Fall 12-12-2018

PHYSELLA ACUTA, COMPARATIVE IMMUNOLOGY AND EVOLUTIONARY ASPECTS OF GASTROPOD IMMUNE FUNCTION

Jonathan H. Schultz
University of New Mexico - Main Campus

Jonathan H. Schultz
candidate

Biology
Department

This dissertation is approved, and is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Dr. Coenraad M. Adema, Chairperson

Dr. Eric S. Loker

Dr. Irene Salinas

Dr. Patrick Hanington

PHYSELLA ACUTA, COMPARATIVE IMMUNOLOGY AND EVOLUTIONARY ASPECTS OF GASTROPOD IMMUNE FUNCTION

By

Jonathan H. Schultz

B.A., Biology, State University of New York at Geneseo, 2012

DISSERTATION

Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy

Ph.D. in Biology

Center for Evolutionary and Theoretical Immunology (CETI)
The University of New Mexico
Albuquerque, New Mexico

May, 2019

DEDICATION

For my nephews Jackson Jonathan and Ian James. The next generation of critical thinkers.

ACKNOWLEDGMENTS

I'd like to especially thank my advisor, colleague, and friend Dr. Coen M. Adema. I knew very little about the scientific endeavor entering graduate school and I thank you for accepting me into your program and elevating my career in science. I look forward to continuing our impassioned discussions of comparative immunobiology.

Thank you to my committee members Dr. Irene Salinas, Dr. Eric S. Loker, and Dr. Patrick Hanington. Irene, thank you for your continued support and guidance, especially the push to attend NACI in North Carolina, which set the course of direction for my postdoctoral studies. Sam, thank you for opening my eyes to the world of evolutionary immunology during my very first semester at UNM and thank you for the support from CETI throughout the years. Patrick, thank you for the enlightening discussions on snail immunity and for supporting me through the end of my PhD.

Thank you to Dr. Susan Bandoni-Muench, my first mentor in science. The drive to pursue a career as a scientist started because of taking your parasitology course at SUNY Geneseo and led me on my path to UNM. I am forever grateful.

Thank you to my lab and office mates near and far: Dr. Janeth Peña, Lijun Lu, Journey Nolan, and Erin Watson. We shared some of the highs and lows of graduate school together and there isn't a better group of people to spend the time with. I wish you all continued success and happiness.

The utmost gratitude to my closest friends Arthur Tackman, Ben Wheaton, Bishoy Kamel, Danielle Griegos, Erica and Dan Ebbs, and Nate Holscher. Thank you for all of the support and friendship that will last beyond my years in Albuquerque. A very special shout out to my roommate and confidant Joshua Lynn. I couldn't ask for a better friend.

Love and thanks to my family. My brother Justin, Nikki, and JJ, thank you for everything, especially the nightly vid chats. There just aren't enough words. To my Mom, Dad, and Aunt Ronni, for your continued support and for travelling all those miles to see me defend.

And lastly, thanks to my ride or die PhD pup and travel buddy, Bo (a.k.a. THE BEANZ), seeing me through it all.

PHYSELLA ACUTA, COMPARATIVE IMMUNOLOGY AND EVOLUTIONARY ASPECTS OF GASTROPOD IMMUNE FUNCTION

By

Jonathan H. Schultz

B.A., Biology, State University of New York at Geneseo, 2012

Ph.D., Biology, University of New Mexico, 2019

ABSTRACT

Gastropod immunobiology has benefitted from investigations focused on the planorbid snail *Biomphalaria glabrata*, intermediate host for the human parasite *Schistosoma mansoni*. Though such concentrated efforts have elucidated fascinating aspects of invertebrate immunity, they have not provided full knowledge regarding the evolution of immune function among other gastropod species. This dissertation presents the importance of making strategic choices regarding which organisms to select for comparative immunology. Herein, the choice was made to investigate the immunobiology of *Physella acuta*, a freshwater snail species of the Physidae, a sister family to Planorbidae to which *B. glabrata* belongs. Benefiting greatly from the use of next-generation sequencing (NGS), the immunobiology of *P. acuta* was studied using 454 pyrosequencing, Illumina RNA-seq, experimental infections with *Echinostoma paraensei* (trematode parasite), and other molecular techniques. These analyses revealed that many components of gastropod immunity have been conserved among physid and planorbid snails. Also, *P. acuta* displays differences in immune function, such as the use of fibrinogen-related proteins in response to trematode parasite exposure. Remarkably, *P. acuta* differentially expressed relatively large immune-relevant gene families (CD109/TEP, dermatopin, GTPase IMAP, among others) after exposure to *E. paraenesi*. Inspection of the individual members that represent these gene families

demonstrated complex transcriptional profiles that suggest parasite influence on host immune function and the capacity of a host to maintain homeostasis while supporting parasite development, an extended phenotype of *E. paraenesi*. These lab-based studies represent the first large scale characterizations of *P. acuta* immune function.

The immune factors described through NGS approaches enable investigations of the ecoimmunology of *P. acuta* snails collected from the field. This approach uncovered many sequences that are differentially expressed by *P. acuta* naturally in the field relative to the lab environment. There is variation in the expression of certain antimicrobial factors and genes governing biological processes. Overall, this dissertation has expanded the scope of gastropod immunity and provides resources and insights that are accessible for continued development and understanding of evolutionary and comparative immunology concepts.

TABLE OF CONTENTS

CHAPTER 1: ……..……1

INTRODUCTION……1

COMPARATIVE IMMUNOLOGY……1

GASTROPOD IMMUNOLOGY…….3

PHYSELLA ACUTA……..5

SIGNIFICANCE……...6

REFERENCES……...10

CHAPTER 2: ……..……15

COMPARATIVE IMMUNOGENOMICS OF MOLLUSCS……..15

ABSTRACT……..16

INTRODUCTION……17

MOLLUSCAN IMMUNITY……..18

BEYOND "FACTOROLOGY", IMMUNOGENOMICS PHASE 1……...22

GENOMES AND NEXT-GENERATION SEQUENCING, IMMUNOGENOMICS PHASE 2…….28

LIMITATIONS AND POTENTIAL, FUTURE IMMUNOGENOMICS……...34

ACKNOWLEDGEMENTS……38

FIGURES AND TABLES……..39

REFERENCES…….42

CHAPTER 3 …….63

COMPARATIVE IMMUNOLOGICAL STUDY OF THE SNAIL *Physella acuta* (HYGROPHILA,PULMONATA) REVEALS SHARED AND UNIQUE ASPECTS OF GASTROPOD IMMUNOBIOLOGY……..…….....63

ABSTRACT……..64

INTRODUCTION……65

MATERIALS AND METHODS……...69

RESULTS……..77

DISCUSSION……...82

ACKNOWLEDGEMENTS……90

APPENDIX A……...90

FIGURES AND TABLES……..91

REFERENCES…….97

CHAPTER 4……...107

CHARACTERIZATION OF THE EARLY HOST RESPONSE OF *Physella acuta* TO THE DIGENETIC TREMATODE *Echinostoma paraensei*…….107

ABSTRACT……108

INTRODUCTION……...…109

MATERIALS AND METHODS…….112

RESULTS……115

DISCUSSION…….119

ACKNOWLEDGEMENTS……...…127

FIGURES AND TABLES……128

REFERENCES……...134

CHAPTER 5: ……145

FROM LAB TO FIELD: INITIAL INVESTIGATION OF THE ECOIMMUNOLOGY OF *Physella acuta* SNAILS……...145

ABSTRACT……146

INTRODUCTION……...…147

MATERIALS AND METHODS ……149

RESULTS……...…..152

DISCUSSION…….155

ACKNOWLEDGEMENTS……...…159

FIGURES AND TABLES……...160

REFERENCES……...171

CHAPTER 6:……175

DISCUSSION…….175

SUMMARY AND CONCLUSIONS……...175

REFERENCES……...180

Chapter 1: Introduction

Comparative Immunology

The origin of comparative immunology as a field of science dates back to the first investigations of phagocytosis in starfish larvae performed by Élie Metchnikoff in the late 1800's (Metchnikoff, 1905). The field has since developed into a formal discipline, exemplified by the establishment of the International Society of Developmental and Comparative Immunology (ISDCI) in 1976. A cursory glance at publications within the namesake journal Developmental and Comparative Immunology (DCI) reveals the broad range of animal diversity studied for immunity. This diversity, in scale and scope, includes the use of non-traditional non-vertebrate model organisms (arthropods, gastropods, ascidians, etc.) and a myriad of cellular and molecular approaches to investigate animal immunobiology. To aptly portray the breadth and diversity of scientific questions and pursuits within the field of comparative immunology would require a life-long effort exceeding the duration of my PhD graduate studies. More important, though, are the reasons such a diversity of scientific pursuits exists within the field and why investigations of comparative immunology should, and will endure.

First and foremost, invoking a comparative framework to investigate immunity forces us to consider the effect that evolution has had in developing, shaping, and selecting metazoan immune mechanisms. By doing so, we can determine which aspects of immunity have been conserved over time and which appear to be taxon-specific novelties of immune function. As a result, we begin to question our preconceived notions of immunity from model (mammalian) organisms and explore alternative ways for how animals have achieved immunoprotection. This notion, in light of the fact that vertebrate organisms constitute only 1% of the living animal diversity of deuterostomes, protostomes, and pre-bilaterians, makes a compelling case for

investigating the immunobiology of non-traditional model organisms (Gourbal et al., 2018). A hallmark of the scientific process is to reject dogmatic perspectives and that should also include perspectives of immunity.

Second, investigation of animal immunity across phylogeny is essential to grasp the impacts that modern anthropogenic phenomena, such as globalization and global climate change, will have on important economic (food) sources and agents of disease transmission (Kim et al., 2018; Loker and Bayne, 2018; Stensgaard et al., 2018). These global changes are capable of compromising immune function, affecting the geographic distribution and abundance, and perhaps leading to the extinction, of essential animal taxa. Investigations of immunity from diverse taxa can help inform policy to better preserve and even bolster our valuable biological resources at a global level. These considerations demonstrate the potential benefit of ecogenomic approaches to investigating invertebrate immunity, which require ecological and immunological viewpoints of investigators.

Lastly, comparative immunology has already yielded impactful discoveries of immunity that traverse animal diversity. These include, but are not limited to, discoveries of Toll receptors in *Drosophila* (Lemaitre et al., 1996), RNA-interference mechanisms in *Caenorhabditis elegans* (Fire et al., 1998), variable lymphocyte receptors (VLRs) in lampreys (Pancer et al., 2004), MHC-like allorecognition in the tunicate *Botryllus schlosseri* (De Tomaso et al., 2005), and somatic diversification of anti-parasite response genes in the snail *Biomphalaria glabrata* (Zhang et al., 2004). The discovery of Toll receptors and MHC-like recognition in invertebrates opened the door for recognizing that dynamic and competent immune capabilities evolved first among these animals (invertebrates) and have been conserved for millions of years and occur throughout phylogeny. RNA-interference (RNAi) was, and continues to be, a cornerstone of experiments for

functional characterization of (immune-relevant) genes of interest. VLRs demonstrate how convergent evolution has led metazoans to develop alternative mechanisms to generate diverse immune receptors. The discovery of diversified anti-parasite genes in *B. glabrata* challenges the notion that invertebrates are incapable of somatically diversifying immune genes, an outdated concept derived from vertebrate-centered perspectives of lymphocyte-based immune systems. In total, these discoveries are epitomized by the concept expressed by Klein that, regarding immune function, "it might be best to forget what we have learned from vertebrates and start a fresh search, unbiased by preconceived ideas" (Klein, 1989). This concept provided me a personal framework to investigate snail immunobiology and is the basis for my dissertation, which I discuss next.

Gastropod Immunology

Gastropods belong to the phylum Mollusca, the second largest phylum of the animal kingdom, and have adopted diverse lifestyles which has contributed to their colonization of terrestrial, marine, and freshwater habitats (Loker, 2010). Interest in gastropod (snail) immunology exists, in large part, because gastropods are hosts for a great diversity of trematode parasites, some of which cause disease in humans (Adema et al., 2012). Also important is the consideration that snails have increasingly become economic resources for food and cosmetics (Draelos, 2017; Loker and Bayne, 2018; Segade et al., 2013). In regard to disease, schistosomiasis, one of the most prevalent neglected tropical diseases (second only to malaria), is mainly transmitted by snails of three genera: *Biomphalaria*, *Bulinus*, and *Oncomelania* (McManus et al., 2018). The distribution, severity, and prevalence of this disease varies and is dependent on the particular flatworm-parasite species with endemic areas mainly determined by the geographic range of snail intermediate hosts. More than 200 million people worldwide have

schistosomiasis and the World Health Organization's (WHO) goal of eliminating schistosomiasis as a health burden by 2025 bears witness to the deleterious impacts of this global disease (https://www.who.int/schistosomiasis/en/). The approaches for eliminating the disease-causing parasites are diverse and include mass drug administration (MDA) of praziquantel, an effective chemotherapy for adult-stage parasites, and snail control efforts (https://www.who.int/schistosomiasis/strategy/en/). The latter includes control of the snail host *B. glabrata*, an intermediate host for *Schistosoma mansoni* (Coustau et al., 2015; Pila et al., 2017). *Biomphalaria glabrata* has been the main subject for investigating snail immunobiology for over half a century. More specifically, studies of the interaction of *B. glabrata* with the trematode parasites *S. mansoni* and *Echinostoma paraensei* have revealed fascinating aspects of gastropod immunity. These include the somatic diversification of immune factors that aid in defense of trematode infection (see Chapter 2) and the development of the concept known as acquired resistance, a phenomenon now considered present among many invertebrate lineages (Gourbal et al., 2018; Lie and Heyneman, 1975). Other discoveries include the existence of a gene-linked resistance locus in *B. glabrata* and functional characterization (via RNAi) of immune-relevant proteins that aid in protection of offspring (Baron et al., 2013; Tennessen et al., 2015). Significantly, over the last 13 years, an international consortium of more than 100 investigators developed and published a full-scale genome of *B. glabrata* that continues to support and extend a robust interest in snail biology, including snail immunity (Adema et al., 2017). This is not an exhaustive list of the efforts and discoveries pertaining to *B. glabrata*. A search on PubMed reveals over 3000 scientific papers since 1947 on *B. glabrata*, and research continues. Therefore, comparative immunology, from a gastropod point-of-view, is mainly informed by investigations of one snail species. However, next-generation sequencing (NGS)

technology is beginning to aid broader investigations of gastropod immunity (see Chapter 2), including the immune systems of the gastropod species *Littorina littorea* (Gorbushin, 2016, 2018; Gorbushin and Borisova, 2015), *Haliotis* spp. (Nam et al., 2016; Priyathilaka et al., 2018; Zhang et al., 2018), and *Oncomelania hupenis* (Zhi-Qiang, et al., 2017). I would be remiss not to mention the snail *Physella acuta*, the main subject of research for my dissertation, which I have developed as a new model for investigating snail immunobiology.

Physella acuta

Physella acuta (Draparnaud, 1805) is a freshwater snail of the family Physidae, a major group of the clade Hygrophila (Dayrat et al., 2011). Like many gastropod species, *P. acuta* is a simultaneous hermaphrodite, but preferentially outcrosses (Escobar et al., 2008; Tsitrone et al., 2003). The unique reproductive behavior of *P. acuta* has spurred investigations of mating behavior and the success of this snail as an invasive species. Also of interest are the population dynamics of *P. acuta*, in connection with *P. acuta's* impressive geographic range, global distribution and success (Ebbs et al., 2018; Lawton et al., 2018). The multitude of common names for *P. acuta*, including "sewage snail", adequately capture the hardiness of this species and spotlight the use of *P. acuta* in many ecotoxicology experiments (Gao et al., 2017; Guo and He, 2014; Martínez-Paz et al., 2017; Morales et al., 2018). Despite these attributes, *P. acuta* has been mostly overlooked as a model for investigating snail immunobiology. As a result, very little is known about the immune capabilities of *P. acuta* or other physid snails in general.

There are many reasons why *P. acuta* was chosen as an additional model for gastropod immunobiology, which are detailed throughout this dissertation. Perhaps the primary motivation is the phylogenetic placement of *P. acuta*. As a member of the family Physidae, sister to the Planorbidae, which houses *B. glabrata*, *P. acuta* can reveal shared and unique differences of

immune capabilities between these closely related snail taxa. As such, the development of *P. acuta* as a lab model for comparative immunology began with a mitogenome project of locally collected (New Mexico) *P. acuta* snails. Nolan et al. (2014) at the University of New Mexico characterized two main isolates (A & B) of *P. acuta* snails, each with uniquely different mitochondrial genomes (>10% nucleotide difference overall). In addition to this intra-species variability, it was also shown that *P. acuta* has an atypical mitochondrial gene order relative to other gastropod species (Nolan et al., 2014). The functional implications of these differences have yet to be determined, albeit, they do demonstrate the propensity for organisms to display taxon-specific differences in biology. Initial efforts to apply NGS to document pathogen-specific responses of *P. acuta* yielded sequence data that helped start my graduate work. I explored the concept that differences of immune capabilities exist between closely related organisms and investigated snail immunobiology, afforded by NGS technology and the local availability of *P. acuta* snails.

Significance

The body of work presented below is categorized into four main chapters. Chapter two is a review, published as part of a special edition of DCI which highlighted the impact of NGS technology on comparative immunogenomics. This chapter provides a historical framework of gastropod immunology and how high throughput sequencing affords us the ability to investigate whole-organismal immunity of non-model organisms and sets the stage for utilizing *P. acuta* in comparative immunological contexts. In this light, we have chosen *P. acuta* to a strategic vantage point to investigate gastropod immunology; investigating the family Physidae, a neighboring family of the Planorbidae (*B. glabrata*), expands the scope of gastropod

immunobiology and helps interpret the evolution of immunity between these two important taxonomic groups.

Chapter three is an investigation of *P. acuta* immune capabilities using 454 pyrosequencing to record transcriptomes of two individual *P. acuta* snails; one snail after septic exposure and one snail un-exposed to pathogens. The main goal of this analysis was to capture and characterize a large portion of the immune repertoire expressed by *P. acuta* and compare this to the immune capabilities of *B. glabrata* that have been documented previously. As expected, *P. acuta* shares many of the same immune features with *B. glabrata*, including a modest catalogue of antimicrobial peptide sequences balanced with an extensive set of diverse antimicrobial protein genes. Also discussed is the relevance of fibrinogen-related domain-containing sequences, which are prominent features of bivalve immunity, and hypothesized to be important for *B. glabrata* and *P. acuta* immunity (see Chapters three, four, and five). Chapter two also presents a compelling finding that, in contrast to *B. glabrata*, *P. acuta* lacks an extensive repertoire of FREP genes and does not somatically diversify these sequences. This finding demonstrates the importance of broadening the scope of gastropod immunology by investigating organisms of different taxa. Lastly, we report for the first time, experimental use of the host-parasite interaction of *P. acuta* and *E. paraensei*. Exposing *P. acuta* to *E. paraensei* helped to confirm that FREPs are not prominent anti-parasite immune components for this snail. Moreover, this host-parasite model has important potential for study of gastropod immunobiology because *E. paraensei* (and related parasites) dramatically interferes with host biology and helped elucidate novel immune mechanisms in *B. glabrata* (Adema et al., 2000; Hanington et al., 2010, 2012; Lie and Heyneman, 1977; Lie et al., 1975, 1976).

Chapter four extends the characterization of the host-parasite interaction of *P. acuta* and *E. paraensei* using classical histological techniques and contemporary Illumina RNA sequencing. The main objective of this analysis was to characterize the early interaction of snail and host (2-8 days post exposure), a time interval that includes the initial battle between host and parasite, leading to immune elimination or establishment of infection, and the extended phenotype displayed by the parasite after takeover of the host. I used whole-organismal histology to demonstrate the development of *E. paraensei* within *P. acuta* and differential expression analyses of RNA sequences to provide insights into the transcriptional profiles of *P. acuta* in response to trematode parasite exposure and infection. A more in-depth analysis of immune-relevant transcripts that are differentially expressed after exposure to *E. paraensei* was performed to catalog putative immune genes and investigate the potential for this parasite to influence host immune function. This investigation indicates that individual members of large families of immune factors of *P. acuta* can have unique expression patterns in response to *E. paraensei*, similar to observations made from *B. glabrata* and *Crassostrea gigas* (Deleury et al., 2012; Zhang et al., 2015).

The fifth chapter aimed to open up the potential for investigating the ecoimmunology of snails by connecting molecular immunology with the immune status of snails in their natural habitats. Physid snails were collected from Shady Lakes in Albuquerque, NM with the goal of comparing transcriptional profiles of lab-maintained and field-derived physid snails. The primary motivation of this work was to determine if field-collected snails display obvious differences of expressed sequences relative to their lab counterparts in order to determine if laboratory models reflect the same biological status. Snails caught from the field were screened for the absence of patent trematode infections (i.e. did not shed cercariae) and transcriptomes of four field-collected

snails and seven lab-collected snails were assembled. As next-generation sequencing technology is likely to become more affordable and increase in use, this analysis highlights the feasibility of using molecular ecoimmunological approaches to investigate immunity (Hawley and Altizer, 2011; Van Straalen and Roelofs, 2010). Overall, an ecoimmunology approach can facilitate a better understanding of the processes preventing disease transmission and merge immunological and ecological questions.

This dissertation demonstrates how traditional and current molecular techniques can facilitate investigating questions of invertebrate immunology. The work presented here is accessible to all investigators interested in evolutionary immunology by providing avenues for addressing ongoing questions of immunity. The NGS datasets also provide a potential resource for investigators of invertebrates within other fields of comparative genomics.

I encourage those interested in immunological questions to consider the incredible diversity of metazoans and expand the scope of investigation beyond traditional model organisms, similar to how I have applied *Physella acuta* for developing research to reveal new aspects in comparative immunology.

References

1. Adema, C.M., Sapp, K.K., Hertel L.A., Loker E.S., 2000. Immunobiology of the Relationship of Echinostomes with Snail Intermediate Hosts. in: Fried, B., Graczyk, T.K. (eds), Echinostomes as Experimental Models for Biological Research. Springer, Dordrecht.

2. Adema, C.M., Bayne, C.J., Bridger, J.M., Knight, M., Loker, E.S., Yoshino, T.P., Zhang, S.M., 2012. Will all scientists working on snails and the diseases they transmit please stand up? PLoS Negl. Trop. Dis. 6, e1835.

3. Adema, C.M., Hillier, L.W., Jones, C.S., Loker, E.S., Knight, M., Minx, P., Oliveira, G., Raghavan, N., Shedlock, A., do Amaral, L.R., Arican-Goktas, H.D., Assis, J.G., Baba, E.H., Baron, O.L., Bayne, C.J., Bickham-Wright, U., Biggar, K.K., Blouin, M., Bonning, B.C., Botka, C., Bridger, J.M., Buckley, K.M., Buddenborg, S.K., Lima, Caldeira, R., Carleton, J., Carvalho, O.S., Castillo, M.G., Chalmers, I.W., Christensens, M., Clifton, S., Cosseau, C., Coustau, C., Cripps, R.M., Cuesta-Astroz, Y., Cummins, S.F., di Stephano, L., Dinguirard, N., Duval, D., Emrich, S., Feschotte, C., Feyereisen, R., FitzGerald, P., Fronick, C., Fulton, L., Galinier, R., Gava, S.G., Geusz, M., Geyer, K.K., Geraldo-Calderón, G.I., de Souza Gomes, M., Gordy, M.A., Gourbal, B., Grunau, C., Hanington, P.C., Hoffmann, K.F., Hughes, D., Humphries, J., Jackson, D.J., Jannotti- Passos, L.K., de, Jesus, Jeremias, W., Jobling, S., Kamel, B., Kapusta, A., Kaur, S., Koene, J.M., Kohn, A.B., Lawson, D., Lawton, S.P., Lìang, D., Limpanont, Y., Liu, S., Lockyer, A.E., Lovasto, T.L., Ludolf, F., Magrini, V., McManus, D.P., Medina, M., Misra, M., Mitta, G., Mkoji, G.M., Montague, M.J., Montelongo, C., Moroz, L.L., Munoz-Torres, M.C., Niazi, U., Noble, L.R., Oliveira, F.S., Pais, F.S., Papenfuss, A.T., Peace, R., Pena, J.J., Pila, E.A., Quelais, T., Raney, B.J., Rast, J.P., Rollinson, D., Rosse, I.C., Rotgans, B., Routledge, E.J., Ryan, K.M., Scholte, L.L.S., Storey, K.B., Swain, M., Tennessen, J.A., Tomlinson, C., Trujillo, D.L., Volpi, E.V., Walker, A.J., Wang, T., Wannaporn, I., Warren, W.C., Wu, X.J., Yoshino, T.P., Yusuf, M., Zhang, S.M., Zhao, M., Wilson, R.K., 2017. Whole genome analysis of a schistosomiasis- transmitting freshwater snail. Nat. Commun. 8, 15451.

4. Baron, O.L., van West, P., Industri, B., Pouchet, M., Dubreuil, G., Gourbal, B., Reichart, J.M., Coustau, C., 2013. Parental transfer of the antimicrobial protein LBP/BPI protects *Biomphalaria glabrata* eggs against oomycete infections. PLoS Pathogens. 9, e1003792.

5. Coustau, C., Gourbal, B., Duval, D., Yoshino, T.P., Adema, C.M., Mitta, G., 2015. Advances in gastropod immunity from the study of the interaction between the snail *Biomphalaria glabrata* and its parasites: a review of research progress over the last decade. Fish Shellfish Immunol. 46, 5-16.

6. Dayrat, B., Conrad, M., Balayan, S., White, T.R., Albrecht, C., Golding, R., Gomes, S.R., Harasewych, M.G., Martins, A.M., 2011. Phylogenetic relationships and evolution of pulmonate gastropods (Mollusca): new insights from increased taxon sampling. Mol. Phylogenet. Evol. 59, 425-437.

7. Deleury, E., Dubreuil, G., Elangovan, N., Wajnberg, E., Reichart, J.M., Gourbal, B., Duval, D., Baron, O.L., Gouzy, J., Coustau, C., 2012. Specific versus non-specific immune responses in an invertebrate species evidenced by a comparative *de novo* sequencing study. PLoS One 7 e32512.

8. De Tomaso, A.W., Nyholm, S.V., Palmeri, K.J., Ishizuka, K.J., Ludington, W.B., Mitchel, K., Weissman, I.L., 2005. Isolation and characterization of a protochordate histocompatibility locus. Nature. 438, 454-459.

9. Draelos, Z.D., 2017. The Role of a Natural Mollusk Egg-Derived Ingredient in Facial Appearance. J. Drugs Dermatol. 16, 678-681.

10. Ebbs, E.T., Loker, E.S., Brant, S.V., 2018. Phylogeography and genetics of the globally invasive snail *Physa acuta* Draparnaud 1805, and its potential to serve as an intermediate host to larval digenetic trematodes. BMC Evol. Biol. 18, 103.

11. Escobar, J.S., Jarne, P., Charmantier, A., David, P., 2008. Outbreeding alleviates senescence in hermaphroditic snails as expected from the mutation-accumulation theory. Curr. Biol. 18, 906-910.

12. Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., Mello, C.C., 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature. 391, 806-811.

13. Gao, L., Doan, H., Nidumolu, B., Kumar, A., Gonzago, D., 2017. Effects of copper on the survival, hatching, and reproduction of a pulmonate snail (*Physa acuta*). Chemosphere. 185, 1208-1216.

14. Gorbushin, A.M., Borisova, E.A., 2015. Lectin-like molecules in transcriptome of *Littorina littorea* hemocytes. Dev. Comp. Immunol. 210-220.

15. Gorbushin, A.M., 2016. Membrane Attack Complex/Perforin domain-containing proteins in a dual-species transcriptome of caenogastropoda *Littorina littorea* and its trematode parasite *Himasthla elongata*. Fish Shellfish Immunol. 54, 254-260.

16. Gourbal, B., Pinaud, S., Beckers, G.J.M., Van Der Meer, J.W.M., Conrath, U., Netea, M.G., 2018. Innate immune memory: An evolutionary perspective. Immunol. Rev. 283, 21-40.

17. Guo, Y., He, H., 2014. Identification and characterization of a goose-type lysozyme from sewage snail *Physa acuta*. Fish Shellfish Immuno. 39, 321-325.

18. Hanington, P.C., Forys, M.A., Dragoo, J.W., Zhang, S.M., Adema, C.M., Loker, E.S., 2010. Role for a somatically diversified lectin in resistance of an invertebrate to parasite infection. PNAS. 107, 21087–21092.

19. Hanington, P.C., Forys, M.A., Loker, E.S., 2012. A somatically diversified defense factor, FREP3, is a determinant of snail resistance to schistosome infection. PLoS Negl. Trop. Dis. 6, e1591.

20. Kim, J.R., Wong, T.M., Curry, P.A., Yeung, N.W., Hayes, K.A., Cowie, R.H., 2018. Modelling the distribution in Hawaii of *Angiostrongylus cantonensis* (rat lungworm) in its gastropod hosts. Parasitology. 21, 1-8.

21. Klein, J., 1989. Are invertebrates capable of anticipatory immune responses? Scand. J. Immunol. 29, 499-505.

22. Lawton, S.P., Allan, F., Hayes, P.M., Smit, N.J., 2018. DNA barcoding of the medically important freshwater snail *Physa acuta* reveals multiple invasion events into Africa. Acta. Trop. 188, 86-92.

23. Lie, K.J., Heyneman, D., 1975. Studies on resistance in snails: a specific tissue reaction to *Echinostoma lindoense* in *Biomphalaria glabrata* snails. Int. J. Parasitol. 5, 621–625.

24. Lie, K.J., Heyneman, D., Lim, H.K., 1975. Studies on resistance in snails: specific resistance induced by irradiated miracidia of *Echinostoma lindoense* in *Biomphalaria glabrata* snails. Int. J. Parasitol. 5, 627-631.

25. Lie, K.J., Heyneman, D., and Jeang, K.H., 1976. Studies on resistance in snails. 7. Evidence of interference with the defense reaction in *Biomphalaria glabrata* by trematode larvae. J. Parasitol. 62, 608-615.

26. Lie, K.J., Heyneman, D., 1977. Studies on resistance in snails: interference by nonirradiated echinostome larvae with natural resistance to *Schistosoma mansoni* in *Biomphalaria glabrata*. J. Invert. Pathol. 29, 118-125.

27. Loker, E.S., 2010. Gastropod immunobiology. Adv. Exp. Med. Biol. 708, 17-43.

28. Loker, E.S., Bayne, C.J., 2018. Molluscan Immunobiology: Challenges in the Anthropocene Epoch, in: Cooper, E.L. (Ed.), Advances in Comparative Immunology. Springer International Publishing AG, Switzerland, pp. 343-407. doi: 10.1016/j.actatropica.2018.09.013.

29. Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J.M., Hoffmann, J.A., 1996. The dorsoventral regulatory gene cassette spätzle/Toll/cactus controls the potent antifungal response in *Drosophila* adults. Cell. 86, 973-983.

30. Martínez-Paz, P., Morales, M., Sánchez-Argüello, P., Morcillo, G., Martínez-Guitarte, J.L., 2017. Cadmium in vivo exposure alters stress response and endocrine-related genes in the freshwater snail *Physa acuta*. New biomarker genes in a new model organism. Environ. Pollut. 220, 1488-1497.

31. McManus, D.P., Dunne, D.W., Sacko, M., Utzinger, J., Vennervald, B.J., Zhou, X.N., 2018. Schistosomiasis. Nat. Rev. Di.s Primers. doi: 10.1038/s41572-018-0013-8.

32. Metchnikoff, E., 1905. Immunity in Infective Disease. Cambridge University Press, Cambridge.

33. Morales, M., Martínez-Paz, P., Sánchez-Argüello, P., Morcillo, G., Martínez-Guitarte, J.L., 2018. Bisphenol A (BPA) modulates the expression of endocrine and stress response genes in the freshwater snail *Physa acuta*. Ecotoxicol. Environ. Saf. 152, 132-138.

34. Nam, B.H., Jung, M., Subramaniyam, S., Yoo, S.I., Markkandan, K., Moon, J.Y., Kim, Y.O., Kim, D.G., An, C.M., Shin, Y., Jung, H.J., Park, J.H., 2016. Transcriptome Analysis Revealed Changes of Multiple Genes Involved in *Haliotis discus hannai* Innate Immunity during *Vibrio parahemolyticus* Infection. PLoS One. 11, e0153474.

35. Nolan, J.R., Bergthorsson, U., Adema, C.M., 2014. *Physella acuta*: atypical mitochondrial gene order among panpulmonates (Gastropoda). J. Mollus. Stud. 80, 388-399.

36. Pancer, Z., Amemiya, C.T., Ehrhardt, G.R., Ceitlin, J., Gartland, G.L., Cooper, M.D., 2004. Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey. Nature. 430, 174-180.

37. Pila, E.A., Li, H., Hambrook, J.R., Wu, X., Hanington, P.C., 2017. Schistosomiasis from a Snail's Perspective: Advances in Snail Immunity. Trends Parasitol. 33, 845-857.

38. Priyathilaka, T.T., Bathige, S.D.N.K., Lee, S., Nam, B.H., Lee, J., 2018. Transcriptome-wide identification, functional characterization, and expression analysis of two novel invertebrate-type Toll-like receptors from disk abalone (*Haliotis discus discus*). Fish Shellfish Immunol. 84, 802-815.

39. Segade, P., García-Estévez, J., Arias, C., Iglesias, R., 2013. Parasitic infections in mixed system-based heliciculture farms: dynamics and key epidemiological factors. Parasitology. 140, 482-497.

40. Stensgaard, A.S., Vounatsou, P., Sengupta, M.E., Utzinger, J., 2018. Schistosomes, snails and climate change: Current trends and future expectations. Acta. Trop. doi:10.1016/j.actatropica.2018.09.013.

41. Tennessen, J.A., Bonner, K.M., Bollmann, S.R., Johnstun, J.A., Yeh, J.Y., Marine, M., Tavalire, H.F., Bayne, C.J., Blouin, M.S., 2015. Genome-wide scan and test of candidate genes in the snail *Biomphalaria glabrata* reveal new locus influencing resistance to *Schistosoma mansoni*. PLoS. Negl. Trop. Dis. 9, e0004077.

42. Tsitrone, A., Jarne, P., David, P., 2003. Delayed selfing and resource reallocations in relation to mate availability in the freshwater snail *Physa acuta*. Am. Nat. 162, 474-488.

43. World Health Organization, Schistosomiasis. https://www.who.int/schistosomiasis/en/ (accessed 21 March 2018).

44. Zhang, S.M., Adema, C.M., Kepler, T.B., Loker, E.S., 2004. Diversification of Ig superfamily genes in an invertebrate. Science 305, 251-254.

45. Zhang, L., Li, L., Guo, X., Litman, G.W., Dishaw, L.J., Zhang, G., 2015. Massive expansion and functional divergence of innate immune genes in a protostome. Sci. Rep. 5, 8693.

46. Zhang, X., Shi, J., Sun, Y., Habib, Y.J., Yang, H., Zhang, Z., Wang, Y., 2018. Integrative transcriptome analysis and discovery of genes involving in immune response of hypoxia/thermal challenges in the small abalone *Haliotis diversicolor*. Fish Shellfish Immunol. 84, 609-626.

47. Zhi-Qiang, Q., Xiao-Jing, M., Wei, G., Shi-Zhu, L., 2017. Study on transcriptome of *Oncomelania hupensis* before and after *Schistosoma japonicum* invasion: *De novo* assembly of data by RNA-Seq. Zhongguo Xue Xi Chong Bing Fang Zhi Za Zhi. 29, 300-304.

Chapter 2: Comparative immunogenomics of molluscs

Schultz, J.H., Adema, C.M., 2017. Comparative immunogenomics of molluscs. Dev. Comp. Immunol. 75, 3-15.

Abstract

Comparative immunology, studying both vertebrates and invertebrates, provided the earliest descriptions of phagocytosis as a general immune mechanism. However, the large scale of animal diversity challenges all-inclusive investigations and the field of immunology has developed by mostly emphasizing study of a few vertebrate species. In addressing the lack of comprehensive understanding of animal immunity, especially that of invertebrates, comparative immunology helps toward management of invertebrates that are food sources, agricultural pests, pathogens, or transmit diseases, and helps interpret the evolution of animal immunity. Initial studies showed that the Mollusca (second largest animal phylum), and invertebrates in general, possess innate defenses but lack the lymphocytic immune system that characterizes vertebrate immunology. Recognizing the reality of both common and taxon-specific immune features, and applying up-to-date cell and molecular research capabilities, in-depth studies of a select number of bivalve and gastropod species continue to reveal novel aspects of molluscan immunity. The genomics era heralded a new stage of comparative immunology; large-scale efforts yielded an initial set of full molluscan genome sequences that is available for analyses of full complements of immune genes and regulatory sequences. Next-generation sequencing (NGS), due to lower cost and effort required, allows individual researchers to generate large sequence datasets for growing numbers of molluscs. RNAseq provides expression profiles that enable discovery of immune genes and genome sequences reveal distribution and diversity of immune factors across molluscan phylogeny. Although computational *de novo* sequence assembly will benefit from continued development and automated annotation may require some experimental validation, NGS is a powerful tool for comparative immunology, especially increasing coverage of the extensive molluscan diversity. To date, immunogenomics revealed new levels of complexity of

molluscan defense by indicating sequence heterogeneity in individual snails and bivalves, and members of expanded immune gene families are expressed differentially to generate pathogen-specific defense responses.

1. Introduction

Historical observations of the association between exposure to disease and subsequent protection from future (human) illness eventually led to development of the smallpox vaccine by Jenner in the late 1700s (Owen et al., 2013), and of the Germ Theory, linking pathogens and disease (Walker et al., 2006). Studies of how immunity against pathogens is achieved culminated in our current understanding of immunology, mostly as it reflects the immune function of vertebrate animals. Some of the early studies of immunity, however, also benefited from use of invertebrate organisms. Most famously, Metchnikoff discovered phagocytes and their role in immunity in starfish larvae (Metchnikoff, 1905). By using one (invertebrate) organism to make predictions of immune function in other (vertebrate) animals Metchnikoff gave rise to a new field of biology: comparative immunology. The power of comparative immunology begotten by investigating invertebrates is evident from landmark characterization of e.g. lectins (Prokop et al., 1968), antimicrobial peptides (Boman and Hultmark, 1987), Toll-like receptors (Lemaitre et al., 1996; Medzhitov et al., 1997), and RNA interference (Fire et al., 1998) that have expanded our understanding of immunity and revealed shared features among animals across phylogeny and evolution.

The study of immune function of molluscs (including snails, bivalves, cephalopods, others, see Fig. 1) is motivated importantly by notions that many molluscs are economically valuable food sources, especially in aquaculture (Carnegie et al., 2016), or may transmit infectious diseases of medical and veterinary relevance (Adema et al., 2012). Moreover, the

highly diverse phylum Mollusca is second in size among animals only to Arthropoda and represents the generally understudied lophotrochozoan protostomes, one of three lineages of metazoan animals, along with ecdysozoan protostomes and deuterostomes (e.g. Erwin et al., 2011). As such, study of molluscs will continue to broaden understanding of the evolution of immune function across the range of metazoan phylogeny, especially because recent insights suggest that molluscs are capable of sophisticated and specific immune responses (Adema and Loker, 2015; Coustau et al., 2015; Guo et al., 2015).

This review briefly discusses the view of molluscan immune capabilities as it developed from investigations before the availability of immunogenomics. We then present the more specific characterization of molluscan immunity that was afforded by studies benefitting from PCR and Sanger sequencing. Following is a discussion of molluscan immune capabilities discovered from genome mining and transcriptome analyses, especially by relatively easily-applied large-scale next-generation sequencing techniques. Lastly, we consider current limitations of utilizing NGS data and discuss the future of molluscan comparative immunology now that large sequence datasets are increasingly available.

2. Molluscan immunity

Historically, molluscan immunology has been studied in a small number of species represented within the diversity of the phylum Mollusca. Practical considerations that included ease of collection, animal size, reliable animal husbandry, selective rearing of genetic lineages, as well as relevance for disease transmission or economical (aquaculture) importance have focused consistent study toward a few species of the Gastropoda, and increasingly so in recent times to some representatives of the classes Bivalvia and Cephalopoda, to the exclusion of other molluscan classes (Fig. 1). Experimental approaches for initial immunological studies included

the monitoring of responses of bivalve and gastropod molluscs following exposure to inorganic material (e.g. Indian ink; Tripp and Kent, 1967), to pathogens, introduced either by bacterial injection or through infection with parasites, notably digenean flatworms like *Schistosoma mansoni* that causes significant infectious disease when transmitted to humans (Tebeje et al., 2016). Snails were observed to rapidly clear bacteria from circulation and survive the exposure, with indications of elevated immunity, a more rapid clearance, after an initial encounter (Bayne, 1980; van der Knaap et al., 1983a, 1981). Some individual snails among populations of otherwise parasite-susceptible *Biomphalaria glabrata* proved naturally resistant to digenetic trematodes, with more rapid responses toward a secondary exposure (Lie and Heyneman, 1979). Susceptibility to parasite infection was determined by the genetic background of snail and parasites (Richards et al., 1992). Professional phagocytic cells termed hemocytes, dwelling in the tissues or circulating with the blood fluid of gastropods and bivalves, phagocytose or encapsulate pathogens, eliminating these with cell- mediated cytotoxicity involving lysosomal enzymes and production of reactive oxygen species (Adema et al., 1991; Granath and Yoshino, 1983; La Peyre et al., 1995; McKerrow et al., 1985; Mohandas et al., 1985; van der Knaap and Loker, 1990). Depending on the species, molluscs may have either a single type or several functionally different categories of hemocytes, and these cells may originate from connective tissue or specialized organs, termed the amoebocyte producing organ (APO) in gastropods (Jeong et al., 1983), or from the white body organ in cephalopods (Claes, 1996; Cowden, 1972). Recognition of nonself and subsequent immune activation is mediated through lectins, initially referred to as agglutinins or cytophilic receptors for foreignness, present as humoral factors or on the surface of hemocytes (Cheng et al., 1984; Michelson and Dubois, 1977; Mullainadhan and Renwrantz, 1986; Renwrantz and Cheng, 1977; Rögener et al., 1985; van der Knaap et al., 1983b). Lectins

are non-enzymatic, non-antibody proteins that function as pattern recognition receptors (PRRs) by binding to repetitive carbohydrate surface determinants that characterize groups of pathogens (pathogen associated molecular patterns, PAMPs) such as lipopolysaccharide (LPS) and peptidoglycans of bacteria (Vasta and Ahmed, 2009) and activate immune responses. Contrary to expectations regarding animal immunity drawn from a vertebrate perspective of immune function, and by the observation of some level of immunological memory in gastropods (Lie and Heyneman, 1979), no indications were found in molluscs, or invertebrates in general, of lymphocytic defenses, i.e. no T-cells, B- cells or the rearranging genes that drive generation of antigen- specific receptors (Warr, 1981). As a consequence, invertebrates were deemed to possess a rather unsophisticated innate-type immunity, with a reliance only on invariable, germline-encoded genes for general broad immune recognition of categories of pathogenic organisms. However, Klein (1989) championed the importance of investigating the immunity of invertebrates from new perspectives that are not myopically biased by norms of vertebrate immunology. While invertebrates may not possess all canonical features of the vertebrate immune system, as a result of a long independent evolutionary development they may bear homologs of immune mechanisms, as well as unique immune features that are specific to their lineage. Through analyses of such immune features, comparative immunology can provide important insights into the evolution of immunity (Marchalonis and Schluter, 1990).

Continued study of molluscan immunology revealed several defense mechanisms, both analogous to aspects of vertebrate immunology and *de novo* from investigation of mollusc-pathogen encounters and the immune factors involved in these interactions. In the absence of specific reagents, comparative studies were often performed with heterologous techniques and reagents, borrowed from vertebrate immunology. The mechanism for production by molluscan

hemocytes of toxic reactive oxygen species toward pathogens (Adema et al., 1994) was proposed to be homologous to that responsible for the cytotoxic respiratory burst of vertebrate phagocytes as both were sensitive to catechol-like phenol inhibitors of the vertebrate NADPH-oxidase enzyme complex that generates superoxide (Adema et al., 1993; Noe€l et al., 1993). Chemical inhibitors and immune-reagents raised against the components of vertebrate signaling pathways were applied to explore and characterize the role of MAPK and other signaling pathways in the regulation of molluscan immunity (Humphries et al., 2001; Walker et al., 2010; Zahoor et al., 2009, 2008; Zelck et al., 2007). A role for vertebrate-type cytokines in immune regulation was proposed, based on altered cell morphology and activity of hemocytes following exposure to (human) cytokines, cross-reactivity of molluscan hemocytes with immune-reagents specific for vertebrate cytokines such as Il-1, Il-6 and TNF, and reduced parasite success after treatment of snails with recombinant human cytokine (Connors et al., 1998, 1995a, 1995b; Granath et al., 1994; Hughes et al., 1992, 1991, 1990; Ottaviani et al., 2000, 1997, 1995a, 1995b, 1993; Ouwe-Missi-Oukem-Boyer et al., 1994; Steelman and Connors, 2009; Stefano et al., 1991). Although the presence of functional cytokine-like factors in molluscs was not unanticipated and heterologous immune reagents may correctly reveal sequence homologs in molluscs (Williams and Gagnon, 1982), the actual identity of such factors remained to be confirmed due to the potential for false positive immune reactivity via the presence of aspecific antibody-binding proteins in some molluscs (Hahn et al., 1996) and the lack of evolutionary sequence conservation among animal cytokines (Beschin et al., 1999).

Study of mollusc-specific immune features, independent of concepts of vertebrate immunology, provided considerable understanding of how lectins function as non-self receptors. Immune activation resulted from a complex interaction between soluble and cell-bound lectin

receptors, enabled by conformational changes after interaction with target epitopes (Cheng and Manzi, 1996; Dam et al., 1992; Renwrantz and Richards, 1992; Richards and Renwrantz, 1991). Along with lectins, hemocytes of *Mytilus edulis* also released cytotoxic factors with lytic activity toward target cells (Leippe and Renwrantz, 1988). The snail *Biomphalaria alexandrina* expresses multimeric hemolymph lectin molecules that are composed of different members of a family of glycan-binding lectins, potentially expanding the range of antigen recognition (Mansour, 1995). Levels of agglutinating activity (lectins) differed in individual Pacific oysters, *Crassostrea gigas* that were either susceptible or resistant to *Perkinsus marinus* parasites (Romestand et al., 2002). The gastropod *B. glabrata* responded to *Echinostoma paraensei* (flatworm) infection with expression of increased levels of diverse hemolymph lectins distributed across several molecular weight ranges (Monroy et al., 1992), referred to as G1M, G2M, and 65 kd lectins (Adema et al., 1997a). These foundational research efforts provided a deeper understanding of molluscan immune capabilities. Due to technical limitations, however, the factors or mechanisms involved could not yet be defined beyond general descriptors like "cross-reactivity with heterologous immune reagents", "sensitivity to inhibitors", "specific carbohydrate binding activity" or "specific molecular weight". Regardless, molluscan immunology was well-positioned to benefit from the more detailed analysis of immune factors and mechanisms afforded by newly available molecular techniques and to develop the potential of comparative immunology for detailed tracking of the evolutionary development of immune function across animal clades.

3. Beyond "factorology", immunogenomics phase 1

The first phase of molluscan immunogenomics began in the early 1990s, benefitting from several technological developments that facilitated routine access to molecular biology (PCR,

cloning and automated dye-terminator Sanger sequencing) and to the internet (computational bioinformatics and rapidly growing public sequence databases) in a manner that made it possible to identify molluscan immune factors through characterization of coding sequences. A strategy that consisted of partial peptide sequencing (Edman degradation) of purified proteins to enable design of PCR primers and amplify cDNA for sequencing led to targeted characterization of the first snail lectins. The elucidation of the encoding sequences promoted these lectins from factors that were broadly defined by molecular weight and carbohydrate specificity to specific, named entities, with function and relationship to other defense factors defined by sequence. The parasite-reactive plasma lectins of *B. glabrata* described by Monroy et al. (1992) were found to include a diverse set of related fibrinogen-related proteins (FREPs), comprising N-terminal immunoglobulin superfamily domains juxtapositioned with a fibrinogen-like domain at the C-terminus (Adema et al., 1997b). The mucus of the slug *Incilaria burgsdorfii* yielded Incilarin A-C, three different C-type lectins that shared considerable sequence identity (44-55% identity of encoded amino acids; Yuasa et al., 1998). Screening of a cDNA library revealed sialic-acid binding activity derived from several lectins from *Limax flavus* (Gastropoda), also highly similar yet with different sequence identities, and each consisting of a single fibrinogen-like domain (Kurachi et al., 1998). Expression of related lectin sequences from multigene families may broaden the range of carbohydrate epitopes that can be bound, also considering that lectins may assemble as multimeric molecules under native conditions (e.g. Adema et al., 1997a; Mansour, 1995). Some of the newly characterized sequences were studied in greater detail. The full-length genome sequences were determined for several of these FREP genes (Leonard et al., 2001; Zhang et al., 2001). Moreover, it was observed that somatic mutation of FREP genes causes *B. glabrata* snails to generate individually unique repertoires of parasite-reactive FREP lectins

(Zhang et al., 2004). Note that FREPs are discussed in detail by Portet et al. elsewhere in this issue. By revealing intronic sequence differences, Goodall et al. (2006) identified allelic variants of *B. glabrata* superoxide dismutase (SOD1) that differentially affected the production of reactive oxygen species in a manner that associated with parasite resistance/susceptibility phenotypes of individual snails.

