Recent advances in crustacean genomics

Jonathon H. Stillman,^{1,*} John K. Colbourne,[†] Carol E. Lee,[‡] Nipam H. Patel,^{§,¶} Michelle R. Phillips,^{||} David W. Towle,[#] Brian D. Eads,[†] Greg W. Gelembuik,[‡] Raymond P. Henry,^{**} Eric A. Johnson,^{††} Michael E. Pfrender^{‡‡} and Nora B. Terwilliger^{||}

*The Romberg Tiburon Center and Department of Biology, San Francisco State University, 3150 Paradise Drive, Tiburon, CA 94920, USA; [†]The Center for Genomics and Bioinformatics, Indiana University, 915 East Third Street, Bloomington, IN 47405-7107, USA; [‡]Department of Zoology, 430 Lincoln Drive, Birge Hall, University of Wisconsin, Madison, WI 53706, USA; [§]Department of Molecular Cell Biology; [¶]Department of Integrative Biology and HHMI, UC Berkeley, Berkeley, CA 94720-3140, USA; [∥]Oregon Institute of Marine Biology, Charleston, OR 97420, USA; [#]Mount Desert Island Biological Laboratory, Salsbury Cove, ME 04672, USA; ^{**}Auburn University, Auburn, AL 36849, USA; ^{††}University of Oregon, Institute of Molecular Biology, Eugene, OR 94703, USA; ^{‡‡}Department of Biology, Utah State University, 5305 Old Main Hill Road, Logan, UT 84322-5305, USA

Synopsis Crustaceans are a diverse and ancient group of arthropods that have long been studied as interesting model systems in biology, especially for understanding animal evolution and physiology and for environmentally relevant studies. Like many model systems, advances in DNA-sequencing methodologies have led to a large amount of genomics-related projects. The purpose of this article is to highlight the genome projects and functional genomics (transcriptomics) projects that are currently underway in crustacean biology. Specifically, we have surveyed the amount of publicly available DNA sequence data (both genomic and EST data) across all crustacean taxa for which a significant number of DNA sequences have been generated. Several ongoing projects are presented including the ecology of invasive species, thermal physiology, ion and water balance, ecology and evolutionary biology, and developmental biology.

Introduction

Background

During the past decade, advancements in DNAsequencing technologies that have coupled increased sequencing throughput and quality with decreased cost are revolutionizing the way in which we think about conducting research in biology. Specifically, the advancement in reduction of DNA-sequencing costs has led to adoption of whole genomes for the study of biological problems in areas outside the realm of biomedical sciences. Genomics approaches are appearing across many disciplines of biology, including physiology, ecology and environmental biology, evolutionary biology, and developmental biology. Additionally, the recent advancements in DNA sequencing have the potential to profoundly impact our understanding of basic physiological mechanisms. This potential arises from the fact that an extraordinary amount of biological knowledge is obtained from a surprisingly small number of species. Our understanding of the genetic regulation of metazoan phenotypes is largely based on studies

using laboratory model organisms like yeast (Saccharomyces cerevisiae), fruit fly (Drosophila melanogaster), roundworm (Caenorhabditis elegans), zebrafish (Danio rerio), and mouse (Mus musculus). For biomedical research, the National Institutes of Health primarily foster 12 species whose research communities have been granted early access to draft genome sequences (http://www.nih.gov/science/ models). Modern DNA-sequencing technologies have replaced the once painstaking hunt for genes associated with a particular phenotype with highthroughput scans for all regions of a genome producing a trait of interest. Even in Drosophila research, which benefits from a century of community-built infrastructures for high-throughput biology, genomic data are transforming an elegant study system for classical and molecular genetics into a powerful model system with applications in all areas of biology. Because ever-increasing speed, everdecreasing costs, and development of automated pipelines to handle the resultant data stream, whole genomes can be sequenced in a short period of time, bringing the ability to adopt genomic approaches to

From the symposium "Recent Advances in Crustacean Genomics" presented at the annual meeting of the Society for Integrative and Comparative Biology, January 2–6, 2008, at San Antonio, Texas.

¹E-mail: stillmaj@sfsu.edu

Integrative and Comparative Biology, pp. 1–17 doi:10.1093/icb/icn096

© The Author 2008. Published by Oxford University Press on behalf of the Society for Integrative and Comparative Biology. All rights reserved. For permissions please email: journals.permissions@oxfordjournals.org.

research on non-model organisms that possess characteristics ideally suited for the biological questions posed.

One such group of organisms is Crustacea. Crustaceans are biologically diverse, ancient, globally distributed, and experimentally tractable (VanHook and Patel 2008). Certain crustaceans (e.g., Daphnia) have been the subjects of intense study for hundreds of years; whereas others have just recently become the focus of intense study (e.g., porcelain crabs). At the 2006 Society for Integrative and Comparative Biology (SICB) annual meeting in Orlando, Florida, crustacean biologists using genomic and proteomic approaches met for a symposium titled "Genomic and Proteomic Approaches in Crustacean Biology". In all reported cases, the efforts to develop crustacean genomic resources (e.g., EST libraries and genome sequence data) were works in progress. Most of the presentations made in the 2006 symposium were published in Volume 46 Issue 6 of the journal Integrative and Comparative Biology. During the 2 years following that first gathering, much progress was made on those projects, and additional crustacean genomics projects were initiated. To provide an update on those projects and to report on additional projects to the crustacean research community, a latebreaking symposium was organized at the 2008 SICB meeting in San Antonio, Texas. This article presents a report of this second symposium and a general review of some current crustacean genomics projects that are underway or have been recently completed. However, as this article does not represent the ultimate communication of those projects, most of which are still being prepared for publication, readers are advised that they must look elsewhere for the definitive presentation of each of the crustacean genomics projects summarized subsequently.

Summary of crustacean genomics projects by taxa

Although DNA sequence data are available across a wide range of crustacean classes, there are only complete genome projects presently underway for two groups of crustaceans, the cladocerans *Daphnia pulex and D. magna*, and the amphipod *Jasus slatteryi* (Table 1). For many groups of crustaceans, only mitochondrial DNA sequence data have been generated by projects focusing on the evolution of diverse arthropods (cite). For most crustacean classes, a substantial EST library has been developed in at least one order. (Table 1), yet the sizes of the libraries are not equally distributed among the orders. Six species account for about 86% of the total EST resources for crustaceans: three branchiopod crustaceans [two species of *Daphnia* and the brine shrimp (35%)] and three decapods [the American lobster, the Pacific white shrimp, and the porcelain crab (51%)] (Table 1). In contrast to the disparities in size of the EST libraries, the number of publicly available gene sequences in the nucleotide database (GenBank nt) is spread more evenly among the different crustacean classes (Table 1). However, when gene models from the genomic projects mentioned above are added to GenBank, this even spread will disappear.

Research questions being addressed

Crustacean biology is fascinating and continues to produce insights across a wide array of biological problems. Many species are important surrogate models for understanding certain medical conditions. However, crustacean genetics lags far behind genetics of insects and other model organisms. Genomic data will rapidly help bridge this disparity. In the subsequent sections are highlights of crustacean research projects employing genomic tools. These projects are in areas spanning ecology and evolutionary biology, biology of invasive species, ion and water balance, thermal physiology, and developmental biology.

Ecology and evolutionary biology

Daphnia genomics spark interdisciplinary research: a brief report on the development and progress of an ecological genomics community

In 2001, the Daphnia genome project was launched to create an ideal model system to study how genomes evolve in response to environmental change, and for analyses of genetic constraints on patterns of species' distribution and abundance. The crustacean Daphnia was chosen because of the longstanding interest in its biology, beginning with its earliest description by Swammerdam as the "water flea" nearly 350 years earlier. The sustained interest in Daphnia is partly due to its abundance in freshwater ponds and lakes around the globe, the ease of manipulation in both laboratory and nature and its central role in food webs. Presently, Daphnia is the focal organism for research in multiple fields, including limnology, life history, physiology, nutrition, predation, parasitology, toxicology, phylogeography, and behavior. The result of this diverse interdisciplinary research is a deeper understanding of the ecology of this species. The goal of the Daphnia genomics project is to build upon this solid foundation in the environmental sciences by developing the Daphnia system to the same level of molecular, cellular and developmental understanding

Table 1	Crustacean	genomics	resources,	arranged	taxonomically	by	Class	and	Order,	with	statistics	on	genomic	data	available	as (of
August	2008 in Gen	Bank ^a															

