ABSTRACT

The *Pomacea* apple snails are amphibious operculate freshwater gastropods. They show a circum-tropical and subtropical distribution and exhibit different stress tolerant abilities as some of them are invasive (e.g., *Pomacea maculata* and *Pomacea canaliculata*) and others are non-invasive species (e.g., *Pomacea diffusa*). With the development of proteomic and transcriptomic techniques and their applications in the environmental studies on non-model organisms, I was able to explore the molecular mechanisms underlying species invasion and adaptive evolution in large scale. By using a comparative approach, I have profiled the perivitelline fluid of *P. maculata*, compared the physiological and proteomic responses of *P. canaliculata* and *P. diffusa* to thermal and hypoxic stressors, and examined their sequence divergences.

*Pomacea maculata* adopts aerial oviposition. Proteins inside the egg perivitelline fluid (PVF) are presumably important for the embryonic development, but little is known about their identities. Using SDS-PAGE coupled LC-MS/MS, I identified 74 proteins from the PVF of *P. maculata*, among which 32 are novel. This collection of egg PVF proteins is the largest among those of gastropods. The annotated proteins were classified into nine functional groups: three major perivitelline subunits, immune response, energy metabolism, protein degradation, oxidation-reduction, signaling and binding, transcription and translation, cytoskeleton and others. Comparison of the gene expression levels between albumen gland (the organ that secretes PVF) and other organs showed that 24 PVF genes were specifically expressed in albumen gland. Base substitution analysis of PVF and housekeeping orthologs between *P. maculata* and its closely related species *P. canaliculata* showed that the PVF genes have a higher mean nonsynonymous substitution rate (Ka), synonymous substitution rate (Ks), and Ka/Ks, indicating these reproductive genes have weaker selective constrains. Phylogenetic analysis of perivitellin subunits suggested an ancient gene duplication during the evolution of *Pomacea*.

Tolerance to extreme environmental conditions is a trait shared by many invasive species, but little is known about the molecular mechanisms behind such tolerance. I therefore compared the heat tolerance, proteomic response to heat stress, and adaptive sequence divergence in the invasive snail *P. canaliculata* and its non-invasive congener *P. diffusa*. The LT50 of *P. canaliculata* was significantly higher than that of *P. diffusa*. More than 3,350 proteins were identified from the hepatopancreas of the snails exposed to acute and chronic thermal stress using iTRAQ-coupled mass spectrometry. Acute exposure (3-h exposure at 37 °C with 25 °C as control) resulted in similar numbers (27 in *P. canaliculata* and 23 in *P. diffusa*) of differentially expressed proteins in the two species. Chronic exposure (3-wk exposure at 35 °C with 25 °C as control) caused differential expression of more proteins (58 in *P. canaliculata* and 118 in *P. diffusa*), with many of them related to restoration of
damaged molecules, ubiquitinating dysfunctional molecules, and utilization of energy reserves in both species; but only in *P. diffusa* there was a shift from carbohydrate to lipid catabolism. Analysis of orthologous genes encoding the differentially expressed proteins revealed two genes having clear evidence of positive selection (Ka/Ks > 1), and seven candidates for more detailed analysis of positive selection (Ka/Ks between 0.5 and 1). These nine genes are related to energy metabolism, cellular oxidative homeostasis, signaling and binding processes.

I also compared the responses of two congeners (*P. canaliculata* and *P. diffusa*) to hypoxia exposure by determining their mortality, analyzing their proteome and conducting an evolutionary analysis of genes encoding proteins that were differently expressed between the two species and between different hypoxia treatments. A 72 h experiment was conducted, which showed that *P. canaliculata* is much more tolerant to hypoxia than *P. diffusa*. The two species were then exposed to four levels of dissolved oxygen (6.7, 4.1, 2.0 and 1.0 mg O$_2$ L$^{-1}$) for 8 h, and their gill proteins were analyzed using iTRAQ-coupled LC-MS/MS. There were striking differences in protein expression profiles between the two species. Compared with *P. diffusa*, the more hypoxia tolerant *P. canaliculata* had more up-regulated signal transduction proteins and down-regulated proteins which are involved in glycolysis and the tricarboxylic acid cycle. Evolutionary analysis revealed three orthologous genes encoding the differentially expressed proteins having clear signal of positive selection, indicating selection has acted on some of the hypoxia responsive genes.