The above targeted characterization of specific immune factors provided detailed sequence information that began to facilitate comparative immunology, i.e. considering the nature of molluscan defense factors relative to other organisms, and distribution of related factors across phylogeny. The experimental work was laborious, however, and progress came at considerable effort and investment. Many more candidate immune factors were identified at a higher rate of discovery by transcriptomic studies that employed random gene discovery via high throughput sequencing of libraries of expressed sequence tags (ESTs). Short cDNA inserts, usually representing partial gene transcripts, generated by a number of strategies, including differential display reverse transcription PCR (DD RT-PCR), suppression subtractive hybridization (SSH), open reading frame EST (ORESTES), were cloned and sequenced to capture gene expression profiles from snails like *Lymnaea stagnalis* (Hoek et al., 1996) and prominently *B. glabrata*, untreated or exposed to pathogens. Aided by computational bioinformatics, EST projects recorded many transcribed sequences that helped reveal aspects of the immune system of gastropods (see Fig. 1). Based on sequence similarity with previously described defense factors, mostly from other animal taxa, *in silico* analyses identified gastropod genes for all components of an internal defense system ranging from non-self recognition (lectins); activation and regulation of immune responses; humoral defense factors (antibacterial proteins and peptides); cellular cytotoxicity (lysosomal enzymes, production and metabolism of

reactive oxygen species), including antiviral responses as afforded by genes encoding the machinery for RNA interference. It is of note that up to 60% of EST data from molluscs cannot be annotated because they have no matches in (current) sequence databases. Designated as novel, unknown, a proportion of these sequences are likely unique to molluscan biology and await functional characterization. For an incomplete, yet representative listing of associated projects see: Bouchut et al., 2006; Hanelt et al., 2008; Ittiprasert et al., 2013; Knight et al., 1999; Lockyer et al., 2007a, 2007b; Miller et al., 2001; Mitta et al., 2005; Nowak et al., 2004; Raghavan et al., 2003; Schneider and Zelck, 2001.

Since, EST-based gene discovery has been applied widely to a range of additional species of gastropods, bivalves and cephalopods. This has provided a wealth of information regarding (partial) expressed gene sequences of molluscs (Fig. 1). To date, EST-library screening of this type continues to provide candidate immune factors for further characterization (e.g. Bai et al., 2009; Borisova and Gorbushin, 2014; Ding et al., 2011; Fleury et al., 2009; Gueguen et al., 2003; Goodson et al., 2005; Liao et al., 2013; Liu et al., 2011; Roberts and Goetz, 2003; Seo et al., 2016; Wang et al., 2009, 2007). For example, in 1999, Escoubas et al. recovered from the oyster *C. gigas* an EST sequence similar to mammalian IkappaB kinase (IKK) proteins that have central roles in cell (immune) signaling through activation of nuclear factor-kappaB (NF-kappaB). The recombinantly-expressed complete sequence effected expression of a NF-kappaB-controlled reporter gene. This provided a first indication of functional NFKb signaling in molluscs. The characterization of bivalve lectins and other defense factors has been driven mostly by initial identification of EST sequences (Kang et al., 2006; Korneev et al., 2002; Song et al., 2006). It is of particular interest that RNA interference is available for functional transcriptomics in gastropod and bivalve molluscs (Jiang et al., 2006; Owens and Malham, 2015; You et al., 2012).

RNAi knockdown of gene transcripts, and thereby of protein expression, effected by either long double-stranded RNA or short interfering RNA has shown phenotypic changes in *B. glabrata* to demonstrate involvement of immune activities of FREP3 (Hanington et al., 2012, 2010), the cytokine MIF (Baeza-Garcia et al., 2010), and antibacterial LBP/BPI (Baron et al., 2013).

The analysis of EST datasets can reveal groups of genes that are functionally linked such that their expression provides evidence for active cellular, humoral and metabolic processes in an organism. However, EST-based expression profiles recorded from insert libraries best inform by presence of particular sequences. This technique provides modest information regarding relative expression levels of (constitutive) immune genes, that are perhaps increased in response to a pathogen or regarding the potential for pathogen-specific immune responses in molluscs. The lack of detection of a particular immune transcript sequence may be due to differential expression or incomplete sampling. Due to the great effort required for sequencing many individual EST clones, even in automated fashion, and the random nature of insert selection for sequencing, it is unlikely that EST projects yield full representation of the total diversity of expressed sequences.

Microarray approaches provide an alternative way to study transcriptomics that affords a more accurate comparison of expression levels of particular (immune) genes by differentially-treated molluscs. Several microarray platforms have been designed and applied to study immunity in gastropods and bivalves of medical or aquaculture relevance (De Zoysa et al., 2012, 2011; Dheilly et al., 2011; Fleury and Huvet, 2012; Jenny et al., 2007; Jia et al., 2011; Leite et al., 2013; Romero et al., 2015; Venier et al., 2011; Wang et al., 2016, 2010; Zhang et al., 2016). The number of gene targets included on initial microarray designs was limited by the modest extent of available sequence data but recent microarrays can harbor upward of 30,000 gene

features such that immune sequences are likely included, even if these were not specifically selected as targets on the array. Microarrays that also contain unannotated gene targets have potential to identify unknown, novel candidate immune genes if these are differentially expressed following pathogen encounters. Microarray studies have shown novel aspects of immunity of particular species of molluscs. For example, different strains of the gastropod *B. glabrata* display distinct baseline expression profiles (Lockyer et al., 2012, 2008; Zahoor et al., 2014); this snail mounts pathogen-specific immune responses to Gram (-), Gram (+) bacteria and metazoan parasites (Adema et al., 2010), and FREP3 was identified as a common feature of successful defense responses (Hanington et al., 2012, 2010).

Such studies provided less insights to broaden comparative immunology. Microarrays are species-specific; they employ previously available sequences for a particular organism as hybridization targets to detect relative amounts of matching gene transcripts when comparing different RNA samples of the same organism. Comparison of results from expression studies for general features of molluscan immunity are difficult because microarrays are available only for a limited number of gastropods and bivalves and the various array platforms may differ considerably in number and representation of gene targets. Nevertheless, the focus of molluscan immunogenomics on gene discovery through sequencing and characterization of (partial) cDNA sequences was highly effective in cataloguing and defining immune capabilities of bivalves and gastropods, even though the techniques available were unlikely to capture complete transcriptomes and thus did not reveal the full extent of gene diversity within molluscan species or the taxonomic distribution of particular molluscan defense genes. Furthermore, in this time of extensive characterization of cDNA sequences, little information had accrued regarding

intron/exon structures, distribution of genes in the genome, general genome architecture, and regulatory sequences for managing gene expression toward immune responses.

4. Genomes and next-generation sequencing, immunogenomics phase 2

Genome sequencing efforts, culminating in completion of the human genome in 2000, had led to development of scientific technology that was able to produce high quality genome sequence assemblies. However, the capacity for genome sequencing was limited. The complex and expensive hardware infrastructure that afforded genome sequencing was available only at a modest number of central facilities. Full genome characterization required a large, costly effort reliant on labor-intensive sequencing methods and challenging computational bioinformatics to organize massive sequence datasets into a genome assembly. Nevertheless, advocacy from (international) consortia of scientists led to inclusion of several molluscs with relevance for biomedical research or aquaculture among the exclusive group of organisms that enjoyed high quality, full genome characterization. Obviously, by providing insight into all aspects of organismal biology, genome sequences yield valuable resources, also for comparative immunology. To better qualify for full genome sequencing, some research communities developed complementary genomic resources. Bacterial artificial chromosome (BAC) libraries were developed for several molluscs like *B. glabrata* (Gastropoda) and the oysters *Crassostrea virginica* and *C. gigas*, to gain initial access to genomic information (Adema et al., 2006; Cunningham et al., 2006; Raghavan et al., 2007). BAC libraries consist of considerable numbers of clones that represent the full genome with large (>100 kb) inserts of molluscan genomic DNA. Selection and detailed analysis of BAC inserts is a somewhat protracted process, but allows for focusing investigative analysis towards genes of interest. Several BACs have been sequenced full-length in the case of *B. glabrata* (Tennessen et al., 2015a; Hanington et al., 2010). This

further confirmed the presence of gene families of some immune genes in the genome of this gastropod: four FREP genes are present within a 115,524 base pair (bp) genomic region (Hanington et al., 2010) and *B. glabrata* BAC clone BG_BBa-10D22 (182,461 bp) contains five peptidoglycan recognition protein (PGRP) genes (GenBank: AC235813.3). Pending full genome assembly, this likely provides an incomplete view of these gene families in *Biomphalaria*. For instance, PGRP gene families comprising up to 12 different genes are present in other molluscs species such as *Bathymodiolus platifrons*, *C. gigas*, *Euprymna scolopes*, *Octopus vulgaris*, and *Solen grandis* (Castellanos- Martinez et al., 2014; Goodson et al., 2005; Itoh and Takahashi, 2008; Wei et al., 2012; Wong et al., 2015; Zhang and Yu, 2013). The use of BAC inserts containing defense/stress response sequences (actin and ferritin) as probes for fluorescent *in situ* hybridization (FISH) showed a repositioning of *B. glabrata* chromosomes in the nucleus following parasite exposure, likely in aid of active transcription of response genes (Knight et al., 2011). The utility of BAC libraries is underscored by continued use of this approach to characterize immune genes of additional molluscs in recent times, with continued observation of clustering of immune factors, e.g. two genes encoding lipopolysaccharide and beta-1,3- glucan binding proteins in the Zhikong scallop *Chlamys farreri* (Kasthuri et al., 2013; Premachandra et al., 2012; Zhang et al., 2008; Zhao et al., 2012a). Large scale genome efforts, aimed at obtaining high genome coverage and stringent computational genome assembly with availability of transcriptomic data for validation of predicted gene models have led to characterization of the full genomes of one bivalve *C. gigas*, the cephalopod *Octopus bimaculoides* and three gastropods *B. glabrata*, *Aplysia californica* and *Lottia gigantea* (see Table 1). Generally, molluscan genome assemblies consist of high numbers of genome scaffolds; concise assembly is challenged by large genome size and high repetitive content. The genome assemblies do provide,

however, good representation of expressed protein-encoding genes and genome size, e.g. 97% and 83%, respectively, for *O. bimaculoides* (Albertin et al., 2015). As such, they give unprecedented access to the most complete gene complements ever for these molluscs, and this allows broad analyses, including those for immune relevant genes.

Accordingly, genome analyses can now identify gene sequences that validate and further develop previous inferences about immune function in molluscs. In the case of molluscan cytokines, previously indicated by time-appropriate, yet rather indirect methods, sequence and functional analyses have confirmed gene sequences for TNF, IL-17 and MIF among bivalves, gastropods and cephalopods (also see Gao et al., 2015; Rosani et al., 2015a). Sequence homologs for Il-1 or Il-6 were not recorded. It remains to be determined whether activities previously ascribed to these cytokines derive from functional analogs or yet other factors (Ottaviani et al., 1995a, 1993).

Molluscs have abundant genes encoding for a diversity of lectins, including many with C-type lectin domains, fibrinogen-related sequences or C1q domains (e.g Gorbushin and Borisova, 2015; Zhang et al., 2012). Molluscs also show expanded gene families of Toll-like receptors (TLR). Diversity of this PRR likely expands non-self recognition capabilities (Buckley and Rast, 2015). The availability of non-coding intergenic genomic sequences further enables exploration of (immune-relevant) regulatory mechanisms involving transcription factors and nuclear receptors (Humphries and Harter, 2015; Kaur et al., 2015). Moreover, molluscs possess the machinery to employ methylation for epigenetic regulation of gene expression (Fneich et al., 2013; Geyer et al., 2011) and genome-wide analysis of *C. gigas* suggests that a nucleotide sequence composition is biased towards GC content in families of inducible genes (like stress and environmental response genes) to facilitate methylation for epigenetic control of gene

expression (Gavery and Roberts, 2010). The genome sequences also provide targets to develop markers for linkage studies. Such analysis of the association of particular SOD1 alleles with resistance to schistosome parasites indicated the presence of a linked cluster of redox genes and possibly other defense genes in the snail *B. glabrata* (Blouin et al., 2013). Additional analyses identified the so-called hyperdiverse Guadeloupe Resistance Cluster (HRC) and several other groups of genes as candidate genomic regions associated with different levels of susceptibility for parasite infection. Recent efforts have been made to identify the specific genes contributing to the differential susceptibility phenotype, but many remain to be characterized (Allan et al., 2017; Tennessen et al., 2015a, 2015b).

As indicated above, the availability of molluscan genomes facilitated comparative analyses of immune properties among diverse species representing different classes of molluscs. Mostly, however, the potential of genomic studies deepened comprehension of immunity unique to single molluscs. Rather than sparking comparative immunology across classes, orders or even phyla, genome sequences tended to focus research inward, for a few model molluscs. This was of course because of the considerable challenge to complete genome characterization for any particular organism, dramatically limiting the number of species for which this was achieved. This situation changed significantly with the availability of next-generation sequencing (NGS) capabilities. Several novel high throughput sequencing technologies have been developed as alternatives to routine Sanger sequencing. Of these, some have already been relinquished (454, Solid) while Illumina technology has become a mainstay (Pettersson et al., 2009). Compared to Sanger sequencing, NGS technology generates relatively short sequence reads (150 nt) that may challenge correct *de novo* assembly, but sequence is collected at massively increased fold-coverage and speed. Significantly, NGS carries a modest cost that now enables individual

researchers to capture and characterize both genome and transcriptome profiles for their organism(s) of choice (Sohn and Nam, 2016). The third (next-next) generation of sequencing technology is already coming on-line, and by enabling significantly longer sequence reads, will improve sequence assembly capabilities (Ansorge, 2009; Rhoads and Au, 2015; Schadt et al., 2010). Of note, the amount of nucleic acid input material required for NGS is in the nanogram to microgram range, far less than required previously for Sanger technology. This enables complete sequencing from few or even individual organisms. With the current NGS capabilities, the number of sequenced genomes and transcriptomes, collected in the form of massive, comprehensive datasets is rapidly increasing (see Fig. 1). Without doubt, many more datasets are awaiting submission to public databases. These developments have rendered the unique status of traditional genome projects obsolete. In fact, some have mockingly declared the death of the genome paper (Smith, 2016), because such high- level prestigious publications are being replaced by reports that announce the availability of novel NGS data (e.g. Huang and Wu, 2015). The scrutiny and effort that yielded the initial set of genomes is unlikely to be paralleled and new genomes will not be assembled or annotated at the same level of quality. Clearly, however, NGS is of great benefit to comparative immunology. The capture of comprehensive (if not complete) genome and transcriptome sequence data, for any initial biological research question, provides an unprecedented data resource that can be mined with computational bioinformatics methods, and also for immune sequences. With NGS, indeed there are effectively no more non- model organisms for gene discovery (Dheilly et al., 2014).

The new sequencing capabilities have already shaped new understanding at different levels of molluscan immunogenomics. NGS characterization of expression profiles of a novel strain of *B. glabrata* increased the number of unique FREP gene sequences previously recorded

from this gastropod, perhaps by also revealing novel alleles. Moreover, the non-selective manner of sequence sampling also identified FREP-related lectin sequences with similar domain structures (upstream IgSF sequences with a C- terminal lectin domain) but displaying either a galectin or a C-type lectin domain. These novel sequences, named galectin-related protein (GREP) and C-type lectin-related protein (CREP), respectively, are now grouped together with FREPs as Variable Immunoglobulin and Lectin domain containing molecules or VIgLs (Dheilly et al., 2015). Gorbushin and colleagues (Gorbushin et al., 2010; Gorbushin and Borisova, 2015) have tracked the distribution of defense factors like FREPs across gastropod phylogeny. Inspection of the genome of the euopisthobranch *A. californica*, and of NGS transcriptome data of hemocytes of the prosobranch *Littorina littorea*, have revealed FREP genes outside of *Biomphalaria*. No FREPs were observed from the genome of the basal gastropod *Lottia gigantea*. Pending analyses of additional NGS datasets from other molluscs, this led to development of the hypothesis that FREPs are an evolutionary innovation that took place after the emergence of class Gastropoda within the Mollusca (Gorbushin et al., 2010). This is especially intriguing because *Mytilus galloprovincialis* (class Bivalvia) was reported to express diversified proteins that contain fibrinogen-related domains, also known as FReDs, thought to function in immunity (Romero et al., 2011). Numerous other FReD-containing proteins, e.g. those consisting of single FBG domains, have been recorded from a range of molluscs but these have not been reported to be diversified (Albertin et al., 2015; Gorbushin and Iakovleva, 2011; Zhang et al., 2012). A phylogenetic analysis of molluscan FReD sequences, including FREPs, does not reveal a clear evolutionary origin for canonical FREPs (Fig. 2a). The FREP sequences from *A. californica*, *B. glabrata*, and *L. littorea* cluster separately among other FReD-containing sequences. Moreover, the fibrinogen-related sequences of *Biomphalaria* and *Mytilus* shown to

undergo extensive sequence diversification (Dheilly et al., 2015; Hanington et al., 2010; Romero et al., 2011; Zhang et al., 2004), fall out in different clades (Fig. 2b and c). This suggests that diversification of FReDs may have occurred independently at different stages during the evolution of molluscs. Of course, future investigations may uncover different instances of diversified FREPs and other FReD-containing proteins. Yet the fact remains that such systems for intra-individual sequence diversification greatly increase the overall FReD diversity that occurs in molluscs (Fig. 2b and c).

An apparently general concept of molluscan immunology was revealed by immunogenomic studies of NGS data involving gene complements and expression profiles from two classes: gastropods and bivalves. Deleury et al. (2012) collected NGS transcription profiles of *B. glabrata* snails that were evoked by different pathogens (yeast, Gram (+) and Gram (-) bacteria). Zhang et al. (2015) independently exposed *C. gigas* oysters to multiple pathogens including (strains of) particular bacteria and a virus, to determine expression profiles using Illumina NGS. Similar results were observed from annotation and analysis of specific expressed sequences: both the gastropod and the bivalve molluscs studied possess expanded immune gene families, and in both molluscs, pathogens are met with generally similar immune processes or signaling pathways. However, individual sequences of particular gene families are differentially expressed to generate pathogen- specific immune responses. This indicates that molluscs possess complex immune systems with extensive discriminatory properties.

5. Limitations and potential, future immunogenomics

NGS has already shown itself to be a powerful tool for immunogenomics. It holds great promise to elevate molluscan immunology from detailed study of a few model species to a broad vantage point for true comparative immunology that can reveal commonly shared as well as

unique immune features across the diversity of the phylum Mollusca. We must also consider carefully the limitations of NGS technology to maximize the future benefits. Current NGS methods generate massive datasets of relatively short read lengths. Computational assembly, especially *de novo* (i.e. without a reference genome or gene models), to reconstruct the genome or the transcript from which they derive, is very challenging. Several software packages are available for assembly but these are each designed with different assumptions and assembly strategies in mind. Comparison of *de novo* genome assemblies with reference-based (human) genome assemblies of Han Chinese and Yourbal individuals showed that the *de novo* assemblies were 16.2% shorter than the reference genomes and were lacking over 2000 protein coding exons (Alkan et al., 2011). When considered at the rate at which data is now being generated, these discrepancies can lead to erroneous conclusions about the absence of immune components for non-model, invertebrate organisms for which reference genomes are typically unavailable. Transcriptomes also succumb to assembly difficulties. The choice of assembly software can mean the difference between generating artificial, chimeric contigs, or producing many short, redundant contigs along with correctly assembled sequence contigs (Mundry et al., 2012).

Annotation of assembled sequence data to infer open reading frames, gene models and especially function, is another challenge for correct interpretation of molluscan immunology. For example, Blast2GO (Conesa et al., 2005) is one of the main tools for informative (automated) annotation of large sequence datasets from non-model organisms. The program relies on blast searches to identify transcripts with sequence similarity to a modest set of previously characterized sequences that have been associated with so-called Gene Ontology annotation to define their functional role in molecular, cellular or biochemical processes. Transcripts that do not share significant blast similarities with those included in the reference database are excluded

from the annotation. Typically this affects a high proportion of molluscan sequences, and although GO annotation is available for non-model organisms through B2G-FAR (Götz et al., 2011), the resulting annotation may still not be suited for correct functional interpretation of molluscan immune factors. For instance, a FREP lectin transcript that contains two functionally distinct IgSF and fibrinogen-like domains may generate similarities to sequences with annotation that indicates antibody-based immune function and fibrinogen-mediated blood clotting, even though neither of these phenomena exist in snail biology. Thus it is of considerable importance to verify at least some results from *in silico* analysis of NGS sequence data, validating assembly of sequences of interest and applying knowledge of specific organismal biology.

Ultimately, immunogenomics can provide an essential mechanistic basis to evaluate the biological relevance of immune phenomena that may be inferred from observations at the organismal level (Hauton and Smith, 2007; Little et al., 2008). Surprisingly, we may uncover mechanistically supported immune phenomena in the lab that fit less with our perceptions of how immunity should benefit fitness. For instance, there is the unexpected observation of immune memory in the snail *B. glabrata* that prevents secondary infection by the parasite *S. mansoni* in snails that continue to sustain a primo infection by the same parasite strain, without obvious benefits to snail fitness (Pinaud et al., 2016).

For developing a representative overview of molluscan immunology, increasingly refined phylogenies are available to guide selection of species that will fill the current gaps in coverage by immunogenomics analysis of molluscan diversity (Kocot et al., 2011; Smith et al., 2011). As evident from Fig. 1, several classes of molluscs deserve particular attention for obtaining NGS data. Individual investigators are encouraged to release any available relevant NGS datasets so that duplication of sequencing efforts can be avoided. The Global Invertebrate Genomics

Alliance (GIGA community of Scientists, 2014; http://giga-cos.org/) aims to generate genomic sequences for 5000 (marine) invertebrates and may offer datasets for molluscan clades yet to be studied with regard to their immune function. Recent depositions of NGS data in public databases indicate that it is feasible to obtain and collect NGS data from molluscs of the classes Scaphopoda, Monoplacophora, Polyplacophora and Aplacophora. The currently available data of these lesser studied molluscs deserves analysis for immune genes. Encouragingly, genome size estimates have been made for several Aplacophora, Polyplacophora and Scaphopoda (Kocot et al., 2015) in light of interest to perform NGS-based genome characterization of representatives of these groups. Given the great diversity of the Mollusca, it is unlikely that full coverage by genome sequencing is ever achieved. However, strategic sampling will likely provide an initial overview that can help to identify taxa that may harbor unique aspects of immunobiology (gene gains or losses) and that are of particular interest for more detailed sampling.

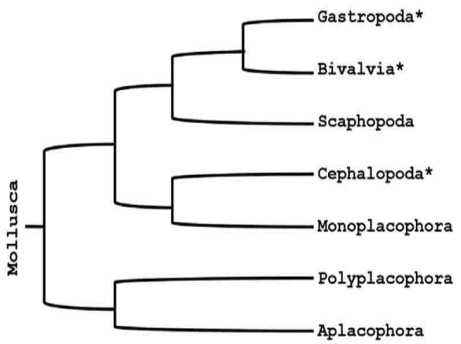
A more comprehensive, inclusive understanding of immune function in molluscs will identify genes and immune mechanisms that can be targeted to bolster immune vigor in species with importance for aquaculture or conservation, or for control efforts aimed at molluscs that negatively impact humanity. Notably, it is worthwhile to consider the potential of the CRISPR/Cas9 system for genome editing (Singh et al., 2017) as a means for modifications of molluscan genomes in light of future functional immunogenomics studies.

In closing, recent studies that track distribution of IL-17 signaling components in over 30 bivalves (Rosani et al., 2015a), or that reveal that LBP/BPI proteins are not routinely conserved across animal phylogeny, yet are import immune factors in molluscs (Baron et al., 2016), show the great potential of modern immunogenomics to push molluscan immunology from detailed

study of limited numbers of model species to true comparative immunology, revealing aspects of evolution of animal immune function across broad ranges of phylogeny.

Acknowledgements

The authors wish to thank E.S Loker and all other colleagues at the Center for Evolutionary and Theoretical Immunology at UNM for helpful discussions. This study was supported by NIH grant P30GM110907 from the National Institute of General Medical Sciences.



	ESTs					SRA	
	1990	2000	2010	2016	# species	2016	# species
Gastropoda*	0	1232	630146	647770	23	1985	182
Bivalvia*	0	0	347393	382612	37	2829	127
Scaphopoda	0	0	0	0	-	7	6
Cephalopoda*	0	0	46081	114034	6	225	38
Monoplacophora	0	0	0	0	-	2	1
Polyplacophora	0	0	498	1548	2	5	5
Aplacophora	0	0	0	1907	2	13	11

Fig. 1. Simplified phylogeny of the classes in the phylum Mollusca adapted from Kocot et al., 2011 and Smith et al., 2011. Asterisks indicate genome availability in public databases. For each class the number and increments of expressed sequence tag (EST) and short read archive (SRA) entries are indicated over time. Note that SRA entries were not available before 2008, coinciding with the advent of next-generation capabilities. Data were obtained by querying the SRA and EST databases of GenBank using name of class and modification dates as search terms.

Table 1
Partial listing of publications reporting on NGS data collection from molluscs. The listing of transcriptomic data only includes immune-related studies. Information for genomic data was retrieved from the Genome database of NCBI and the Animal Genome Size Database (Gregory, 2017). Publicly available genomes without a corresponding publication are indicated with an asterix.

Transcriptomic Data

Class	Organism	NGS Platform	Reference
Gastropoda	*Arion vulgaris*	Illumina HiSeq2000	Bulat et al., 2016
Gastropoda	*Biomphalaria glabrata*	Illumina GAIIx	Pinaud et al., 2016
Gastropoda	*Biomphalaria glabrata*	Illumina HiSeq2000	Kenny et al., 2016
Gastropoda	*Conus geographus*	Illumina HiSeq2000	Safavi-Hemami et al., 2016
Gastropoda	*Haliotis discus hannai*	Illumina HiSeq2000	Nam et al., 2016
Gastropoda	*Littorina littorea*	Illumina HiSeq2000	Gorbushin and Borisova, 2015
Gastropoda	*Oncomelania hupensis*	Illumina HiSeq2000	Zhao et al., 2015
Gastropoda	*Biomphalaria glabrata*	Illumina GAIIx	Deleury et al., 2012
Gastropoda	*Thais clavigera*	454 GS 20	Rhee et al., 2012
Bivalvia	*Crassostrea virginica*	Illumina GAIIx	McDowell et al., 2016
Bivalvia	*Cristaria plicata*	Illumina HiSeq2500	Patnaik et al., 2016
Bivalvia	*Mercenaria mercenaria*	Illumina HiSeq2000	Wang et al., 2016b
Bivalvia	*Mytilus chilensis*	Illumina HiSeq2000	Detree et al., 2016
Bivalvia	*Pinctada fucata*	Illumina HiSeq2000	Wang et al., 2016c
Bivalvia	*Saccostrea glomerata*	Illumina HiSeq2000	Ertl et al., 2016
Bivalvia	bivalvia		Rosani et al., 2015a
Bivalvia	*Bathymodiolus platifrons*	Illumina HiSeq2000	Wong et al., 2015
Bivalvia	*Chlamys farreri*	Illumina HiSeq2000	Hu et al., 2015
Bivalvia	*Crassostrea gigas*	Ion Torrent Proton	Chen et al., 2015
Bivalvia	*Crassostrea gigas*	Illumina GAIIx	He et al., 2015
Bivalvia	*Crassostrea gigas*	Illumina HiSeq2000	Rosani et al., 2015b
Bivalvia	*Crassostrea gigas*	Illumina GAIIx	Zhao et al., 2015
Bivalvia	*Patinopecten yessoensis*	Illumina HiSeq2000	Ding et al., 2015
Bivalvia	*Crassostrea gigas Crassostrea hongkongensis*	Illumina HiSeq2000	Zhao et al., 2014
Bivalvia	*Crassostrea virginica*	Illumina GAIIx	McDowell et al., 2014
Bivalvia	*Crassostrea virginica*	Illumina GAIIx	Zhang et al., 2014
Bivalvia	*Mytilus galloprovincialis*	Illumina GAIIx	Gerdol et al., 2014
Bivalvia	*Pecten maximus*	Illumina HiSeq2000	Pauletto et al., 2014
Bivalvia	*Mytilus chilensis*	Illumina HiSeq2000	Nuñez-Acuña and Gallardo-Ascárate, 2013
Bivalvia	*Mytilus edulis*	454 GS FLX Titanium	Philipp et al., 2012
Bivalvia	*Pinctada martensii*	Illumina HiSeq2000	Zhao et al., 2012b
Bivalvia	*Patinopecten yessoensis*	454 GS FLX Titanium	Hou et al., 2011
Bivalvia	*Bathymodiolus azoricus*	454 GS FLX Titanium	Bettencourt et al., 2010
Cephalopoda	*Euprymna tasmanica*	454 GS FLX+	Salazar et al., 2015
Cephalopoda	*Octopus vulgaris*	Illumina GAIIx	Castellanos-Martinez et al., 2014

Genomic Data

Class	Organism	GenBank Assembly Accession	Reference	Genome Size (bp)/database source
Gastropoda	*Biomphalaria glabrata*	GCA_000457365.1	Adema et al., 2017: In prep*	9.16 × 10^8 /GenBank
Gastropoda	*Conus tribblei*	GCA_001262575.1	Barghi et al., 2016	2.16 × 10^9/GenBank
Gastropoda	*Lymnaea stagnalis*	GCA_900036025.1	Davison et al., 2016	8.33 × 10^8/GenBank
Gastropoda	*Aegista diversifamilia Dolicheulota formosensis*	(SRA Accessions) SRR1918809 SRR1920140	Huang and Wu, 2015	Data not available
Gastropoda	*Aplysia californica*	GCA_000002075.2	Moroz and Kandel, 2006*	9.27 × 10^8/GenBank
Gastropoda	*Lottia gigantea*	GCA_000327385.1	Simakov et al., 2013	3.59 × 10^8/GenBank
Bivalvia	*Corbicula fluminea*	GCA_001632725.1	Peñarrubia et al., 2016*	Data not available
Bivalvia	*Dreissena polymorpha*	GCA_000806325.1	Peñarrubia et al., 2015	1.66 × 10^9/Animal Genome Size Database
Bivalvia	*Mytilus galloprovincialis*	GCA_000715055.1	Nguyen et al., 2014*	1.62 × 10^9/GenBank
Bivalvia	*Crassostrea gigas*	GCA_000297895.1	Zhang et al., 2012	5.57 × 10^8/GenBank
Cephalopoda	*Octopus bimaculoides*	GCA_001194135.1	Albertin et al., 2015	2.33 × 10^9/GenBank

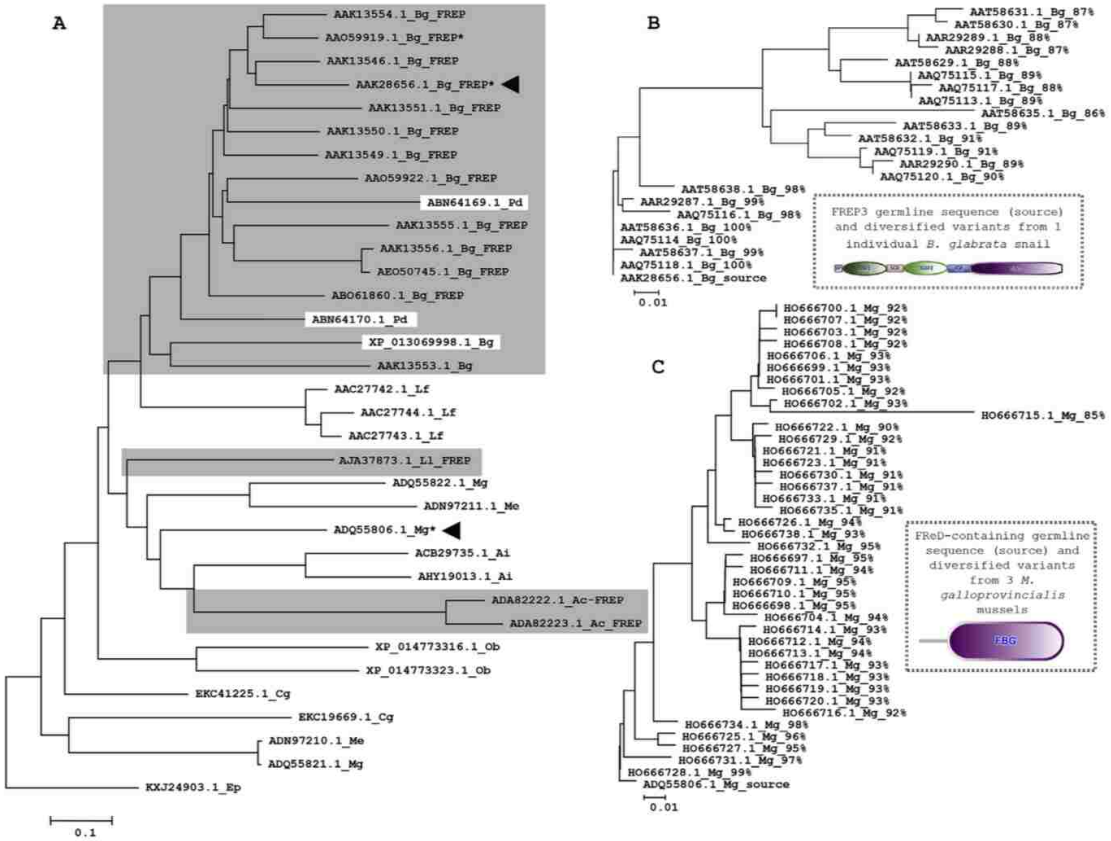


Fig. 2. Phylogenetic reconstruction of molluscan fibrinogen-related domain (FReD) sequences. (A) Phylogenetic tree of FReD-containing sequences from three molluscan classes: Gastropoda (5 species), Bivalvia (4 species), Cephalopoda (1 species). A FReD sequence from the cnidarian (pre-bilaterian) *Exaiptaisa pallida* (Ep) is used as outgroup. The alignment of these FReD sequences included 103 amino acids (aa). Grey boxes highlight canonical fibrinogen-related proteins (FREPs: sequences containing 1 or 2 N-terminal IgSF domains and a C-terminal FBG domain). Asterisks indicate FReD-containing sequences reported to be extensively diversified (Dheilly et al., 2015; Hanington et al., 2010; Romero et al., 2011; Zhang et al., 2004). Arrowheads mark germ-line encoded sequences used in Fig. 2b and c. (B) Sequence diversity of a 127 aa region of the fibrinogen-related sequence of FREP3 (canonical structure indicated) recorded from the germline encoded (source) sequence and diversified variants from one individual *B. glabrata* snail. For each percentage value the number of different amino acids residues relative to the source sequence (127aa) are as follows: 100% = 0aa, 99% = 1aa, 98% = 2aa, 91% = 11aa, 90% = 12aa, 89% = 13aa, 88% = 14-15aa, 87% = 16aa, 86% = 17aa. (C) Sequence diversity of a 129 aa region of the fibrinogen-related sequence of a single FReD-containing lectin (see structure) recorded from the germline encoded (source) sequence and diversified variants from three individual *M. galloprovincialis* mussels. For each percentage value the number of different amino acids residues relative to the source sequence (129aa) are as follows: 100% = 0aa, 99% = 1aa, 98% = 2aa, 97% = 3aa, 96% = 4-5aa, 95% = 6aa, 94% = 7aa, 93% = 8-9aa, 92% = 10aa, 91% = 11aa, 90% = 12aa, 85% = 19aa. *Aplysia californica* (Ac), *Argopecten irradians* (Ai), *Biomphalaria glabrata* (Bg), *Crassostrea gigas* (Cg), *Limax flavus* (Lf), *Littorina littorea* (Ll), *Mytilus edulis* (Me), *Mytilus galloprovincialis* (Mg), *Octopus bimaculoides* (Ob), *Planorbella duryi* (Pd). All accession numbers are provided within the trees. Analyses were conducted in MEGA6 (Tamura et al., 2013) using the Neighbor-Joining (Saitou and Nei, 1987) and Poisson correction methods (Zuckerkandl and Pauling, 1965). Scale bars: units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated.

References

1. Adema, C.M., van der Knaap, W.P.W., Sminia, T., 1991. Molluscan hemocyte- mediated cytotoxicity: the role of reactive oxygen intermediates. Rev. Aquat. Sci. 4, 201-223.

2. Adema, C.M., van Deutekom-Mulder, E.C., van der Knaap, W.P.W., Sminia, T., 1993. NADPH-oxidase activity: the probable source of reactive oxygen intermediate generation in hemocytes of the gastropod *Lymnaea stagnalis*. J. Leukoc. Biol. 54, 379-383.

3. Adema, C.M., van Deutekom-Mulder, E.C., van der Knaap, W.P.W., Sminia, T., 1994. Schistosomicidal activities of *Lymnaea stagnalis* haemocytes: the role of oxygen radicals. Parasitology 109, 479-485.

4. Adema, C.M., Hertel, L.A., Loker, E.S., 1997a. Infection with *Echinostoma paraensei* (Digenea) induces parasite-reactive polypeptides in the hemolymph of the gastropod host *Biomphalaria glabrata*. In: Beckage, N.E., Parasites, Pathogens (Eds.), Effects on Host Hormones and Behavior. Chapman & Hall, New York, pp. 76-98.

5. Adema, C.M., Hertel, L.A., Miller, R.D., Loker, E.S., 1997b. A family of fibrinogen-related proteins that precipitates parasite-derived molecules is produced by an invertebrate after infection. Proc. Natl. Acad. Sci. U. S. A. 94, 8691-8696.

6. Adema, C.M., Luo, M.Z., Hanelt, B., Hertel, L.A., Marshall, J.J., Zhang, S.M., DeJong, R.J., Kim, H.R., Kudrna, D., Wing, R.A., Soderlund, C., Knight, M., Lewis, F.A., Caldeira, R.L., Jannotti-Passos, L.K., Carvalho, Odos S., Loker, E.S., 2006. A bacterial artificial chromosome library for *Biomphalaria glabrata*, intermediate snail host of *Schistosoma mansoni*. Mem. Inst. Oswaldo Cruz 101, 167-177.

7. Adema, C.M., Hanington, P.C., Lun, C.M., Rosenberg, G.H., Aragon, A.D., Stout, B.A., Lennard Richard, M.L., Gross, P.S., Loker, E.S., 2010. Differential transcriptomic responses of *Biomphalaria glabrata* (Gastropoda, Mollusca) to bacteria and metazoan parasites, *Schistosoma mansoni* and *Echinostoma paraensei* (Digenea, Platyhelminthes). Mol. Immunol. 47, 849-860.

8. Adema, C.M., Bayne, C.J., Bridger, J.M., Knight, M., Loker, E.S., Yoshino, T.P., Zhang, S.M., 2012. Will all scientists working on snails and the diseases they transmit please stand up? PLoS Negl. Trop. Dis. 6, e1835.

9. Adema, C.M., Loker, E.S., 2015. Digenean-gastropod host associations inform on aspects of specific immunity in snails. Dev. Comp. Immunol. 48, 275-283. Albertin, C.B., Simakov, O., Mitros, T., Wang, Z.Y., Pungor, J.R., Edsinger-Gonzales, E., Brenner, S., Ragsdale, C.W., Rokhsar, D.S., 2015. The octopus genome and the evolution of cephalopod neural and morphological novelties. Nature 524, 220-224.

10. Alkan, C., Sajjadian, S., Eichler, E.E., 2011. Limitations of next-generation genome sequence assembly. Nat. Methods 8, 61-65.

11. Allan, E.R., Tennessen, J.A., Bollmann, S.R., Hanington, P.C., Bayne, C.J., Blouin, M.S., 2017. Schistosome infectivity in the snail, *Biomphalaria glabrata*, is partially dependent on the expression of Grctm6, a Guadeloupe Resistance Complex protein. PLoS Negl. Trop. Dis. 11, e0005362.

12. Ansorge, W.J., 2009. Next-generation DNA sequencing techniques. N. Biotechnol. 25, 195e203.

13. Baeza Garcia, A., Pierce, R.J., Gourbal, B., Werkmeister, E., Colinet, D., Reichart, J.M., Dissous, C., Coustau, C., 2010. Involvement of the cytokine MIF in the snail host immune response to the parasite *Schistosoma mansoni*. PLoS Pathog. 6, e1001115.

14. Bai, Z., Yin, Y., Hu, S., Wang, G., Xhang, X., Li, J., 2009. Identification of genes involved in immune response, microsatellite, and SNP markers from expressed sequence tags generated from hemocytes of freshwater pearl mussel (*Hyriopsis cumingii*). Mar. Biotechnol. (NY) 11, 520-530.

15. Barghi, N., Concepcion, G.P., Olivera, B.M., Lluisma, A.O., 2016. Structural features of conopeptide genes inferred from partial sequences of the *Conus tribblei* genome. Mol. Genet. Genomics 291, 411-422.

16. Baron, O.L., van West, P., Industri, B., Ponchet, M., Dubreuil, G., Gourbal, B., Reichart, J.M., Coustau, C., 2013. Parental transfer of the antimicrobial protein LBP/BPI protects *Biomphalaria glabrata* eggs against oomycete infections. PLoS Pathog. 9, e1003792.

17. Baron, O.L., Deleury, E., Reichart, J.M., Coustau, C., 2016. The LBP/BPI multigenic family in invertebrates: evolutionary history and evidences of specialization in mollusks. Dev. Comp. Immunol. 57, 20-30.

18. Bayne, C.J., 1980. Molluscan immunity: induction of elevated immunity in the land snail (*Helix*) by injections of bacteria (*Pseudomonas aeruginosa*). Dev. Comp. Immunol. 4, 43-54.

19. Beschin, A., Bilej, M., Brys, L., Torreele, E., Lucas, R., Magez, S., De Baetselier, P., 1999. Convergent evolution of cytokines. Nature 400, 627-628.

20. Bettencourt, R., Pinheiro, M., Egas, C., Gomes, P., Afonso, M., Shank, T., Santos, R.S., 2010. High-throughput sequencing and analysis of the gill tissue transcriptome from the deep-sea hydrothermal vent mussel *Bathymodiolus azoricus*. BMC Genomics 11, 559.

21. Blouin, M.S., Bonner, K.M., Cooper, B., Amarasinghe, V., O'Donnell, R.P., Bayne, C.J., 2013. Three genes involved in the oxidative burst are closely linked in the genome of the snail, *Biomphalaria glabrata*. Int. J. Parasitol. 43, 51-55.

22. Boman, H.G., Hultmark, D., 1987. Cell-free immunity in insects. Annu. Rev. Microbiol. 41, 103-126.

23. Borisova, E.A., Gorbushin, A.M., 2014. Molecular cloning of a-2-macroglobulin from hemocytes of common periwinkle *Littorina littorea*. Fish Shellfish Immunol. 39, 136-137.

24. Bouchut, A., Roger, E., Coustau, C., Gourbal, B., Mitta, G., 2006. Compatibility in the *Biomphalaria glabrata*/*Echinostoma caproni* model: potential involvement of adhesion genes. Int. J. Parasitol. 36, 175-184.

25. Buckley, K.M., Rast, J.P., 2015. Diversity of animal immune receptors and the origins of recognition complexity in the deuterostomes. Dev. Comp. Immunol. 49, 179-189.

26. Bulat, T., Smidak, R., Sialana, F.J., Jung, G., Rattei, T., Bilban, M., Sattmann, H., Lubec, G., Aradska, J., 2016. Transcriptomic and proteomic analysis of *Arion vulgaris*-Proteins for probably successful survival strategies? PLoS One 11, e0150614.

27. Carnegie, R.B., Arzul, I., Bushek, D., 2016. Managing marine mollusc diseases in the context of regional and international commerce: policy issues and emerging concerns. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 371.

28. Castellanos-Martinez, S., Arteta, D., Catarino, S., Gestal, C., 2014. *De novo* transcriptome sequencing of the *Octopus vulgaris* hemocytes using Illumina RNA-Seq technology: response to the infection by the gastrointestinal parasite *Aggregata octopiana*. PLoS One 9, e107873.

29. Chen, H., Wang, L., Zhou, Z., Hou, Z., Liu, Z., wang, W., Gao, D., Gao, Q., Wang, M., Song, L., 2015. The comprehensive immunomodulation of NeurimmiRs in haemocytes of oyster *Crassostrea gigas* after acetylcholine and norepinephrine stimulation. BMC Genomics 16, 942.

30. Cheng, T.C., Marchalonis, J.J., Vasta, G.R., 1984. Role of molluscan lectins in recognition processes. Prog. Clin. Biol. Res. 157, 1-15.

31. Cheng, T.C., Manzi, J.J., 1996. Correlation between the presence of lathyrose with the absence of *Haplosporidium nelsoni* in *Crassostrea virginica* from two South Carolina tributaries where *Perkinsus marinus* also inhibits hemocyte agglutination by the *Lathyrus odoratus* lectin. J. Shellfish Res. 15, 391-394.

32. Claes, M.F., 1996. Functional morphology of the white bodies of the cephalopod mollusc *Sepia officinalis*. Acta Zool. 77, 173-190.

33. Conesa, A., Götz, S., García-Gomez, J.M., Terol, J., Talo, M., Robles, M., 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics. Bioinformatics 21, 3674-3676.

34. Connors, V.A., de Buron, I., Granath Jr., W.O., 1995a. *Schistosoma mansoni*: interleukin-1 increases and decreases output of cercariae in schistosome-susceptible *Biomphalaria glabrata*. Exp. Parasitol. 80, 139-148.

35. Connors, V.A., de Buron, I., Granath Jr., W.O., 1995b. *Schistosoma mansoni* interleukin-1 increases phagocytosis and superoxide production by hemocytes and decrease output of cercariae in schistosome-susceptible *Biomphalaria glabrata*. Ex. Parasitol. 80, 139-148.

36. Connors, V.A., de Buron, I., Jourdane, J., Theron, A., Agner, A., Granath Jr., W.O., 1998. Recombinant human interleukin-1-mediated killing of *Schistosoma mansoni* primary sporocysts in *Biomphalaria glabrata*. J. Parasitol. 84, 920-926.

37. Coustau, C., Gourbal, B., Duval, D., Yoshino, T.P., Adema, C.M., Mitta, G., 2015. Advances in gastropod immunity from the study of the interaction between the snail *Biomphalaria glabrata* and its parasites: a review of research progress over the last decade. Fish Shellfish Immunol. 46, 5-16.

38. Cowden, R.R., 1972. Some cytological and cytochemical observations on the leucopoietic organs, the "white bodies" of *Octupus vulgaris*. J. Invertebr. Pathol 19, 113-119.

39. Cunningham, C., Hikima, J., Jenny, M.J., Chapman, R.W., Fang, G.C., Saski, C., Lundqvist, M.L., Wing, R.A., Cupit, P.M., Gross, P.S., Warr, G.W., Tomkins, J.P., 2006. New resources for marine genomics: bacterial artificial chromosome libraries for the Eastern and Pacific oysters (*Crassostrea virginica* and *C. gigas*). Mar. Bitotechnol. (NY) 8, 521-533.

40. Dam, T.K., Sarkar, M., Ghosal, J., Choudhury, A., 1992. A novel galactosyl-binding lectin from the plasma of the blood clam, *Anadara granosa* (L) and a study of its combining site. Mol. Cell Biochem. 117, 1-9.

41. Davison, A., McDowell, G.S., Holden, J.M., Johnson, H.F., Koutsovoulos, G.D., Liu, M.M., Hulpiau, P., Van Roy, F., Wade, C.M., Banerjee, R., Yang, F., Chiba, S., Davey, J.W., Jackson, D.J., Levin, M., Blaxter, M.L., 2016. Formin is associated with left-right asymmetry in the pond snail and the frog. Curr. Biol. 26, 654-660.

42. Deleury, E., Dubreuil, G., Elangovan, N., Wajnberg, E., Reichart, J.M., Gourbal, B., Duval, D., Baron, O.L., Gouzy, J., Coustau, C., 2012. Specific versus non-specific immune responses in an invertebrate species evidenced by a comparative *de novo* sequencing study. PLoS One 7, e32512.

43. Detree, C., Núñez-Acuña, G., Roberts, S., Gallardo-Escarate, C., 2016. Uncovering the complex transcriptome response of *Mytilus chilensis* against saxitoxin: implications of harmful algal blooms on mussel populations. PLoS One 11, e0165231.

44. De Zoysa, M., Nikapitiya, C., Oh, C., Lee, Y., Whang, I., Lee, J.S., Choi, C.Y., Lee, J., 2011. Microarray analysis of gene expression in disk abalone *Haliotis discus discus* after bacterial challenge. Fish Shellfish Immunol. 30, 661-673.

45. De Zoysa, M., Nikapitiya, C., Oh, C., Whang, I., Shin, H.J., Lee, J., 2012. cDNA microarray analysis of disk abalone genes in gills and hemocytes after viral hemorrhagic septicemia virus (VHSV) challenge. Fish Shellfish Immunol. 32, 1205-1215.

46. Dheilly, N.M., Lelong, C., Huvet, A., Favrel, P., 2011. Development of a Pacific oyster (*Crassostrea gigas*) 31,918-feature microarray: identification of reference genes and tissue-enriched expression patterns. BMC Genomics 12, 468.

47. Dheilly, N.M., Adema, C., Raftos, D.A., Goutbal, B., Grunau, C., Du Pasquier, l, 2014. No more non-model species: The promise of next generation sequencing for comparative immunology. Dev. Comp. Immunol. 45, 56-66.

48. Dheilly, N.M., Duval, D., Mouahid, G., Emans, R., Allienne, J.F., Galinier, R., Genthin, C., Dubois, E., Du Pasquier, L., Adema, C.M., Grunau, C., Mitta, G., Gourbal, B., 2015. A family of variable immunoglobulin and lectin domain containing molecules in the snail *Biomphalaria glabrata*. Dev. Comp. Immunol. 48, 234-243.

49. Ding, J., Li, J., Bao, Y., Li, L., Wu, F., Zhang, G., 2011. Molecular characterization of a mollusk chicken-type lysozyme gene from *Haliotis discus hannai* Ino, and the antimicrobial activity of its recombinant protein. Fish Shellfish Immunol. 30, 163-172.

50. Ding, J., Zhao, L., Chang, Y., Zhao, W., Du, Z., Hao, Z., 2015. Transcriptome sequencing and characterization of Japanese scallop *Patinopecten yessoensis* from different shell color lines. PLoS One 10, e0116406.

51. Ertl, N.G., O'Connor, W.A., Papanicolaou, A., Weigand, A.N., Elizur, A., 2016. Transcriptome analysis of the Sydney rock oyster *Saccostrea glomerata*: Insights into molluscan immunity. PLoS One 11, e0156649.

52. Erwin, D.H., Laflamme, M., Tweedt, S.M., Sperling, E.A., Pisani, D., Peterson, K.J., 2011. The Cambrian conundrum: early divergence and later ecological success in the early history of animals. Science 334, 1091-1097.