Class	Order (Suborder)	Genus species	Genomic	Archived Data (GenBank)			
(Subclass, Infraclass)	[common name]	(common name)	data types ^b	ESTs ^c	Nucleotide ^d	GEO DataSets ^c	
Branchiopoda (Phyllopoda)	Cladocera (Anomopoda)	Daphnia pulex	G, M, E	152,687	3212	4	
	[water flea]	Daphnia magna	G, E	13,183	197 Mbp ^f	8	
	Notostraca [tadpole shrimp]	Triops cancriformis	Μ	1	639	0	
		Triops longicaudatus	М	1			
Branchiopoda (Sarsostraca)	Anostraca [brine shrimp]	Artemia franciscana	Μ	37,618	410	0	
Cephalocarida	Brachypoda	Hutchinsoniella macracantha	М	0	14	0	
Ostracoda	Myodocopida	Vargula hilgendorfii	М	0	1016	0	
Remipedia	Nectiopoda	Speleonectes tulumenis	М	0	81	0	
Maxillopoda (Copepoda)	Calanoida [copepod]	Eurytemora affinis	E	0	1852	0	
		Calanus finmarchicus	E	10,962		0	
	Siphonostomatoida [copepod]	Lepeophtheirus salmonis	Μ	48,301	1190	0	
	Harpacticoida [copepod]	Tigriopus japonicus	Μ	0	646	0	
		Tigriopus californicus	М	0		0	
Maxillopoda (Thecostraca, Cirripedia)	Pedunculata [gooseneck barnacle]	Pollicipes mitella	Μ	0	618	0	
		Pollicipes polymerus	М	0		0	
	Sessilia [barnacle]	Tetraclita japonica	М	0	2895	0	
		Megabalanus volcano	М	0		0	
Maxillopoda (Branchiura)	Arguloida [sea louse]	Argulus americanus	Μ	0	33	0	
Malacostraca (Eumalacostraca)	Decapoda (Brachyura) [true crab]	Cancer magister (Dungeness crab)	E	1137 ^e	4634	0	
		Carcinus maenas (European green crab)	E	15,558		0	
		Callinectes sapidus (blue crab)	E, M	10,563		14	
		Uca pugilator (fiddler crab)	E	3646		0	
	Decapoda (Anomura) [half crabs]	Petrolisthes cinctipes (porcelain crab)	E	97,806	2193	2	
	Decapoda (Dendrobranchiata) [shrimp]	Litopenaeus vannamei (white shrimp)	E, M	155,411	3515	4	
		Penaeus monodon (tiger prawn)	E, M	8398		2	
	Decapoda (Astacidea) [clawed lobsters]	Homarus americanus (American lobster)	E	29,558	234	5	
	Decapoda (Palinura) [spiny lobsters]	Panulirus japonicus (spiny lobster)	Μ	0	1094	0	
	Amphipoda [amphipods]	Parhyale hawaiensis	G, E	11	3319	0	
		Jasus slatteryi	G	1		0	
		Gammarus pulex	E	12,345		0	
	Euphausiacea [krill]	Euphausia superba	E	1770	169	0	

Table 1 Continued

Class	Order (Suborder)	Genus species	Genomic	Archived Data (GenBank)				
(Subclass, Infraclass)	[common name]	(common name)	data types ^b	ESTs ^c	Nucleotide ^d	GEO DataSets ^c		
Malacostraca (Hoplocaridea)	Stomatopoda [mantis shrimp]	Squilla mantis	М	0	122	0		
		Harpiosquilla harpax	Μ	0		0		
		Squilla empusa	М	0		0		
		Lysiosquillina maculata	Μ	0		0		
		Gonodactylus chiragra	Μ	0		0		
Malacostraca (Phyllocarida)	Leptostraca	Nebalia hessleri		0	70	0		
Total				586,625	27,956	39		

^aGenBank (http://www.ncbi.nlm.nih.gov/Genbank/index.html). Additional sequence data may be available for some crustacean taxa at the Marine Genomics Project (http://www.marinegenomics.org).

^bFor genomic data types, M, mitochondrial genome sequence; G, nuclear genome and/or BAC sequence; E, EST libraries.

^cESTs and GEO DataSets are for individual taxa.

^dNucleotide sequences are for each Order, not for individual component taxa.

^eThe 11,904 *C. magister* clones discussed in this article have not yet been submitted to GenBank.

^fThe first draft assembly of Daphnia pulex genome sequence (197,300,000 bp) is available at wFleaBase (http://wfleabase.org).

as other model species, but with the added advantage of being able to interpret observations in the context of ecological challenges.

Unlike the research communities for biomedical models-which are primarily trained in molecular genetics and thus predisposed to incorporating genomic approaches into their research-Daphnia biologists initially faced the challenges of integrating large genetic data sets into their research (in many cases for the first time) and initiating high-throughput molecular investigations. Reacting to both the promise and hurtle in espousing genomics for an organism amenable to environmental studies, a large community is forming to establish the infrastructures required to link genomic data to the unique biology of Daphnia. Within 5 years of developing data and tools from this first crustacean genome sequence, genomics is rapidly impacting Daphnia research, most profoundly by uniting researchers from multiple disciplines. Although many avenues are possible for building genomic resources for a new model species, the route taken by the Daphnia community is proving to be rewarding in both expected and unforeseen ways.

Why a model system for ecological and evolutionary genomics?

While the relationship between phenotypic variation and the environmental context of organisms has long been a central focus of investigation for ecological and evolutionary biologists, the most significant recent advances have come from studies of laboratory model species coupling genes and their interactions to cellular and organismal traits. These advances are largely a consequence of the emergence of high-throughput genomics in these model species. Research tools derived from draft genome sequences produce comprehensive data sets that are interpreted in light of rich genetic information, which is archived and disseminated through electronic databases. In standard model systems, these key resources are driven by large and established research communities with a deep understanding of molecular and developmental biology. Yet, despite the genetic strengths of these traditional plant and animal systems, little is known about the natural environmental factors that contribute to their genetic and phenotypic evolution. Moreover, there is growing appreciation for the importance of phenotypic plasticity (including the process of acclimation that allows organisms to cope with diverse and changing environmental conditions), the effects of species interactions within communities and the mechanism of adaptive evolution in natural populations. These processes in natural populations complicate the interpretation of observations from genetic studies within laboratory settings that commonly use inbred strains. While phenotypic plasticity has traction in both the theoretical and empirical literature, our understanding of the genetic basis of genomeenvironment interactions remains limited. Major

advances have been made in the study of the genetic basis of adaptive phenotypic variation, but a comparable understanding of plasticity remains elusive. The development of modern genomic approaches and the emergence of appropriate model systems for ecological genomic studies hold the promise of rapidly increasing our understanding in this critical area.

An important application of plasticity studies includes the regulation of toxicants that threaten ecosystem health and possibly human safety. The process and limits of both acclimation and adaptation in conferring greater tolerance to environmental chemicals are poorly understood. Daphnia is the choice sentinel invertebrate for assessing the toxicity of environmental pollutants and plays a critical role in establishing regulatory criteria by governmental agencies for protecting freshwater ecosystems. Yet, current techniques for risk assessment are arguably slow, costly, and provide no information on the mechanism by which chemicals have toxic effects. By consequence, The National Toxicology Program estimates that only 6% of the \sim 80,000 chemicals manufactured in the United States have undergone even rudimentary testing. Furthermore, even fewer studies have extended observations made on the influence of chemicals on individuals within the laboratory to their short-term and long-term effects on natural populations. Environmental managers and regulatory agencies are anxious for the emergence of novel molecular technologies applied to toxicological models, promising data that are more sensitive, more quickly obtained, less expensive, and better able to detect the physiological and population-level effects of manufactured compounds.

Genomics applied to Daphnia has the potential to be a key model for a comparative study of developmental and disease processes in a first non-insect arthropod. Genome sequence data for over 19 insect genera (some of the best research organisms for classical and molecular genetics) are allowing the elucidation of how the evolution of genome structure creates phenotypic diversity and disorders. Crustaceans are the closest relatives to insects, both groups evolving from a shared pancrustacean ancestor. The Daphnia genome sequence polarizes the structural features that vary among insect genomes, by rooting polymorphic traits, such as the expansion and organization of homeotic gene families or the expansion of gene families linked to detecting and defending against parasites. There is a growing body of work related to regulation of developmental genes (Shiga et al. 2002; Sagawa et al. 2005), and the genetic basis and evolutionary ecology of parasites' resistance and

immunity (Little 2002; Little et al. 2002, 2003, 2004; Ebert et al. 2004; Chadwick and Little 2005; Mitchell et al. 2005).

The Daphnia Genomics Consortium success story

The challenges faced during the initial development of the Daphnia system included the lack of genetic tools and the lack of a large research community equipped to use such tools. Rapid progress has been made in these areas by the coordinated activities of the Daphnia Genomics Consortium (DGC, http:// daphnia.cgb.indiana.edu/). The DGC is dedicated to building a distributed set of genomics resources through international collaboration. The Consortium was initially formed to acquire a complete genome sequence for Daphnia plus the associated databases and functional genomic tools to rapidly identify genetic networks of ecological importance. In 2001, the genome-sequencing project required substantial financial investments and the commitment of a specialized team that (1) routinely produces data using hundreds of automated sequencers, (2) unambiguously assembles the millions of sequence reads into long-overlapping segments, (3) combines various sources of information for identifying and annotating genes, and (4) disseminates the results in easily accessible electronic databases. In 2002, the DGC was actively recruiting members from around the world by directly communicating with researchers who could contribute preliminary genomic tools, data and biocomputing expertise, or seek funding for such activities. Yet, success at recruiting the consortium's main body-composed of researchers who could make use of these tools for ecological research-was primarily achieved by extending an open invitation via the consortium's website, its electronic mailing list, and regular scientific meetings. In October, 2002, at the first DGC meeting in Bloomington, Indiana, the structure and goals of the consortium were ultimately defined, and plans were drafted to win broad support from academic and governmental agencies.

