in general.

It is hypothesized that early expression of gene products involved in forming a molecular memory of TH should exhibit TH-dependent modulation at nonpermissive environmental temperatures, while expression of gene products that further propagate the developmental

program or that are directly involved in metamorphic changes are likely refractory. Previous hepatic studies have detected a slight increase in *thrb* expression 3 weeks after TH treatment at nonpermissive temperature, which suggests that it may constitute part of the molecular memory (Atkinson et al. 1996; Helbing 1993). The breadth of this potential role remains to be evaluated in other tadpole organs and tissues, as well as the contribution of other genes to the molecular memory of TH formed at low temperature. While the transcriptomic resources that would be leveraged for this venture using *Rana catesbeiana* are scarce in comparison to those available for the *Xenopus* species, the increased accessibility of high-throughput RNA sequencing (RNA-seq) and sequence reconstruction techniques that do not rely on a reference genome effectively reduces this limitation (Francis et al. 2013; Hornett et al. 2012; Martin et al. 2011).

Characterization of the regulatory factors that mediate TH-dependent initiation of tissue-specific gene expression programs during metamorphosis have been extensively studied in *Xenopus laevis*, although this species experiences markedly different environmental conditions in its natural habitat and laboratory experiments employed higher temperature exposure regimens (20ºC – 22ºC) along with supraphysiological levels of TH (Buckbinder et al. 1992; Wang et al. 1993). The genomes of both *Xenopus* species have been sequenced and annotated to some extent, but they are separated from Ranids by approximately 200 million years of evolution (Sumida et al. 2004). As a consequence, the degree of sequence variation is such that *Xenopus* genomic and transcriptomic resources are poorly suited for use in Ranid species (Helbing 2012). With its natural temperature control of development, *Rana catesbeiana* provides an exemplary biological platform to investigate the initial steps of TH-regulated gene expression within anuran metamorphosis.

1.4 Objectives

The main objectives of this thesis are:

1) *to evaluate expression profile of canonical direct-response genes at low temperature and investigate sensitivity to chemical disruption of the TH molecular memory*

2) *to increase resources available for transcriptomic analysis in bullfrog*

The first data chapter addresses the first objective. In Chapter 2, the TH-response profile of several direct-response and nondirect-response genes in liver, back skin, tail fin, lung, and brain at low temperature is determined. This provides insight into the relative contribution of these genes to the molecular memory of TH formed at low temperature in different tissues. Exposure to two model EDCs, triclosan and ibuprofen, in conjunction with TH at low temperature complements these results by demonstrating which, if any, of these genes are sensitive to disruption at low temperature and whether or not an EDC signal may weigh on the TH molecular memory.

Chapters 3 and 4 describe the results of *de novo* sequence assembly approaches to address the second objective. Increasing transcriptome sequence information for *R. catesbeiana* will not only allow future studies to identify additional genes that may be involved in the low temperature TH memory phenomenon, but will be an enabling resource for this organism that will affect a multitude of research avenues.

2 Influence of temperature on thyroid hormone signaling and endocrine disruptor action in *Rana (Lithobates) catesbeiana* tadpoles

Abstract

THs are essential for normal growth, development, and metabolic control in vertebrates. Their absolute requirement during amphibian metamorphosis provides a powerful means to detect and assess the impact of environmental contaminants on TH signaling in the field and laboratory. As poikilotherms, frogs can experience considerable temperature fluctuations. Previous work demonstrated that low temperature prevents precocious TH-dependent induction of metamorphosis. However, a shift to a permissive higher temperature allows resumption of the induced metamorphic program regardless of whether or not TH remains. We investigated the impact of temperature on the TH-induced gene expression programs of premetamorphic *Rana (Lithobates) catesbeiana* tadpoles following a single injection of 10 pmoles/g body wet weight T_3. Abundance profiles of several T_3-responsive mRNAs in liver, brain, lung, back skin, and tail fin were characterized under permissive (24°C), nonpermissive (5°C), or temperature shift (5 to 24°C) conditions. While responsiveness to T_3 was retained to varying degrees at nonpermissive temperature, T_3 modulation of *thibz* occurred in all tissues at 5°C suggesting an important role for this transcription factor in initiation of T_3-dependent gene expression programs. Low temperature immersion of tadpoles in water containing 10 nM T_3 and the nonsteroidal anti-inflammatory drug, ibuprofen, or the antimicrobial agent, triclosan, perturbed some aspects of the gene expression programs of tail fin and back skin that was only evident upon temperature shift. Such temporal uncoupling of chemical exposure and resultant biological effects in developing frogs necessitates a careful evaluation of environmental temperature influence in environmental monitoring programs.

2.1 Introduction

T_4 and T_3 are essential for vertebrate growth, development, and metabolic control (Forrest et al. 2013; Mullur et al. 2014). The most dramatic illustration of the importance of THs is during amphibian postembryonic development where the tadpole undergoes metamorphosis into a frog. Anuran metamorphosis is characterized by a complex temporal synchronization across diverse tissues resulting in extensive changes in body plan and organ function including tail resorption, maturation of epidermis and lungs, and reprogramming/reorganization of brain and liver (Shi 2000). Such exquisite sensitivity to and absolute dependence upon THs is the basis for the Organization for Economic Cooperation and Development (OECD) amphibian metamorphosis assay for detection of thyroid disrupting compounds (OECD, 2009). This assay is based upon *Xenopus laevis*, a laboratory species with limited natural range, and there is a desire for expanding the ability to detect TH-disrupting activities to locally-relevant sentinel species in the field. Since amphibia are poikilotherms and often experience considerable thermal fluctuations, knowledge regarding the influence of temperature on sentinel biology represents an important consideration when evaluating endocrine disruption.

For the majority of amphibians in temperate latitudes, development from egg to juvenile occurs within the warmer seasons (spring/summer) of a given year and frogs typically overwinter. However, in some species, such as *Rana (Lithobates) catesbeiana,* the tadpoles display a larval period that extends over 2-3 years (Cecil et al. 1979). Thus, the postembryonic developmental program of this particular anuran species has evolved to include marked fluctuations in environmental temperature as an influencing factor in the life stage transition.

The importance of temperature in *Rana catesbeiana* development has been clearly demonstrated under laboratory conditions. *Rana catesbeiana* tadpoles acclimated to nonpermissive cold temperatures (4-5°C) will not undergo natural metamorphosis, nor will they

undertake precociously-induced development when exposed to exogenous TH (Atkinson et al. 1996; Frieden et al. 1965). Developmental stasis in animals injected with TH at low temperature can be maintained for an extended period with no detectible hormone remaining 60-80 days following treatment. Remarkably, a shift to permissive temperature (e.g. 23±1°C) at 80-110 days results in accelerated metamorphosis (Ashley et al. 1968; Frieden et al. 1965; Yamamoto et al. 1966). The occurrence of progression through development coincident with the lack of circulating TH was hypothesized to reflect the actions of a hormone-dependent "imprint" or memory retained in treated animals (Frieden 1968; Frieden et al. 1965).

At the molecular level, commitment to and execution of anuran metamorphosis involves TH-dependent modulation of tissue-specific components within the tadpole transcriptome and proteome (Das et al. 2006; Domanski et al. 2007; Veldhoen et al. 2002; Veldhoen et al. 2006). Central to commencement of the metamorphic program are the actions of nuclear THRs that interact with TH and modulate expression levels of hormone-responsive genes, including those encoding the TRs (Shi 2000). Prior investigation of the interplay between low environmental temperature and the hepatic postembryonic developmental program in *Rana catesbeiana* tadpoles has demonstrated maintenance of a noninduced state for TH-responsive genes. Exposure to T_3 at permissive temperature leads to a rapid increase in the levels of *thra* and *thrb* mRNA encoding the two TR isoforms (Atkinson et al. 1996; Helbing et al. 1992; Schneider et al. 1991), while hormone exposure at low temperature results in no change in transcript abundance up to 21 days (Atkinson et al. 1996; Helbing et al. 1994; Mochizuki et al. 2012). Expression profiles of additional hepatic genes encoding the urea cycle enzymes *cps1* and *otc* can also reflect temperature-associated control of hormone induction (Atkinson et al. 1996; Helbing et al. 1992; Mochizuki et al. 2012).

Modulation of metamorphosis in *Rana catesbeiana* tadpoles by environmental temperature raises the possibility that the nature and timing of biological effects from exposure to EDCs may be affected. We investigated the impact of environmental temperature on TH action as well as the ability of two common water-borne contaminants and known EDCs to alter hormone-induced changes in the tadpole transcriptome. The nonsteroidal anti-inflammatory drug, ibuprofen (IBF) and the antimicrobial agent, triclosan (5-chloro-2-(2,4-dichlorophenoxy)phenol; TCS) are ubiquitous aquatic micropollutants that have been detected in surface waters worldwide at levels known to alter TH-regulated gene expression and development in *Rana catesbeiana* and *Pseudacris regilla* tadpoles (Dann et al. 2011; Hinther et al. 2011; Luo et al. 2014; Marlatt et al. 2013; Osachoff et al. 2014; Veldhoen et al. 2006; Veldhoen et al. 2014).

Using qPCR assessment, we determined the influence of environmental temperature on the status of select TH-responsive and stress-related gene transcripts in several tissues of premetamorphic *Rana catesbeiana* tadpoles induced to undergo precocious metamorphosis. Changes in the transcriptome of liver, brain, tail fin, back skin, and lung were evaluated; each tissue encompassing diverse developmental outcomes experienced during anuran postembryonic development. The importance of temperature with respect to modulating EDC influence on the TH response was also highlighted following exposure of T_3-treated tadpoles to IBF and TCS. Our observations support the existence of temperature control over progression through postembryonic development of *Rana catesbeiana* at the level of establishment of tissue-specific metamorphic gene expression programs. Such uncoupling of induction and execution of metamorphosis with potential temporal shift in EDC effects has important implications towards environmental toxicology assessments that utilize this sentinel as well as for other wildlife species with a similar paradigm.

2.2 Materials and methods

2.2.1 Experimental animals

TK stage III-XI premetamorphic *Rana catesbeiana* tadpoles (Taylor et al. 1946) were caught locally in Victoria (BC, Canada). Tadpoles were housed in the University of Victoria Outdoor Aquatic Unit in 100 gallon covered fiberglass tanks containing recirculated dechlorinated municipal water at 15°C with pH 7.1-7.2 and 81-82% dissolve oxygen (DO) and fed daily with spirulina (Aquatic ELO-systems, Inc., FL, USA). The care and treatment of animals was in accordance with guidelines established by the Canadian Council on Animal Care and approved by the Animal Care Committee of the University of Victoria. All chemical exposures were performed in the Outdoor Aquatic Unit and resulted in no mortalities prior to termination of the experiments.

2.2.2 Animal exposures and tissue culture

2.2.2.1 General considerations

Tadpole exposures were performed in 8 L of dechlorinated municipal water in 12 L high density polyethylene buckets immersed but still standing in temperature-controlled, circulating water in large indoor tanks. Animals were not fed during the course of the chemical exposure experiments. Water quality parameters (dissolved oxygen, pH, temperature, nitrate, nitrite, and ammonia) were routinely measured and fell within acceptable limits of the Canadian Council of Ministers of the Environment guidelines on Water Quality for the Protection of Aquatic Life (http://ceqg-rcqe.ccme.ca/). Animals were injected intraperitoneally via the tail muscle (Helbing et al., 1992) with 1 µL/g body weight (bw) of 10 µM T_3 (CAS no. 6106-07-6; Sigma-Aldrich Canada Ltd., Oakville, ON, Canada) prepared in 800 µM NaOH (CAS no. 1310-73-2; ACP Chemicals Inc., Montréal, QC, Canada) leading to a final *in vivo* chemical concentration of 10

pmoles/g bw T_3. Treatment control animals were injected with a corresponding volume of 800 μM NaOH vehicle.

2.2.2.2 *In vivo* exposures

TK stage VIII-XI tadpoles were selected for the whole animal exposures followed by multiple tissue evaluation. Animals included in the permissive temperature exposure series were maintained in 24±1°C water throughout the experiment and not fed two days prior to injection. Tadpoles were subsequently weighed and injected with T_3 or vehicle solution and placed in buckets containing 8 L of fresh water. These buckets were immersed but remained free standing in a large tank containing 24±1°C water. Forty-eight hours later the tadpoles were euthanized and tissues collected as indicated below.

Tadpoles included in the nonpermissive temperature exposure series were added to buckets containing 8 L of water at an average density of 11 g/L and placed in an 890 L tank containing 10±1°C water in which the buckets were immersed but remained free standing. The water temperature was gradually reduced to 5±1°C over three days using a cooling pump and then maintained at that temperature throughout the experiment. On experimental day one, 18 cold-acclimated animals were injected with T_3 or vehicle solution as indicated above and transferred to fresh 5±1°C water that was changed every two days. After seven days, half of the exposed animals (n=9 for each treatment) were transferred to buckets containing 24±1°C water (Figure 2.1). The remaining animals were maintained at 5±1°C. Twenty-four hours later the tadpoles were euthanized and tissues collected as indicated below.

2.2.2.3 Tissue collection

Animals were euthanized by immersion in a solution of 0.1% (w/v) tricaine methanesulfonate (Syndel Laboratories, Nanaimo, Canada) prepared in 25 mM sodium bicarbonate (Sigma-Aldrich

Canada) and maintained at the appropriate experimental temperature. All reagents used in animal euthanasia and tissue collection were preincubated at the appropriate experimental temperature to eliminate thermal fluctuations. Tadpole tissues were dissected immediately after euthanasia followed by immersion in RNA*later* (Life Technologies Inc., Burlington, ON, Canada) which was also preincubated at the appropriate experimental temperature. One mL of the tissue preservative was used for each 8 mm^3 liver, 25 mm^2 back skin and 12 mm^2 tail fin sample; 0.5 mL for each 27 mm^3 brain and single lung lobe sample. Tissues from the animals exposed under nonpermissive temperature were collected in a cold room, while those from animals treated at permissive temperature were harvested at ambient temperature. After incubation at 4°C for 24 h, *RNAlater* preserved tissues were transferred to -20°C until isolation of total RNA.

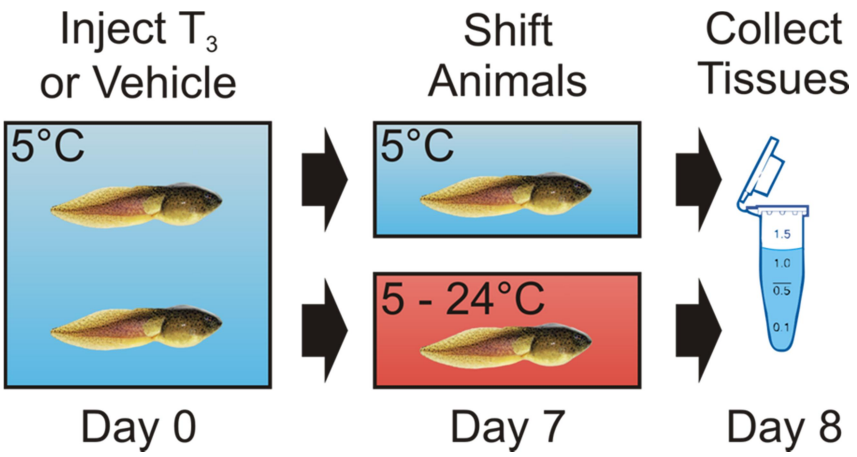

Figure 2.1 Experimental design for the 8 day exposure. Note: data from animals exposed to TH at 24°C was drawn from Maher *et al.*2015, so these animals are not depicted in this design.

2.2.2.4 *In vivo* exposure to T_3 ± IBF or TCS at nonpermissive temperature followed by incubation of tail fin biopsies at nonpermissive and permissive temperatures

Experiments that included IBF or TCS were performed by animal immersion in chemical treatment followed by culturing of tadpole back skin and tail fin tissue. Animal preconditioning included determination of body wet weight followed by placement in buckets of 15°C water

(n=8 animals/treatment condition at average density of 2.5 g/L) placed in a larger tank containing 5±1°C water for a 48 h acclimation period. Dissolved oxygen, temperature, and ammonia were routinely measured and were within acceptable limits as noted earlier. T_3, IBF (CAS no. 14883; Sigma-Aldrich Canada), and TCS (Irgasan, CAS no. 3380-34-5; Sigma-Aldrich Canada) were prepared in NaOH vehicle as 10,000x concentrated stocks and applied to the exposure buckets to final concentrations of 10 nM T_3, 10 nM T_3 + 15 µg/L IBF, or 10 nM T_3 + 3 µg/L TCS along with a 1.6 µM NaOH vehicle control (Figure 2.2). Tadpoles were euthanized following 48 h of treatment as described in section 2.2.3. Tissue biopsies of back skin and dorsal tail fin were collected using a 4 mm dermal punch (Miltex, Integra LifeSciences Corporation, York, PA, USA) from each exposed animal positioned on a chilled dissection platform and immediately placed into individual wells of a 96-well culture plate (Primaria, Corning Incorporated, Corning, NY, USA) preincubated at either at 5 or 24°C containing 200 µL of 70% Leibovitz's L-15 medium (Life Technologies) supplemented with 10 mM HEPES pH 7.5, 50 units/mL penicillin G sodium, and 50 mg/mL streptomycin sulfate (Life Technologies). Cultured tissues were then incubated at 5 or 24°C for 24 h and subsequently transferred to temperature-matched 100 µL RNA*later* for preservation.

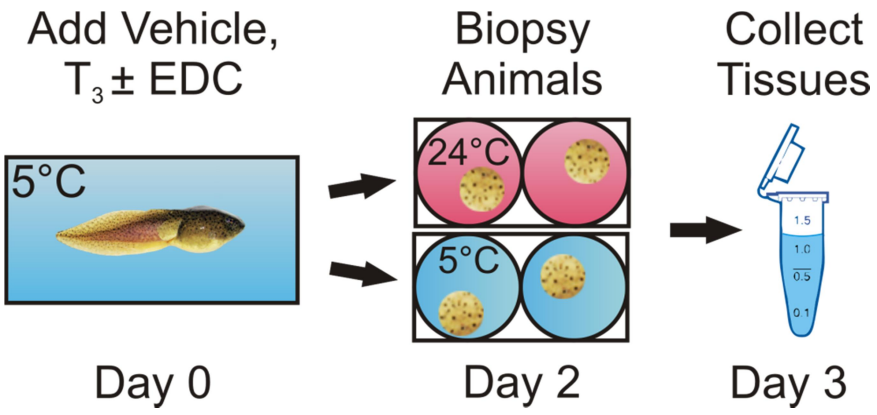

Figure 2.2 Experimental design for the low temperature hormone and EDC exposure.

2.2.3 Isolation of total RNA and cDNA preparation

All tissue samples were placed into 0.5 mL Safe-Lock tubes (Eppendorf Canada, Mississauga, ON, Canada) containing 500 µL TRIzol reagent (Life Technologies) and a 3 mm diameter tungsten-carbide bead. Samples were shaken twice for 3 min at 20 Hz in a Retsch MM301 Mixer Mill (Thermo Fisher Scientific, Markham, ON, Canada) with mixing chambers rotated 180° between cycles. Twenty µg of glycogen (Roche Diagnostics, Laval, QC, Canada) was added to the aqueous fraction of all tissue homogenates, except for liver, as a carrier prior to isopropanol precipitation to maximize RNA yield. Isolated RNA was resuspended (liver and brain, 60 µL; lung, 30 µL; back skin and tail fin, 40 µL) in diethyl pyrocarbonate-treated water (Sigma-Aldrich Canada) and stored at -80°C. For each sample, cDNA was synthesized from 1 µg total RNA using the High Capacity cDNA Reverse Transcription Kit (Life Technologies) following the manufacturer's instructions where the reaction mixture was incubated at 25°C for 10 min, 42°C for 2 h and then 5 min at 85°C. The resulting cDNA was diluted 20-fold in PCR-grade water with storage at -20°C.

2.2.4 Quantitation of mRNA abundance

Abundance of select mRNA transcripts was measured using a MX3005P real-time qPCR system (Agilent Technologies Canada Inc., Mississauga, ON, Canada) or a CFX Connect real-time PCR detection system (Bio-Rad Laboratories Inc., Hercules, CA, USA). To eliminate potential machine bias, tissue-specific samples were dedicated to either system and were not split between machine types. Amplification primers and TaqMan hydrolysis probes were designed using Primer Premier Version 5 (Premier Biosoft, Palo Alto, CA, USA) and ordered from Integrated DNA Technologies (Coralville, IA, USA).

TaqMan-based multiplex combinations and qPCR reaction assembly were as previously described and included detection of ribosomal protein L8 (*rpl8*)*, thra, and thrb,* heat shock

protein 30 (*hsp30*), catalase (*cat*), superoxide dismutase (*sod*), eukaryotic elongation factor 1A (*eef1a*) and ribosomal protein S10 (*rps10*) mRNA (Hammond et al. 2013; Wojnarowicz et al. 2013). SYBR-based qPCR reactions were performed as described previously (Hammond et al. 2013; Mochizuki et al. 2012; Veldhoen et al. 2014; Veldhoen et al. 2014) and included detection of iodothyronine deiodinases 2 and 3 (*dio2* and *dio3*), TH-induced bZip protein (*thibz*), Krüppel-like factor 9 (*klf9*), *cps1*, *otc*, *Rana* larval keratin 1 (*rlk1*), *rpl8*, *eef1a*, *rps10*, cold-inducible RNA binding protein (*cirbp*) mRNA. Validation of qPCR primer sets for use on each tissue type was performed as described previously (Veldhoen et al. 2014). Application of specific qPCR tools towards evaluation of each *Rana catesbeiana* tissue is delineated in Appendix 1. As the total RNA was not DNase I treated prior to cDNA preparation, absence of genomic DNA gDNA contamination was confirmed through analysis of the qPCR denaturation profiles of *rlk1* and *rps10* amplification reactions. qPCR amplicons generated from a gDNA source produce denaturation profiles different from that generated from cDNA template (Veldhoen et al. 2014). No contamination of qPCR data from a genomic DNA-derived signal was observed. QPCR reactions were performed in quadruplicate for each sample and gene transcript combination with cycle threshold (C_t) values averaged. All qPCR reactions included 2 µL of diluted cDNA sample. Two assay controls were run for each qPCR plate run: a negative control lacking cDNA template to ensure target specificity in fluorescent signal and a positive control reaction containing tissue-specific cDNA standard to assess interplate assay performance. Normalization to the geometric mean of *rpl8*, *rps10*, and *eef1a* C_t values was performed and mRNA levels quantified using the $2^{-\Delta\Delta Ct}$ method (Livak et al. 2001).

2.2.5 Statistical analyses

Relative fold difference data obtained from the qPCR assay were not normally distributed (Shapiro-Wilk test) and displayed unequal variances (Levene's test). Therefore, nonparametric Kruskal-Wallis and Mann-Whitney U tests were performed on all data sets using R Studio software (Gastwirth et al. 2013; R Core Team 2013; Wickham 2009). Statistical significance was considered at p-value ≤ 0.05.

2.3 Results

2.3.1 General indicators of temperature or stress responsiveness in premetamorphic tadpole tissues

No overt changes in morphology were observed in control tadpoles under any temperature condition and TH-injected animals maintained at 5°C. A similar degree of reddening of the hindlimb due to intense vascularization was apparent in the TH-injected animals maintained at permissive temperature only and TH-injected animals shifted from nonpermissive to permissive temperature (data not shown).

Evaluation of change in *cirbp* mRNA levels, an indication of exposure to cold stress, in vehicle control animals showed an increase in the nonpermissive condition relative to permissive temperature (Figure 2.3). This cold-induced elevation of *cirbp* mRNA abundance relative to

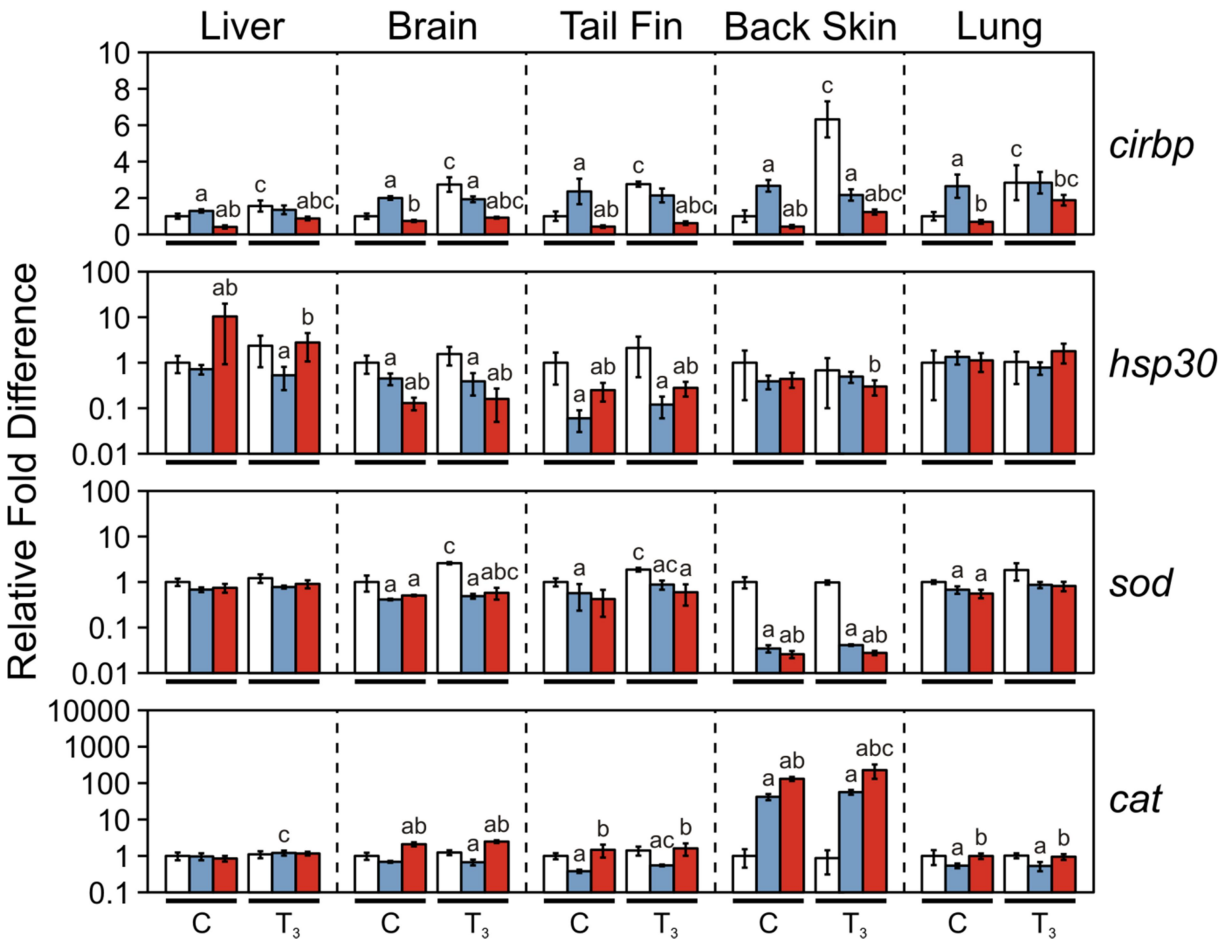

Figure 2.3 Impact of environmental temperature on TH modulation of select mRNA across tissues in the premetamorphic *Rana catesbeiana* tadpole. Animals were acclimated to 24°C (white bars) or 5°C (blue bars) prior to NaOH vehicle (C) or 10 nM T_3 exposure. Seven days post-injection, a subset of 5°C animals was transferred to 24°C water (red bars) and all animals maintained for an additional 24 h. The median fold difference of *cirbp*, *hsp30*, *sod*, and *cat* mRNA relative to the 24°C vehicle control is shown for each treatment group with whiskers denoting associated median absolute deviation (MAD). For each treatment, statistically significant difference ($p \leq 0.05$) comparing 5°C (blue bar) or temperature-shift (red bar) with 24°C (white bar) tadpoles is indicated by "a". Also for each treatment, statistical difference between temperature-shift (red bar) relative to 5°C (blue bar) is indicated by "b". Finally, significance between T_3 exposure and their temperature-matched control is shown by "c".

permissive controls was observed in all tissues tested to a maximum increase up to 3-fold ($p \leq 0.02$; Figure 2.3) with a marked reduction up to 6-fold ($p \leq 0.001$) in abundance levels relative to cold-induced levels upon shift from a nonpermissive to permissive temperature (Figure 2.3). This general pattern was also observed in TH-treated animals undergoing temperature shift for all tissues (Figure 2.3), but the responses were found to be modulated by the hormone itself under permissive only temperature conditions. Hormone responsiveness was attenuated at nonpermissive temperature and subsequently restored upon shift to permissive temperature (Figure 2.3). In all tissues, T_3 treatment significantly increased *cirbp* mRNA abundance relative to vehicle controls up to 6-fold ($p \leq 0.013$; Figure 2.3) showing, for the first time, that this transcript is hormone-responsive.