53. Escoubas, J.M., Briant, L., Montagnani, C., Hez, S., Devaux, C., Roch, P., 1999. Oyster IKK-like protein shares structural and functional properties with its mammalian homologues. FEBS Lett. 453, 293-298.

54. Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., Mello, C.C., 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature 391, 806-811.

55. Fleury, E., Huvet, A., Lelong, C., de Lorgeril, J., Boulo, V., Gueguen, Y., Bachere, E., Tanguy, A., Moraga, D., Fabioux, C., Lindeque, P., Shaw, J., Reinhardt, R., Prunet, P., Davey, G., Lapegue, S., Sauvage, C., Corporeau, C., Moal, J., Gavory, F., Wincker, P., Moreews, F., Klopp, C., Mathieu, M., Boudry, P., Favrel, P., 2009. Generation and analysis of a 29,745 unique expressed sequence tags from the Pacific oyster (*Crassostrea gigas*) assembled into a publicly accessible database: the Gigas Database. BMC Genomics 10, 341.

56. Fleury, E., Huvet, A., 2012. Microarray analysis highlights immune response of pacific oysters as a determinant of resistance to summer mortality. Mar. Biotechnol. (NY) 14, 203-217.

57. Fneich, S., Dheilly, N., Adema, C., Rognon, A., Reichelt, M., Bulla, J., Grunau, C., Cosseau, C., 2013. 5-methyl-cytosine and 5-hydroxy-methyl-cytosine in the genome of *Biomphalaria glabrata*, a snail intermediate host of *Schistosoma mansoni*. Parasit. Vectors 6, 167.

58. Gao, D., Qiu, L., Gao, Q., Hou, Z., Wang, L., Song, L., 2015. Repertoire and evolution of TNF superfamily in *Crassostrea gigas*: implications for expansion and diversification of this superfamily in Mollusca. Dev. Comp. Immunol. 51, 251-260.

59. Gavery, M.R., Roberts, S.B., 2010. DNA methylation patterns provide insight into epigenetic regulation in the Pacific oyster (*Crassostrea gigas*). BMC Genomics 11, 483.

60. Gerdol, M., De Moro, G., Manfrin, C., Milandri, A., Riccardi, E., Beran, A., Venier, P., Pallavicini, A., 2014. RNA sequencing and *de novo* assembly of the digestive gland transcriptome in *Mytilus galloprovincialis* fed with toxinogenic and non- toxic strains of *Alexandrium minutum*. BMC Res. Notes 7, 722.

61. Geyer, K.K., Rodríguez Lopez, C.M., Chalmers, I.W., Munshi, S.E., Truscott, M., Heald, J., Wilkinson, M.J., Hoffmann, K.F., 2011. Cytosine methylation regulates oviposition in the pathogenic blood fluke *Schistosoma mansoni*. Nat. Commun. 2, 424.

62. GIGA Community of Scientists, Bracken-Grissom, H., Collins, A.G., Collins, T., Crandall, K., Distel, D., Dunn, C., Giribet, G., Haddock, S., Knowlton, N., Martindale, M., Medina, M., Messing, C., O'Brien, S.J., Paulay, G., Putnam, N., Ravasi, T., Rouse, G.W., Ryan, J.F., Schulze, A., Wo€rheide, G., Adamska, M., Bailly, X., Breinholt, J., Browne, W.E., Diaz, M.C., Evans, N., Flot, J.F., Fogarty, N., Johnston, M., Kamel, B., Kawahara, A.Y., Laberge, T., Lavrov, D., Michonneau, F., Moroz, L.L., Oakley, T., Osborne, K., Pomponi, S.A., Rhodes, A., Santos, S.R., Satch, N., Thacker, R.W., Van de Peer, Y., Voolstra, C.R., Welch, D.M., Winston, J., Zhou, X., 2014. The global invertebrate genomics alliance (GIGA): developing community resources to study diverse invertebrate genomes. J. Hered. 105, 1-18.

63. Goodall, C.P., Bender, R.C., Brooks, J.K., Bayne, C.J., 2006. *Biomphalaria glabrata* cytosolic copper/zinc superoxide dismutase (SOD1) gene: association of SOD1 alleles

with resistance/susceptible to *Schistosoma mansoni*. Mol. Biochem. Parasitol 147, 207-210.

64. Goodson, M.S., Kojadinovic, M., Troll, J.V., Scheetz, T.E., Casavant, T.L., Soares, M.B., McFall-Ngai, M.J., 2005. Identifying components of the NF-kappaB pathway in the beneficial *Euprymna scolopes*-*Vibrio fischeri* light organ symbiosis. Appl. Environ. Microbiol. 71, 6934-6946.

65. Gorbushin, A.M., Panchin, Y.V., Iakovleva, N.V., 2010. In search of the origin of FREPs: characterization of *Aplysia californica* fibrinogen-related proteins. Dev. Comp. Immunol. 34, 465-473.

66. Gorbushin, A.M., Iakovleva, N.V., 2011. A new family of single fibrinogen domain lectins in *Mytilus*. Fish Shellfish Immunol. 30, 434-438.

67. Gorbushin, A.M., Borisova, E.A., 2015. Lectin-like molecules in transcriptome of *Littorina littorea* hemocytes. Dev. Comp. Immunol. 48, 210-220.

68. Götz, S., Arnold, R., Sebastian-Leon, P., Martín-Rodríguez, S., Tischler, P., Jehl, M.A., Dopazo, J., Rattei, T., Conesa, A., 2011. B2G-FAR, a species-centered GO annotation repository. Bioinformatics 27, 919-924.

69. Granath Jr., W.O., Yoshino, T.P., 1983. Lysosomal enzyme activities in susceptible and refractory strains of *Biomphalaria glabrata* during the course of infection with *Schistosoma mansoni*. J. Parasitol. 69, 1018-1026.

70. Granath Jr., W.O., Connors, V.A., Tarleton, R.L., 1994. Interleukin 1 activity in haemolymph from strains of the snail *Biomphalaria glabrata* varying in susceptibility to the human blood fluke, *Schistosoma mansoni*: presence, differential expression, and biological function. Cytokine 6, 21-27.

71. Gregory, T.R., 2017. Animal Genome Size Database. http://www.genomesize.com.

72. Gueguen, Y., Cadoret, J.P., Flament, D., Barreau-Roumiguiere, C., Girardot, A.L., Garnier, J., Hoareau, A., Bachere, E., Escoubas, J.M., 2003. Immune gene discovery by expressed sequence tags generated from hemocytes of the bacteria-challenged oyster, *Crassostrea gigas*. Gene 303, 139-145.

73. Guo, X., He, Y., Zhang, L., Lelong, C., Jouaux, A., 2015. Immune and stress responses in oysters with insights on adaptation. Fish Shellfish Immunol. 46, 107-119.

74. Hahn, U.K., Fryer, S.E., Bayne, C.J., 1996. An invertebrate (Molluscan) plasma protein that binds to vertebrate immunoglobulins and its potential for yielding false- positives in antibody-based detection systems. Dev. Comp. Immunol. 20, 39-50.

75. Hanelt, B., Lun, C.M., Adema, C.M., 2008. Comparative ORESTES-sampling of transcriptomes of immune-challenged *Biomphalaria glabrata* snails. J. Invertebr. Pathol 99, 192-203.

76. Hanington, P.C., Forys, M.A., Dragoo, J.W., Zhang, S.M., Adema, C.M., Loker, E.S., 2010. Role for a somatically diversified lectin in resistance of an invertebrate to parasite infection. Proc. Natl. Acad. Sci. U. S. A. 107, 21087-21092.

77. Hanington, P.C., Forys, M.A., Loker, E.S., 2012. A somatically diversified defense factor, FREP3, is a determinant of snail resistance to schistosome infection. PLoS Negl. Trop. Dis. 6, e1591.

78. Hauton, C., Smith, V., 2007. Adaptive immunity in invertebrates: a straw house without mechanistic foundation. Bioessays 29, 1138-1146.

79. He, Y., Jouaux, A., Ford, S.E., Lelong, C., Sourdaine, P., Mathieu, M., Guo, X., 2015. Transcriptome analysis reveals strong and complex antiviral response in a mollusc. Fish Shellfish Immunol. 46, 131-144.

80. Hoek, R.M., Smit, A.B., Frings, H., Vink, J.M., de Jong-Brink, M., Geraerts, W.P., 1996. A new Ig-superfamily member, molluscan defence molecule (MDM) from *Lymnaea stagnalis*, is down-regulated during parasitosis. Eur. J. Immunol. 26, 939-944.

81. Hou, R., Bao, Z., Wang, S., Su, H., Li, Y., Du, H., Hu, J., Wang, S., Hu, X., 2011. Transcriptome sequencing and *de novo* analysis for Yesso scallop (*Patinopecten yessoensis*) using 454 GS FLX. PLoS One 6, e21560.

82. Hu, F., Pan, L., Cai, Y., Liu, T., Jin, Q., 2015. Deep sequencing of the scallop *Chlamys farreri* transcriptome response to tetrabromobisphenol A (TBBPA) stress. Mar. Genomics 31-38.

83. Huang, C., Wu, W., 2015. Genomic resources of two land snails, *Aegista diversifamilia* and *Dolicheulota formoensis*, generated by Illumina paired-end sequencing [version 1; referees: 1 approved].F1000Research. 4, 106.

84. Hughes, T.K., Smith, E.M., Chin, R., Cadet, P., Sinisterra, J., Leung, M.K., Shipp, M.A., Scharrer, B., Stefano, G.B., 1990. Interaction of immunoreactive monokines (interleukin 1 and tumor necrosis factor) in the bivalve mollusc *Mytilus edulis*. Proc. Natl. Acad. Sci. U. S. A. 87, 4426-4429.

85. Hughes, T.K., Smith, E.M., Barnett, J.A., Charles, R., Stefano, G.B., 1991. LPS stimulated invertebrate hemocytes: a role for immunoreactive TNF and IL-1. Dev. Comp. Immunol. 15, 117-122.

86. Hughes, T.K., Smith, E.M., Leung, M.K., Stefano, G.B., 1992. Evidence for the conservation of an immunoreactive monokine network in invertebrates. Ann. N. Y. Acad. Sci. 650, 74-80.

87. Humphries, J.E., Elizondo, L., Yoshino, T.P., 2001. Protein kinase C regulation of cell spreading in the molluscan *Biomphalaria glabrata* embryonic (Bge) cell line. Biochim. Biophys. Acta 1540, 243-252.

88. Humphries, J., Harter, B., 2015. Identification of nuclear factor kappaB (NF-KB) binding motifs in *Biomphalaria glabrata*. Dev. Comp. Immunol. 53, 366-370.

89. Itoh, N., Takahashi, K.G., 2008. Distribution of multiple peptidoglycan recognition proteins in the tissues of Pacific oyster, *Crassostrea gigas*. Comp. Biochem. Physiol. B Miochem. Mol. Biol. 150, 409-417.

90. Ittiprasert, W., Miller, A., Su, X.Z., Mu, J., Bhusudsawang, G., Ukoskit, K., Knight, M., 2013. Identification and characterisation of functional expressed sequence tags-derived simple sequence repeat (eSSR) markers for genetic linkage mapping of *Schistosoma mansoni* juvenile resistance and susceptibility loci in *Biomphalaria glabrata*. Int. J. Parasitol. 43, 669-677.

91. Jenny, M.J., Chapman, R.W., Mancia, A., Chen, Y.A., McKillen, D.J., Trent, H., Lang, P., Escoubas, J.M., Bachere, E., Boulo, V., Liu, Z.J., Gross, P.S., Cunningham, C., Cupit, P.M., Tanguy, A., Guo, X., Moraga, D., Boutet, I., Huvet, A., De Guise, S., Almeida, J.S., Warr, G.W., 2007. A cDNA microarray for *Crassostrea virginica* and *C. gigas*. Mar. Biotechnol. 9, 577-591.

92. Jeong, K.H., Lie, J.K., Heyneman, D., 1983. The ultrastructure of the amebocyte-producing organ in *Biomphalaria glabrata*. Dev. Comp. Immunol. 7, 217-228.

93. Jia, X., Zou, Z., Wang, G., Wang, S., Wang, Y., Zhang, Z., 2011. Gene expression profiling in respond to TBT exposure in small abalone *Haliotis diversicolor*. Fish Shellfish Immunol. 31, 557-563.

94. Jiang, Y., Loker, E.S., Zhang, S.M., 2006. *In vivo* and *in vitro* knockdown of FREP2 gene expression in the snail *Biomphalaria glabrata* using RNA interference. Dev. Comp. Immunol. 30, 855-866.

95. Kang, Y.S., Kim, Y.M., Park, K.I., Kim Cho, S., Choi, K.S., Cho, M., 2006. Analysis of EST and lectin expressions in hemocytes of Manila clams (*Ruditapes philippinarum*) (Bivalvia: Mollusca) infected with *Perkinsus olseni*. Dev. Comp. Immunol. 30, 1119-1131.

96. Kasthuri, S.R., Premachandra, H.K., Umasuthan, N., Whang, I., Lee, J., 2013. Structural characterization and expression analysis of a beta-thymosin homologue (Tb) in disk abalone, *Haliotis discus discus*. Gene 527, 376-383.

97. Kaur, S., Jobling, S., Jones, C.S., NOble, L.R., Routledge, E.J., Lockyer, A.E., 2015. The nuclear receptors of *Biomphalaria glabrata* and *Lottia gigantea*: implications for developing new model organisms. PLoS One 10, e0121259.

98. Kenny, N.J., Truchado-García, M., Grande, C., 2016. Deep, multi-stage transcriptome of the schistosomiasis vector *Biomphalaria glabrata* provides platform for understanding molluscan disease-related pathways. BMC Infect. Dis. 16, 618.

99. Klein, J., 1989. Are invertebrate capable of anticipatory immune responses? Scand. J. Immunol. 29, 499-505.

100. Knight, M., Miller, A.N., Patterson, C.N., Rowe, C.G., Michaels, G., Carr, D., Richards, C.S., Lewis, F.A., 1999. The identification of markers segregating with resistance to *Schistosoma mansoni* infection in the snail *Biomphalaria glabrata*. Proc. Natl. Acad. Sci. U. S. A. 96, 1510-1515.

101. Knight, M., Ittiprasert, W., Odoemelam, E.C., Adema, C.M., Miller, A., Raghavan, N., Bridger, J.M., 2011. Non-random organization of the *Biomphalaria glabrata* genome in interphase Bge cells and the spatial repositioning of activated genes in cells co-cultured with *Schistosoma mansoni*. Int. J. Parasitol. 41, 61-70.

102. Kocot, K.M., Cannon, J.T., Todt, C., Citarella, M.R., Kohn, A.B., Meyer, A., Santos, S.R., Schandler, C., Moroz, L.L., Lieb, B., Halanych, K.M., 2011. Phylogenomics reveals deep molluscan relationships. Nature 477, 452-456.

103. Kocot, K.M., Jeffrey, N.W., Mulligan, K., Halanych, K.M., Gregory, T.R., 2015. Genome size estimates for Aplacophora, Polyplacophora, and Scaphopoda: small solengasters and sizeable scaphopods. J. Mollus. Stud. 82, 216-219.

104. Korneev, S.A., Kemenes, I., Straub, V., Staras, K., Korneeva, E.I., Kemenes, G., Benjamin, P.R., O'Shea, M., 2002. Suppression of nitric oxide (NO)-dependent behavior by double-stranded RNA-mediated silencing of a neuronal NO synthase gene. J. Neurosci. 22, RC227.

105. Kurachi, S., Song, Z., Takagaki, M., Yang, Q., Winter, H.C., Kurachi, K., Goldstein, I.J., 1998. Sialic-acid-binding lectin from the slug *Limax flavus*-cloning, expression of the polypeptide, and tissue localization. Eur. J. Biochem. 254, 217-222.

106. La Peyre, J.F., Chu, F.L., Vogelbein, W.K., 1995. *In vitro* interaction of *Perkinsus marinus* merozoites with eastern and Pacific oyster hemocytes. Dev. Comp. Immunol. 19, 291-304.

107. Leippe, M., Renwrantz, L., 1988. Release of cytotoxic and agglutinating molecules by *Mytilus* hemocytes. Dev. Comp. Immunol. 12, 297-308.

108. Leite, R.B., Milan, M., Coppe, A., Bortoluzzi, S., dos Anjos, A., Reinhart, R., Saavedra, C., Patarnello, T., Cancela, M.L., Bargelloni, L., 2013. mRNA-Seq and microarray development for the Grooved Carpet shell clam, *Ruditapes decussatus*: a functional approach to unravel host-parasite interaction. BMC Geno- mics 14, 741.

109. Lemaitre, B., Nicholas, E., Michaut, L., Reichart, J.M., Hoffmann, J.A., 1996. The dorsoventral regulatory gene cassette spätzle/Toll/cactus controls the potent antifungal response in *Drosophila* adults. Cell 86, 973-983.

110. Leonard, P.M., Adema, C.M., Zhang, S.M., Loker, E.S., 2001. Structure of two FREP genes that combine IgSF and fibrinogen domains, with comments on diversity of the FREP gene family in the snail *Biomphalaria glabrata*. Gene 269, 155-165.

111. Liao, Z., Wang, X.C., Liu, H.H., Fan, M.H., Sun, J.J., Shen, W., 2013. Molecular characterization of a novel antimicrobial peptide from *Mytilus coruscus*. Fish Shellfish Immunol. 34, 610-616.

112. Lie, K.J., Heyneman, D., 1979. Acquired resistance to echinostomes in four *Biomphalaria glabrata* strains. Int. J. Parasitol. 9, 533-537.

113. Little, T.J., Colegrave, N., Sadd, B.M., Schmid-Hempel, P., 2008. Studying immunity at the whole organism level. Bioessays 30, 404-405.

114. Liu, L., Yang, J., Qiu, L., Wang, L., Zhang, H., Wang, M., Vinu, S.S., Song, L., 2011. A novel scavenger receptor-cysteine-rich (SRCR) domain containing scavenger receptor identified from mollusk mediated PAMP recognition and binding. Dev. Comp. Immunol. 35, 227-239.

115. Lockyer, A.E., Spinks, J., Noble, L.R., Rollinson, D., Jones, C.S., 2007a. Identification of genes involved in interactions between *Biomphalaria glabrata* and *Schistosoma mansoni* by suppression subtractive hybridization. Mol. Biochem. Parasitol 151, 18-27.

116. Lockyer, A.E., Spinks, J.N., Walker, A.J., Kane, R.A., Noble, L.R., Rollinson, D., Dias-Neto, E., Jones, C.S., 2007b. *Biomphalaria glabrata* transcriptome: identification of cell-signalling, transcriptional control and immune-related genes from open reading frame expressed sequence tags (ORESTES). Dev. Comp. Immunol. 31, 763-782.

117. Lockyer, A.E., Spinks, J., Kane, R.A., Hoffmann, K.F., Fitzpatrick, J.M., Rollinson, D., Noble, L.R., Jones, C.S., 2008. *Biomphalaria glabrata* transcriptome: cDNA microarray profiling identifies resistant- and susceptible-specific gene expression in haemocytes from snail strains exposed to *Schistosoma mansoni*. BMC Genomics 9, 634.

118. Lockyer, A.E., Emery, A.M., Kane, R.A., Walker, A.J., Mayer, C.D., Mitta, G., Coustau, C., Adema, C.M., Hanelt, B., Rollinson, D., Noble, L.R., Jones, C.S., 2012. Early differential gene expression in haemocytes from resistant and susceptible *Biomphalaria glabrata* strains in response to *Schistosoma mansoni*. PLoS One 7, e51102.

119. Mansour, M.H., 1995. Evidence for a family of schistosome glycan-binding lectins in *Biomphalaria alexandrina*. Dev. Comp. Immunol. 19, 365-376.

120. Marchalonis, J.J., Schluter, S.F., 1990. Origins of immunoglobulins and immune recognition molecules. Bioscience 40, 758-768.

121. McDowell, I.C., Nikapitiya, C., Aguiar, D., Lane, C.E., Istrail, S., Gomez-Chiarri, M., 2014. Transcriptome of american oysters, *Crassostrea virginica*, in response to bacterial challenge: insights into potential mechanisms of disease resistance. PLoS One 9, e105097.

122. McDowell, I.C., Modak, T.H., Lane, C.E., Gomez-Chiarri, M., 2016. Multi-species protein similarity clustering reveals novel expanded immune gene families in the eastern oyster *Crassostrea virginica*. Fish Shellfish Immunol. 53, 13-23.

123. McKerrow, J.H., Jeong, K.H., Beckstead, J.H., 1985. Enzyme histochemical comparison of *Biomphalaria glabrata* amebocytes with human granuloma macrophages. J. Leukoc. Biol. 37, 341-347.

124. Medzhitov, R., Preston-Hurlburt, P., Janeway, C.A., 1997. A human homologue of the *Drosophila* Toll protein signals activation of adaptive immunity. Nature 388, 394-397.

125. Metchnikoff, E., 1905. Immunity in Infective Disease. Cambridge University Press, Cambridge.

126. Michelson, E.H., Dubois, L., 1977. Agglutinins and lysins in the molluscan family Planorbidae: a survey of hemolymph, egg-masses, and albumen-gland extracts. Biol. Bull. 153, 219-227.

127. Miller, A.N., Raghavan, N., FitzGerald, P.C., Lewis, F.A., Knight, M., 2001. Differential gene expression in haemocytes of the snail *Biomphalaria glabrata*: effects of *Schistosoma mansoni* infection. Int. J. Parasitol. 31, 687-696.

128. Mitta, G., Galinier, R., Tisseyre, P., Allienne, J.F., Girerd-Chambaz, Y., Guillou, F., Bouchut, A., Coustau, C., 2005. Gene discovery and expression analysis of immune-relevant genes from *Biomphalaria glabrata* hemocytes. Dev. Comp. Immunol. 29, 393-407.

129. Mohandas, A., Cheng, T.C., Cheng, J.B., 1985. Mechanism of lysosomal enzyme release from *Mercenaria mercenaria* granulocytes: a scanning electron microscope study. J. Invertebr. Pathol 46, 189-197.

130. Monroy, F., Hertel, L.A., Loker, E.S., 1992. Carbohydrate-binding plasma proteins from the gastropod *Biomphalaria glabrata*: strain specificity and the effects of trematode infection. Dev. Comp. Immunol. 16, 355-366.

131. Mullainadhan, P., Renwrantz, L., 1986. Lectin-dependent recognition of foreign cells by hemocytes of the mussel, *Mytilus edulis*. Immunobiology 171, 263-273.

132. Mundry, M., Bornberg-Bauer, E., Sammeth, M., Feulner, P.G., 2012. Evaluating characteristics of *de novo* assembly software on 454 transcriptome data: a simulation approach. PLoS One 7, e31410.

133. Nam, B.H., Jung, M., Subramaniyam, S., Yoo, S.I., Markkandan, K., Moon, J.Y., Kim, Y.O., Kim, D.G., An, C.M., Shin, Y., Jung, H.J., Park, J.H., 2016. Transcriptome analysis revealed changes of multiple genes involved in *Haliotis discus hannai* innate immunity during *Vibrio parahemolyticus* infection. PLoS One 11, e0153474.

134. Noël, D., Bachere, E., Mialhe, E., 1993. Phagocytosis associated chemiluminescence of hemocytes in *Mytilus edulis* (Bivalvia). Dev. Comp. Immunol. 17, 483-493.

135. Nowak, T.S., Woodards, A.C., Jung, Y., Adema, C.M., Loker, E.S., 2004. Identification of transcripts generated during the response of resistant *Biomphalaria glabrata* to *Schistosoma mansoni* infection using suppression subtractive hybridization. J. Parasitol. 90, 1034-1040.

136. Nun~ez-Acun~a, G., Gallardo-Ascarate, C., 2013. Identification of immune-related SNPs in the transcriptome of *Mytilus chilensis* through high-throughput sequencing. Fish Shellfish Immunol. 35, 1899-1905.

137. Ottaviani, E., Franchini, A., Franceschi, C., 1993. Presence of several cytokine-like molecules in molluscan hemocytes. Biochem. Biophys. Res. Commun. 195, 984-988.

138. Ottaviani, E., Caselgrandi, E., Franceschi, C., 1995a. Cytokines and evolution: *In vitro* effects of IL-1, IL-1B, TNF-a and TNF-B on an ancestral type of stress response. Biochem. Biophys. Res. Commun. 207, 288-292.

139. Ottaviani, E., Franchini, A., Cassanelli, S., Genedani, S., 1995b. Cytokines and molluscan immune responses. Biol. Cell. 85, 87-91.

140. Ottaviani, E., Sassi, D., Kletsas, D., 1997. PDGF- and TGF-beta-induced changes in cell shape of invertebrate immunocytes: effect of calcium entry blockers. Eur. J. Cell Biol. 74, 336-341.

141. Ottaviani, E., Franchini, A., Malagoli, D., Genedani, S., 2000. Immunomodulation by recombinant human interleukin-8 and its signal transduction pathways in invertebrate hemocytes. Cell Mol. Life Sci. 57, 506-513.

142. Ouwe-Missi-Oukem-Boyer, O., Porchet, E., Capron, A., Dissous, C., 1994. Characterization of immunoreactive TNF alpha molecules in the gastropod *Biomphalaria glabrata*. Dev. Comp. Immunol. 18, 211-218.

143. Owen, J., Punt, J., Stranford, S., 2013. Kuby Immunology, seventh ed. W.H. Freeman and Company, New York.

144. Owens, L., Malham, S., 2015. Review of the RNA interference pathway in molluscs including some possibilities for use in bivalves in aquaculture. J. Mar. Sci. Eng. 3, 87-99.

145. Patnaik, B.B., Wang, T.H., Kang, S.W., Hwang, H.J., Park, S.Y., Park, E.B., Chung, J.M., Song, D.K., Kim, C., Kim, S., Lee, J.S., Han, Y.S., Park, H.S., Lee, Y.S., 2016. Sequencing, *de novo* assembly, and annotation of the transcriptome of the endangered freshwater pearl bivalve, *Cristaria plicata*, provides novel insights into functional genes and marker discovery. PLoS One 11, e0148622.

146. Pauletto, M., Milan, M., Moreira, R., Novoa, B., Figueras, A., Babbucci, M., Patarnello, T., Bargelloni, L., 2014. Deep transcriptome sequencing of *Pecten maximus* hemocytes: a genomic resource for bivalve immunology. Fish Shellfish Immunol. 31, 154-165.

147. Peñarrubia, L., Sanz, N., Pla, c., Vidal, O., Vin~as, J., 2015. Using massive parallel sequencing for the development, validation, and application of population genetics markers in the invasive bivalve zebra mussel (*Dreissena polymorpha*). PLoS One 10, e0120732.

148. Pettersson, E., Lundeberg, J., Ahmadian, A., 2009. Generations of sequencing technologies. Genomics 93, 105-111.

149. Philipp, E.E., Kraemer, L., Melzner, F., Poustka, A.J., Thieme, S., Findeisen, U., Schreiber, S., Rosenstiel, P., 2012. Massively parallel RNA sequencing identifies a complex immune gene repertoire in the lophotrochozoan *Mytilus edulis*. PLoS One 7, e33091.

150. Pinaud, S., Portela, J., Duval, D., Nowacki, F.C., Olive, M.A., Allienne, J.F., galinier, R., Dheilly, N.M., Kieffer-Jaquinod, S., Mitta, G., Theron, A., Gourbal, B., 2016. A shift from cellular to humoral responses contributes to innate immune memory in the vector snail *Biomphalaria glabrata*. PLoS Pathog. 12, e1005361.

151. Premachandra, H.K., Wan, Q., Elvitigala, D.A., De Zoysa, M., Choi, C.Y., Whang, I., Lee, J., 2012. Genomic characterization and expression profiles upon bacterial infection of a novel cystatin B homologue from disk abalone (*Haliotis discus discus*). Dev. Comp. Immunol. 38, 495-504.

152. Prokop, O., Uhlenbruck, G., Ko€hler, W., 1968. A new source of antibody-like substances having anti-blood group specificity. A discussion on the specificity of *Helix* agglutinins. Vox. Sang. 14, 321-333.

153. Raghavan, N., Miller, A.N., Gardner, M., FitzGerald, P.C., Kerlavage, A.R., Johnston, D.A., Lewis, F.A., Knight, M., 2003. Comparative gene analysis of *Biomphalaria*

glabrata hemocytes pre- and post-exposure to miracidia of *Schistosoma mansoni*. Mol. Biochem. Parasitol 126, 181-191.

154. Raghavan, N., Tettelin, H., Miller, A., Hostetler, J., Tallon, L., Knight, M., 2007. Nimbus (BgI): an active non-LTR retrotransposon of the *Schistosoma mansoni* snail host *Biomphalaria glabrata*. Int. J. Parasitol. 37, 1307-1318.

155. Renwrantz, L.R., Cheng, T.C., 1977. Agglutinin-mediated attachment erythrocytes to hemocytes of *Helix pomatia*. J. Invertebr. Pathol 29, 97-100.

156. Renwrantz, L.R., Richards, E.H., 1992. Recognition of beta-glucuronidase by the calcium-independent phosphomannosyl surface receptor of haemocytes from the gastropod mollusc, *Helix pomatia*. Dev. Comp. Immunol. 16, 251-256.

157. Rhee, J.S., Kim, B.M., Jeong, C.B., Horiguchi, T., Lee, Y.M., Kim, I.C., Lee, J.S., 2012. Immune gene mining by pyrosequencing in the rockshell, *Thais clavigera*. Fish Shellfish Immunol. 32, 700-710.

158. Rhoads, A., Au, K.F., 2015. PacBio sequencing and its applications. Genomics Proteomics Bioinforma. 13, 278-289.

159. Richards, E.H., Renwrantz, L.R., 1991. Two lectins on the surface of *Helix pomatia* haemocytes: a Ca^{2+}-dependent, GalNac-specific lectin and a Ca^{2+}-independent, mannose 6-phosphate-specific lectin which recognises activated homologous opsonins. J. Comp. Physiol. B 161, 43-54.

160. Richards, C.S., Knight, M., Lewis, F.A., 1992. Genetics of *Biomphalaria glabrata* and its effect on the outcome of *Schistosoma mansoni* infection. Parasitol. Today 8, 171-174.

161. Roberts, S.B., Goetz, F.W., 2003. Expressed sequence tag analysis of genes expressed in the bay scallop, *Argopecten irradians*. Biol. Bull. 205, 227-228.

162. Rögener, W., Renwrantz, L., Uhlenbruck, G., 1985. Isolation and characterization of a lectin from the hemolymph of the cephalopod *Octopus vulgaris* (Lam.) inhibited by alpha-D-lactose and N-acetyl-lactosamine. Dev. Comp. Immunol. 9, 605-616.

163. Romero, A., Dios, S., Poisa-Beiro, L., Costa, M.M., Posada, D., Figueras, A., Novoa, B., 2011. Individual sequence variability and functional activities of fibrinogen-related proteins (FREPs) in the Mediterranean mussel (*Mytilus galloprovincialis*) suggest ancient and complex immune recognition models in invertebrates. Dev. Comp. Immunol. 35, 334-344.

164. Romero, A., Forn-Cuní, G., Moreira, R., Milan, M., Bargelloni, L., Figueras, A., Novoa, B., 2015. An immune-enriched oligo-microarray analysis of gene expression in Manila clam (*Venerupis philippinarum*) haemocytes after a *Perkinsus olseni* challenge, 43, 275-286.

165. Romestand, B., Corbier, F., Roch, P., 2002. Protease inhibitors and haemagglutinins associated with resistance to the protozoan parasite, *Perkinsus marinus*, in the Pacific oyster, *Crassostrea gigas*. Parasitology 125, 323-329.

166. Rosani, U., Varotto, L., Gerdol, M., Pallavicini, A., Venier, P., 2015a. IL-17 signaling components in bivalves: comparative sequence analysis and involvement in the immune responses. Dev. Comp. Immunol. 52, 255-268.

167. Rosani, U., Varotto, L., Domeneghetti, S., Arcangeli, G., Pallavicini, A., Venier, P., 2015b. Dual analysis of host and pathogen transcriptomes in ostreid herpes-virus 1-positive *Crassostrea gigas*. Environ. Microbiol. 17, 4200-4212.

168. Safavi-Hemami, H., Li, Q., Jackson, R.L., Song, A.S., Boomsma, W., Bandyopadhyay, P.K., Gruber, C.W., Purcell, A.W., Yandell, M., Olivera, B.M., Ellgaard, L., 2016. Rapid expansion of the protein disulfide isomerase gene family facilitates the folding of venom peptides. Proc. Natl. Acad. Sci. U. S. A. 113, 3227-3232.

169. Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evo 4, 406-425.

170. Salazar, K.A., Joffe, N.R., Dinguirard, N., Houde, P., Castillo, M.G., 2015. Transcriptome analysis of the white body of the squid *Euprymna tasmanica*. PLoS One 10, e0119949.

171. Schadt, E.E., Turner, Kasarskis, A., 2010. A window into third-generation sequencing. Hum. Mol. Genet. 19, R227-R240.

172. Schneider, O., Zelck, U.E., 2001. Differential display analysis of hemocytes from schistosome-resistant and schistosome-susceptible intermediate hosts. Parasitol. Res. 87, 489-491.

173. Seo, J.K., Go, H.J., Kim, C.H., Nam, B.H., Park, N.G., 2016. Antimicrobial peptide, hdMolluscidin, purified from the gill of the abalone, *Haliotis discus*. Fish Shellfish Immunol. 52, 289-297.

174. Simakov, O., Marletaz, F., Cho, S.J., Edsinger-Gonzales, E., Havlak, P., Hellsten, U., Kuo, D.H., Larsson, T., Lv, J., Arendt, D., Savage, R., Osoegawa, K., de Jong, P., Grimwood, J., Chapman, J.A., Shapiro, H., Aerts, A., Otillar, R.P., terry, A.Y., Boore, J.L., Grigoriev, I.V., Lindberg, D.R., Seaver, E.C., Weisblat, D.A., Putnam, N.H., Rokhsar, D.S., 2013. Insights into bilaterian evolution from three spiralian genomes. Nature 493, 526-531.

175. Singh, V., Braddick, D., Dhar, P.K., 2017. Exploring the potential of genome editing CRISPR-Cas9 technology. Gene 599, 1-18.

176. Smith, S.A., Wilson, N.G., Goetz, F.E., Feehery, C., Andrade, S.C.S., Rouse, G.W., Giribet, G., Dunn, C.W., 2011. Resolving the evolutionary relationships of molluscs with phylogenomic tools. Nature 480, 364-367.

177. Smith, D.R., 2016 Jun 23. Goodbye genome paper, hello genome report: the increasing popularity of 'genome announcements' and their impact on science. Brief. Funct. Genomics pii: elw026.

178. Sohn, J.I., Nam, J.W., 2016 Oct 14. The present and future of *de novo* whole-genome assembly. Brief. Bioinform. pii: bbw096.

179. Song, L., Xu, W., Li, C., Li, H., Wu, L., Xiang, J., Guo, X., 2006. Development of expressed sequence tags from the bay scallop, *Argopecten irradians irradians*. Mar. Biotechnol. (NY) 8, 161-169.

180. Steelman, B.N., Connors, V.A., 2009. Chemokinetic effect of interleukin-1 beta on cultured *Biomphalaria glabrata* embryonic cells. J. Parasitol. 95, 772-774.

181. Stefano, G.B., Smith, E.M., Hughes, T.K., 1991. Opioid induction of immunoreactive interleukin-1 in *Mytilus edulis* and human immunocytes: an interleukin-1-like substance in invertebrate neural tissue. J. Neuroimmunol. 32, 29-34.

182. Tamura, K., Stecher, G., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evo 30, 2725-2729.

183. Tebeje, B.M., Harvie, M., You, H., Loukas, A., McManus, D.P., 2016. Schistosomiasis vaccines: where do we stand? Parasit. Vectors 9, 528.

184. Tennessen, J.A., Theron,A., Marine, M., Yeh, J.Y., Rognon, A., Blouin, M.S., 2015a. Hyperdiverse gene cluster in snail host conveys resistance to human schistosome parasites. PLoS Genet. 11, e1005067.

185. Tennessen, J.A., Bonner, K.M., Bollmann, S.R., Johnstun, J.A., Yeh, J.Y., Marine, M., Tavalire, H.F., Bayne, C.J., Blouin, M.S., 2015b. Genome-wide scan and test of candidate genes in the snail *Biomphalaria glabrata* reveal new locus influencing resistance to *Schistosoma mansoni*. PLoS. Negl. Trop. Dis. 9, e0004077.

186. Tripp, M.R., Kent, V.E., 1967. Studies on oyster cellular immunity. In Vitro 3, 129-135.

187. van der Knaap, W.P.W., Sminia, T., Kroese, F.G.M., Dikkeboom, R., 1981. Elimination of bacteria from the circulation of the pond snail *Lymnaea stagnalis*. Dev. Comp. Immunol. 5, 21-32.

188. van der Knaap, W.P.W., Boots, A.M., van Asselt, L.A., Sminia, T., 1983a. Specificity and memory in increased defence reactions against bacteria in the pond snail *Lymnaea stagnalis*. Dev. Comp. Immunol. 7, 435-443.

189. van der Knaap, W.P.W., Sminia, T., Schutte, R., Boerrigter-Barendsen, L.H., 1983b. Cytophilic receptors for foreignness and some factors which influence phagocytosis by invertebrate leucocytes: in vitro phagocytosis by amoebocytes of the snail *Lymnaea stagnalis*. Immunology 48, 377-383.

190. van der Knaap, W.P.W., Loker, E.S., 1990. Immune mechanisms in trematode-snail interactions. Parasitol. Today 6, 175-182.

191. Vasta, G.R., Ahmed, H., 2009. Introduction to animal lectins. In: Vasta, G.R., Ahmed, H. (Eds.), Animal Lectins: a Functional View. CRC Press, Boca Raton, pp. 3-9.

192. Venier, P., Varotto, L., Rosani, U., Millino, C., Celegato, B., Bernante, F., Lanfranchi, G., Novoa, B., Roch, P., Figueras, A., Pallavicini, A., 2011. Insights into the innate immunity of the Mediterranean mussel *Mytilus galloprovincialis*. BMC Genomics 12, 69.

193. Walker, L., Levine, H., Jucker, M., 2006. Koch's postulates and infectious proteins. Acta Neuropathol. 112, 1-4.

194. Walker, A.J., Lacchini, A.H., Sealey, K.L., Mackintosh, D., Davies, A.J., 2010. Spreading by snail (*Lymnaea stagnalis*) defence cells is regulated through integrated PKC, FAK, and Src signaling. Cell Tissue Res. 34, 131-145.

195. Wang, H., Song, L., Li, C., Zhao, J., Zhang, H., Ni, D., Xu, W., 2007. Cloning and characterization of a novel C-type lectin from Zhikong scallop *Chlamys farreri*. Mol. Immunol. 44, 722-731.

196. Wang, L., Song, L., Zhao, J., Qiu, L., Zhang, H., Xu, W., Li, H., Li, C., Wu, L., Guo, X., 2009. Expressed sequence tags from the zhikong scallop (*Chlamys farreri*): discovery and annotation of host-defense genes. Fish Shellfish Immunol. 26, 744-750.

197. Wang, S., Peatman, E., Liu, H., Bushek, D., Ford, S.E., Kucuktas, H., Quilang, J., Li, P., Wallace, R., Wang, Y., Guo, X., Liu, Z., 2010. Microarray analysis of gene expression in eastern oyster (*Crassostrea virginica*) reveals a novel combination of antimicrobial and oxidative stress host responses after derm (*Perkinsus marinus*) challenge. Fish Shellfish Immunol. 29, 921-929.

198. Wang, K., Pales Espinosa, E., Tanguy, A., Allam, B., 2016. Alterations of the immune transcriptome in resistant and susceptible hard clams (*Mercenaria mercenaria*) in response to Quahog Parasite Unknown (QPX) and temperature. Fish Shellfish Immunol. 49, 163-176.

199. Wang, K., del Castillo, C., Corre, E., Pales Espinosa, E., Allam, B., 2016b. Clam focal and systemic immune responses to QPX infection revealed by RNA-seq technology. BMC Genomics 17, 146.

200. Wang, Z., Wang, B., Chen, G., Jian, J., Lu, Y., Xu, Y., Wu, Z., 2016c. Transcriptome analysis of the pearl oyster (*Pinctada fucata*) hemocytes in response to *Vibrio alginolyticus* infection. Gene 575, 421-428.

201. Warr, G.W., 1981. Immunity in invertebrates. J. Invert. Pathol 38, 311-314.

202. Wei, X., Yang, J., Yang, D., Xu, J., Liu, X., Yang, J., Fang, J., Qiao, H., 2012. Molecular cloning and mRNA expression of two peptidoglycan recognition protein (PGRP) genes from mollusk *Solen grandis*. Fish Shellfish Immunol. 32, 178-185.

203. Williams, A.F., Gagnon, J., 1982. Neuronal cell Thy-1 glycoprotein: homology with immunoglobulin. Science 216, 696-703.

204. Wong, Y.H., Sun, J., He, L.S., Chen, L.G., Qiu, J.W., Qian, P.Y., 2015. High-throughput transcriptome sequencing of the cold seep mussel *Bathymodiolous platifrons*. Sci. Rep. 5, 16597.

205. You, Y., Huan, P., Liu, B., 2012. RNAi assay in primary cells: a new method for gene function analysis in marine bivalve. Mol. Biol. Rep. 39, 8209-8216.

206. Yuasa, H.J., Furuta, E., Nakamura, A., Takagi, T., 1998. Cloning and sequencing of three C-type lectins from body surface mucus of the land slug, *Incilaria fruhstorferi*. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 119, 479-484.

207. Zahoor, Z., Davies, A.J., Kirk, R.S., Rollinson, D., Walker, A.J., 2008. Disruption of ERK signalling in *Biomphalaria glabrata* defence cells by *Schistosoma mansoni*: implications for parasite survival in the snail host. Dev. Comp. Immunol. 32, 1561-1571.

208. Zahoor, Z., Davies, A.J., Kirk, R.S., Rollinson, D., Walker, A.J., 2009. Nitric oxide production by *Biomphalaria glabrata* haemocytes: effects of *Schistosoma mansoni* ESPs and regulation through the extracellular signal-regulated kinase pathway. Parasit. Vectors 2, 18.

209. Zahoor, Z., Lockyer, A.E., Davies, A.J., Kirk, R.S., Emery, A.M., Rollinson, D., Jones, C.S., Noble, L.R., Walker, A.J., 2014. Differences in the gene expression profiles of haemocytes from schistosome-susceptible and -resistant *Biomphalaria glabrata* exposed to *Schistosoma mansoni* excretory-secretory products. PLoS One 9, e93215.

210. Zelck, U.E., Gege, B.E., Schmid, S., 2007. Specific inhibitors of mitogen-activated protein kinase and P13-K pathways impair immune responses by hemocytes of trematode intermediate host snails. Dev. Comp. Immunol. 31, 321-331.

211. Zhang, S.M., Leonard, P.M., Adema, C.M., Loker, E.S., 2001. Parasite-responsive IgSF members in the snail *Biomphalaria glabrata*: characterization of novel genes with tandemly arranged IgSF domains and a fibrinogen domain. Immunogenetics 53, 992.

212. Zhang, S.M., Adema, C.M., Kepler, T.B., Loker, E.S., 2004. Diversification of Ig superfamily genes in an invertebrate. Science 305, 251-254.

213. Zhang, Y., Zhang, X., Scheuring, C.F., Zhang, H.B., Huan, P., Li, F., Xiang, J., 2008. Construction and characterization of two bacterial artificial chromosome libraries of Zhikong scallop, *Chlamys farreri* Jones et Preston, and identification of BAC clones containing the genes involved in its innate immune system. Mar. Biotechnol. (NY) 10, 358-365.

214. Zhang, G., Fang, X., Guo, X., Li, L., Luo, R., Xu, F., Yang, P., Zhang, L., Wang, X., qi, H., Xiong, Z., Que, H., Xie, Y., Holland, P.W., Paps, J., Zhu, Y., Wu, F., Chen, Y., Wang, J., Peng, C., Meng, J., Yang, L., Liu, J., Wen, B., Zhang, N., Huang, Z., Zhu, Q., Feng, Y., Mount, A., Hedgecock, D., Xu, Z., Liu, Y., Domazet-Loso, T., Du, Y., Sun, X., Zhang, S., Liu, B., Cheng, P., Jiang, X., Li, J., Fan, D., Huang, R., Yang, H., Du, X., Chen, L., Yang, M., Gaffney, P.M., Wang, S., Luo, L., She, Z., Ming, Y., Huang, W., Zhang, S., Huang, B., Zhang, Y., Qu, T., Ni, P., Miao, G., Wang, J., Wang, Q., Steinberg, C.E., Wang, H., Li, N., Qian, L., Zhang, G., Li, Y., Yang, H., Liu, X., Wang, J., Yin, Y., Wang, J., 2012. The oyster genome reveals stress adaptation and complexity of shell formation. Nature 409, 49-54.

215. Zhang, Y., Yu, Z., 2013. The first evidence of positive selection in peptidoglycan recognition protein (PGRP) genes of *Crassostrea gigas*. Fish Shellfish Immunol. 34, 1352-1255.

216. Zhang, L., Li, L., Zhu, Y., Zhang, G., Guo, X., 2014. Transcriptome analysis reveals a rich gene set related to innate immunity in the eastern oyster. Mar. Biotechnol. 16, 17-33.

217. Zhang, L., Li, L., Guo, X., Litman, G.W., Dishaw, L.J., Zhang, G., 2015. Massive expansion and functional divergence of innate immune genes in a protostome. Sci. Rep. 5, 8693.

218. Zhang, S.M., Loker, E.S., Sullivan, J.T., 2016. Pathogen-associated molecular patterns activate expression of genes involved in cell proliferation, immunity, and detoxification in the amebocyte-producing organ of the snail *Biomphalaria glabrata*. Dev. Comp. Immunol. 56, 25-36.

219. Zhao, C., Zhang, T., Zhang, X., Hu, S., Xiang, J., 2012a. Sequencing and analysis of four BAC clones containing innate immune genes from the Zhikong scallop (*Chlamys farreri*). Gene 502, 9-15.

220. Zhao, X., Wang, Q., Jiao, Y., Huang, R., Deng, Y., Wang, H., Du, X., 2012b. Identification of genes potentially related to biomineralization and immunity by transcriptome analysis of pearl sac in pearl oyster *Pinctada martensii*. Mar. Biotechnol. 14, 730-739.

221. Zhao, X., Yu, H., Kong, L., Liu, S., Li, Q., 2014. Comparative transcriptome analysis of two oysters, *Crassostrea gigas* and *Crassostrea hongkongensis* provides insights into adaptation to hypo-osmotic conditions. PLoS One 9, e111915.

222. Zhao, Q.P., Xiong, T., Xu, X.J., Jiang, M.S., Dong, H.F., 2015. *De novo* transcriptome analysis of *Oncomelania hupensis* after molluscicide treatment by next-generation sequencing: implications for biology and future snail interventions. PLoS One 10, e0118673.

223. Zuckerkandl, E., Pauling, L., 1965. Evolutionary divergence and convergence in proteins. In: Bryson, V., Vogel, H.J. (Eds.), Evolving Genes and Proteins. Academic Press, New York, pp. 97-166.

Chapter 3: Comparative immunological study of the snail *Physella acuta* (Hygrophila, Pulmonata) reveals shared and unique aspects of gastropod immunobiology

Schultz, J.H., Bu, L., Adema, C.M., 2018. Comparative immunological study of the snail *Physella acuta* (Hygrophila, Pulmonata) reveals shared and unique aspects of gastropod immunobiology. Mol. Immunol. 101, 108-119.

Abstract

The freshwater snail *Physella acuta* was selected to expand the perspective of comparative snail immunology. Analysis of *Physella acuta*, belonging to the Physidae, taxonomic sister family to Planorbidae, affords family-level comparison of immune features characterized from *Biomphalaria glabrata*, the model snail often used to interpret general gastropod immunity. To capture constitutive and induced immune sequences, transcriptomes of an individual *Physella acuta* snail, 12 h post injection with bacteria (Gram –/+) and one sham-exposed snail were recorded with 454 pyrosequencing. Assembly yielded a combined reference transcriptome containing 24,288 transcripts. Additionally, genomic Illumina reads were obtained (∼15-fold coverage). Recovery of transcripts for two macin-like antimicrobial peptides (AMPs), 12 aplysianins, four LBP/BPIs and three physalysins indicated that *Physella acuta* shares a similar organization of antimicrobial defenses with *Biomphalaria glabrata*, contrasting a modest AMP arsenal with a diverse set of antimicrobial proteins. The lack of predicted transmembrane domains in all seven *Physella acuta* PGRP transcripts supports the notion that gastropods do not employ cell-bound PGRP receptors, different from ecdysozoan invertebrates yet similar to mammals (vertebrate deuterostomes). The well-documented sequence diversification by *Biomphalaria glabrata* FREPs (immune lectins comprising immunoglobulin superfamily domains and fibrinogen domains), resulting from somatic mutations of a large FREP gene family is hypothesized to be unique to Planorbidae; *Physella acuta* revealed just two bonafide FREP genes and these were not diversified. Furthermore, the flatworm parasite *Echinostoma paraensei*, confirmed here to infect both snail species, did not evoke from *Physella acuta* the abundant expression of FREP proteins at 2, 4 and 8 days post exposure that was previously observed from *Biomphalaria glabrata*. The *Physella acuta* reference transcriptome also revealed 24 unique

transcripts encoding proteins consisting of a single fibrinogen-related domain (FReDs), with a short N-terminal sequence encoding either a signal peptide, transmembrane domain or no predicted features. The *Physella acuta* FReDs are candidate immune genes based on implication of similar sequences in immunity of bivalve molluscs. Overall, comparative analysis of snails of sister families elucidated the potential for taxon-specific immune features and investigation of strategically selected species will provide a more comprehensive view of gastropod immunity.