Overall, my study has revealed the considerable investment of *P. maculata* to ensure the development of its offspring in aerially deposited eggs. The proteomic and base substitution rate analyses indicate genetic basis of differential resistance to heat and hypoxia stress between *P. canaliculata* and *P. diffusa*. My combined proteomics and comparative evolutionary analysis provides a framework for studying the genetic basis of species invasion and predicting their further expansion in a changing climate for non-model species.
# Table of contents

## Chapter 1. General introduction

1. Apple snails: the family Ampullariidae

2. Biology of *Pomacea*
   - 1.2.1 Habitats and distribution
   - 1.2.2 Species invasion in *Pomacea*
   - 1.2.3 Physiological characteristics of *Pomacea*: amphibiousness and aerial oviposition

3. Perivitelline fluid in *Pomacea*
   - 1.3.1 Origin of perivitelline fluid in *Pomacea*
   - 1.3.2 *Pomacea maculata*

4. Stress resistance in *Pomacea*
   - 1.4.1 *Pomacea canaliculata*
   - 1.4.2 *Pomacea diffusa*
   - 1.4.3 Comparative study on stress tolerance

5. Overview of thesis

## Chapter 2 Identification and evolutionary analysis of egg perivitelline fluid proteins

2.1 Introduction

2.2 Materials and methods
   - 2.2.1 Snail culture
   - 2.2.2 Transcriptome sequencing, assembly and annotation
   - 2.2.3 Egg mass collection and protein extraction
   - 2.2.4 SDS-PAGE and LC–MS/MS Analysis
   - 2.2.5 Database search
   - 2.2.6 Transcriptomic expression analysis
   - 2.2.7 Bioinformatic analysis
   - 2.2.8 Evolutionary analysis

2.3 Results
Chapter 3 Genetic basis of differential heat resistance between two congeneric freshwater snails

3.1 Introduction

3.2 Materials and methods

3.2.1 Animal collection and acclimation

3.2.2 Tolerance to acute thermal exposure

3.2.3 Protein expression

3.2.4 Base substitution rate analysis

3.3 Results

3.3.1 Tolerance to acute thermal exposure

3.3.2 Proteome analysis

3.3.3 Identification of orthologous proteins, and estimating Ka/Ks of differentially expressed orthologs

3.3.4 Pathway enrichment analysis

3.4 Discussion

3.4.1 Molecular chaperones
3.4.2 Oxidative stress related proteins........................................69
3.4.3 Energy metabolism related proteins....................................70
3.4.4 Immune responsive proteins...........................................72
3.4.5 Comparison of the orthologous proteins between congener: challenges and insights......................................................................73

Chapter 4 Molecular mechanisms of tolerance to hypoxia stress in congeneric snails..........................................................82
4.1 Introduction.............................................................................82
4.2 Material and methods.............................................................86
  4.2.1 Snail maintenance and hypoxia exposure system.................86
  4.2.2 Experiment 1. Differential mortality under hypoxia..............87
  4.2.3 Experiment 2. Responses of the proteome to acute hypoxia exposure..............................................................................88
    4.2.3.1 Exposure to hypoxia, protein extraction and labelling......88
    4.2.3.2 Fractionation and LC-MS/MS analysis........................89
    4.2.3.3 Protein identification.............................................90
    4.2.3.4 Determination of differentially expressed proteins.........91
    4.2.3.5 Functional classification.........................................92
    4.2.3.6 Evolutionary analysis of orthologous protein-coding genes93
4.3 Results.....................................................................................94
  4.3.1 Differential tolerance to hypoxia.......................................94
  4.3.2 Proteome analysis.............................................................94
  4.3.3 Estimation of Ka/Ks........................................................97
4.4 Discussion..............................................................................97
  4.4.1 Hypoxia sensing and signal transduction............................99
  4.4.2 Protein synthesis arrest..................................................100
  4.4.3 Energy metabolism.......................................................101
  4.4.4 Effects of different oxygen concentrations......................102
  4.4.5 Evolutionary analysis of orthologs....................................104

Chapter 5 General conclusions and perspectives........................114
Literature cited

Appendixes

Appendix 1. A python script used for filtering unpaired scans and separating data generated by the CID and HCD scans in mass spectrometry

Appendix 2. Permission of reusing published work of Journal of Proteome Research in author’s thesis

CURRICULUM VITAE