Elevation of *hsp30* mRNA abundance in liver tissue, an indicator of thermal stress, is noted upon temperature shift compared to either permissive or nonpermissive conditions. In contrast, *hsp30* mRNA levels were largely unchanged in lung and back skin and were reduced under low temperature or temperature-shifted conditions for brain and tail fin relative to permissive temperature ($p \leq 0.006$; Figure 2.3).

We also assessed the expression levels of the TH-responsive indicators of oxidative stress, *sod* and *cat*. Brain and tail showed a significant increase in *sod* mRNAs ($p \leq 0.008$) upon T_3 treatment at 24°C while no effect was observed for liver, back skin, or lung (Figure 2.3). While low temperature exposure (shifted or not) resulted in a reduction in *sod* transcripts irrespective of hormone status in all tissues except liver; the most significant effect was a 30-fold reduction of this transcript in the back skin ($p=0.001$; Figure 2.3). *Cat* mRNAs remained largely unaffected by T_3 treatment but showed a low temperature repression in brain, tail fin and lung ($p \leq 0.001$; Figure 2.3). The exception to this was observed in the back skin where *cat* mRNA levels

increased at least 60-fold ($p \leq 0.001$). Shifting the temperature from nonpermissive to permissive resulted in a hormone-independent increase in *cat* mRNAs in all tissues except liver (Figure 2.3).

2.3.2 Effects of environmental temperature on the metamorphic gene expression program

The greatest effect of reduced environmental temperature on mRNA abundance in control animals was an approximate 10- to 100-fold increase relative to permissive temperature controls noted for *thra*, *klf9*, and *cebp1* ($p=0.001$; Figure 2.4) in back skin; although the levels of these transcripts were also significantly affected in tail fin and brain (*thra* and *klf9*) and lung and liver (*cebp1*). A marked temperature-associated reduction in hepatic mRNA levels was observed for *thibz* (9-fold; $p=0.002$) compared to animals assessed at permissive environmental temperature.

Vehicle-treated tadpoles shifted from nonpermissive to permissive temperature conditions demonstrated elevated levels of *thra* and *thrb* mRNA in nearly all tissues examined with more pronounced temperature-associated change noted across tissues for the latter gene transcript (up to 23-fold for *thrb* compared with 4-fold for *thra*, $p \leq 0.014$ for both; Figure 2.4). Temperature-shift also impacted mRNA abundance for *klf9* for all tissues except liver with a reduction of up to 6-fold across the tadpole tissues examined ($p \leq 0.022$) compared to the low temperature animals. *Thibz* mRNA levels returned to levels measured at 24°C in liver upon temperature-shift ($p=0.002$; Figure 2.4).

In the context of TH exposure, the levels of five mRNA encoding regulators of RNA transcription (*thra*, *thrb*, *thibz*, *klf9*, and *cebp1*) are augmented across a number of tadpole tissues following exposure to T_3 at low temperature similar to what is observed at the permissive temperature (Figure 2.4). Of particular note is *thibz* which maintains its substantial increase in all five tissues examined and *klf9* which is induced by T_3 in all tissues but liver (Figure 2.4). The

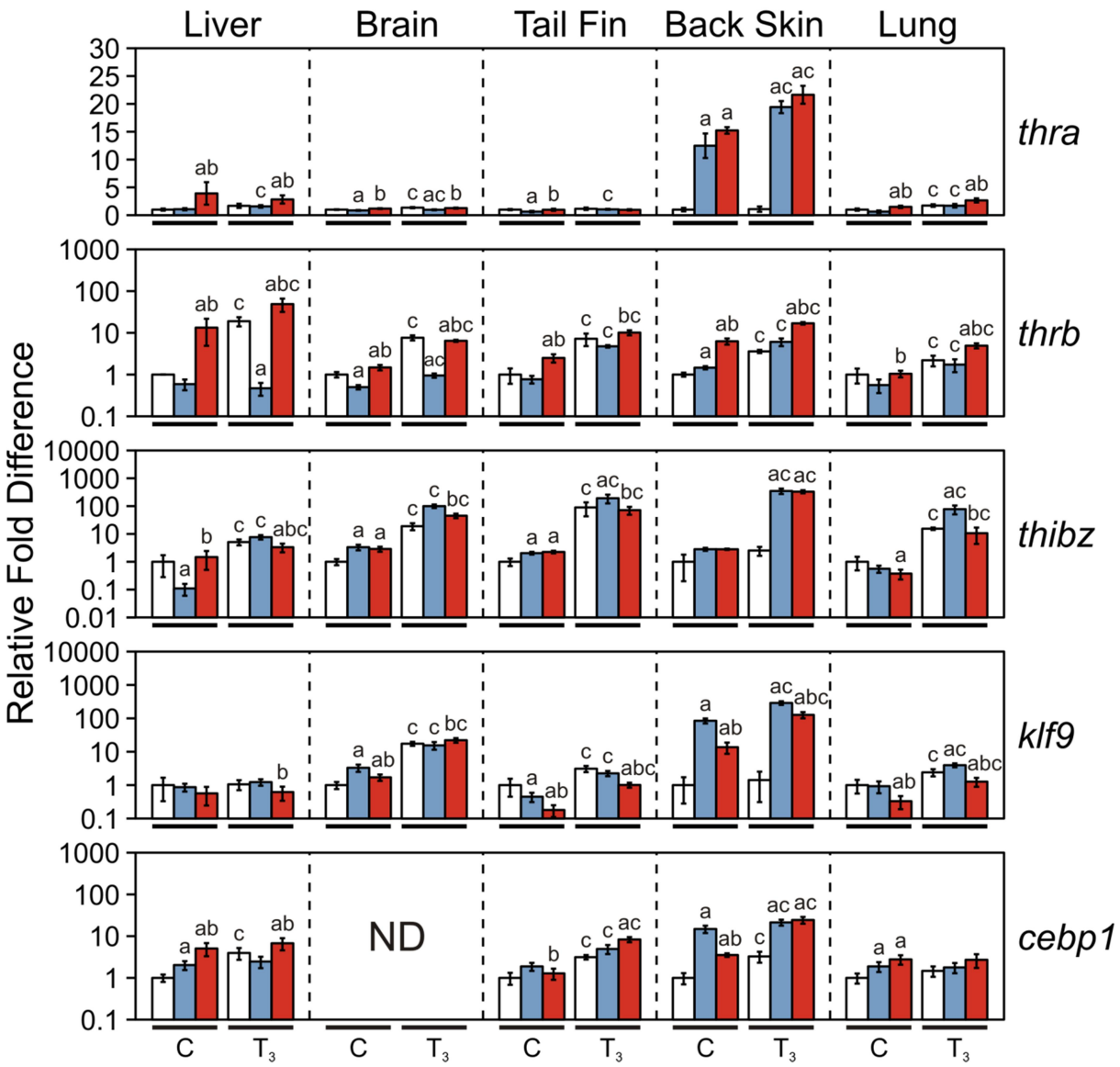

Figure 2.4 Effects of environmental temperature on the relative abundance of *thra*, *thrb*, *thibz*, *klf9* and *cebp1* mRNA within different tissues of T_3-treated premetamorphic *Rana catesbeiana* tadpoles. See Figure 1 legend for plot details. ND= not detected.

greatest magnitude of change in mRNA abundance following TH-treatment was detected for *thibz* under both temperature conditions (permissive temperature in tail fin, 90-fold; nonpermissive in lung, 138-fold; $p<0.001$ for both; Figure 2.4). In contrast, however, the T_3-induced response was eliminated at nonpermissive temperature for *thrb* and *cebp1* in the liver (Figure 2.4) in a manner observed for *cirbp* (Figure 2.3).

We also examined three additional known TH-responsive genes that have restricted expression reflecting their specific roles in select tissues. Of the two urea cycle genes measured, *cps1* showed a 2.5-fold increase in mRNA abundance in animals exposed to T_3 and held at 5°C ($p = 0.01$), while no change in the level of *otc* transcript was detected (Figure 2.5). The gene transcript encoding *rlk1* has expression restricted to the brain and skin. It is not responsive to TH treatment in the brain although animals held at the nonpermissive temperature show an increase in mRNA levels ($p\leq0.017$; Figure 2.5). In the tail fin and back skin, *rlk1* levels are reduced by as much as 20-fold upon exposure to T_3 at permissive temperature ($p\leq0.006$). This response is eliminated in animals held at the nonpermissive temperature (Figure 2.5).

Animals treated with T_3 at 5°C and subsequently shifted to the permissive temperature showed maintenance or reinstatement of TH-dependent modulation of mRNA levels for almost all of these genes similar to the levels measured in tadpoles exposed under permissive conditions alone. These included *thrb* ($p=0.005$) and *otc* ($p=0.001$) in liver, *rlk1* in tail fin and back skin ($p\leq0.001$), and *cirbp* in all tissues ($p\leq0.022$) compared with vehicle treated controls (Figures 2.3, 2.4, and 2.5). In certain instances, while the T_3-associated response was comparable between temperature-shifted animals and the higher temperature counterparts, the resulting mRNA levels were dissimilar in magnitude. A marked example is noted for *rlk1* mRNA in back skin, where T_3-mediated reduction in transcript levels under permissive temperature is 20-fold ($p=0.001$)

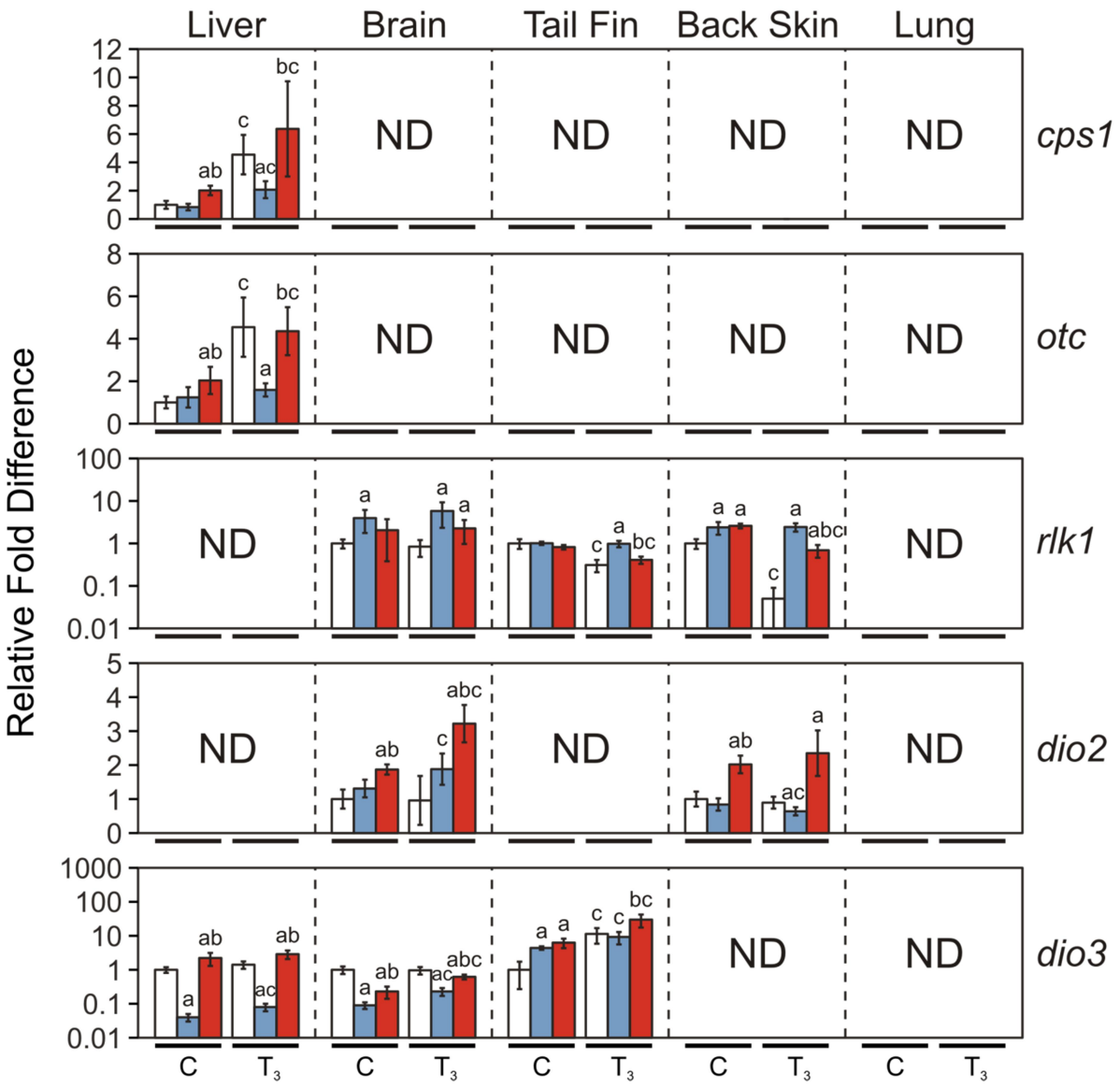

Figure 2.5 Evaluation of environmental temperature-mediated alteration in relative abundance of *cps1*, *otc*, *rlk1*, *dio2*, and *dio3* mRNA across different tissues of T_3-treated premetamorphic *Rana catesbeiana* tadpoles. See Figure 2.1 legend for plot details. ND= not detected.

while shifted animals show a 4-fold decrease ($p<0.001$) (Figure 2.5). Lack of recovery in TH-dependent modulation of gene transcript abundance was shown for hepatic *cebp1* mRNA ($p=0.189$) compared to the permissive condition ($p=0.005$), likely a result of a temperature-associated increase in background levels as demonstrated for vehicle controls (Figure 2.4).

Lastly, we examined the effect of temperature upon the transcripts encoding the iodothyronine deiodinases Dio2 and Dio3. Expression of these transcripts were also restricted to brain and back skin (*dio2*) and liver, brain and tail fin (*dio3*) at the premetamorphic developmental stage used in the present study. Both transcripts were not responsive to T_3 at the permissive temperature with the exception of an approximate 10-fold increase in *dio3* mRNA abundance in the tail fin ($p<0.001$; Figure 2.5). Interestingly, cold temperature enabled a significant response to T_3 in both *dio2* and *dio3* in all of the tissues where these transcripts are expressed with a concomitant reduction in hepatic *dio3* mRNA levels observed (17-fold; $p=0.001$) compared to animals assessed at permissive environmental temperature (Figure 2.5). *Dio3* transcripts also displayed a reduced abundance in brain at nonpermissive temperature (11-fold; $p=0.001$; Figure 2.5). *Dio2* and *dio3* levels largely returned to or exceeded levels measured at permissive temperature in liver, brain, and tail fin upon temperature-shift ($p\leq0.043$; Figure 2.5).

2.3.3 Contribution of environmental temperature to EDC exposure effects

Gene transcript profiling was performed on back skin and tail fin cultured at either 5°C or 24°C following whole animal exposure at nonpermissive (5°C) temperature to T_3 in the presence and absence of the endocrine disruptors, IBF and TCS. In this experiment, it is important to note that the tissues derived from exposed animals were subsequently cultured in the absence further chemical treatment and that biopsies originating from the same animal were maintained either at

the nonpermissive (5°C) temperature or placed into medium at the permissive (24°C) temperature.

Tissues from vehicle-injected control animals shifted to the permissive temperature showed no effect on the abundance levels of *thra, thrb, thibz, dio2, sod*, or *cat* mRNAs (Figures 2.6 and 2.7). A marked reduction in the levels of *klf9, cebp1, rlk1*, and *cirbp* mRNAs were observed while a significant increase was noted for *dio3* and *hsp30* mRNAs (Figures 2.6 and 2.7). Of these transcripts, the transcript profiles of *thra, rlk1, dio2, dio3, cirbp, sod,* and *cat* were identical for T_3 treated animals with or without added IBF or TCS (Figure 2.6 and 2.7) suggesting that the temperature effect was the overriding influence on these transcripts.

However, T_3-specific responses were detected for *thibz, thrb, klf9* and *cebp1* with the latter three showing indications of perturbation by IBF and/or TCS. Both tail fin and back skin demonstrated an approximately 2-fold increase in *thibz* mRNAs at the nonpermissive temperature that was maintained regardless of the additional IBF or TCS exposure ($p \leq 0.002$; Figure 2.6). All cultured tissue samples from animals exposed to $T_3 \pm$ IBF or TCS showed augmentation of *thibz* mRNA levels of 3 to 5-fold -upon shift to the permissive temperature when compared to the low temperature condition (Figure 2.6). Cultured tissues derived from T_3 exposed animals displayed increased abundance of *thrb* mRNA compared to the vehicle control at the permissive temperature (3-fold in back skin; 5-fold in tail fin; $p \leq 0.003$ for both tissues). Coexposure with T_3 and TCS resulted in a reduction of this hormone-mediated elevation of *thrb* transcript levels in tail fin by 1.5-fold ($p=0.041$) whereas T_3+IBF treatment was not different from T_3 alone (Figure 2.6). The T_3-induced response at permissive temperature was not affected by IBF or TCS in the back skin (Figure 2.6).

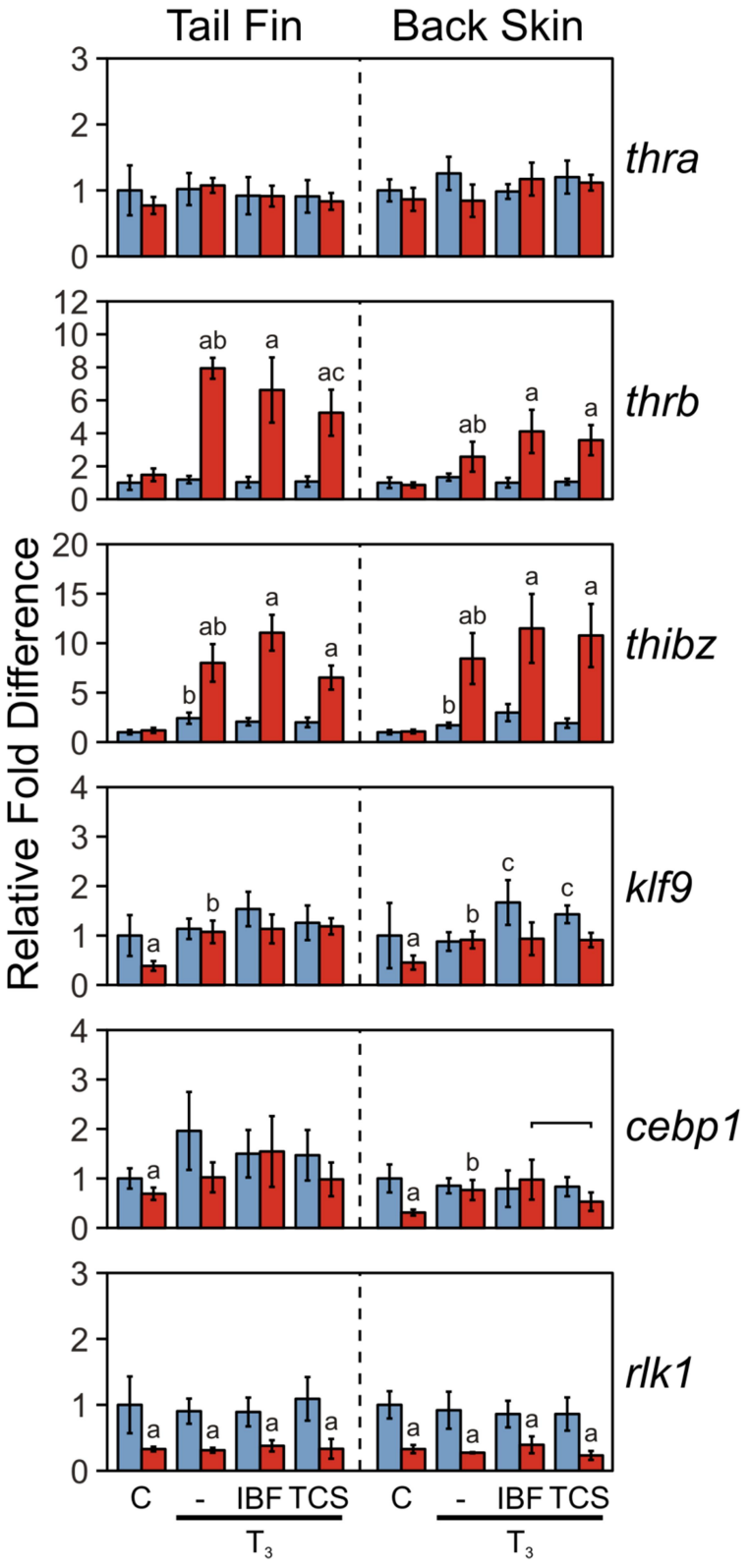
Tail Fin
Back Skin
thra
thrb
thibz
klf9
cebp1
rlk1
Relative Fold Difference
C
-
IBF
TCS
T3

Figure 2.6 Influence of environmental temperature on mRNA abundance profiles in cultured back skin and tail fin isolated from *Rana catesbeiana* tadpoles exposed to TH in the presence or absence of IBF or TCS. Premetamorphic tadpoles were exposed to NaOH vehicle (C), 10 nM T_3, 10 nM T_3 + 15 µg/L IBF, or 10 nM T_3 + 3 µg/L TCS at 5°C for 48 h. Two tissue biopsies were subsequently isolated from each animal from the indicated tissue and cultured either at 5°C (blue bars) or 24°C (red bars) for 24 h in the absence of further chemical treatment. Median fold difference of *thra*, *thrb*, *thibz*, *klf9*, *cebp1*, and *rlk1* mRNA levels relative to the vehicle control maintained at 5°C is shown for each treatment group with whiskers denoting associated MAD. Statistically significant difference ($p \leq 0.05$) of biopsies cultured at 24°C relative to those cultured at 5°C is indicated by "a". Significant difference between biopsies from animals treated with T_3 alone relative to vehicle controls is indicated by "b". Within a given temperature, significance of T_3+IBF or T_3+TCS conditions compared to T_3 alone is shown by "c" and between the T_3+IBF and T_3+TCS treatments by a square bracket.

Temperature shift resulted in a reduction of *klf9* in tail fin and *cebp1* transcripts in both tissues up to 3-fold relative to the nonpermissive temperature ($p \leq 0.018$; Figure 2.6). However, animals treated with T_3 displayed mRNA abundance in temperature shifted tissues similar to that determined for vehicle controls maintained at low temperature (Figure 2.6). This attenuation was independent of EDC treatment. At the nonpermissive temperature, a ~2-fold elevation of *klf9* in the back skin from animals coexposed to T_3 and IBF or TCS relative to T_3 alone was observed ($p \leq 0.031$; Figure 2.6). Finally, despite the general 3-fold induction of *hsp30* mRNA levels due to the temperature shift, we detected a significant 7-fold induction of this transcript in the combination T_3 + TCS treatment ($p = 0.031$; Figure 2.7).

2.4 Discussion

Metamorphosis of a larval tadpole into a juvenile frog involves coordination between environmental cues, endocrine systems, and tissue-specific developmental programs that orchestrate changes in form and function required to transition from an aquatic to a terrestrial existence. In temperate regions, many frog species accomplish this postembryonic

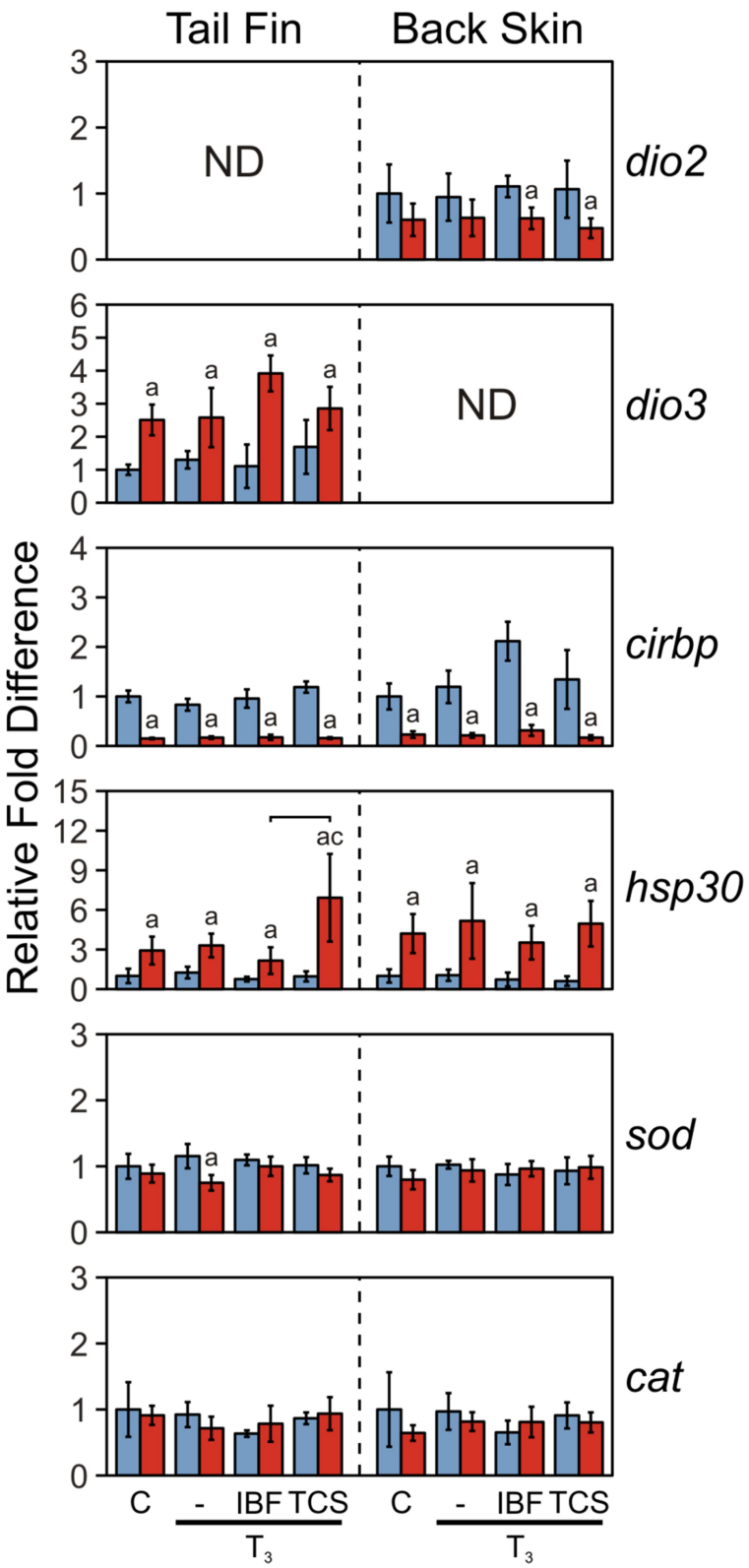

Figure 2.7 Influence of environmental temperature on abundance of *dio2*, *dio3*, *cirbp*, *hsp30*, *sod*, and *cat* mRNA in cultured back skin and tail fin isolated from tadpoles treated with TH and IBF or TCS. Further plot details are outlined in the Figure 2.4 legend. ND= not detected.

developmental process in a single spring-summer period. However, *Rana catesbeiana* overwinters as a tadpole and, consequently, has evolved biological mechanisms allowing for stasis in progression through metamorphosis until favorable environmental conditions are restored. Prior investigations have pointed towards environmental temperature as a key regulator of this process (Atkinson et al. 1996; Frieden et al. 1965). For example, cellular uptake of T_3 in *Rana catesbeiana* red blood cells is reduced at 4°C relative to 23°C. However, nuclear T_3 levels and the ability of T_3 to associate with TH receptors are not temperature-dependent (Murata et al. 2005).

Remarkably, tadpoles maintained under nonpermissive low temperature conditions for months are refractory to the overt inductive action of TH, while still able to undergo metamorphosis well after the original stimulus has been cleared (Ashley et al. 1968; Frieden et al. 1965; Yamamoto et al. 1966). Thus, there exists a retention of developmental program information in the *Rana catesbeiana* tadpole that is under tight regulation by temperature (Frieden 1968; Frieden et al. 1965). Our current investigation evaluates the mRNA abundance profiles of several TH-responsive and stress-related genes across different tadpole tissues under alternate temperature conditions in an attempt to further define the impact of environmental temperature on tissue-specific gene expression programs.