In 2003, scientists representing the U.S. Department of Energy's Joint Genome Institute agreed to collaborate with the DGC in generating a genome sequence of *D. pulex*. This event marked the beginning developments of the community infrastructure to analyze and utilize the complete genome sequence data for both basic and applied interdisciplinary research. Early on, it became apparent that, to capitalize on this unique opportunity, efforts by many people on multiple fronts were needed to (1) identify a suitably homozygous natural isolate for sequencing, (2) create and sequence a large and diverse set of cDNA libraries for defining genes (over 100,000 cDNA isolates were sequenced from 37 separate libraries representing expressed genes under 24 distinct ecological conditions and three developmental stages), (3) produce a molecular marker-based genetic map for placing assembled parts of the sequenced genome along chromosomes define their recombinational landscapes, and (4) predict genes de novo using annotation tools and experimentally validate these findings, (5) construct high-throughput functional genomic tools like microarrays and proteomic assays to discover condition-specific patterns of gene co-expression, (6) continuously pull together new results into publically available electronic databases, and (7) coordinate a consortium-wide investigation of the genome by pairing experts in bioinformatics with biologists specializing in Daphnia and arthropods or with experts in ecology and evolution to give biological meaning to the data.

At the last Daphnia genomics meeting in summer 2007, the draft genome sequence assembly plus results from related research was released as early access to the public on two web-based data portals: wFleaBase (Colbourne et al. 2005; http://wfleabase. org) and the Joint Genome Institute (JGI) Daphnia genome annotation portal (http://genome.jgi-psf.org/ Dappu1). This rapid achievement is credited to numerous investigators that form the core group of the DGC. Yet, after 5 years, the consortium surprisingly boasts over 300 researchers from 17 countries (still growing). Many are volunteering their expert interpretation of how the genome sequence reflects crustacean biology, including aquatic lifestyle, modes of development, nutritional needs, or physiological responses to ecological stress. Others who study the molecular evolution and function of particular gene families offer to describe the particularities of Daphnia's gene losses and gains when compared to other sequenced arthropods. This broad interest in the sequence data sparked a series of training sessions on manual gene annotation by telephone conferencing, by hosting local annotation jamborees and by streaming video emphasized by detailed instructions posted on the DGC Collaboration Wiki and the JGI web pages. As a result, the genome database is vastly improved by community curation while simultaneously identifying the initial gene sets of greatest interest to crustacean biologists.

Daphnia genomics resources available through the DGC and Indiana University Center for Genomics and Bioinformatics (IU-CGB) includes both supporting data and functional genomic tools such as cDNA

libraries, EST sequences, and Microarrays. The current D. pulex microarrays include two designs forming a versatile platform for integrated interdisciplinary research. Design-1 arrays contain 10,000 probes (long oligonucleotides; 70-mers) unique and specific to gene transcripts identified from the cDNA-sequencing project. Design-2 arrays, printed in collaboration with Roche NimbleGen, are composed of DNA probes tightly packed by a Digital Micromirror Device (DMD, Texas Instruments) oligonucleotide synthesizer. The DMD is an array of millions of micron-sized mirrors positioned to reflect a UV light source on coordinated positions within the array, and can synthesize millions of userdefined oligonucleotide probes with high yield and accuracy. Using DMD, we print the entire Daphnia genome on a single microarray slide, represented as 4.2 million isothermal oligonucleotides of length 50-70 bases that span the sequenced genome at average intervals of 30 bases. Included are 225,000 probes of random sequences, sharing base compositions equivalent to the Daphnia sequences, plus other negative and spiking controls. The random probes are used to set appropriate thresholds that measure significant hybridization signals over the background. Experiments conducted on this tiling array platform are being used to validate the frozen gene sets of the current genome annotation, improve the predicted gene structures by empirically determining UTRs and intron-exon boundaries, identifying missing upstream exons and alternative transcripts, for proposing gene structure models in regions containing no predicted genes, and to delineate transcriptionally active regions of the genome from intergenic regions.

Daphnia as a model for evolutionary, ecological, and environmental genomics

A growing area of research interest is the relationship between genome function and environment. For example, recent analyses of whole genome sequences show that there are often many more predicted genes than there are genes with known function. One possibility is that the expression, regulation, and function of many genes may be highly context-dependent and only manifest in particular environments (Gasch et al. 2000). The importance of understanding the genetic basis of interactions between genotype and environment is reflected in a renewed interest in phenotypic plasticity and its relationship to adaptive evolution (West-Eberhard 2003; Miner et al. 2005; Pigliucci 2005). This shift in focus is also reflected by an increase in research on the relationship between environmental factors and epistatic interactions among genes, which may make a substantial contribution to variation in complex traits such as disease susceptibility (Carlborg and Haley 2004). Increasingly, QTL studies are focusing on regulatory networks of polymorphic genes based on a combination of QTL analysis and microarray expression profiles (eQTLs) (Bing and Hoeschele 2005; Carlborg et al. 2005; de Koning et al. 2005). The combination of these two methods, referred to as genetical genomics (Jansen and Nap 2001; de Koning and Haley 2005) is a powerful approach for inferring a gene's transcriptional relationships (Li et al. 2005) and has been utilized to demonstrate that regulation of many genes has a heritable basis (Cheung and Spielman 2002; Morley et al. 2004; Hubner et al. 2005). A genetical genomic approach, applied in an organism such as Daphnia for which environmental conditions can be accurately and systematically manipulated, will significantly advance our understanding of the relationship between the phenotype and the underlying genotypic and environmental effects (Eads et al. 2008).

Thermal physiology

Porcelain crabs, genus *Petrolisthes*, as a model crustacean for thermal physiology

Temperature is a critically important environmental factor because of its effects on all biological processes (Hochachka and Somero 2002). Overwhelming evidence suggests that climatic change is occurring and that global temperatures are higher than ever during the past century (Karl and Trenberth 2003; IPCC 2007). Changes in habitat temperature appear to be causing increasing levels of physiological stress in organisms that live close to their physiological limits in nature, as is evidenced by coral bleaching events (Coles and Brown 2003; Hughes et al. 2003) and species shifts correlated with increases in habitat temperature (Barry et al. 1995; Parmesan et al. 1999; Sagarin et al. 1999; Hughes 2000; Crozier 2003; Chevaldonne and Lejeusne 2003).

How natural populations respond to climatic change will depend on both the magnitude of change in the temperature of the habitat, and each species' physiological response to that change. A central challenge for ecological physiologists is to understand the physiological responses of organisms to changes in temperature of the habitat and to use this information to develop predictions of how climatic change will impact species abundance and distribution limits (Helmuth et al. 2002; Clarke 2003; Chevaldonne and Lejeusne 2003). A powerful approach to address this challenge combines organismal physiology and transcriptome profiling to examine integrated physiological responses and high-resolution molecular responses to thermal acclimation and thermal stress. Porcelain crabs (genus Petrolisthes) have proved valuable in the study of physiological adaptation to temperature. Petrolisthes are diverse with over 100 species worldwide and 46 species in the eastern Pacific Ocean (Haig 1960; Stillman and Reeb 2001). In the eastern Pacific, congeneric Petrolisthes species are found from Alaska to southern Chile, and within each biogeographic province individual species live in specific vertical zones, from the subtidal to the upper intertidal (Stillman and Reeb 2001). Across this distributional range, crab body temperatures can vary from sub-zero to over 40°C. Species living in the upper intertidal zone are eurythermal, experiencing and tolerating higher and lower temperatures during low tide than do sympatric subtidal congeners (Stillman 2002).

Our studies of porcelain crab thermal physiology has demonstrated that warm-adapted porcelain crabs live near their thermal limits in nature, and have very low capacities for adjusting their thermal limits as compared to species living in cooler habitats (Stillman 2003). From these results, we predict that small increases in habitat temperature will dramatically increase thermal stress for those species living in the warmest habitats. Why is it that some species can adjust their thermal limits more than other species? How will these species respond to thermal stress? Addressing these questions using a combination of physiology and functional genomics is a powerful approach that is yielding novel insights and framing new directions for elucidation of basic mechanisms setting thermal phenotype.