1. Introduction

Despite the considerable biological diversity of gastropods, most investigations of gastropod immunity have utilized the model organism *Biomphalaria glabrata*. In-depth characterization of *B. glabrata* as the primary representative species is begotten by its role as an intermediate host for the flatworm parasite *Schistosoma mansoni*, one of the causative agents of human schistosomiasis, a neglected tropical disease afflicting and requiring treatment for more than 200 million people each year (Egesa et al., 2017). Consequently, extensive investigations have led to a better understanding of the immune capabilities and host/parasite dynamics of *B. glabrata* and its cognate pathogens (Adema and Loker, 2015; Coustau et al., 2015; Knight et al., 2016; Yoshino et al., 2012). The recent full characterization of the *B. glabrata* genome now helps to further interpret these findings (Adema et al., 2017). Functional assays of immunity in *B. glabrata* have revealed proteins that aid in defense against pathogens, such as macrophage migration inhibitory factor (MIF), toll-like receptor (TLR), granulin, lipopolysaccharide-binding protein/bactericidal permeability-increasing protein (LBP/ BPI), and biomphalysin (Baeza Garcia et al., 2010; Baron et al., 2013; Galinier et al., 2013; Hathaway et al., 2010; Pila et al., 2016a, 2016b). Also, gene discovery efforts have yielded peptidoglycan-recognition protein (PGRP) and aplysianin as candidate defense factors of *B. glabrata* (Bayne, 2009; Bouchut et al.,

2007; Zhang et al., 2007). With the exception of aplysianin, these snail sequences are homologs of deuterostome and ecdysozoan immune factors, indicating broad commonalities within the evolution of immune function in metazoans. However, we must also consider the potential for differences and lineage-specific aspects of immunity. Transcriptomic and genomic analyses of *B. glabrata* reveal species-specific immune responses and complexities of immune gene complements (Adema et al., 2017; DeLeury et al., 2012). In addition, bivalve molluscs (sister-taxon to gastropods) display class-unique expansion and diversity of pathogen-response genes (Zhang et al., 2015). Other differences include the relatively limited repertoire of antimicrobial peptides (AMPs) recorded from *B. glabrata*. Confined to a single family of macin-like sequences, the paucity of these factors is striking when considering the diversity of AMPs identified from bivalve molluscs, such as *Mytilus galloprovincialis*, or arthropods (Leoni et al., 2017; Mitta et al., 2000; Mylonakis et al., 2016). Given the limited number of observed AMP sequences within the *B. glabrata* genome, it is reasonable to hypothesize that antimicrobial immunity is afforded by a reliance on other immune factors, such as antimicrobial proteins. In fact, Baron et al. (2016) recorded a unique expansion and specialization of LBP/BPI sequences in *B. glabrata* compared to other animal phyla. Furthermore, a genome-level survey indicated the presence of two achacin and 21 biomphalysin genes (Adema et al., 2017). Indeed, we must remain open to the potential for the discovery of new immune genes, gene-family expansions, and novel immune functions in distinct branches of phylogeny.

Biomphalaria glabrata also possesses more than 20 fibrinogen-related protein (FREP) genes within its genome (Adema et al., 2017; Dheilly et al., 2015). FREPs are carbohydrate-binding lectins that were first discovered in *B. glabrata* in response to infection with the trematode parasite *Echinostoma paraensei* (Adema et al., 1997; Lie and Basch, 1967). Further

investigations showed that FREPs play an important role in *B. glabrata* defense against *S. mansoni* and similar responses were ascribed to another planorbid snail, *Helisoma trivolvis* (Adema et al., 1999; Hanington et al., 2010). In addition, qPCR recorded differential expression of FREPs, especially after *B. glabrata* exposure to *E. paraensei* (Hertel et al., 2005). FREP genes are somatically diversified via point mutation and gene conversion (Zhang et al., 2004). The ability of *B. glabrata* to achieve this type of shotgun immunity (consider Müller et al., 2017) suggests the potential for anticipatory immunity of other lophotrochozoan invertebrates. This discovery has prompted investigators to search for FREPs in other gastropods. Gorbushin et al. (2010) investigated the FREP repertoire of the euopisthobranch *Aplysia californica*. Euopisthobranchs are an evolutionary sister-group to panpulmonates, like *B. glabrata* (Jörger et al., 2010; White et al., 2011). Two FREP sequences were identified within the genome of *A. californica*, but these were not diversified above levels determined from control genes (actin, FMRFamide) suggesting absence of somatic diversification. In addition, a search in the genome of *Lottia gigantaea*, member of the most basal gastropod clade, did not reveal any FREP sequences (Gorbushin et al., 2010). More recently, an RNAseq-based analysis of lectin-like molecules in the common periwinkle *Littorina littorea* (Caenogastropoda) revealed a single FREP gene and no sequence diversification (Gorbushin and Borisova, 2015). These analyses indicate that the somatic diversification of FREPs in *B. glabrata* is different from other taxa and may have originated after the evolutionary split of Euopisthobranchia and Panpulmonata (Gorbushin et al., 2010). Interestingly, recent work by Wu et al. (2017) identified fibrinogen-related domain-containing proteins (FReDs) as parasite-binding plasma proteins of *B. glabrata*. These sequences contain a fibrinogen domain but lack upstream immunoglobulin superfamily (IgSF) domains, a canonical feature of bonafide FREPs. Similar immune-related

sequences collectively designated as FReDs, have been recorded from *M. galloprovincialis* and these transcripts showed considerable sequence variation at inter-and intra-individual levels (Romero et al., 2011).

Comparative investigations of gastropod immunity reveal aspects of evolution among metazoans, yet such studies are limited in quantity (Schultz and Adema, 2017). It is important to determine to what extent aspects of *B. glabrata* defense, including the somatic diversification of immune molecules (FREPs), characterize gastropod immunity or represent taxon-specific capabilities. Indeed, next-generation sequencing (NGS) facilitates true comparative immunology of the extensive diversity of gastropods, especially when using strategic choices of species at relevant branches of phylogeny to investigate the applicability of our current understanding of *B. glabrata*-centered gastropod immunity (Schultz and Adema, 2017).

Physella acuta was selected to investigate taxonomic distribution of immune capabilities described from *B. glabrata*. As a member of the family Physidae (Bouchet and Rocroi, 2005), sister family to the Planorbidae, *P. acuta* is more closely related to *B. glabrata* than other molluscs for which immune-related research exists and both of these snail families belong to the taxonomic clade Hygrophila (Jörger et al., 2010; Nolan et al., 2014). Physid snails, and in particular *P. acuta*, have received considerable attention in ecotoxicology experiments, investigations of reproductive biology, and phylogenetic analyses (Camargo and Alonos, 2017; Dillon et al., 2011; Martínez-Paz et al., 2017; Nolan et al., 2014; Wethington and Lydeard, 2007). However, few reports of physid immunity exist (Grimm-Jørgensen, 1987, 1983; Guo and He, 2014; Lee et al., 2011; Tanveer, 1991).

In this study, we have employed *P. acuta* to record immune genes from the family Physidae for comparison to the immunome of *B. glabrata*. For investigation of a role of FREPs

in *P. acuta*, we re-established viable host/parasite interactions of *P. acuta* with the digenean trematode *E. paraensei*, a notion originally stated by Lie and Basch (1967). This parasite, that evokes abundant FREP expression in the plasma of *B. glabrata*, now afforded us the ability to assess FREP expression in the plasma of *E. paraensei*-infected *P. acuta* snails. We also employed next-generation sequencing (454 and Illumina) to record baseline and induced innate immune defenses of *P. acuta* snails, to gain a comprehensive view of the extent of known antimicrobial response factors, investigate FREP expression and diversity, and aid annotation of *P. acuta* FREPs.

2. Materials and methods

2.1. Organisms and experimental treatments

Physella acuta (isolate A) snails, originally collected in 2010 from Stubblefield Lake in northern New Mexico, are maintained at the University of New Mexico (UNM, Nolan et al., 2014). Adult snails (4–6 mm shell length) were used in this study.

Gram-negative *Escherichia coli* (ATCC 14948) and Gram-positive *Micrococcus luteus* (ATCC 9341) bacteria were obtained from the microbiology teaching facility at UNM, cultured overnight in LB, and diluted (OD600 = 1). One individual snail was injected with 50 μL of a 1:1 bacterial suspension (*E. coli* and *M. luteus*) into the headfoot using a G27 hypodermic needle. The bacterially-injected snail (B) was placed into a well of a 24-well plate with artificial spring water (ASW) for 12h (Adema et al., 1999; Hanelt et al., 2008). An individual sham-exposed (S) snail, which did not receive an injection, was also placed into a well of a 24-well plate with ASW for 12 h.

The trematode parasite *E. paraensei* is maintained at UNM, utilizing M Line *B. glabrata* snails and hamsters as intermediate and definitive hosts, respectively (Loker et al., 1992).

Individual *P. acuta* snails were placed in 12-well plates with ASW and exposed to ten *E. paraensei* miracidia overnight. The following day, wells were checked for the absence of free-swimming parasites as a measure of infection. Exposed snails were placed in aquaria and fed red-leaf lettuce *ad libitum*. Infections were confirmed at 2, 4, and 8 days post exposure (DPE) via observation through the shell of parasite sporocysts in the heart of *P. acuta* using a stereomicroscope.

2.2. Nucleic acid extraction and purification

RNA was extracted from whole-body tissues of *P. acuta* snails using Trizol (Invitrogen, Carlsbad, CA) following the manufacturer's protocol and treated with DNA-freeTM (DNA Removal Kit, Ambion, Inc., Austin, TX) to remove residual genomic DNA. RNA quality was checked using a bioanalyzer (Agilent 2100, nano kit). For 454 pyrosequencing, mRNA was isolated from the B and S snails (Ambion® MicroPoly(A) Purist Kit, Life Technologies).

DNA was extracted from two *P. acuta* snails following the CTAB method of Winnepenninckx et al. (1993). For removal of RNA, the samples were treated with RNAse A for 1 h at 37 °C (ThermoFischer, Waltham, MA).

2.3. 454 pyrosequencing

Messenger RNA was subjected to 454 pyrosequencing on the GS FLX + sequencing platform (Research and Testing Laboratory, Lubbock, TX). Reads were submitted to the Short Read Archive (SRA) of GenBank (accessions SRX1041852, SRX1041811).

2.4. Illumina genomic DNA sequencing

DNA samples from two *P. acuta* snails were used to produce Nextera libraries (Illumina), which were then sequenced (NextSeq 500) at the CETI core facility of UNM. Genomic reads (150 bp, paired-end) from both snails were submitted to the SRA database of GenBank

(accessions SRX3347695, SRX3347220). Genome coverage was estimated (15X) relative to *Physa rubra* (1.17 Gbp; Animal Genome Size Database; Gregory, 2017) using the total number of nucleotides recorded from the *P. acuta* genome.

2.5. Assembly of *P. acuta* reference transcriptome

Prior to transcriptome assembly, reads were assembled with publicly available ITS sequences to reconstruct a *P. acuta* rDNA cassette (accessions KF316327, KF316329, KF316328, KF316326). Reads were screened to remove sequences derived from rDNA cassettes, mtDNA (Nolan et al., 2014; accession NC_023253), and *E. coli* and *M. luteus* genomes (accessions NC_000913, NC_012803) using SeqClean software (http://sourceforge.net/projects/seqclean/). Considering that the Newbler software, provided by 454 LifeSciences, uses a lower stringency for merging reads (Mundry et al., 2012), thereby increasing risk of assembling chimeric contigs, reads from the B and S datasets were assembled individually using MIRA 4 (default parameters, v4.0.2, Chevreux et al., 1999). The assemblies were then merged (99% sequence identity, 50 nt minimum overlap) in Sequencher (v4.10.1, Gene Codes Corporation) to generate a combined reference transcriptome that captured baseline and evoked immune responses. The reference transcriptome was uploaded to a local blastable database (ncbi-blast- 2.6.0+) for downstream analyses.

2.6. General analysis and annotation of *P. acuta* reference transcriptome

Completeness of the 454 reference transcriptome was assessed relative to the transcriptome of *B. glabrata*, a closely related snail species for which considerable sequence data is available (VectorBase, BglaB1.6), using gVolante (Nishimura et al., 2017) to evaluate the re-presentation of 303 conserved eukaryotic core genes. All predicted transcripts of *P. acuta* were

subjected to InterProScan (Finn et al., 2017) and DIAMOND automated blastx analysis (Buchfink et al., 2015) for annotation of protein-level sequences.

2.7. Identification of antimicrobial peptides (AMPs)

To search for antimicrobial peptides expressed by *P. acuta*, the reference transcriptome was queried against the Antimicrobial Peptide Database (Wang et al., 2009). Sequence similarity matches with an evalue ≤ 1E-6 were considered significant. Recovered AMP-like sequences were then used as protein queries (tblastn) against the reference transcriptome of *P. acuta* to search for additional related sequences. Phylogenetic reconstruction of protostome macin-type antimicrobial peptides was performed using MEGA 6 (Maximum-like- lihood, Tamura et al., 2013) and the Whelan and Goldman (WAG) protein substitution model, employing only alignable regions shared by all predicted transcripts (amino acid). Sequences were retrieved from the AMP database (*Achatina fulica* AP02160, *Hirudo medicinalis* AP02025, *Hydra magnipapillata* APO1552, *Theromyzon tessulatum* APO15565) and GenBank (*Ancylostoma ceylancium* EYB89210, *Ancylostoma duodenale* KIH43895, *A. californica* NP_001191629, *Diploscapter pachys* PAV63472, *Eisenia fetida* AKO70613, *Folsomia candida* XP_ 021953172, *Haemonchus contortus* CDJ83884, *Helobdella robusta* XP_ 009018889, *Hydra vulgaris* AFQ20833, *Hyriopsis cumingii* AEC50045, *M. galloprovincialis* CCC15015-CCC15019, *Necator americanus* XP_ 013305848, *Oesophagostomum dentatum* KHJ90660, *Orbicella faveolata* XP_020621503, *Orchesella cincta* ODM88634, *Pristionchus pacificus* PDM62724 *Ruditapes philippinarum* ALJ03313, *Teladorsagia circumcincta* PIO76669). *Biomphalaria glabrata* macins (biomphamacins) were retrieved from the *B. glabrata* genome project (Adema et al., 2017).

2.8. Identification of antimicrobial proteins

The reference transcriptome was queried against the non-redundant protein (nr) database of NCBI via blastx. Transcripts of antimicrobial proteins were recovered using keyword searches of blast results for aplysianin/achacin, lipopolysaccharide-binding protein/bactericidal permeability-increasing protein (LBP/BPI), and peptidoglycan-recognition protein (PGRP). Additionally, a sequence similarity search (tblastn) of the *P. acuta* reference transcriptome was performed using *B. glabrata* biomphalysin (accession KC012466.1). A representative sequence from each antimicrobial protein category was used to query (tblastn) the reference transcriptome to search for additional related sequences. Transcripts included in the downstream analyses were full- length (i.e. containing start and stop codons) continuous open-reading frames (ORFs). The bioinformatics tools used to analyze and annotate predicted transcripts of *P. acuta* at the amino acid level included: ExPASy translate (Gasteiger et al., 2003), SignalP 4.1 (Petersen et al., 2011), SMART (Letunic and Bork, 2017), and blast (Altschul et al., 1990).

2.9. RT-PCR validation of predicted antimicrobial proteins

RNA from individual *P. acuta* snails was reverse transcribed (Omniscript RT kit, Qiagen) with polyT primers followed by PCR (Platinum Taq High Fidelity, ThermoFischer) using gene specific primers (Table S1). Appropriately sized amplicons were sequenced directly on both strands using the QIAquick PCR Purification Kit (Qiagen), BigDye v3.1 (Invitrogen, Carlsbad, CA) and PCR primers. Extension products were recorded (ABI 3130x) at the Molecular Biology Facility of UNM. Chromatograms were edited by eye in Sequencher (v4.10.1) and compared to *in silico* predicted transcripts of the 454 reference transcriptome. RT-PCR validated sequences were submitted to GenBank under the accessions MF787214, MF978286, MF978287, MF787215, MF787217, MF787218, and MF978288.

2.10. FREP analysis

Blast-generated annotations of all sequences in the reference transcriptome were queried for FReD-like sequences using fibrinogen, FBG, and FREP as keywords. The resulting *in silico* predicted transcripts were analyzed for complete ORFs and canonical features of bonafide FREPs, including IgSF and FBG domains (Hanington and Zhang, 2011). Selected transcripts were then translated and used to query (tblastn) the reference transcriptome to broaden the search of *P. acuta* FREPs. Secondary protein structures of IgSF domains of *P. acuta* FREPs were predicted using JPred 4 (Drozdetskiy et al., 2015). The genomic reads recorded from *P. acuta* were uploaded to a local blast database and used to map intron/exon splice sites on the FREP transcripts and determine the intron/exon splice sequences. Size and full genomic sequences of introns were not characterized.

To gauge diversification of *P. acuta* FREPs, RT-PCR amplification of an IgSF and FBG-containing region of Pa_FREP1 and Pa_FREP2 (accessions MF787215, MF787216) was performed using cDNA from an individual *P. acuta* snail using primers designed from the predicted transcripts (Table S1) and high fidelity Platinum SuperFi DNA polymerase (Invitrogen). Amplicons were cloned (Zero-blunt TOPO, Invitrogen) into pCR 2.1 (Invitrogen, Carlsbad, CA) and 24 inserts were sequenced (BigDye, ABI). As a control, 24 inserts of the housekeeping genes EIF2a and COX1 were cloned and sequenced (accessions MF978287, MF978286). Partial sequence of an allele of Pa_FREP2 was also recovered (accession MH358391). Sequences were aligned for comparison (Sequencher, v4.10.1) and inspected for nucleotide polymorphisms.

To check protein level FREP response to an infection, plasma (cell- free hemolymph) was collected from control (unexposed) and *E. paraensei*-infected snails (16 total, 4–6 mm in length). Live snails unexposed and 2, 4, and 8 DPE were placed within a 600 µL microcentrifuge tube perforated with a small hole at the bottom and were lightly crushed with a dissecting probe. Each tube was placed within a 1.7 mL collection microcentrifuge tube and spun at 500 x g for ten minutes at 4 °C. With this approach, body tissues remained in the upper microcentrifuge tube and hemolymph was collected in the collection tube, separated into a pellet of hemocytes and 2–5 µL of cell-free plasma. Two µL of plasma was placed in 18 µL of 2X Laemmli Buffer (Bio-Rad, Hercules, CA) and 5% 2-Mercaptoethanol, boiled (ten minutes) to denature proteins and then spun (16,000 x g) for 5min to remove any debris. Samples were separated on 5–20% gradient SDS-PAGE gels (Couch et al., 1990; Loker and Hertel, 1987). Protein gels were stained with Coomassie Blue (0.25% w/v R-250, 10% v/v glacial acetic acid, 45% v/ v methanol, 45% v/v H2O) by heating the gel for ten seconds in a microwave (adapted from ThermoFisher Coomassie R-250 staining protocol), followed by an incubation for ten minutes at room temperature, a rinse with H2O, and destained (10% glacial acetic acid, 45% methanol, 45% H2O) for several hrs. The gels were then photodocumented (BioRad Gel Doc XR+).

To determine if FREPs were upregulated at RNA level after exposure to *E. paraensei*, differential expression analysis of *P. acuta* transcripts was performed. A total of 16 *P. acuta* snails were used: seven control snails unexposed to *E. paraensei*, five snails at 2 DPE, and four snails at 8 DPE. Individual *P. acuta* snails were placed in 12-well plates with ASW and exposed to ten *E. paraensei* miracidia overnight. Infections and RNA extractions were performed as described previously. RNA samples were used to generate TruSeq (v2) libraries (Illumina) which were then sequenced (NextSeq 500, 150bp, paired-end). Raw reads were processed using

Trimmomatic (Bolger et al., 2014) and aligned (Bowtie2) to the *Echinostoma caproni* genome (GCA_000950555.1) to filter out reads contributed by the parasite. Filtered reads (paired-end) were then assembled using Trinity *de novo* assembler version 2.4.0 (Grabherr et al., 2011) and contig-redundancy was minimized using CD-HIT at a 95% clustering threshold (Li and Godzik, 2006). Finally, read abundance was quantified with RSEM and comparisons of control and experimental transcripts were performed using DESeq2 (Li and Dewey, 2011; Love et al., 2014). The differentially expressed sequences and associated annotations are provided (Table S2; Data S1).

2.11. FReD identification and analysis

Predicted transcripts containing fibrinogen-related domains, that were not FREPs (i.e. lacking IgSF domains), were recovered using InterProScan to identify fibrinogen (FBG) domains and by using the amino acid sequence of the FBG domain of Pa_FREP1(accession MF787215) as a query (tblastn) against the reference transcriptome. All recovered sequences were inspected for ORFs (containing start and stop codons) and subjected to protein domain analysis using SMART (Letunic and Bork, 2017) and SignalP 4.1 (Petersen et al., 2011) to identify transmembrane domains, signal peptides, and other putative domains.

The FReD sequences (amino acid levels) were aligned in Clustal X (v2.1, Larken et al., 2007; Data S2). Phylogenetic analyses were performed using MEGA 6 (Maximum-likelihood, Tamura et al., 2013) and the Whelan and Goldman (WAG) protein substitution model, employing only alignable regions shared by all predicted transcripts. FReDs of *B. glabrata* (accessions: XP_013069966, XP_013071685) and *M. galloprovincialis* (accessions: ADQ55806, ADQ55821, ADQ55822) were recovered from GenBank. A percent similarity and identity

matrix of *P. acuta* FReDs was generated via pairwise alignments between FReD amino acid sequences (blastp).

3. Results

3.1. Experimental treatments

The bacterial exposure, implemented by injection of 50μL of a bacterial suspension, and sham exposure, did not negatively impact snail viability; both B and S snails were actively moving 12 h post-exposure.

Preliminary experiments demonstrated that *E. paraensei* exposure yielded patent infections, culminating in the release of viable cercariae from *P. acuta*. For this study, two replicate exposures (2 × 12 P. acuta snails) to ten *E. paraensei* miracida were performed. All wells lacked observable miracidia upon inspection after 24 h of exposure, suggesting successful parasite penetration of the snail body. Twenty-three of the 24 *P. acuta* snails exposed to the parasite survived. Viable infections were confirmed by the presence of parasite sporocysts in the snail heart at 2, 4, and 8 DPE or by snails shedding cercariae (patent infection). Infection prevalence of 52.6% was achieved (12 of 23 snails).

3.2. *P. acuta* reference transcriptome

The 454 reads generated from snail S (548,871 reads) and snail B (635,344 reads) were assembled individually using MIRA and merged in Sequencher yielding a total of 24, 288 predicted transcripts (reference transcriptome, Data S3). gVolante analysis showed that general annotation features were similar between *B. glabrata* and *P. acuta* including N50 length, AT-richness, and mean sequence length (Fig. S1A). The *P. acuta* reference transcriptome included 63% (partial and complete sequences) of the 303 eukaryotic core genes. By comparison, the *B. glabrata* transcriptome recorded 71% (partial and complete sequences; Fig. S1B).

InterProScan and automated blastx searches provided general annotation of the *P. acuta* transcriptome (Table S3).

3.3. Antimicrobial peptides (AMPs)

Sequence similarity searches against the nr database of NCBI and the AMP database (Wang et al., 2009) identified two *P. acuta* transcripts as potential antimicrobial peptides of the macin family of AMPs, now referred to as physamacins (Pa_macin_1, Pa_macin_2; Data S4). Both *P. acuta* macins most closely resemble Mytimacin-6 of the bivalve *M. galloprovincialis* (blastp, accession AHG59339). Phylogenetic analysis showed *P. acuta* macins cluster with other lophotrochozoan (annelids and molluscs) macins (Fig. 1), separate from ecdysozoan (arthropods and nematodes) macins. Additionally, pre-bilaterian (cnidarian) macins more closely group with lophotrochozoan macin sequences.

3.4. Antimicrobial proteins

The reference transcriptome of *P. acuta* revealed a myriad of known gastropod antimicrobial protein sequences including 12 aplysianin/achacin, four LBP/BPI, seven PGRP, and three physalysins (Table 1). Based on distinct coding sequences and untranslated regions (UTRs), these transcripts were considered to originate from unique genes (Sequencher 99% identity, 50 nt minimum overlap). The full-length coding sequences (ORFs) of all these antimicrobial proteins are provided (Data S4) and a randomly selected representative partial sequence from each immune-gene category was validated by RT-PCR using gene-specific primers (Table S1). The predicted transcript of Pa_aplysianin1, for which partial sequence was validated, most resembles an L-amino acid oxidase called ink toxin 1 of *Aplysia punctata* (blastx, accession AAR14185) and *Aplysianin-A* of *Aplysia kurodai* (blastp, accession Q17043) with 38% identity for both analyses. LBP/BPI1 from *P. acuta* was significantly similar to LBP/BPI3.2

of *B. glabrata* with 38% identity (blastp, accession AKM45823). PGRP1 from the *P. acuta* reference transcriptome has 100% identity to an independently reported PGRP of *P. acuta* (accession AEH26026). In addition, all PGRP sequences recorded from the reference transcriptome do not harbor transmembrane domains and therefore code for short-type regulatory PGRPs. Lastly, blastx analysis showed that physalysin1 resembles most an uncharacterized sequence from *A. californica* (accession XP_005093610) and both sequences resemble an aerolysin-like factor similar to *B. glabrata* biomphalysin (accession AGG38744).

3.5. FREPs

Two full-length FREP transcripts were recorded from the reference transcriptome (Data S5). Analysis of the genomic Illumina reads from the *P. acuta* genome (15-fold coverage) did not indicate the presence of additional FREP gene sequences. Experimental validation by RT-PCR and sequencing confirmed the computationally predicted sequences of Pa_FREP1 and Pa_FREP2. Pa_FREP1 has 51% sequence identity (blastp) and 65% sequence similarity to fibrinogen-related protein A precursor (accession AQX34557) of *B. glabrata*. Pa_FREP2 has 44% sequence identity and 61% similarity to fibrinogen-related protein J2 (accession AQX3454). Canonical FREP features were evident from *P. acuta* FREP1 and FREP2 (Fig. 2A). Additional to a signal peptide (1-21aa FREP1, 1- 23aa FREP2), these include an N-terminal immunoglobulin superfamily (IgSF) domain (26-145aa FREP1, 32-141aa FREP2) identified by conserved cysteine residues and predicted secondary protein structures (β sheets; Léonard et al., 2001) and a C-terminal fibrinogen (FBG) domain (203-414aa FREP1, 297- 498aa FREP2). The predicted molecular weights are 43.3 kDa and 53.3 kDa for FREP1 and FREP2, respectively. The amino acid alignment indicates considerable sequence differences between the two genes,

including sequence length (421 aa FREP1, 499aa FREP2), mainly accounted for in the region between the IgSF and FBG domains.

Using genomic Illumina sequence data, intron/exon splice sites for both *P. acuta* FREPs were found to be Y/GT and HAG/B, similar to Léonard et al., 2001 (Fig. 2B). The FREP1 gene spans three exons and two introns whereas the FREP2 gene comprises six exons and five introns (Fig. 2B). These intron/exon structures differ from those of FREP genes reported from *B. glabrata* and *A. californica* (Adema et al., 2017; Gorbushin et al., 2010).

The subcloning and sequencing of RT-PCR amplicons (24 each) of a 649 nt region of *P. acuta* FREP1 and a 735 nt region of FREP2, which span the IgSF and FBG domains, did not reveal sequences modified via single nucleotide polymorphisms (Table 2). However, one variant sequence of FREP2 was recovered from two independent inserts and is designated as an allele of FREP2 (accession MH358391). Sequence analysis of cloned fragments (24 each) of two control genes, EIF2a and COX1 (Data S5), revealed one variant sequence among the COX1 sequences, likely achieved via polymerase error, and the EIF2a sequence was not polymorphic (Table 2).

Coomassie-stained SDS-PAGE gels revealed that plasma samples of *P. acuta* snails harboring viable infections (Fig. 3A) at 2, 4, and 8 DPE to *E. paraensei* miracidia were not discernibly different from plasma samples of control *P. acuta* snails unexposed to the parasite (Fig. 3B). In particular, the predicted size range (43.3–53.3 kDa) for *P. acuta* FREPs did not reveal obvious changes in protein composition. These results contrast a repeated observation of increased expression of FREPs in *B. glabrata* after exposure to *E. paraensei*, evident as prominent and broad protein bands in plasma samples from *B. glabrata* at 3 DPE (Adema et al., 1999). Also, differential expression analysis of Illumina transcriptome data did not reveal any upregulated FREP sequences of *P. acuta* snails 2 and 8 DPE when compared to unexposed

control *P. acuta* snails (Table S2; Data S1), further indicating the lack of a FREP response in *P. acuta* snails after exposure to the parasite *E. paraensei*.

3.6. FReDs

Twenty-four full-length single-FBG FReD sequences were recovered from the reference transcriptome of *P. acuta*. These sequences contain predicted C-terminal FBG domains with short upstream sequences encoding signal peptide domains, transmembrane domains, or neither of these domains. Phylogenetic analysis of the FBG domains of these FReDs yielded five clusters of related sequences (Fig. 4A). Two clusters (1 and 5) contain FReD sequences with predicted signal peptides, indicative of secreted proteins. Cluster 3 includes FReD sequences with predicted transmembrane domains. The remaining two clusters (2 and 4) contain five sequences without signal peptides or transmembrane domains, identifying putative cytosolic factors. As shown (Fig. 4B), extensive sequence diversity exists among these structurally-related sequences. For sequences with high percent identity (red boxes above 100% diagonal), additional full-length nucleotide sequence pairwise alignments were performed. With the exception of the two contigs Pa_FReD23 and Pa_FReD24, no two sequences had fewer than 13 variant nucleotides when compared, suggesting that these FReD sequences likely originate from unique genes. Blast searches indicate that *B. glabrata* also expressed a diversity of similar FReD proteins. Phylogenetic analysis of predicted amino acid sequences of gastropod and bivalve FReDs shows that a group of secreted FReD sequences of *P. acuta* is most closely related to three immune-relevant FReDs of *M. galloprovincialis* (Fig. 4C). *Biomphalaria glabrata* FReDs, also proposed to be immune-relevant (Wu et al., 2017), cluster with a different subset of secreted *P. acuta* FReDs (Fig. 4C).

4. Discussion

The current view of gastropod immunity is primarily focused through a lens provided by extensive studies of *B. glabrata*. It is pertinent to investigate whether this single species of lung-breathing freshwater snails (Gastropoda, Hygrophila, Planorbidae) provides a representative characterization of immunity for the diverse Gastropoda, largest class within the phylum Mollusca. Next-generation sequencing (NGS) has removed previous limitations for comparative immunology by providing easy access to organismal genomes and facilitating broad exploration of immune capabilities among phylogenetically distinct animals, including molluscs (Dheilly et al., 2014; Litman and Cooper, 2007; Schultz and Adema, 2017). Knowledge derived from model organisms like *B. glabrata* can now be considered to inform the evolution of immune function within larger taxonomic groups. Previously available NGS datasets from a somewhat random selection of snail species may be mined for comparative gastropod immunity, however, strategic choices of gastropod species that represent distinct evolutionary branches of phylogeny will develop answers more directly. For example, *B. glabrata* employs somatic mutation of immune genes encoding fibrinogen-related proteins (FREPs) to achieve individually unique immune identities (Dheilly et al., 2015; Galinier et al., 2017; Hanington et al., 2010; Moné et al., 2010; Zhang et al., 2004). Analyses, mostly of NGS data, involving basal gastropods (*Lo. gigantea*, *Li. littorea*), the phylogenetically distant *A. californica* (euophistobranch) and *H. trivolvis*, representing a sister genus to *Biomphalaria* (Adema, 2015; Adema et al., 1999; Gordy et al., 2015; Gorbushin and Borisova, 2015; Gorbushin et al., 2010) do not provide the phylogenetic resolution to clarify origin and distribution of somatically diversified FREPs as gastropod immune factors. This study presents *P. acuta* as a strategic choice because it represents the Physidae, sister family of Planorbidae (that includes *B. glabrata*) within the hygrophila, for

family-level comparative analyses of gastropod immune function, including FREP biology relative to the *Biomphalaria* model.

gVolante analysis indicates that the reference transcriptome of *P. acuta* relative to the transcriptome of *B. glabrata* (Fig. S1A) provides a comprehensive record of transcripts. Although not all of the 303 eukaryotic core genes are recorded from either snail species (Fig. S1B), it is possible that some of the core genes are defined too widely to relevantly capture the biology of molluscs. In addition, protein coding sequences of one particular mollusc species (*Biomphalaria pfeifferi*) may not necessarily be recorded from other mollusc transcriptomes and genomes (Buddenborg et al., 2017). This indicates the potential for incomplete sampling by RNAseq or it may reflect differences due to phylogenetic distance of gene expression, or even presence (absence), of particular genes. The use of whole-body tissue prevents bias for genes uniquely expressed in particular organs. Both a sham-exposed and a bacterially-injected snail were sampled to capture both constitutive and induced defense factors. Exposure to bacteria and bacterial pathogen-associated molecular patterns (PAMPs) can increase abundance of immune transcripts such that they can be more readily detected (Adema et al., 2010; Hanelt et al., 2008; Sullivan and Belloir, 2014). The genomic reads captured for *P. acuta* (15-fold coverage) provided a valuable resource for mapping transcripts to the genome to confirm gene predictions and infer gene number and annotate intron/exon structure of select genes.

Compared to other metazoans, the AMP arsenal of *B. glabrata* is restricted to a gene family of macins (Adema et al., 2017), with one macin known to be upregulated in *B. glabrata* after exposure to *S. mansoni* miracidia (Ittiprasert et al., 2010). This particular macin sequence was most similar to theromacin originally described from the leech *Theromyzon tessulatum* (Tasiemski et al., 2004). Similarly, just two macin-like AMP sequences were recorded for

P. acuta. Macins have been recorded from other molluscs, but in bivalves multiple families of AMP sequences are present (Gerdol et al., 2012). A broader view of the phylogenetic distribution of modest representation of AMP genes will require analysis of other gastropod species outside the Hygrophila. Gastropods likely depend on other aspects of immunity, such as cellular immune function (Coustau et al., 2015; Humphries and Yoshino, 2003) or other, larger humoral factors, like antimicrobial proteins (see below) to deal with infection.

The *B. glabrata* genome and other investigations of *B. glabrata* immunity revealed multiple antimicrobial protein genes (Adema et al., 2017; Baron et al., 2016). Likewise, the *P. acuta* reference transcriptome yielded 12 aplysianin, four LBP/BPI, three physalysin, and seven PGRP sequences. Aplysianin, first described from *A. kurodai*, is an L-amino acid oxidase (LAAO) that produces hydrogen peroxide, an antibacterial reactive oxygen species (Kamiya et al., 1986; Kasai et al., 2015). Furthermore, *B. glabrata* snails resistant to the trematode parasite *E. caproni* expressed LAAOs after exposure to miracidia (Bouchut et al., 2007). A proteomics analysis of the egg mass fluid (EMF) of *B. glabrata* identified several LAAOs, suggesting parental investment in immunoprotection of offspring (Hathaway et al., 2010). These observations suggest an immune role for the 12 LAAO transcripts in *P. acuta*. Four LBP/BPI sequences were recorded from the *P. acuta* reference transcriptome, similar to the *B. glabrata* genome that includes five genes encoding LBP/BPI. These are known defense factors of *B. glabrata* and some are parentally-invested in EMF to aid in defense against pathogenic water molds known as oomycetes (Baron et al., 2013; Hathaway et al., 2010). In addition, *B. glabrata* expresses unique members of the LBP/BPI family to tailor a defense response to different pathogens (Baron et al., 2016). Three physalysin transcripts were detected in the reference transcriptome of *P. acuta*. Pore-forming toxins are ubiquitous across all kingdoms of life and are

thought to have arisen in metazoans via horizontal gene transfer from bacteria (Galinier et al., 2013; Podobnik and Anderluh, 2017). Twenty-one biomphalysins (similar pore-forming toxins) were catalogued from the *B. glabrata* genome (Adema et al., 2017). Biomphalysins responded to diverse types of pathogens including, fungi, Gram-negative and Gram-positive bacteria, and trematode parasites (Tetreau et al., 2017). In the latter case, biomphalysins bind directly to the surface of sporocysts and exert cytotoxic activity (Galinier et al., 2013). The apparent modest diversity of physalysins, relative to the number of biomphalysins of *B. glabrata*, may represent another difference in immunobiology of these two species. Seven PGRP transcripts were found within the reference transcriptome of *P. acuta*, all of which are short-form variants that contain predicted signal peptide sequences. PGRPs belong to two classes of sequences: short-form and long-form. Short-form PGRPs are secreted proteins that bind peptidoglycan directly, ultimately resulting in activation of the Toll and prophenoloxidase pathways in *Drosophila* melanogaster (Park et al., 2007). Long-form PGRPs usually contain transmembrane domains, are membrane-bound, and regulate the IMD pathway in *Drosophila* (Steiner, 2004; Zaidman-Rémy et al., 2006). PGRPs have been characterized from other gastropods, including *Haliotis discus discus* (Premachandra et al., 2014) and *B. glabrata*. The genome of the latter contains eight PGRP genes (Adema et al., 2017), including a long form PGRP gene that yields splice variants (Zhang et al., 2007). Analysis of the *P. acuta* transcriptome and genome did not reveal any long-form PGRPs. Potentially *P. acuta* and *B. glabrata* differ in their utilization of PGRPs; *P. acuta* may only rely on secreted, short-form PGRP types. Furthermore, the absence of transmembrane domains in *B. glabrata* long-form PGRPs (accessions EF079962, EF452348, EF452347, EF079963, EF452349) suggests that gastropod PGRPs (short and long-forms) are not membrane-bound. Usage of non cell-bound PGRPs proposes that this aspect of gastropod immunity more

closely reflects vertebrate (mammalian) immunity (Lu et al., 2006), rather than ecdysozoan immunobiology (Royet et al., 2011).

Two full-length transcripts encoding canonical FREPs were identified within the reference transcriptome of *P. acuta*. Sequence analysis of the genomic Illumina reads identified matching genomic sequences for the two transcripts and did not indicate additional FREP genes in *P. acuta*. This is in stark contrast relative to *B. glabrata*, which has a FREP gene family of more than 20 genes (Adema et al., 2017; Dheilly et al., 2015). The modest FREP gene complement of *P. acuta* is more similar to that of other gastropods like *A. californica*, where two FREP genes have been identified, (Gorbushin et al., 2010) and *L. littorea*, where one FREP transcript was expressed (Gorbushin and Borisova, 2015). *Physella acuta* only has FREPs with a single IgSF domain upstream of the FBG domain whereas other gastropods have FREPs containing two IgSF domains (*A. californica* and *L. littorea*) or two types of FREPs with either one or two IgSF domains (*B. glabrata*). Additionally, the unique intron/exon structures of *P. acuta* FREPs, especially the presence of introns in the FBG domain (Pa_FREP2; Fig. 2B) suggests a unique evolutionary history that differs from other gastropod FREPs (Adema et al., 2017; Gorbushin et al., 2010). Using the same high fidelity PCR and sequencing methods that demonstrated FREP sequence diversification in *B. glabrata* (Dheilly et al., 2015; Zhang et al., 2004), no FREP sequence diversification was recorded from *P. acuta* FREPs relative to controls (Table 2). Combined, these observations from *P. acuta* present dramatic departures from the biology of *B. glabrata* FREPs, perhaps indicating that somatic diversification and gene-family expansion of FREPs are specific features of planorbid immunity. Investigations of additional hygrophilid gastropods (lymnaeids and acroloxids) will help develop this hypothesis. FREPs were initially discovered due to increased abundance in the plasma of *B. glabrata* after parasite

infection, specifically from *E. paraensei*-infected *B. glabrata* snails and similar responses were recorded from *H. trivolvis* (Adema et al., 1997, 1999; Hanington et al., 2010). Serendipitously, we confirmed that *E. paraensei* is compatible with *P. acuta* snails. *Echinostoma paraensei* infection did not evoke detectable expression of *P. acuta* FREPs at protein levels, signifying a striking difference from *B. glabrata* immunobiology. The analysis of protein-level response showing lack of FREP involvement was corroborated independently by differential expression analysis of transcripts of *E. paraensei* exposed *P. acuta* snails at 2 and 8 DPE. However, not all *P. acuta* snails get infected after experimental exposure to the parasite. This implies that some *P. acuta* snails can mount effective defense responses to *E. paraensei*. Current efforts, consisting of comprehensive analyses of NGS data are underway to characterize the immune response of *P. acuta* specifically to determine, if not FREPs, what other aspects of *P. acuta* immunity are at play in this novel host/parasite interaction.

Previously Gorbushin et al. (2010) have proposed that FREP diversity and diversification are features that evolved with panpulmonate gastropods. Our analysis of *P. acuta*, a hygrophilid snail of a sister-family to *B. glabrata*, leads us to further propose that FREP sequence diversification, gene expansion (diversity), and FREP protein expression are immune features that may be restricted to planorbid snails. Indeed, we should expect phylogenetic distances and distinct evolutionary histories to contribute to differences in immunobiology among species, including the use of FREPs in anti-trematode responses (Suwannatrai et al., 2016). Of course, FREPs are immune determinants for planorbid snails that transmit schistosome parasites that cause human disease (Hanington et al., 2012, 2010), and FREPs remain the only invertebrate immune factors known to be somatically diversified (Ghosh et al., 2011). Given the diversity of sequences encoding recognition proteins across phylogeny, and the need to immunologically

contend with rapidly evolving pathogens, novel immune factors are likely to exist within *P. acuta* (Buckley and Rast, 2015; Jack, 2015).

The search for FREPs identified another category of 24 FBG containing sequences from the reference transcriptome of *P. acuta*, known as FReDs. These FReD-type sequences are not well known as immune factors of gastropods, but blast searches indicate a number of FReDs expressed by *B. glabrata* with two sequences found to bind parasites (Wu et al., 2017). FReDs are also present in bivalves, where a range of two to seven genes have been recorded from *Mytilus edulis*, *Mytilus californianus*, and *M. galloprovincialis* (Gorbushin and Iakovleva, 2011). The FReDs of *M. galloprovincialis* are immune-relevant and diversified at the mRNA level (Romero et al., 2011). An immune role for FReDs is also indicated by the large family of 58 FReD sequences in the ecdysozoan *Anopheles gambiae*, also considering that these sequences are upregulated in response to malaria infection (Christophides et al., 2002; Dimopoulos et al., 2002; Zdobnov et al., 2002). The *P. acuta* FReDs vary considerably at amino acid level, yet the FBG domains cluster by groups of FReD sequences that share similar upstream protein domains. Two groups of FReDs contain predicted signal peptide sequences, which identifies secreted proteins (Fig. 4A). One group of FReDs have predicted transmembrane domains, a novel type of molluscan FReD (Fig. 4A). The previously mentioned immune-relevant bivalve FReD sequences cluster in a gene tree with a group of *P. acuta* secreted FReD types (Fig. 4C). Indeed, both groups of FReDs from *P. acuta* and *M. galloprovincialis* contain predicted upstream signal peptides. As such, this group of *P. acuta* FReDs are promising candidates as immune-relevant sequences. Interestingly, the parasite-reactive *B. glabrata* FReDs (Wu et al., 2017) share sequence similarities to the FBG domain of a subset of secreted *P. acuta* FReDs (Fig. 4C). Potentially these secreted *P. acuta* FReDs are also parasite antigen interactive molecules. The

remaining groups of *P. acuta* FReDs are putative cytosolic proteins because they lack signal peptides and transmembrane domains. The extensive variability in sequence length and high levels of nucleotide polymorphisms between *P. acuta* FReDs suggest that these diverse transcripts are likely derived from unique genes, rather than through somatic diversification mechanisms identified in *B. glabrata* or the uncharacterized mechanisms proposed to drive FReD diversity in bivalves (Romero et al., 2011; Zhang et al., 2004). Clearly, FReDs deserve more research regarding their potential as immune factors in *P. acuta* and other gastropods.

We characterized aspects of the understudied immunobiology of *P. acuta* and performed a family-level comparison of gastropod immunomes to determine the extent to which the well-studied model *B. glabrata* represents gastropod immunity. Considering snails of the families Physidae and Planorbidae, it seems that immunity is organized similarly with modest arsenals of antimicrobial peptides and several multigenic families of antimicrobial proteins. There are, however, dramatic differences in FREP biology between these gastropod species, leading us to speculate that this is a unique feature of planorbid immunobiology, rather than a general feature of gastropod immunity. The considerable diversity of *P. acuta* FReDs and presence of FReDs in *B. glabrata* illuminates a potential novel feature of gastropod immunobiology. Ongoing investigations of *P. acuta* immune responses to parasite exposure will contribute a more functional understanding of *P. acuta* defenses. Additional investigations of hygrophilid immunity are needed to expand our phylogenetic resolution of gastropod immune capabilities. Lastly, strategic choices of organisms used for comparative immunology will help identify taxon-specific and broad aspects of immune function within the diverse phylum Mollusca and help bridge gaps in our knowledge on the evolution of immunity among metazoans.

Acknowledgments

This study was supported by NIH grant P20GM103452 from the National Institute of General Medical Sciences (NIGMS). The authors wish to acknowledge the molecular biology facility (MBF) at the University of New Mexico for sequencing and technical support and the Center for Evolutionary and Theoretical Immunology (CETI) for NGS data. JHS recognizes the Alvin R. and Caroline G. Grove scholarship. JHS also acknowledges Benjamin Wheaton for valuable manuscript edits.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.molimm.2018.05.029.

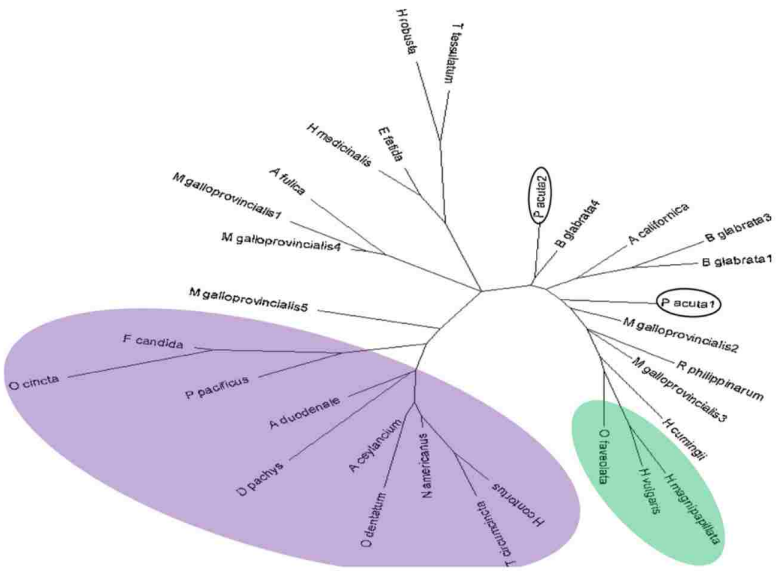


Fig. 1. *P. acuta* antimicrobial peptides: physamacins.
Unrooted gene tree displaying some of the diversity of protostome metazoan macin-like sequences. Sequences (amino acid) were retrieved from the Antimicrobial Peptide Database (Wang et al., 2016) and GenBank. Accession numbers and complete species names are provided in materials and methods. A total of 31 sequences were aligned and phylogenetic reconstruction yielded broad evolutionary groups of ecdysozoan macins (purple), lophotrochozoan macins (unshaded), and pre-bilaterian macins (cnidarians, green). Pre-bilaterian macins are more similar to lophotrochozoan macin sequences. Note that *P. acuta* macin sequences (circled) group together with other mollusc macins. Sequence analysis suggests that physamacin1 is more similar to gsastropod macins, whereas physamacin2 is more similar to bivalve macins (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

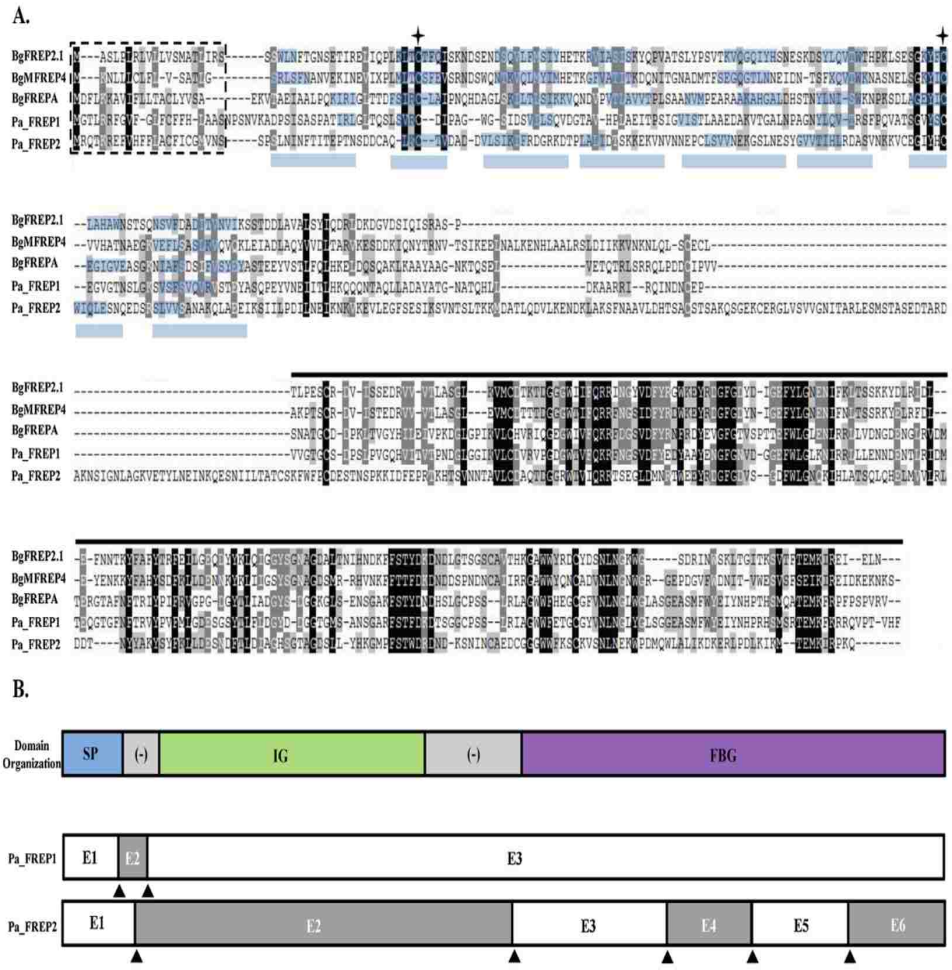


Fig. 2. *P. acuta* fibrinogen-related proteins (FREPs).