We hypothesized that gene expression profiles may fall into two general categories: 1) genes whose encoded proteins play a role in establishment of the TH-mediated metamorphic program and that are refractory to temperature control and 2) genes whose translated products are involved in execution of the metamorphic program required for developmental progression and that are affected by environmental temperature. Our current mRNA profiling suggests that the ability of key gene transcripts to respond to TH action and which encode transcription factors

(*thra*, *thrb*, *thibz*, and *klf9*) involved in induction of metamorphosis remains unaffected by reduced environmental temperature in select tissues. Such mRNA may represent vital elements associated with establishment of a commitment phase within the transcriptome under nonpermissive temperature conditions. Early work performed on the mechanistic nature of anuran metamorphosis provided evidence that aspects of initiation and execution of this developmental program could be uncoupled by inhibition of RNA transcription, controllers of the transcription complex, and protein translation or phosphorylation suggesting that a commitment point exists early on where subsequent reversal of the TH-mediated process cannot occur and all effectors required for completion of metamorphosis are established (Buckbinder et al. 1992; Ji et al. 2007; Kanamori et al. 1993; Skirrow et al. 2008). It presently remains unclear whether such molecular mechanisms are involved in temperature-dependent regulation of metamorphosis.

Epigenetic marks associated with active transcription in *Rana catesbeiana* liver have been detected in the *thrb* gene at nonpermissive temperature (Mochizuki et al. 2012) suggesting that the hormone-dependent increase in mRNA levels we observed may reflect maintenance of an active gene promoter complex. How this is accomplished under low environmental temperature conditions is not known. However, it is clear that prometamorphic gene expression programs in multiple tissues are underway to some extent at nonpermissive temperature. The data presented herein suggest that a central component in this process may involve the up-regulation of *thibz*. It would be interesting to examine the impact of shorter term temperature modulations across multiple tadpole tissues to further define the kinetics of transcriptional events in play and the associated epigenetics established with a view towards defining pivotal aspects of temperature control.

The later phase of postembryonic development includes modulation of genes encoding effectors of the metamorphic program (Brown et al. 1996; Kanamori et al. 1993; Wang et al. 1993). These include additional transcription regulators, transporters, proteases, structural proteins, and metabolic enzymes involved in cell proliferation, apoptotic events, and tissue remodeling (Brown et al. 1996; Das et al. 2006; Helbing et al. 2003; Searcy et al. 2012; Skirrow et al. 2007). In a biphasic metamorphic process under temperature regulation, such effectors would be expected to display attenuated response to TH at low temperature conditions. In general, this is supported by the cross-tissue mRNA profiles for *cirbp*, *rlk1*, and *otc* and, to a lesser extent, by *sod* and *hsp30*; the latter two genes being less responsive to TH. Gene transcripts *cps1* and *otc*, encoding components of the urea cycle, are effectors previously demonstrated to be nonresponsive to T_3 treatment at nonpermissive temperature (Atkinson et al. 1996). Our present work recapitulated temperature-dependent control of the TH-response for *otc*. However, we noted a moderate effect of TH on *cps1* mRNA levels under nonpermissive temperature which likely reflects the increased sensitivity in detection afforded by the qPCR assay.

The ability of *Rana catesbeiana* tadpoles to reduce TH levels at nonpermissive temperature as revealed by Yamamoto *et al.* (1966) is also reinforced by our present observations of the capacity of T_3 to increase abundance of *dio3* mRNA across all tissues displaying measurable transcript levels. Removal of continual hormone stimulus is likely an important aspect of maintaining arrest in developmental progression over longer periods at nonpermissive temperature. Once TH is metabolized, overwintering tadpoles would no longer require the presence of substantial Dio3 and establish a reduced background level of *dio3* mRNA.

An interesting observation derived from mRNA profiling information is that tadpole tissues display nonequivalence in temperature-related modulation. Of particular note are comparisons between the effects of nonpermissive temperature on back skin and tail fin. In postembryonic development of Ranidae, the tail is resorbed while the back skin matures into several defined layers (Fox 1974). Programmed cell death is required in both outcomes with tail resorption including apoptosis of both muscle and dermal cells and the remodeling of body skin involving apoptosis of apical and skein cells and dedifferentiation of a basal cell subpopulation to form the adult epidermis (Suzuki et al. 2002; Yoshizato 2007). The two tissues show differing degrees of sensitivity to low temperature in expression of *cat* and *sod* mRNA suggestive of alternate redox conditions required under a cold environment in these tissue types independent of hormone status. In the context of metamorphosis, reactive oxygen species have been proposed to act in establishing a TH-mediated proapoptotic signal pathway in the tadpole tail (Hanada et al. 1997; Kashiwagi et al. 1999). However, we note that a change in environmental temperature does not markedly alter the pattern of *cat* or *sod* mRNA in tadpole tail fin. In stark contrast to the tail fin, back skin is especially sensitive to temperature with respect to oxidative defense pathway components that enhance protection against hydrogen peroxide accumulation in this circumstance. A shift to permissive environmental temperature of cold-treated animals fails to reinstate the *sod* or *cat* mRNA levels observed with tadpoles originally maintained under permissive conditions, at least within the limited 24 h observation period employed.

In contrast to indicators of oxidative stress, Rlk1, an identified early marker of apoptosis in tadpole epithelial tissues (Domanski et al. 2007), demonstrates similar transcript abundance profiles between back skin and tail fin associated with changes in environmental temperature. Colder conditions eliminate the TH-associated reduction in *rlk1* mRNA levels in both tadpole

tail fin and back skin lending support for an anti-apoptotic cellular environment established under nonpermissive temperatures.

The hepatic transcriptome and its modulation by TH was particularly sensitive to temperature with hormone-mediated increases in *thrb*, *cebp1*, and *otc* mRNA eliminated under cold conditions consistent with previous observations (Atkinson et al. 1996). Temperature-dependent control of TH action was, however, not absolute with TH maintaining the ability to alter the levels of *thibz* and *cps1* gene transcripts. The response profile observed in the present study for hepatic *thra* transcripts was in contrast to observations made in previous work (Atkinson et al. 1996; Helbing 1993). Our assessment shows TH-responsiveness at nonpermissive temperature with transfer from cold-treatment to permissive temperature augmenting *thra* mRNA abundance in a hormone-independent fashion, while the earlier investigation indicated a significant decrease in mRNA abundance at 5°C and an increase upon temperature shift. Such incongruency may be the result of a supraphysiological T_3 concentration used in the earlier work. Regardless, these observations indicate that control over the status of liver is crucial during an extended overwintering period with severe restrictions to prevent aberrant TH action reflected within the transcriptome. Such temperature-associated control would be expected to maintain ammonotelism, among other critical metabolic processes in the liver, in overwintering animals.

The prospect of altering tissue-specific gene expression programs in animals maintained at nonpermissive environmental temperature that would manifest a detectable biological change upon a return to more favorable conditions was examined with IBF and TCS, environmental chemical contaminants known to display disruptive activity towards TH action (Marlatt et al. 2013; Veldhoen et al. 2006; Veldhoen et al. 2014). Cultured tail fin from tadpoles exposed *in vivo* to T_3 and TCS that was maintained in hormone- and EDC-free culture under low

environmental temperature demonstrated altered levels of *thrb* and *hsp30* mRNA following a shift to permissive conditions providing clear evidence of an EDC-mediated alteration in the TH response (see Figures 2.6 and 2.7). In contrast, *klf9* transcript abundance was augmented in back skin cultures maintained under chemical-free and nonpermissive temperature conditions that originated from tadpoles exposed *in vivo* to T_3 plus either EDC. Such observations lend support for the ability of an EDC to alter anuran biology under low environmental temperature conditions with the bioactive chemical changing the TH-induced "molecular memory" in the transcriptome of developing *Rana catesbeiana*.

The precise mechanism(s) of action of IBF and TCS on TH-mediated anuran development at environmentally-relevant concentrations is presently not known. The therapeutic mechanism of action of IBF is attributed to cyclooxygenase inhibition leading to a reduction in prostaglandin synthesis (Rainsford 2009). However, evidence for other mechanisms involving impact on transcription factor activity exists (Stuhlmeier et al. 1999; Tegeder et al. 2001; Veldhoen et al. 2014). TCS also appears to target transcription factors (Dann et al. 2011; Marlatt et al. 2013; Yueh et al. 2014). The present observation that the gene expression program can be altered at low temperature leading to a noticeable effect at permissive temperature is consistent with modulation of transcription factor status for both TCS and IBF; the associated mechanisms require further investigation.

The identification of temperature-dependent regulation of anuran postembryonic development and the capability to include a temporal uncoupling of induction from execution of metamorphosis has important ramifications towards environmental monitoring of chemical contaminant exposure events that may pose a risk for indigenous *Rana catesbeiana* populations. Observations of altered transcriptome profile at permissive temperature following an earlier low

temperature exposure of tadpoles with TH and EDC reinforce the potential of environmental contaminants to manifest their effects well after the original point source has dispersed. Given that many species exhibit a dependence upon temperature as an important modulator of development, a new dimension of EDC impact must be considered. With respect to temperature-associated effects on EDC action, attention towards anthropogenic changes in climatic conditions as it relates towards effects-based environmental monitoring must be assessed in order to ensure continued efficacy of current management practices.

2.5 Conclusions

Low temperature suspension of TH-induced metamorphosis in *Rana catesbeiana* tadpoles allows the examination of molecular events involved in progression through postembryonic development by exploiting a natural regulatory mechanism rather than the use of chemical inhibitors of TH production, RNA transcription, or protein translation. TH-mediated modulation in the abundance of mRNA encoding select transcriptional regulatory proteins is unaffected by environmental temperature and may represent components of the transcriptome involved in establishment of competence to progress through metamorphosis upon renewal of permissive temperature conditions. Further investigation into this TH-dependent process during nonpermissive conditions will help determine underlying molecular mechanisms involved in the temporal uncoupling of tissue-specific TH-modified transcriptomes from execution of the metamorphic program and provide important insight into the potential impact of EDCs on amphibia that demonstrate developmental stasis at low temperatures. Elucidation of the particular genes that contribute to this molecular memory will require expansion of the currently scarce transcriptomic resources available for *R. catesbeiana*; RNA-seq techniques that do not rely on an established genome offer precisely the tools to do so.

3 *De novo* transcriptome assemblies of *Rana* (*Lithobates*) *catesbeiana* and *Xenopus laevis* tadpole livers for comparative genomics without reference genomes

Abstract

In this work we studied the liver transcriptomes of two frog species, *Rana* (*Lithobates*) *catesbeiana* and *Xenopus laevis*. We used RNA-seq data to assemble and annotate these transcriptomes, and compared how their baseline expression profiles change when tadpoles of the two species are exposed to TH. We generated more than 1.5 billion RNA-seq reads in total for the two species under two conditions as treatment/control pairs. We *de novo* assembled these reads using Trans-ABySS to reconstruct reference transcriptomes, obtaining over 350,000 and 130,000 putative transcripts for *Rana catesbeiana* and *Xenopus laevis*, respectively. Using available genomics resources for *Xenopus laevis*, we annotated over 97% of our *Xenopus laevis* transcriptome contigs, demonstrating the utility and efficacy of our methodology. Leveraging this validated analysis pipeline, we also annotated the assembled *Rana catesbeiana* transcriptome. We used the expression profiles of the annotated genes of the two species to examine the similarities and differences between the tadpole liver transcriptomes. We also compared the gene ontology terms of expressed genes to measure how the animals react to a challenge by TH.

Our study reports three main conclusions. First, *de novo* assembly of RNA-seq data is a powerful method for annotating and establishing transcriptomes of non-model organisms. Second, the liver transcriptomes of the two frog species, *Rana catesbeiana* and *Xenopus laevis*, show many common features and the distribution of their gene ontology profiles are statistically indistinguishable. Third, although they broadly respond the same way to the presence of TH in their environment, their receptor/signal transduction pathways display marked differences.

3.1 Introduction

The inordinate value in having genomic information for a given species is uncontested, providing information regarding inheritance, disease, proteins, metabolites, and other regulatory molecules. However, only a miniscule fraction of eukaryotic species has more than the most basic genetic information available. For example, 8,624 genomes to date are listed in the Genomes Online Database (Pagani et al. 2012) (www.genomesonline.org, Accession date: August 22, 2104) that are on-going or completed, out of an estimated 8.7±1.3 million eukaryotic species worldwide (Mora et al. 2011).

Despite concerted efforts to obtain representative genomes such as the Genome 10K project (Genome 10K Community of Scientists 2009), sequencing the genomes of organisms of interest presents a considerable challenge in bioinformatics, and requires substantial resources that are beyond the budgets of most projects. On the other hand, RNA-seq has the potential to rapidly and economically transform a species' molecular knowledgebase. Transcriptomes provide information on complex biological processes, and give a picture on how the static genome behaves in dynamic environments. They enable the assessment of environmental impact factors, such as climate change – and of particular interest to our study – pollutants.

RNA-seq data are typically generated as tens to hundreds of millions of paired 75-mer to 150-mer reads. These reads are often mapped onto a completed, annotated genome scaffold for identification and quantitation (Trapnell et al. 2010). In the absence of a reference genome or when the reference is poorly reconstructed or annotated, *de novo* assembly of RNA-seq reads (Birol et al. 2009; Grabherr et al. 2011; Robertson et al. 2010; Schulz et al. 2012) is an enabling technology to study the transcriptomes of non-model species (Bazinet et al. 2013; Frischkorn et al. 2014; Mehr et al. 2013; Ribeiro et al. 2014; Sookruksawong et al. 2013). Although the analysis of the results of a *de novo* transcriptome assembly is a nontrivial task (Nakasugi et al.

2014; Xu et al. 2014), coupled with a quality controlled and streamlined bioinformatics pipeline, it is a cost-effective strategy to glean biological insights. In the present work, we demonstrate the use and value of the technology to compare the liver transcriptomes of two amphibian species when exposed to TH.

Amphibians are among the most threatened vertebrates on the planet (AmphibiaWeb 2012). They are also the only group where most of its members exhibit a life history that includes distinct independent aquatic larval and terrestrial juvenile/adult phases. The transition between the larval and juvenile phases requires substantial or complete remodelling of the organism (metamorphosis) in anticipation of a terrestrial lifestyle. This places amphibians in a unique position for the assessment of toxicological effects in both aquatic and terrestrial environments. Amphibians have an undeniable role as sentinel species, as a food source, and in insect control; yet over 60% of about 7,000 extant amphibian species are currently threatened or declining in numbers (AmphibiaWeb 2012).

Despite their established importance, only one completed amphibian genome is currently available from a model diploid laboratory frog, *Xenopus (Silurana) tropicalis* (Hellsten et al. 2010) – a species whose natural habitat is restricted to parts of Africa (AmphibiaWeb 2012; Hellsten et al. 2010). The most extensively used laboratory frog is *Xenopus laevis*, and among amphibians, it has the largest proportion of cDNA resources available on public databases (James-Zorn et al. 2013). However, due to challenges in dealing with its allotetraploid nature, only recently have portions of this species' genome become available. This species also has a natural range limited to Africa with some accidental introductions in the United States (AmphibiaWeb 2012).

On the other hand, several relatives of the most numerous amphibian species, the "true frogs" (*Ranidae*), can be found worldwide as native or introduced species. Despite their importance as environmental sentinels and ecosystem service providers, there are only limited genomic resources for these species (Kiemnec-Tyburczy et al. 2012; Savage et al. 2014).

Xenopus diverged from the true frogs over 200 million years ago (Sumida et al. 2004). This evolutionary divergence is accentuated by their differing life histories, behavior, and markedly different sex differentiation systems (Eggert 2004). Recent evidence suggests that the innate immune system of *Xenopus* is also fundamentally different from three frog families including the *Ranidae* (Kiemnec-Tyburczy et al. 2012). Furthermore, the current genomics tools developed for *Xenopus spp* are unfortunately inadequate for use with *Ranids*; *Xenopus tropicalis* reference genome representing an ineffective genome scaffold to study *Ranid* transcriptomes due incomplete or inaccurate annotation and extensive interspecies sequence divergence. A higher level of transcript annotation exists for a related species, *Xenopus laevis*, due to its long-standing use as a developmental model. Despite earlier cDNA sequencing efforts (James-Zorn et al. 2013), recent *de novo* RNA-seq experiments on *Xenopus laevis* embryos discovered thousands of novel transcripts (Blower et al. 2013), further accentuating the paucity of genomic resources even for this well-studied amphibian species.

The gap in knowledge regarding postembryonic development in amphibia is even more pronounced. Previous work has established a clear dependence of THs in amphibian metamorphosis, and although several landmark studies have characterized the genetic programs involved in select tissues (e.g. Atkinson et al. 1996; Das et al. 2002; Denver et al. 1997; Helbing 2012; Mukhi et al. 2010), changes involving the liver have been less studied (Atkinson et al. 1996; Helbing 2012).

The liver as the largest internal organ performs essential metabolic, exo- and endocrine functions, including bile production, metabolism of dietary compounds, detoxification, carbohydrate metabolism, and production of blood clotting factors and serum proteins (Zorn 2008). The precise role that the liver serves in each of these biological processes changes dramatically during the TH-dependent genetic reprogramming of this organ from larval to juvenile form (Atkinson et al. 1996; Mukhi et al. 2010). Biochemical changes include induction of the urea cycle, albumin synthesis, and globin switching, in addition to changes in immune system function that impact the liver (Gilbert et al. 1996).

Here we examine the TH response of the tadpole liver tissues of two frog species, *Rana catesbeiana* and *Xenopus laevis* using RNA-seq assays. Although the species under study lack finished reference genomes, we demonstrate that *de novo* transcriptome assembly methods constitute an enabling technology for comparative transcriptomics. Reconstructing the liver transcript sequences for the two species, we annotated lists of putative protein encoding mRNA, and performed comparative analyses on gene ontology to identify common and species-specific functional processes as they relate to TH status.

3.2 Materials and Methods

3.2.1 Ethics approval

Premetamorphic *Rana catesbeiana* tadpoles were locally caught (Victoria, BC, Canada) under a BC Ministry of Forests, Lands and Natural Resource Operations permit VI11-71459 while *Xenopus laevis* tadpoles were raised at the University of Victoria Aquatics Facility. Animal husbandry was carried out in accordance with the guidelines of the Canadian Council on Animal Care. The University of Victoria animal care committee specifically approved this study.

3.2.2 Sample collection

Premetamorphic *Xenopus laevis* (NF stage 54) and *Rana catesbeiana* (TK stage VI-X) tadpoles were exposed to 400 µM NaOH (ACP Chemicals Inc.) or to 10 nM T_3 (Sigma-Aldrich Canada Ltd.) in NaOH vehicle for 48 h as described previously (Veldhoen et al., 2014). Treated tadpoles of both species were euthanized using 0.1% (w/v) tricaine methanesulfonate in dechlorinated municipal water containing 25 mM sodium bicarbonate. Liver tissue was isolated and preserved in RNA*later* following the manufacturer's protocol (Life Technologies Inc.) and hepatic total RNA extracted using TRIzol reagent as described previously (Veldhoen et al. 2014). RNA-seq libraries were generated from 8 to 11µg of total RNA with RIN scores in the range from 6.6 to 7.2, using superscript random primer for double strand cDNA synthesis (SPCL Kit, Invitrogen, Waltham, MA). Generated libraries were sequenced using the HiSeq 2000 paired-end sequencing platform as per the manufacturer's instructions (Illumina Inc., San Diego, CA) to generate 2x75 base pair reads.

3.2.3 Transcriptome assembly

Each of the four libraries was assembled separately with Trans-ABySS (Robertson et al. 2010) over a range of k-mer sizes (every other k between 38 and 74 bp). Trans-ABySS was selected for its robust and competitive performance in assembling transcriptomes (see Birol et al. 2015b for a recent comparison to other available transcriptome assemblers). The final assemblies of the two conditions (T_3 and NaOH) were concatenated separately for each species. In these assemblies redundant sequences were removed by aligning contigs to themselves (with BWA (Li et al. 2009), default parameters); allowing no alignment gaps. Shorter contigs were tagged for removal when they had at least 95% match to a longer contig. Further, SGA (Simpson et al. 2012) was used on the final set of contigs to improve the contiguity and remove remaining redundant sequences by performing overlap-based contig extensions.

We used BioBloom Tools (Chu et al. 2014) to screen for possible bacterial and viral contaminations in the original reads, and filtered out contigs associated with these reads. The remaining contigs were used as the final set of assemblies for annotation and differential expression analysis. Assembly statistics in Table 3.1 refer to these final assemblies.

Tool versions, parameters and command lines for transcriptome assembly are detailed in Birol et al. 2015a, which includes scripts to reproduce our analysis.

3.2.4 Open reading frame (ORF) analysis

We used TransDecoder (Haas et al. 2013) with the default parameters to predict the ORFs in the final assemblies for both species.

The *sensitive* option of Bowtie (Langmead et al. 2012) was used to map the putative *Xenopus laevis* transcripts to the Xenbase *Xenopus laevis* 7.1 genome assembly and the transcriptome assemblies from a study on *Xenopus laevis* embryogenesis (Blower et al. 2013). We required at least half of the entire length of our *Xenopus laevis* transcriptome contigs to be used in the alignments.

3.2.5 Differential expression analysis

Bowtie (Langmead et al. 2012) with *multiple alignment* option, was used to separately align the control and treatment raw reads to the final assemblies in each species. To estimate the relative expression levels of reconstructed transcripts at gene level, we generated the raw median fold-coverage for all contigs based on these read alignments. Using DESeq (Anders et al. 2010) in *blind* mode and with *fit-only* option, we selected the final set of differentially expressed genes with p-value threshold 5%.

3.2.6 Transcript annotation

Contigs were aligned using the BLASTx or BLASTn programs from the BLAST+ software package (Camacho et al. 2009) against three reference databases: (1) the NCBI non-redundant (NR) database (retrieved 24 March 2014); (2) a collection of all amphibian transcript sequences, including mRNA from *Xenopus tropicalis* and *Xenopus laevis*, as well as *Rana catesbeiana* expressed sequence tags (EST) and transcriptome shotgun assembly (TSA) data from the green frog, *Rana clamitans*, in GenBank (retrieved 25 February 2014); and (3) the Ensembl *Homo sapiens* cDNA database from the GRCh38 genome assembly (retrieved 9 June 2014). To identify sequence orthologs between the species, tBLASTx was used with the *Rana catesbeiana* contigs to query the *Xenopus laevis* contigs generated from the livers in the present study. Alignments with a minimum E-value of $1x10^{-5}$ were considered, and the top alignment from each of the three databases was retained and merged. Preference was given to the Ensembl result for identification purposes due to the extensiveness of its associated annotations, while the result from the set of amphibian transcript sequences provided independent confirmation of the veracity of *de novo* assembled sequences.

3.2.7 Gene ontology (GO) analysis

GO analyses were performed using the Ensembl-annotated data. The Ensembl IDs were mapped to the corresponding UniProt IDs using ID mapper (www.uniprot.org). The unique UniProt IDs were used as input to retrieve the GO data for molecular function, cellular component, and biological process categories and subcategories within. In performing the gene ontology analysis, we used the UniProt database release 2014_07, dated 9 July 2014.

The normalized stacked bar plots in Figures 3.1 and 3.3 show the relative frequencies of categories over three domains of classification: (1) Biological processes describe molecular activities within cells, tissues and organs (15 categories); (2) Molecular functions relate to

chemical events, such as catalysis or enzyme activities (21 categories); and (3) Cellular components refer to where in a cell or its extracellular environment a particular transcript is active (nine categories). In plotting the figures, for molecular functions, we filtered out categories that had less than five representative transcripts for either species. The other two domains had all their categories meeting this criterion. In performing statistical significance tests, we considered all categories in each domain.

To establish statistical similarities between the relative frequencies of categories in both species, considering all or differentially expressed transcripts, we performed a series of X^2 tests. We applied a Bonferroni correction (Benjamini et al. 1995) factor of 6, for three domains of classification and two sets of comparisons (for all and for differentially expressed transcripts). We applied the same coefficient when we excluded the receptor activity from the list of categories in molecular functions. We performed statistical similarity tests for the relative frequencies of 45 categories for two sets of comparisons. However, seven of them had zero occurrence in the differentially expressed list of transcripts, hence were not tested for significance. To reflect the number of tests we performed, we used a Bonferroni correction factor of 83. Frequencies of observation of each category in all three domains, and the results of our hypothesis tests are provided in the Supplementary Materials.

3.2.8 Pathway analysis

Biological pathways were generated using the Ingenuity Pathway Analysis (IPA®, QIAGEN Redwood City). Uniprot IDs of differentially expressed genes and their associated log fold change ratios are used as the input for IPA. In the case of redundancies, which occurred when multiple contigs are annotated with the same ID, log fold change of the longest contig is considered representative for that gene. For each pathway, a p-value is calculated using the

Fisher's exact test against the null hypothesis that genes from that pathway are appearing in our list randomly. These p- values are then adjusted using the Benjamini-Hochberg correction. The analysis was performed with the IPA 2014 Spring release.

3.2.9 qPCR validation

Independent validation was performed on vehicle and T_3-exposed tadpoles (n=7-16) using qPCR as described previously in Veldhoen et al., 2014. All primers were subjected to a rigorous three tier QC procedure to ensure accurate detection and optimum performance (Veldhoen et al., 2014). The new immune system components primers and qPCR assay conditions are referred to in Appendix 3. Statistical significance was determined using the Mann-Whitney U test using R Studio software (R Core Team 2013).

3.2.10 Data availability

The generated reads and assembled transcripts are available through NCBI repositories through BioSample accession numbers SAMN03274104, SAMN03274105, SAMN03274106 and SAMN03274107, and BioProject ID PRJNA271360.

3.3 Results

We sequenced liver tissue transcriptomes of *Rana catesbeiana* and *Xenopus laevis* tadpole liver tissues exposed to T_3, and to the vehicle control NaOH. Previous analysis of these transcriptomes indicated that they exhibited an appropriate response to T_3, and, thus, they are representative of the TH-mediated induction process (Helbing 2012; Veldhoen et al. 2014). We assembled transcriptomes for the two species using Trans-ABySS (Robertson et al. 2010) following the protocol described in the Methods section. The putative *Rana catesbeiana* and *Xenopus laevis* transcript fragments were reconstructed in 353,253 and 134,203 contigs, respectively. These draft liver transcriptomes formed the basis of our expression level analyses

and functional gene annotation. Collected sequencing data and summary statistics for transcriptome assemblies are shown in Table 3.1.

Table 3.1 RNA-seq data and transcriptome assembly results.

Species		# RNA-seq reads (10^6)	Average fragment length (bp)	# Contigs	Contig N50 (bp)	Transcriptome size (Mbp)	GC content (%)
Rana catesbeiana	Control	347	375	353,253	1,815	429	43.3
	T_3	472	360				
Xenopus laevis	Control	396	394	134,203	1,596	151	41.5
	T_3	428	376				

Sequences were generated using 75 bp paired end reads.

We assessed the quality of the final transcriptome assemblies using 248 highly conserved core eukaryotic genes (Parra et al. 2009) (CEGs), and showed that we were able to reconstruct all CEGs for *Rana catesbeiana*, and miss only one for *Xenopus laevis*. Also, using BWA (Li et al. 2009) and allowing for multiple alignments, we were able to align 96.1% and 94.6% of raw reads back to the assembled *Rana catesbeiana* and *Xenopus laevis* transcriptomes, respectively. These results indicate a good quality assembly, suitable for downstream biological analyses.

To examine the extent of ortholog overlap between the two species, tBLASTx was used to compare the contig sets from *Rana catesbeiana* and *Xenopus laevis*. We observed that 41% of the *Rana catesbeiana* contigs had orthologous sequences in the assembled *Xenopus laevis* liver transcriptome. Also, ORF analysis using TransDecoder (http://transdecoder.sf.net) indicated that 15% (51,720) of *Rana catesbeiana* contigs had putative coding potential, similar to 14% (18,328) of *Xenopus laevis* contigs.

Using BLAST tools (Altschul et al. 1990), we compared our putative liver transcripts with three potential annotation resources: a collection of *Rana catesbeiana* ESTs, all amphibian cDNA sequences from NCBI nucleotide database (Pruitt et al. 2007), and *Rana clamitans*

transcriptome shotgun assembly (TSA) sequences from NCBI; the non-redundant NR database of NCBI (Pruitt et al. 2007); and the human protein database of Ensembl (Flicek et al. 2008). The NR database is a curated sequence database of genomes, transcripts and proteins, and aims to provide a non-redundant representation of those sequences. The amphibian cDNAs we collected from NCBI are not yet curated into NR, but most of the Ensembl human proteins are represented in NR.

The database that yielded the highest number of hits was the amphibian cDNA collection. Using those sequences, we were able to map 257,793 (73%) of *Rana catesbeiana* and 105,030 (78%) of *Xenopus laevis* contigs with BLASTn. Although BLASTx alignments to NR yielded fewer hits (129,326 for *Rana catesbeiana* and 69,582 for *Xenopus laevis*) they did provide 20,813 unique additional annotations for the *Rana catesbeiana* contigs, and 4,607 unique additional annotations for the *Xenopus laevis* contigs. Finally, BLASTx alignments to Ensembl human proteins returned 92,524 hits for *Rana catesbeiana* and 60,243 hits for *Xenopus laevis* with an expected limited number of unique hits for both species (150 and 11, respectively). Overall, using these three resources, we were able to annotate 278,799 (79%) of the *Rana catesbeiana* and 109,649 (82%) of the *Xenopus laevis* contigs in our transcriptome assemblies (Table 3.2).