Mechanistic bases of thermal phenotype

The mechanistic bases that create the great diversity of thermophysiological phenotypes (e.g., thermal limits, thermal breadth, and thermal plasticity) of organisms (Huey and Kingsolver 1993; Feder et al. 2000; Hochachka and Somero 2002), as well as the mechanisms by which thermal phenotypes are adjusted during thermal acclimation remain an area of active study. Single-gene approaches to an understanding of the mechanisms of thermal physiology have focused largely on the heat-shock response (Feder and Hofmann 1999) and have illustrated how expression of hsps correlate with differences in heat-tolerance limits (Tomanek and Somero 1999; Hofmann et al. 2002), and with heat stress (Hofmann and Somero 1995, 1996; Roberts et al. 1997;

Feder et al. 1997; Tomanek 2002). However, multiple gene studies have taught that hsps are not the only genes that respond to environmental stress. As a result of the advances in genome science, we can now examine the expression of thousands of genes simultaneously using DNA microarrays (Brown and Botstein 1999). Microarrays have led to novel insights about mechanisms driving both ultimate (Ferea et al. 1999; Oleksiak et al. 2002) and proximate (Gracey et al. 2001; Gracey and Cossins 2003, Podrabsky and Somero 2004, Teranishi and Stillman 2007) changes in organismal physiology. For example, in yeast, a defined set of genes collectively known as the environmental stress response (ESR) is differentially expressed during environmental stresses (Gasch et al. 2000). The ESR includes induction of hsps, as well as induction of genes involved in energy metabolism and repression of genes involved in growth of cells (Gasch et al. 2000).

To build porcelain crab microarrays, we constructed non-normalized (Stillman et al. 2006) and normalized mixed tissue, life stage, and environmental exposure cDNA libraries for the porcelain crab *P. cinctipes*, a common intertidal inhabitant in the northeastern Pacific. In conjunction with the JGI, we sequenced over 60,000 cloned cDNAs, generating 97,806 high quality ESTs (Genbank nos FE742652– FE840457) that represent about 30K unique consensus sequences and 19K unique clusters. Consensus sequences as well as annotation and microarray data are hosted on the Porcelain Crab Array Database (http://array.sfsu.edu).

Microarray-based studies have great potential for generation of novel hypotheses regarding genes that are important in complex physiological responses to environmental stress. In the majority of these studies, model organisms have been used to examine transcriptome responses to stresses such as temperature, nutritional state, and hypoxia that elicit widespread, general physiological responses that are regulated by a large number of genes (Ferea et al. 1999; Gasch et al. 2000; Kaan et al. 2002; Oleksiak et al. 2002; Becerra et al. 2003). Studies of non-model organisms that have diverse natural histories may reveal biological responses to environmental stresses not observed in model species (Gracey et al. 2001; Gracey and Cossins 2003, Podrabsky and Somero 2004, Teranishi and Stillman 2007; Place et al. 2008).

Profiles of gene expression during thermal acclimation

To identify the mechanisms that alter the limits of cardiac thermal performance during thermal acclimation, and to examine responses to thermal stress, studies are being conducted at both organismal and molecular levels. We acclimated crabs to different thermal regimes, measured heat and cold tolerance limits of cardiac function during exposure to thermally stressful temperatures, and used cDNA microarrays to examine transcriptome profiles in differently acclimated or stressed crabs. Foci of current research are focused on (1) determining the changes in gene expression that accompany changes in cardiac tolerances to heat and cold during thermal acclimation to high and low constant temperatures, (2) characterizing gene expression responses to acute thermal stress in crabs acclimated to different temperatures, (3) characterizing gene expression changes that occur during physiological acclimation to fluctuating thermal regimes, and (4) characterizing seasonal and latitudinal variation in gene expression and response to acute thermal stress (e.g., Teranishi and Stillman 2007) in crabs acclimatized to natural habitat conditions.

Ion and water balance

Effects of salinity on gene expression patterns in gills of the green crab *Carcinus maenas*

The green crab C. maenas, an invasive species originating in Europe and now established in many coastal and estuarine regions around the world, is remarkably tolerant of environmental stresses, including changes in salinity. We determined transcriptional changes associated with hypotonic salinity stress in an effort to further understand the osmoregulatory mechanisms in C. maenas. Microarray analysis utilized a 4462-feature oligonucleotide array derived from 15,637 expressed sequence tags identified in a normalized multiple-tissue cDNA library (Towle and Smith 2006). Two time-course experiments sampled posterior gills at intervals following transfer from full-strength seawater to 10% diluted seawater, a condition under which the crabs effectively hyperosmoregulate their hemolymph (Siebers et al. 1982). Total RNA extracts were reverse transcribed in the presence of Cy3-dCTP or Cy5-dCTP and the labeled cDNAs were hybridized to the arrays, with each experimental sample paired with a control.

Many of the features on the array remain unannotated due to the paucity of functional information for crustacean genomes, and some of these unannotated genes showed marked changes in expression following dilution to lower salinities. Among the features for which reliable annotations are available, Na^+/K^+ -ATPase α subunit and cytoplasmic carbonic anhydrase transcripts were significantly induced, confirming previous evidence obtained by quantitative PCR in C. maenas and other species (Henry et al. 2006; Jayasundara et al. 2007). Other membrane proteins showing increased mRNA expression included y-amino butyric acid and dopamine receptors, sodium/glucose cotransporter, and chloride/bicarbonate exchanger, while aquaporin-encoding transcripts decreased. Among transcripts encoding stress-related proteins, ubiquitin-conjugating enzyme E2 and proteasome α -subunit increased, while heatshock cognate protein 70 decreased, suggesting that a shift in protein metabolic machinery accompanied the response to decreased salinity. A large suite of mitochondrial proteins exhibited increased mRNA expression, including cytochromes b and c, ATP synthase subunits, and cytochrome c oxidase, confirming the previously noted proliferation of mitochondria-rich cells in gills of euryhaline crabs challenged by low salinity (Goodman and Cavey 1990; Luquet et al. 2002). Transiently upregulated transcription factors included MAX binding protein, methyl-CpG binding domain protein 4, and transcription factor IIe. The role of these transcriptional factors remains elusive but provides a rich source of data for designing additional experiments. It appears that the transcriptional response to low salinity in C. maenas is accompanied by a complex pattern of changes that involves much more than the transporters themselves.

Molt cycle

Developmental and osmotically induced changes in gene expression in megalopas and juvenile instars of the Dungeness crab, *Cancer magister*

The ability of a marine organism to tolerate, acclimate, and adapt to its surroundings is an essential component of survival and fitness in the everchanging oceanic environment. The decapod crustacean C. magister, a vital commercial species, inhabits estuaries and nearshore waters along the western coast of North America. Its complex life cycle involves major morphological transformations and multiple habitat changes throughout development, from a pelagic lifestyle in the open ocean as a zoea to a benthic lifestyle in the tidal flats as a juvenile. These relocations, and the dynamic nature of the estuarine environment, expose C. magister to substantial fluctuations in salinity, temperature, and dissolved oxygen concentrations throughout its development. The robust population of C. magister clearly demonstrates the adaptive physiology of this species and its development of the necessary molecular mechanisms for maintaining homeostasis when

faced with environmental challenges. However, while our knowledge of the physiology and development of these animals and how they respond to their environment is rapidly advancing, there is still minimal understanding of how these responses are integrated at the molecular level and the effects of natural biological cycles.

Using a functional genomics approach, we are investigating the transcriptional phenotype of C. magister during molting and in response to shortterm, ecologically relevant periods of osmotic stress. We have created a cDNA library for C. magister that comprises a diverse representation of crabs from different life history and molting stages subjected to a range of environmental conditions, and contains transcripts from whole animals and tissues including gills, hypodermis, heart, hepatopancreas, skeletal muscle, and hemocytes. Additionally, we designed 60-mer to 70-mer oligonucleotides, which specifically represented 11 previously described C. magister genes, including members of the hemocyanin gene family, which play prominent roles in aerobic respiration, molting, and immune responses (Terwilliger et al. 2006). These oligonucleotides and a total of 11,904 randomly chosen clones from the cDNA library were used to construct cDNA microarrays at the University of Oregon Genomics Facility.

Molting is an important and on-going process of physiological change in the life history of all crustaceans, but the contribution of this crucial trait to protein expression and gene expression has been largely ignored in the literature (Shechter et al. 2007). This challenging process requires meticulous control of regulatory proteins and genes in order to form and harden a new exoskeleton, as well as to shed the old one. Using the C. magister microarrays, we characterized global transcriptional patterns that changed as a function of the molt cycle from late premolt megalopa to second instar ecdysis. A cohort of megalopas was collected and raised in the laboratory through the second instar juvenile stage. The molt stage of all individuals was monitored, and four juveniles were collected at the same time each day throughout the molt cycle until a majority of individuals had molted into second instars, a process spanning ~ 14 days.