(a) Alignment of the coding sequences of *P. acuta* FREPs 1 and 2 and single IgSF domain-containing FREPs of *B. glabrata* (accessions: ADE45330, AAk13551, AQX34557). Signal peptides of FREP sequences are boxed. Stars designate conserved cysteine residues of the IgSF domains. Secondary protein structure predictions were performed using JPred v4 on the IgSF domains of *B. glabrata* and *P. acuta* FREPs to identify β sheet organization, canonical protein structures of IgSF domains (highlighted in blue). Blue bars beneath the sequence alignment show the general locations of β sheet regions of gastropod FREP IgSF domains. The FBG domains are indicated with a solid line above the amino acid alignment. (b) Top, domain organization of *P. acuta* FREPs including signal peptide (SP), immunoglobulin superfamily domain (IG), and fibrinogen domain (FBG). Sequences without similarity to known domains (-) separate the SP, IG domains and IG, FBG domains. Bottom, intron/exon structures of *P. acuta* FREPs 1 and 2. Exons are numbered and depicted as alternating white and grey boxes, arrowheads indicate locations of introns (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

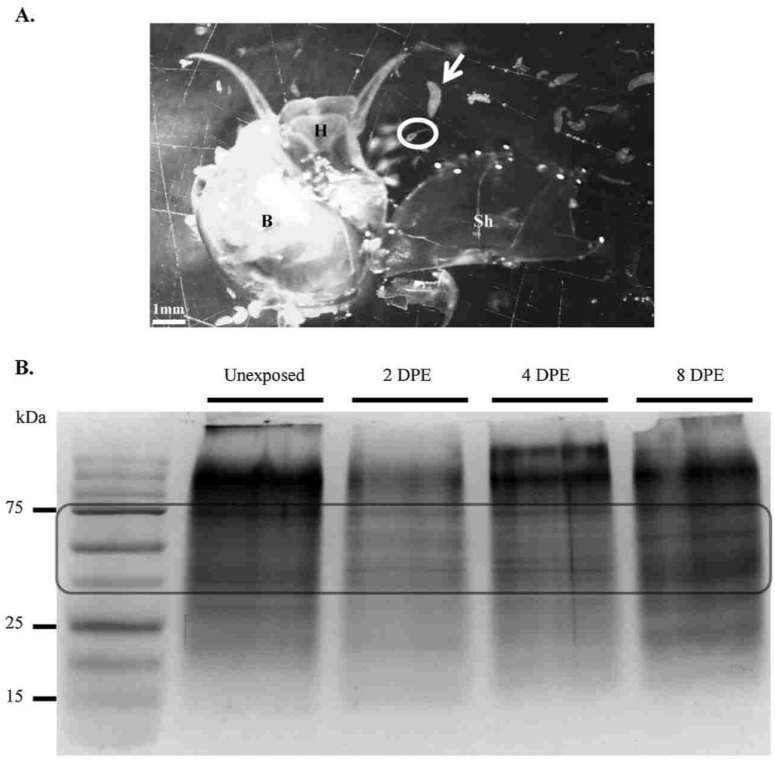


Fig. 3. *P. acuta*: viable infection with the trematode parasite *E. paraensei*.
(a) Stereomicroscope image of a patently infected *P. acuta* snail four weeks post exposure to ten *E. paraensei* miracidia. The upper right of the image shows parasite sporocysts (arrow) and cercariae (circle), bottom left shows a partially dissected *P. acuta* snail with shell removed (B, snail body; H, headfoot; Sh, shell). (b) SDS/PAGE gradient (5–20%) gel of cell-free plasma from three individual *P. acuta* snails at 2, 4, and 8 DPE to ten *E. paraensei* miracidia and one unexposed snail. The range of predicted molecular weights (43.3 and 53.53 kDa) of *P. acuta* FREPs (box) does not show obvious changes in abundance of plasma proteins.

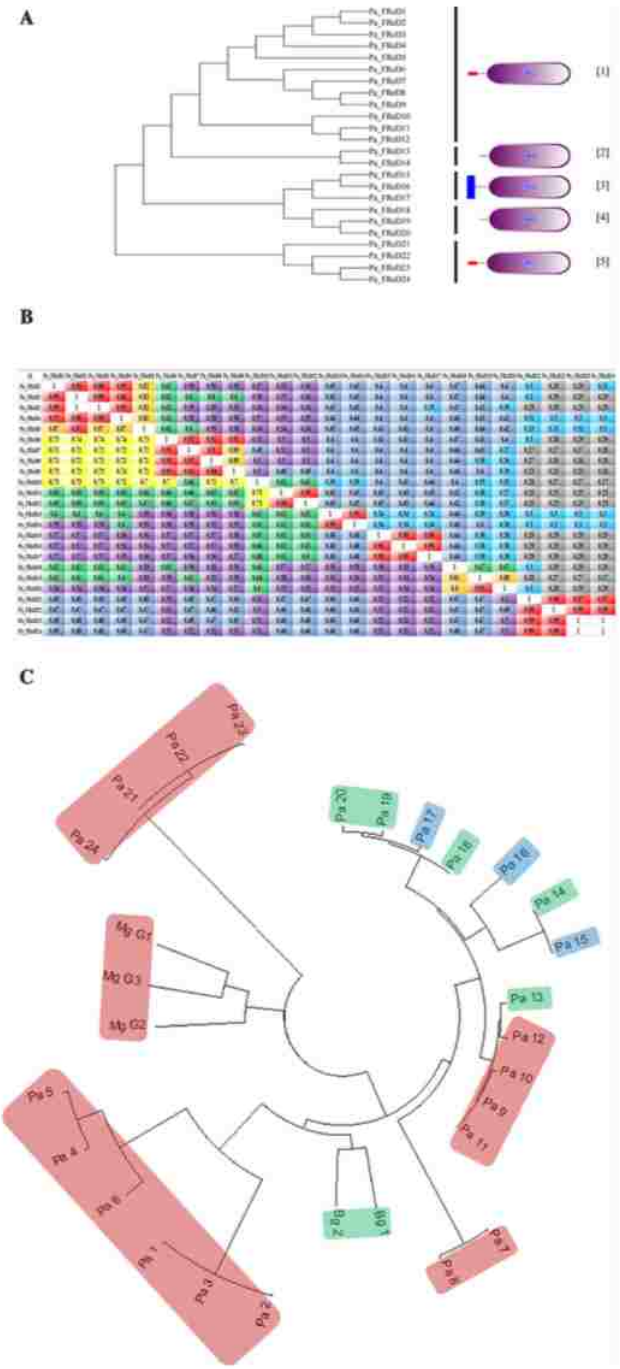


Fig. 4. Fibrinogen domain-containing sequences (FReDs) of *P. acuta*.

(a) Phylogenetic reconstruction of *P. acuta* FReDs. Twenty-four amino acid sequences, representing regions of the fibrinogen domain, were aligned. To the right of the phylogenetic tree are graphical depictions of the full-length predicted amino acid sequence domains: purple: FBG domain, red: signal peptide domain, blue: transmembrane domain. Sequences lacking either predicted signal peptide or transmembrane domains represent putative cytosolic FReD proteins. Numbers in brackets correspond to clusters of related amino acid sequences. (b) Matrix showing percent similarity and percent identity of pair wise alignments of *P. acuta* FReD amino acid sequences. Percent identity is indicated above diagonal (white) and percent similarity below the diagonal. Range of percent similarities and identities are as follows: 90–100%, red; 80–89%, orange; 70–79%, yellow; 60–69%, green; 50–59%, purple; 40–49%, dark blue; 30–39%, light blue; 20–29%, grey. (c) Phylogenetic analysis of the 24 *P. acuta* FReD sequences, two FReD sequences of *B. glabrata* (Bg) and three FReD sequences of the bivalve *M. galloprovincialis* (Mg). See materials and methods for accession numbers. Secreted-type FReDs (red), cell-bound FReDs (blue), cytosolic FReDs (green) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Table 1
Number of unique transcripts of select antimicrobial proteins predicted from the reference transcriptome of *P. acuta*.

Immune Factor	# of predicted sequences
Aplysianin/Achacin	12
FREPs	2
LBP/BPI	4
PGRP	7
Physalysin	3

Table 2
Number of consensus sequences detected after alignment of 24 sub-cloned nucleotide sequences of the housekeeping genes *cytochrome oxidase 1* (cox1) and *eukaryotic translation initiation factor 2A* (EIF2a) and *fibrinogen-related protein 1* (Pa_FREP1) and *fibrinogen-related protein 2* (Pa_FREP2) of *P. acuta*.

Gene	# of non-consensus sequences recovered among 24 cloned amplicons
COX1	1
EIF2a	0
Pa_FREP1	0
Pa_FREP2	2 (allelic difference, not polymorphic sequences)

References

1. Adema, C.M., 2015. Fibrinogen-related proteins (FREPs) in mollusks. In: Hsu, E., Du Pasquier, L. (Eds.), Pathogen-Host Interactions: Antigenic Variation v. Somatic Adaptations. Springer International Publishing, Switzerland, pp. 111–129.

2. Adema, C.M., Loker, E.S., 2015. Digenean-gastropod host associations inform on aspects of specific immunity in snails. Dev. Comp. Immunol. 48, 275–283.

3. Adema, C.M., Hertel, L.A., Miller, R.D., Loker, E.S., 1997. A family of fibrinogen-related proteins that precipitates parasite-derived molecules is produced by an invertebrate after infection. PNAS 94, 8691–8696.

4. Adema, C.M., Hertel, L.A., Loker, E.S., 1999. Evidence from two planorbid snails of a complex and dedicated response to digenean (echinostome) infection. Parasitology 119, 395–404.

5. Adema, C.M., Hanington, P.C., Lun, C.M., Rosenberg, G.H., Aragon, A.D., Stout, B.A., Richard, M.L.L., Gross, P.S., Loker, E.S., 2010. Differential transcriptomic responses of *Biomphalaria glabrata* (Gastropoda, Mollusca) to bacteria and metazoan parasites, *Schistosoma mansoni* and *Echinostoma paraensei* (Digenea, Platyhelminthes). Mol. Immunol. 47, 849–860.

6. Adema, C.M., Hillier, L.W., Jones, C.S., Loker, E.S., Knight, M., Minx, P., Oliveira, G., Raghavan, N., Shedlock, A., do Amaral, L.R., Arican-Goktas, H.D., Assis, J.G., Baba, E.H., Baron, O.L., Bayne, C.J., Bickham-Wright, U., Biggar, K.K., Blouin, M., Bonning, B.C., Botka, C., Bridger, J.M., Buckley, K.M., Buddenborg, S.K., Lima, Caldeira, R., Carleton, J., Carvalho, O.S., Castillo, M.G., Chalmers, I.W., Christensens, M., Clifton, S., Cosseau, C., Coustau, C., Cripps, R.M., Cuesta-Astroz, Y., Cummins, S.F., di Stephano, L., Dinguirard, N., Duval, D., Emrich, S., Feschotte, C., Feyereisen, R., FitzGerald, P., Fronick, C., Fulton, L., Galinier, R., Gava, S.G., Geusz, M., Geyer, K.K., Geraldo-Calderón, G.I., de Souza Gomes, M., Gordy, M.A., Gourbal, B., Grunau, C., Hanington, P.C., Hoffmann, K.F., Hughes, D., Humphries, J., Jackson, D.J., Jannotti- Passos, L.K., de, Jesus, Jeremias, W., Jobling, S., Kamel, B., Kapusta, A., Kaur, S., Koene, J.M., Kohn, A.B., Lawson, D., Lawton, S.P., Lìang, D., Limpanont, Y., Liu, S., Lockyer, A.E., Lovasto, T.L., Ludolf, F., Magrini, V., McManus, D.P., Medina, M., Misra, M., Mitta, G., Mkoji, G.M., Montague, M.J., Montelongo, C., Moroz, L.L., Munoz-Torres, M.C., Niazi, U., Noble, L.R., Oliveira, F.S., Pais, F.S., Papenfuss, A.T., Peace, R., Pena, J.J., Pila, E.A., Quelais, T., Raney, B.J., Rast, J.P., Rollinson, D., Rosse, I.C., Rotgans, B., Routledge, E.J., Ryan, K.M., Scholte, L.L.S., Storey, K.B., Swain, M., Tennessen, J.A., Tomlinson, C., Trujillo, D.L., Volpi, E.V., Walker, A.J., Wang, T., Wannaporn, I., Warren, W.C., Wu, X.J., Yoshino, T.P., Yusuf, M., Zhang, S.M., Zhao, M., Wilson, R.K., 2017. Whole genome analysis of a schistosomiasis- transmitting freshwater snail. Nat. Commun. 8, 15451.

7. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

8. Baeza Garcia, A., Pierce, R.J., Gourbal, B., Werkmeister, E., Colinet, D., Reichart, J.M., Dissous, C., Coustau, C., 2010. Involvement of the cytokine MIF in the snail host immune response to the parasite *Schistosoma mansoni*. PLoS Pathogens 6 e1001115.

9. Baron, O.L., van West, P., Industri, B., Pouchet, M., Dubreuil, G., Gourbal, B., Reichart, J.M., Coustau, C., 2013. Parental transfer of the antimicrobial protein LBP/BPI protects *Biomphalaria glabrata* eggs against oomycete infections. PLoS Pathogens 9 e1003792.

10. Baron, O.L., Deleury, E., Reichart, J.M., Coustau, C., 2016. The LBP/BPI multigenic family in invertebrates: evolutionary history and evidences of specialization in mollusks. Dev. Comp. Immunol. 57, 20–30.

11. Bayne, C.J., 2009. Successful parasitism of vector snail *Biomphalaria glabrata* by the human fluke (trematode) *Schistosoma mansoni*: a 2009 assessment. Mol. Biochem. Parasitol. 165, 8–18.

12. Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics 30, 2112–2120.

13. Bouchet, P., Rocroi, J.P., 2005. Classification and nomenclator of gastropod families. Malacologia 397.

14. Bouchut, A., Coustau, C., Gourbal, B., Mitta, G., 2007. Compatibility in the *Biomphalaria glabrata*/*Echinostoma caproni* model: new candidate genes evidenced by a suppressive subtractive hybridization approach. Parasitology 134, 575–588.

15. Buchfink, B., Xie, C., Huson, D.H., 2015. Fast and sensitive protein alignment using DIAMOND. Nat. Methods 12, 59–60.

16. Buckley, K.M., Rast, J.P., 2015. Diversity of animal immune receptors and the origins of recognition complexity in the deuterostomes. Dev. Comp. Immunol. 49, 179–189.

17. Buddenborg, S.K., Bu, L., Zhang, S.M., Schilkey, F.D., Mkoji, G.M., Loker, E.S., 2017. Transcriptomic responses of *Biomphalaria pfeifferi* to *Schistosoma mansoni*: Investigation of a neglected African snail that supports more *S. mansoni* transmission than any other snail species. PLoS Negl. Trop. Dis. 11 e0005984.

18. Camargo, J.A., Alonos, Á., 2017. Ecotoxilogical assessment of the impact of fluoride (F-) and turbidity on the freshwater snail *Physella acuta* in a polluted river receiving an industrial effluent. Environ. Sci. Pollut. Res. Int. 24, 15667–15677.

19. Chevreux, B., Wetter, T., Suhai, S., 1999. Genome sequence assembly using trace signals and additional sequence information. Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB) 99, 45–56.

20. Christophides, G.K., Zdobnov, E., Barillas-Mury, C., Birney, E., Blandin, S., Blass, C., Brey, P.T., Collins, F.H., Danielli, A., Dimopoulos, G., Hetru, C., Hoa, N.T., Hoffmann, J.A., Kanzok, S.M., Letunic, I., Levashina, E.A., Loukeris, T.G., Lycett, G., Meister, S., Michel, K., Moita, L.F., Müller, H.M., Osta, M.A., Paskewitz, S.M., Reichart, J.M., Rzhetsky, A., Troxler, L., Vernick, K.D., Vlachou, D., Volz, J., von Mering, C., Xu, J., Zheng, L., Bork, P., Kafatos, F.C., 2002. Immunity-related genes and gene families in *Anopheles gambiae*. Science 298, 159–165.

21. Couch, L., Hertel, L.A., Loker, E.S., 1990. Humoral response of the snail *Biomphalaria glabrata* to trematode infection: observations on a circulating hemagglutinin. J. Exp. Zool. 255, 340–349.

22. Coustau, C., Gourbal, B., Duval, D., Yoshino, T.P., Adema, C.M., Mitta, G., 2015. Advances in gastropod immunity from the study of the interaction between the snail *Biomphalaria glabrata* and its parasites: a review of research progress over the last decade. Fish. Shellf. Immunol. 46, 5–16.

23. Deleury, E., Dubreuil, G., Elangovan, N., Wajnberg, E., Reichart, J.M., Gourbal, B., Duval, D., Baron, O.L., Gouzy, J., Coustau, C., 2012. Specific versus non-specific immune responses in an invertebrate species evidenced by a comparative *de novo* sequencing study. PLoS One 7 e32512.

24. Dheilly, N.M., Adema, C.M., Raftos, D.A., Gourbal, B., Grunau, C., Du Pasquier, L., 2014. No more non-model species: the promise of next generation sequencing for comparative immunology. Dev. Comp. Immunol. 45, 56–66.

25. Dheilly, N., Duval, D., Mouahid, G., Emans, R., Allienne, J.F., Galinier, R., Genthon, C., Dubois, E., Du Pasquier, L., Adema, C.M., Grunau, C., Mitta, G., Gourbal, B., 2015. A family of variable immunoglobulin and lectin domain containing molecules in the snail *Biomphalaria glabrata*. Dev. Comp. Immunol. 48, 234–243.

26. Dillon Jr, R.T., Wethington, A.R., Lydeard, C., 2011. The evolution of reproductive isolation in a simultaneous hermaphrodite, the freshwater snail *Physa*. BMC Evol. Biol. 11, 144.

27. Dimopoulos, G., Christophides, G.K., Meister, S., Schultz, J., White, K.P., Barillas-Mury, C., Kafatos, F.C., 2002. Genome expression analysis of *Anopheles gambiae*: responses to injury, bacterial challenge, and malaria infection. Proc. Natl. Acad. Sci. U. S. A. 99, 8814–8819.

28. Drozdetskiy, A., Cole, C., Procter, J., Barton, G.J., 2015. JPred4: a protein secondary structure prediction server. Nucleic Acids Res. 43, W389–W394.

29. Egesa, M., Hoffmann, K.F., Hokke, C.H., Yazdanbakhsh, M., Cose, S., 2017. Rethinking schistosomiasis vaccine development: synthetic vesicles. Trends Parasitol. 30187–30193.

30. Finn, R.D., Attwood, T.K., Babbitt, P.C., Bateman, A., Bork, P., Bridge, A.J., Chang, H.Y., Dosztányi, Z., El-Gebali, S., Fraser, M., Gough, J., Haft, D., Holliday, G.L., Huang, H., Huang, X., Letunic, I., Lopez, R., Lu, S., Marchler-Bauer, A., Mi, H., Mistry, J., Natale, D.A., Necci, M., Nuka, G., Orengo, C.A., Park, Y., Pesseat, S., Piovesan, D., Potter, S.C., Rawlings, N.D., Redaschi, N., Richardson, L., Rivoire, C., Sangrador-Vegas, A., Sigrist, C., Sillitoe, I., Smithers, B., Squizzato, S., Sutton, G., Thanki, N., Thomas, P.D., Tosatto, S.C., Wu, C.H., Xenarios, I., Yeh, L.S., Young, S.Y., Mitchell, A.L., 2017. InterPro in 2017-beond protein family and domain annotations. Nucleic Acids Res. 45, D190–D199.

31. Galinier, R., Portela, J., Moné, Y., Allienne, J.F., Henri, H., Delbecq, S., Mitta, G., Gourbal, B., Duval, D., 2013. Biomphalysin, a new β pore-forming toxin involved in *Biomphalaria glabrata* immune defense against *Schistosoma mansoni*. PLoS Pathogens 9, e1003216.

32. Galinier, R., Roger, E., Moné, Y., Duval, D., Portet, A., Pinaud, S., Chaparro, C., Grunau, C., Genthon, C., Dubois, E., Rognon, A., Arancibia, N., Dejean, B., Théron, A., Gourbal, B., Mitta, G., 2017. A multistrain approach to studying the mechanisms underlying compatibility in the interaction between *Biomphalaria glabrata* and *Schistosoma mansoni*. PLoS Negl. Trop. Dis. 11 e0005398.

33. Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Apple, R.D., Bairoch, A., 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 31, 81–84.

34. Gerdol, M., De Moro, G., Manfrin, C., Venier, P., Pallavicini, A., 2012. Big defensins and mytimacins, new AMP families of the Mediterranean mussel *Mytilus galloprovincialis*. Dev. Comp. Immunol. 36, 390–399.

35. Ghosh, J., Lun, C.M., Majeske, A.J., Sacchi, S., Schrankel, C.S., Smith, L.C., 2011. Invertebrate immune diversity. Dev. Comp. Immunol. 35, 959–974.

36. Gorbushin, A.M., Borisova, E.A., 2015. Lectin-like molecules in transcriptome of *Littorina littorea* hemocytes. Dev. Comp. Immunol. 48, 210–220.

37. Gorbushin, A.M., Iakovleva, N.V., 2011. A new gene family of single fibrinogen domain lectins in *Mytilus*. Fish. Shellf. Immunol. 30, 434–438.

38. Gorbushin, A.M., Panchin, Y.V., Iakovleva, N.V., 2010. In search of the origin of FREPs: characterization of californica fibrinogen-related proteins. Dev. Immunol. 34, 465–473.

39. Gordy, M.A., Pila, E.A., Hanington, P.C., 2015. The role of fibrinogen-related proteins in the gastropod immune response. Fish. Shellf. Immunol. 46, 39–49.

40. 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., Friendman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652.

41. Gregory, T.R., 2017. Animal Genome Size Database. http://www.genomesize.com.

42. Grimm-Jørgensen, Y., 1983. Immunoreactive somatostatin in two pulmonate gastropods. Gen. Comp. Endocrinol. 49, 108–114.

43. Grimm-Jørgensen, Y., 1987. Somatostatin and calcitonin modulate gastropod internal defense mechanisms. Dev. Comp. Immunol. 11, 487–499.

44. Guo, Y., He, H., 2014. Identification and characterization of a goose-type lysozyme from sewage snail *Physa acuta*. Fish. Shellfish. Immunol. 39, 321–325.

45. Hanelt, B., Lun, C.M., Adema, C.M., 2008. Comparative ORESTES-sampling of transcriptomes of immune-challenged *Biomphalaria glabrata* snails. J. Invertebr. Pathol. 99, 192–203.

46. Hanington, P.C., Forys, M.A., Dragoo, J.W., Zhang, S.M., Adema, C.M., Loker, E.S., 2010. Role for a somatically diversified lectin in resistance of an invertebrate to parasite infection. PNAS 107, 21087–21092.

47. Hanington, P.C., Forys, M.A., Loker, E.S., 2012. A somatically diversified defense factor, FREP3, is a determinant of snail resistance to schistosome infection. PLoS Negl. Trop. Dis. 6, e1591.

48. Hanington, P.C., Zhang, S.M., 2011. The primary role of fibrinogen-related proteins in invertebrates is defense, not coagulation. J. Innate Immun. 3, 17–27.

49. Hathaway, J.J.M., Adema, C.M., Stout, B.A., Mobarak, C.D., Loker, E.S., 2010. Identification of protein components of egg masses indicates parental investment in immunoprotection of offspring by *Biomphalaria glabrata* (Gastropoda, Mollusca). Dev. Comp. Immunol. 34, 425–435.

50. Hertel, L.A., Adema, C.M., Loker, E.S., 2005. Differential expression of FREP genes in two strains of *Biomphalaria glabrata* following exposure to the digenetic trematodes *Schistosoma mansoni* and *Echinostoma paraensei*. Dev. Comp. Immunol. 29, 295–303.

51. Humphries, J.E., Yoshino, T.P., 2003. Cellular receptors and signal transduction in molluscan hemocytes: connections with the innate immune system of vertebrates. Integr. Comp. Biol. 43, 305–312.

52. Ittiprasert, W., Miller, A., Myers, J., Nene, V., El-Sayed, N.M., Knight, M., 2010. Identification of immediate response genes dominantly expressed in juvenile resistant and susceptible *Biomphalaria glabrata* snails upon exposure to *Schistosoma mansoni*. Mol. Biochem. Parasitol. 169, 27–39.

53. Jack, R.S., 2015. Evolution of immunity and pathogens. In: Hsu, E., Du Pasquier, L. (Eds.), Pathogen-host Interactions: Antigenic variation v. Somatic Adaptations. Springer International Publishing, Switzerland, pp. 1–20.

54. Jörger, K.M., Stöger, I., Kano, Y., Fukuda, H., Knebelsberger, T., Schrödl, M., 2010. On the origin of Acochlidia and other enigmatic euthyneuran gastropods, with implications for the systematics of Heterobranchia. BMC Evol. Biol. 10, 323.

55. Kamiya, H., Muramoto, K., Yamazaki, M., 1986. aplysianin-A, an antibacterial and antineoplastic glycoprotein in the albumen gland of A sea hare, *Aplysia kurodai*. Experientia 42, 1065–1067.

56. Kasai, K., Ishikawa, T., Nakamura, T., Miura, T., 2015. Antibacterial properties of L-amino acid oxidase: mechanisms of action and perspectives for therapeutic applications. Appl. Microbiol. Biotechnol. 99, 7847–7857.

57. Knight, M., Ittiprasert, W., Arcian-Goktas, H.D., Bridger, J.M., 2016. Epigenetic modulation, stress and plasticity in susceptibility of the snail host, *Biomphalaria glabrata*, to *Schistosoma mansoni* infection. Int. J. Parasitol. 46, 389–394.

58. Larken, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947–2948.

59. Lee, Y.S., Lee, S., Kang, S.W., Jeong, J.E., Baek, M.K., Choi, S., Chae, S., Jo, Y.H., Han, Y.S., Park, H., 2011. Expressed sequence tag analysis of *Physa acuta*: a freshwater pulmonate in Korea. J. Shellf. Res. 30, 127–132.

60. Léonard, P.M., Adema, C.M., Zhang, S.M., Loker, E.S., 2001. Structure of two FREP genes that combine IgSF and fibrinogen domains, with comments on diversity of the FREP gene family in the snail *Biomphalaria glabrata*. Gene 269, 155–165.

61. Leoni, G., De Poli, A., Mardirossian, M., Gambato, S., Florian, F., Venier, P., Wilson, D.N., Tossi, A., Pallavicini, A., Gerdol, M., 2017. Myticalins: A novel multigenic family of linear, cationic antimicrobial peptides from marine mussels (*Mytilus* spp.). Mar. Drugs 15 E261.

62. Letunic, I., Bork, P., 2017. 20 years of the SMART protein domain annotation resource. Nucleic Acids Res. http://dx.doi.org/10.1093/nar/gkx922.

63. Li, B., Dewey, C.N., 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12, 323.

64. Li, W., Godzik, A., 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22, 1658–1659.

65. Lie, K.J., Basch, P.F., 1967. The life history of *Echinostoma paraensei* sp. N. (Trematoda: Echinostomatidae). J. Parasitol. 53, 1192–1199.

66. Litman, G.W., Cooper, M.D., 2007. Why study the evolution of immunity? Nat. Immunol. 8, 547–548.

67. Loker, E.S., Hertel, L.A., 1987. Alterations in *Biomphalaria glabrata* plasma induced by infection with the digenetic trematode *Echinostoma paraensei*. J. Parasitol. 73, 503–513.

68. Loker, E.S., Cimino, D.F., Hertel, L.A., 1992. Excretory-secretory products of *Echinostoma paraensei* sporocysts mediate interference with *Biomphalaria glabrata* hemocyte functions. J. Parasitol. 78, 104–115.

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

70. Lu, X., Wang, M., Qi, J., Wang, H., Li, X., Gupta, D., Dziarski, R., 2006. Peptidoglycan recognition proteins are a new class of human bactericidal proteins. J. Biol. Chem. 281, 5895–5907.

71. Martínez-Paz, P., Morales, M., Sánchez-Argüello, P., Morcillo, G., Martínez-Guitarte, J.L., 2017. Cadmium in vivo exposure alters stress response and endocrine-related genes in the freshwater snail *Physa acuta*. New biomarker genes in a new model organism. Environ. Pollut. 220, 1488–1497.

72. Mitta, G., Vandenbulcke, F., Roch, P., 2000. Original involvement of antimicrobial peptides in mussel innate immunity. FEBS Lett. 486, 185–190.

73. Moné, Y., Gourbal, B., Duval, D., Du Pasquier, L., Kieffer-Jaquinod, S., Mitta, G., 2010. A large repertoire of parasite epitopes matched by a large repertoire of host immune receptors in an invertebrate host/parasite model. PLoS Negl. Trop. Dis. 4 e813.

74. Müller, V., de Boer, R.J., Bonhoeffer, S., Szathmáry, E., 2017. An evolutionary perspective on the systems of adaptive immunity. Biol. Rev. Camb. Philos. Soc. http://dx.doi.org/10.1111/brv.12355.

75. Mundry, M., Bornberg-Bauer, E., Sammeth, M., Feulner, P.G.D., 2012. Evaluating characteristics of *de novo* assembly software on 454 transcriptome data: a simulation approach. PLoS One 7 e31410.

76. Mylonakis, E., Podsiadlowski, L., Muhammed, M., Vilcinskas, A., 2016. Diversity, evolution and medical applications of insect antimicrobial peptides. Philos. Trans. R. Soc. Lond. B Biol. Sci. 371.

77. Nishimura, O., Hara, Y., Kuraku, S., 2017. gVolante for standardizing completeness assessment of genome and transcriptome assemblies. Bioinformatics. http://dx.doi.org/10.1093/bioinformatics/btx445.

78.

79. Nolan, J.R., Bergthorsson, U., Adema, C.M., 2014. *Physella acuta*: atypical mitochondrial gene order among panpulmonates (Gastropoda). J. Mollus. Stud. 1–12.

80. Park, J.W., Kim, C.H., Kim, J.H., Je, B.R., Roh, K.B., Kim, S.J., Lee, H.H., Ryu, J.H., Lim, J.H., Oh, B.H., Lee, W.J., Ha, N.C., Lee, B.L., 2007. Clustering of peptidoglycan recognition protein-SA is required for sensing lysine-type peptidoglycan in insects. Proc. Natl. Acad. Sci. U. S. A. 104, 6602–6607.

81. Petersen, T.N., Brunak, S., von Heijne, G., Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods 8, 785–786.

82. Pila, E.A., Tarrabain, M., Kabore, A.L., Hanington, P.C., 2016a. A novel toll-like receptor (TLR) influences compatibility between the gastropod *Biomphalaria glabrata*, and the digenean trematode *Schistosoma mansoni*. PLoS Pathog. 12 e1005513.

83. Pila, E.A., Gordy, M.A., Phillips, V.K., Kabore, A.L., Rudko, S.P., Hanington, P.C., 2016b. Endogenous growth factor stimulation of hemocyte proliferation induces resistance to *Schistosoma mansoni* challenge in the snail host. Proc. Natl. Acad. Sci. U. S. A. 113, 5305–5310.

84. Podobnik, M., Anderluh, G., 2017. Pore-forming toxins in Cnidaria. Semin. Cell Dev. Biol. http://dx.doi.org/10.1016/j.semcdb.2017.07.026.

85. Premachandra, H.K., Elvitigala, D.A., Whang, I., Lee, J., 2014. Identification of a novel molluscan short-type peptidoglycan recognition protein in disk abalone (*Haliotis discus discus*) involved in host antibacterial defense. Fish. Shellf. Immunol. 39, 99–107.

86. Romero, A., Dios, S., Poisa-Beiro, L., Costa, M.M., Posada, D., Figueras, A., Novoa, B., 2011. Individual sequence variability and functional activities of fibrinogen-related proteins (FREPs) in the Mediterranean mussel (*Mytilus galloprovincialis*) suggest ancient and complex immune recognition models in invertebrates. Dev. Comp. Immunol. 35, 334–344.

87. Royet, J., Gupta, D., Dziarski, R., 2011. Peptidoglycan recognition proteins: modulators of the microbiome and inflammation. Nat. Rev. Immunol. 11, 837–851.

88. Schultz, J.H., Adema, C.M., 2017. Comparative immunogenomics of mollusks. Dev. Comp. Immunol. 75, 3–15.

89. Steiner, H., 2004. Peptidoglycan recognition proteins: on and off switches for innate immunity. Immunol. Rev. 198, 83–96.

90. Sullivan, J.T., Belloir, J.A., 2014. Activation of an innate immune response in the schistosome-transmitting snail *Biomphalaria glabrata* by specific bacterial PAMPs. Dev. Comp. Immunol. 42, 256–260.

91. Suwannatrai, K., Suwannatrai, A., Tabsripair, P., Welbat, J.U., Tangkawattana, S., Cantacessi, C., Mulvenna, J., Tesana, S., Loukas, A., Sotillo, J., 2016. Differential protein expression in the hemolymph of *Bithynia siamensis goniomphalos* infected with *Opisthorchis viverrini*. PLoS Negl. Trop. Dis. 10 e0005104.

92. Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.

93. Tanveer, A., 1991. Studies on the morphology of hemocytes of *Physa acuta* and *Bellamaya bengalensis*. Punjab Univ. J. Zool. 6, 7–18.

94. Tasiemski, A., Vandenbulcke, F., Mitta, G., Lemoine, J., Lefebvre, C., Sautière, P.E., Salzet, M., 2004. Molecular characterization of two novel antibacterial peptides inducible upon bacterial challenge in an annelid, the leech *Theromyzon tessulatum*. J. Biol. Chem. 279, 30973–30982.

95. Tetreau, G., Pinaud, S., Portet, A., Galinier, R., Gourbal, B., Duval, D., 2017. Specific pathogen recognition by multiple innate immune sensors in an invertebrate. Front. Immunol. 8, 1249.

96. Wang, G., Li, X., Wang, Z., 2009. APD2: the updated antimicrobial peptide database and its application in peptide design. Nucleic Acids Res. 37, 933–937.

97. Wethington, A.R., Lydeard, C., 2007. A molecular phylogeny of Physidae (Gastropoda: Basommatophora) based on mitochondrial DNA sequences. J. Molluscan Stud. 73, 241–257.

98. White, T.R., Conrad, M.M., Tseng, R., Balayan, S., Golding, R., de Frias Martins, A.M., Dayrat, B.A., 2011. Ten new complete mitochondrial genomes of pulmonates (Mollusca: Gastropoda) and their impact on phylogenetic relationships. BMC Evol. Biol. 11, 295.

99. Winnepenninckx, B., Backeljau, T., De Wachter, R., 1993. Extraction of high molecular weight DNA from molluscs. Trends Genet. 9, 407.

100. Wu, X.J., Dinguirard, N., Sabat, G., Lui, H.D., Gonzalez, L., Gehring, M., Bickham-Wright, U., Yoshino, T.P., 2017. Proteomic analysis of *Biomphalaria glabrata* plasma proteins with binding affinity to those expressed by early developing larval *Schistosoma mansoni*. PLoS Pathog. 13 e1006081.

101. Yoshino, T.P., Wu, X.J., Liu, H., Gonzalez, L.A., Deelder, A.M., Hokke, C.H., 2012. Glycotope sharing between snail hemolymph and larval schistosomes: larval transformation products alter shared glycan patterns of plasma proteins. PLoS Negl. Trop. Dis. 6, e1569.

102. Zaidman-Rémy, A., Hervé, M., Poidevin, M., Pili-Floury, S., Kim, M.S., Blanot, D., Oh, B.H., Ueda, R., Mengin-Lecreulx, D., Lemaitre, B., 2006. The *Drosophila* amidase PGRP-LB modulates the immune response to bacterial infection. Immunity 24, 463–473.

103. Zdobnov, E.M., von Mering, C., Letunic, I., Torrents, D., Suyama, M., Copley, R.R., Christophides, G.K., Thomasova, D., Holt, R.A., Subramanian, G.M., Mueller, H.M., Dimopoulus, G., Law, J.H., Wells, M.A., Birney, E., Charlab, R., Halpern, A.L., Kokoza, E., Kraft, C.L., Lai, Z., Lewis, S., Louis, C., Barillas-Mury, C., Nusskern, D., Rubin, G.M., Salzberg, S.L., Sutton, G.G., Topalis, P., Wides, R., Wincker, P., Yandell, M., Collins, F.H., Ribeiro, J., Gelbart, W.M., Kafatos, F.C., Bork, P., 2002. Comparative genome and proteome analysis of *Anopheles gambiae* and *Drosophila melanogaster*. Science 298, 149–159.

104. Zhang, S.M., Adema, C.M., Kepler, T.B., Loker, E.S., 2004. Diversification of Ig superfamily genes in an invertebrate. Science 305, 251–254.

105. Zhang, S.M., Zeng, Y., Loker, E.S., 2007. Characterization of immune genes from the schistosome host snail *Biomphalaria glabrata* that encode peptidoglycan recognition proteins and gram-negative bacteria binding protein. Immunogenetics 59, 883–898.

106. Zhang, L., Li, L., Guo, X., Litman, G.W., Dishaw, L.J., Zhang, G., 2015. Massive expansion and functional divergence of innate immune genes in a protostome. Sci. Rep. 5, 8693.

Chapter 4: Characterization of the early host response of *Physella acuta* to the digenetic trematode *Echinostoma paraensei*

Jonathan H. Schultz, Lijing Bu, Bishoy Kamel, and Coen M. Adema

Center of Evolutionary and Theoretical Immunology (CETI), Department of Biology, University of New Mexico, New Mexico, USA

Abstract

This study analyzed the early response of the snail *Physella acuta* to *Echinosoma paraensei*, a previously described compatible snail/digenean trematode association, during the initial phases of infection, 2 and 8 days post exposure (DPE). RNA-seq Illumina data were collected and assembled from 5 snails at 2DPE, 4 snails at 8DPE and 7 unexposed (control) snails. A reference transcriptome (325,563 transcripts, 98% representation of eukaryotic universal single copy orthologs; BUSCO) was assembled to guide RNAseq analyses, and available genomic Illumina reads were applied toward a *P. acuta* genome assembly comprising 827,188 scaffolds (includes 88% BUSCO genes). Relative to controls, parasite-exposure of *P. acuta* led to 10,195 differentially expressed (DE) genes at 2DPE and 8,876 DE genes at 8DPE. Few sequence homologs (18% of upregulated, and 22% of downregulated sequences) were shared between 2 and 8DPE snails. Gene Ontology (GO) analysis provided annotation of only 1.2% of DE genes to interpret the response of *P. acuta* to the parasite. Increased insights were achieved by BLAST identification of 460 immune-relevant transcripts among the DE profiles. *Physella acuta* expressed many transcripts of the CD109/TEP, GTPase IMAP and *Limulus* agglutination factor (dermatopontin) immune gene families. Newly observed sequences indicated that *P. acuta* has at least 82 different members of the FReD gene family and disclosed transcripts that uniquely (among Metazoa) combined C-LECT and C1q domains. Notably, *P. acuta* expressed mostly different assemblages of transcripts from particular immune gene families between 2 and 8DPE. This shift in transcriptomic profile, especially involving immune factors, after parasite establishment in compatible *P. acuta*, raises the notion that *E. paraensei* may manipulate the snail host to express immune genes not targeting the parasite itself, but only counter other types of (opportunistic) pathogens. We propose that this represents an extended phenotype of

E. paraensei, imposed upon *P. acuta* following successful parasite establishment in the snail after 2DPE. The extensive genomic and transcriptomic resources provided will aid future investigations to further resolve the molecular underpinnings of the compatible interaction between *P. acuta* and *E. paraensei*.

Introduction

The Physidae, a family of freshwater panpulmonte snails, like many gastropods serve as obligatory intermediate hosts to a great diversity of trematode parasites (Adema et al., 2009; Barragán-Sáenz et al., 2009; Kraus et al., 2014; Ebbs et al., 2018; Muñoz-Antolí et al., 2000; Muñoz-Antolí et al., 2008; Toledo et al., 1998). Overall, snail-parasite compatibility displays varying degrees of specificity such that the host range for a particular parasite is restricted to a subset of snail taxa (Adema and Loker, 2015). This suggests that snail responses are tailored for distinct parasites and prevent most attempts at foreign invasion. Many efforts to elucidate the molecular mechanisms that determine snail-parasite associations have been motivated by the goal to curb transmission of human-relevant snail-transmitted diseases (Adema et al., 2012; Lu et al., 2018; World Health Organization, 2012). Investigations with non-human parasites transmitted by snails, such as echinostome trematodes however, have elucidated novel aspects of gastropod immunity that help interpret the evolution of immune function among protostome invertebrates and better understand host suitability for parasites (Loker et al., 2004). As an example, the concept of acquired resistance, a mechanism by which an invertebrate organism exhibits a more pronounced and effective immune response to a pathogen after a previous encounter, was described in snails after exposure to echinostome parasites (Lie and Heyneman, 1979; Lie et al., 1982, 1983). These findings instigated investigations of the cellular interactions of the snail host *B. glabrata* and its cognate echinostomes (Adema et al., 1994; Ataev and

Coustau, 1999; DeGaffé and Loker, 1998; Humbert and Coustau, 2001; Jeong et al., 1984; Joky et al., 1985; Lie et al., 1981; Loker et al., 1986; Loker et al., 1989; Loker et al., 1992; Mounkassa and Jourdane, 1990; Noda and Loker, 1989a, 1989b; Uchikawa and Loker, 1992). Study of the humoral response of *B. glabrata* to echinostomes have been pivotal in understanding these host-parasite combinations (Couch et al., 1990; Loker and Hertel, 1987; Monroy et al., 1992, 1993; Loker et al., 1994; Hertel et al., 1994; Tetraeu et al., 2017; Vergote et al., 2005). In particular, *B. glabrata* is capable of somatically diversifying carbohydrate-binding lectins known as fibrinogen-related proteins (FREPs) that are anti-trematode defense factors, thought to increase the range of non-self recognition (Adema et al., 1997, 1999, 2015; Hanington et al., 2010a, 2012; Moné et al., 2010; Zhang et al., 2004). Overall, these *B. glabrata*-echinostome parasite models contributed significantly to our understanding of gastropod immunobiology and elucidated novel invertebrate immune capabilities. It is important, however, to consider that there exists a great diversity of immune capabilities among animals (Buckley and Rast, 2015) and that even closely related species exhibit taxon-specific differences in how immunity is achieved (e.g. Schultz and Adema, 2017; Schultz et al., 2018). As such, investigation of diverse host-parasite models remain relevant and important to expand the scope of interpreting gastropod immunity.

Parasites and their snail hosts are locked in evolutionary arm races (Mitta et al., 2017) which have manifested in specific associations (Adema and Loker, 2015). However, some trematode parasites are generalists that exhibit a wider range of suitable snail hosts. The digenetic trematode *Echinostoma paraensei* is capable of infecting multiple families of snails across gastropod phylogeny, including the Lymnaeidae, Physidae, and aforementioned Planorbidae (Lie and Basch, 1967; Maldonado et al., 2001). Indeed, we recently demonstrated that a lab-strain of *E. paraensei*, originally collected from the field in Brazil (Lie and Basch,

1967) and maintained with *B. glabrata* (Loker et al., 1992), is capable of infecting the freshwater snail *Physella acuta* (Schultz et al., 2018). Remarkably, this new host-parasite combination showed that the response of *P. acuta* to *E. paraensei* is different from that of *B. glabrata* (Schultz et al., 2018). Typically, the fate of trematode parasites (*Schistosoma mansoni* and *E. paraensei*) in *B. glabrata* is determined early after exposure; changes in transcriptomic profiles show that the initial host immune response is downregulated 2 days post-exposure (DPE) (Hanington et al., 2010b). These changes occur during parasite development from miracaida to sporocysts (within hours after exposure) and at the onset of increased FREP expression (Adema et al., 1999; Hanington et al., 2010a). However, *P. acuta* does not increase FREP gene or protein expression after parasitic exposure and *P. acuta* does not somatically diversify FREP genes (Schultz et al., 2018). Investigations of *P. acuta* immunology are few in number (Grimm-Jørgensen, 1987, 1983; Guo and He, 2014; Lee et al., 2011; Schultz et al., 2018; Tanveer, 1991) and both nature and molecular basis of host interactions of *P. acuta* with *E. paraensei* during the early phase of infection remain unknown. In this study we aim to characterize the response of *P. acuta* to *E. paraensei* at 2DPE when *P. acuta* mounts efforts to counter colonization by the parasite and at 8DPE when the parasite has successfully negated the host immune response and established a viable infection. In the early stage of interaction, this approach will elucidate anti-digenean responses of *P. acuta*, ultimately to expand the view of gastropod immune function and characterize the general interaction between parasite and host.

RNA-seq supported by NGS genomic assembly afforded unbiased and comprehensive characterization of transcriptomic activities and used to survey the snail host responses of *P. acuta* 2 and 8DPE to the parasite *E. paraensei* with particular attention to the role of immune function of *P. acuta* to combat trematode infection (Schultz and Adema, 2017).

Materials and Methods

Snail and parasite strains

Physella acuta snails (isolate A; Nolan et al., 2014) and *E. paraensei* parasites (Loker et al., 1992) were utilized in this study and are maintained at the University of New Mexico.

Parasite exposures

For RNA-seq, adult *P. acuta* snails (3-4mm) were placed in 12-well plates with artificial spring water (ASW) and exposed to approximately 10 *E. paraensei* miracidiae overnight. The following day, wells were checked for the absence of free-swimming miracidae as a measure of infection. Snails were then placed in a new aquarium with ASW and fed red-leaf lettuce *ad libitum*. Prior to RNA extraction (Trizol, Carlsbad, CA) infections were confirmed at 2 and 8 DPE via observation through the shell of parasite sporocysts in the heart of *P. acuta* snails. Additionally, presence of echinostome DNA was confirmed from snail tissue by PCR amplification using mitochondrial ND1 primers (Morgan and Blair, 1998), Platinum™ Taq DNA Polymerase (Thermo Fisher, Waltham, MA) and Trizol DNA-extraction method (Trizol, Carlsbad, CA).

Genome assembly

The genome of *P. acuta* was previously sequenced (Illumina NextSeq 500, 150bp PE reads) utilizing two individual snails (Schultz et al., 2018; accessions SRX3347695, SRX3347220). The genomic reads were processed using Trimmomatic (Bolger et al., 2014) to remove low quality nucleotides and adaptor sequences added during Nextera (Illumina) library preparation. The reads were assembled using MEGAHIT v1.0 (Li et al., 2016) with a k-mer range 21-251. This primary assembly was improved by performing multiple iterations of Rascaf (Song et al., 2016), using RNAseq data (see below), SSPACE (Boetzer et al., 2011), and

GapCloser (Luo et al., 2012). Completeness of assembly was tested using gVolante (Nishimura et al., 2017) and BUSCO v2/v3 to identify core eukaryotic genes within the genome. The genome assembly was deposited in GenBank (RDRX00000000).

Transcriptome assembly and differential expression analyses

Previously collected RNAseq data (Illumina NextSeq 500, 150bp PE reads; Schultz et al., 2018; SRA accession SRR8080746) retrieved from 7 control (non-exposed), 5 2-DPE, and 4 8-DPE snails to 10 *E. paraensei* miracidia were used to generate a reference transcriptome. Briefly, raw reads were processed using Trimmomatic (Bolger et al., 2014). To filter out reads contributed by the parasite, the reads were aligned using Bowtie2 (Langmead and Salzberg, 2012) to the *Echinostoma caproni* genome (GCA_000950555.1), a close relative of *E. paraensei*. Filtered reads (paired-end) were assembled using Trinity *de novo* version 2.5.1 (Grabherr et al., 2011). Additionally, the contigs were mapped to the *P. acuta* genome using megablast (Altschul et al., 1990) to ascertain that transcripts were of snail origin. Contigs that aligned to the genome were processed using CD-HIT at a 95% clustering threshold to minimize contig-redundancy (Li and Godzik, 2006). Completeness assessment of the transcriptome was performed using gVolante and the BUSCO v2/v3 ortholog search pipeline (Nishimura et al., 2017). Finally, read abundance was quantified with RSEM (Li and Dewey, 2011) and comparisons of control and experimental transcripts for differential expression were performed using EBSeq using default parameters (Leng and Kendziorski, 2015).

Annotation of differentially expressed transcripts

Nucleotide sequences of *in silico* predicted transcripts, that were differentially expressed (DE) (up-regulated, down-regulated) at 2 and 8 DPE, were translated using TransDecoder v5.3.0

(https://github.com/TransDecoder/TransDecoder/archive/TransDecoder-v5.3.0) to obtain the longest open reading frames (ORFs, minumum of 100 amino acids encoded).

Translated sequences were assigned Gene Ontology (GO) identifiers via AgBase-GOanna (McCarthy et al., 2006) using the blastp algorithm, Ag-Uniprot database, and default parameters (except e-value threshold was set to $10e^{-06}$). Graphical interpretations of GO annotations were achieved via WEGO 2.0 (http://wego.genomics.org.cn/).

Nucleotide sequences of *in silico* predicted transcripts were annotated via DIAMOND blastx (Buchfink et al., 2015) with an e-value threshold of $10e^{-03}$. Annotations were assessed manually. The annotation of the BLAST hit with the highest similarity score was retained for each transcript. Immune-relevant transcript sequences were analysed in greater detail (see below).

Translated nucleotide sequences were also annotated using InterPro (Finn et al., 2017) to explore functional domains in sequences without significant blast hits, in order to broaden annotation of DE sequences additional to GO and BLAST annotation. Only immune-relevant domains were quantified and collated from up and downregulated sequences. Immune-relevant domains were identified by previous investigations of gastropod immunity (Hanelt et al., 2008).