Although the contribution of Ensembl human proteins was marginal in this scheme, alignments to this database were instrumental in functional analysis using GO terms, as the collection represents the most extensively annotated data.

Table 3.2 Annotation of assembled transcriptome contigs. Three annotation approaches were taken: using BLASTn against a collection of nucleotide sequences of select amphibia, BLASTx against the non-redundant NCBI protein database, and using BLASTx against the Ensembl (*Homo sapiens*) database for subsequent GO analysis.

Contig annotations with	*Rana catesbeiana*	*Xenopus laevis*
BLASTn (amphibian)	257,793	105,030
BLASTx (NR)	129,326	69,582
BLASTx (Ensembl *Homo sapiens*)	92,524	60,243
Total annotated combining all three methods	278,756	109,649

With the described annotation process, we were not able to observe sequence similarities between 24,554 putative *Xenopus laevis* transcripts and the three resources used. We compared this remaining set with the Xenbase *Xenopus laevis* 7.1 genome assembly (http://gbrowse.xenbase.org/fgb2/gbrowse/xl7_1/), and with the data reported in a recent publication evaluating *Xenopus laevis* embryogenesis (Blower et al. 2013). Of these outstanding transcripts, 21,115 matched to Xenbase genome assembly, and a further 610 had counterparts in the embryogenesis transcriptome assemblies. Thus, combining the annotation results from all resources, out of 134,203 putative *Xenopus laevis* transcripts in the present study, 97.84% (131,311) matched to at least one of the relevant external resources available for this species. The remaining 2,492 putative transcripts that lack support in the literature are mostly composed of shorter sequences (N50 of 529 bp, compared to an overall *Xenopus laevis* assembly N50 of 1596bp). We note that shorter sequences are relatively harder for proper alignment and annotation, and a subset of these may represent *bona fide Xenopus laevis* transcripts yet to be annotated.

We performed enrichment analysis of GO terms on our assembled transcriptomes using the UniProtKB database (www.uniprot.org) with three perspectives: (1) Biological Processes, (2) Molecular Functions, and (3) Cell Components (Figure 3.1). As indicated above, for the functional annotations we used alignment of Ensembl human protein information to the putative

transcripts. We subsequently mapped the Ensembl-derived human protein IDs onto UniProt entries.

Using this approach, 19,514 and 17,831 unique UniProt accession identifiers (AC IDs) were generated from the 92,527 and 60,243 original Ensembl hits for *Rana catesbeiana* and *Xenopus laevis*, respectively, indicating that the relative annotations were highly similar.

We observed that the profiles comprising the Biological Processes and Cellular Components were not significantly different between the two species (Bonferroni-corrected (Benjamini et al. 1995) X^2 p-value threshold of 5%) (Figure 3.1 and Supplementary Table 1 (Birol et al. 2015a)). An example is the Immune Function category, where 1,426 and 1,174 Uniprot AC IDs were identified for *Rana catesbeiana* and *Xenopus laevis*, respectively (Supplementary Table 2 (Birol et al. 2015a)). Of these, 1,072 were common to both species representing 75% and 91% of the identified hits in this category, respectively (Supplementary Table 2 (Birol et al. 2015a)).

In contrast, initial evaluation of Molecular Functions revealed significant interspecies differences (Supplementary File 1 (Birol et al. 2015a)). However, when we removed the category of receptor activity from the list, the distributions of the two species were again statistically indistinguishable. The receptor activity category revealed a higher proportion of independent UniProt AC hits in *Rana catesbeiana* (713; Supplementary Table 3 (Birol et al. 2015a)) compared to *Xenopus laevis* liver (471; Supplementary Table 3 (Birol et al. 2015a)). It is interesting to note that 410 of these hits are common to both species representing 56% and 87% of the identified hits in this category for *Rana catesbeiana* and *Xenopus laevis*, respectively (Supplementary Table 3 (Birol et al. 2015a)). About half of these are linked to developmental processes including angiogenesis and homeostatic processes such as protein, nucleotide, lipid, and carbohydrate metabolism. Closer inspection of those UniProt AC IDs that differ

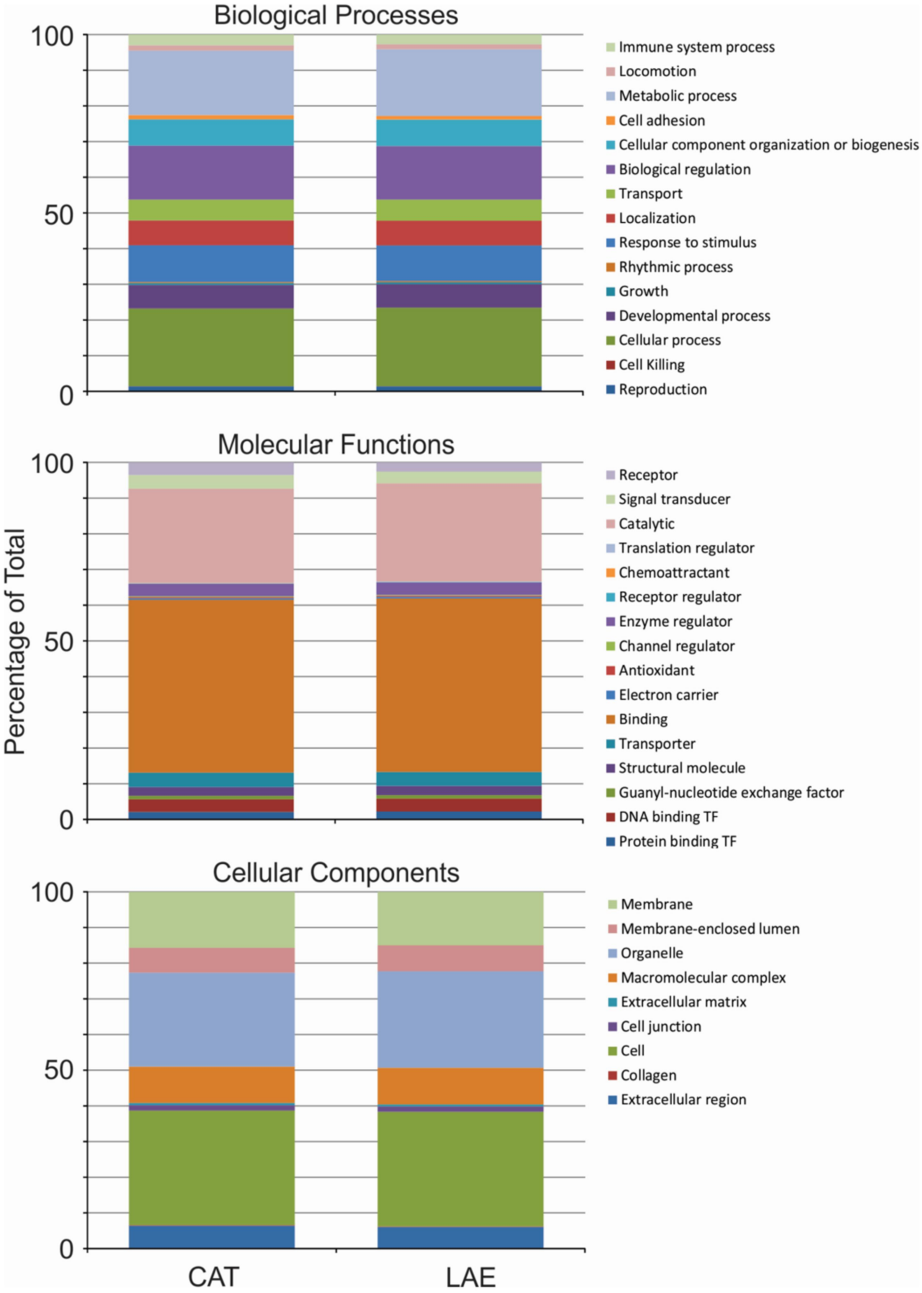

Figure 3.1 GO classification of all reconstructed *Rana catesbeiana* and *Xenopus laevis* liver transcripts with UniProtKB AC numbers. The two series in each stacked bar plot correspond to all *Rana catesbeiana* (CAT) and *Xenopus laevis* (LAE) transcripts.

between the species indicated increased representation in *Rana catesbeiana* of signal transduction components (246 versus 51, respectively), particularly transmembrane G protein-coupled receptors.

Using the aligned reads and the DESeq software (Anders et al. 2010) set at three stringency levels (p=0.2%, 2% and 5%), we analyzed differential expression between treatment-control pairs for each species (Figure 3.2). At a p-value threshold of 5% *Rana catesbeiana* and *Xenopus laevis* transcriptomes had 11,453 and 3,016 differentially expressed transcripts, respectively. As an important RNA-seq validation step, seven known T_3-responsive gene transcripts were evaluated in both species by qPCR (Appendix 2). While the RNA-seq data was derived from n=1 per treatment, the qPCR data was derived from a minimum of 7 individuals per treatment group. There was strong concordance between the fold change values of the two independent methods supporting the integrity of the RNA-seq observations.

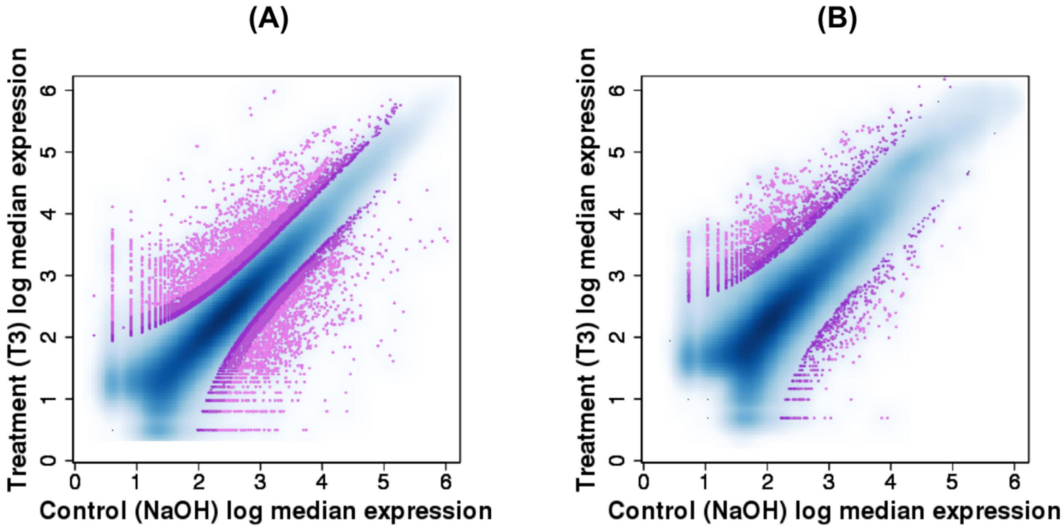

Figure 3.2 Differential expression of assembled transcripts for (A) *Rana catesbeiana* and (B) *Xenopus laevis*. Three shades of purple designate p-values of differential expression estimates 0.05, 0.02 and 0.002, lighter colours indicating lower thresholds.

We further analyzed the biological functions of the differentially expressed transcripts using their inferred roles in GO and pathway databases. Performing GO analysis on the differentially expressed transcripts, we observed statistically significant differences in biological processes and metabolic functions, the latter only marginally falling below the Bonferroni-corrected threshold (Figure 3.3, Supplementary Table 1 (Birol et al. 2015a)). We further performed z-tests on the relative frequencies of categories to observe that differentially expressed transcripts with functions in immune system processes differ between the two species. Of the 106 and 24 UniProt AC IDs (*Rana catesbeiana* and *Xenopus laevis*, respectively) identified by DESeq as T_3-responsive, only eight were common (Supplementary Table 5 (Birol et al. 2015a)). This strongly contrasts with the high degree of commonality identified in the overall Uniprot AC ID profiles of the two species (Supplementary Table 2 (Birol et al. 2015a)). In the receptor activity category, a significant difference between species was maintained upon T_3 exposure. However the nature of the profile differed with *Rana catesbeiana* exhibiting 10 times more transcripts in this category compared to *Xenopus laevis* (50 for *Rana catesbeiana versus* five for *Xenopus laevis*; Supplementary Table 6 (Birol et al. 2015a)). Only two sequences were common between the two species in this category (Supplementary Table 6 (Birol et al. 2015a)). See Birol et al. (2015a) Additional File 1 for detailed reports on the GO analysis and the hypothesis tests performed.

Further examination of the impact of T_3 exposure on the liver transcriptomes was accomplished by analyzing the UniProt AC annotated transcripts of the two species with the IPA tool (Qiagen). Figures 3.4 and 3.5 present a prioritized list of the top 25 significantly-impacted pathways of the *Rana catesbeiana* and *Xenopus laevis* liver transcriptomes organized according to the number of observed pathway components. Over half of these pathways are shared between the two species and include acute phase response signaling, several pathways involving Rxr

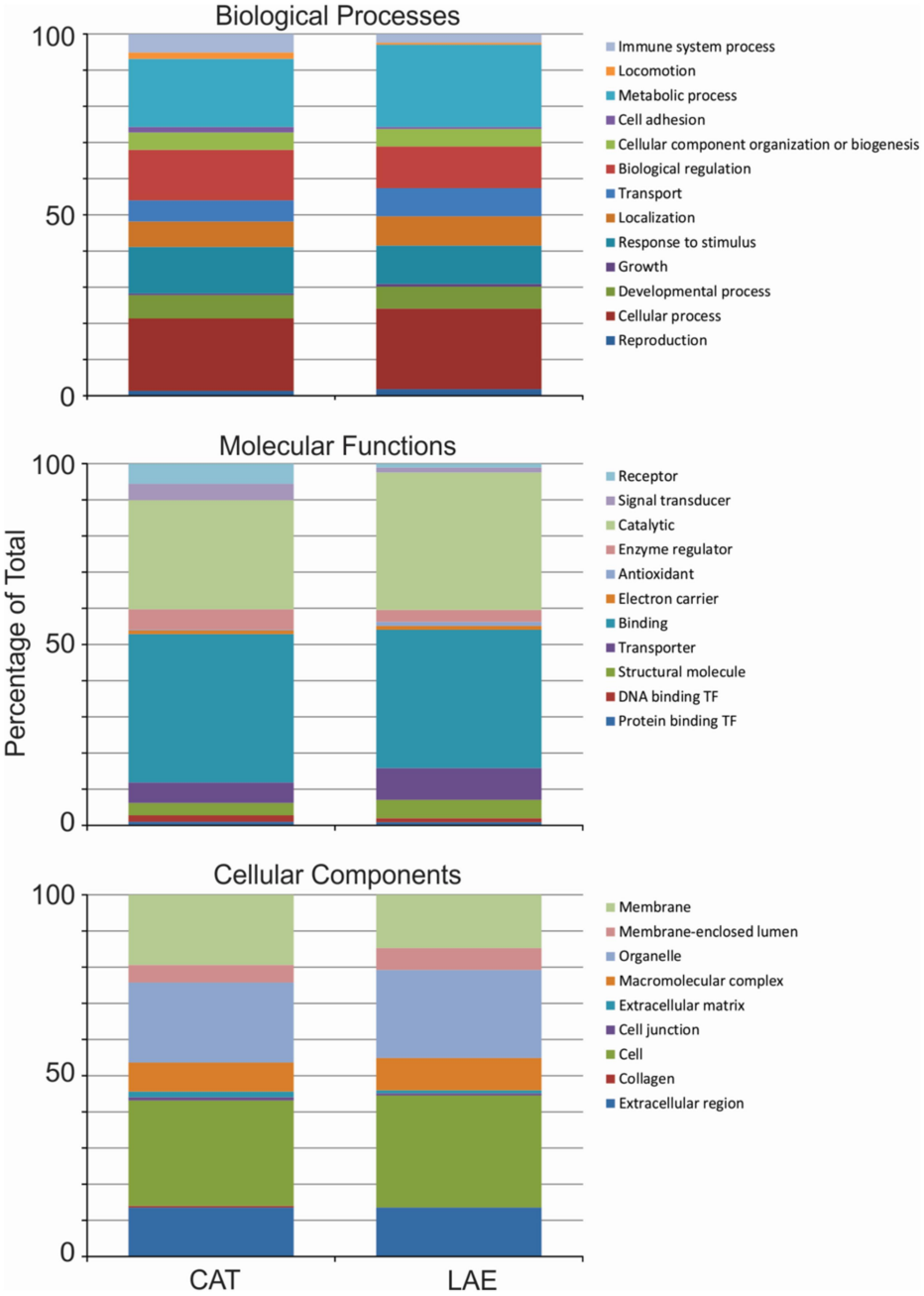

Figure 3.3 GO classification of DESeq-selected, TH-responsive *Rana catesbeiana* and *Xenopus laevis* liver transcripts with UniProtKB AC numbers. The two series in each stacked bar plot correspond to differentially expressed *Rana catesbeiana* (CAT) and *Xenopus laevis* (LAE) transcripts, with a p-value threshold of 5%.

function, lipid metabolism, melatonin/seratonin degradation, urea cycle, and estrogen biosynthesis (Figures 3.4 and 3.5). Notable differences included the marked alteration of the complement and coagulation systems and antigen presentation pathways in *Rana catesbeiana* compared to *Xenopus laevis* (Figure 3.4) and the greater involvement of the eukaryotic initiation factor 2 (EIF2) protein translation pathway and pathways involving glycosaminoglycans and cholesterol metabolism (Figure 3.5). The involvement of the immune system in the T_3 response of *Rana catesbeiana* was independently supported by qPCR analysis on select immune-related gene transcripts (Figure 3.6).

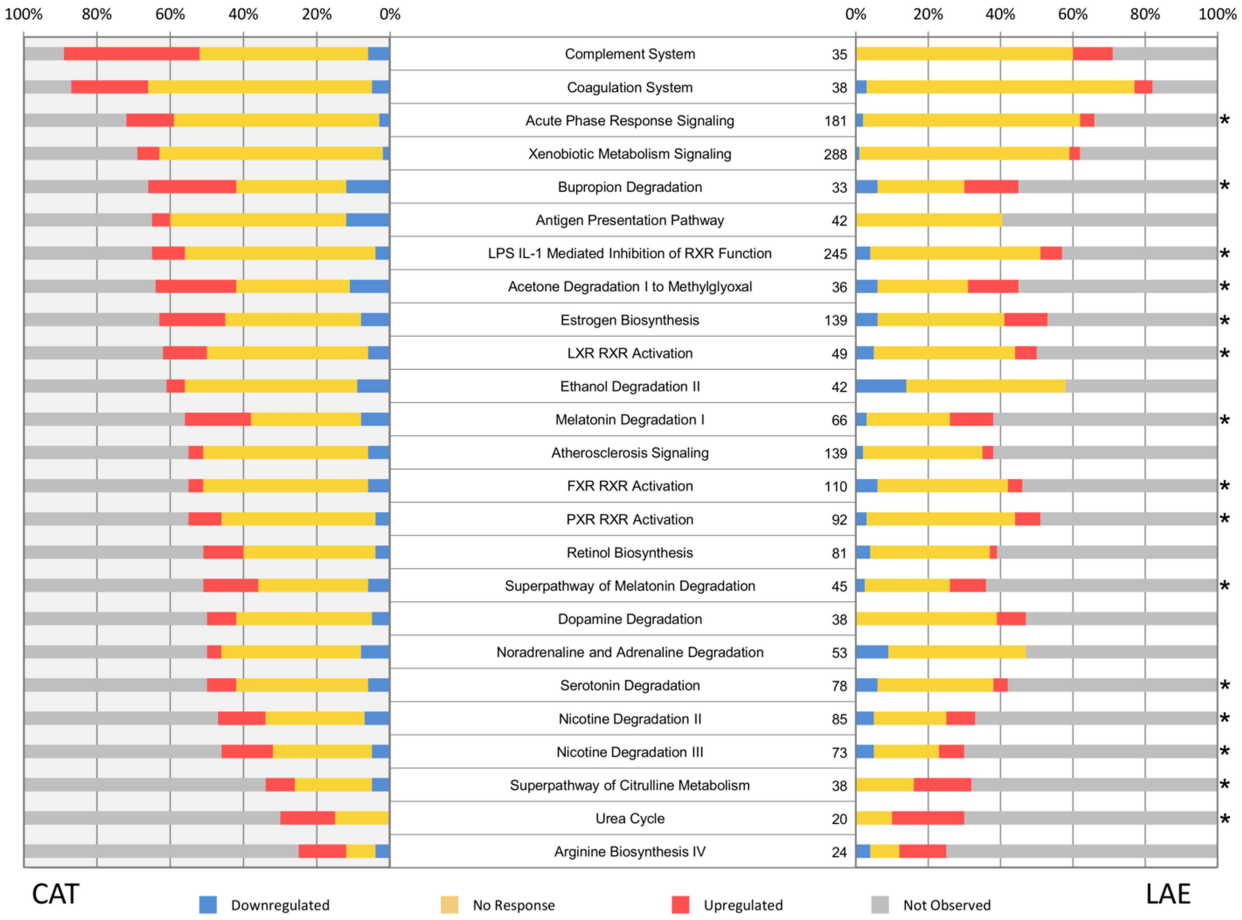

Figure 3.4 Pathway analysis for liver transcripts from *Rana catesbeiana* (CAT) and *Xenopus laevis* (LAE). Top 25 impacted pathways after TH treatment for Rana catesbeiana ranked by the highest proportion of overall observed genes. The pathway names are indicated in the center of the figure with the total number of genes known in each IGA pathway indicated. The asterisk indicates those pathways that are found in the top 25 list of Xenopus laevis. The colour coded bar plots illustrate the percentage of the total number of gene transcripts in a pathway that are downregulated (blue), non-responsive (yellow), upregulated (red) or not observed in the experiment (gray) relative to the control condition. Differentially expressed transcripts were determined using a p-value threshold of 5%.

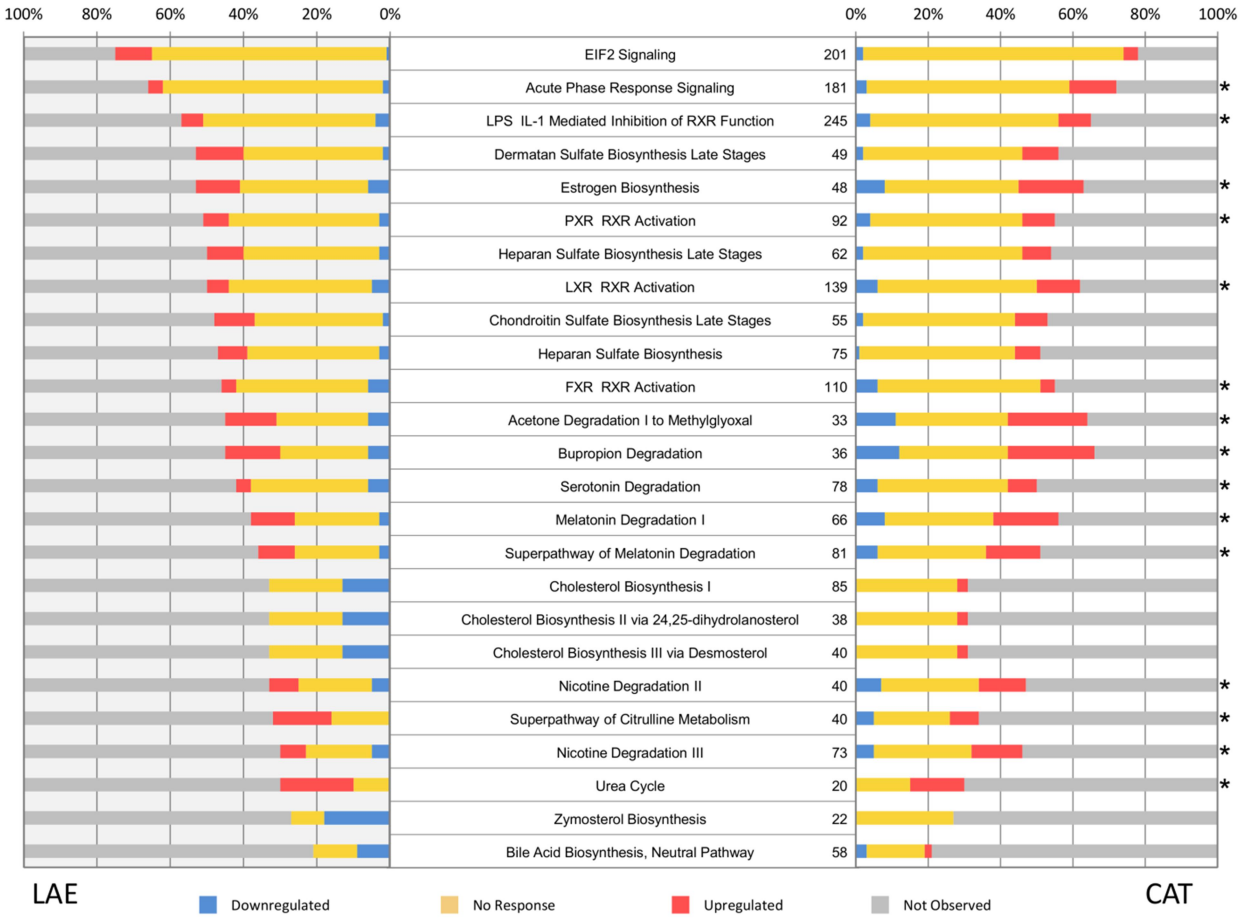

Figure 3.5 Pathway analysis for liver transcripts from *Xenopus laevis* (LAE) and *Rana catesbeiana* (CAT). Top 25 impacted pathways after TH treatment for Xenopus laevis ranked by the highest proportion of overall observed genes. The asterisk indicates those pathways that are found in the top 25 list of Rana catesbeiana. Plot details are as in the Figure 3.4 legend.

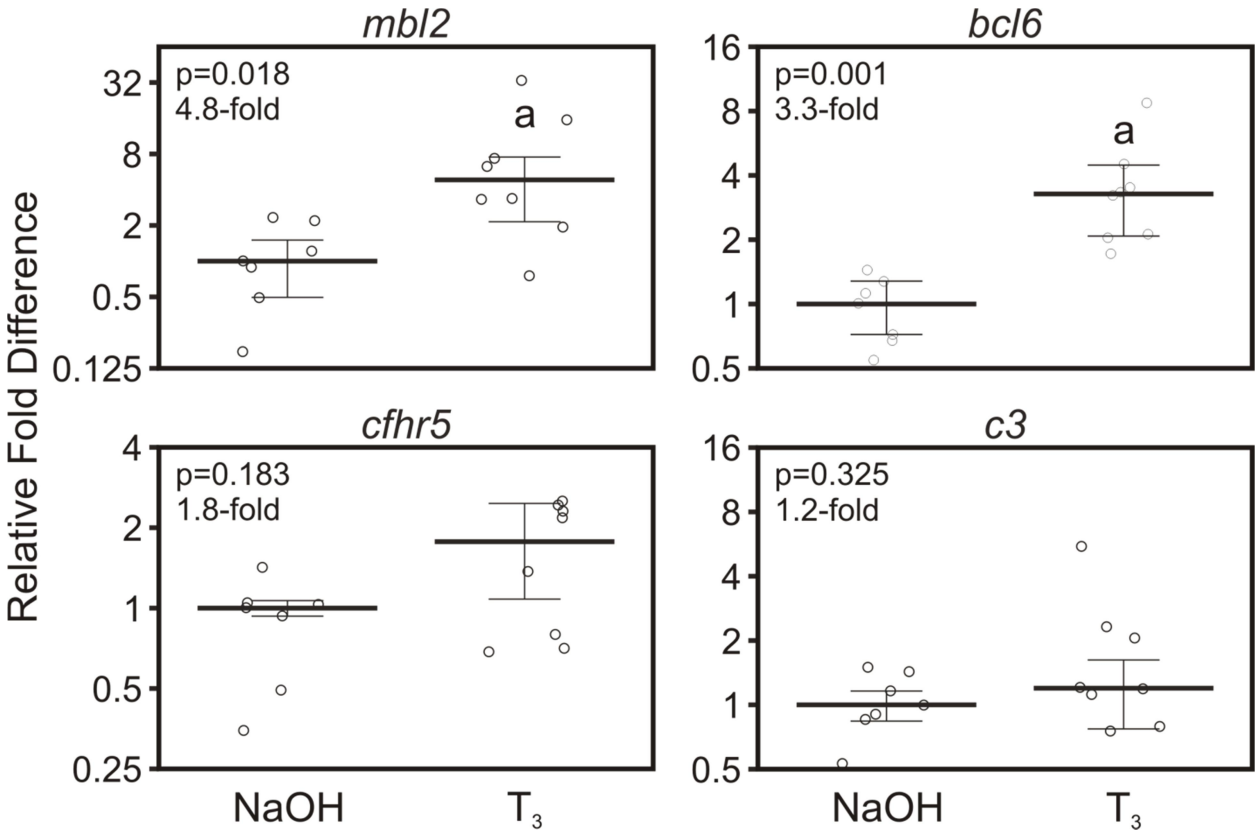

Figure 3.6 qPCR analysis of immune system components in *Rana catesbeiana* treated with vehicle control (NaOH) or 10 nM T_3. Individual animals are depicted by open circles with the median values (horizontal black bars) and median absolute deviations (whiskers) shown.

