Evolutionary Mechanisms of Rapid Adaptation During Freshwater Invasions by the Saline Copepod *Eurytemora affinis*

By

Marijan Posavi

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Zoology)

at the

UNIVERSITY OF WISCONSIN-MADISON

2015

Date of final oral examination: 5/4/15

The dissertation is approved by the following members of the Final Oral Committee:

Carol E. Lee, Professor, Zoology
Bret R. Larget, Professor, Statistics and Botany
Yuseob Kim, Associate Professor, Life Science, Ewha Womans University, Korea
John E. Pool, Assistant Professor, Genetics
Colin N. Dewey, Assistant Professor, Biostatistics and Medical Informatics, and Computer Science

Table of Contents

Abstract

Evolutionary mechanisms of rapid adaptation to environmental change are poorly understood. I use inbred lines generated from ancestral saline and derived freshwater populations of the copepod *Eurytemora affinis* to study evolutionary mechanisms of maintenance of genetic variation under varying conditions (Chapter 1), and the extent and nature of regulatory evolution during invasions to freshwater habitats (Chapters 2 and 3). In the first chapter, I provide a rare empirical example of complete beneficial reversal of dominance associated with environmental change and evidence for marginal overdominance in salinity tolerance. These mechanisms might be crucial for maintaining genetic variation in salinity tolerance in *E. affinis*, allowing rapid adaptation to salinity changes during habitat invasions. In the second chapter, I report striking shifts in gene expression between freshwater and saline lines, out of which relatively small number was due to a plastic response (acclimation), while a majority of gene expression shifts were evolutionary (heritable differences). This indicates that regulatory evolution might play an important role in rapid adaptation during habitat invasions in *E. affinis*. In the third chapter, I characterize the relative contribution of *cis* and *trans*-regulatory evolution associated with freshwater invasions. Majority of gene expression shifts between freshwater and saline lines were due to *trans*, or combined effects of both *cis* and *trans*-regulatory effects, while several important genes, such as Na^+/H^+ exchanger and acidic chitinase, showed evidence of *cis* regulatory evolution only. These results contribute significantly to our understanding of regulatory evolution during rapid adaptation in *E. affinis*.

List of Tables and Figures

Chapter 1

occurs when $0 \leq [h_1, h_2] < 0.5$ (Curtsinger et al. 1994). If $h_1 = h_2 = 0$ (complete dominance), beneficial reversal of dominance would result in marginal overdominance in Wallace's sense (1968), where the arithmetic mean of heterozygotes is greater than that of homozygotes (arithmetic mean overdominance; Wills 1976; Felsenstein 1976). Marginal overdominance could also occur when the geometric or harmonic mean fitness of the heterozygote across environments is greater than that of both homozygotes (Gillespie 1973; Levene 1953).

dominance and marginal overdominance in salinity tolerance for the copepod *E. affinis*. (**A**) Juveniles from all four saline and freshwater parental inbred lines (FW_1, FW_2, SW_1 and SW_2) were gradually transferred to a common salinity of 2.5 PSU, and reared at this salinity until they became sexually mature. (**B**) Upon reaching sexual maturity, 16 mating combinations were formed (as shown in Table 1) to obtain three different types of F1 offspring: (a) parental inbred lines (SW_1, SW_2, FW_1, and FW_2), (b) reciprocal within-salinity F1 crosses (SW-F1 and FW-F1), and (c) reciprocal between-salinity F1 crosses (SWxFW-F1). (**C**) After successful mating, egg sacs (F1 offspring) were removed from females and each egg sac was split across three salinities (0, 2.5 and 15 PSU) and reared until adulthood to measure survival and infer changes in dominance of salinity tolerance across salinities. The common-garden experiment was conducted in two blocks at two different time periods.

Figure 3. Maximum-likelihood estimates of probabilities of survival from hatching to adult for (**a**) two freshwater and two saltwater parental inbred lines, (**b**) reciprocal within-salinity F1 crosses, and (**c**) reciprocal between-salinity F1 crosses. Survival of the between-salinity F1 crosses (c, purple lines) Page 50

was not significantly different from survival of freshwater F1 crosses (b, blue dashed lines) under freshwater conditions or survival of saltwater F1 crosses (b, red dashed lines) under saltwater conditions. This pattern of survival strongly supported the presence of beneficial reversal of dominance (also see Tables 3 and 4).