Expanded and novel candidate immune-gene analysis

Sequences were selected based on keyword searches of annotations of DE transcripts. Only transcripts with the appropriate annotation and containing an open reading frame of at least 100 amino acids were included in downstream analysis. Translated sequences (amino acid level) were subjected to SMART (Letunic and Bork, 2017) protein domain prediction software. Sequences (aa level) of a particular gene family were aligned in Clustal X (Larken et al., 2007). Resulting sequence alignments were visually inspected and the terminal ends of sequences were

trimmed to reduce gaps within the alignment. Best-fit protein model predictions for phylogenetic analysis were performed with MEGA7 (Tamura et al., 2013). With the exception of the alignment of FReD amino acid sequences, all gaps and/or missing data were excluded from model predictions and phylogenetic tree reconstruction. All phylogenetic reconstructions utilized the Maximum Likelihood method, using the evolutionary models predicted by MEGA. The LG model was used for the CD109/TEP gene family, the WAG model for the complement C1q gene family, the JTT+G+I model for the dermatopontin gene family, the LG+G model for the GTPase IMAP gene family, and the WAG+G model for the FReD gene family.

Results

Parasitic exposure

All *P. acuta* snails survived the overnight exposure to *E. paraensei* and displayed normally active behavior until sampling.

P. acuta genome

The initial genome assembly (not shown) yielded 827,188 scaffolds with an N50 length of 1,289nt including 47.19% complete sequences of 303 eukaryotic core genes (BUSCO v2/v3). After 6 iterations of genome scaffold improvement (see methods), the final *P. acuta* genome (Fig. 1A) yielded 799,945 sequences with an N50 length of 1,358nt. The improved genome recorded 63.37% core eukaryotic genes and yielded 223,011 scaffolds greater than 1kb in length (Fig. 1A). The GC-content of the *P. acuta* genome (36.32%) is similar to that of other gastropods, which are AT-rich (Adema et al., 2017).

P. acuta reference transcriptome

The Illumina RNA-Seq reads generated from 16 *P. acuta* snails (isolate A, control and experimental samples) were assembled yielding a final reference transcriptome of 325,563

sequences with an N50 length of 1,262nt (Fig. 1B). Analysis of completeness (gVolante and BUSCO) of the reference transcriptome recorded 98% of the core eukaryotic genes and 225 transcripts greater than 10kb in length, a considerable improvement compared to a transcriptome generated using 454 NGS technology for *P. acuta* reported previously (Schultz et al., 2018). The two largest transcripts were 26,660 and 26,045 nucleotides in length with open reading frames of 9,378 and 8,392 amino acids long, respectively. Blastp analysis assigned best-hit annotations of uncharacterized proteins from *Aplysia californica* (XP_012944848.1, XP_012940635.1). Similar to the genome of *P. acuta*, the reference transcriptome also displayed AT-richness (Fig. 1B, GC-content 36.91%).

Differential expression

Relative to control snails, 2DPE snails yielded 10,195 DE transcripts of which 5,846 were upregulated and 4,349 were downregulated (Fig. 2A). A total of 8,876 transcripts (5,275 up-regulated, 3,601 down-regulated) were differentially expressed in 8DPE snails (Fig. 2A). Six transcripts up-regulated at 2DPE were subsequently down-regulated at 8DPE and 1,641 remained up-regulated at 8DPE (Fig. 2B). One down-regulated transcript at 2DPE was subsequently up-regulated at 8DPE and 1,469 remained down-regulated at 8DPE (Fig. 2B). In total, 4,205/2,985 transcripts were uniquely up-regulated at 2 and 8DPE respectively and 2,880/2,132 transcripts were uniquely down-regulated at 2 and 8DPE respectively.

GO annotation of DE sequences

Due to reliance on blast-similarity and a paucity of gastropod-derived entries in GenBank, GO analysis assigned annotation to only 1.2% (229) of DE sequences (Fig. 3A), providing 133 GO terms (level 2) assigned to biological process, 168 to cellular component, and 137 to molecular function (Fig. 3B,C). A total of 101 and 107 GO categories were assigned to

transcripts downregulated at 2 and 8DPE, respectively; and 94/136 GO categories were assigned to transcripts upregulated at 2 versus 8DPE (Fig. 3C). The categories “translation regulator activity”, “molecular carrier activity”, “transcription regulator activity”, “presynaptic process involved in chemical synaptic transmission”, and “cell proliferation” were assigned exclusively to upregulated sequences. In contrast, “signal transducer activity” was assigned uniquely to downregulated sequences.

Profiles of immune-relevant DE transcripts

A total of 460 DE immune-relevant transcripts were identified by BLAST from both experimental groups, of which 230 were uniquely DE at 2DPE, 141 were uniquely DE at 8DPE, and 89 were shared. Heatmap plots revealed individual transcripts that were differentially expressed 2 and 8DPE versus controls and were sorted by transcript name (Fig. 4A) and consecutively sorted by log2 fold changes of 2DPE transcripts followed by 8DPE transcripts (Fig. 4B). Only transcripts with a minimum fold change of 4 were included in the analyses. Note that variation in expression patterns of individual sequences does not necessarily reflect the expression patterns of the entire gene-family. As an example, the peroxiredoxin gene family (Fig. 4) is represented by five sequences. Despite each having their own unique expression profiles, there is always representation of some members of the peroxiredoxin gene family regardless of infection status.

InterPro domain annotation, performed to gain further insight into the function of transcripts not identified by GO or blast, yielded 28 categories of immune-relevant domains (Table 1). Four domain categories: alpha-2-macroglobulin, alpha-macroglobulin, serpin, and SRCR (highlighted in red), were absent from downregulated transcripts at 2 and 8DPE (Table 1). Immune-relevant domains consistently and abundantly present (C-type lectin, fibrinogen,

immunoglobulin, leucine-rich repeat, Sushi/SCR/CCP) among all experimental groups are highlighted in green (Table 1).

Expanded and novel candidate immune-gene repertoires

Select categories of DE immune-relevant genes were chosen for in depth analysis based on large numbers of transcripts representing gene families or novel combination of protein domains. These included five gene families of sequences encoding CD109/TEP, complement C1q, dermatopontin, GTPase immunity-associated proteins (IMAP), and fibrinogen-related domain (FReD) transcripts (Fig. 5). The sequences include complete and partial transcripts.

The CD109/TEP gene family of *P. acuta* is represented by 18 sequences (Fig. 5A); some with signal peptides and others lacking any discernible domain upstream of the thioester protein/alpha 2 macroglobulin (A2M) domains.

The dermatopontin gene family (proposed agglutination factors; Fujii et al., 1992) is represented by 23 DE sequences in *P. acuta* and includes transcripts containing signal peptides and transmembrane domains upstream of the dermatopontin domain (Fig. 5C).

The GTPase IMAP gene family of *P. acute* includes 36 sequences with a mix of sequences containing signal peptides, transmembrane domains, and no predicted localization domains (Fig 5D). Additionally, one sequence displays C2H2 zinc-finger domains downstream of the GTPase domain. Lastly, 58 unique transcripts of the FReD gene family in *P. acuta* were recorded, yielding a total of 82 sequences (complete and partial) when including 24 FReD sequences retrieved from a previous investigation of bacterially challenged and control *P. acuta* snails (Schultz et al., 2018; Fig 5E). *Physella acuta* FReDs contain either signal peptides, transmembrane domains or no N-terminal localization domains.

Some sequences combined a novel arrangement of concurrent N-terminal C1q-like domains and C-terminal C-type lectin domains, represented by a family of 6 sequences of the complement C1q gene family (Fig. 5B).

Discussion

Parasite-host compatibility can be determined by a multitude of factors including host physiology and immune competence in combinations with parasite capabilities to evade or interfere with host defenses. Despite such restrictions, some parasites (echinostomes) are generalists that can infect and thrive in hosts of diverse phylogenetic background (Toledo et al., 2009). In such cases it is intriguing to consider whether the parasite employs one single approach to infect and develop or, alternatively, the parasite requires several strategies to effect compatibility, tailored to different environments encountered inside different hosts.

The digenetic trematode *E. paraensei,* most often studied from interactions with the planorbid snail *B. glabrata* as intermediate snail host, is one such generalist that develops in snails of other families of freshwater gastropods, including the Lymnaeidae and as recently shown also the Physidae (Lie and Basch, 1967; Maldonado et al., 2001; Pinheiro et al., 2004; Schultz et al., 2018). Schultz et al. (2018) reported that *E. paraensei* encounters a considerably different snail host immune response whether infecting *P. acuta* or *B. glabrata.* Whereas *B. glabrata*, the main model to interpret gastropod immune function, possesses an immune gene family of over 20 fibrinogen-related protein (FREP)-encoding genes to express a diverse array of FREP proteins that co-determine anti-parasite resistance (Adema et al. 2017), *P. acuta* has only few FREP genes and does not express these in response to *E. paraensei.* Thus, characterization of the host response in one particular host-parasite interaction may not define determinants of resistance or compatibility for the interaction of that parasite with another host species. This

provides intriguing questions regarding host suitability and the immune mechanisms that dictate compatibility with cognate pathogens.

To better understand the adaptability of *E. paraensei*, this study characterized the host response of *P. acuta* during initial stages (2 and 8 DPE) of infection by this generalist parasite. Trematode parasites usually attract snail immune reactions immediately following penetration, however, compatibility with the snail-first intermediate host is generally decided at 4DPE (Alba et al., 2018; Loker and Adema, 1995; Nowak et al., 2004). At this time, either the parasite suffers immuno-elimination in a resistant snail, or the parasite escapes/renders ineffective the host defenses and proceeds to alter snail host (immuno-)physiology to support a permanent, viable infection (Hanington et al., 2010b). Accordingly, the time points selected include an initial hostile environment defined by anti-parasite host defenses (2DPE) that transitions to a parasite-modulated intramolluscan environment that is supportive of *E. paraensei* parasite development (8DPE).

The compatible interaction of *E. paraensei* and *P. acuta* was confirmed via PCR amplification of parasite genomic sequences recovered among nucleic acid samples extracted from whole-body snail tissue. Additionally, transcripts expressed by *E. paraenesei* were recovered from the Illumina RNA-seq data of *P. acuta* 2 and 8 DPE to the parasite. In conjuntion with a previous analysis (Schultz et al., 2018), which demonstrated this compatible host-parasite interaction, NGS methods provided molecular insights of host responses, including host immunity. Illumina technology enabled comprehensive capture and investigation of the initial snail defense response to *E. paraensei* that transitions to a transcriptomic profile associated with an intramollusan host environment that helps the parasite infection come to fruition.

Analysis of the transcriptomic data collected was supported by generating a genome assembly of *P. acuta* (isolate A) to enable transcript mapping. This initial genome assembly (estimated 15-fold coverage) is highly fragmented (> 800,000 scaffolds) and will benefit from future improvements, however, it is available publicly (GenBank accession RDRX00000000) as a resource to interpret the biology of *P. acuta*, also as vector of digenetic trematode parasites. The Illumina RNAseq datasets collected here from *P. acuta,* unexposed snails at 2 and 8DPE to *E. paraensei* (GenBank SRA/TSA accession SRR8080746/GHAL00000000) are comprehensive also, as indicated by 99% representation of the BUSCO set of 303 conserved eukaryotic genes of metazoa (Nishimura et al., 2017, Fig. 1B), appropriate for comparison of differentially expressed (DE) transcripts toward characterization of the early interaction of *P. acuta* and *E. paraensei*.

Clearly, *E. paraensei* infection alters the transcription profiles of *P. acuta* relative to unexposed control snails with DE transcript numbers detected via EBSeq and somewhat more abundant at 2DPE compared to 8DPE (Fig. 2A). Whereas the overall transcript numbers do not display great differences, the gene expression profiles differ considerably at 2 and 8 DPE in that they share no more than 18% of transcripts with increased expression and 22% of transcripts with lowered expression (Fig. 2B). More dramatically, many transcripts with high expression levels at 2DPE are expressed at low levels at 8DPE and vice versa. Overall, it is apparent that the molecular underpinnings of *P. acuta* biology change significantly during the transition from the initial response to parasite invasion to the establishment of a compatible *E. paraensei* infection.

Gene Ontology (GO) analysis was applied to begin interpreting the functional implications of the overall differential gene expression patterns in *P. acuta* following parasite exposure. Although routinely applied in transcriptomic studies, this method is not optimal because biological function is assigned to sequences based on blast-similarity, which is not

frequently detected for transcripts from non-model organisms. Although only 1.2% of DE sequences from *P. acuta* were assigned a GO annotation (Fig. 3A), this approach identified some functional differences within the DE transcript profiles (Fig. 3C). For instance, low "signal transducer activity" may indicate parasite interference with signaling pathways of *P. acuta* in order to establish and develop. Also, increased "cell-proliferation" at 8DPE suggests tissue repair and perhaps even defense activities against other pathogens after the parasite has established in the host.

To further develop insights concerning how *P. acuta* manages exposure to *E. paraensei* and how the parasite may influence the biology of its host, attention was directed onto DE transcripts with immune function. Use of previously applied criteria (Hanelt et al., 2008; Pinaud et al., 2016; Tetreau et al., 2017) led to identification of 460 immune-relevant DE sequences (Fig. 4). These belong to a wide range of functionally diverse sequences including antimicrobial factors like antimicrobial peptides (AMPs), and L-amino acid oxidases (LAAOs); receptors for pathogen associated molecular patterns (PAMPs), like peptidoglycan recognition proteins (PGRP) and toll-like receptors (TLR) and sequences involved in detoxification of free radicals like glutathione S-transferases.

Analyses of the immune components of the transcriptome profiles of *P. acuta* following exposure to *E. paraensei* identified five categories of immune-genes that provide new insight into immune function in *P. acuta* (Fig. 5), based on a role in innate immunity in other molluscs, gene number and diversity, and in one case, a unique combination of protein domains. The latter applies to a family of C1q domain-containing sequences (Fig. 5B), characteristic for highly abundant complement-like immune factors in the genome of the bivalve *Crassostrea gigas* (Zhang et al., 2015a) and also described for other molluscs (Gorbushin, 2018; Zhang et al.,

2016a). Most remarkably, these sequences from *P. acuta* show a unique organization of C1q domains upstream of C-type lectin domains, also immune-relevant (Brown et al., 2018). To our knowledge this combination is unique among Metazoa, and may represent taxon-specific immune factors of *P. acuta*.

Physella acuta also yielded a family of 25 DE transcripts with similarity to *Limulus* agglutination factor-like (LAF) sequences, containing a dermatopontin domain (Fig. 5C). LAF proteins of the horseshoe crab *Limulus polyphemus* (Fujii et al., 1992) facilitate agglutination of immune cells called amebocytes. Differential expression at 2 and 8DPE suggests a potential involvement of the *P. acuta* LAF-like sequences in immunity, especially considering that Bouchut et al. (2006) demonstrated increased expression of two hemocyte-derived dermatopontin-like genes from resistant *B. glabrata* snails following response to *Echinostoma caproni* and dermatopontin-like plasma proteins of *B. glabrata* that bind to proteins of *S. mansoni* parasites (Wu et al., 2017). The presence of signal peptides and transmembrane domains in the LAF-like sequences of *P. acuta* suggests that these proteins may serve as putative cell receptors within *P. acuta*.

In total, 36 differentially expressed immune-relevant sequences expressed by *P. acuta* after exposure to *E. paraensei* belonged to the GTPase IMAP family (Fig. 5D). Predicted signal peptides and transmembrane domains suggest that *P. acuta* GTPase IMAPs function as cell surface proteins. Much remains to be learned about this gene family in invertebrate organisms but GTPase IMAPs have been implicated in an immune context for *Biomphalaria* snails in response to LPS stimulation (Zhang et al., 2016b) and after exposure to *S. mansoni* (Buddenborg et al., 2017), and a large family of GTPase IMAP sequences was differentially expressed by the eastern oyster *Crassostrea virginica* after bacterial exposure

(McDowell et al., 2016). It is plausible that these sequences are surveillance factors after pathogen exposure that regulate downstream defense mechanisms. How this is achieved is unknown for *P. acuta* and other molluscs, warranting further investigation.

The largest group of DE sequences (2 and 8DPE) consisted of a family of 58 fibrinogen domain-containing sequences (FreDs; Fig. 4, 5E), that are both up and downregulated at 2 and 8DPE (Table 1). This group included a variety of secreted sequences (containing signal peptides), membrane-bound factors, and possible cytosolic factors. Only two of these FreDs (designated as FReD 16 and 20) were previously recorded by Schultz et al. (2018) among 24 full length FReD sequences from *P. acuta* (unexposed and bacterially challenged). It is potentially more likely that these two FreDs are expressed in response to bacteria, however the function of the remaining 56 FReDs is less clear. Notably, two of the DE FReD sequences were identified as FREPs due to the presence of upstream IgSF domains, increasing the number of FREP genes of *P. acuta* to four (Schultz et al., 2018). Other than in *B. glabrata*, where FREPs contribute prominently to anti-parasite responses, this study further confirms the conclusion (Schultz et al., 2018) that *P. acuta* does not employ increased FREP expression to respond to *E. paraensei*. In fact, only two FREP sequences were detected and these were downregulated at 2DPE and expressed at baseline level (same as control) at 8DPE (Fig. 4). The considerable diversity of 82 FreD sequences (including 2 FREPs) in *P. acuta* (Fig. 5E) adds to an increasing number of studies from diverse arthropods and molluscs which indicate an immune role for expanded families of FReD genes in invertebrate immunity (Christophides et al., 2002; Dimopoulos et al., 2002; Gerdol et al., 2017; Romero et al., 2011; Wu et al., 2017; Zdobnov et al., 2002; Zhang et al., 2015a).

While members of all categories of immune gene families were always represented (Fig. 4), variation was evident in the level of expression of these immune-gene families relative to control snails. A subset of various immune sequences that belong to multiple different gene families was expressed at baseline (control) level at 2DPE and subsequently differentially expressed at either higher or lower levels at 8DPE (Fig. 4). There are intriguing categories of immune transcripts present in the DE profiles at 2 and 8DPE to *E. paraensei*. For instance, there are 4 categories of immune-relevant domains that were detected only among upregulated sequences (Table 1). These include alpha macroglobulin, alpha-2 macroglobulin, serpin, and SRCR (scavenger receptor cysteine rich). Alpha and alpha-2 macroglobulin (A2M) domains constitute a structural component of thioester-containing proteins (TEPs) and CD109 proteins that associate with innate immune defenses of molluscs and arthropods (Falade and Otarigho, 2018; Portet et al., 2018). Expression of BgTEP in hemocytes (defense cells) of *B. glabrata* is upregulated after exposure to the trematode parasite *S. mansoni* (Portet et al., 2018), and secreted BgTEP is proposed to be a component of the FREP-*S. mansoni* immune complexes that mediates immune activation after binding parasite antigens (Moné et al., 2010). Serpin-like transcripts may function in regulation of development, DNA binding, but also in immunomodulation (Law et al., 2006). Through binding a variety of ligands, including microbial surface determinants, SRCR-containing sequences may be involved in development and immunity (Buckley and Rast, 2015).

Increased differential expression in *P. acuta* of genes from the above immune-gene families after exposure to *E. paraensei* (Fig. 4) at 2 and 8DPE, even as the parasite thrives within *P. acuta*, suggests that some of these factors, expressed at 2DPE respond against *E. paraensei*, but that others are important for snail homeostasis during times of injury and repair. Indeed,

molluscs possess expanded immune-gene families consisting of members that differ in immune efficacy toward particular pathogens; through selective expression profiles gastropods can tailor immune response to be optimally effective to particular pathogens or parasites (Adema et al., 2010; Deleury et al., 2012; Zhang et al., 2015a). The most upregulated sequence (916-fold) in *P. acuta* harboring *E. paraensei* at 8DPE is a G-type lysozyme-6-like protein that likely functions in antibacterial defense by cleaving glycosidic bonds in peptidoglycan layers of bacterial cell walls (Guo and He, 2014). It is unlikely that such a type of immune factor is upregulated to combat trematode parasites, and even more unlikely after the initial response to *E. paraensei* has transitioned to that of a snail that harbors a successfully developing parasite infection. The transcriptome of *P. acuta* at 8DPE to *E. paraensei*, however, does include a multitude of defense genes and these likely serve not to counter the thriving *E. paraensei* but to protect against other (opportunistic) parasites and pathogens.

The findings from this study are interpreted as follows. The molecular mechanisms that underlie the compatibility of *E. paraensei* with *P. acuta* are highly complex. At the onset of this interaction, both parasite and host mount opposing determinants of compatibility. *Physella acuta* expresses a broad set of select members of several immune-gene families, selected for maximum efficiency to counter invading trematode parasites. Simultaneously, *E. paraensei* applies as yet unknown mechanisms to evade or interfere with the host defense response. The outcome of this interaction leads to polar extremes: either *P. acuta* proves resistant and eliminates the parasite or *E. paraensei* establishes a successful infection and modifies the snail host to facilitate parasite development and reproduction. Intriguingly, after transitioning from defense to a situation where the parasite has achieved compatibility, the *P. acuta* host continues to elaborate immune activities, although a different subset of each immune gene family is expressed. It is

hypothesized that this situation reflects the extended phenotype of the parasite imposed on the snail host. Briefly, the concept of the extended phenotype poses that traits of one organism (parasite) many manifest themselves apart the physical body of the organism itself (Dawkins, 1982). Assuming that individual sequences of a large immune-gene family from *P. acuta* are pathogen-specific, at 8DPE *E. paraensei* controls the snail host to selectively express immune genes that do not impact the parasite, but rather provide immune protection against other pathogens, effectively becoming an extended phenotype of the parasite. NGS approaches further revealed the immune armament of *P. acuta* to respond to digenean parasite infection and helped postulate that *E. paraensei* exhibits an extended phenotype in the snail intermediate host for immunoprotection. This finding also put into focus the question of how compatible parasites are able to effect such changes in the snail host for future research.

Acknowledgements

The authors wish to acknowledge the molecular biology facility (MBF) at the University of New Mexico for sequencing and technical support and Lijun Lu and Sam Loker for NGS sequencing (NIH R37AI101438).

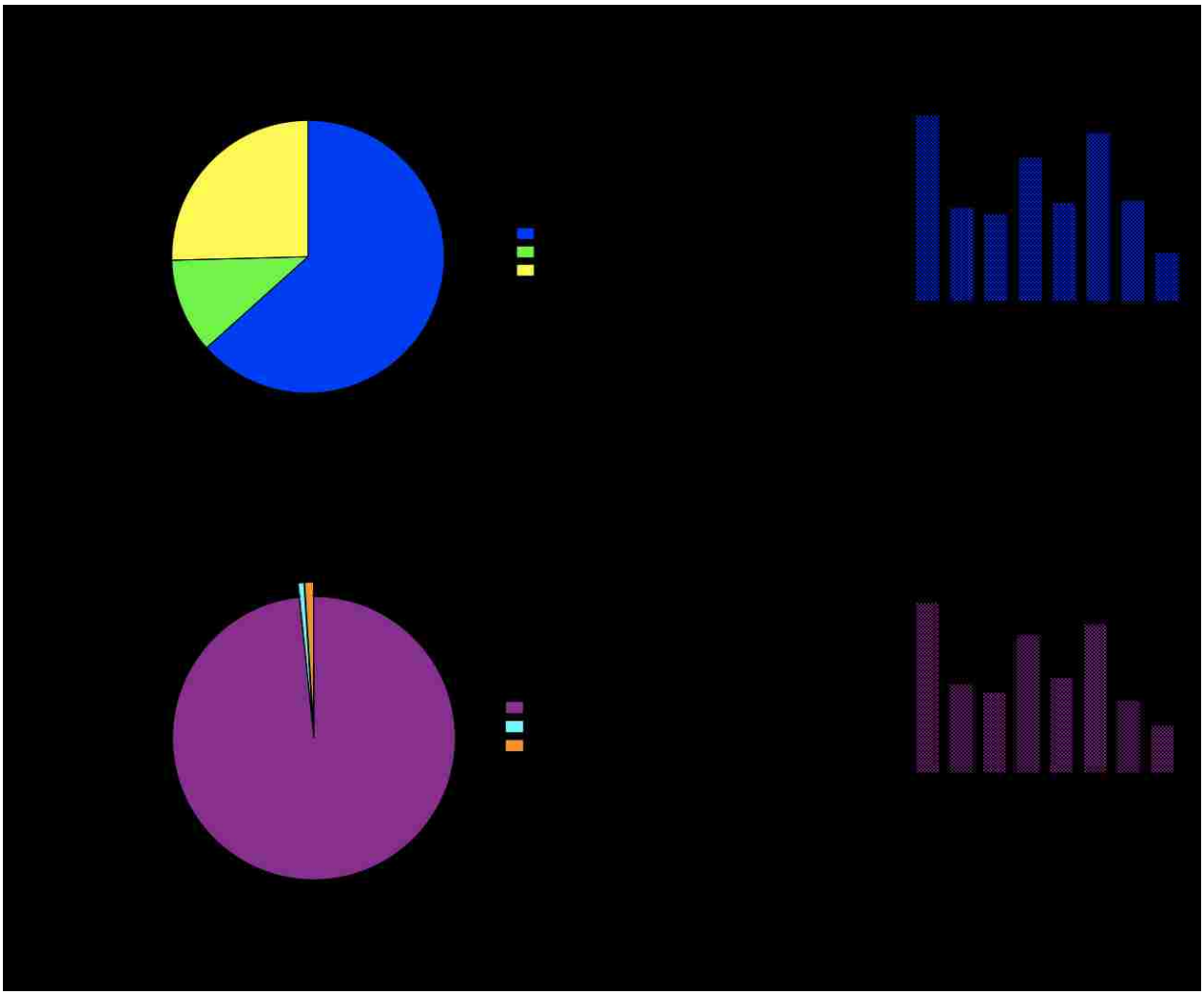


Fig. 1. gVolante analyses of *P. acuta* genome and transcriptome.
(a) Assessment of genome completeness and general annotation features of the *P. acuta* genome portraying the percentages of 303 BUSCO eukaryotic core genes recorded. Annotation features are labeled on the x-axis and raw counts using a log2 scale are presented on the y-axis. Note: GC-content is represented as a percentage. (b) Assessment of transcriptome completeness portraying the percentages of 303 BUSCO eukaryotic core genes recorded from the *P. acuta* transcriptome. Annotation features are the same as in (a). Partial: sequences identified as incomplete eukaryotic core genes (i.e. not full-length). Not detected: sequences of the gVolante core set that are not identified.

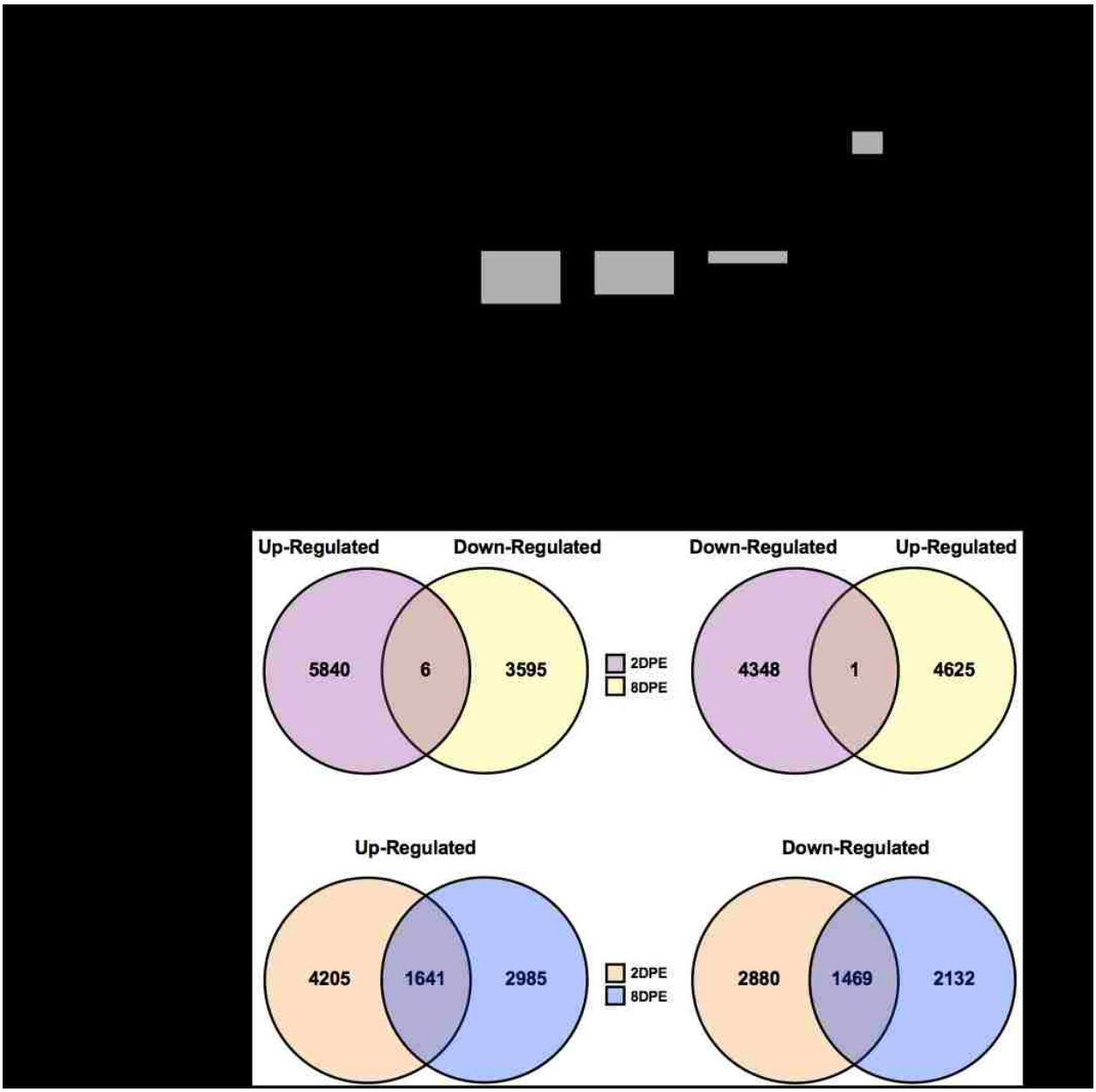


Fig. 2. Analysis of *P. acuta* DE sequences 2 and 8DPE to *E. paraensei* relative to controls. (a) Transcriptome profiles of differentially expressed sequences after exposure to *E. paraensei*. Comparison of control and experimental groups are labeled on the x-axis. Total transcript numbers are presented on the y-axis. (b) Comparative analysis of the number of sequences shared between DE groups at 2 and 8DPE. DE groups are defined by sequences upregulated or downregulated at various time points.

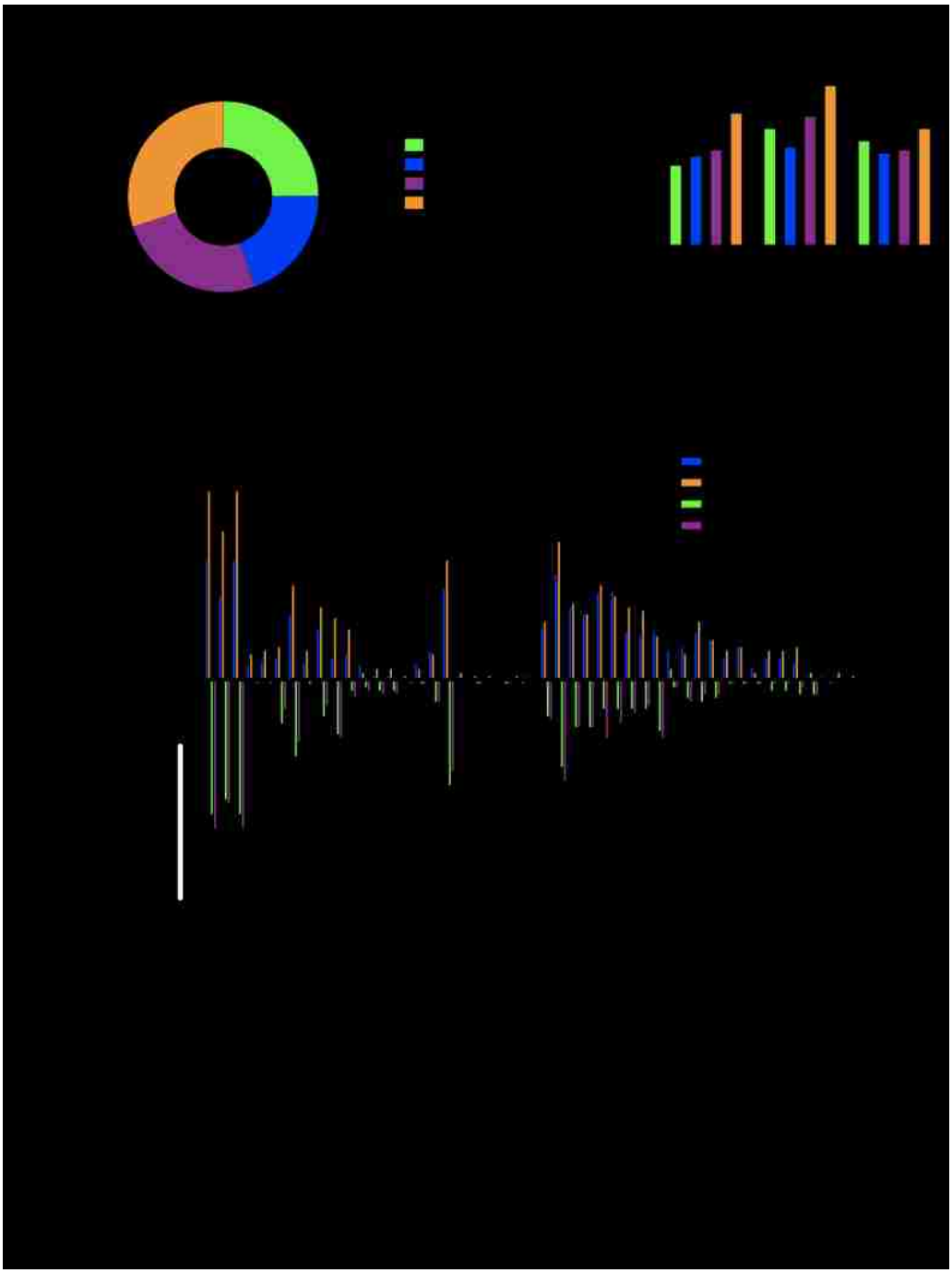


Fig. 3. Gene ontology analysis of *P. acuta* DE sequences after parasite exposure.
(a) Representation (DE group) of the number of DE sequences assigned GO annotation. (b) Number of level-2 GO terms assigned to each DE group within the hierarchical GO categories *biological process*, *cellular component*, and *molecular function*. (c) Graphical display of level-2 GO terms assigned to DE sequences. Level-2 GO terms are labelled along the x-axis and gene number is displayed on the y-axis.

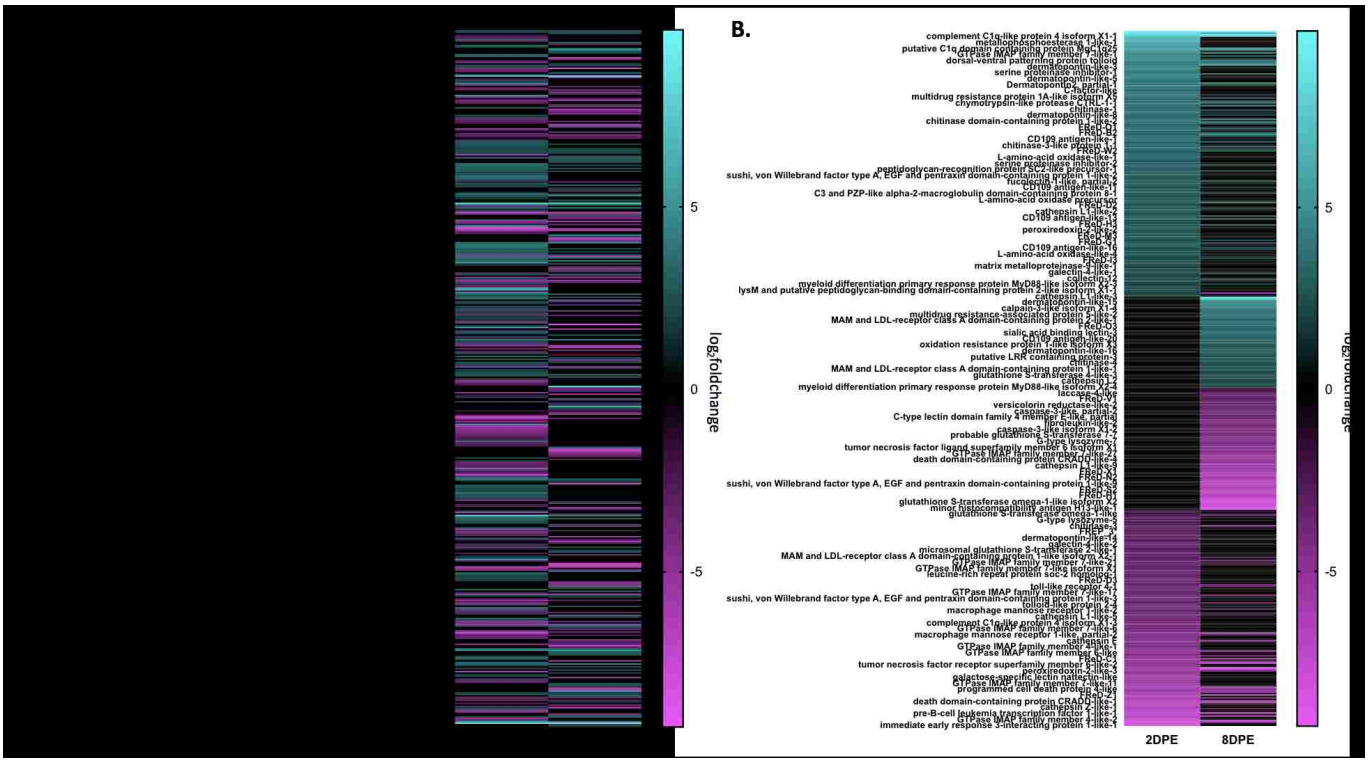


Fig. 4. Heatmaps of immune-relevant DE sequences at 2 and 8DPE.
(a) Heatmap of immune-relevant DE sequences sorted by assigned annotation name along the y-axis. (b) Heatmap of immune-relevant DE sequences consecutively sorted first by the $\log_2$ fold change of sequences DE at 2DPE then by the $\log_2$ fold change of sequences DE at 8DPE. Note: not all genes are listed along the y-axis.

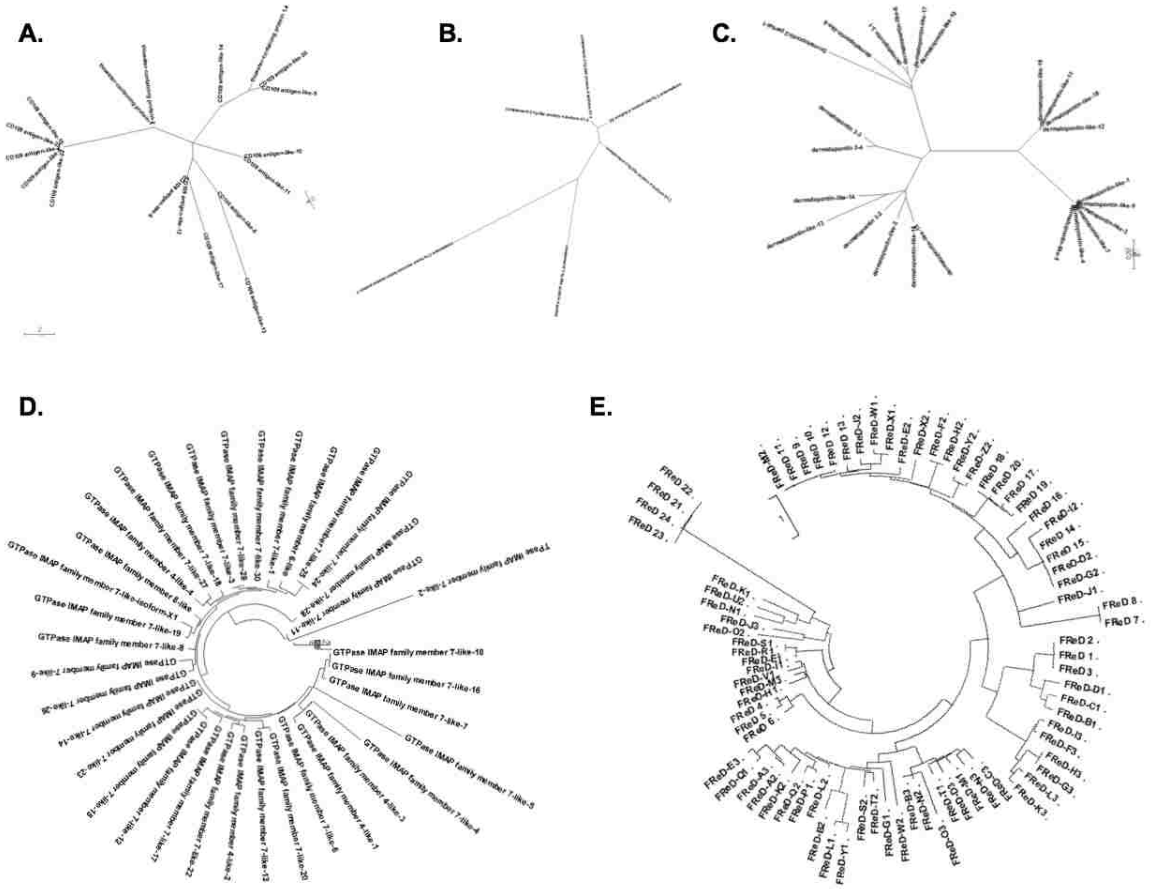


Fig. 5. Unrooted gene trees of DE immune-relevant sequences after parasite exposure belonging to gene families of *P. acuta*.

(a) Eighteen DE sequences that represent the CD109-TEP gene family of *P. acuta*. (b) Six DE sequences that represent the complement C1q gene-family of *P. acuta*. These sequences were selected because of the unique combination of C1q and C-LECT domains within each transcript. (c) Twenty-four DE sequences that represent the dermatopontin gene-family of *P. acuta.* Note: functional characterization a protein containing dermatopontin domains revealed hemagglutination activity (see *Limulus* agglutination factor (LAF; Fujii et al., 1992). (d) Thirty-six DE sequences that represent the GTPase IMAP gene-family of *P. acuta*. (e) Eighty-two sequences that represent the FreD gene-family of *P. acuta*. Fifty-eight of these sequences were DE after parasite exposure (contain letters in their gene name) and 24 were previously recorded from a 454 transcriptome generated from 2 individual *P. acuta* snails (see Schultz et al., 2018).

Table 1. Number of immune-relevant domains, predicted by InterProScan, from DE sequences of *P. acuta* 2 and 8DPE to *E. paraensei*.

Immune-relevant Domains	2DPE_UP	8DPE_UP	2DPE_DOWN	8DPE_DOWN
Aerolysin	26	15	6	2
Alpha-2-macroglobulin	32	21	0	0
Alpha-macroglobulin	27	7	0	0
C-type Lectin	196	161	275	189
C1q	72	67	93	16
C2	152	130	6	4
Concanavalin A-like lectin/glucanase domain superfamily	82	42	19	12
DEAD/DEAH box helicase domain	8	6	1	3
Death	28	9	32	40
Fibrinogen	218	69	102	102
Fucolectin tachylectin-4 pentraxin-1	2	6	1	2
Galectin, carbohydrate recognition domain	122	37	7	13
Glutathione S-transferase	0	30	32	51
H-type lectin	3	14	38	17
immunoglobulin	1064	659	175	155
Kazal	232	19	13	60
Leucine-rich repeat	178	160	124	111
Macrophage migration inhibitory factor	0	0	0	4
Peroxiredoxin	3	0	1	0
Programmed cell death protein	0	0	0	2
Serine protease	134	69	34	24
Serpin	39	27	0	0
SRCR	114	15	0	0
Superoxide dismutase	11	9	8	10
Sushi/SCR/CCP	322	128	261	111
Thioredoxin	21	33	16	36
Toll/interleukin-1 receptor homology (TIR)	11	11	23	18
Tumor necrosis factor	51	64	196	85

References

1. Adema, C. M., Arguello, D. F., Stricker, S. A., E. S. Loker. 1994. A time-lapse study of interactions between *Echinostoma paraensei* intramolluscan larval stages and adherent hemocytes from *Biomphalaria glabrata* and *Helix aspersa*. Journal of Parasitology 80: 719-727.

2. Adema, C. M., Hertel, L. A., Miller, R. D., E.S. Loker. 1997. A family of fibrinogen-related proteins that precipitates parasite-derived molecules is produced by an invertebrate after infection. Proceedings of the National Academy of Sciences. U. S. A. 94: 8691-8696.

3. Adema, C.M., Hertel, L.A., E. S. Loker. 1999. Evidence from two planorbid snails of a complex and dedicated response to digenean (echinostome) infection. Parasitology 119: 395-404.

4. Adema, C. M., Lun, C. M., Hanelt, B., and R. S. Seville. 2009. Digenean trematode infections of native freshwater snails and invasive *Potamopyrgus antipodarum* in the Grand Teton National Park/John D. Rockefeller Memorial Parkway Area. Journal of Parasitology 95: 224-227.

5. Adema, C. M., Hanington, P. C., Lun, C. M., Rosenberg, G. H., Aragon, A. D., Stout, B. A., Lennard Richard, M. L., Gross, P. S., E. S. Loker. 2010. Differential transcriptomic responses of *Biomphalaria glabrata* (Gastropoda, Mollusca) to bacteria and metazoan parasites, *Schistosoma mansoni* and *Echinostoma paraensei* (Digenea, Platyhelminthes). Molecular Immunology 47: 849-860.

6. Adema, C. M., Bayne, C. J., Bridger, J. M., Knight, M., Loker, E. S., Yoshino, T. P., S.M. Zhang. 2012. Will all scientists working on snails and the diseases they transmit please stand up? PLoS Neglected Tropical Diseases 6: e1835.

7. Adema, C. M. 2015. Fibrinogen-related proteins (FREPs) in mollusks. Pathogen-Host Interactions: Antigenic Variation v. Somatic Adaptations. E. Hsu and L. Du Pasquier (eds.). Springer International Publishing, Switzerland, p. 111–129.

8. Adema, C. M., E. S. Loker. 2015. Digenean-gastropod host associations inform on aspects of specific immunity in snails. Developmental and Comparative Immunology 48: 275-283.

9. Adema, C. M., Hillier, L. W., Jones, C. S., Loker, E. S., Knight, M., Minx, P., Oliveira, G., Raghavan, N., Shedlock, A., do Amaral, L. R., Arican-Goktas, H. D., Assis, J. G., Baba, E. H., Baron, O. L., Bayne, C. J., Bickham-Wright, U., Biggar, K. K., Blouin, M., Bonning, B. C., Botka, C., Bridger, J. M., Buckley, K. M., Buddenborg, S. K., Lima, Caldeira, R., Carleton, J., Carvalho, O. S., Castillo, M. G., Chalmers, I. W., Christensens, M., Clifton, S., Cosseau, C., Coustau, C., Cripps, R. M., Cuesta-Astroz, Y., Cummins, S.F., di Stephano, L., Dinguirard, N., Duval, D., Emrich, S., Feschotte, C., Feyereisen, R.,

FitzGerald, P., Fronick, C., Fulton, L., Galinier, R., Gava, S.G., Geusz, M., Geyer, K. K., Geraldo-Calderón, G. I., de Souza Gomes, M., Gordy, M. A., Gourbal, B., Grunau, C., Hanington, P. C., Hoffmann, K. F., Hughes, D., Humphries, J., Jackson, D. J., Jannotti-Passos, L. K., de, Jesus, Jeremias, W., Jobling, S., Kamel, B., Kapusta, A., Kaur, S., Koene, J. M., Kohn, A. B., Lawson, D., Lawton, S. P., Lìang, D., Limpanont, Y., Liu, S., Lockyer, A. E., Lovasto, T. L., Ludolf, F., Magrini, V., McManus, D. P., Medina, M., Misra, M., Mitta, G., Mkoji, G. M., Montague, M. J., Montelongo, C., Moroz, L. L., Munoz-Torres, M. C., Niazi, U., Noble, L. R., Oliveira, F. S., Pais, F. S., Papenfuss, A. T., Peace, R., Pena, J. J., Pila, E. A., Quelais, T., Raney, B. J., Rast, J. P., Rollinson, D., Rosse, I. C., Rotgans, B., Routledge, E. J., Ryan, K. M., Scholte, L. L. S., Storey, K. B., Swain, M., Tennessen, J. A., Tomlinson, C., Trujillo, D. L., Volpi, E. V., Walker, A. J., Wang, T., Wannaporn, I., Warren, W. C., Wu, X. J., Yoshino, T. P., Yusuf, M., Zhang, S. M., Zhao, M., R. K. Wilson. 2017. Whole genome analysis of a schistosomiasis-transmitting freshwater snail. Nature Communications 8: 15451.

10. Alba, A., Vázquez, A. A., Sánchez, J., Duval, D., Hernández, H. M., Sabourin, E., Vittecoq, M., Hurtrez-Boussés, S., B. Gourbal. *Fasciola hepatica-Pseudosuccinea columella* interaction: effect of increasing parasite doses, successive exposures and geographical origin on the infection outcome of susceptible and naturally-resistant snails from Cuba. Parasites and Vectors 11: 559.

11. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., D. J. Lipman .1990. Basic local alignment search tool. Journal of Molecular Biology 215: 403-410.

12. Ataev, G. L., Fournier, A., C. Coustau. 1998. Comparison of *Echinostoma caproni* mother sporocyst development *in vivo* and *in vitro* using *Biomphalaria glabrata* snails and a *B. glabrata* embryonic cell line. Journal of Parasitology 84: 227-235.

13. Ataev, G. L., C. Coustau. 1999. Cellular response to *Echinostoma caproni* infection in *Biomphalaria glabrata* strains selected for susceptibility/resistance. Developmental and Comparative Immunology 23: 187-198.

14. Barragán-Sáenz, F. A., Sánchez-Nava, P., Hernández-Gallegos, O., and G. Salgado-Maldonado. 2009. Larval stages of trematodes in gastropods from Lake Chicnahuapan, State of Mexico, Mexico. Parasitology Research 105: 1163-1167.