3.4 Discussion

Application of *de novo* sequencing and assembly approaches has enabled the direct comparison of the expressed gene information contained in two distantly related amphibian species. Both *Xenopus laevis* and *Rana catesbeiana* are pivotal species of interest in examination of the role of THs during the postembryonic metamorphic process. Although similarities in the response patterns have been implied, the extent of similarity and differences linked to their species specific natural histories have not been directly investigated. For example, *Xenopus laevis* frogs remain aquatic, while *Rana catesbeiana* frogs adopt a semi-terrestrial lifestyle around the riparian areas they inhabit. Moreover, it is known that these two species have different sex determination systems: the chicken-like ZW for *Xenopus laevis*, and the XY for *Rana catesbeiana*, similar with humans (Eggert 2004). Such differences may influence their sensitivity to environmental contaminants that act through disruption of endocrine systems (Helbing 2012).

Our study presents a high quality assembly of the liver transcriptome of developmentally-matched premetamorphic tadpoles of both species. The quality of the transcriptomes produced is accentuated by the fact that 98% of contigs we assemble have counterparts when compared to the currently available genomic resources (www.xenbase.org) (Blower et al. 2013). Also, the contig sizes obtained in the present study are longer than those published previously (Blower et al. 2013; Savage et al. 2014). Most next generation sequencing data available are currently derived from *Xenopus tropicalis* due to its amenable genetics and genome resources, and the current estimated number of genes for this species is ~28,000 with an average transcript length of 1,300 bp (Nordberg et al. 2014). Our observations indicated that the transcript N50 lengths for both *Rana catesbeiana* (1,815 bp) and *Xenopus laevis* (1,596 bp) species were highly consistent

with this estimate. These contig sizes are approximately three-fold longer than those reported from pyrosequencing and *de novo* assembled transcriptomes of two *Ranid* species (Savage et al. 2014). Even though the number and size of reconstructed transcripts varied between the two species, the fact that we were able to obtain close to 20,000 UniProt AC IDs suggests that the generated contig assemblies are most likely a comprehensive representation of the tadpole liver transcriptome. We think the differences in the assembly statistics are confounded by the variable size and complexity of the two transcriptomes, as well as the biology of the two species.

The results of the current study explore the limits of utility of *de novo* transcriptome assemblies for functional annotation. While the existence of the vast majority of contigs as *bona fide* transcripts is now clear, relatively few could actually be annotated sufficiently for subsequent GO analyses. Moreover, roughly 85% of the transcripts did not have putative ORFs in both species; consistent with an observation recently made by Blower *et al*. during *Xenopus laevis* embryogenesis (Blower et al. 2013). This remarkable cross-species consistency supersedes gene ploidy levels and, in fact, has been observed in investigations of human, mouse, and zebrafish where long noncoding RNAs (lncRNAs) are spliced and polyadenylated but lack protein coding ability (Guttman et al. 2009; Guttman et al. 2011; Guttman et al. 2010; Ulitsky et al. 2011). Therefore more attention should be given to such ORF-negative RNAs whose implied biological significance is currently not known.

In contrast to the relative richness of genomic and transcriptomic information available for *Xenopus* species, relatively few resources are available for the *Ranidae*. The present work fills a critical knowledge gap, and provides relevant comparative information regarding the gene expression profiles of the liver in the premetamorphic tadpole with or without exposure to the metamorphosis-inducing TH. It was of great interest to establish the degree of similarity of this

tissue's responsiveness between species, and we observed that *Rana catesbeiana* and *Xenopus laevis* liver transcriptomes were highly similar to each other yielding essentially superimposable GO profiles.

The liver undergoes a substantial genetic reprogramming during TH-dependent metamorphosis, but the life histories of these two frog species suggest that the liver is likely to display marked differences in some pathways linked to their diverse ecological niches as frogs. We found that the vast majority of the transcript profiles of the livers of premetamorphic tadpoles are highly comparable to each other with the exception of the receptor/signal transduction-related pathways; particularly with respect to G protein signaling. This may translate into downstream differences in developmental outcomes, but could also contribute to differential species sensitivities to toxicants and environmental contaminants (Relyea et al. 2009).

In the context of precociously-induced metamorphosis, exposure of premetamorphic tadpoles to TH results in a substantial reprogramming of the liver in both species. However, the remarkable similarity of response to hormone action extended across most pathways including lipid signaling and metabolism, the urea cycle, and pathways involving cytochrome P450 mixed function oxidases (*e.g*. acute phase response signaling, lipid metabolism, melatonin/seratonin degradation, and estrogen biosynthesis). All of these pathways have previously been identified as TH and/or metamorphic targets (Atkinson et al. 1996; Duarte-Guterman et al. 2010; Duarte-Guterman et al. 2010; Duarte-Guterman et al. 2011; Ichu et al. 2014; Langlois et al. 2010; Langlois et al. 2011; McMenamin et al. 2014; Rollins-Smith 1998).

Despite these marked similarities across the two amphibian species, species-specific differences in receptor/signal transduction pathways were maintained following TH induction.

However, the most prominent species difference observed was the degree of involvement of the immune system in response to T_3 exposure. Although a large number of transcripts associated with immune system functions in complement, coagulation, and antigen presentation pathways were detected in both species, there was a high degree of primarily elevated responses found in *Rana catesbeiana* compared to *Xenopus laevis*. This could represent a fundamental difference in the TH-mediated maturation of the immune system (Rollins-Smith 1998). *Xenopus* species have proven to be excellent models for the comparative and developmental study of the immune system in vertebrates (Robert et al. 2009). Unfortunately though, the functional genomics of the immune system are still poorly understood for thousands of other frog species, many of whom transition to a terrestrial lifestyle post metamorphosis, in contrast to the completely aquatic *Xenopus*. In fact, recent work on the immunomes of two frog species, *Espadarana prosoblepon* and *Lithobates yavapaiensis*, identified significant divergence in inflammatory response and acquired immunity transcripts relative to innate immunity and immune system development transcripts (Savage et al. 2014). The implications of the findings presented herein and in the Savage *et al*. study (Savage et al. 2014) suggest that there are different propensities for responding to immunological cues; an important factor since many amphibian (and an increasing number of reptilian and fish) populations are severely threatened by emerging infectious diseases including *Batrachochytrium dendrobatidis* (Bd) fungus and iridoviruses (Daszak et al. 1999; Fisher et al. 2012; Grayfer et al. 2012). Susceptibility to disease is known to be developmental stage dependent (Rollins-Smith 1998), but the factors that determine whether an organism lives or dies as a result of infection are not known. Recent work using microarrays on Bd-infected *Xenopus tropicalis* (Rosenblum et al. 2009) and iridovirus-infected fathead minnows (Cheng et al. 2014) reveal that there remains much to learn about the interplay between infection and

immunity, and why some animal populations are more susceptible than others to these diseases (McMahon et al. 2014). Our present work adds additional insight towards the interplay between hormone action and immune system maturation in the context of anuran development.

3.5 Conclusions

Rana catesbeiana, and *Xenopus laevis* represent 260 million years of evolutionary divergence with very different genome sizes and ploidy levels. The present study demonstrates that the *de novo* transcriptome sequencing and assembly method we report is a successful approach that provides a valuable foundation from which to tackle critical biological issues that can extend beyond what can be gleaned from a restricted number of laboratory-reared species available, as well as support investigations related to comparative biology. As such, our methodology can act as a guideline for studying species that currently lack an associated reference genome.

4 *De novo* assembly and synthesis of a shared reference transcriptome from replicate *Rana* (*Lithobates*) *catesbeiana* back skin samples

Abstract

Motivation: RNA-seq has provided comprehensive transcriptome profiling for many species with established reference genomes, and *de novo* sequence assembly techniques extend this technology to non-model organisms with sparse genomic resources. Despite of this appealing feature of *de novo* RNA-seq assembly, analyzing these *de novo* assembled transcripts remains a complex and computationally challenging task. This problem is especially aggravated for comparative transcriptome studies in which several RNA-seq datasets need to be analyzed and compared together.

Results: We present a streamlined method for assembling multiple *de novo* RNA-seq assemblies into a single reference transcriptome that contains a representative of all expressed transcripts, and demonstrate its utility by applying it to replicate back skin Illumina read libraries from control and TH-treated *Rana catesbeiana* assembled by Trans-ABySS. Our approach improved assembly contiguity of the reference back skin transcriptome, and identified expression of RNA processing and cell cycle-associated genes as those most changed by TH treatment.

4.1 Introduction

RNA-seq is a powerful technique for characterizing and quantifying transcriptomes. It produces tens to hundreds of millions of reads and has a broad range of applications including detection of splice variants, identifying post-transcriptional RNA editing events and gene/transcript expression analysis (Ozsolak et al. 2011; Wang et al. 2009). However, the primary objective of RNA-seq in many biological studies is to perform comparative

transcriptome studies between several different samples, individuals and/or biological conditions (Oshlack et al. 2010). Unlike PCR or microarrays, certain RNA-seq techniques do not require any prior knowledge of the organism being studied and therefore RNA-seq is an enabling technology for organisms with limited or nonexistent genomic resources (Francis et al. 2013; Hornett et al. 2012; Martin et al. 2011).

Several programs exist for *de novo* RNA-seq assembly. These include Trans-ABySS (Robertson et al. 2010), Oases (Schulz et al. 2012), and Trinity (Grabherr et al. 2011). These tools start by loading all the subsequences of length *k*, or *k-mers*, of the reads into a de Bruijn graph data structure and then extend the k-mer sequences based on unambiguous overlaps between them; this way they assemble the RNA-seq data into transcript sequences, called *contigs*. While *de novo* assembly of RNA-seq is a powerful technology, utilizing these *de novo* assembled transcriptomes for downstream analysis remains a daunting and computationally expensive task (Nakasugi et al. 2014; Vijay et al. 2013).

The difficulty of this task is exacerbated in comparative transcriptomic studies where several samples with different features must be *de novo* assembled to a single reference transcriptome that includes a representative for all expressed transcripts across all samples (Davidson et al. 2014). Creating these shared reference transcriptomes is essential for multi-sample scenarios because each sample must be compared to the same reference for meaningful analyses to be performed. The typical approach to generate these references is to combine all the raw reads and assemble them together (Haas et al. 2013). While this approach is simple and straightforward, it is inefficient due to the size and complexity of the de Bruijn graph that results from the large pool of k-mers. By combining sequence data from multiple genetically distinct individuals, it is also prone to chimeric misassemblies due to allelic variation. While most *de novo* transcriptome

assemblers use strategies to resolve these ambiguities to some extent, this is a computationally challenging problem that is exacerbated when ambiguities accumulate (Góngora-Castillo et al. 2013).

An alternate approach that does not suffer from these shortcomings is to generate a shared reference transcriptome from independent *de novo* assembly of each samples' reads. Herein we introduce a method to generate a shared reference transcriptome from multiple RNA-seq samples and automatically calculate expression of each contig per sample. It may be applied to the output of any *de novo* assembler and does not change the assembled sequences (see Raymond *et al*. for a recent comparison between *de novo* transcriptome assemblers). We demonstrate the power of this sequence processing strategy in improving replicate RNA-seq assemblies and its efficiency and value in performing analyses that yield novel biological insights into the TH-induced metamorphic gene expression program of *Rana catesbeiana*.

4.2 Methods

Each set of reads is first assembled independently and then iteratively concatenated each assembly and removes redundant sequences to create a shared reference that describes all expressed transcripts while avoiding over-representation. The original reads are aligned to the reference using BWA (Li et al. 2009) and differentially expressed transcripts are identified using the DESeq (Anders et al. 2010) or DESeq2 (Love et al. 2014) Bioconductor packages (whichever desired by the user) and produces publication-ready visualization of the results in multiple formats.

4.2.1 Pipeline overview

Here we present an overview of the methods used in different stages of the pipeline, which are summarized in Figure 4.1.

4.2.1.1 Generating reference transcriptomes from separate *de novo* assemblies

The individual assemblies of each sample are processed iteratively. At each stage, the two libraries are concatenated and end-to-end gap-free alignments are produced using *bowtie2* (Langmead et al. 2012) allowing for a very small number of mismatches (*very-sensitive* option in *bowtie2* with at least 98% sequence match). After each alignment, each set of related transcripts is analyzed and only the longest contig from each set of alignment-related sequences is retained. In the process of generating the reference transcriptome none of the assembled sequences are modified, but rather individual sequences are either retained or removed from the final set based on particular criteria.

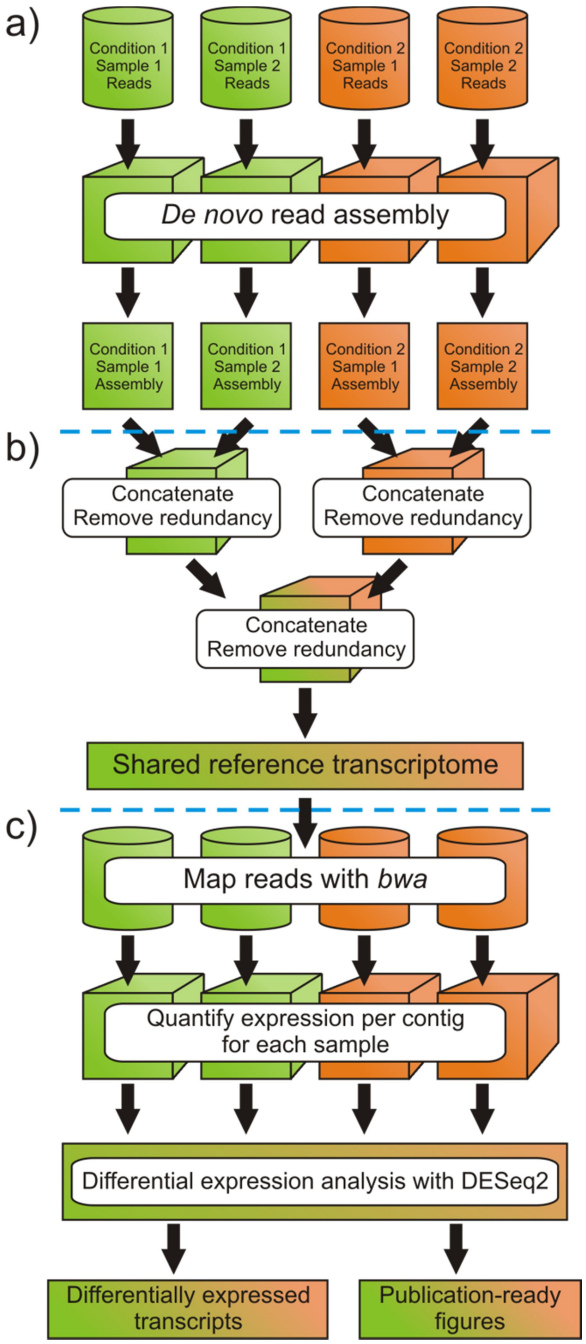

Figure 4.1 Data preparation and flow through the pipeline. In this example we have two conditions with 2 samples per condition. (a) RNA-seq libraries are de novo assembled separately. (b) The pipeline generates a shared reference transcriptome over several rounds of concatenation and redundancy removal. (c) It then quantifies the expression level of each contig for each sample and invokes DESeq2 to perform differential expression analysis.

4.2.1.2 Read abundance quantification

Read abundance calculations are necessary for estimating transcriptome expression profiles. BWA (Li et al. 2009) is used to align the original reads in each input RNA-seq library to the reference transcriptome. Using these alignments, it computes the number of RNA-seq reads aligned to each transcript per sample.

While counting, reads that can be aligned to two or more transcripts with equal mapping quality are discarded, which makes definitive identification of which transcript the read originated from impossible. These multi-mapped reads can be drawn from *bona fide* transcript splice variants and isoforms due to their partial sequence overlap, or from lingering redundantly-assembled sequences. By discarding the multi-mapped the total number of reads for some transcripts is undercounted, but because we are discarding the same fraction of reads between samples the relative counts between those similar transcripts will still be correct (Anders et al. 2015).

Discarding these reads is especially important for reducing false positives during differential expression analysis. This is illustrated by the hypothetical scenario of two transcripts with a region of identical sequence, where one transcript is highly expressed in one treatment condition while the other is refractory. If the multi-mapped reads are not discarded, the erroneous sharing of reads from the over-expressed transcript will artefactually increase the counts for the other transcript as well. If severe enough, this over-count could lead to falsely identifying the refractory transcript as differentially expressed.

In addition to raw counts, other measures of transcript abundance for each transcript per sample are output, including Reads Per Kilobase of transcript length per Million reads mapped (RPKM) and Fragments Per Kilobase of exon per Million fragments mapped (FPKM). Alternatively, the pipeline can generate the median coverage for each transcript, which is the

depth of coverage in the middle point of each transcript. For the latter feature, *Bedtools* (Quinlan et al. 2010) is used to calculate the coverage across the transcripts for each sample.

4.2.1.3 Differential expression analysis

After computing abundance levels across samples in different units, these numbers can be directly used to detect differentially expressed genes/transcripts across samples or conditions. There are several freely available methods that can perform differential analysis without using a reference genome, such as DESeq and DESeq2. These software packages require R for statistical modeling to identify differentially expressed transcripts based on the abundances generated, for which raw counts should be used to satisfy the model they employ. The expression measures generated by our pipeline are formatted for direct input to these or other packages to estimate differential gene expression. To detect the differentially expressed transcripts, DESeq2, or DESeq if the newer version is unavailable, is used to generate publication-ready visualizations of the results (*MA* plot and Volcano plot).

4.2.2 Implementation

We used a combination of Bash, Python and R to develop our pipeline. Bash handles the decision logic to call different software packages and scripts throughout the pipeline's flow. Calculations and statistical analyses are done in Python or R.

4.2.3 Functional analysis of shared reference transcripts

To render biological conclusions from the analysis of RNA-seq data, it is crucial to functionally annotate the transcripts with putative gene identifications. The common approach for annotation of *de novo* assemblies is to search for sequence homology with already well annotated and biologically related data sets. To annotate the assembled transcripts, one can do BLASTx or BLASTn alignments against related and well-annotated datasets. Examples for how

to run such BLASTx or BLASTn homology search are included in the Supplementary Material of Birol et al. (2015a). BLASTx alignment of an RNA-seq assembly to a large database such as the NCBI NR set can be computationally taxing. To accelerate annotation of this and future *Rana catesbeiana* assemblies, nine *Rana catesbeiana* Trans-ABySS assemblies were combined, processed to reduce redundant sequences, and annotated to create a single reference transcriptome. Construction of this Bullfrog Annotation Resource for the Transcriptome (BART) and its use for annotation of the shared reference transcriptome is described in section 4.2.4. After assigning putative gene names to the transcripts, the results can be used for further downstream analysis such as searching for gene ontology and/or pathway analysis using free and commercial tools such as DAVID (Huang et al. 2009) and IPA (Qiagen), respectively.

The shortage of high quality annotations for many species and genera, especially in the case of non-model organisms, limits the effectiveness of homology-based approach. Therefore, in addition to this approach, it is common to perform *ab initio* protein coding predictions on the assembled transcripts. There are several freely available tools for this purpose such as TransDecoder (Haas et al. 2013).

4.2.4 BART construction

A set of annotated *Rana catesbeiana* contig sequences with minimal sequence redundancy was constructed to accelerate annotation of subsequent *Rana catesbeiana* RNA-seq assemblies. Data from two experiments were utilized to create this resource. For the first, premetamorphic *Rana catesbeiana* were exposed to dilute NaOH vehicle control or 10 nM T_3 at 24°C, sacrificed 48 h later, the entire brain collected, and RNA purified as described in Veldhoen et al. (2014) and chapter 2. For the second, premetamorphic *Rana catesbeiana* were acclimated to 5°C over 72 h, then were exposed to dilute NaOH vehicle control or 10 nM T_3. Seven days later, a set of control

or T_3-treated animals was transferred to 24°C, and then 24 h later all animals were sacrificed and tail fin, back skin, brain, and lung tissues collected and RNA purified as described in Veldhoen et al. (2014) and chapter 2. See Appendix 7 for additional details. Seventy-five bp paired-end single-stranded RNA-seq reads were generated from a total of 18 tissue samples using the Illumina 2500 HiSeq (San Diego, CA, USA) platform (see Appendix 7). Each read library was assembled independently using Trans-ABySS 1.5.0 with k=42. Rather than simple concatenation of the libraries, which would preserve identical or similar transcripts reconstructed in separate libraries, the incidence of redundant sequences was reduced according to a mixed serial-parallel scheme presented in a general form in Appendix 8. BLASTn was used to identify contigs that were at least 95% identical to another contig over 60% or more of each of their lengths. The longest member of each group of such similar contigs was retained to the next stage of the iterative sequence redundancy reduction. The final set of 581,998 sequences, termed BART, was annotated as described in Section 4.2.3. The *R. catesbeiana* Trans-ABySS back skin reference trasncriptome described in the main text was aligned to BART using BLASTn, and the annotation of the top-aligning BART contig with an e-value less than $1x10^{-5}$ was then associated with each back skin contig.

4.2.5 *Rana catesbeiana* samples and data

To demonstrate the effectiveness of our pipeline at processing replicate transcriptomes of an organism with a paucity of genomic data, we sequenced transcriptomes from the back skin of three individual *Rana catesbeiana* tadpoles that were injected with 10 pmol/g body weight of T_3 (Sigma-Aldrich Canada Ltd.) prepared in dilute NaOH (ACP Chemicals Inc.) and sacrificed 48 h post-injection. A matched group of vehicle only-injected tadpoles consisted of an additional group of 3 individual animals. Details of the exposures and evidence of tissue responsiveness to

T_3 treatment using qPCR of these animals can be found in Maher *et al.* (submitted) and in Appendix 5. These samples were also used by Maher *et al.*, but within the context of a separate study with distinct analyses.

Single-stranded RNA-seq libraries were generated from these six samples individually using Illumina HiSeq 2500 paired-end sequencing platform (San Diego, CA, USA) and 100 base pair (bp) paired end sequencing protocol following manufacturer's instructions. Information on the three control (C) and three treated (T_3) read libraries is shown in Table 4.1.

Table 4.1 Collected *Rana catesbeiana* sequence data. Two Illumina HiSeq 2500 lanes were sequenced per sample.

Condition	# RNA-seq read pairs (10^6)	Average fragment length (bp)
C	67.42	394.61
C	78.06	374.03
C	88.91	380.01
T_3	79.02	388.79
T_3	70.30	396.19
T_3	80.52	382.83

The RNA-seq observations were independently supported using qPCR primers designed using Primer Premier 5 (Premier Biosoft), ordered from Integrated DNA technologies, and validated through the quality control procedure described previously (Veldhoen *et al.,* 2014). The qPCR assays were applied to tissue obtained from 8-9 animals per hormone condition from the experiment reported in Maher *et al.* (submitted) using reaction conditions described in Veldhoen *et al.* (2014). See Appendix 6 for primer sequences and additional information.

4.3 Results

Rana catesbeiana is a cosmopolitan representative of the *Ranidae* (or "true frogs"). This family of frogs can be found world-wide as native or introduced species, but despite their importance as environmental sentinels and ecosystem service providers, there are limited genomic resources

available for these species (Kiemnec-Tyburczy et al. 2012; Savage et al. 2014). Currently, no well-annotated genomic reference for *Rana catesbeiana* or any other true frog exists. Therefore, *de novo* approaches must be used to study the transcriptome of this organism. *Rana catesbeiana* genome size is estimated around 5 Gbp, which is among the larger Anuran genomes (Helbing 2012).

THs are responsible for the postembryonic transformation of a tadpole into a juvenile frog and this metamorphosis involves substantial changes to tissue transcriptomes (Atkinson et al. 1996; Birol et al. 2015; Das et al. 2002; Denver et al. 1997; Mukhi et al. 2010). Young, premetamorphic tadpoles are functionally athyroid yet possess a thyroid gland. This gland begins producing THs during prometamorphosis reaching maximal levels at metamorphic climax; the phase characterized by rapid organismal remodeling from a tadpole to frog morphology (Gilbert et al. 1996).

Tadpoles can be precociously induced to undergo metamorphosis by administration of low levels of THs. One tissue that undergoes substantial remodeling essential for the organism's survival as it emerges from an aquatic to a terrestrial environment is the back skin. Tadpole skin consists of three cell types: apical, skein, and basal cells (Robinson et al. 1987). The apical and skein cells, which make up the outermost cell layers, undergo apoptosis during metamorphosis, while the basal cells generate the cells of the adult skin (Fox 1974; Suzuki et al. 2002). This repopulation is in contrast to the breakdown and resorption of the tail fin and its skin, which is of similar structure and organization to the body skin (Ishizuya-Oka et al. 2010; Yoshizato 2007). Using our pipeline we define the changes in the back skin transcriptome in response to TH-treatment and identify RNA processing and cell cycle control components as key processes affected by this hormone.

4.3.1 Comparison of assembly strategies

Table 4.2 shows the summary of the assembly statics on the final reference transcriptomes generated with the different tools and methods compared. Neither of the tools was able to complete assembly of the pooled reads with minimum k-mer coverage of 1 or 2 due to memory exhaustion, but both were successful when the minimum was set to 3. The minimum k-mer coverage that produced successful assemblies for the tools following the separate assembly scheme was 2.

Table 4.2 Summary of assembly statistics for different assembly methods for generating a reference transcriptome from the 6 RNA-seq libraries. The reference transcriptomes were created by pooling the read libraries before assembly or assembling them separately followed by concatenation of the assembled sequence files and processing by our pipeline to produce the final reference transcriptomes.

Method	# contigs	N50 (bp)	N20 (bp)	Min length (bp)	Max length (bp)	Total assembled sequences (Mbp)
Trans-ABySS (pooled)	833,804	1,042	2,462	200	23,063	600.9
Trans-ABySS (separate)	1,836,602	1,172	3,130	200	22,989	1,310
Trans-ABySS (separate) + pipeline	720,446	1,393	3,634	200	22,989	562.6

The assembled contigs produced by combining Trans-ABySS with our pipeline for sequence redundancy reduction are more contiguous than those produced from the combined reads by these tools (Table 4.2). Additionally, the N50 and N20 are several hundreds of base-pairs larger than the ones generated with the same tools from the combined dataset (Table 4.2).

The cost of these assembly improvements is an increased runtime (Table 4.3). The greater amount of memory required by Trans-ABySS to assemble the combined six read libraries reflects the burden of handling the high number of k-mers during assembly. Trans-ABySS

displayed a modest difference in total runtime between the combined and separate assembly schemes (Table 4.3); the wall time to assemble each individual library was 5 h. The rest of the run time noted in Table 4.3 includes total run-time for the separate assemblies of the 6 libraries.

We assessed the completeness of the final assemblies using the CEGMA tool (Parra, Bradnam, & Korf, 2007). Trans-ABySS fully assembled 246 CEGs in both the combined and separate assembly cases (Table 4.3). Moreover, ORF analysis using TransDecoder (Haas *et al*., 2013) showed that our approach improved the share of assembled transcripts with complete predicted ORFs greater than 300 nucleotides in length (Table 4.3).

Overall, our pipeline generated more complete and contiguous assemblies from the same RNA-seq data compared to the approach of combining the read libraries by adding a small overhead to the run time.

Table 4.3 Summary of performance and assembly completeness measures.