Figure 4. Survival from hatching to adult (maximum-likelihood estimates, numbers in or above the bars) in four inbred lines and their reciprocal crosses under (A) saltwater conditions (15 PSU) and (B) freshwater conditions (0 PSU). Red arrows (middle) indicate the increase in survival due to heterosis alone in the saltwater within-salinity crosses (SW-F1, light red bars), calculated as the difference between mean survival of saltwater within-salinity crosses (SW-F1) and saltwater parental inbred lines (SW_1 and SW_2, red bars). Blue arrows (left) indicate the increase in survival due to heterosis in the freshwater within-salinity crosses (FW-F1, blue striped bars), calculated as the difference between mean survival of FW-F1 and freshwater parental inbred lines (FW_1 and FW_2, blue bars). Purple arrows (right) indicate the increase in survival in the between-salinity F1 crosses (SWxFW-F1, purple bars) due to reversal of dominance, calculated as the Page 51

Chapter 2

Na^+/H^+ exchanger) might be involved in ionic regulation in crustaceans. Location of ion transport enzymes (apical *vs.*basolateral) might be crucial for effective osmoregulation. ATP levels needed for higher activity of ion transport activity in freshwater environment might be maintained by arginine kinase. Under saline conditions high ionic concentration of water relative to that of the intracellular fluid allows diffusion of Na^+ into the cell. Na^+/K^+-ATPase (red), located on basolateral membrane than uptakes the Na+ ions. On the contrary, ionic concentration of the freshwater is orders of magnitude lower than that of saline. The crucial transporter and enzyme, involved in osmoregulation, might be Na^+/H^+ exchanger (blue) and V-type H^+-ATPase (teal).

Figure 2. A common garden experiment scheme that allowed distinguishing between environmentally induced and genetically based differences in gene expression between saline and freshwater inbred lines. **(A)** Juveniles from all four saline and freshwater parental inbred lines (FW1, FW2, SW1 and SW2) were gradually transferred to a common salinity of 5 PSU, and reared at this salinity until they became sexually mature. **(B)** The newly produced offspring was separated from parents and reared at 5 PSU until metamorphosis (~15 days of age) **(C)** When offspring reached metamorphosis each sample was split across two salinities (0 and 15 PSU). (**D**) Juveniles

Page 111

were reared at final salinities (either 0 or 15 PSU) for the next 16-18 days i.e., until they became adults. The total of 50 adult copepods (25 females and 25 males) were randomly selected from each sample for total RNA extraction. Samples were sequenced in three batches allowing only FW1 *vs*. SW2 and FW2 *vs*. SW1 pairwise comparisons.

Figure 3. An overview of a protocol used to detect DE genes. Page 112

RNA-seq data, complete genome sequence (fasta format), and gene annotation (GTF file) were used as input to TopHat which uses Bowtie as alignment engine. Mapped reads were than assembled in Cufflinks (Trapnell et. 2012) producing transcript assemblies for all 22 samples collected from four inbred lines. In the next step, Cuffmerge utility was used to merge all 22 transcript assemblies with original gene annotation to produce single improved GTF file. This improved GTF file was then used as input to RSEM (Li and Dewey 2011) to generate reference transcript sequences and to estimate transcript abundances at gene and isoform-level. RSEM abundance estimates (estimated number of fragments for a given gene or isoform) were used to test for differential expression applying Generalized Linear Models (GLM) framework using Bioconductor Package EdgeR (McCarthy et al. 2012). Employed statistical model accommodated for complex design of common garden experiment.

differentially expressed between FW and SW inbred lines in freshwater and saline conditions, respectively. Genes that were differentially expressed (based on FDR < 0.01 from negative binomial GLM) across both FW1 *vs*. SW2 and FW2 *vs*. SW1 inbred lines comparisons are denoted by **, while genes that were DE in only one comparison are denoted by *. Heat map of TMM normalized gene counts of subset of genes that were grouped into GO category — sodium ion transport (containing 60 genes in total) under (**A**) freshwater and (**B**) saline conditions. Although GO — sodium ion transport was significantly enriched (FDR corrected P value = 0.00036) only 8 genes were differentially expressed between FW and SW inbred lines in freshwater and saline conditions, respectively. Genes that were differentially expressed (based on FDR < 0.01 from negative binomial GLM) across both FW1 *vs*. SW2 and FW2 *vs*. SW1 inbred lines comparisons are denoted by **, while genes that were DE in only one comparison are denoted by *.