15. Boetzer, M., Henkel, C. V., Jansen, H. J., Butler, D., W. Pirovano. 2011. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics 27: 578-579.

16. Bolger, A. M., Lohse, M., B. Usadel. 2014. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics 30: 2112-2120.

17. Bouchut, A., Roger, E., Coustau, C., Gourbal, B., G. Mitta. 2006. Compatibility in the *Biomphalaria glabrata/Echinostoma caproni* model: potential involvement of adhesion genes. International Journal of Parasitology 36: 175-184.

18. Brown, G. D., Willment, J. A., L. Whitehead. 2018. C-type lectins in immunity and homeostasis. Nature Reviews Immunology 18: 374-389.

19. Buckley, K. M. and J. P. Rast. 2015. Diversity of animal immune receptors and the origins of recognition complexity in the deuterostomes. Developmental Comparative Immunology 49:179-189.

20. Buchfink, B., Xie, C., D. H. Huson. 2015. Fast and sensitive protein alignment using DIAMOND. Nature Methods 12: 59-60.

21. Buddenborg, S. K., Bu, L., Zhang, S. M., Schilkey, F. D., Mkoji, G. M., E. S. Loker. 2017. Transcriptomic responses of *Biomphalaria pfeifferi* to *Schistosoma mansoni*: Investigation of a neglected African snail that supports more *S. mansoni* transmission than any other snail species. PLoS Neglected Tropical Diseases 11: e0005984.

22. Christophides, G. K., Zdobnov, E., Barillas-Mury, C., Birney, E., Blandin, S., Blass, C., Brey, P. T., Collins, F. H., Danielli, A., Dimopoulos, G., Hetru, C., Hoa, N. T., Hoffmann, J. A., Kanzok, S. M., Letunic, I., Levashina, E. A., Loukeris, T. G., Lycett, G., Meister, S., Michel, K., Moita, L. F., Müller, H. M., Osta, M. A., Paskewitz, S. M., Reichart, J. M., Rzhetsky, A., Troxler, L., Vernick, K. D., Vlachou, D., Volz, J., von Mering, C., Xu, J., Zheng, L., Bork, P., F. C. Kafatos. 2002. Immunity-related genes and gene families in *Anopheles gambiae*. Science 298, 159–165.

23. Couch, L., Hertel, L. A., E. S. Loker. 1990. Humoral response of the snail *Biomphalaria glabrata* to trematode infection: observations on a circulating hemagglutinin. Journal of Experimental Zoology 255: 340-349.

24. Dawkins, R. 1982. The Extended Phenotype. Oxford University Press, Oxford, United Kingdom, 30.

25. DeGaffé, G., E. S. Loker. 1998. Susceptibility of *Biomphalaria glabrata* to infection with *Echinostoma paraensei*: correlation with the effect of parasite secretory-excretory products on host hemocyte spreading. Journal of Invertebrate Pathology 71: 64-72.

26. Deleury, E., Dubreuil, G., Elangovan, N., Wajnberg, E., Reichart, J. M., Gourbal, B., Duval, D., Baron, O. L., Gouzy, J., C. Coustau. 2012. Specific versus non-specific immune responses in an invertebrate species evidenced by a comparative *de novo* sequencing study. PLoS One 7: e32512.

27. Dimopoulos, G., Christophides, G. K., Meister, S., Schultz, J., White, K. P., Barillas-Mury, C., F. C. Kafatos. 2002. Genome expression analysis of *Anopheles gambiae*: responses to injury, bacterial challenge, and malaria infection. Proceedings of the National Academy of Sciences. U. S. A. 99: 8814-8819.

28. Ebbs, E.T., Loker, E.S., S. V. Brant. 2018. Phylogeography and genetics of the globally invasive snail *Physa acuta* Draparnaud 1805, and its potential to serve as an intermediate host to larval digenetic trematodes. BMC Evolutionary Biology 18: 103.

29. Falade, M. O. and B. Otarigho. 2018. Comparative Functional Study of Thioester-containing Related Proteins in the Recently Sequenced Genome of *Biomphalaria glabrata*. Iran Journal of Parasitology 13: 79-88.

30. Finn, R. D., Attwood, T. K., Babbitt, P. C., Bateman, A., Bork, P., Bridge, A. J., Chang, H. Y., Dosztányi, Z., El-Gebali, S., Fraser, M., Gough, J., Haft, D., Holliday, G. L., Huang, H., Huang, X., Letunic, I., Lopez, R., Lu, S., Marchler-Bauer, A., Mi, H., Mistry, J., Natale, D. A., Necci, M., Nuka, G., Orengo, C. A., Park, Y., Pesseat, S., Piovesan, D., Potter, S. C., Rawlings, N. D., Redaschi, N., Richardson, L., Rivoire, C., Sangrador-Vegas, A., Sigrist, C., Sillitoe, I., Smithers, B., Squizzato, S., Sutton, G., Thanki, N., Thomas, P. D., Tosatto, S. C., Wu, C. H., Xenarios, I., Yeh, L. S., Young, S. Y., Mitchell, A. L., 2017. InterPro in 2017-beond protein family and domain annotations. Nucleic Acids Research 45: D190-D199.

31. Fujii, N., Minetti, C. A., Nakhasi, H. L., Chen, S. W., Barbehenn, E., Nunes, P. H., N. Y. Nguyen. 1992. Isolation, cDNA cloning, and characterization of an 18-kDa hemagglutinin and amebocyte aggregation factor from *Limulus polyphemus*. Journal of Biological Chemistry 267: 22452-22459.

32. Gerdol, M., Luo, Y.J., Satoh, N., A. Pallavicini. 2017. Genetic and molecular basis of the immune system in the brachiopod *Lingula anatina*. Developmental and Comparative Immunology 82: 7-30.

33. Gorbushin, A.M. 2018. Immune repertoire in the transcriptome of *Littorina littorea* reveals new trends in lophotrochozoan proto-complement evolution. Developmental and Comparative Immunology 84: 250-263.

34. 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., Friendman, N., A. Regev. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnology 29: 644-652.

35. Grimm-Jørgensen, Y. 1983. Immunoreactive somatostatin in two pulmonate gastropods. General and Comparative Endocrinology 49: 108-114.

36. Grimm-Jørgensen, Y. 1987. Somatostatin and calcitonin modulate gastropod internal defense mechanisms. Developmental and Comparative Immunology 11: 487-499.

37. Guo, Y., H. He. 2014. Identification and characterization of a goose-type lysozyme from sewage snail *Physa acuta*. Fish and Shellfish Immunology 39: 321-325.

38. Hanelt, B., Lun, C. M., C. M. Adema. 2008. Comparative ORESTES-sampling of transcriptomes of immune-challenged *Biomphalaria glabrata* snails. Journal of Invertebrate Pathology 99: 192-203.

39. Hanington, P. C., Forys, M. A., Dragoo, J. W., Zhang, S. M., Adema, C. M., E. S. Loker. 2010a. Role for a somatically diversified lectin in resistance of an invertebrate to parasite infection. Proceedings of the National Academy of Sciences. U. S. A. 107: 21087-21092.

40. Hanington, P. C., Lun, C. M., Adema, C. M., E. S. Loker. 2010b. Time series analysis of the transcriptional responses of *Biomphalaria glabrata* throughout the course of intramolluscan development of *Schistosoma mansoni* and *Echinostoma paraensei*. International Journal of Parasitology 40: 819-831.

41. Hanington, P. C., Forys, M. A., E.S. Loker. 2012. A somatically diversified defense factor, FREP3, is a determinant of snail resistance to schistosome infection. PLoS Neglected Tropical Diseases 6: e1591.

42. Hertel, L. A., Stricker, S. A., Monroy, F. P., Wilson, W. D., E. S. Loker. 1994. *Biomphalaria glabrata* hemolymph lectins: binding to bacteria, mammalian erythrocytes, and to sporocysts and rediae of *Echinostoma paraensei*. Journal of Invertebrate Pathology 64: 52-61.

43. Humbert, E., C. Coustau. 2001. Refractoriness of host haemocytes to parasite immunosuppressive factors as a putative resistance mechanism in the *Biomphalaria glabrata-Echinostoma caproni* system. Parasitology 122: 651-660.

44. Jeong, K. H., Lie, K. J., D. Heyneman. 1984. An ultrastructural study on ventricular encapsulation reactions in *Biomphalaria glabrata* exposed to irradiated echinostome parasites. International Journal of Parasitology 14: 127-133.

45. Joky, A., Matricon-Gondran, M., J. Benex. 1985. Response to the amoebocyte-producing organ of sensitized *Biomphalaria glabrata* after exposure to *Echinostoma caproni* miracidia. Journal of Invertebrate Pathology 45: 28-33.

46. Kraus, T. J., Brant S. V., C. M. Adema. 2014. Characterization of Trematode Cercariae from *Physella acuta* in the Middle Rio Grande. Comparative Parasitology 81: 105-109.

47. Langmead, B. and S. L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods 9: 357-359.

48. Larken, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T. J., D. G. Higgins. 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947-2948.

49. Law, R. H., Zhang, Q., McGowan, S., Buckle, A. M., Silverman, G. A., Wong, W., Rosado, C. J., Langendorf, C. G., Pike, R. N., Bird, P. I., J. C. Whisstock. 2006. An overview of the serpin superfamily. Genome Biology 7: 216.

50. Leng, N., Li, Y., McIntosh, B. E., Nguyen, B. K., Duffin, B., Tian, S., Thomson. J. A., Dewey, C. N., Stewart, R., C. Kendziorski. 2015. EBSeq-HMM: a Bayesian approach for identifying gene-expression changes in ordered RNA-seq experiments. Bioinformatics 31: 2614-2622.

51. Letunic, I. and P. Bork. 2017. 20 years of the SMART protein domain annotation resource. Nucleic Acids Research. http://dx.doi.org/10.1093/nar/gkx922.

52.

53. Li, D., Liu, C. M., Luo, R., Sadakane, K., T. W. Lam. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct *de Bruijn* graph. Bioinformatics 31: 1674-1676.

54. Li, W. and A. Godzik. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22: 1658-1659.

55. Li, B. C. N. Dewey. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12: 323.

56. Lie, K. J. and P. F. Basch. 1967. The life history of *Echinostoma paraensei* sp. N. (Trematoda: Echinostomatidae). Journal of Parasitology 53: 1192-1199.

57. Lie, K. J. and D. Heyneman. 1979. Acquired resistance to echinostomes in four *Biomphalaria glabrata* strains. International Journal of Parasitology 42: 533-537.

58. Lie, K. J., Jeong, K. H., D. Heyneman. 1981. Selective interference with granulocyte function induced by *Echinostoma paraensei* (Trematoda) larvae in *Biomphalaria glabrata* (Mollusca). Journal of Parasitology 67: 790-796.

59. Lie, K. J., Jeong, K. H., D. Heyneman. 1982. Further characterization of acquired resistance in *Biomphalaria glabrata*. Journal of Parasitology 68: 529-531.

60. Lie, K. J., Jeong, K. H., D. Heyneman. 1983. Acquired resistance in snails. Induction of resistance to *Schistosoma mansoni* in *Biomphalaria glabrata*. International Journal of Parasitology 13: 301-304.

61. Loker, E. S., Bayne, C. J., M. A. Yui. 1986. *Echinostoma paraensei*: hemocytes of *Biomphalaria glabrata* as targets of echinostome mediated interference with host snail resistance to *Schistosoma mansoni*. Experimental Parasitology 62: 149-154.

62. Loker, E. S., L. A. Hertel. 1987. Alterations in *Biomphalaria glabrata* plasma induced by infection with the digenetic trematode *Echinostoma paraensei*. Journal of Parasitology 73: 503-513.

63. Loker, E. S., Boston, M. E., C. J. Bayne. 1989. Differential adherence of M line *Biomphalaria glabrata* hemocytes to *Schistosoma mansoni* and *Echinostoma paraensei* larvae, and experimental manipulation of hemocyte binding. Journal of Invertebrate Pathology 54: 260-280.

64. Loker, E. S., Cimino, D. F., L. A. Hertel. 1992. Excretory-secretory products of *Echinostoma paraensei* sporocysts mediate interference with *Biomphalaria glabrata* hemocyte functions. Journal of Parasitology 78: 104-115.

65. Loker, E. S., Couch, L., L. A. Hertel. 1994. Elevated agglutination titres in plasma of *Biomphalaria glabrata* exposed to *Echinostoma paraensei*: characterization and functional relevance of a trematode-induced response. Parasitology 108: 17-26.

66. Loker, E. S. and C. M. Adema. 1995. Schistosomes, Echinostomes and Snails: Comparative Immunobiology. Trends in Parasitology 11: 120-124.

67. Loker, E. S., Adema, C. M., Zhang, S., T. B. Kepler. 2004. Invertebrate immune systems–not homogeneous, not simple, not well understood. Immunological Reviews 198: 10-24.

68. Lu, X. T., Gu, Q. Y., Limpanont, Y., Song, L. G., Wu, Z. D., Okanurak, K., Z. Y. Lv. Snail-borne parasitic diseases: an update on global epidemiological distribution, transmission interruption and control methods. Infectious Disease of Poverty 7: 28.

69. Luo et al. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. GigaScience: 1-18.

70. Maldonado, A. Jr, Vieira, G. O., Garcia, J. S., Rey, L., R. M. Lanfredi. 2001. Biological aspects of a new isolate of *Echinostoma paraensei* (Trematoda: Echinostomatidae): susceptibility of sympatric snails and the natural vertebrate host. Parasitology Research 87: 853-859.

71. McCarthy, F. M., Wang, N., Magee, G. B., Nanduri, B., Lawrence, M. L., Camon. E. B., Barrell, D. G., Hill, D. P., Dolan, M. E., Williams, W. P., Luthe, D. S., Bridges, S. M., S.C. Burgess. 2006. AgBase: a functional genomics resource for agriculture. BMC Genomics 7: 229.

72. McDowell, I. C., Modak, T. H., Lane, C. E., M. Gomez-Chiarri. 2016. Multi-species protein similarity clustering reveals novel expanded immune gene families in the eastern oyster *Crassostrea virginica*. Fish and Shellfish Immunology 53: 13-23.

73. Mitta, G., Gourbal, B., Grunau, C., Knight, M., Bridger, J. M., A. Théron. 2017. The Compatibility Between *Biomphalaria glabrata* Snails and *Schistosoma mansoni*: An Increasingly Complex Puzzle. Advances in Parasitology 97: 111-145.

74. Moné, Y., Gourbal, B., Duval, D., Du Pasquier, L., Kieffer-Jaquinod, S., G. Mitta. 2010. A large repertoire of parasite epitopes matched by a large repertoire of host immune receptors in an invertebrate host/parasite model. PLoS Neglected Tropical Diseased 4: e813.

75. Monroy, F., Hertel, L. A., E. S. Loker. 1992. Carbohydrate-binding plasma proteins from the gastropod *Biomphalaria glabrata*: strain specificity and the effects of trematode infection. Developmental and Comparative Immunology 16: 355-366.

76. Monroy, F. P., E. S. Loker. 1993. Production of heterogeneous carbohydrate-binding proteins by the host snail *Biomphalaria glabrata* following exposure to *Echinostoma paraensei* and *Schistosoma mansoni*. Journal of Parasitology 79: 416-423.

77. Morgan, J. A. and D. Blair. 1998. Relative merits of nuclear ribosomal internal transcribed spacers and mitochondrial CO1 and ND1 genes for distinguishing among *Echinostoma* species (Trematoda). Parasitology 116: 289-297.

78. Mounkassa, J. B., J. Jourdane. 1990. Dynamics of the leukocytic response of *Biomphalaria glabrata* during the larval development of *Schistosoma mansoni* and *Echinostoma liei*. Journal of Invertebrate Pathology 55: 306-311.

79. Muñoz-Antolí, C., Toledo, R., J. G. Esteban. 2000. The life cycle and transmission dynamics of the larval stages of *Hypoderaeum conoideum*. Journal of Helminthology 74: 165-172.

80. Muñoz-Antoli, C., Marin, A., Vidal, A., Toledo, R., J. G. Esteban. 2008. *Euparyphium albuferensis* and *Echinostoma friedi* (Trematoda: Echinostomatidae): experimental cercarial transmission success in sympatric snail communities. Folia Parasitologica 55: 122-126.

81. Nishimura, O., Hara, Y., S. Kuraku. 2017. gVolante for standardizing completeness assessment of genome and transcriptome assemblies. Bioinformatics. http://dx.doi.org/10.1093/bioinformatics/btx445.

82. Noda, S. and E. S. Loker. 1989a. Phagocytic activity of hemocytes of M-line *Biomphalaria glabrata* snails: effect of exposure to the trematode *Echinostoma paraensei*. Journal of Parasitology 75: 261-269.

83. Noda, S. and E. S. Loker. 1989b. Effects of infection with *Echinostoma paraensei* on the circulating haemocyte population of the host snail *Biomphalaria glabrata*. Parasitology 98: 35-41.

84. Nolan, J. R., Bergthorsson, U., C. M. Adema. 2014. *Physella acuta*: atypical mitochondrial gene order among panpulmonates (Gastropoda). Journal of Molluscan Studies 80: 1-12.

85. Nowak, T. S., Woodards, A. C., Jung, Y., Adema, C. M., E. S. Loker. Identification of transcripts generated during the response of resistant *Biomphalaria glabrata* to *Schistosoma mansoni* infection using suppression subtractive hybridization. Journal of Parasitology 90: 1034-1040.

86. Pinaud, S., Portela, J., Duval, D., Nowacki, F. C., Olive, M. A., Allienne, J. F., Galinier, R., Dheilly, N. M., Kieffer-Jaquinod, S., Mitta, G., Theron, A., B. Gourbal. 2016. A shift from cellular to humoral responses contributes to innate immune memory in the vector snail *Biomphalaria glabrata*. PLoS Pathogens 12: e1005361.

87. Pinheiro, J., Maldonado, A. J., Attias, M., R. M., Lanfredi. 2004. Morphology of the rediae of *Echinostoma paraensei* (Trematoda: Echinostomatidae) from its intermediate host *Lymnaea columella* (Mollusca, Gastropoda). Parasitology Research 93: 171-177.

88. Portet, A., Galinier, R., Pinaud, S., Portela, J., Nowacki, F., Gourbal, B., D. Duval. 2018. BgTEP: An Antiprotease Involved in Innate Immune Sensing in *Biomphalaria glabrata*. Frontiers in Immunology 9: 1206.

89. Romero, A., Dios, S., Poisa-Beiro, L., Costa, M. M., Posada, D., Figueras, A., B. Novoa. 2011. Individual sequence variability and functional activities of fibrinogen-related proteins (FREPs) in the Mediterranean mussel (*Mytilus galloprovincialis*) suggest ancient and complex immune recognition models in invertebrates. Developmental and Comparative Immunology 35: 334-344.

90. Schultz, J. H. and C. M. Adema. 2017. Comparative immunogenomics of mollusks. Developmental and Comparative Immunology 75: 3-15.

91. Schultz, J. H., Bu, L., C. M. Adema. Comparative immunological study of the snail *Physella acuta* (Hygrophila, Pulmonata) reveals shared and unique aspects of gastropod immunobiology. Molecular Immunology 101: 108-119.

92. Song, L., Shankar, D. S., L. Florea. 2016. Rascaf: Improving Genome Assembly with RNA Sequencing Data. Plant Genome 9.

93. Tamura, K., Stecher, G., Peterson, D., Filipski, A., S. Kumar. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution 30: 2725-2729.

94. Tanveer, A. 1991. Studies on the morphology of hemocytes of *Physa acuta* and *Bellamaya bengalensis*. Punjab University Journal of Zoology 6: 7-18.

95. Tetreau, G., Pinaud, S., Portet, A., Galinier, R., Gourbal, B., D. Duval. 2017. Specific Pathogen Recognition by Multiple Innate Immune Sensors in an Invertebrate. Frontiers in Immunology 8: 1249.

96. Toledo, R., Muñoz-Antolí, C., Pérez, M., J. G. Esteban. 1998. Larval trematode infections in freshwater gastropods from the Albufera Natural Park in Spain. Journal of Helminthology 72: 79-82.

97. Toledo, R., Esteban, J. G., B. Fried. 2009. Chapter 3. Recent advances in the biology of echinostomes. Advances in Parasitology 69: 147-204.

98. Uchikawa, R., E. S. Loker. 1992. *Echinostoma paraensei* and *Schistosoma mansoni*: adherence of unaltered or modified latex beads to hemocytes of the host snail *Biomphalaria glabrata*. Experimental Parasitology 75: 223-232.

99. Vergote, D., Bouchut, A., Sautière, P. E., Roger, E., Galinier, R., Rognon, A., Coustau, C., Salzet, M., G. Mitta. 2005. Characterisation of proteins differentially present in the plasma of *Biomphalaria glabrata* susceptible or resistant to *Echinostoma caproni*. International Journal of Parasitology 35: 215-224.

100. World Health Organization: https://www.who.int/schistosomiasis/en/

101. Wu, X. J., Dinguirard, N., Sabat, G., Lui, H. D., Gonzalez, L., Gehring, M., Bickham-Wright, U., T. P. Yoshino. 2017. Proteomic analysis of *Biomphalaria glabrata* plasma proteins with binding affinity to those expressed by early developing larval *Schistosoma mansoni*. PLoS Pathogens 13: e1006081.

102. Zdobnov, E. M., von Mering, C., Letunic, I., Torrents, D., Suyama, M., Copley, R. R., Christophides, G. K., Thomasova, D., Holt, R. A., Subramanian, G. M., Mueller, H. M., Dimopoulus, G., Law, J. H., Wells, M. A., Birney, E., Charlab, R., Halpern, A. L., Kokoza, E., Kraft, C. L., Lai, Z., Lewis, S., Louis, C., Barillas-Mury, C., Nusskern, D., Rubin, G. M., Salzberg, S. L., Sutton, G. G., Topalis, P., Wides, R., Wincker, P., Yandell, M., Collins, F. H., Ribeiro, J., Gelbart, W. M., Kafatos, F. C., P. Bork. 2002. Comparative genome and proteome analysis of *Anopheles gambiae* and *Drosophila melanogaster*. Science 298: 149-159.

103. Zhang, S. M., Adema, C. M., Kepler, T. B., E.S. Loker. 2004. Diversification of Ig super-family genes in an invertebrate. Science 305: 251–254.

104. Zhang, L., Li, L., Guo, X., Litman, G. W., Dishaw, L. J., G. Zhang. 2015a. Massive expansion and functional divergence of innate immune genes in a protostome. Scientific Reports 5: 8693.

105. Zhang, L., Li, L., Zhu, Y., Zhang, G., X. Guo. 2016a Transcriptome analysis reveals a rich gene set related to innate immunity in the Eastern oyster (*Crassostrea virginica*). Marine Biotechnology 16: 17-33.

106. Zhang, S. M., Loker, E. S., J. T. Sullivan. 2016b. Pathogen-associated molecular patterns activate expression of genes involved in cell proliferation, immunity and detoxification in

the amebocyte-producing organ of the snail *Biomphalaria glabrata*. Developmental and Comparative Immunology 56: 25-36.

Chapter 5: From lab to field: Initial investigation of the ecoimmunology of *Physella acuta* snails

Jonathan H. Schultz and Coen M. Adema

Center of Evolutionary and Theoretical Immunology (CETI), Department of Biology, University of New Mexico, New Mexico, USA

Abstract

This study initiated the application of lab research-derived insights into molecular immunology of physid snails toward ecoimmunology of *Physella acuta*. A comparative analysis of lab-reared (L) and field (F) snails collected from Shady Lakes, a field-site in Albuquerque, NM was applied to investigate the immune status of snails in natural environments. Transcriptomes were sequenced with Illumina RNA-seq technology, and assembled to facilitate comparison of gene expression profiles of reference-lab snails (L; 7 lab-reared *P. acuta*) and four transcriptomes from individual field-derived (F) *P. acuta* snails. Gene ontology (GO) annotation of transcripts suggested a generally similar biological status for (L) and (F) snails, based on expression of similar numbers of genes and accommodation of the same level-2 GO category biological functions. Multi-dimensional scaling analysis (MDS), however, disclosed that (L) and (F) snails are functionally distinct due to expression of homologous genes at different levels (more extreme, higher or lower). Additionally, annotation of differentially expressed (DE) genes, with an emphasis on immune-relevant sequences, showed that (F) snails emphasize different categories of gene families to accommodate certain biological functions. Toward antibacterial immunity, (F) snails express Bactericidal permeability-increasing proteins (BPIs) and L-amino-acid oxidases (LAAOs) at similar or elevated levels, whereas macin-like and lysozyme sequences are expressed at the same or lower levels relative to (L) snails. Furthermore, (L) and (F) snails expressed immune sequences belonging to the same gene-categories, but utilized different (gene) members of each category. For example, fibrinogen-related domain containing sequences (FReDs) were recovered from both (L) and (F) snails, but (F) expressed a greater number of (uniquely different) FReD transcripts, indicative of importance of this category of genes for

P. acuta snails in natural settings. Overall, some differences exist between transcriptomic profiles of (L) and (F) *P. acuta*, but molecular immunology can be applied to assess the immune status of *P. acuta* in nature and enable future ecoimmunological studies to characterize nuances of (immune) processes that govern snail survival and persistence in the field where there is an abundance of diverse pathogens and parasites.

Introduction

The 'omics' era, driven by advancements in sequencing technology, has resulted in multifaceted approaches for investigating molecular aspects of organismal biology (Horner et al., 2010; Schultz and Adema, 2017). Investigators of diverse fields, like ecology and immunology, can benefit from these technological advancements, such that diverse biological questions are answered with the same tools. In particular, next-generation sequencing (NGS) facilitates analysis of the functional genomics (transcriptomics) of organisms in their natural habitats (Fuess et al., 2018). Merging ecological and immunological questions (ecoimmunology) helps to interpret how organisms interact with their biotic environments and how they survive threats from diverse pathogens (Demas and Nelson, 2012). These questions will also help interpret the persistence and distribution of snails in natural environments and elucidate the evolutionary histories of snail species and populations. Indeed, NGS aids characterization of the molecular underpinnings of immune defenses of animals that can be considered in an ecological context (Bowden et al., 2017).

Invertebrates make up most of the biodiversity, both in number of species and number of individual animals, and represent essential components of ecosystems across the globe (More et al., 2011). A considerable proportion of this biodiversity consists of gastropods (snails), members of the second largest animal phylum Mollusca. There are approximately 60,000

gastropod species that occupy diverse habitats including freshwater, marine, and terrestrial environments (Loker and Bayne, 2018). Gastropods serve a number of important ecological roles including as sources of food (Carnegie et al., 2016), and as intermediate hosts for trematode parasites, some that are medically relevant (Adema et al., 2012). The importance of exploring the ecoimmunology of these organisms is underscored by the need to track the impacts that globalization and global climate change will have on these economic sources and agents of disease transmission.

Globalization has encouraged the spread of invasive gastropods species, including the Giant African Snail (*Lissachatina fulica*; Shackelton et al., 2019) and the freshwater snail *Physella acuta* (Ebbs et al., 2018; Lawton et al., 2018). Of the two major isolates (Nolan et al., 2014), isolate A *P. acuta* snails exist globally whereas isolate B snails are restricted to the United States (Ebbs et al., 2018). Successful dispersal and colonization of invaded ranges by *P. acuta* can be attributed to many factors including possession of robust immune systems to combat newly encountered pathogens. Conversely, *Physella acuta* serves as intermediate host to a diversity of trematode parasites and may alter the endemic distribution of medical and veterinary relevant infectious diseases (Dreyfuss et al., 2002; Kanev, 1994). *Physella acuta* also serves as host for the laboratory model organism *Echinostoma paraensei* and recent investigations of *P. acuta* immunobiology after exposure to *E. paraensei* recorded unique aspects of gastropod immunity that revealed evolutionary development of taxon specific immune features for this snail species (Schultz et al., 2018a). In light of previous insights that invertebrates possess complex and sophisticated immune systems (Loker et al., 2004; Ghosh et al., 2011), and characterization of *P. acuta* immune repertoires we now begin to consider how immune function shapes the ecology of *P. acuta* and connect our understanding of gastropod immunity from

laboratory experiments to the field. However, there are no investigations of immunity of *P. acuta* snails in an ecoimmunological context and there are few examples utilizing snails from natural habitats to investigate immunity (Buddenborg et al., 2017; Zhao et al., 2017). Ecoimmunology investigations that utilize transcriptomics will provide a more comprehensive understanding of the interactions of host immune responses (gene expression) with native pathogens and parasites (Bankers et al., 2017).

Here, we initiate characterization of the ecoimmunology of *P. acuta* snails collected from Shady Lakes, a field site within Albuquerque, New Mexico. We assembled and characterized the transcriptomes of locally collected field (F) and lab-reared (L) *P. acuta* snails. Transcriptome profiles of individual *P. acuta* snails from Shady Lakes were annotated and compared to a transcriptome assembled from (L) snails employing general annotation with gene ontology predictions. Research of lab-maintained snails, assumed to live under artificially favorable conditions for snails, can reveal the mechanistic basis for immune processes and stress responses that should be applicable to track the immune status of snails in the field. To test this assumption, we determined differential expression of immune-relevant sequences from field-collected snails relative to lab snails. Beyond the analyses presented, the transcriptomic data collected here provides a valuable resource for future investigations that aim to decipher the immune capacities (and other biological features) that contribute to the success of a globally invasive gastropod.

Materials and Methods

Organisms

Two groups of *P. acuta* snails were used in this study. Lab-reared *P. acuta* snails (isolate A), housed in aquaria and fed red-leaf lettuce *ad libitum*, are maintained at the University of New Mexico. Eighty-four physid snails (7-9mm) from the field, morphologically identified by

presence of slender tentacles and sinistral shells, were collected from Shady lakes (35.2144° N, 106.5956° W) every two weeks during June-August 2015. Seeking non-infected individuals, snails were screened for patent trematode parasite infections. Snails were placed in wells of a 24-well plate with added artificial spring water (ASW) and left overnight to allow the release of larval stages of trematode parasites (cercariae). The following day, snails were screened for patent infections and preserved in RNAlater (ThermoFisher) for downstream analyses. Prior to nucleic acid extraction, whole-body tissues were slightly disrupted and inspected under a stereomicroscope to further confirm the absence of parasites.

Nucleic acid extraction and NGS sequencing

For downstream 16S PCR amplification and sequencing, DNA was extracted from a part of the headfoot tissues of field-collected snails using the CTAB method of Winnepenninckx et al. (1993). Extracted DNA samples were treated with RNAse A for 1 h at 37 °C (ThermoFisher).

RNA was extracted from whole-body tissues of (L) and (F) *P. acuta* snails using Trizol (Invitrogen) following the manufacturer's protocol. For removal of residual genomic DNA, RNA samples were treated with DNA-free™ (DNA Removal Kit, Ambion, Inc.). RNA quality was checked using a bioanalyzer (Agilent 2100, nano kit). All RNA samples were prepared for Illumina mRNA sequencing using Kappa Library Amplification kits (Kapabiosystems) with TruSeq (v2, ILLUMINA) adaptors and then sequenced on the Illumina NextSeq 500 platform (150bp, paired-end). The read data were submitted to the SRA database of GenBank (accession pending).

Species identification

PCR amplification of 16S rDNA sequences from ten field-collected snails was performed using DNA samples from snail headfoot tissue and conserved primers (F: 5'

ACGTGATCTGAGTTCAGACCGG 3', R: 5' CGCCTGTTTATCAAAAACAT 3'; Wethington and Lydeard, 2007). Amplicons were sequenced directly on both strands (BigDye v3.1, ABI 3130x) at the Molecular Biology Facility of UNM and chromatograms were edited by eye in Sequencher (v4.10.1). 16S nucleotide sequences were aligned (Clustal X; Larken et al., 2007) with 47 16S sequences of hygrophilid snails (PopSet: 164430551; Wethington and Lydeard, 2007). A phylogenetic reconstruction of 16S ribosomal genes was performed using the maximum-likelihood method and HKY nucleotide substitution model in MEGA 6 (bootstrap at 100 iterations). The lymnaeid snail *Stagnicola elodes* was used as outgroup.

Transcriptome assemblies

Illumina RNA-seq reads from four (F) *P. acuta* snails and seven (L) *P. acuta* snails were processed using Trimmomatic (Bolger et al., 2014). Reads from the seven (L) snails were combined to generate a reference lab assembly (RL) using Trinity *de novo* version 2.5.1 (Grabherr et al., 2011) with default parameters. Assemblies were also generated for each of the four (F) *P. acuta* snails. Transcripts from each assembly were processed using CD-HIT (Li and Godzik, 2006) at a 95% clustering threshold to minimize contig-redundancy and then mapped to the *P. acuta* genome assembly (GenBank accession: RDRX00000000) using megablast to obtain transcripts of snail origin. To identify transcripts that were differentially expressed in (F) snails relative to (L) snails, reads from all samples were combined to generate a comprehensive transcriptome (needed for mapping reads) using the same methods described above. Read abundance was quantified with RSEM (Li and Dewey, 2011) and differential expression (DE) was performed using EdgeR (Robinson et al., 2010) with default parameters.

Annotation of transcripts

Nucleotide sequences of *in silico* predicted transcripts of the (RL) transcriptome and four (F) transcriptomes were translated using TransDecoder v5.3.0 to obtain open reading frames (ORFs) with a minimum length of 100 amino acids. Predicted protein sequences were annotated with Gene Ontology (GO) identifiers using AgBase-GOanna (McCarthy et al., 2006), the blastp algorithm, Ag-Uniprot database, and default parameters (e-value threshold = 10e-06).

The (L) and (F) transcriptomes of *P. acuta* snails were annotated with DIAMOND (Buchfink et al., 2015) blastx (e-value threshold of 10e-06) using the non-redundant (nr) protein database of GenBank. Keyword searches for "cytochrome" and "16S" were performed to identify any transcripts that might indicate the presence of parasites within the snails. Additionally, transcriptomes were subjected to a blastx similarity search against trematode, nematode, and mollusc protein sequences retrieved from NCBI (October 2018) to confirm the absence of parasites from (F) snails.

Annotation of DE nucleotide sequences was performed with DIAMOND blastx with an e-value threshold of 10e-03. The annotation of the BLAST hit with the highest similarity score was retained for each transcript. DE sequences were also annotated with GO identifiers using the same methods listed above.

Results

Field collected snails, presence of parasites

504 Physids were collected from the field over a three-month interval during the Summer of 2015. These snails had a trematode infection prevalence of 4%, shedding a diverse assemblage of strigeid and echinostome parasites. The field snails (SL_1, Sl_2, SL_4, SL_9) used in this study did not shed any cercariae after incubation overnight in ASW and there were no obvious

signs of parasite infection within dissected snail tissues. In addition, bioinformatic analyses did not yield any transcripts of parasite origin from these (F) snails.

Species identification

Phylogenetic reconstruction using 16S genetic markers (GenBank accessions pending) identified five of ten field-collected physid snails as *P. acuta* that group within the *P. acuta* clade (Fig. 1.). Field-snail sample SL_4 is more closely related to *P. acuta* isolate B, whereas field samples SL_1, SL_2, Sl_9, and SL_10 were closer to *P. acuta* isolate A, the invasive lineage of *P. acuta* (Fig. 1). Note that the phylogenetic reconstruction also identified an uncharacterized clade of physid snails present at Shady Lakes that includes field samples SL_3, Sl_6, SL_7, and SL_8, grouping separate from *Aplexa spp.* and the branches that represent named species of physids (Fig. 1.).

GO annotation

A total of 1,752 translated nucleotide sequences from the (RL) transcriptome were assigned level-2 GO annotations (Fig. 2A.). 1,513 Translated nucleotide sequences of sample SL_1 received level-2 GO annotations, 1,262 for sample SL_2, 1,276 for sample SL_4, and 1,198 for sample SL_9 (Fig. 2A). The number of assignments belonging to the three GO categories biological process, cellular component and molecular function are shown in Figure 2A. From each transcriptome, the number of genes for each specific level-2 GO category is plotted (Fig. 2B). The processes supported by the transcriptomic profiles of (L) and (F) snails were generally similar; they were assigned mostly the same assemblage of level-2 categories. A few differences were evident, however; “detoxification” was not assigned to sample SL_1 and “carbohydrate utilization” was assigned only to samples SL-2 and SL_4 (Fig. 2B).

Differential expression

Despite highly similar GO profiles, 5,410 sequences were differentially expressed in the (F) snails (combined) relative to the (L) snails. Of these, 2,161 sequences were upregulated and 3,249 were downregulated. Multidimensional scaling (MDS) of gene expression (log_2foldchange) separated (L) and (F) snails into two distinct clusters due to variation in the level of expression of homologous genes, with more extreme (up or down) expression levels in (F) snails (Fig. 3).

Blastx similarity searches yielded annotation for 39% of upregulated sequences and 47% of downregulated sequences from (F) snails. Excluding sequences that remained without annotation, the 100-most upregulated and 100-most downregulated sequences are shown (Table 1). Twenty-three categories of immune-relevant sequences were recovered among the annotated DE genes, and the number of sequences that represent each of these immune categories is listed in Table 2. Unique transcript sequences representing bactericidal permeability-increasing proteins, baculoviral IAP-repeat proteins, CD109/TEP proteins, glutathione S-transferases, L-amino-acid oxidases, and sialic-acid binding lectins were only expressed at the same or higher levels in (F) snails relative to (L) snails (Table 2). Conversely, homologous sequences of the immune-gene families beta-1,3-glucan-binding protein-like, caspase, G-type lysozyme, macin-like, and peroxiredoxin were expressed at similar or lower levels by (F) snails relative to (L) snails (Table 2). The number of sequences for each immune-gene category recorded from the four individual F transcriptomes is also presented (Table 2). Each individual (F) *P. acuta* snail expressed all of the orthologues of the immune-relevant sequences in each category.

GO annotation of DE genes revealed the absence of upregulated genes that were assigned "transcription regulator activity" and "cell proliferation" level-2 GO identifiers. Downregulated

sequences did not include transcripts for the GO categories“synapse part”, “molecular function regulator”, “locomotion”, and “cell killing” (Fig. 4). Also, there is variation in gene expression of sequences belonging to most GO categories, such that (F) snails utilize different sequences relative to (L) snails.

Discussion

Employing ecoimmunology approaches to organismal biology allows for determination of immune status and how homeostasis is achieved by animals in natural environments, often faced with increased threats. Shady lakes (Albuquerque, New Mexico) is a field site with high levels of biodiversity including imported aquatic plants, trout fishes, aquatic crustaceans, and diverse gastropod species. The latter includes introduced caenogastropod species (*Bellamya chinensis*; Johnson et al., 2009) and members of the freshwater panpulmonate clade Hygrophila, including lymnaeid snails (*Pseudosuccinea columella*; GenBank accessions MG595730-MG595732, MG595738-MG595740), planorbid snails (including *Planorbella duryi*, Schultz et al., 2018b), ancylids, and physid snails. Taxonomic identification of the latter revealed a diversity of snail species, including *P. acuta* and an uncharacterized clade of physids (Fig. 1). Additional research is needed to investigate this potentially new species in light of the possibility for yet other unique mitochondrial gene orders to be discovered in physid snails (Nolan et al., 2014) and the capacity for this clade to serve as host for parasites (Adema et al., 2009; Ebbs et al., 2016), including digenetic trematodes, as observed from snails collected for this study at Shady Lakes, with a trematode infection prevalence of 4%. To eliminate parasite infection as a confounding variable, *P. acuta* snails that did not harbor infections were selected to compare the transcriptomic profiles of field-collect (F) snails and (L) lab-reared snails.

Overall, general annotation of the four (F) transcriptomes and one (RL) transcriptome (comprised of sequenced Illumina reads from 7 lab *P. acuta* snails) indicated that all these snails exhibited similar biological functioning, as revealed by GO annotation (Fig. 2B). These similarities include the number of genes expressed by each snail for each level-2 GO category (Fig. 2B) and also the same GO categories that were represented equally within each transcriptome. One conspicuous difference includes the absence of *in silico* predicted "carbohydrate utilization" activity from the (L) snails and (F) snails SL_1 and SL_9. Gene ontology of "carbohydrate utilization" includes functions such as sugar transport and sugar metabolism (within an organism) that are activated if carbohydrates, like glucose, are limiting. This suggests the potential for variation in the availability of nutrients among field snails, and may indicate differences in resource allocations for (F) snails in a natural setting. This also suggests that (L) snails receive adequate nutrients in lab conditions. Surprisingly, there were few gene annotations for "detoxification" in each (F) transcriptome, and these were completely absent from (F) snail SL_1. It was anticipated that (F) snails would express more genes associated with detoxification processes, but the (L) transcriptome expressed more genes (4) for this process. It is possible that increased gene number for detoxification processes in the (L) transcriptome is an artifact of using more snail sequences to generate the (L) transcriptome (7 samples), increasing the likelihood of recording more genes with this function. However, it is also possible that lab culture conditions, which are void of environmental perturbation, like precipitation or wind that circulate water, foster environments that require detoxification processes to take place in (L) *P. acuta* snails.

Despite overall similarity in biological activities engaged by (L) and (F) snails (as suggested by GO analysis), transcript-level analysis identified differential expression of 5,410

sequences in (F) snails relative to (L) *P. acuta* snails. Furthermore, multidimensional scaling of fold changes of expression of homologous genes separates (L) and (F) snails into groups (Fig. 9). There are quantifiable differences in gene expression that do indicate biological differences of these two groups. Inspection of immune-relevant sequences from the 100 most-upregulated sequences showed a higher level of expression of particular FReDs in (F) snails relative to lab snails (Table 1,2). Gastropods and bivalves express sequences containing fibrinogen-related domains in response to bacterial and parasite exposure (Schultz et al., 2018a; Romero et al., 2011; Wang et al., 2019). It is proposed that (F) snails effect increased expression (level and number) of these types of sequences to deal with more diverse pathogen threats in natural habitats. Also, genes associated with sequestering harmful effects of oxygen radicals (glutathione S-transferases) are among the 100 most-upregulated sequences in (F) snails and none are downregulated relative to lab snails. Again, GO annotation did not indicate major differences in detoxification processes among (L) and (F) snails, and closer inspection reveals that (F) snails are likely experiencing increased levels of cellular stress (Sharma et al., 2004). Conversely, cathepsin-like (lysosomal proteases) sequences are among the most downregulated in (F) snails when compared to (L) snails (Table 1,2). It is challenging to evaluate the biological significance of this DE for such a functionally broad group of sequences; one explanation may be that there is a greater availability of resources for (F) snails compared to lab counterparts, and (F) snails do not utilize protein degradation for supplemental nutrition.

More remarkable still, is that chitinases and PGRPs, involved in putative antifungal and antibacterial recognition, are among the most downregulated sequences in (F) snails (Table 1). Other sequences encoding proteins with antimicrobial activity (G-type lysozyme and macin-like) are also expressed at low levels in (F) snails as compared to (L) snails (Table 2). It is evident that

lab-reared *P. acuta* snails are not free of microbial threats, but instead face a consistent burden of bacterial pathogens, perhaps abundant uniquely in lab conditions. Indeed, there are (F) snails that display high levels of expression of antimicrobial sequences belonging to different categories including bactericidal permeability-increasing proteins (BPI) and L-amino-acid oxidases (LAAOs). The use of antimicrobial sequences of different immune categories may indicate that the immune activities of *P. acuta* are tailored to prominent pathogens faced in the field versus a lab environment. Importantly, there is selective use (DE) of individual members of immune-relevant gene families, including FReDs, dermatopontins, galectins, among others. Previous investigation of the response of *P. acuta* to the trematode parasite *E. paraensei* demonstrated a similar pattern, such that any gene members of an immune-relevant gene family can be DE in response to parasite exposure (Schultz et al., in prep). These findings illustrate the capability of *P. acuta* to discern and respond to diverse assemblages of pathogens, concordant with the capabilities of other mollusks (Deleury et al., 2012; Zhang et al., 2015).

The results of these analyses demonstrate that general annotation (GO analysis) of sequences from (F) and (L) *P. acuta* snails may fail to offer readily apparent differences of biological function and demonstrate the need to use bioinformatic software judiciously (Schultz and Adema, 2017). However, analysis of the differential expression of the immune-relevant sequence repertoires of (L) and (F) *P. acuta* snails, does reveal functional differences between these groups. It seems that (L) and (F) snails maintain unique immune statuses by DE (levels and members of immune-gene families) to deal with stressors. Furthermore, the abundant diversity and expression of FReD sequences from field snails (Table 2) further supports the use of these types of proteins for immunoprotection against pathogens. The ecoimmunology approach of investigating *P. acuta* immunity begins to facilitate interpretation of the persistence of snails in

natural environments and identifies putative biomarkers that can identify the types of stressors experienced. Identification and characterization of immune-relevant sequences identified DE in (F) snails provide transcriptional profiles that define the molecular nature of field caught and lab-reared *P. acuta* snails (Figure 5).

Continued investigation of the specific responses of *P. acuta* in natural settings will aid better understanding of the complex molecular mechanisms that underlie the survival and success of gastropods. The insights garnered from laboratory experiments, such as the differential use (by molluscs) of sequences within immune-gene families to meet unique pathogen burdens, are supported by this study utilizing field snails. Ecoimmunology approaches, coupled with analysis of lab and field organisms, are likely to provide the most accurate interpretations of immune status.

Acknowledgements

The authors wish to acknowledge the molecular biology facility (MBF) at the University of New Mexico for sequencing and technical support and the Center for Evolutionary and Theoretical Immunology (CETI) for NGS data.

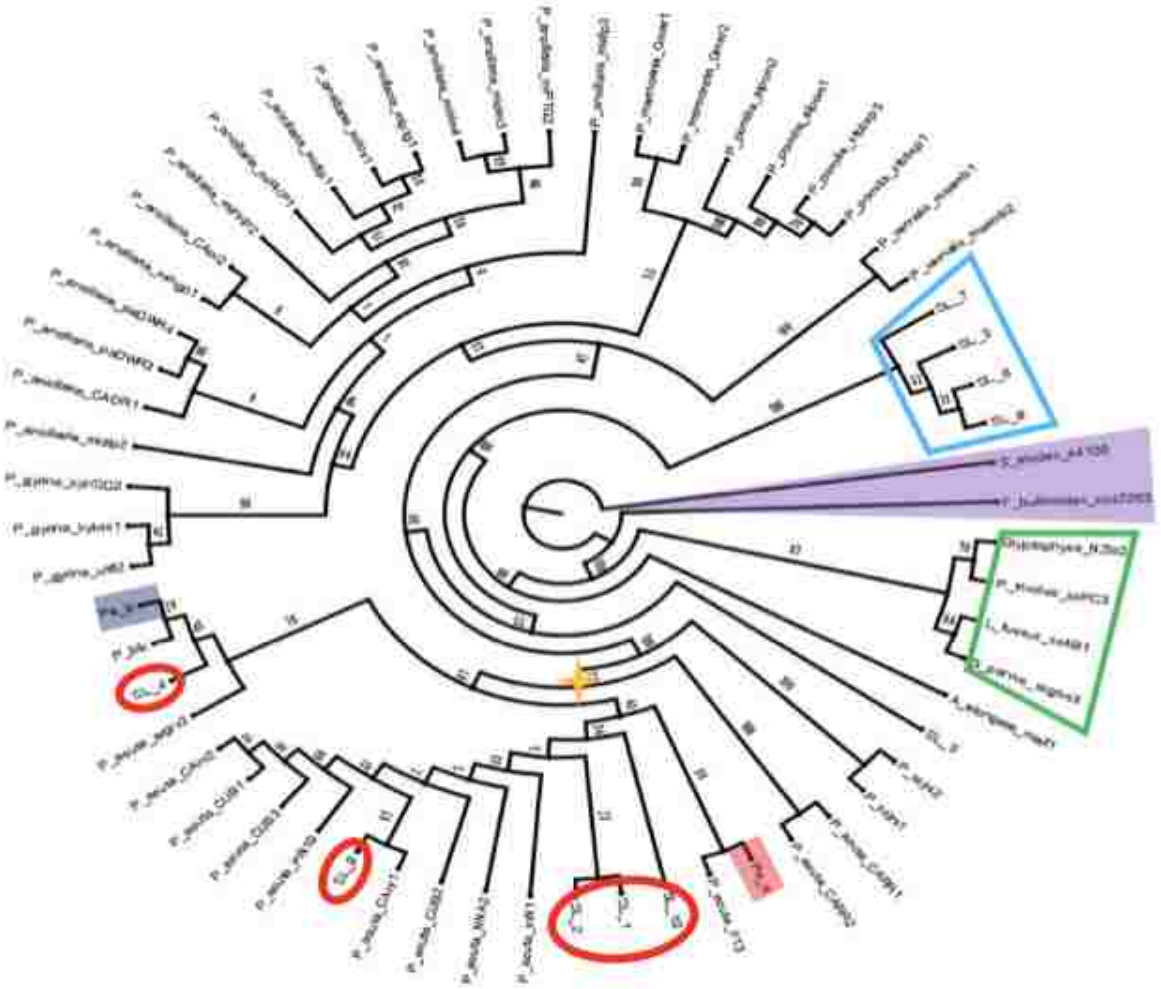


Fig. 1. Phylogenetic reconstruction of 16s ribosomal genes of physid snails collected from Shady lakes and other hygrophilid snails obtained from GenBank.
Phylogenetic tree was generated using the maximum-likelihood and HKY nucleotide substitution model (bootstrap 100 iterations). Yellow star marks node of *P. acuta* clade; red circles identify *P. acuta* snails collected from Shady Lakes; Red and blue shading identify *P. acuta* isolates A & B, respectively (Nolan et al., 2014); Uncharacterized clade of physids outlined in blue; Planorbid snails outlined in green; Outgroup (lymnaeid snails) highlighted in purple. Remaining physid snails are not highlighted.

Fig. 2. Gene ontology analysis of transcriptomes of field-collected *P. acuta* snails and transcriptome of lab-derived (control) *P. acuta* snails.
(a) Number of genes from each transcriptome assigned annotation to three main GO categories biological process, cellular component, and molecular function. (b) Number of genes from each transcriptome assigned to specific level-2 GO categories.