Method	Max memory (G)	Wall clock run time	Reconstructed full length CEGs (out of 248)	Contigs with complete ORF
Trans-ABySS (pooled)	98.8	31:33	246	6.09%
Trans-ABySS (separate)	58.8	36:50*	246	8.89%

* Includes the entire run for generating individual assemblies and generating shared reference transcriptome

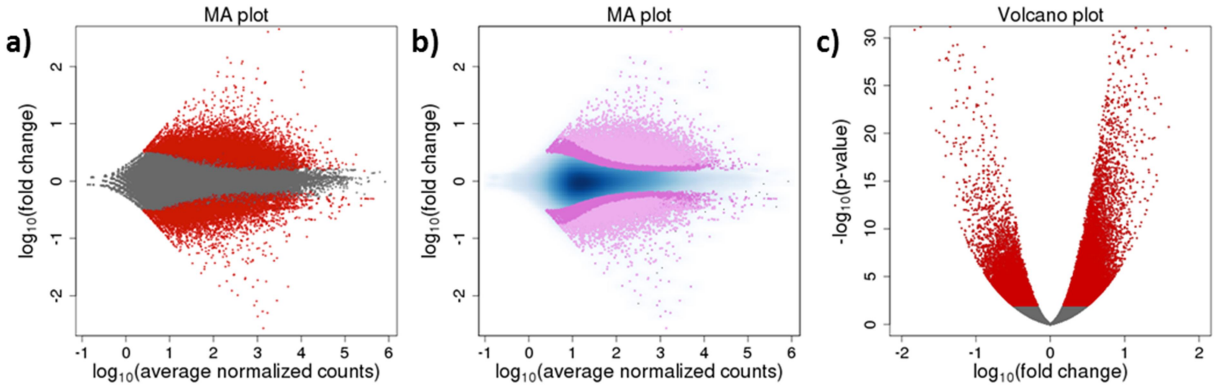

Figure 4.2 Differential expression result visualizations the *Rana catesbeiana* results. (a) MA plot comparing log-transformed average fold change to the log-transformed average normalized count. Red dots = transcript with $p < 0.01$; gray dots = transcript with $p \geq 0.01$. (b) Alternate depiction of (a) demonstrating the effect of alternate p-value thresholds, as well as transcript density. Blue represents density, where darker shades indicate higher density. Two shades of purple designate p-value thresholds of 0.0001 or 0.01, where lighter colour indicates the lower threshold. (c) Volcano plot describing relationship between log-transformed p-value and log-transformed average fold change. Colours are as in panel (a).

4.3.2 *Rana catesbeiana* transcriptome results

The original 6 RNA-seq datasets were aligned back to the generated reference transcriptome separately, and the alignments parsed to calculate total raw counts as described in section 4.2.1.2. The wall-clock run time of this stage with our datasets was 6h 21m, including 5h 53m for alignments and the rest for calculations and parsing the files. Next, a lower limit on contig length of 500 bases was applied, and DESeq2 and the median-corrected read counts generated in the previous step were used for differential expression analysis using. Of the 282,238 contigs 500 bases or longer, this approach identified 18,224 differentially expressed transcripts using a p-value threshold of 0.01. Figure 4.2 shows three visualizations of the differential expression results using p-value thresholds of 0.01 or 0.0001.

The contigs were annotated via BLASTn alignment to the BART contig set, where the annotation of the best-aligned BART contig was transferred to the query contig. This approach successfully annotated 97% of the contigs; the remaining 8543 were annotated independently. The BART annotations, and those of the contigs that did not align to BART, were derived from BLASTx alignment to the NCBI NR database (retrieved February 10, 2015) and the Ensembl human protein database (release 78, retrieved February 11, 2015), and by BLASTn alignment to a set of amphibian mRNA sequences downloaded from NCBI GenBank (retrieved February 11, 2015). Alignments with an e-value less than $1x10^{-5}$ and whose length represented at least 10% of the database sequence were retained, and the top hit from each of the three series of alignments was selected. The annotation results are summarized in Table 4.4.

Table 4.4 Summary of transcript annotation results.

Annotation method	# annotated contigs
BLASTn (amphibian)	203,785
BLASTx (NR)	131,200
BLASTx (Ensembl *Homo sapiens*)	101,830
Total annotated combining all 3 methods	223,353

Ensembl annotations of the differentially expressed transcripts selected by DESeq2 with p-value threshold of 0.01 yielded 4,145 unique Ensembl gene IDs. To characterize the reaction of the animals to a challenge by TH, we performed both gene ontology enrichment analysis and pathway analysis based on these Ensembl IDs. We used DAVID to perform enrichment analysis of GO terms on this list from perspective of Biological Processes (Huang et al. 2009). Then, we applied REVIGO (Supek et al. 2011) on the list of GO terms produced by DAVID to cluster and

summarize the results. Figure 4.3 summarizes this final clustering generated by REVIGO 'tabletop' feature. Note that individual Ensembl IDs may be related to multiple GO terms.

The GO analysis indicated that genes associated with RNA processing and the cell cycle were most affected by TH treatment (Figure 4.4). See Figure 4.4 and Appendix 6 for independent qPCR confirmation of increased expression of select RNA processing-associated genes in response to T_3 treatment. Activation of new gene expression programs and shutdown of obsolete ones is an integral part of the transition from larval to adult skin. The wave of new transcripts that results from TH treatment must be spliced, capped with 7-methylguanosine, and polyadenylated. As the level of circulating TH increases in the tadpole, expression of key cell cycle control genes changes to regulate the proliferation of skin cells, including cyclin C and cyclin B (Buchholz et al. 2007; Skirrow et al. 2008; Suzuki et al. 2009).

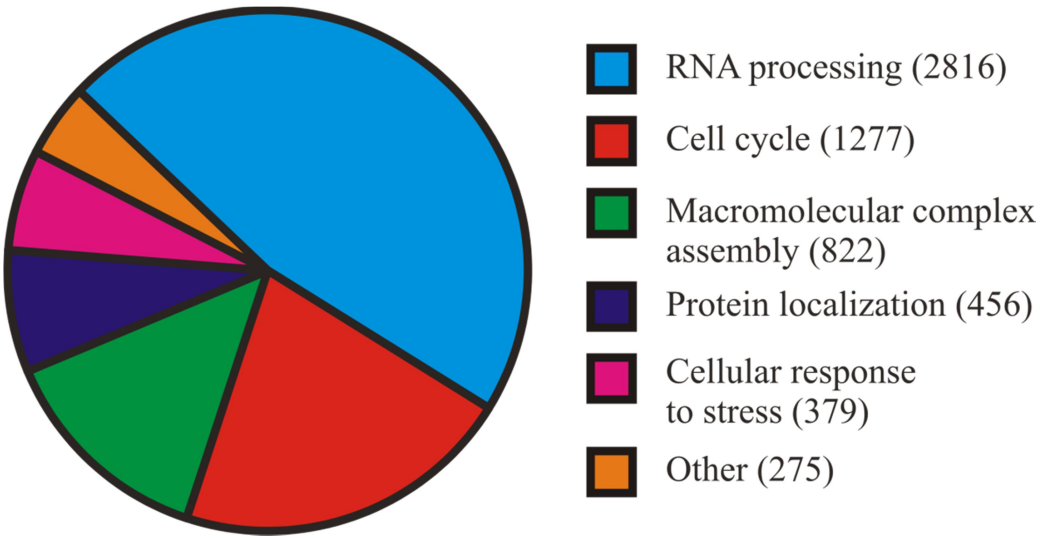

Figure 4.3 Biological process GO classification of DESeq2-selected TH-responsive *Rana catesbeiana* back skin transcripts annotated by Ensembl IDs. The numbers in parentheses are the number of individual IDs found in each GO class.

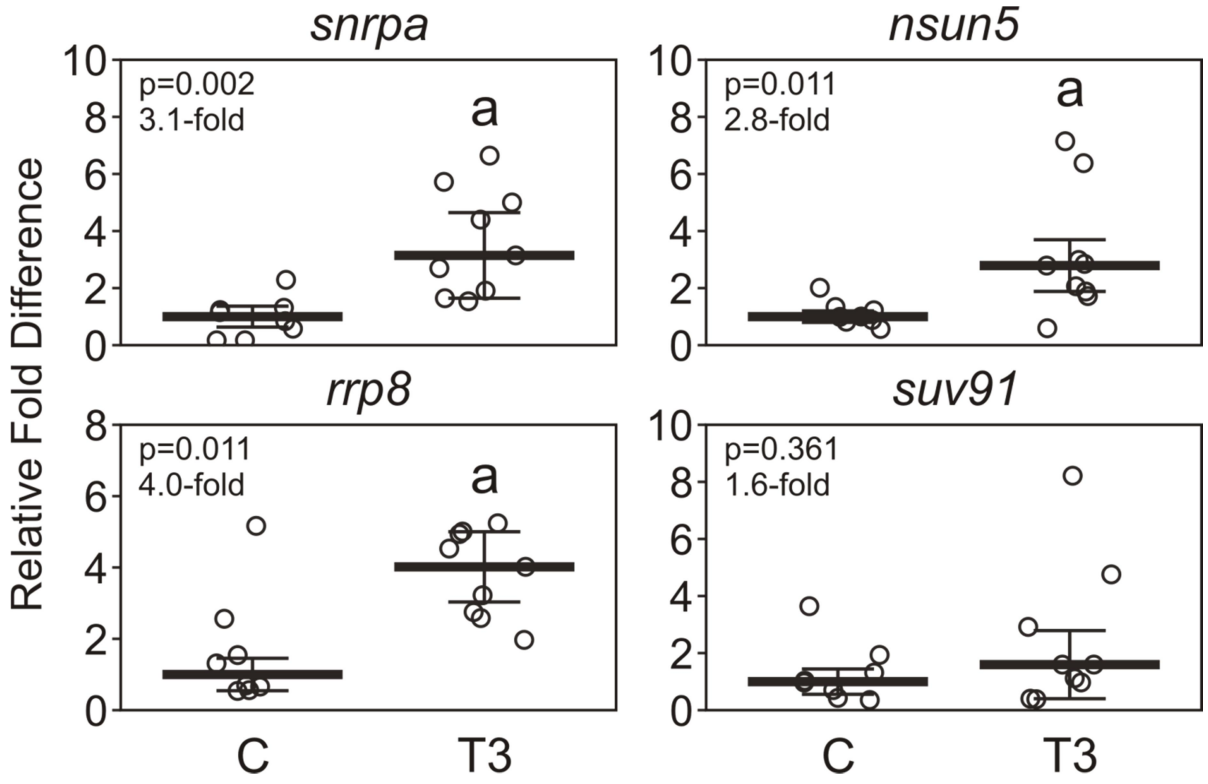

Figure 4.4 qPCR analysis of select RNA processing-related genes in *Rana catesbeiana* tadpoles treated with vehicle control (C) or 10 nM T_3. Individual animals are depicted by the open circles with the median (horizontal black bar) and median absolute deviation (whiskers) shown.

4.4 Conclusion

Separate assembly of the read libraries and subsequent processing through our pipeline improved assembly completeness and enabled consistent comparison of the six *R. catesbeiana* back skin samples, which revealed RNA processing and cell cycle-related genes as the categories most affected by TH treatment in this tissue. Streamlined approaches that maximize assembly quality while reducing the demand on computational resources represent opportunities for transcriptomic studies of organisms without annotated genomes to be undertaken by groups with limited familiarity with high performance computing.

5 Synthesis

Understanding the regulation of the TH-induced gene expression network is critical to comprehension of amphibian metamorphosis, which is an excellent model system for study of TH action in vertebrates. Past studies have uncoupled execution of the metamorphic program from its initiation by chemical interference with RNA transcription or protein translation or phosphorylation (Buckbinder et al. 1992; Ji et al. 2007; Kanamori et al. 1993; Skirrow et al. 2008), although such treatment can negatively impact animal health during the experiment (Wang et al. 1993). Investigations utilizing post embryonic development of *Rana catesbeiana* and temperature control allows for a nonchemical means to achieve uncoupling of the phasic program of metamorphosis (Atkinson et al. 1996; Hammond et al. 2015; Mochizuki et al. 2012). This method is unsuitable for the two *Xenopus spp.* as their native tropical habitats did not demand development of mechanisms to handle these low temperatures. The work in this thesis, combined with the previous chemical-reliant studies on other frogs, on the program of metamorphosis in *Rana catesbeiana* allow an overall view of the flow of molecular information

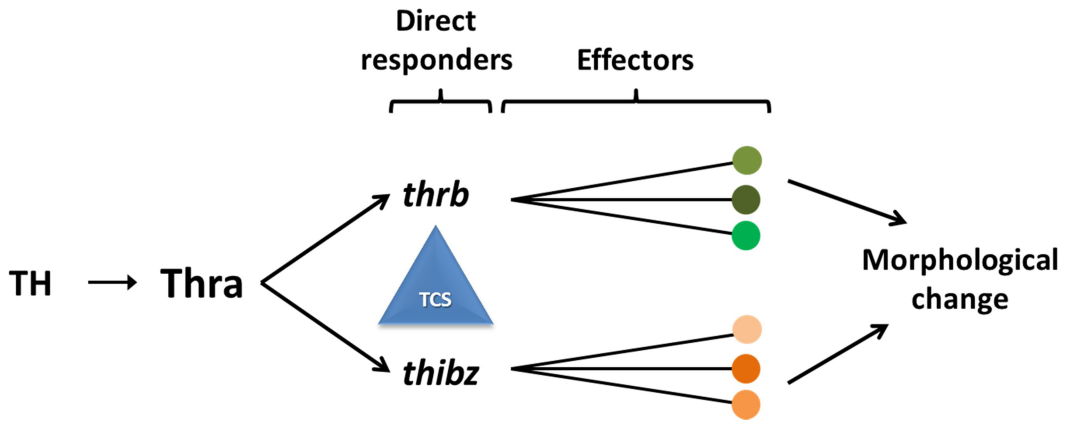

Figure 5.1 Overview of flow of molecular information from TH-mediated signaling and modification of the transcriptome through to execution of metamorphosis. Blue triangle signifies disruption of *thrb* expression by TCS described in Chapter 2.

from TH-mediated signaling and modification of the transcriptome through to execution of metamorphosis (Figure 5.1). This thesis used low environmental temperature to uncouple the initiation phase of the TH-regulated metamorphic gene expression program from its execution to evaluate the contribution of *thra, thrb, klf9, cepb1,* and *thibz* to the particular program of five *Rana catesbeiana* tissues. It then described creation of new resources for *Rana catesbeiana* that will transform transcriptomic analyses in this important model species.

5.1 Possible role of *thibz* in TH molecular memory suggests important contribution to metamorphic program

Each of the five directly-responsive transcription factors showed evidence that they may be involved in forming the low temperature TH memory in at least one of the *Rana catesbeiana* tissues examined, but *thibz* was the sole gene that may be substantially involved in every tissue. While the *thibz* promoter is well-characterized in *Xenopus laevis*, its features have not yet been described in *Rana catesbeiana*, and it is unknown which genes it regulates in either species. It has been implicated as participating in both apoptosis of larval-specific cell types and proliferation of progenitors to form adult tissues, so it may be involved in mediating the TH response upstream of cell fate-specific genes (Ishizuya-Oka et al.). Investigation into the actions of Thibz protein using targeted chromatin immunoprecipitation followed by massively parallel sequencing of genomic regions isolated from tadpole tissues under low temperature TH-treatment will help further elucidate regulatory events and restrictions in the genetic program of metamorphosis. Additionally, determination of the expression status of *thibz* and these other directly-responsive transcription factors under this TH exposure scenario would provide further insight into the TH regulatory regime.

5.2 Improved *de novo* sequence assembly approaches and new *Rana catesbeiana* transcriptomic resources

The results from chapter 2 provided valuable but limited insight into the gene expression changes involved in forming the molecular memory of TH exposure at low temperature. Before a whole-transcriptome view of the gene expression changes within this exposure scenario can be obtained, the suitability of RNA-seq techniques that do not rely on a sequenced and annotated reference genome needed to be evaluated.

Using Trans-ABySS, we were able to reconstruct more than 350 000 *Rana catesbeiana* hepatic transcripts to produce a snapshot of gene expression before and after T_3 exposure. My multi-level annotation scheme was able to putatively identify nearly 80% of these sequences. Prior to this work, there were less than 35 000 *Rana catesbeiana* sequences previously available, representing approximately 15 000 unique transcripts. This > 10-fold increase in *Rana catesbeiana* transcriptome information is truly transformative, and has largely done away with the previously necessary gene-by gene approach.

The veracity of this assembly approach was supported by several lines of evidence: the excellent agreement between the *Rana catesbeiana* and *Xenopus laevis* gene ontologies, the latter of which closely agreed with previous reports from other *Xenopus laevis* tissues (Das et al. 2006) and *Xenopus tropicalis* tadpoles (Tan et al. 2013); concordance of both datasets with the findings of Blower *et al.* (2013) and Tan *et al.* (2013) with respect to the proportion of assembled sequences containing putative ORFs; and observation of the expected response to TH of established TH-sensitive genes. In addition to yielding this trove of new *Rana catesbeiana* information, this assembly and annotation approach could be applied to other organisms with insufficient genomic data, so the work reported in chapter 3 has far-reaching consequences for transcriptomics in general.

Although of high-quality, the data yielded by the *de novo* assembly approach described in chapter 3 was limited by its generation from only one sample per treatment condition. Chapter 4 described Trans-ABySS assembly of replicate control and T_3-treated *Rana catesbeiana* back skin samples and processing of the independent assemblies through the TRENCH pipeline to mitigate redundant sequences. As RNA-seq costs decrease, replicate samples for organisms without extensive genomic resources will become more prevalent, and the combined Trans-ABySS – TRENCH pipeline will aid future transcriptomic studies in *Rana catesbeiana* and innumerable other species.

A product of the work in chapter 4 that benefits *Rana catesbeiana* research in particular is BART, a transcriptomic resource with several applications. It is a set of representative sequences from several *Rana catesbeiana* tissues and TH exposure conditions that can be mined for tissue-specific transcript information that may not be found in the liver assembly described in chapter 3. Another advantage of this set over the liver assembly is that all of the BART contigs were reconstructed from strand-specific reads, which is necessary to confidently identify antisense transcripts. BART accelerates annotation of newly-assembled *Rana catesbeiana* sequences because alignment of these new sequences to BART is performed solely in nucleotide space, rather than the six-frame translations required for nucleotide to protein alignments using BLASTx. Preliminary experiments using BART as a reference transcriptome to which RNA-seq reads are directly aligned to obtain expression data of each representative BART sequence have been performed and have yielded promising results. In the future, read libraries generated from animals treated with vehicle or T_3 at nonpermissive temperature could be applied to BART in this fashion to identify other genes that may be involved in the TH molecular memory formed at low temperature.

A limitation of BART is that it is not a true reference transcriptome. The method used in its construction was not intended to retain all unique sequences, and likely excluded many splice variants while including incompletely assembled sequences. One of the sources of this property is the alignment overlap criteria, namely that the alignment length had to represent at least 60% of the length of both contigs. Therefore a contig that is sufficiently smaller than another to which it aligns will not be removed. The length criterion was applied to avoid removing transcripts that were only highly similar over a small region, which could represent a conserved sequence motif. This approach has the advantage of catching both contigs that are eclipsed by a larger related contig, and those where each overlaps the other on one end. Future refinement of the BART set could focus on identifying and removing contigs that are highly similar but where one is much smaller than the other, as in the case of small incompletely-assembled contigs. However, the effort required to protect this process from inappropriately favouring overlarge artefactual sequences may make that endeavour untenable with the impending availability of the bullfrog genome.

5.3 Conclusion

The tissue-specific expression profiles obtained for the directly TH-responsive transcription factors accentuates the diversity of TH signaling that is modulated by environmental temperature, and reinforces the potential significance of *thibz* in the molecular memory of TH. Beyond its potential as a nonchemical means to study early events of TH signaling, disruption of this molecular memory by TCS and IBF demonstrates the need to understand the mechanism of its formation. Further investigations into the genes involved in this phenomenon have been empowered by the sequenced and annotated *R. catesbeiana* transcriptomes and the techniques

and software used to produce them, which can be applied to the study of other resource-poor wildlife species to elevate that research as well.

Bibliography

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J., 1990. Basic local alignment search tool. J Mol Biol 215(3): 403-410.

AmphibiaWeb. (2012). "AmphibiaWeb: Information on amphibian biology and conservation." Retrieved January 12, 2012, from www.amphibiaweb.org.

Anders, S. and Huber, W., 2010. Differential expression analysis for sequence count data. Genome Biol 11(10): R106.

Anders, S., Pyl, P. T. and Huber, W., 2015. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 31(2): 166-169.

Aranda, A. and Pascual, A., 2001. Nuclear hormone receptors and gene expression. Physiol Rev 81(3): 1269-1304.

Ashley, H., Katti, P. and Frieden, E., 1968. Urea excretion in the bullrog tadpole: effect of temperature, metamorphosis, and thyroid hormones. Dev Biol 17(3): 293-307.

Astapova, I., Lee, L. J., Morales, C., Tauber, S., Bilban, M. and Hollenberg, A. N., 2008. The nuclear corepressor, NCoR, regulates thyroid hormone action *in vivo*. Proc Natl Acad Sci U.S.A. 105(49): 19544-19549.

Atkinson, B. G., Helbing, C. and Chen, Y. 1996. Reprogramming of gene expression during amphibian metamorphosis. Metamorphosis: postembryonic reprogramming of gene expression in amphibian and insect cells. L. I. Gilbert, J. Tata and B. G. Atkinson. San Diego, Academic Press**:** 539-566.

Atlas, M., 1935. The effect of temperature on the development of *Rana pipiens*. Physiol Zool 8(3): 290-310.

Bassett, J., Harvey, C. B. and Williams, G. R., 2003. Mechanisms of thyroid hormone receptor-specific nuclear and extra nuclear actions. Mol Cell Endocrinol 213: 1-11.

Bazinet, A. L., Cummings, M. P., Mitter, K. T. and Mitter, C. W., 2013. Can RNA-Seq Resolve the Rapid Radiation of Advanced Moths and Butterflies (Hexapoda: Lepidoptera: Apoditrysia)? An Exploratory Study. PLoS One 8(12).

Benjamini, Y. and Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological): 289-300.

Bianco, A. C., Salvatore, D., Gereben, B., Berry, M. J. and Larsen, P. R., 2002. Biochemistry, cellular and molecular biology, and physiological roles of the iodothyronine selenodeiodinases. Endocr Rev 23(1): 38-89.

Birol, I., Behsaz, B., Hammond, S. A., Kucuk, E., Veldhoen, N. and Helbing, C. C., 2015a. *De novo* transcriptome assemblies of *Rana* (*Lithobates*) *catesbeiana* and *Xenopus laevis* tadpole livers for comparative genomics without reference genomes. PLoS One 10(6): e0130720.

Birol, I., Jackman, S. D., Nielsen, C. B., Qian, J. Q., Varhol, R., Stazyk, G., Morin, R. D., Zhao, Y., Hirst, M., Schein, J. E., Horsman, D. E., Connors, J. M., Gascoyne, R. D., Marra, M. A. and Jones, S. J., 2009. *De novo* transcriptome assembly with ABySS. Bioinformatics 25(21): 2872-2877.

Birol, I., Raymond, A., Chiu, R., Nip, K. M., Jackman, S. D., Kreitzman, M., Docking, T. R., Ennis, C. A., Robertson, A. G. and Karsan, A., 2015b. Kleat: cleavage site analysis of transcriptomes. Pac Symp Biocomput 20: 347-358.

Bizhanova, A. and Kopp, P., 2009. Minireview: The sodium-iodide symporter NIS and pendrin in iodide homeostasis of the thyroid. Endocrinology 150(3): 1084-1090.

Björnsson, B. T., Stefansson, S. O. and McCormick, S. D., 2011. Environmental endocrinology of salmon smoltification. Gen Comp Endocrinol 170(2): 290-298.

Blower, M. D., Jambhekar, A., Schwarz, D. S. and Toombs, J. A., 2013. Combining different mRNA capture methods to analyze the transcriptome: analysis of the Xenopus laevis transcriptome. PLoS One 8(10): e77700.

Bodenner, D. L., Mroczynski, M. A., Weintraub, B. D., Radovick, S. and Wondisford, F. E., 1991. A detailed functional and structural analysis of a major thyroid hormone inhibitory element in the human thyrotropin β-subunit gene. J Biol Chem 266(32): 21666-21673.

Brix, K., Führer, D. and Biebermann, H., 2011. Molecules important for thyroid hormone synthesis and action - known facts and future perspectives. Thyroid Res 4 Suppl 1: S9.

Brown, D. D., Wang, Z., Furlow, J. D., Kanamori, A., Schwartzman, R. A., Remo, B. F. and Pinder, A., 1996. The thyroid hormone-induced tail resorption program during *Xenopus laevis* metamorphosis. Proc Natl Acad Sci U.S.A. 93: 1924-1929.

Brown, G. W., Jr., Brown, W. R. and Cohen, P. P., 1959. Comparative biochemistry of urea synthesis. II. Levels of urea cycle enzymes in metamorphosing *Rana catesbeiana* tadpoles. J Biol Chem 234(7): 1775-1780.

Buchholz, D. R., Heimeier, R. A., Das, B., Washington, T. and Shi, Y. B., 2007. Pairing morphology with gene expression in thyroid hormone-induced intestinal remodeling and identification of a core set of TH-induced genes across tadpole tissues. Dev Biol 303(2): 576-590.

Buckbinder, L. and Brown, D. D., 1992. Thyroid hormone-induced gene expression changes in the developing frog limb. J Biol Chem 267: 25786-25791.

Buisine, N., Ruan, X., Bilesimo, P., Grimaldi, A., Alfama, G., Ariyaratne, P., Mulawadi, F., Chen, J., Sung, W. K., Liu, E. T., Demeneix, B. A., Ruan, Y. and Sachs, L. M., 2015. *Xenopus tropicalis* genome re-scaffolding and re-annotation reach the resolution required for *in vivo* ChIA-PET analysis. PLoS One 10(9): e0137526.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. and Madden, T. L., 2009. BLAST+: architecture and applications. BMC Bioinformatics 10: 421.

Cao, X. Y., Jiang, X. M., Dou, Z. H., Rakeman, M. A., Zhang, M. L., O'Donnell, K., Ma, T., Amette, K., DeLong, N. and DeLong, G. R., 1994. Timing of vulnerability of the brain to iodine deficiency in endemic cretinism. N Engl J Med 331(26): 1739-1744.

Carr, F. E. and Wong, N. C., 1994. Characteristics of a negative thyroid hormone response element. J Biol Chem 269(6): 4175-4179.

Cecil, S. G. and Just, J. J., 1979. Survival rate, population density, and development of a naturally occurring anuran larvae (*Rana catesbeiana*). Copeia(3): 447-453.

Chassande, O., Fraichard, A., Gauthier, K., Flamant, F., Legrand, C., Savatier, P., Laudet, V. and Samarut, J., 1997. Identification of transcripts initiated from an internal promoter in the c-erbA alpha locus that encode inhibitors of retinoic acid receptor-alpha and triiodothyronine receptor activities. Mol Endocrinol 11(9): 1278-1290.

Chen, Y. and Atkinson, B. G., 1997. Role for the *Rana catesbeiana* homologue of C/EBP alpha in the reprogramming of gene expression in the liver of metamorphosing tadpoles. Dev Genet 20: 152-162.

Cheng, K., Escalon, B. L., Robert, J., Chinchar, V. G. and Garcia-Reyero, N., 2014. Differential transcription of fathead minnow immune-related genes following infection with frog virus 3, an emerging pathogen of ectothermic vertebrates. Virology 456-457: 77-86.

Chiamolera, M. I., Sidhaye, A. R., Matsumoto, S., He, Q., Hashimoto, K., Ortiga-Carvalho, T. M. and Wondisford, F. E., 2012. Fundamentally distinct roles of thyroid hormone receptor isoforms in a thyrotroph cell line are due to differential DNA binding. Mol Endocrinol 26(6): 926-939.

Chu, J., Sadeghi, S., Raymond, A., Jackman, S. D., Nip, K. M., Mar, R., Mohamadi, H., Butterfield, Y. S., Robertson, A. G. and Birol, I., 2014. BioBloom tools: fast, accurate and memory-efficient host species sequence screening using bloom filters. Bioinformatics. 30(23): 3402-3404.

Cordeiro, A., Souza, L. L., Einicker-Lamas, M. and Pazos-Moura, C. C., 2013. Non-classic thyroid hormone signalling involved in hepatic lipid metabolism. J Endocrinol 216(3): R47-57.

Dann, A. B. and Hontela, A., 2011. Triclosan: environmental exposure, toxicity and mechanisms of action. J Appl Toxicol 31(4): 285-311.

Das, B., Cai, L., Carter, M. G., Piao, Y. L., Sharov, A. A., Ko, M. S. and Brown, D. D., 2006. Gene expression changes at metamorphosis induced by thyroid hormone in Xenopus laevis tadpoles. Dev Biol 291(2): 342-355.

Das, B., Schreiber, A. M., Huang, H. and Brown, D. D., 2002. Multiple thyroid hormone-induced muscle growth and death programs during metamorphosis in Xenopus laevis. Proc Natl Acad Sci U.S.A. 99(19): 12230-12235.

Das, B., Schreiber, A. M., Huang, H. and Brown, D. D., 2002. Multiple thyroid hormone-induced muscle growth and death programs during metamorphosis in *Xenopus laevis*. Proc Natl Acad Sci U.S.A. 99: 12230-12235.

Daszak, P., Berger, L., Cunningham, A. A., Hyatt, A. D., Green, D. E. and Speare, R., 1999. Emerging infectious diseases and amphibian population declines. Emerg Infect Dis 5(6): 735-748.

Davidson, N. M. and Oshlack, A., 2014. Corset: enabling differential gene expression analysis for *de novo* assembled transcriptomes. Genome Biol 15(7): 410.

Davis, P. J. and Davis, F. B., 1996. Nongenomic actions of thyroid hormone. Thyroid 6(5): 497-504.