Our data demonstrate that large suites of genes exhibit coordinated expression at distinct times during the molt cycle. These data, in combination with immunohistological data for specific *C. magister* genes (Terwilliger et al. 2005), indicate that premolt, postmolt, and intermolt periods may begin before recognizable morphological changes take place, the current method for determining stages in the molt cycle. Furthermore, genes that changed as a function of molting were involved in multiple physiological processes, such as immunity, cell-cycle regulation, morphogenesis, protein processing, and synthesis and deposition of the cuticle. Several cuticular protein genes that have been characterized in other crustaceans (Andersen 1999; Faircloth and Shafer 2007) showed marked differences in transcript levels during the molt cycle of C. magister. Many of these genes had highest transcript levels during postmolt or premolt stages, as expected if they are involved in structural or regulatory functions in the crustacean cuticle (Marlowe and Dillaman 1994; Shafer et al. 1994, 1995; Willis 1999). However, others also had peak transcript levels in the middle of the molt cycle, a previously undescribed pattern of expression. This suggests that cuticular protein genes are temporally diverse in expression and may have species-specific functions.

Changes in the expression patterns of cryptocyanin and the six independently regulated hemocyanin subunits during C. magister development have been described (Terwilliger and Ryan 2001; Terwilliger et al. 2005). The oligonucleotides we designed for these genes showed expected patterns of expression based on differences in protein expression and gene expression measured across the molt cycle of C. magister (Terwilliger et al. 2005). Additionally, our data for the hemocyanin subunits confirm established developmental stage-specific differences (Durstewitz and Terwilliger 1997), with subunit 6 present only in adult crabs and subunit 4 having lower expression in first instars. Thus, these genes function as a good internal control on our microarrays and confirm that other patterns of gene expression are likely to be physiologically relevant.

In addition to identifying genes that exhibited increased transcript levels during specific periods of the molt cycle, we were also interested in genes that functioned during environmental challenge, such as high and low salinities, and wanted to ascertain if these genes played similar or different roles at specific times during molting. To examine the transcriptional control of osmoregulation in C. magister, we stressed individual megalopas, first instar juveniles and adult crabs at ambient 32 ppt salinity (control) with acute transfer to 16 ppt, 24 ppt and 35 ppt salinity for 8 h, to approximate conditions that the crabs experience in the field during a tidal cycle (Brown and Terwilliger 1992). Tissue-specific gene expression in adult C. magister was compared to global expression patterns in the other life history stages.

Many of the genes we identified in our molting experiments also responded to exposure to osmotic stress. For instance, several of the cuticular protein genes demonstrated differences in transcript abundance in response to high or low salinities. Notably, these genes also showed diverse tissue expression in adult *C. magister*, an unexpected result since previously they have been described only in hypodermal tissue (Faircloth and Shafer 2007), an epithelium that generates the new exoskeleton during molting. Some of the cuticular protein genes also showed life stage-specific responses to stress.

Our work provides a broad view of the complex transcriptional patterns that take place during ecologically relevant osmotic stress and molting in C. magister, a commercially important crab species. Natural biological cycles such as molting and reproduction are an important area of research in crustaceans and the contribution of these cycles should be considered when exploring the expressions of genes and proteins. At this time, we have identified more than 100 new genes in C. magister, significantly increasing the amount of genetic data available for this organism. Several of these genes have interesting patterns that change during the molt cycle and in response to osmotic stress, providing new opportunities for research and comparative studies. We are currently exploring hypotheses related to specific genes involved in these processes in C. magister and sequencing additional genes from our cDNA library. The comprehensive nature of our cDNA library and exploration of differential transcriptional levels related to molting and environmental stress makes this project unique among crustacean genomics projects in progress elsewhere, and it is our hope that the data generated will be useful for future studies of C. magister and comparisons to on-going work in other crustaceans.

Evolutionary developmental biology

The molecular bases of body plans in amphipod crustaceans

Crustaceans provide an excellent system for investigating the evolution of developmental mechanisms and the molecular changes that have led to morphological evolution. Their relatively close relationship to the genetic model insect system, *D. melanogaster*, and the extreme morphological variation found within the Crustacea contribute to the usefulness of crustaceans for comparative developmental studies. The amphipod *Parhyale hawaiensis* has emerged as a particularly promising system for various "evo-devo" investigations. This amphipod species has proven highly amenable to experimental manipulation, is straightforward to rear in the laboratory, and large numbers of embryos are available year around. A detailed staging system has been developed to characterize the entire embryogenesis of P. hawaiensis, and robust protocols exist for collection and fixation of all embryonic stages, in situ hybridization to study mRNA localization, and immunocytochemistry to study protein localization. Microinjection of blastomeres enables detailed cell-lineage analyses, transient and transgenic (Pavlopoulos and Averof 2005) introduction of recombinant genetic material, and targeted knockdowns of gene function using either RNAi or morpholino approaches.

A number of characteristics of P. hawaiensis have made them particularly amenable to embryological and molecular genetic manipulation. Females produce embryos every two weeks once they reach sexual maturity. Embryogenesis is relatively short, lasting about 10 days at 26°C. Close examination of the embryonic development of P. hawaiensis has produced a very detailed staging system (Browne et al. 2005). Embryogenesis has been divided into 30 discrete stages that are readily identifiable with living animals or by means of common molecular markers on fixed specimens. As direct developers, hatchlings possess a complete complement of segments and appendages morphologically similar to those of adults. Females normally brood the embryos in a ventral brood pouch. Embryos can be rapidly and easily removed from the brood pouch and maintained in seawater. Eggs collected can be hatched individually and the mature animals can subsequently be used in pair-wise sister-brother or mother-son matings to generate inbred lines. Fertilized eggs can be removed from females prior to their first cleavage and are sufficiently large to perform microinjections (Gerberding et al. 2002) and isolate blastomeres (Extavour 2005) with relative ease. In addition, the embryos from an individual female develop synchronously. Developing P. hawaiensis embryos are optically clear, allowing for both detailed microscopic analyses in situ and the use of fluorescently tagged tracer molecules in live embryos. The yolk, while opaque, is sequestered early in development to the center of the developing egg and then later to the developing midgut of the embryo. In addition, early cleavage is holoblastic (total), allowing the fates of individual early cells to be explored through experimental manipulation (Gerberding et al. 2002; Extavour 2005).

The genome of *P. hawaiensis* is estimated to be 3.6 Gb in size (A. Aboobaker and N. Patel, submitted for publication). The JGI has undertaken both EST and directed BAC sequencing of *P. hawaiensis*, and both of these sequencing projects are nearing completion. In addition, the related amphipod crustacean, *Jassa slatteryi*, whose genome size is 690 Mb, has been selected for $5 \times$ genome sequencing, paired end reads from 100,000 cDNA clones, as well as $6 \times$ BAC end sequencing.

EST sequencing

Using a collection of P. hawaiensis embryos spanning the first two-thirds of embryogenesis (0–156 h), a normalized total embryonic cDNA library with an average insert size of 2.2 kb was constructed for EST sequencing. A total of 29,566 cDNA clones were end sequenced using traditional methods of Sanger sequencing. Ninety percent of the sequences passed QC to cluster, yielding a total of 47,732 unique ESTs with a mean read length of 670 bp. These ESTs were parsed into 13,366 clusters of highly related sequence. The resulting cluster distribution suggests that the source cDNA library successfully normalized gene coverage of the embryonic transcriptome: 68% of the total clusters contained sequence from only a single clone and only 66% of passing clones formed clusters with greater than one clone per cluster. A searchable, web-based, database of the P. hawaiensis EST sequences is being developed at JGI, which will provide the scientific community ready access to these data. To supplement these EST data, JGI will perform an additional single 454-FLX run on sheared cDNA to generate roughly 80 MB of data in ~200 bp reads (400,000-600,000 clones). JGI will cluster the 454 ESTs with the ESTs already run on the Sanger platform.

Directed BAC sequencing

A *P. hawaiensis* BAC library (F. Poulin and N. Patel, in collaboration with the C. Amemiya laboratory, submitted for publication) was constructed from nuclear DNA isolated from a isofemale line "iso2" that was established about 6 years ago. With an average insert size of 120–150 Kb, the library was designed to cover the entire genome ~5 times. As the size of the *Parhyale* genome is estimated to be 3.6 Gb, an array of 129,024 clones was plated onto seven high-density filters. Screens with single copy genes confirmed coverage of $5-7\times$.

Probes made from cDNA sequences from roughly 60 developmentally important genes were used to screen the BAC library filters, using standard techniques for radioactive hybridization. Positive BAC clones for each probe were sized, screened by PCR, and restriction mapped to ensure regions of interest were centered before being sent for sequencing. The Stanford Human Genome Center is in the process of sequencing 70 BAC clones.