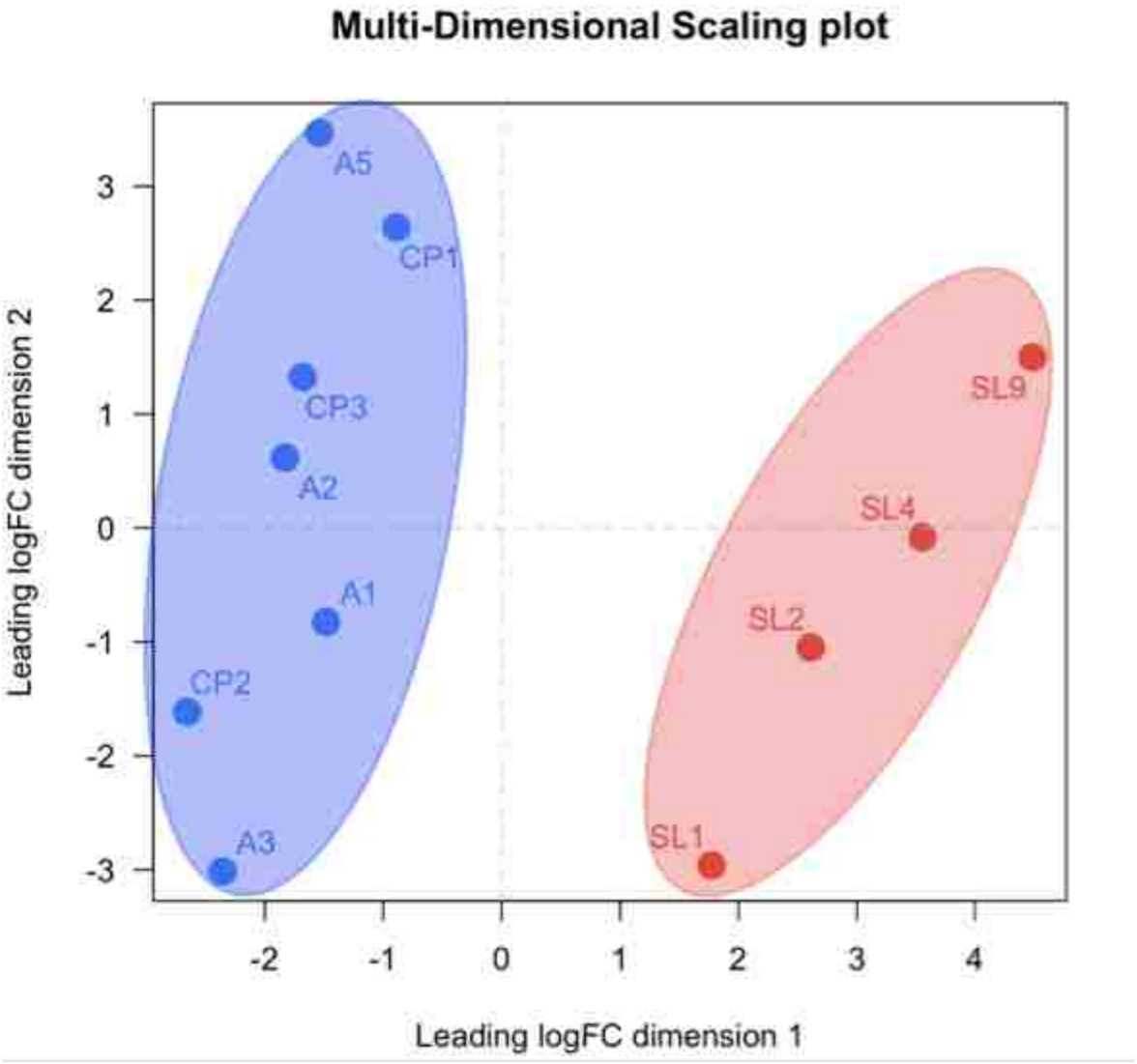


Fig. 3. Multidimensional scaling of pairwise comparisons of differences of the log$_2$foldchange of homologous genes between lab-mainted and field-collected *P. acuta* snails.
Despite variation of differential expression within groups (along y-axis), lab-maintained (highlighted in blue) and field-collected snails (highlighted in red) are separated distinctly along the x-axis.

Fig. 4. Gene ontology analysis of field-collected *P. acuta* DE sequences.
Graphical display of level-2 GO terms assigned to DE sequences of (F) *P. acuta* snails relative to (L) *P. acuta* snails . Level-2 GO terms are labelled along the x-axis and gene number is displayed on the y-axis.

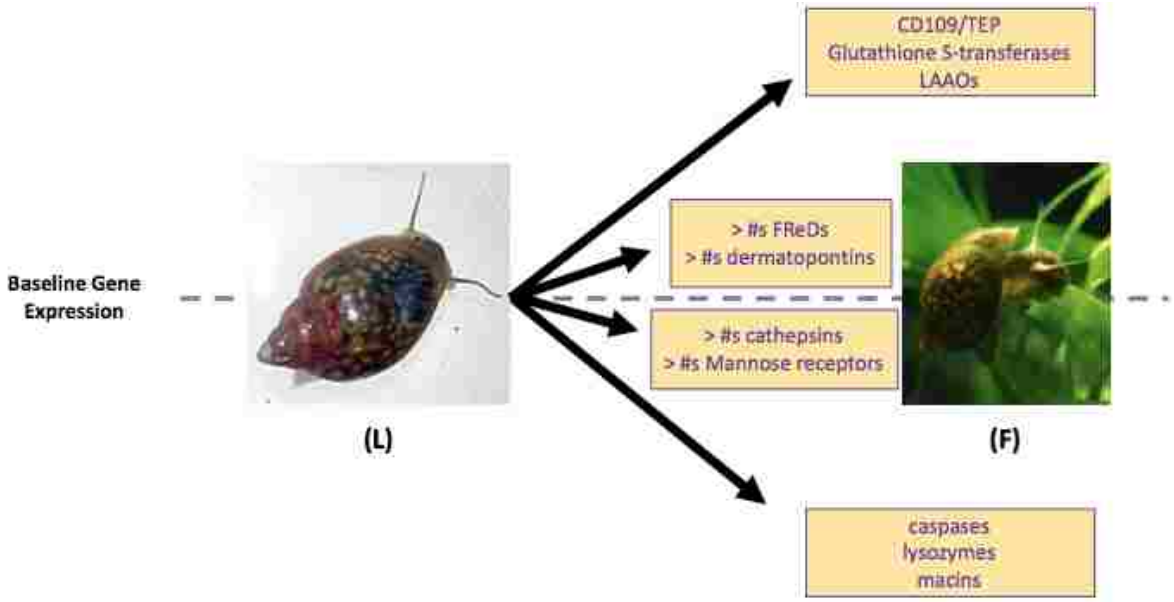


Fig. 5. DE transcription profiles distinguish field-collected and lab-reared *P. acuta* snails. Analysis of (F) *P. acuta* sequences differentially expressed relative to (L) *P. acuta* snails identifies immune-relevant sequences that inform the biological activity of field-caught *P. acuta*. (F) snails express CD109/TEP sequences at same or elevated levels relative to (L) snails and differentially express greater numbers of FReD and dermatopontin-like sequences. Conversely, fewer numbers of cathepsin and mannose receptor-like sequences are DE and (F) snails express caspases, lysozymes, and macins at same or lower levels relative to (L) snails.

Table 1. Annotation of the 100-most up and downregulated sequences of (F) *P. acuta* snails relative to (L) *P. acuta* snails. Log_2foldchange is listed adjacent to sequence annotations. Colors designate immune-relevant sequences of different functional categories: Red, FREDs; Green, glutathione S-transferases; Purple, cathepsins; Blue, pathogenesis-related sequences; Orange, GTPase IMAPs; Yellow, PGRP/chitinases; Light blue, cellular differentiation factor.

log2FoldChange	Upregulated Sequence annotation	log2FoldChange2	Downregulated sequence annotation
15.103	AJP62180.1 cytochrome c oxidase subunit 3 (mitochondrion) [Radix swinhoei]	-16.003	AFE62779.1 NADH dehydrogenase subunit 4L (mitochondrion) [Physella acuta]
14.45	XP_005107584.1 PREDICTED: 60S ribosomal protein L19 isoform X1 [Aplysia californica]	-15.127	XP_005090789.1 PREDICTED: lipoxygenase homology domain-containing protein 1-like isoform X1 [Aplysia californica]
14.226	AFE62782.1 NADH dehydrogenase subunit 2 (mitochondrion) [Physella acuta]	-14.757	YP_008994234.1 cytochrome c oxidase subunit II (mitochondrion) [Physella acuta]
12.722	XP_013062476.1 PREDICTED: signal recognition particle subunit SRP72-like [Biomphalaria glabrata]	-13.166	EFX82968.1 hypothetical protein DAPPUDRAFT_100890 [Daphnia pulex]
12.407	CDW75723.1 UNKNOWN [Stylonychia lemnae]	-13.12	YP_008994234.1 cytochrome c oxidase subunit II (mitochondrion) [Physella acuta]
12.198	XP_013067909.1 PREDICTED: heat shock protein Hsp-12.2-like [Biomphalaria glabrata]	-13.116	XP_013080229.1 PREDICTED: protein AIG1-like [Biomphalaria glabrata]
12.153	XP_013069998.1 PREDICTED: ficolin-2-like [Biomphalaria glabrata]	-13.052	XP_013088257.1 PREDICTED: uncharacterized protein LOC106072429 [Biomphalaria glabrata]
12.121	XP_013787111.1 perivitellin-2 67 kDa subunit-like [Limulus polyphemus]	-13.043	XP_021139940.1 endonuclease/exonuclease/phosphatase family domain-containing protein 1, partial [Columba livia]
12.015	XP_005101950.1 PREDICTED: COP9 signalosome complex subunit 8-like [Aplysia californica]	-12.675	XP_013062270.1 PREDICTED: E3 ubiquitin-protein ligase TRIM33-like isoform X1 [Biomphalaria glabrata]
11.944	PVD26222.1 hypothetical protein C0Q70_13892 [Pomacea canaliculata]	-12.478	XP_021139940.1 endonuclease/exonuclease/phosphatase family domain-containing protein 1, partial [Columba livia]
11.853	XP_012943100.1 PREDICTED: uncharacterized protein LOC101862895 [Aplysia californica]	-12.273	XP_021139940.1 endonuclease/exonuclease/phosphatase family domain-containing protein 1, partial [Columba livia]
11.813	WP_079377292.1 hypothetical protein [Acinetobacter baumannii]	-12.217	KRH93531.1 hypothetical protein M153_7960002914 [Pseudoloma neurophilia]
11.524	XP_013076736.1 PREDICTED: integrin beta-PS-like [Biomphalaria glabrata]	-12.1	XP_022080279.1 endonuclease/exonuclease/phosphatase family domain-containing protein 1-like isoform X2 [Acanthaster planci]
11.508	XP_013062909.1 PREDICTED: glutathione S-transferase omega-1-like [Biomphalaria glabrata]	-12.067	XP_012940968.1 PREDICTED: cAMP-regulated D2 protein-like [Aplysia californica]
11.505	EJY65597.1 hypothetical protein OXYTRI_14248 (macronuclear) [Oxytricha trifallax]	-11.939	ABK90856.1 cathepsin-L-like cysteine peptidase, partial [Radix peregra]

11.291	XP_013091464.1 PREDICTED: uncharacterized protein LOC106075092 [Biomphalaria glabrata]	-11.93	XP_013075352.1 PREDICTED: G patch domain-containing protein 3-like [Biomphalaria glabrata]
10.954	XP_012942591.1 PREDICTED: aldehyde dehydrogenase, mitochondrial-like [Aplysia californica]	-11.653	KRX07431.1 Protein kinase-like domain [Pseudocohnilembus persalinus]
10.888	XP_005101068.2 PREDICTED: uncharacterized protein LOC101863352 [Aplysia californica]	-11.615	XP_005094471.1 PREDICTED: methionine synthase reductase-like [Aplysia californica]
10.804	XP_004516261.2 PREDICTED: protein TAR1-like [Cicer arietinum]	-11.559	XP_012941557.1 PREDICTED: pre-B-cell leukemia transcription factor 1-like [Aplysia californica]
10.734	AAS86709.1 unknown, partial [Lymnaea stagnalis]	-11.53	XP_021139940.1 endonuclease/exonuclease/phosphatase family domain-containing protein 1, partial [Columba livia]
10.603	ADK11409.1 fibrinogen-related protein 2.27 [Biomphalaria glabrata]	-11.518	XP_013073227.1 PREDICTED: protein AIG1-like [Biomphalaria glabrata]
10.596	XP_013076566.1 PREDICTED: tenascin-like [Biomphalaria glabrata]	-11.48	XP_005097791.2 PREDICTED: CD82 antigen-like [Aplysia californica]
10.564	XP_013066622.1 PREDICTED: centaurin-gamma-1A-like [Biomphalaria glabrata]	-11.296	XP_005110158.1 PREDICTED: cytosolic Fe-S cluster assembly factor narfl-like [Aplysia californica]
10.504	XP_013081771.1 PREDICTED: cystathionine gamma-lyase-like [Biomphalaria glabrata]	-11.266	XP_013065887.1 PREDICTED: 1-acyl-sn-glycerol-3-phosphate acyltransferase alpha-like [Biomphalaria glabrata]
10.5	XP_013093076.1 PREDICTED: ficolin-1-like [Biomphalaria glabrata]	-11.249	XP_013088548.1 PREDICTED: dnaJ homolog subfamily B member 9-like [Biomphalaria glabrata]
10.465	XP_013069998.1 PREDICTED: ficolin-2-like [Biomphalaria glabrata]	-11.174	XP_013073227.1 PREDICTED: protein AIG1-like [Biomphalaria glabrata]
10.401	XP_013074715.1 PREDICTED: pathogenesis-related protein PR-1-like [Biomphalaria glabrata]	-11.15	XP_013094731.1 PREDICTED: LOW QUALITY PROTEIN: carbamoyl-phosphate synthase [ammonia], mitochondrial-like [Biomphalaria glabrata]
10.385	XP_005094856.1 PREDICTED: ubiquitin-like domain-containing CTD phosphatase 1 [Aplysia californica]	-11.061	XP_012936974.1 PREDICTED: cathepsin L1-like [Aplysia californica]
10.351	KIK49843.1 hypothetical protein GYMLUDRAFT_78612 [Gymnopus luxurians FD-317 M1]	-11.036	XP_013069704.1 PREDICTED: A-kinase anchor protein 14-like [Biomphalaria glabrata]
10.344	XP_013079396.1 PREDICTED: uncharacterized protein LOC106065180 [Biomphalaria glabrata]	-11.034	XP_013096901.1 PREDICTED: 5-AMP-activated protein kinase catalytic subunit alpha-2-like isoform X1 [Biomphalaria glabrata]
10.313	XP_013095240.1 PREDICTED: uncharacterized protein LOC106078775 [Biomphalaria glabrata]	-10.963	XP_005091119.1 PREDICTED: uncharacterized protein LOC101855924 [Aplysia californica]
10.292	XP_013076253.1 PREDICTED: glutamine synthetase-like [Biomphalaria glabrata]	-10.92	XP_013073227.1 PREDICTED: protein AIG1-like [Biomphalaria glabrata]
10.287	XP_013068537.1 PREDICTED: organic solute transporter subunit alpha-like [Biomphalaria glabrata]	-10.909	XP_013083326.1 PREDICTED: uncharacterized protein LOC106068505 [Biomphalaria glabrata]
10.283	XP_013088844.1 PREDICTED: alpha-aminoadipic semialdehyde synthase, mitochondrial-like, partial [Biomphalaria glabrata]	-10.882	XP_012941174.1 PREDICTED: uncharacterized protein LOC101864166 [Aplysia californica]

10.24	XP_013084778.1 PREDICTED: uncharacterized protein LOC106069617 [Biomphalaria glabrata]	-10.863	XP_002128544.1 PREDICTED: E-selectin-like [Ciona intestinalis]
10.16	XP_013069998.1 PREDICTED: ficolin-2-like [Biomphalaria glabrata]	-10.855	PVD26238.1 hypothetical protein C0Q70_13908 [Pomacea canaliculata]
10.099	XP_013071551.1 PREDICTED: glutathione S-transferase 3-like [Biomphalaria glabrata]	-10.84	XP_013077685.1 PREDICTED: collagen alpha-6(VI) chain-like [Biomphalaria glabrata]
9.996	XP_012941987.1 PREDICTED: glutamate--cysteine ligase catalytic subunit-like [Aplysia californica]	-10.819	XP_013090604.1 PREDICTED: uncharacterized protein LOC106074385 isoform X2 [Biomphalaria glabrata]
9.966	XP_013787111.1 perivitellin-2 67 kDa subunit-like [Limulus polyphemus]	-10.818	XP_013093592.1 PREDICTED: 40S ribosomal protein S6 [Biomphalaria glabrata]
9.819	XP_009062522.1 hypothetical protein LOTGIDRAFT_166832 [Lottia gigantea]	-10.783	XP_005100229.1 PREDICTED: COP9 signalosome complex subunit 6-like [Aplysia californica]
9.811	CDW75723.1 UNKNOWN [Stylonychia lemnae]	-10.775	XP_013092519.1 PREDICTED: mitogen-activated protein kinase kinase kinase 3-like, partial [Biomphalaria glabrata]
9.781	XP_013061045.1 PREDICTED: PAX3- and PAX7-binding protein 1-like [Biomphalaria glabrata]	-10.758	XP_013081676.1 PREDICTED: uncharacterized protein LOC106067101 [Biomphalaria glabrata]
9.735	XP_017312137.1 PREDICTED: E-selectin-like isoform X1 [Ictalurus punctatus]	-10.678	XP_013060574.1 PREDICTED: GTPase IMAP family member 7-like [Biomphalaria glabrata]
9.718	XP_013061143.1 PREDICTED: retinal dehydrogenase 1-like [Biomphalaria glabrata]	-10.652	YP_008994234.1 cytochrome c oxidase subunit II (mitochondrion) [Physella acuta]
9.665	XP_013094774.1 PREDICTED: perivitellin-2 67 kDa subunit-like [Biomphalaria glabrata]	-10.541	XP_012940944.1 PREDICTED: low choriolytic enzyme-like [Aplysia californica]
9.634	XP_013092109.1 PREDICTED: uncharacterized protein LOC106075762 [Biomphalaria glabrata]	-10.507	XP_013072237.1 PREDICTED: cathepsin Z-like [Biomphalaria glabrata]
9.497	XP_022304386.1 kielin/chordin-like protein [Crassostrea virginica]	-10.375	XP_022085246.1 S-adenosylmethionine mitochondrial carrier protein-like isoform X1 [Acanthaster planci]
9.433	XP_013064603.1 PREDICTED: calmodulin-2/4-like [Biomphalaria glabrata]	-10.344	XP_006819549.1 PREDICTED: fidgetin-like protein 1-like [Saccoglossus kowalevskii]
9.393	XP_012938075.1 PREDICTED: Kv channel-interacting protein 4-like [Aplysia californica]	-10.319	EJD40057.1 PR-1-like protein [Auricularia subglabra TFB-10046 SS5]
9.346	XP_018645332.1 unnamed protein product [Schistosoma mansoni]	-10.294	XP_012938639.1 PREDICTED: transmembrane protein 214-B-like [Aplysia californica]
9.244	XP_013088844.1 PREDICTED: alpha-aminoadipic semialdehyde synthase, mitochondrial-like, partial [Biomphalaria glabrata]	-10.281	XP_012943827.1 PREDICTED: copper transport protein CTR2-like isoform X2 [Aplysia californica]
9.203	XP_013082098.1 PREDICTED: uncharacterized protein LOC106067464 [Biomphalaria glabrata]	-10.233	XP_013086565.1 PREDICTED: PHD and RING finger domain-containing protein 1-like [Biomphalaria glabrata]
9.186	XP_013079004.1 PREDICTED: serine-rich adhesin for platelets-like [Biomphalaria glabrata]	-10.214	AJQ21525.1 membrane-bound C-type lectin [Mytilus galloprovincialis]
9.116	XP_017948702.1 PREDICTED: E-selectin [Xenopus tropicalis]	-10.174	XP_005090375.1 PREDICTED: very-long-chain 3-oxoacyl-CoA reductase-like [Aplysia californica]
9.081	XP_013238305.1 hypothetical protein DI09_258p10 [Mitosporidium daphniae]	-10.157	XP_013079138.1 PREDICTED: peroxisomal biogenesis factor 3-like [Biomphalaria glabrata]

9.063	XP_002949364.1 hypothetical protein VOLCADRAFT_104287 [Volvox carteri f. nagariensis]	-10.15	XP_005107087.1 PREDICTED: dynein regulatory complex protein 1-like [Aplysia californica]
9.018	XP_013062913.1 PREDICTED: glutathione S-transferase omega-1-like [Biomphalaria glabrata]	-10.141	XP_013082283.1 PREDICTED: glycoprotein 3-alpha-L-fucosyltransferase A-like isoform X2 [Biomphalaria glabrata]
8.969	XP_013094215.1 PREDICTED: pathogenesis-related protein 1-like [Biomphalaria glabrata]	-10.116	ABK90856.1 cathepsin-L-like cysteine peptidase, partial [Radix peregra]
8.949	XP_005092070.1 PREDICTED: coiled-coil domain-containing protein 177-like isoform X1 [Aplysia californica]	-10.11	XP_013072214.1 PREDICTED: LOW QUALITY PROTEIN: circularly permutated Ras protein 1-like [Biomphalaria glabrata]
8.862	AAS86709.1 unknown, partial [Lymnaea stagnalis]	-10.106	XP_005101695.1 PREDICTED: uncharacterized protein LOC101852775 [Aplysia californica]
8.808	XP_013089668.1 PREDICTED: 3-ketodihydrosphingosine reductase-like isoform X2 [Biomphalaria glabrata]	-10.105	XP_005096114.2 PREDICTED: uncharacterized protein LOC101863806 [Aplysia californica]
8.781	XP_013071059.1 PREDICTED: zinc finger CCCH domain-containing protein 13-like isoform X2 [Biomphalaria glabrata]	-10.1	XP_013060574.1 PREDICTED: GTPase IMAP family member 7-like [Biomphalaria glabrata]
8.74	XP_013091641.1 PREDICTED: solute carrier family 13 member 2-like [Biomphalaria glabrata]	-10.075	XP_013069926.1 PREDICTED: OCIA domain-containing protein 1-like [Biomphalaria glabrata]
8.685	XP_013089827.1 PREDICTED: uncharacterized protein LOC106073743 [Biomphalaria glabrata]	-10.061	XP_013071165.1 PREDICTED: sodium/potassium-transporting ATPase subunit alpha-like [Biomphalaria glabrata]
8.631	XP_013095408.1 PREDICTED: probable 28S rRNA (cytosine(4447)-C(5))-methyltransferase isoform X1 [Biomphalaria glabrata]	-10.045	XP_013063573.1 PREDICTED: low-density lipoprotein receptor-related protein 6-like [Biomphalaria glabrata]
8.482	XP_022529974.1 putative ferric-chelate reductase 1 [Astyanax mexicanus]	-10.018	XP_012943585.1 PREDICTED: annexin A11-like [Aplysia californica]
8.45	XP_013063887.1 PREDICTED: LOW QUALITY PROTEIN: uncharacterized protein LOC106052963 [Biomphalaria glabrata]	-9.974	XP_012938377.1 PREDICTED: uncharacterized protein LOC101857719 [Aplysia californica]
8.365	XP_013061493.1 PREDICTED: ganglioside GM2 activator-like [Biomphalaria glabrata]	-9.949	XP_013094598.1 PREDICTED: GTPase IMAP family member 6-like [Biomphalaria glabrata]
8.314	XP_013069998.1 PREDICTED: ficolin-2-like [Biomphalaria glabrata]	-9.905	NP_001298195.1 uncharacterized LOC106076952 precursor [Biomphalaria glabrata]
8.276	XP_013083707.1 PREDICTED: uncharacterized protein LOC106068784 [Biomphalaria glabrata]	-9.905	XP_013063573.1 PREDICTED: low-density lipoprotein receptor-related protein 6-like [Biomphalaria glabrata]
8.246	XP_005097791.2 PREDICTED: CD82 antigen-like [Aplysia californica]	-9.856	XP_005103924.1 PREDICTED: fibrinogen-like protein 1 isoform X2 [Aplysia californica]
8.238	XP_013080817.1 PREDICTED: complex I assembly factor TIMMDC1, mitochondrial-like [Biomphalaria glabrata]	-9.856	YP_008994234.1 cytochrome c oxidase subunit II (mitochondrion) [Physella acuta]
8.233	XP_013061332.1 PREDICTED: uncharacterized protein LOC106050829 [Biomphalaria glabrata]	-9.816	XP_013091037.1 PREDICTED: cytochrome c oxidase assembly factor 6 homolog [Biomphalaria glabrata]
8.184	XP_013094890.1 PREDICTED: perivitellin-2 67 kDa subunit-like [Biomphalaria glabrata]	-9.777	XP_005105793.2 PREDICTED: uncharacterized protein LOC101863532 [Aplysia californica]
8.183	XP_013089663.1 PREDICTED: uncharacterized protein LOC106073600 [Biomphalaria glabrata]	-9.77	XP_013070226.1 PREDICTED: uncharacterized protein LOC106057535 [Biomphalaria glabrata]
8.149	XP_013069998.1 PREDICTED: ficolin-2-like [Biomphalaria glabrata]	-9.709	OPL21320.1 hypothetical protein AM593_00633, partial [Mytilus galloprovincialis]
8.14	XP_012935251.1 PREDICTED: signal recognition particle subunit SRP72-like [Aplysia californica]	-9.708	XP_022287001.1 estradiol 17-beta-dehydrogenase 11-like [Crassostrea virginica]
8.088	PVD26222.1 hypothetical protein C0Q70_13892 [Pomacea canaliculata]	-9.69	AEH26026.1 peptidoglycan recognition protein [Physella acuta]

8.081	XP_013065775.1 PREDICTED: zinc transporter ZIP14-like [Biomphalaria glabrata]	-9.686	XP_013084266.1 PREDICTED: GTPase IMAP family member 7-like, partial [Biomphalaria glabrata]
7.925	XP_005103148.1 PREDICTED: ribosome biogenesis protein WDR12 homolog [Aplysia californica]	-9.683	XP_013063598.1 PREDICTED: monocyte to macrophage differentiation factor 2-like [Biomphalaria glabrata]
7.836	XP_013094890.1 PREDICTED: perivitellin-2 67 kDa subunit-like [Biomphalaria glabrata]	-9.67	XP_013073193.1 PREDICTED: geranylgeranyl transferase type-1 subunit beta-like [Biomphalaria glabrata]
7.784	XP_013066470.1 PREDICTED: uncharacterized protein LOC106054929 [Biomphalaria glabrata]	-9.668	XP_013096244.1 PREDICTED: uncharacterized protein LOC106079598 isoform X1 [Biomphalaria glabrata]
7.778	XP_013090583.1 PREDICTED: poly [ADP-ribose] polymerase 14-like [Biomphalaria glabrata]	-9.622	XP_013068263.1 PREDICTED: putative gamma-glutamyltransferase YwrD isoform X1 [Biomphalaria glabrata]
7.75	XP_013094706.1 PREDICTED: U11/U12 small nuclear ribonucleoprotein 25 kDa protein-like [Biomphalaria glabrata]	-9.619	XP_013072829.1 PREDICTED: serine/threonine-protein kinase RIO1-like [Biomphalaria glabrata]
7.729	XP_013087668.1 PREDICTED: N-acetylated-alpha-linked acidic dipeptidase 2-like [Biomphalaria glabrata]	-9.609	XP_013075417.1 PREDICTED: LOW QUALITY PROTEIN: peroxisomal sarcosine oxidase-like [Biomphalaria glabrata]
7.715	AAO83839.1 voltage-dependent L-type calcium channel alpha-1 subunit isoform b [Lymnaea stagnalis]	-9.584	BAV14502.1 chitinase [Lymnaea stagnalis]
7.693	XP_022735763.1 uncharacterized protein LOC111289175 [Durio zibethinus]	-9.581	XP_012938639.1 PREDICTED: transmembrane protein 214-B-like [Aplysia californica]
7.684	CDW75723.1 UNKNOWN [Stylonychia lemnae]	-9.568	XP_021353692.1 centrosomal protein of 112 kDa-like isoform X2 [Mizuhopecten yessoensis]
7.656	XP_012940656.1 PREDICTED: uncharacterized protein LOC101858564 [Aplysia californica]	-9.564	PSN29075.1 hypothetical protein C0J52_28073 [Blattella germanica]
7.643	XP_005109862.1 PREDICTED: selenocysteine lyase-like [Aplysia californica]	-9.487	XP_013078574.1 PREDICTED: mitogen-activated protein kinase kinase kinase 3-like [Biomphalaria glabrata]
7.622	XP_013081493.1 PREDICTED: poly [ADP-ribose] polymerase 12-like [Biomphalaria glabrata]	-9.477	XP_013089724.1 PREDICTED: MIP18 family protein CG30152-like [Biomphalaria glabrata]
7.574	XP_013091357.1 PREDICTED: SCO-spondin-like [Biomphalaria glabrata]	-9.462	XP_013068891.1 PREDICTED: UPF0544 protein C5orf45 homolog [Biomphalaria glabrata]
7.534	XP_009062522.1 hypothetical protein LOTGIDRAFT_166832 [Lottia gigantea]	-9.459	XP_012935793.1 PREDICTED: uncharacterized protein LOC101853227 [Aplysia californica]
7.507	XP_013066482.1 PREDICTED: uncharacterized protein LOC106054940 [Biomphalaria glabrata]	-9.436	NP_001191497.1 60S ribosomal protein L37 [Aplysia californica]
7.494	XP_012946756.1 PREDICTED: unconventional myosin-VIIa-like [Aplysia californica]	-9.435	XP_013072593.1 PREDICTED: serine/threonine-protein phosphatase 6 regulatory ankyrin repeat subunit B-like [Biomphalaria glabrata]
7.467	XP_013085374.1 PREDICTED: cell wall protein DAN4-like [Biomphalaria glabrata]	-9.41	XP_005100955.1 PREDICTED: 39S ribosomal protein L14, mitochondrial-like isoform X2 [Aplysia californica]
7.389	XP_013065388.1 PREDICTED: serine--tRNA ligase, cytoplasmic-like [Biomphalaria glabrata]	-9.407	XP_011412676.1 PREDICTED: chondroitin proteoglycan 2 isoform X2 [Crassostrea gigas]
7.387	XP_013083283.1 PREDICTED: peroxisomal membrane protein 2-like [Biomphalaria glabrata]	-9.366	XP_013072744.1 PREDICTED: uncharacterized protein LOC106059623 [Biomphalaria glabrata]
7.373	XP_013079232.1 PREDICTED: solute carrier family 28 member 3-like [Biomphalaria glabrata]	-9.356	XP_013086487.1 PREDICTED: probable G-protein coupled receptor 21 [Biomphalaria glabrata]

Table 2. Number of immune-sequences of (F) *P. acuta* snails DE relative to (L) *P. acuta* snails and number of sequences from each immune-gene category expressed within individual transcriptomes of *P. acuta* snails collected from the field (SL_1,2,4,9). Colors identify groups of sequences of (F) snails that share the same expression profiles when compared to (L) snails. Note: FReD sequence numbers identified from field-collected transcriptomes are highlighted in red.

Immune Gene Annotation	DOWN	UP	Usage relative to lab-reared snails	Field Snails	SL 1	SL 2	SL 4	SL 9
beta-1,3-glucan-binding protein-like	1	0	same or down		4	4	2	3
caspase	4	0	same or down		45	41	35	29
G-type Lysozyme	3	0	same or down		11	8	14	4
macin-like	2	0	same or down		8	6	5	5
peroxiredoxin	1	0	same or down		10	10	11	9
bactericidal permeability-increasing protein (BPI)	0	1	same or up		11	8	11	11
baculoviral IAP-repeat	0	2	same or up		73	60	45	35
CD109/TEP	0	7	same or up		94	77	45	34
glutathione S-transferase	0	5	same or up		89	55	68	50
L-amino-acid oxidases (LAAO)	0	3	same or up		10	6	6	5
sialic acid-binding lectin	0	1	same or up		3	0	1	0
C-type Lectin	6	1	selective use of gene-family members		21	11	13	9
cathepsin	22	1	selective use of gene-family members		96	63	72	46
chitinase	2	2	selective use of gene-family members		46	27	34	25
complement-like	3	2	selective use of gene-family members		23	23	20	19
dermatopontin	1	9	selective use of gene-family members		28	24	26	27
Fibrinogen-related domain-containing (FReD)	8	34	selective use of gene-family members		301	246	234	147
galectin	2	3	selective use of gene-family members		35	17	29	13
GTPase IMAP	18	7	selective use of gene-family members		147	116	112	67
mannose receptor	11	3	selective use of gene-family members		64	43	41	32
Pathogenesis-related protein	3	5	selective use of gene-family members		42	33	31	26
Peptidoglycan-recognition protein (PGRP)	2	1	selective use of gene-family members		23	19	15	11
Toll-like receptor (TLR)	1	1	selective use of gene-family members		116	74	74	54

References

1. Adema, C.M., Lun, C.M., Hanelt, B., Seville, R.S., 2009. Digenean trematode infections of native freshwater snails and invasive *Potamopyrgus antipodarum* in the Grand Teton National Park/John D. Rockefeller Memorial Parkway Area. J. Parasitol. 95, 224-227.

2. Adema, C.M., Bayne, C.J., Bridger, J.M., Knight, M., Loker, E.S., Yoshino, T.P., Zhang, S.M., 2012. Will all scientists working on snails and the diseases they transmit please stand up? PLoS Negl. Trop. Dis. 6, e1835.

3. Bankers, L., Fields, P., McElroy, K.E., Boore, J.L., Logsdon, J.M. Jr., Neiman, M., 2017. Genomic evidence for population-specific responses to co-evolving parasites in a New Zealand freshwater snail. 26, 3663-3675.

4. Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics 30, 2112–2120.

5. Bowden, R.M., French, S.S., Demas, G.E., 2017. Introduction to ecoimmunology: An integrative approach. J. Exp. Zool. A. Ecol. Integr. Physiol. 327, 219-221.

6. Buddenborg, S.K., Bu, L., Zhang, S.M., Schilkey, F.D., Mkoji, G.M., Loker, E.S., 2017. Transcriptomic responses of *Biomphalaria pfeifferi* to *Schistosoma mansoni*: Investigation of a neglected African snail that supports more *S. mansoni* transmission than any other snail species. PLoS Negl. Trop. Dis. 11 e0005984.

7. Buchfink, B., Xie, C., Huson, D.H., 2015. Fast and sensitive protein alignment using DIAMOND. Nat. Methods 12, 59–60.

8. Carnegie, R.B., Arzul, I., Bushek, D., 2016. Managing marine mollusc diseases in the context of regional and international commerce: policy issues and emerging concerns. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 371.

9. Deleury, E., Dubreuil, G., Elangovan, N., Wajnberg, E., Reichart, J.M., Gourbal, B., Duval, D., Baron, O.L., Gouzy, J., Coustau, C., 2012. Specific versus non-specific immune responses in an invertebrate species evidenced by a comparative *de novo* sequencing study. PLoS One 7 e32512.

10. Demas, G.E., Nelson, R.J., 2012. Ecoimmunology. Oxford University Press, Inc. New York, New York.

11. Dreyfuss, G., Vignoles, P., Abrous, M., Rondelaud, D., 2002. Unusual snail species involved in the transmission of *Fasciola hepatica* in watercress beds in central France. 9, 113-120.

12. Ebbs, E.T., Loker, E.S., Davis, N.E., Flores, V., Veleizan, A., Brant, S.V., 2016. Schistosomes with wings: how host phylogeny and ecology shape the global distribution of *Trichobilharzia querquedulae* (Schistosomatidae). Int. J. Parasitol. 46, 669-677.

13. Ebbs, E.T., Loker, E.S., Brant, S.V., 2018. Phylogeography and genetics of the globally invasive snail *Physa acuta* Draparnaud 1805, and its potential to serve as an intermediate host to larval digenetic trematodes. BMC Evol. Biol. 18, 103.

14. Fuess, L.E., Mann, W.T., Jinks, L.R., Brinkhuis, V., Mydlarz, L.D., 2018. Transcriptional analyses provide new insight into the late-stage immune response of a diseased Caribbean coral. R. Soc. Open Sci. 5, 172062.

15. Ghosh, J., Lun, C.M., Majeske, A.J., Sacchi, S., Schrankel, C.S., Smith, L.C., 2011. Invertebrate immune diversity. Dev. Comp. Immunol. 35, 959–974.

16. 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., Friendman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652.

17. Horner, D.S., Pavesi, G., Castrignanò, T., De Meo, P.D., Liuni, S., Sammeth, M., Picardi, E., Pesole, G., 2010. Bioinformatics approaches for genomics and post genomics applications of next-generation sequencing. Brief Bioinform. 11, 181-197.

18. Johnson, P.T., Olden, J.D., Solomon, C.T., Vander Zanden, M.J., 2009. Interactions among invaders: community and ecosystem effects of multiple invasive species in an experimental aquatic system. Oecologia. 159, 161-170.

19. Kanev, I., 1994. Life-cycle, delimitation and redescription of *Echinostoma revolutum* (Froelich, 1802) (Trematoda: Echinostomatidae). Syst. Parasitol. 28, 124-144.

20. Larken, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947–2948.

21. Lawton, S.P., Allan, F., Hayes, P.M., Smit, N.J., 2018. DNA barcoding of the medically important freshwater snail *Physa acuta* reveals multiple invasion events into Africa. Acta Trop. 188, 86-92.

22. Li, W., Godzik, A., 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22, 1658–1659.

23. Li, B., Dewey, C.N., 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12, 323.

24. Loker, E.S., Adema, C.M., Zhang, S., Kepler, T.B., 2004. Invertebrate immune systems–not homogeneous, not simple, not well understood. Immunol. Rev. 198, 10-24.

25. Loker, E.S., Bayne, C.J., 2018. Molluscan Immunobiology: Challenges in the Anthropocene Epoch, in: Cooper, E.L. (Ed.), Advances in Comparative Immunology. Springer International Publishing AG, Switzerland, pp. 343-407.

26. McCarthy, F.M., Wang, N., Magee, G.B., Nanduri, B., Lawrence, M.L., Camon, E.B., Barrell, D.G., Hill, D.P., Dolan, M.E., Williams, W.P., Luthe, D.S., Bridges, S.M., Burgess, S.C., 2006. AgBase: a functional genomics resource for agriculture. BMC Genomics. 7, 229.

27. Mora, C., Tittensor, D.P., Adl, S., Simpson, A. G. B., Worm, B., 2011. How Many Species Are There on Earth and in the Ocean? PLoS Biol. 9, e1001127.

28. Nolan, J.R., Bergthorsson, U., Adema, C.M., 2014. *Physella acuta*: atypical mitochondrial gene order among panpulmonates (Gastropoda). J. Mollus. Stud. 1–12.

29. Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 26, 139-140.

30. Romero, A., Dios, S., Poisa-Beiro, L., Costa, M.M., Posada, D., Figueras, A., Novoa, B., 2011. Individual sequence variability and functional activities of fibrinogen-related proteins (FREPs) in the Mediterranean mussel (*Mytilus galloprovincialis*) suggest ancient and complex immune recognition models in invertebrates. Dev. Comp. Immunol. 35, 334–344.

31. Schultz, J.H., Adema, C.M., 2017. Comparative immunogenomics of mollusks. Dev. Comp. Immunol. 75, 3–15.

32. Schultz, J.H., Bu, L., Adema, C.M., 2018a. Comparative immunological study of the snail *Physella acuta* (Hygrophila, Pulmonata) reveals shared and unique aspects of gastropod immunobiology. Mol. Immunol. 101, 108-119.

33. Schultz, J.H., Bansbach, L.M., Bremmer, J.A., Dimmler, K.E., Forde, Q.A., Gagliano, E.M., Glenn, E.M., Greengrass, C.M., Hayes, J.P., Kraus, A.L., Larsen, L.I., Lucero, E., McClendon, M.T., Mercer, H.L., Mims, K.C., Patel, K.N., Patsalis, F.I., Peterson, D.E., Platero, J.M., Rizvi, M.M., Serna, K.I., Steele, T.E., Turner, N.E., Bu, L., Lu, L., Adema, C.M., 2018b. The mitochondrial genome of the planorbid snail *Planorbella duryi.* Mitochondrial DNA Part B. 3, 972-973.

34. Shackleton, R.T., Shackleton, C.M., Kull, C.A., 2019. The role of invasive alien species in shaping local livelihoods and human well-being: A review. 229, 145-157.

35. Sharma, R., Yang, Y., Sharma, A., Awasthi, S., Awasthi, Y.C., 2004. Antioxidant role of glutathione S-transferases: protection against oxidant toxicity and regulation of stress-mediated apoptosis. Antioxid. Redox Signal. 6, 289-300.

36. Wang, L., Zhang, H., Wang, M., Zhou, Z., Wang, W., Liu, R., Huang, M., Yang, C., Qiu, L., Song, L., 2019. The transcriptomic expression of pattern recognition receptors: Insight into molecular recognition of various invading pathogens in Oyster *Crassostrea gigas*. Dev. Comp. Immunol. 91, 1-7.

37. Wethington, A.R., Lydeard, C., 2007. A molecular phylogeny of Physidae (Gastropoda: Basommatophora) based on mitochondrial DNA sequences. J. Molluscan Stud. 73, 241–257.

38. Winnepenninckx, B., Backeljau, T., De Wachter, R., 1993. Extraction of high molecular weight DNA from molluscs. Trends Genet. 9, 407.

39. Zhang, L., Li, L., Guo, X., Litman, G.W., Dishaw, L.J., Zhang, G., 2015. Massive expansion and functional divergence of innate immune genes in a protostome. Sci. Rep. 5, 8693.

40. Zhao, J.S., Wang, A.Y., Zhao, H.B., Chen, Y.H., 2017. Transcriptome sequencing and differential gene expression analysis of the schistosome-transmitting snail *Oncomelania hupensis* inhabiting hilly and marshland regions. Sci. Rep. 7, 15809.

Chapter 6: Discussion

Summary and Conclusions

Comparative immunology is not restricted to investigation of any one organism and there is a multitude of cellular and molecular techniques available to investigate animal immunity (Cooper, 2018). Long evolutionary histories have led to the development of an extensive, and as yet unfathomed diversity of immune capabilities among living animals within and across phylogenetic branches. Historical approaches to comparative immunology involved detailed analysis of immune function for representative organisms of large phylogenetic groups. Current approaches utilizing new technological advancements (see below) have demonstrated that animals within the same taxonomic groups can display unique aspects of immunity. It is realized that there are no longer constraints imposed by a lack of available data, but rather constraints on the amount of time required to analyze it. As such, investigators of comparative immunology need to decide where to look (phylogenetically speaking) and how. The incredible breadth of existent biodiversity prevents investigation of immunity for all metazoans and it is necessary to make strategic choices when choosing an animal for study.

Chapter 2

Perhaps concluding with a few select organisms (*Biomphalaria glabrata*), the days of publishing full-scale single genome papers in prestigious journals, like Science and Nature, are numbered (Smith, 2016). This is because next-generation sequencing (NGS) has dramatically "changed the game", such that sequencing entire genomes, transcriptomes, and even epigenomes is no longer limited to traditional model organisms, and small groups or even individual investigators can aquire large datasets (Nolwenn et al., 2014). There are many (even unrelated) projects that generate data to reveal the immune capabilities of animals. To maximize the

benefits of NGS, it is imperative that investigators of comparative immunology target select representative taxa for sequencing that can help interpret the evolution of immunity. Despite the need to critically address and validate the interpretations of immunity gleaned from use of NGS, and the many computational issues associated with this technology, including software that is often tailored for investigations of model organisms (Gene Ontology), the NGS world does and will continue to rule. Choosing organisms of phylogenetic branches depauperate of immune inquiry are likely to reveal unique aspects of immunobiology that fill gaps in our current perspectives of immunity.

Chapter 3

This chapter utilized NGS for investigation of immunity of *Physella acuta*, chosen because this species was representative of a taxon understudied for immunobiology, and therefore, could instruct on the evolution of immunity by enabling a comparison to *B. glabrata*, a well-studied and closely related snail species. Analysis of a reference transcriptome assembled from bacterially-exposed and unexposed *P. acuta* snails demonstrated conserved and unique aspects of gastropod immunity. *Physella acuta* employed a relatively small repertoire of antimicrobial peptides (AMPs) in control and bacterially-elicited transcriptomes, expressing only two sequences of the macin-like family of AMPs. This is evidenced to be a conserved feature among panpulmonate gastropods because the genome of *B. glabrata* also suggests an AMP repertoire limited to macin-like sequences. An extensive toolkit of antimicrobial proteins was expressed by *P. acuta*, a feature similarly shared by *B. glabrata*. However, the prominent use of somatically diversified fibrinogen-related proteins (FREPs) as anti-parasite factors in *B. glabrata* is not a shared feature of gastropod immunity, as demonstrated by investigating the composition of FREP genes and the anti-trematode (*Echinostoma paraenesi*) response of *P. acuta*. This

finding cautions against making general assumptions of immunity for all gastropods based on the findings one (*B. glabrata*) snail species and exemplifies the importance of investigating representative taxa in a comparative immunological fashion. This analysis also showed that *P. acuta* snails can be experimentally infected by the trematode parasite *E. paraenesi*, which supports the notion that *E. paraenesi* is a generalist-type parasite, finding host suitability within and across major families of gastropods.

Chapter 4

Characterization of the recently described host-parasite model, *P. acuta* and *E. paraensei*, presented valuable considerations of next-generation sequencing, gastropod immunity, and parasite influence on host responses. First, NGS is a powerful tool for investigating non-traditional organisms (see Chapter 2), but many of the current bioinformatic software packages, which were mainly developed to annotate sequences of model organisms, are less applicable to study (non-model) invertebrate organisms, like gastropods. This was demonstrated by the limited annotation of differentially expressed (DE) sequences of *P. acuta* after exposure to *E. paraensei* using Gene Ontology (GO) tools. Rather than consider this hindering for study of comparative immunology, it requires us to get to "know" our model organisms in a more detailed manner. BLAST annotation of DE sequences of *P. acuta* yielded an exhaustive list of sequences from which I scanned for immune-relevant transcripts. I discovered expanded families of immune-relevant sequences such as fibrinogen-related domain-containing (FReD) sequences, GTPase IMAP, and C1q-domain containing sequences (among others). This type of analysis is only possible if considered in an comparative framework with other gastropods and mollusc species. *Physella acuta* differentially expressed an extensive diversity of immune-relevant sequences, belonging to different immune-gene families, after exposure to *E.*

paraensei. Furthermore, individual members of these families show variation in expression patterns, such that some sequences may be up or downregulated at 2 and 8DPE to *E. paraensei*. This finding is reminiscent of gene expression patterns identified in *B. glabrata* and *Crassostrea gigas* (Deleury et al., 2012; Zhang et al., 2015). These results led me to postulate the hypothesis that variation in expression of individual snail sequences may be part of an extended phenotype of *E. paraenesi* and helps to interpret how a snail host can survive by maintaining immunoprotection from other pathogens without responding immunologically, and perhaps even supporting, development of a parasite.

Chapter 5

Investigations of invertebrate immunobiology have benefitted extensively from laboratory experiments that manipulate the environmental conditions (type and dosage of pathogens exposed) of laboratory animals to compare responses to unaffected controls. In particular, gastropod immunology has been primarily investigated using laboratory approaches to decipher the immune mechanisms that confer immunoprotection to the lab model *B. glabrata*. We have expanded the scope of gastropod immunology by utilizing the snail *P. acuta* in similar experiments to reveal some of its immune capabilities, mainly through use of NGS. This chapter investigated the ecoimmunology of *P. acuta* snails, applying Illumina RNA-seq to field snails and lab-reared snails uninfected by parasites (trematode and nematode), to uncover inherent differences in gene expression between these two groups. Our results indicate that overall gene expression, as determined by using gene ontology tools, is similar for lab and field collected *P. acuta* snails. However, in depth analysis of genes differentially expressed indicates that differences in gene expression can cluster snails into groups by location (field or lab). A closer look at the expression profiles of immune-relevant DE genes among these two groups reveals

differential use of categories of antimicrobial proteins, such that some categories including BPIs and LAAOs are upregulated in field snails and others (lysozymes and macins) are downregulated. We hypothesize that these differences can be attributed to host-specific responses to pathogen burdens experienced by lab and field snails. A great number of transcripts encoding fibrinogen domain-containing sequences (FReDs) were upregulated in field-collected *P. acuta* snails, supporting the notion of FReD involvement in the immune surveillance responses of *P. acuta*. Despite this difference between field and lab snails regarding expression of FReD sequences, the transcriptomes generated from four individual field snails and one transcriptome generated from lab snails, both contained many expressed FReD sequences. Overall, *P. acuta* snails employ selective use of individual members of immune-relevant gene families, consistent with previous investigations of *P. acuta* immunity (Chapter 4). Continued investigations of the ecoimmunology of *P. acuta* snails will benefit from lab-based research while providing a touchstone of "real-world" biology to refine the understanding of how immune function may shape the success and evolution of snails in the field.

References

1. Cooper, E.L., 2018. Advances in comparative immunology. Springer International Publishing AG, Switzerland.

2. Deleury, E., Dubreuil, G., Elangovan, N., Wajnberg, E., Reichhart, J.M., Gourbal, B., Duval, D., Baron, O.L., Gouzy, J., Coustau, C., 2012. Specific versus non-specific immune responses in an invertebrate species evidenced by a comparative *de novo* sequencing study. PLoS One. 7, e32512.

3. Dheilly, N.M., Adema, C., Raftos, D.A., Gourbal, B., Grunau, C., Du Pasquier, L., 2014. No more non-model species: The promise of next generation sequencing for comparative immunology. Dev. Comp. Immunol. 45, 56-66.

4. Smith, D.R., 2016. Goodbye genome paper, hello genome report: the increasing popularity of 'genome announcements' and their impact on science. Brief. Funct. Genomics pii: elw026.

5. Zhang, L., Li, L., Guo, X., Litman, G.W., Dishaw, L.J., Zhang, G., 2015. Massive expansion and functional divergence of innate immune genes in a protostome. Sci. Rep. 5, 8693.