Davis, P. J., Leonard, J. L. and Davis, F. B., 2008. Mechanisms of nongenomic actions of thyroid hormone. Front Neuroendocrinol 29(2): 211-218.

Dent, J. N. 1968. Survey of amphibian metamorphosis. Metamorphosis. W. Etkin and L. I. Gilbert. New York, Appleton-Century-Crafts**:** 271-311.

Denver, R. J., 2013. Neuroendocrinology of amphibian metamorphosis. Curr Top Dev Biol 103: 195-227.

Denver, R. J., Pavgi, S. and Shi, Y. B., 1997. Thyroid hormone-dependent gene expression program for Xenopus neural development. J Biol Chem 272(13): 8179-8188.

Dodd, M. and Dodd, J. 1976. The biology of metamorphosis. Physiology of the Amphibia. B. Lofts. New York, Academic Press**:** 467-599.

Domanski, D. and Helbing, C. C., 2007. Analysis of the *Rana catesbeiana* tadpole tail fin proteome and phosphoproteome during T3-induced apoptosis: identification of a novel type I keratin. BMC Dev Biol 7: 94.

Duarte-Guterman, P., Langlois, V. S., Pauli, B. D. and Trudeau, V. L., 2010. Expression and T3 regulation of thyroid hormone- and sex steroid-related genes during *Silurana* (*Xenopus*) *tropicalis* early development. Gen Comp Endocrinol 166(2): 428-435.

Duarte-Guterman, P. and Trudeau, V. L., 2010. Regulation of Thyroid Hormone-, Oestrogen- and Androgen-Related Genes by Triiodothyronine in the Brain of *Silurana tropicalis*. J Neuroendocrinol 22(9): 1023-1031.

Duarte-Guterman, P. and Trudeau, V. L., 2011. Transcript profiles and triiodothyronine regulation of sex steroid- and thyroid hormone-related genes in the gonad-mesonephros complex of *Silurana tropicalis*. Mol Cell Endocrinol 331(1): 143-149.

Eggert, C., 2004. Sex determination: the amphibian models. Reprod Nutr Dev 44(6): 539-549.

Fisher, M. C., Henk, D. A., Briggs, C. J., Brownstein, J. S., Madoff, L. C., McCraw, S. L. and Gurr, S. J., 2012. Emerging fungal threats to animal, plant and ecosystem health. Nature 484(7393): 186-194.

Flicek, P., Aken, B. L., Beal, K., Ballester, B., Caccamo, M., Chen, Y., Clarke, L., Coates, G., Cunningham, F., Cutts, T., Down, T., Dyer, S. C., Eyre, T., Fitzgerald, S., Fernandez-Banet, J., Graf, S., Haider, S., Hammond, M., Holland, R., Howe, K. L., Howe, K., Johnson, N., Jenkinson, A., Kahari, A., Keefe, D., Kokocinski, F., Kulesha, E., Lawson, D., Longden, I., Megy, K., Meidl, P., Overduin, B., Parker, A., Pritchard, B., Prlic, A., Rice, S., Rios, D., Schuster, M., Sealy, I., Slater, G., Smedley, D., Spudich, G., Trevanion, S., Vilella, A. J., Vogel, J., White, S., Wood, M., Birney, E., Cox, T., Curwen, V., Durbin, R., Fernandez-Suarez, X. M., Herrero, J., Hubbard, T. J., Kasprzyk, A., Proctor, G., Smith, J., Ureta-Vidal, A. and Searle, S., 2008. Ensembl 2008. Nucleic Acids Res 36(Database issue): D707-714.

Fondell, J. D., 2013. The Mediator complex in thyroid hormone receptor action. Biochim Biophys Acta 1830(7): 3867-3875.

Forrest, D. and Visser, T. J. 2013. Thyroid hormone signalling. Biochim Biophys Acta 1830(7): 3859-4008.

Fox, H., 1974. The epidermis and its degeneration in the larval tail and adult body of *Rana temporaria* and *Xenopus laevis* (Amphibia: Anura). J Zool 174(2): 217-235.

Francis, W. R., Christianson, L. M., Kiko, R., Powers, M. L., Shaner, N. C. and SH, D. H., 2013. A comparison across non-model animals suggests an optimal sequencing depth for *de novo* transcriptome assembly. BMC Genomics 14: 167.

Frieden, E. 1968. Biochemistry of amphibian metamorphosis. Metamorphosis: a problem in developmental biology. E. a. G. Etkin, L. I. New York, Appleton-Century-Crofts.

Frieden, E., Wahlborg, A. and Howard, E., 1965. Temperature control of response of tadpoles to triiodothyronine. Nature 205(4977): 1173-1176.

Frischkorn, K., Harke, M., Gobler, C. and Dyhrman, S. T., 2014. De novo assembly of Aureococcus anophagefferens transcriptomes reveals diverse responses to the low nutrient and low light conditions present during blooms. Front Microbiol 5.

Galay-Burgos, M., Power, D. M., Llewellyn, L. and Sweeney, G. E., 2008. Thyroid hormone receptor expression during metamorphosis of Atlantic halibut (*Hippoglossus hippoglossus*). Mol Cell Endocrinol 281(1-2): 56-63.

Galton, V. A. 1983. Thyroid hormone action in amphibian metamorphosis. Molecular Basis of Thyroid Hormone Action. J. H. Oppenheimer and H. H. Samuels. New York, Academic Press.

Galton, V. A., 1986. Thyroxine and 3,5,3'-triiodothyronine bind to the same putative receptor in hepatic nuclei of Rana catesbeiana tadpoles. Endocrinology 118(3): 1114-1118.

Galton, V. A., 1992. The role of thyroid hormone in amphibian metamorphosis. Trends Endocrinol Metab 3(3): 96-100.

Galton, V. A., de Waard, E., Parlow, A. F., St Germain, D. L. and Hernandez, A., 2014. Life without the iodothyronine deiodinases. Endocrinology 155(10): 4081-4087.

Galton, V. A., Schneider, M. J., Clark, A. S. and St Germain, D. L., 2009. Life without thyroxine to 3,5,3'-triiodothyronine conversion: studies in mice devoid of the 5'-deiodinases. Endocrinology 150(6): 2957-2963.

Gastwirth, J. L., Gel, Y. R., Hui, W. L. W., Lyubchich, V., Miao, W. and Noguchi, K., 2013. lawstat: An R package for biostatistics, public policy, and law.

Genome 10K Community of Scientists, 2009. Genome 10K: a proposal to obtain whole-genome sequence for 10,000 vertebrate species. J Hered 100(6): 659-674.

Gilbert, L. I. and Frieden, E. 1981. Metamorphosis: a problem in developmental biology. New York, New York, Plenum Press.

Gilbert, L. I., Tata, J. R. and Atkinson, B. G. 1996. Metamorphosis: postembryonic reprogramming of gene expression in amphibian and insect cells. San Diego, Academic Press.

Góngora-Castillo, E. and Buell, C. R., 2013. Bioinformatics challenges in *de novo* transcriptome assembly using short read sequences in the absence of a reference genome sequence. Nat Prod Rep 30(4): 490-500.

Gosner, K. L., 1960. A simplified table for staging anuran embryos and larvae with notes on identification. Herpetologica 16: 183-190.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N. and Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29(7): 644-652.

Grayfer, L., Andino, F. D., Chen, G. C., Chinchar, G. V. and Robert, J., 2012. Immune evasion strategies of ranaviruses and innate immune responses to these emerging pathogens. Viruses 4(7): 1075-1092.

Grimaldi, A., Buisine, N., Miller, T., Shi, Y. B. and Sachs, L. M., 2013. Mechanisms of thyroid hormone receptor action during development: lessons from amphibian studies. Biochim Biophys Acta 1830(7): 3882-3892.

Gudernatsch, J., 1912. Feeding experiments on tadpoles. I. The influence of specific organs given as food on growth and differentiation: a contribution to the knowlege of organs with internal secretion. Wilhelm Roux Archiv der Entwicklungsmechanismus der Organismen 35: 457-483.

Guenther, M. G., Barak, O. and Lazar, M. A., 2001. The SMRT and N-CoR corepressors are activating cofactors for histone deacetylase 3. Mol Cell Biol 21(18): 6091-6101.

Guttman, M., Amit, I., Garber, M., French, C., Lin, M. F., Feldser, D., Huarte, M., Zuk, O., Carey, B. W., Cassady, J. P., Cabili, M. N., Jaenisch, R., Mikkelsen, T. S., Jacks, T., Hacohen, N., Bernstein, B. E., Kellis, M., Regev, A., Rinn, J. L. and Lander, E. S., 2009. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 458(7235): 223-227.

Guttman, M., Donaghey, J., Carey, B. W., Garber, M., Grenier, J. K., Munson, G., Young, G., Lucas, A. B., Ach, R., Bruhn, L., Yang, X. P., Amit, I., Meissner, A., Regev, A., Rinn, J. L., Root, D. E. and Lander, E. S., 2011. lincRNAs act in the circuitry controlling pluripotency and differentiation. Nature 477(7364): 295-U260.

Guttman, M., Garber, M., Levin, J. Z., Donaghey, J., Robinson, J., Adiconis, X., Fan, L., Koziol, M. J., Gnirke, A., Nusbaum, C., Rinn, J. L., Lander, E. S. and Regev, A., 2010. Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. Nature Biotechnol 28(5): 503-510.

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B., Lieber, M., Macmanes, M. D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C. N., Henschel, R., Leduc, R. D., Friedman, N. and Regev, A., 2013. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc 8(8): 1494-1512.

Hammond, S. A., Carew, A. C. and Helbing, C. C., 2013. Evaluation of the effects of titanium dioxide nanoparticles on cultured *Rana catesbeiana* tailfin tissue. Front Genet 4: 251.

Hammond, S. A., Veldhoen, N. and Helbing, C. C., 2015. Influence of temperature on thyroid hormone signaling and endocrine disruptor action in *Rana (Lithobates) catesbeiana* tadpoles. Gen Comp Endocrinol 219: 6-15.

Hammond, S. A., Veldhoen, N., Kobylarz, M., Webber, N. R., Jordan, J., Rehaume, V., Boone, M. D. and Helbing, C. C., 2013. Characterization of Gene Expression Endpoints During Postembryonic Development of the Northern Green Frog (*Rana clamitans melanota*). Zool Sci 30(5): 392-401.

Hanada, H., Kashiwagi, A., Takehara, Y., Kanno, T., Yabuki, M., Sasaki, J., Inoue, M. and Utsumi, K., 1997. Do reactive oxygen species underlie the mechanism of apoptosis in the tadpole tail? Free Radic Biol Med 23(2): 294-301.

Heinzel, T., Lavinsky, R. M., Mullen, T. M., Soderstrom, M., Laherty, C. D., Torchia, J., Yang, W. M., Brard, G., Ngo, S. D., Davie, J. R., Seto, E., Eisenman, R. N., Rose, D. W., Glass, C. K. and Rosenfeld, M. G., 1997. A complex containing N-CoR, mSin3 and histone deacetylase mediates transcriptional repression. Nature 387(6628): 43-48.

Helbing, C. 1993. Thyroid Hormone-Induced Changes in Gene Expression in the Bullfrog, *Rana catesbeiana*, Tadpole Liver. PhD, University of Western Ontario.

Helbing, C. C., 2012. The metamorphosis of amphibian toxicogenomics. Front Genet 3: 37.

Helbing, C. C. and Atkinson, B. G., 1994. 3,5,3'-Triiodothyronine-induced carbamyl phosphate synthetase gene expression is stabilized in the liver of *Rana catesbeiana* tadpoles during heat shock. J Biol Chem 269: 11743-11750.

Helbing, C. C., Bailey, C. M., Ji, L., Gunderson, M. P., Zhang, F., Veldhoen, N., Skirrow, R. C., Mu, R., Lesperance, M., Holcombe, G. W., Kosian, P. A., Tietge, J., Korte, J. J. and Degitz, S. J., 2007. Identification of gene expression indicators for thyroid axis disruption in a *Xenopus laevis* metamorphosis screening assay. Part 1. Effects on the brain. Aquat Toxicol 82(4): 227-241.

Helbing, C. C., Gergely, G. and Atkinson, B. G., 1992. Sequential up-regulation of thyroid hormone b receptor, ornithine transcarbamylase and carbamyl phosphate synthetase mRNAs in the liver of *Rana catesbeiana* tadpoles during spontaneous and thyroid hormone-induced metamorphosis. Dev Genet 13: 289-301.

Helbing, C. C., Ji, L., Bailey, C. M., Veldhoen, N., Zhang, F., Holcombe, G. W., Kosian, P. A., Tietge, J., Korte, J. J. and Degitz, S. J., 2007. Identification of gene expression indicators for thyroid axis disruption in a *Xenopus laevis* metamorphosis screening assay. Part 2. Effects on the tail and hindlimb. Aquat Toxicol 82(4): 215-226.

Helbing, C. C., Werry, K., Crump, D., Domanski, D., Veldhoen, N. and Bailey, C. M., 2003. Expression profiles of novel thyroid hormone-responsive genes and proteins in the tail of *Xenopus laevis* tadpoles undergoing precocious metamorphosis. Mol Endocrinol 17(7): 1395-1409.

Hellsten, U., Harland, R. M., Gilchrist, M. J., Hendrix, D., Jurka, J., Kapitonov, V., Ovcharenko, I., Putnam, N. H., Shu, S., Taher, L., Blitz, I. L., Blumberg, B., Dichmann, D. S., Dubchak, I., Amaya, E., Detter, J. C., Fletcher, R., Gerhard, D. S., Goodstein, D., Graves, T., Grigoriev, I. V., Grimwood, J., Kawashima, T., Lindquist, E., Lucas, S. M., Mead, P. E., Mitros, T., Ogino, H., Ohta, Y., Poliakov, A. V., Pollet, N., Robert, J., Salamov, A., Sater, A. K., Schmutz, J., Terry, A., Vize, P. D., Warren, W. C., Wells, D., Wills, A., Wilson, R. K., Zimmerman, L. B., Zorn, A. M., Grainger, R., Grammer, T., Khokha, M. K., Richardson, P. M. and Rokhsar, D. S., 2010. The genome of the Western clawed frog *Xenopus tropicalis*. Science 328(5978): 633-636.

Hertwig, O., 1897. Über den Einfluss der Temperatur auf die Entwicklung von *Rana fusca* und *Rana esculenta*. Archiv für mikroskopische Anatomie 51(1): 319-381.

Hinther, A., Bromba, C. M., Wulff, J. E. and Helbing, C. C., 2011. Effects of triclocarban, triclosan, and methyl triclosan on thyroid hormone action and stress in frog and mammalian culture systems. Environ Sci Technol 45(12): 5395-5402.

Hinther, A., Domanski, D., Vawda, S. and Helbing, C. C., 2010. C-fin: a cultured frog tadpole tail fin biopsy approach for detection of thyroid hormone-disrupting chemicals. Environ Toxicol Chem 29(2): 380-388.

Hinther, A., Edwards, T. M., Guillette, L. J., Jr. and Helbing, C. C., 2012. Influence of nitrate and nitrite on thyroid hormone responsive and stress-associated gene expression in cultured *Rana catesbeiana* tadpole tail fin tissue. Front Genet 3: 51.

Hornett, E. A. and Wheat, C. W., 2012. Quantitative RNA-Seq analysis in non-model species: assessing transcriptome assemblies as a scaffold and the utility of evolutionary divergent genomic reference species. BMC Genomics 13: 361.

Hourdry, J., L'Hermite, A. and Ferrand, R., 1996. Changes in the digestive tract and feeding behavior of anuran amphibians during metamorphosis. Physiol Zool 69(2): 219-251.

Huang, D. W., Sherman, B. T. and Lempicki, R. A., 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4(1): 44-57.

Huang, E. Y., Zhang, J., Miska, E. A., Guenther, M. G., Kouzarides, T. and Lazar, M. A., 2000. Nuclear receptor corepressors partner with class II histone deacetylases in a Sin3-independent repression pathway. Genes Dev 14(1): 45-54.

Huxley, J. S., 1929. Thyroid and temperature in cold-blooded vertebrates. Nature 123: 712.

Ichu, T. A., Han, J., Borchers, C. H., Lesperance, M. and Helbing, C. C., 2014. Metabolomic insights into system-wide coordination of vertebrate metamorphosis. Bmc Developmental Biology 14.

Ishizuya-Oka, A., Hasebe, T. and Shi, Y. B., 2010. Apoptosis in amphibian organs during metamorphosis. Apoptosis 15(3): 350-364.

Ishizuya-Oka, A., Ueda, S. and Shi, Y.-B., 1997. Temporal and spatial regulation of a putative transcriptional repressor implicates it as playing a role in thyroid hormone-dependent organ transformation. Dev Genet 20: 329-337.

Isorna, E., Obregon, M. J., Calvo, R. M., Vazquez, R., Pendon, C., Falcon, J. and Munoz-Cueto, J. A., 2009. Iodothyronine deiodinases and thyroid hormone receptors regulation during flatfish (*Solea senegalensis*) metamorphosis. J Exp Zool B Mol Dev Evol 312B(3): 231-246.

James-Zorn, C., Ponferrada, V. G., Jarabek, C. J., Burns, K. A., Segerdell, E. J., Lee, J., Snyder, K., Bhattacharyya, B., Karpinka, J. B., Fortriede, J., Bowes, J. B., Zorn, A. M. and Vize, P. D., 2013. Xenbase: expansion and updates of the Xenopus model organism database. Nucleic Acids Res 41(D1): D865-D870.

Ji, L., Domanski, D., Skirrow, R. C. and Helbing, C. C., 2007. Genistein prevents thyroid hormone-dependent tail regression of *Rana catesbeiana* tadpoles by targetting protein kinase C and thyroid hormone receptor alpha. Dev Dyn 236(3): 777-790.

Kanamori, A. and Brown, D., 1993. Cultured cells as a model for amphibian metamorphosis. Proc Natl Acad Sci U.S.A. 90: 6013-6017.

Kashiwagi, A., Hanada, H., Yabuki, M., Kanno, T., Ishisaka, R., Sasaki, J., Inoue, M. and Utsumi, K., 1999. Thyroxine enhancement and the role of reactive oxygen species in tadpole tail apoptosis. Free Radic Biol Med 26(7-8): 1001-1009.

Kiemnec-Tyburczy, K. M., Richmond, J. Q., Savage, A. E., Lips, K. R. and Zamudio, K. R., 2012. Genetic diversity of MHC class I loci in six non-model frogs is shaped by positive selection and gene duplication. Heredity 109: 146-155.

Klein, I. and Danzi, S., 2007. Thyroid disease and the heart. Circulation 116(15): 1725-1735.

Koenig, R. J., Lazar, M. A., Hodin, R. A., Brent, G. A., Larsen, P. R., Chin, W. W. and Moore, D. D., 1989. Inhibition of thyroid hormone action by a non-hormone binding c-erbA protein generated by alternative mRNA splicing. Nature 337(6208): 659-661.

Langlois, V. S., Duarte-Guterman, P., Ing, S., Pauli, B. D., Cooke, G. M. and Trudeau, V. L., 2010. Fadrozole and finasteride exposures modulate sex steroid- and thyroid hormone-related gene expression in *Silurana (Xenopus) tropicalis* early larval development. Gen Comp Endocrinol 166(2): 417-427.

Langlois, V. S., Duarte-Guterman, P. and Trudeau, V. L., 2011. Expression profiles of reproduction- and thyroid hormone-related transcripts in the brains of chemically-induced intersex frogs. Sexual Dev 5(1): 26-32.

Langmead, B. and Salzberg, S. L., 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9(4): 357-359.

Larsen, D. A., Swanson, P. and Dickhoff, W. W., 2011. The pituitary-thyroid axis during the parr-smolt transformation of Coho salmon, *Oncorhynchus kisutch*: quantification of TSH β mRNA, TSH, and thyroid hormones. Gen Comp Endocrinol 171(3): 367-372.

Lazar, M. A. and Chin, W. W., 1990. Nuclear thyroid hormone receptors. J Clin Invest 86(6): 1777-1782.

Leloup, J. and Buscaglia, M., 1977. La triiodothyronine, hormone de la metamorphose des amphibiens. Comptes Rendues de l'Academie de les Sciences Paris Serie D 284: 2261-2263.

Levy, O., Dai, G., Riedel, C., Ginter, C. S., Paul, E. M., Lebowitz, A. N. and Carrasco, N., 1997. Characterization of the thyroid Na+/I- symporter with an anti-COOH terminus antibody. Proc Natl Acad Sci U.S.A. 94(11): 5568-5573.

Li, H. and Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25(14): 1754-1760.

Lien, R. J. and Siopes, T. D., 1993. The relationship of plasma thyroid hormone and prolactin concentrations to egg laying, incubation behavior, and molting by female turkeys exposed to a one-year natural daylength cycle. Gen Comp Endocrinol 90(2): 205-213.

Livak, K. J. and Schmittgen, T. D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ Method. Methods 25(4): 402-408.

Love, M. I., Huber, W. and Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15(12): 550.

Luo, Y., Guo, W., Ngo, H. H., Nghiem, L. D., Hai, F. I., Zhang, J., Liang, S. and Wang, X. C., 2014. A review on the occurrence of micropollutants in the aquatic environment and their fate and removal during wastewater treatment. Sci Total Environ 473-474: 619-641.

Maher, S. K., Wojnarowicz, P., Ichu, T.-A., Veldhoen, N., Lu, L., Lesperance, M., Propper, C. R. and Helbing, C. C., 2015. Rethinking the biological relationships of the thyroid hormones, L-thyroxine and 3,5,3'-triiodothyronine. Gen. Comp. Endocrinol. Submitted.

Mangelsdorf, D. J., Thummel, C., Beato, M., Herrlich, P., Schutz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P. and Evans, R. M., 1995. The nuclear receptor superfamily: the second decade. Cell 83(6): 835-839.

Marlatt, V. L., Veldhoen, N., Lo, B. P., Bakker, D., Rehaume, V., Vallee, K., Haberl, M., Shang, D., van Aggelen, G. C., Skirrow, R. C., Elphick, J. R. and Helbing, C. C., 2013. Triclosan exposure alters postembryonic development in a Pacific tree frog (*Pseudacris regilla*) Amphibian Metamorphosis Assay (TREEMA). Aquat Toxicol 126: 85-94.

Martin, J. A. and Wang, Z., 2011. Next-generation transcriptome assembly. Nat Rev Genet 12(10): 671-682.

McMahon, T. A., Sears, B. F., Venesky, M. D., Bessler, S. M., Brown, J. M., Deutsch, K., Halstead, N. T., Lentz, G., Tenouri, N., Young, S., Civitello, D. J., Ortega, N., Fites, J. S., Reinert, L. K., Rollins-Smith, L. A., Raffel, T. R. and Rohr, J. R., 2014. Amphibians acquire resistance to live and dead fungus overcoming fungal immunosuppression. Nature 511(7508): 224-227.

McMenamin, S. K., Bain, E. J., McCann, A. E., Patterson, L. B., Eom, D. S., Waller, Z. P., Hamill, J. C., Kuhlman, J. A., Eisen, J. S. and Parichy, D. M., 2014. Thyroid hormone–dependent adult pigment cell lineage and pattern in zebrafish. Science 345(6202): 1358-1361.

Mehr, S. F. P., DeSalle, R., Kao, H. T., Narechania, A., Han, Z., Tchernov, D., Pieribone, V. and Gruber, D. F., 2013. Transcriptome deep-sequencing and clustering of expressed isoforms from *Favia* corals. BMC Genomics 14.

Mendoza, A., Navarrete-Ramírez, P., Hernández-Puga, G., Villalobos, P., Holzer, G., Renaud, J. P., Laudet, V. and Orozco, A., 2013. 3,5-T_2 is an alternative ligand for the thyroid hormone receptor β1. Endocrinology 154(8): 2948-2958.

Menéndez-Hurtado, A., Santos, A. and Pérez-Castillo, A., 2000. Characterization of the promoter region of the rat CCAAT/enhancer-binding protein α gene and regulation by thyroid hormone in rat immortalized brown adipocytes. Endocrinology 141: 4164-4170.

Mochizuki, K., Goda, T. and Yamauchi, K., 2012. Gene expression profile in the liver of *Rana catesbeiana* tadpoles exposed to low temperature in the presence of thyroid hormone. Biochem Biophy Res Commun 420(4): 845-850.

Mora, C., Tittensor, D. P., Adl, S., Simpson, A. G. and Worm, B., 2011. How many species are there on Earth and in the ocean? PLoS Biology 9(8): e1001127.

Mukhi, S., Cai, L. and Brown, D. D., 2010. Gene switching at *Xenopus laevis* metamorphosis. Dev Biol 338(2): 117-126.

Mullur, R., Liu, Y. Y. and Brent, G. A., 2014. Thyroid hormone regulation of metabolism. Physiol Rev 94(2): 355-382.

Murata, T. and Yamauchi, K., 2005. Low-temperature arrest of the triiodothyronine-dependent transcription in *Rana catesbeiana* red blood cells. Endocrinology 146(1): 256-264.

Nakasugi, K., Crowhurst, R., Bally, J. and Waterhouse, P., 2014. Combining transcriptome assemblies from multiple *de novo* assemblers in the allo-tetraploid plant *Nicotiana benthamiana*. PLoS One 9(3): e91776.

Nieuwkoop, P. D. and Faber, J. 1956. Normal table of *Xenopus laevis*. Amsterdam, North Holland Publishing.

Nordberg, H., Cantor, M., Dusheyko, S., Hua, S., Poliakov, A., Shabalov, I., Smirnova, T., Grigoriev, I. V. and Dubchak, I., 2014. The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. Nucleic Acids Res 42(D1): D26-D31.

Nussey, S. and Whitehead, S. 2001. The Thyroid Gland. Oxford, BIOS Scientific Publishers.

Oetting, A. and Yen, P. M., 2007. New insights into thyroid hormone action. Best Pract Res Clin Endocrinol Metab 21(2): 193-208.

Oppenheimer, J. H., 1968. Role of plasma proteins in the binding, distribution and metabolism of the thyroid hormones. N Engl J Med 278(21): 1153-1162.

Ortiga-Carvalho, T. M., Shibusawa, N., Nikrodhanond, A., Oliveira, K. J., Machado, D. S., Liao, X. H., Cohen, R. N., Refetoff, S. and Wondisford, F. E., 2005. Negative regulation by thyroid hormone receptor requires an intact coactivator-binding surface. J Clin Invest 115(9): 2517-2523.

Osachoff, H. L., Mohammadali, M., Skirrow, R. C., Hall, E. R., Brown, L. L., van Aggelen, G. C., Kennedy, C. J. and Helbing, C. C., 2014. Evaluating the treatment of a synthetic wastewater containing a pharmaceutical and personal care product chemical cocktail: compound removal efficiency and effects on juvenile rainbow trout. Water Res 62: 271-280.

Oshlack, A., Robinson, M. D. and Young, M. D., 2010. From RNA-seq reads to differential expression results. Genome Biol 11(12): 220.

Ozsolak, F. and Milos, P. M., 2011. RNA sequencing: advances, challenges and opportunities. Nat Rev Genet 12(2): 87-98.

Pagani, I., Liolios, K., Jansson, J., Chen, I. M., Smirnova, T., Nosrat, B., Markowitz, V. M. and Kyrpides, N. C., 2012. The Genomes OnLine Database (GOLD) v.4: status of genomic and metagenomic projects and their associated metadata. Nucleic Acids Res 40(Database issue): D571-579.

Paik, W. K. and Cohen, P. P., 1960. Biochemical studies on amphibian metamorphosis. I. The effect of thyroxine on protein synthesis in the tadpole. J Gen Physiol 43: 683-696.

Paquette, M. A., Atlas, E., Wade, M. G. and Yauk, C. L., 2014. Thyroid hormone response element half-site organization and its effect on thyroid hormone mediated transcription. PLoS One 9(6): e101155.

Parker, F., Jr., Robbins, S. L. and Loveridge, A., 1947. Breeding, rearing and care of the South African clawed frog (*Xenopus laevis*). Am Nat 81(796): 38-49.

Parra, G., Bradnam, K., Ning, Z., Keane, T. and Korf, I., 2009. Assessing the gene space in draft genomes. Nucleic Acids Res 37(1): 289-297.

Pascual, A. and Aranda, A., 2013. Thyroid hormone receptors, cell growth and differentiation. Biochim Biophys Acta 1830(7): 3908-3916.

Préau, L., Fini, J. B., Morvan-Dubois, G. and Demeneix, B., 2015. Thyroid hormone signaling during early neurogenesis and its significance as a vulnerable window for endocrine disruption. Biochim Biophys Acta 1849(2): 112-121.

Pruitt, K. D., Tatusova, T. and Maglott, D. R., 2007. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res 35(Database issue): D61-65.

Quinlan, A. R. and Hall, I. M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26(6): 841-842.

R Core Team, 2013. R: A language and environment for statistical computing. Vienna, Austria, R Foundation for Statistical Computing.