Related species

Jassa slatteryi and P. hawaiensis are relatively closely related amphipods-from cDNA sequence comparisons, we estimate that roughly the same molecular distance separates the two amphipods as separates D. melanogaster and D. virilis. The embryonic development of J. slatteryi and P. hawaiensis appear largely identical; if not for the smaller size of early I. slatteryi embryos, they would be difficult to distinguish morphologically from P. hawaiensis embryos. What does distinguish Jassa is the relatively small size of its genome. At 690 Mb, the Jassa genome is only 10 Mb larger than the smallest recorded genome within the malacostracan crustaceans. Unlike P. hawaiensis, however, J. slatteryi has proven difficult to culture in the lab. Obtaining wildcaught Jassa adults is reasonably straightforward, as they live along the northern Californian coastline and in San Francisco Bay. Efforts are underway to develop an inbred laboratory population of J. slatteryi. Jassa slatteryi is an approved sequencing target of the National Human Genome Research Institute's (NHGRI) Large-Scale Genome Sequencing Program (http://www.genome.gov/10002154 and http://www.genome.gov/Pages/Research/Sequencing/ SeqProposals/EcdysozoaProposalFinalPDF.pdf). The J. slatteryi genome is scheduled to be sequenced with $5 \times$ coverage by the Broad Institute at MIT. In addition, J. slatteryi BAC and cDNA libraries are being constructed.

Invasive species biology

Genomic targets of selection during habitat invasions in copepods

The copepod *Eurytemora affinis* offers an exceptional model for exploring mechanisms of niche evolution. This species is a major grazer of algae and is a food source for important commercial fish, such as salmon. Within the past few decades, *E. affinis* has independently invaded freshwater habitats multiple times through human activity (Lee 1999). These invasions have implications for the transmission of disease, as *E. affinis* hosts a variety of waterborne pathogens, such as *Vibrio cholerae* (Huq et al. 1983; Heidelberg et al. 2002). With its small size (1.5 mm), short generation time (10–20 days) (Heinle and Flemer 1975; Lee 2003), and small genome size

(Rasch et al. 2004), *E. affinis* has the potential to become a model system to address a variety of problems in aquatic systems.

During invasion of freshwater habitats, *E. affinis* has rapidly evolved key physiological and fitnessrelated traits, including dramatic evolutionary shifts in physiological tolerance and life history traits (Lee 2003; Lee et al. 2007). Freshwater populations showed increased tolerance to freshwater and reduced tolerance to saltwater relative to their saline progenitors, with negative genetic correlations indicating tradeoffs between freshwater and saltwater tolerances (Lee et al. 2003; Lee 2007). These shifts appear to have arisen through selection of alternative genotypes within saline source populations (Lee et al. 2003) rather than through acclimation (Lee 2003; Lee and Petersen 2002).

Custom cDNA microarrays printed with E. affinis subtracted EST libraries (using representational difference analysis), allowed analysis of a small number of functionally important genes. We have sequenced approximately 3000 unique ESTs, focusing mainly on those that showed patterns of parallel differential gene expression across independent invasions of freshwater habitats. We found altered expression of genes in the freshwater populations across a broad range of functional classes, including those involved in ionic regulation. For example, expression of V-type H⁺ ATPase was significantly upregulated in the freshwater populations relative to saline populations at 0 PSU (freshwater). In contrast, most ion uptake genes (Na⁺,K⁺-ATPase, Ca⁺²-ATPase, etc.) were downregulated in the freshwater populations, relative to saline populations under freshwater conditions. These results were consistent with our assays of V-type H⁺ ATPase and Na⁺,K⁺-ATPase activity. Under freshwater conditions, V-type H⁺ ATPase is hypothesized to be crucial for uptake of ions from a dilute environment against a concentration gradient, while Na, K-ATPase might be rate limiting under saline conditions (Wieczorek et al. 1999; Tsai and Lin 2007). We also found evolutionary shifts in expression of genes putatively involved in integument permeability, including several cuticularproteins. One particular gene was homologous to a cuticular protein that is expressed during the larval molt cycle (Mathelin et al. 1998) and serves as a precursor for an osmoregulatory hormone (Antidiuretic Factor A) (Eigenheer et al. 2002). Our analysis suggests that, as a class, genes involved in extracellular protein secretion are significantly upregulated in the freshwater populations, potentially reflecting selection for enhanced deposition of cuticle.

Analyses of cDNA microarrays are a first step toward exploring genes that might be implicated in adaptation. However, only a subset of the differentially expressed genes would constitute actual targets of selection. Many of the observed shifts in expression would have resulted from trans-regulatory changes, representing changes in DNA sequence elsewhere in the genome (e.g. altered activity of a transcription factor). Only cis-regulatory changes in expression located at the gene (e.g. promoters) would represent potential evidence of selection directly targeting expression of a given gene. In addition, analysis of gene expression would not detect many types of structural changes in the gene, such as substitutions of amino acids.

Thus, in order to determine causal mechanisms of freshwater adaptation during invasions, we are adopting a dual approach of genomic scans for selection integrated with QTL mapping. Our goal is to identify specific genetic targets of selection, focusing on evolutionary shifts that are due to cisregulatory changes. QTL mapping links traits of interest (freshwater survival, integument permeability) with regions in the genome, while scans for selective sweeps reveal whether candidate genes contain genetic signatures of selection (Nielsen 2005). This dual approach will allow us to determine whether genes linked to physiological traits also show signatures of selection. As selection will tend to act most strongly on positions in the genome that underlie traits of functional importance (Nielsen 2005), probing for signatures of selection could reveal the genes that are critical for survival in freshwater. We are currently sequencing the differentially expressed and other candidate genes in ancestral saline and invading freshwater populations for multiple clades to detect genetic signatures of selection. Our preliminary results suggest potential signatures of selection at some of the candidate genes, such as cuticular proteins, and Enhancer of Yellow.

The short generation time of *E. affinis* facilitates the development of inbred lines, which is an important resource for genetic model organisms. The inbred lines we have developed for QTL mapping provide powerful tools that could be used for a myriad of genomic studies, including SNP discovery (Marth et al. 1999), pyrosequencing to distinguish between cis- versus trans-regulatory evolutionary changes in gene expression (Wittkopp et al. 2004), and studies of dominance of traits (Wills 1975).

Uncovering genomic mechanisms of invasions will offer insights into constraints on species'

distributions and limits to phenotypic evolution, and whether shared genetic mechanisms govern the propensity to invade. Insights we gain on genetic mechanisms might allow predictions of (1) which populations will invade and (2) the extent of their potential expansions of range. Our studies have relevance for the numerous brackish water invaders into the Great Lakes and inland waters, as osmoregulation is a common challenge for many of these invaders (Taylor and Harris 1986; Dietz et al. 1996; Lee and Bell 1999). The approaches used here are both mechanistic and synthetic with the goal of offering insights into some of the most intractable problems facing contemporary ecosystems.

Summary

Recent advances in DNA sequencing capabilities and reduction in costs have resulted in the ability of individual researchers to generate an equivalent amount of sequence data in one day as 5 years ago were generated by large sequencing centers (such as the Joint Genome Institute) over the span of months and requiring a large team of personnel. These advances in sequencing technology and bioinformatics approaches to assembly now make feasible high-throughput genome projects targeting a larger diversity of crustaceans. Generating multiple genome sequences from a phenotypically diverse group of phylogenetically resolved species will allow the systematic characterization of functional elements of genomes that are most relevant to an organism's ability to respond to ecological and anthropogenic changes. A larger representation of genome sequences in crustaceans will significantly increase our ability to define and understand the evolution of regulatory elements and will open a path for evaluating adaptive evolution, by developing a whole genome perspective of the rates of divergence and the levels of conservation in functional elements.

Many crustacean draft genome sequences are expected from this miniaturization and proliferation of genome-sequencing facilities. However, because of broad implications for the comparative study of arthropod genome biology and because of important rewards obtained from successfully building a research community of diverse scientists around the first project, our plan is to continue engaging experts from a variety of disciplines to collaborate in annotating and interpreting new genome data. With greater DNA-sequencing capacity there comes a greater need for community-level coordination for obtaining and interpreting the data. As such, this symposium sought to engage crustacean biologists to be a part of such a coordinated community of scientists.

Acknowledgments

We thank the National Science Foundation (IOS-0758513 to J.H.S.) and the Society of Integrative and Comparative Biology, Division of Comparative Physiology and Biochemistry (SICB-DCPB) for their support of this symposium at the 2008 SICB meeting in San Antonio Texas.

Funding

National Science Foundation (ESS-0533920 to J.H.S., IOB-543860 to D.W.T., and IBN-0230005 to R.P.H.); National Institutes of Health (P20 RR016463), and Department of Energy Joint Genome Institute (CSP, 05-SE-14 to J.H.S.) grants.