Rainsford, K. D., 2009. Ibuprofen: pharmacology, efficacy and safety. Inflammopharmacology 17(6): 275-342.

Refetoff, S., Robin, N. I. and Fang, V. S., 1970. Parameters of thyroid function in serum of 16 selected vertebrate species: a study of PBI, serum T_4, free T_4, and the pattern of T_4 and T_3 binding to serum proteins. Endocrinology 86(4): 793-805.

Regard, E., Taurog, A. and Nakashima, T., 1978. Plasma thyroxine and triiodothyronine levels in spontaneously metamorphosing *Rana catesbeiana* tadpoles and in adult anuran amphibia. Endocrinology 102: 674-684.

Relyea, R. A. and Jones, D. K., 2009. The toxicity of Roundup Original Max to 13 species of larval amphibians. Environ Toxicol Chem 28(9): 2004-2008.

Ribeiro, J. M. C., Chagas, A. C., Pham, V. M., Lounibos, L. P. and Calvo, E., 2014. An insight into the sialome of the frog biting fly, *Corethrella appendiculata*. Insect Biochem Mol Biol 44: 23-32.

Robert, J. and Ohta, Y., 2009. Comparative and Developmental Study of the Immune System in *Xenopus*. Dev Dyn 238(6): 1249-1270.

Robertson, G., Schein, J., Chiu, R., Corbett, R., Field, M., Jackman, S. D., Mungall, K., Lee, S., Okada, H. M., Qian, J. Q., Griffith, M., Raymond, A., Thiessen, N., Cezard, T., Butterfield, Y. S., Newsome, R., Chan, S. K., She, R., Varhol, R., Kamoh, B., Prabhu, A. L., Tam, A., Zhao, Y., Moore, R. A., Hirst, M., Marra, M. A., Jones, S. J., Hoodless, P. A. and Birol, I., 2010. *De novo* assembly and analysis of RNA-seq data. Nat Methods 7(11): 909-912.

Robinson, D. H. and Heintzelman, M. B., 1987. Morphology of ventral epidermis of *Rana catesbeiana* during metamorphosis. Anat Rec 217(3): 305-317.

Rollins-Smith, L. A., 1998. Metamorphosis and the amphibian immune system. Immunol Rev 166: 221-230.

Rosenblum, E. B., Poorten, T. J., Settles, M., Murdoch, G. K., Robert, J., Maddox, N. and Eisen, M. B., 2009. Genome-wide transcriptional response of *Silurana (Xenopus) tropicalis* to infection with the deadly chytrid fungus. PLoS One 4(8).

Sakurai, A., Miyamoto, T. and DeGroot, L. J., 1992. Cloning and characterization of the human thyroid hormone receptor β1 gene promoter. Biochem Biophys Res Commun 185(1): 78-84.

Sap, J., Muñoz, A., Damm, K., Goldberg, Y., Ghysdael, J., Leutz, A., Beug, H. and Vennström, B., 1986. The c-erb-A protein is a high-affinity receptor for thyroid hormone. Nature 324(6098): 635-640.

Savage, A. E., Kiemnec-Tyburczy, K. M., Ellison, A. R., Fleischer, R. C. and Zamudio, K. R., 2014. Conservation and divergence in the frog immunome: pyrosequencing and *de novo* assembly of immune tissue transcriptomes. Gene 542(2): 98-108.

Schneider, M. J., Davey, J. C. and Galton, V. A., 1993. *Rana catesbeiana* tadpole red blood cells express an α, but not a β, c-erbA gene. Endocrinology 133(6): 2488-2495.

Schneider, M. J. and Galton, V. A., 1991. Regulation of c-erbA-alpha messenger RNA species in tadpole erythrocytes by thyroid hormone. Mol Endocrinol 5(2): 201-208.

Schulz, M. H., Zerbino, D. R., Vingron, M. and Birney, E., 2012. Oases: robust *de novo* RNA-seq assembly across the dynamic range of expression levels. Bioinformatics 28(8): 1086-1092.

Searcy, B. T., Beckstrom-Sternberg, S. M., Beckstrom-Sternberg, J. S., Stafford, P., Schwendiman, A. L., Soto-Pena, J., Owen, M. C., Ramirez, C., Phillips, J., Veldhoen, N., Helbing, C. C. and Propper, C. R., 2012. Thyroid hormone-dependent development in *Xenopus laevis*: A sensitive screen of thyroid hormone signaling disruption by municipal wastewater treatment plant effluent. Gen Comp Endocrinol 176: 481-492.

Shi, Y.-B. 2000. Amphibian Metamorphosis: From morphology to molecular biology. New York, Wiley-Liss.

Shi, Y.-B., Liang, V. C., Parkison, C. and Cheng, S. Y., 1994. Tissue-dependent developmental expression of a cytosolic thyroid hormone protein gene in *Xenopus*: Its role in the regulation of amphibian metamorphosis. FEBS Lett 355(1): 61-64.

Sillar, K. T., Combes, D., Ramanathan, S., Molinari, M. and Simmers, J., 2008. Neuromodulation and developmental plasticity in the locomotor system of anuran amphibians during metamorphosis. Brain Res Rev 57(1): 94-102.

Simpson, J. T. and Durbin, R., 2012. Efficient *de novo* assembly of large genomes using compressed data structures. Genome Res 22(3): 549-556.

Sirakov, M., Skah, S., Nadjar, J. and Plateroti, M., 2013. Thyroid hormone's action on progenitor/stem cell biology: new challenge for a classic hormone? Biochim Biophys Acta 1830(7): 3917-3927.

Skirrow, R. C. and Helbing, C. C., 2007. Decreased cyclin-dependent kinase activity promotes thyroid hormone-dependent tail regression in *Rana catesbeiana*. Cell Tissue Res 328(2): 281-289.

Skirrow, R. C., Veldhoen, N., Domanski, D. and Helbing, C. C., 2008. Roscovitine inhibits thyroid hormone-induced tail regression of the frog tadpole and reveals a role for cyclin C/Cdk8 in the establishment of the metamorphic gene expression program. Dev Dyn 237(12): 3787-3797.

Sookruksawong, S., Sun, F. Y., Liu, Z. J. and Tassanakajon, A., 2013. RNA-Seq analysis reveals genes associated with resistance to Taura syndrome virus (TSV) in the Pacific white shrimp *Litopenaeus vannamei*. Dev Comp Immunol 41(4): 523-533.

Steinfelder, H. J. and Wondisford, F. E., 1997. Thyrotropin (TSH) β-subunit gene expression--an example for the complex regulation of pituitary hormone genes. Exp Clin Endocrinol Diabetes 105(4): 196-203.

Stuhlmeier, K. M., Li, H. and Kao, J. J., 1999. Ibuprofen: new explanation for an old phenomenon. Biochem Pharmacol 57(3): 313-320.

Sugrue, M. L., Vella, K. R., Morales, C., Lopez, M. E. and Hollenberg, A. N., 2010. The thyrotropin-releasing hormone gene is regulated by thyroid hormone at the level of transcription *in vivo*. Endocrinology 151(2): 793-801.

Sumida, M., Kato, Y. and Kurabayashi, A., 2004. Sequencing and analysis of the internal transcribed spacers (ITSs) and coding regions in the EcoR I fragment of the ribosomal DNA of the Japanese pond frog *Rana nigromaculata*. Genes Genet Sys 79(2): 105-118.

Supek, F., Bošnjak, M., Škunca, N. and Šmuc, T., 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One 6(7): e21800.

Suzuki, K., Machiyama, F., Nishino, S., Watanabe, Y., Kashiwagi, K., Kashiwagi, A. and Yoshizato, K., 2009. Molecular features of thyroid hormone-regulated skin remodeling in *Xenopus laevis* during metamorphosis. Dev Growth Differ 51(4): 411-427.

Suzuki, K., Utoh, R., Kotani, K., Obara, M. and Yoshizato, K., 2002. Lineage of anuran epidermal basal cells and their differentiation potential in relation to metamorphic skin remodeling. Dev Growth Differ 44(3): 225-238.

Tan, M. H., Au, K. F., Yablonovitch, A. L., Wills, A. E., Chuang, J., Baker, J. C., Wong, W. H. and Li, J. B., 2013. RNA sequencing reveals a diverse and dynamic repertoire of the *Xenopus tropicalis* transcriptome over development. Genome Res 23(1): 201-216.

Taurog, A. 1996. Hormone synthesis: Thyroid iodine metabolism. Werner and Ingbar's The Thyroid: A fundamental and clinical text. L. E. a. U. Braverman, R.D. Philadelphia, Lippincott-Raven**:** 47-81.

Taurog, A., Dorris, M. L. and Doerge, D. R., 1996. Mechanism of simultaneous iodination and coupling catalyzed by thyroid peroxidase. Arch Biochem Biophys 330(1): 24-32.

Taylor, A. C. and Kollros, J. J., 1946. Stages in the normal development of *Rana pipiens* larvae. Anatomical Record 94: 7-24.

Tegeder, I., Pfeilschifter, J. and Geisslinger, G., 2001. Cyclooxygenase-independent actions of cyclooxygenase inhibitors. FASEB J 15(12): 2057-2072.

Thormeyer, D. and Baniahmad, A., 1999. The v-erbA oncogene (review). Int J Mol Med 4(4): 351-358.

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., Salzberg, S. L., Wold, B. J. and Pachter, L., 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol 28(5): 511-515.

Ulitsky, I., Shkumatava, A., Jan, C. H., Sive, H. and Bartel, D. P., 2011. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. Cell 147(7): 1537-1550.

Veldhoen, N., Crump, D., Werry, K. and Helbing, C., 2002. Distinctive gene profiles occur at key points during natural metamorphosis in the *Xenopus laevis* tadpole tail. Dev Dyn 225: 457-468.

Veldhoen, N. and Helbing, C. C. 2005. Monitoring gene expression in *Rana catesbeiana* tadpoles using a tail fin biopsy technique and its application to the detection of environmental endocrine disruptor effects in wildlife species. Techniques in Aquatic Toxicology. Volume 2. G. K. Ostrander. Boca Raton, CRC Press**:** 315-327.

Veldhoen, N., Propper, C. R. and Helbing, C. C., 2014. Enabling comparative gene expression studies of thyroid hormone action through the development of a flexible real-time quantitative PCR assay for use across multiple anuran indicator and sentinel species. Aquat Toxicol 148: 162-173.

Veldhoen, N., Skirrow, R., Osachoff, H., Wigmore, H., Clapson, D., Gunderson, M., van Aggelen, G. and Helbing, C., 2006. The bactericidal agent triclosan modulates thyroid hormone-associated gene expression and disrupts postembryonic anuran development. Aquat toxicol 80: 217-227.

Veldhoen, N., Skirrow, R. C., Brown, L. L., van Aggelen, G. and Helbing, C. C., 2014. Effects of acute exposure to the non-steroidal anti-inflammatory drug ibuprofen on the developing North American bullfrog (*Rana catesbeiana*) tadpole. Environ Sci Technol. 48(17): 10439-10447.

Veldhoen, N., Skirrow, R. C., Ji, L., Domanski, D., Bonfield, E. R., Bailey, C. M. and Helbing, C. C., 2006. Use of heterologous cDNA arrays and organ culture in the detection of thyroid hormone-dependent responses in a sentinel frog, *Rana catesbeiana*. Comp Biochem Physiol D 1: 187-199.

Vella, K. R. and Hollenberg, A. N., 2009. The ups and downs of thyrotropin-releasing hormone. Endocrinology 150(5): 2021-2023.

Vezina, F., Gustowska, A., Jalvingh, K. M., Chastel, O. and Piersma, T., 2009. Hormonal correlates and thermoregulatory consequences of molting on metabolic rate in a northerly wintering shorebird. Physiol Biochem Zool 82(2): 129-142.

Vijay, N., Poelstra, J. W., Kunstner, A. and Wolf, J. B., 2013. Challenges and strategies in transcriptome assembly and differential gene expression quantification. A comprehensive *in silico* assessment of RNA-seq experiments. Mol Ecol 22(3): 620-634.

Wang, D. Q., Xia, X. M., Liu, Y., Oetting, A., Walker, R. L., Zhu, Y. L., Meltzer, P., Cole, P. A., Shi, Y. B. and Yen, P. M., 2009. Negative regulation of TSHα target gene by thyroid hormone involves histone acetylation and corepressor complex dissociation. Mol Endocrinol 23(5): 600-609.

Wang, Z. and Brown, D. D., 1993. Thyroid hormone-induced gene expression program for amphibian tail resorption. J Biol Chem 268: 16270-16278.

Wang, Z., Gerstein, M. and Snyder, M., 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10(1): 57-63.

Webb, P. and Baxter, J. D. 2007. Introduction to endocrinology. Greenspan's Basic & Clinical Endocrinology, Eighth Edition. D.G. Gardner and D. Shoback (Eds.). New York, McGraw Hill Medical.

White, B. A. and Nicoll, C. S. 1981. Hormonal control of amphibian metamorphosis. Metamorphosis: a problem in developmental biology. L. I. Gilbert and E. Frieden. New York, Plenum Publishing.

Wickham, H. 2009. ggplot2: elegant graphics for data analysis, Springer New York.

Williams, G. R., 2000. Cloning and characterization of two novel thyroid hormone receptor β isoforms. Mol Cell Biol 20(22): 8329-8342.

Wojnarowicz, P., Ogunlaja, O. O., Xia, C., Parker, W. J. and Helbing, C. C., 2013. Impact of wastewater treatment configuration and seasonal conditions on thyroid hormone disruption and stress effects in *Rana catesbeiana* tailfin. Environ Sci Technol 47(23): 13840-13847.

Wolff, J., Winand, R. J. and Kohn, L. D., 1974. The contribution of subunits of thyroid stimulating hormone to the binding and biological activity of thyrotropin. Proc Natl Acad Sci U.S.A. 71(9): 3460-3464.

Wondisford, F. E., Farr, E. A., Radovick, S., Steinfelder, H. J., Moates, J. M., McClaskey, J. H. and Weintraub, B. D., 1989. Thyroid hormone inhibition of human thyrotropin β-subunit gene expression is mediated by a cis-acting element located in the first exon. J Biol Chem 264(25): 14601-14604.

Xu, G. R., Strong, M. J., Lacey, M. R., Baribault, C., Flemington, E. K. and Taylor, C. M., 2014. RNA CoMPASS: A dual approach for pathogen and host transcriptome analysis of RNA-Seq datasets. PLoS One 9(2).

Yamamoto, K., Kanski, D. and Frieden, E., 1966. Uptake and excretion of thyroxine, triiodothyronine and iodide in bullfrog tadpoles after immersion or injection at 25° and 6°C. Gen Comp Endocrinol 6(3): 312-324.

Yaoita, Y. and Brown, D. D., 1990. A correlation of thyroid hormone receptor gene expression with amphibian metamorphosis. Genes Dev 4(11): 1917-1924.

Yen, P. M., 2001. Physiological and molecular basis of thyroid hormone action. Physiol Rev 81(3): 1097-1142.

Yoshizato, K., 2007. Molecular mechanism and evolutional significance of epithelial-mesenchymal interactions in the body- and tail-dependent metamorphic transformation of anuran larval skin. Int Rev Cytol 260: 213-260.

Yueh, M. F., Taniguchi, K., Chen, S., Evans, R. M., Hammock, B. D., Karin, M. and Tukey, R. H., 2014. The commonly used antimicrobial additive triclosan is a liver tumor promoter. Proc Natl Acad Sci U.S.A. 111(48): 17200-17205.

Zhang, F., Degitz, S. J., Holcombe, G. W., Kosian, P. A., Tietge, J., Veldhoen, N. and Helbing, C. C., 2006. Evaluation of gene expression endpoints in the context of a *Xenopus laevis* metamorphosis-based bioassay to detect thyroid hormone disruptors. Aquat Toxicol 76: 24-36.

Zorn, A. M. 2008. Liver development. StemBook. Cambridge (MA).

Appendix

Appendix 1 Characteristics of qPCR assay reagents used in assessment of TH-induced precocious metamorphosis in *Rana catesbeiana*

Gene	GenBank Acc #	Primer Name	Primer Sequence[c]	Amplicon size (bp)	qPCR Type	Tissue Type[f]	Primer Amt (pmol/rxn)	T_a (°C)
rpl8	AY452063	AMM1	AGGCAGGTCGTGCNTACCA	89	TaqMan	All	1.5	64[b]
		AMM2	GGGATGTTCTACAGGATTCATAGC				1.5	
		AMM3	Cy5-AAACTGCTGGCCACGTGTCCGT-IABk				1.5	
thra	L06064	AMM4	TGATAAGGCCACAGGRTACCACTA	141	TaqMan	All	4.5	64[b]
		AMM5	CGGGTGATCTTGTCGATRA				4.5	
		AMM6	FAM-ACTATCCAGAAGAACCTGCACCCCTC-IABk				1.5	
thrb	L27344	AMM7	CTCATAGAAGAAAACAGAGAAAARAGA	237	TaqMan	All	4.5	64[b]
		AMM8	GAAGGCTTCTAAGTCCACTTTTCC				4.5	
		AMM9	HEX-CATGTGGCCACCAATGCACAGG-IABk				1.5	
hsp30	U44894	AMM13	GCCTCCACCAGACTTACCA	238	TaqMan	All	4.5	64[b]
		AMM14	TCTGTCTCCCTTTTCTTGTCG				4.5	
		AMM15	HEX-CCACCGCCCCTCAAGACAAATC-IABk				1.5	
cat	GQ222411	AMM16	GAATGGTTACGGCTCACACA	176	TaqMan	All	1.5	64[b]
		AMM17	TGGCAATGGCTTCATACAGAT				1.5	
		AMM18	Cy5-CAGGGCATCAGGAATCTGACGGT-IABk				1.5	
sod	BT081775	AMM19	CGAGCAGGAAGAAGATGGA	323	TaqMan	All	4.5	64[b]
		AMM20	CGCCTTTTCCCAAGTCATC				4.5	
		AMM21	FAM-ATTTCAACCCCCAAGGCAAGACC-IABk				1.5	
eef1a	BT081920	AMM43	GCTGCTGGTGTTGGTGART	239	TaqMan	L,B,T	1.5	60[b]
		AMM44	AGCATGTTGTCACCRTTCC				1.5	
		AMM45	HEX-TACATCAAGAAGATTGGTTACAACCC-IABk				1.5	
rps10	BT081704	AMM37	GCYTGGCGTCACTTTTACTG	289	TaqMan	L,B,T	1.5	60[b]
		AMM38	CACGTCCAAAYCCTCCTCTAA				1.5	
		AMM39	FAM-AAGGCTGAGGCTGGWGCTGGAG-IABk				1.5	
dio2	L42815	TAX12aup	CCTGGCTCTSTAYGACTC	295	SYBR	B,BS	5.0	62[a]
		TAX12ddn	RGCTGATCCRAARTTGAC				5.0	

thibz	CAT_THBZIP[d]	TAX15a up TAX15a dn	ASCTCCRCAGAAYCAGCA TCACGTACCAGGCCAAAA	354	SYBR	All	5.0 5.0	62[b]
klf9	CAT_KLF9[d]	TAX16a up TAX16a dn	CYGCTCAGTGTCTGGTGT ARGGGCCGGTACTTGTTT	250	SYBR	All	5.0 5.0	62[b]
cps1	U05193	FM013 up FM013 dn	CCATAGTTGTCGCTCCTT ACATCCTCAGGGCTTTCT	459	SYBR	L	5.0 5.0	55[a]
otc	U26355	TAX28 up TAX28 dn	YATGACYGATGCTGTTCTAG CATAWCCCTTTGGTGTTGC	272	SYBR	L	5.0 5.0	60[b]
rlk1	EF156435	DDKerF3 DDKerRQ	GTTGGCGTTGGTGTTAGCGC GGCACTGCTTCTTGCAACTTG	336	SYBR	B,T,BS	5.0 5.0	55[a]
eef1a	BT081920	TAX23 up TAX23 dn	GCTGCTGGTGTTGGTGART AGCATGTTGTCACCRTTCC	257	SYBR	BS,Lu	5.0 5.0	60[a]
rps10	BT081704	TAX7 up TAX7 dn	TTTGCYTGGCGKCACTTTT ARCRGCACTGCGYCTGTA	213	SYBR	BS,Lu	5.0 5.0	60[a]
cirbp[e]	AB025351	CATCIRBP up CATCIRBP dn	ATCCAAGAAGTGGTTGTGGTG AACACGAATCTGACGACCATC	144	SYBR	All	5.0 5.0	60[a]
dio3[e]	L41731	CATDIO3 up CATDIO3 dn	ATCAATGCTTGGGTCAGAAC TGGATAGCACCTTTCCCAGTT	88	SYBR	L,B,T	5.0 5.0	60[a]
cebp1[e]	U08604	CATCEBP1 up CATCEBP1 dn	AAAGTTGAGGAAGAGGGTGGA ACAGTTGCCCATCACTTTGAC	106	SYBR	All	5.0 5.0	60[a]

[a]The qPCR thermoprogram included 40 cycles of 15 sec (denaturation), 30 sec (annealing), and 45 sec (extension). [b]The qPCR thermoprogram included 40 cycles of 15 sec (denaturation), 30 sec (annealing), 30 sec (extension). [c]Underlined primer sequence denotes the presence of an internal ZEN quencher. [d]Sequence information derived from unpublished *Rana catesbeiana* hepatic RNAseq analysis. [e]qPCR primer sequences obtained from Mochizuki et al. [f]All = all tissues examined, L = liver, B = brain, T = tail fin, BS = back skin, Lu = lung.

Appendix 2 Comparison of liver transcriptome RNA-seq with qPCR results

		qPCR[a]			RNA-seq[b]		
Species	Gene transcript	Control[c]	T_3[c]	Fold Change	Control[c]	T_3[c]	Fold Change
CAT[d]	*thra*	1.00±0.31	1.59±0.50	1.59	52	144	2.77
	thrb	1.00±0.00	20.05±9.80	20.05	1	15	15.00
	col1a2	1.00±0.49	2.10±0.73	2.10	1940	1778	0.92
	dio2	ND[e]	ND	ND	ND	ND	ND
	klf9	1.00±0.83	1.38±0.46	1.38	3242	2707	0.83
	timm50	1.00±0.46	2.33±0.99	2.33	139	186	1.34
	otc	1.00±0.22	5.11±1.37	5.11	1490	24805	16.65
LAE[f]	*thra*	1.00±0.21	1.43±0.27	1.43	23	28	1.22
	thrb	1.00±0.09	8.25±1.03	8.25	2	23	11.50
	col1a2	1.00±0.09	1.07±0.13	1.07	330	784	2.38
	dio2	ND	ND	ND	ND	ND	ND
	klf9	1.00±0.30	4.77±1.03	4.77	209	986	4.72
	timm50	1.00±0.29	0.42±0.12	0.42	59	73	1.24
	otc	1.00±0.13	1.90±0.19	1.90	12022	62495	5.20

[a]From Veldhoen et al., 2014 (n=7-16)
[b]Samples were selected from those previously analysed in Veldhoen et al. (2014) and results are expressed as total counts. See section 3.2 for details.
[c]Median±MAD
[d]Comparison of CAT qPCR to RNA-seq Cronbach's alpha = 0.829
[e]ND=not detected
[f]Comparison of LAE qPCR to RNA-seq Cronbach's alpha = 0.947

Appendix 3 *Rana catesbeiana* immune system qPCR assay primer characteristics and thermocycle conditions

Gene Target	Gene	Primer Name	Primer Sequence	Annealing Temp.[a]	Amplicon (bp)	ΔΔCt Criterion[b]
B-cell CLL/ lymphoma 6	*bcl6*	CIA13 up	GGCAATGTTACCTCCAGATGT	64	474	0.08
		CIA13 dn	GATGTCCGCTCCTCCTCTT			
Complement component 3	*c3*	CIA4 up	GACCATAGTCAAAGCCATCA	64	117	0.06
		CIA4 dn	TACTCCCGTCCTTTCTGC			
Complement factor H-related 5	*cfhr5*	CIA10 up	CCGAGATGTTATGCTGGAC	60	264	0.03
		CIA10 dn	AGGGCTCACTGTTAGTGTTGA			
Mannose-binding lectin (protein C) 2	*mbl2*	CIA6 up	CTTGAGCCAGATGGTACTTACA	64	194	0.04
		CIA6 dn	TCATTATTTACCCGCTAGACAG			

[a]Thermocycle program = 40 cycles of elongation (15 sec), annealing (30 sec), and denaturation (30 sec).
[b]The absolute value of the slope of the difference in Ct (target - rpl8) versus log2 cDNA dilution. A value of ≤ 0.15 supports the use of the comparative Ct method (Veldhoen et al., 2014).

Appendix 4 Comparison of *Rana catesbeiana* back skin reference transcriptome RNA-seq with previous qPCR results.

	qPCR[a]			RNA-seq[b]		
Gene transcript	Control[c]	T_3[c]	Fold Change	Control[c]	T_3[c]	Fold Change
thra	1.00±0.24	1.34±0.48	1.34	5013±673	4496±569	0.90
thrb	1.00±0.29	6.44±1.80	6.44	2558±479	8931±339	3.49
col1a2	1.00±0.40	2.35±1.40	2.35	181053±67507	436542±111757	2.41
dio2	1.00±0.38	0.86±0.14	0.86	1441±168	904±293	0.63
klf9	1.00±0.84	1.25±0.99	1.25	3713±2276	5418±1231	1.46
timm50	1.00±0.27	5.17±1.24	5.17	590±50	2327±34	3.94

[a]From Veldhoen et al., 2014 (n=7-16)
[b]Results expressed as total counts. See Methods.
[c]Median±MAD
[d]Comparison of qPCR to RNA-seq Cronbach's alpha = 0.894

Appendix 5 Sequences and information of qPCR primers used in chapter 4.

Gene	Primer Name	Primer Sequence (5`- 3`)	Amplicon Size (bp)	Annealing Temperature (˚C)	Amount per reaction (pmol)
snrpa	150110	TCCCAGAAGAGACAAACGAG	211	64	5
	150111	GCAGGCTACTTTTTGGCAA			5
nsun5	150112	TTGTTGACTTCTATGGGTGG	245	64	5
	150113	TTATGGAAATCCGTGTTCG			5
rrp8	150114	TGACTCTGCGTTCCCGTAT	254	64	5
	150115	AGCATCACCACAGCCAAA			5
suv91	150116	AAATGCGGATTACTACTG	248	60	5
	150117	CTCCAAATGAGTTAGGGT			5

Appendix 6 Comparison of back skin transcriptome RNA-seq with qPCR results for select RNA-processing targets.

	qPCR[a]			RNA-seq[b]		
Gene transcript	Control[c]	T_3[c]	Fold Change	Control[c]	T_3[c]	Fold Change
snrpa	1.00±0.37	3.14±1.50	3.14	1.00±0.02	8.82±.21	8.82
nsun5	1.00±0.20	2.79±0.91	2.79	1.00±0.03	5.93±0.20	5.93
rrp8	1.00±0.45	4.02±0.98	4.02	1.00±0.04	5.06±0.09	5.06
suv91	1.00±0.44	1.60±1.19	1.60	1.00±0.00	8.67±0.00	8.67

[a]Samples from Veldhoen et al., 2014 (n=8-9)
[b]Results expressed as control-normalized median fold change
[c]Median±MAD

Appendix 7 RNA-seq read libraries used to construct BART. Single-stranded 75 bp reads were generated on an Illumina HiSeq 2500 sequencing platform and assembled using Trans-ABySS version 1.5.0 and k = 42.

Tissue	Hormone Condition	Exposure temperature (°C)	Final temperature (°C)	#RNA-seq read pairs (10^6)
Back skin	Control	5	5	139
Back skin	Control	5	24	90
Back skin	T_3	5	5	121
Back skin	T_3	5	24	136
Tail fin	Control	5	5	96
Tail fin	Control	5	24	101
Tail fin	T_3	5	5	193
Tail fin	T_3	5	24	122
Lung	Control	5	5	108
Lung	Control	5	24	114
Lung	T_3	5	5	125
Lung	T_3	5	24	115
Brain	Control	5	5	110
Brain	Control	5	24	100
Brain	Control	24	24	98
Brain	T_3	5	5	116
Brain	T_3	5	24	101
Brain	T_3	24	24	126

Appendix 8 General schemes for creating an annotation resource for transcriptomes from *de novo* assembled sequences. Panel A: serial scheme. Panel B: parallel scheme. Boxed A-E = individual libraries of assembled contigs; Encircled A^1-E^1 = intermediate library of contigs with reduced sequence redundancy that results from self-alignment of individual libraries and retention of the longest member of each sequence-related group; open circles = intermediate library of contigs with reduced sequence redundancy that results from alignment of two other intermediate libraries; BART = bullfrog annotation resource for the transcriptome, a set of R. catesbeiana contigs with reduced sequence redundancy from multiple input libraries.