References

- Andersen SO. 1999. Exoskeletal proteins from the crab, Cancer pagarus. Comp Biol Physiol A 123:203–11.
- Barry JP, Baxter CH, Sagarin RD, Gilman SE. 1995. Climaterelated, long-term faunal changes in a california rocky intertidal community. Science 267:672–5.
- Becerra M, Lombardia LJ, Gonzalez-Siso MI, Rodriguez-Belmonte E, Hauser NC, Cerdan ME. 2003. Genome-wide analysis of the yeast transcriptome upon heat and cold shock. Comp Funct Genomics 4:366–75.
- Bing N, Hoeschele I. 2005. Genetical genomics analysis of a yeast segregant population for transcription network inference. Genetics 170:533–42.
- Brown AC, Terwilliger NB. 1992. Developmental changes in ionic and osmotic regulation in the Dungeness crab, *Cancer magister*. Biol Bull 182:270–7.
- Brown PO, Botstein D. 1999. Exploring the new world of the genome with DNA microarrays. Nat Genet 21:33–7.
- Browne WE, Price AL, Gerberding M, Patel NH. 2005. Stages of Embryonic development in the Amphipod Crustacean, *Parhyale hawaiensis*. Genesis 42:124–49.
- Carlborg O, Haley CS. 2004. Epistasis: too often neglected in complex trait studies? Nat Rev Genet 5:618–25.
- Carlborg O, de Koning DJ, Manly KF, Chesler E, Williams RW, Haley CS. 2005. Methodological aspects of the genetic dissection of gene expression. Bioinformatics 21:2383–93.
- Chadwick W, Little TJ. 2005. A parasite-mediated life-history shift in *Daphnia magna*. Proc R Soc Lond B Biol Sci 272:505–9.
- Cheung VG, Spielman RS. 2002. The genetics of variation in gene expression. Nat Genet 32:522–55.
- Chevaldonne P, Lejeusne C. 2003. Regional warming-induced pecies shift in north-west Mediterranean marine caves. Ecol Lett 6:371–9.

- Clarke A. 2003. Costs and consequences of evolutionary temperature adaptation. Trends Ecol Evol 18:573–81.
- Colbourne JK, Singan VR, Gilbert D. 2005. wFleaBase: the *Daphnia* genome database. BMC Bioinformatics 6:45.
- Coles SL, Brown BE. 2003. Coral bleaching capacity for acclimatization and adaptation. Adv Mar Biol 46: 183–223.
- Crozier L. 2003. Winter warming facilitates range expansion: cold tolerance of the butterfly *Atalopedes campestris*. Oecologia 135:648–56.
- de Koning DJ, Haley CS. 2005. Genetical Genomics in humans and model organisms. Trends Genet 21:377–81.
- de Koning DJ, Carlborg O, Haley CS. 2005. The genetic dissection of immune response using gene-expression studies and genome mapping. Vet Immunol Immunopathol 105:343–52.
- Dietz TH, Wilcox SJ, Byrne RA, Lynn JW, Silverman H. 1996. Osmotic and ionic regulation of North American zebra mussels (*Dreissena polymorpha*). Am Zool 36:364–72.
- Durstewitz G, Terwilliger NB. 1997. Developmental changes in hemocyanin expression in the Dungeness crab, *Cancer magister*. J Biol Chem 272:4347–50.
- Eads BD, Andrews J, Colbourne JK. 2008. Ecological genomics in *Daphnia*: stress responses and environmental sex determination. Heredity 100:184–90.
- Ebert D, Carius HJ, Little T, Decaestecker E. 2004. The evolution of virulence when parasites cause host castration and gigantism. Am Nat 164:S19–32.
- Eigenheer RA, Nicolson SW, Schegg KM, Hull JJ, Schooley DA. 2002. Identification of a potent antidiuretic factor acting on beetle Malpighian tubules. Proc Natl Acad Sci USA 99:84–9.
- Extavour CG. 2005. The fate of isolated blastomeres with respect to germ cell formation in the amphipod crustacean *Parhyale hawaiensis*. Dev Biol 277:387–402.
- Faircloth LM, Shafer TH. 2007. Differential expression of eight transcripts and their roles in the cuticle of the blue crab, *Callinectes sapidus*. Comp Biol Physiol B 146:370–83.
- Feder ME, Hofmann GE. 1999. Heat-shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. Ann Rev Physiol 61:243–82.
- Feder ME, Blair N, Figueras H. 1997. Natural thermal stress and heat-shock protein expression in Drosophila larvae and pupae. Funct Ecol 11:90–100.
- Ferea TL, Botstein D, Brown PO, Rosenzweig RF. 1999. Systematic changes in gene expression patterns following adaptive evolution in yeast. Proc Natl Acad Sci USA 96:9721–6.
- Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, Brown PO. 2000. Genomic expression programs in the response of yeast cells to environmental changes. Mol Biol Cell 11: 4241–57.
- Gerberding M, Browne WE, Patel NH. 2002. Cell lineage analysis of the amphipod crustacean *Parhyale hawaiensis*

reveals an early restriction of cell fates. Development 129:5789-801.

- Goodman SH, Cavey MJ. 1990. Organization of a phyllobranchiate gill from the green shore crab *Carcinus maenas* (Crustacea, Decapoda). Cell Tiss Res 260:495–505.
- Gracey AY, Cossins AR. 2003. Application of microarray technology in environmental and comparative physiology. Ann Rev Physiol 65:231–59.
- Gracey AY, Troll JV, Somero GN. 2001. Hypoxia-induced gene expression profiling in the euryoxic fish *Gillichthys mirabilis*. Proc Natl Acad Sci USA 98:1993–8.
- Haig J. 1960. The porcellanidae (Crustacea Anomura) of the Eastern Pacific. Los Angeles, California: University of Southern California Press.
- Heidelberg JF, Heidelberg KB, Colwell RR. 2002. Bacteria of the γ -subclass Proteobacteria associated with zooplankton in Chesapeake Bay. Appl Environ Microbiol 68:5498–507.
- Heinle DR, Flemer DA. 1975. Carbon requirements of a population of the estuarine copepod *Eurytemora affinis*. Mar Biol 31:235–47.
- Helmuth B, Harley CDG, Halpin PM, O'Donnell M, Hofmann GE, Blanchette CA. 2002. Climate change and latitudinal patterns of intertidal thermal stress. Science 298:1015–7.
- Henry RP, Thomason KL, Towle DW. 2006. Quantitative changes in branchial carbonic anhydrase activity and expression in the euryhaline green crab, *Carcinus maenas*, in response to low salinity exposure. J Exp Zool A 305:842–50.
- Hochachka PW, Somero GN. 2002. Biochemical adaptation: mechanism and process in physiological evolution. New York: Oxford University Press.
- Hofmann GE, Buckley BA, Place SP, Zippay ML. 2002. Molecular chaperones in ectothermic marine animals: Biochemical function and gene expression. Int Comp Biol 42:808–14.
- Hubner N, et al. 2005. Integrating transcriptional profiling and linkage analysis for identification of genes underlying disease. Nat Genet 37:243–53.
- Huey RB, Kingsolver JG. 1993. Evolution of resistance to high-temperature in ectotherms. Am Nat 142:S21–46.
- Hughes L. 2000. Biological consequences of global warming: in the signal already apparent? Trends Ecol Evol 15:56–61.
- Hughes TP, et al. 2003. Climate change, human impacts, and the resilience of coral reefs. Science 301:929–33.
- Huq A, Small EB, West PA, Huq MI, Rahman R, Colwell RR. 1983. Ecological relationships between *Vibrio cholerae* and planktonic crustacean copepods. Appl Environ Microbiol 45:275–83.
- Jansen RC, Nap JP. 2001. Genetical genomics: the added value from segregation. Trends Genet 17:388–91.
- IPCC (Intergovernmental Panel on Climate Change). 2007. (http://www.ipcc.ch/pdf/assessment-report/ar4/syr/ar4_syr.pdf).

- Jayasundara N, Towle DW, Weihrauch D, Spanings-Pierrot C. 2007. Gill-specific transcriptional regulation of Na⁺/K⁺ -ATPase alpha-subunit in the euryhaline shore crab *Pachygrapsus marmoratus*: sequence variants and promoter structure. J Exp Biol 210:2070–81.
- Kaan T, Homuth G, Mader U, Bandow J, Schweder T. 2002. Genome-wide transcriptional profiling of the *Bacillus subtilis* cold-shock response. Microbiology 148:3441–55.
- Karl TR, Trenberth KE. 2003. Modern global climate change. Science 302:1719–23.
- Lee CE. 1999. Rapid and repeated invasions of fresh water by the saltwater copepod *Eurytemora affinis*. Evolution 53:1423–34.
- Lee CE. 2000. Global phylogeography of a cryptic copepod species complex and reproductive isolation between genetically proximate "populations". Evolution 54:2014–27.
- Lee CE. 2003. Effects of developmental acclimation on adult salinity tolerance in the freshwater-invading copepod *Eurytemora affinis*. Physiol Biochem Zool 76:296–301.
- Lee CE, Frost BW. 2002. Morphological Stasis in the *Eurytemora affinis* species complex (Copepoda: Temoridae). Hydrobiologia 480:111–28.
- Lee CE, Petersen CH. 2002. Genotype-by-environment interaction for salinity tolerance in the freshwater invading copepod *Eurytemora affinis*. Physiol Biochem Zool 75:335–44.
- Lee CE, Bell MA. 1999. Causes and consequences of recent freshwater invasions by saltwater animals. Trends Ecol Evol 14:284–8.
- Lee CE, Remfert JL, Gelembiuk GW. 2003. Evolution of physiological tolerance and performance during freshwater invasions. Int Comp Biol 43:439–49.
- Lee CE, Remfert JL, Chang Y-M. 2007. Response to selection and evolvability of invasive populations. Genetica 129: 179–92.
- Li H, Lu L, Manly KF, Chesler EJ, Bao L, Wang J, Zhou M, Williams RW, Cui Y. 2005. Inferring gene transcriptional modulatory relations: a genetical genomics approach. Hum Mol Genet 14:1119–25.
- Little TJ. 2002. The evolutionary significance of parasitism: do parasite-driven genetic dynamics occur ex silico? J Evol Biol 15:1–9.
- Little TJ, O'Connor B, Colegrave N, Watt K, Read AF. 2003. Maternal transfer of strain-specific immunity in an invertebrate. Curr Biol 13:489–92.
- Little TJ, Carius HJ, Sakwinska O, Ebert D. 2002. Competitiveness and life-history characteristics of *Daphnia* with respect to susceptibility to a bacterial pathogen. J Evol Biol 15:796–802.
- Little TJ, Colbourne JK, Crease TJ. 2004. Molecular evolution of *Daphnia* immunity genes: polymorphism in a gram-negative binding protein gene and an alpha-2macroglobulin gene. J Mol Evol 59:498–506.
- Luquet CM, Genovese G, Rosa GA, Pellerano GN. 2002. Ultrastructural changes in the gill epithelium of the crab *Chasmagnathus granulatus* (Decapoda: Grapsidae) in diluted and concentrated seawater. Mar Biol 141:753–60.

- Marlowe RL, Dillaman RM. 1994. Lectin binding by crustacean cuticle: the cuticle of *Callinectes sapidus* throughout the molt cycle, and the intermolt cuticle of *Procambarus clarkii* and *Ocypode quadrata*. J Crust Biol 14:231–46.
- Marth GT, Korf I, Yandell MD, Yeh RT, Gu Z, Zakeri H, Stitziel NO, et al. 1999. A general approach to single-nucleotide polymorphism discovery. Nat Genet 23:452–6.
- Mathelin J, Quennedey B, Bouhin H, Delachambre J. 1998. Characterization of two new cuticular genes specifically expressed during the post-ecdysial molting period in *Tenebrio molitor*. Gene 211:351–9.
- Mitchell SE, Rogers ES, Little TJ, Read AF. 2005. Hostparasite and genotype-by-environment interactions: temperature modifies potential for selection by a sterilizing pathogen. Evolution 59:70–80.
- Morley M, Molony CM, Weber TM, Delvin JL, Ewens KG, Spielman RS, Cheung VG. 2004. Genetic analysis of genome-wide expression variation in human gene expression. Nature 430:743–7.
- Nielsen R. 2005. Molecular signatures of natural selection. Ann Rev Genet 39:197–218.
- Oleksiak MF, Churchill GA, Crawford DL. 2002. Variation in gene expression within and among natural populations. Nat Genet 32:261–6.
- Parmesan C, et al. 1999. Poleward shifts in geographical ranges of butterfly species associated with regional warming. Nature 399:579–83.
- Pavlopoulos A, Averof M. 2005. Establishing genetic transformation for comparative developmental studies in the crustacean Parhyale hawaiensis. Proc Natl Acad Sci USA 102:7888–93.
- Pigliucci M. 2005. Evolution of phenotypic plasticity: where are we going now? Trends Ecol Evol 20:481–6.
- Place SP, O'Donnell MJ, Hofmann GE. 2008. Gene expression in the intertidal mussel *Mytilus californianus*: physiological response to environmental factors across a biogeographic scale. Mar Ecol Prog Ser 356:1–14.
- Podrabsky JE, Somero GN. 2004. Changes in gene expression associated with acclimation to constant temperatures and fluctuating daily temperatures in an annual killifish *Austrofundulus limnaeus*. J Exp Biol 207:2237–54.
- Rasch EM, Lee CE, Wyngaard GA. 2004. DNA-feulgen cytophotometric determination of genome size for the freshwater-invading copepod *Eurytemora affinis*. Genome 47:559–64.
- Roberts DA, Hofmann GE, Somero GN. 1997. Heat-shock protein expression in *Mytilus californianus*: acclimatization (seasonal and tidal-height comparisons) and acclimation effects. Biol Bull 192:309–20.
- Sagarin RD, Barry JP, Gilman SE, Baxter CH. 1999. Climaterelated change in an intertidal community over short and long time scales. Ecol Monogr 69:465–90.
- Sagawa K, Yamagata H, Shiga Y. 2005. Exploring embryonic germ line development in the water flea, *Daphnia magna*, by zinc-finger-containing VASA as a marker. Gene Expr Patterns 5:699–678.

- Shafer TH, Roer RD, Midgette-Luther C, Brookins TA. 1995. Postedysial cuticle alteration in the blue crab, *Callinectes sapidus*: synchronous changes in lycoprotein's and mineral nucleation. J Exp Zool 271:171–82.
- Shafer TH, Roe RD, Miller CG, Dillaman RM. 1994. Postecdysial changes in the protein and glycoprotein composition of the cuticle of the blue crab, *Callinectes sapidus*. J Crust Biol 14:210–9.
- Shechter A, Tom M, Yudkovski Y, Weil S, Chang SA, Chang ES, Chalifa-Caspi V, Berman A, Sagi A. 2007. Search for hepatopancreatic ecdysteroid-responsive genes during the crayfish molt cycle: from a single gene to multigenicity. J Exp Biol 210:3525–37.
- Shiga Y, Yasumoto R, Yamagata H, Hayashi S. 2002. Evolving role of antennapedia protein in arthropod limb patterning. Development 129:3555–61.
- Siebers D, Leweck K, Markus H, Winkler A. 1982. Sodium regulation in the shore crab *Carcinus maenas* as related to ambient salinity. Mar Biol 69:37–43.
- Stillman JH. 2000. Causes and consequences of thermal tolerance limits in rocky intertidal porcelain crabs, genus *Petrolisthes*. Int Comp Biol 42:790–6.
- Stillman JH. 2003. Acclimation capacity underlies susceptibility to climate change. Science 301:65.
- Stillman JH, Reeb CA. 2001. Molecular phylogeny of Eastern Pacific porcelain crabs, genera *Petrolisthes* and *Pachycheles*, based on the mtDNA 16S rDNA sequence: phylogeographic and systematic implications. Mol Phylog Evol 19:236–45.
- Stillman JH, Teranishi KS, Tagmount A, Lindquist EA, Brokstein PB. 2006. Construction and characterization of EST libraries from the porcelain crab, *Petrolisthes cinctipes*. Int Comp Biol 46:919–30.
- Taylor PM, Harris RR. 1986. Osmoregulation in *Corophium curvispinum* (Crustacea: Amphipoda), a recent coloniser of freshwater I. Sodium ion regulation. J Comp Physiol B 156:323–9.
- Teranishi KS, Stillman JH. 2007. A cDNA microarray analysis of the response to heat stress in hepatopancreas tissue of the porcelain crab *Petrolisthes cinctipes*. Comp Biochem Physiol D 2:53–62.
- Terwilliger NB, Ryan MC, Phillips MR. 2006. Crustacean hemocyanin gene family and microarray studies of expression change during eco-physiological stress. Int Comp Biol 46:991–9.
- Terwilliger NB, Ryan MC, Towle D. 2005. Evolution of novel functions: cryptocyanin helps build new exoskeleton in *Cancer magister*. J Exp Biol 208:2467–74.
- Terwilliger NB, Ryan MC. 2001. Ontogeny of crustacean respiratory proteins. Amer Zool 41:1057–67.
- Tomanek L, Somero GN. 1999. Evolutionary and acclimationinduced variation in the heat-shock responses of congeneric marine snails (genus *Tegula*) from different thermal habitats: implications for limits of thermotolerance and biogeography. J Exp Biol 202:2925–36.
- Towle DW, Smith CM. 2006. Gene discovery in *Carcinus maenas* and *Homarus americanus* via expressed sequence tags. Integ Comp Biol 46:912–8.

- Tsai J-R, Lin H-C. 2007. V-type H+-ATPase and Na⁺,K⁺-ATPase in the gills of 13 euryhaline crabs during salinity acclimation. J Exp Biol 210:620–7.
- VanHook A, Patel NH. 2008. Crustaceans. Curr Biol 18:R547–50.
- West-Eberhard MJ. 2003. Developmental plasticity and evolution. Oxford: Oxford University Press.
- Wieczorek H, Brown D, Grinstein S, Ehrenfeld J, Harvey WR. 1999. Animal plasma membrane energization by protonmotive V-ATPases. Bioessays 21:637–48.
- Willis J. 1999. Cuticular proteins in insects and crustaceans. Amer Zool 39:600–9.
- Wills C. 1975. Marginal overdominance in *Drosophila*. Genetics 81:177–89.
- Wittkopp PJ, Haerum BK, Clark AG. 2004. Evolutionary changes in cis and trans gene regulation. Nature 430:85–8.