Chapter 3

sex was determined females and males were paired to produce three different type of offspring: **SW x SW** (saline parental line), **FW x FW** (freshwater parental line) and ♀**SW** x ♂**FW** (F1 crosses i.e., hybrids). The newly hatched offspring was separated from parents and reared at 5 PSU until metamorphosis (~15 days of age). **(B)** When offspring reached metamorphosis each sample was transferred to freshwater (0 PSU). (**C**) Juveniles were kept under freshwater conditions for the next 16-18 days i.e., until they became adults. The total of 50 adult copepods (25 females and 25 males) were randomly selected from each sample for total RNA extraction.

Figure 2. Hypothetical regulatory divergence scenarios in the copepod *E. affinis* (adopted from McManus et al. 2010). Top panel: in hybrids, a *cis*-regulatory mutation in freshwater (FW) population (pink box) decreases affinity for both the conserved saline (SW) and freshwater (FW) transcription factors (red circles). Bottom panel: a *trans*-regulatory mutation in a freshwater (FW) population transcription factor (pink circle) reduces its binding affinity for the conserved freshwater (FW) and saline (SW) *cis*-regulatory regions (red boxes). Page 166

Chapter 1

Testing for beneficial reversal of dominance during salinity shifts in the invasive copepod *Eurytemora affinis*, and implications for the maintenance of genetic variation

This work was published in 2014 in Evolution.

Posavi, M., G. W. Gelembiuk, B. Larget, and C. E. Lee. 2014. Testing for beneficial reversal of dominance during salinity shifts in the invasive copepod *Eurytemora affinis*, and implications for the maintenance of genetic variation. Evolution. 68:3166–3183.

ABSTRACT

Maintenance of genetic variation at loci under selection has profound implications for adaptation under environmental change. In temporally and spatially varying habitats, non-neutral polymorphism could be maintained by heterozygote advantage across environments (marginal overdominance), which could be greatly increased by beneficial reversal of dominance across conditions. I tested for reversal of dominance and marginal overdominance in salinity tolerance in the salt-to-freshwater invading copepod *Eurytemora affinis.* I compared survival of F1 offspring generated by crossing saline and freshwater inbred lines (between-salinity F1 crosses) relative to within-salinity F1 crosses, across three salinities. I found evidence for both beneficial reversal of dominance and marginal overdominance in salinity tolerance. In support of reversal of dominance, survival of between-salinity F1 crosses was not different from that of freshwater F1 crosses under freshwater conditions and saltwater F1 crosses under saltwater conditions. In support of marginal overdominance, between-salinity F1 crosses exhibited significantly higher survival across salinities relative to both freshwater and saltwater F1 crosses. This study provides a rare empirical example of complete beneficial reversal of dominance associated with environmental change. This mechanism might be crucial for maintaining genetic variation in salinity tolerance in *E. affinis* populations, allowing rapid adaptation to salinity changes during habitat invasions.

Introduction

The performance of a population under selection depends on levels of genetic variation underlying relevant phenotypic traits (Crow and Kimura 1970). When populations are invading novel environments, adaptation at critical traits is often required for the populations to survive and persist (Reznick and Ghalambor 2001; Lee et al. 2003; Phillips et al. 2006; Lee et al. 2007; Simons 2007; Keller and Taylor 2008; Lee and Gelembiuk 2008; Prentis et al. 2008; Lee et al. 2011; Nielsen et al. 2012). Given the waiting time required for *de novo* mutations, it is thought that rapid adaptation during invasions relies predominantly on standing genetic variation (Barrett and Schluter 2008; Lee and Gelembiuk 2008; Prentis et al. 2008). Theoretical studies indicate that high standing genetic variation in source populations greatly facilitates adaptation into novel stressful habitats (Gomulkiewicz et al. 1999; Boulding and Hay 2001; Holt et al. 2003; Innan and Kim 2004; Kim and Gulisija 2010; Holt and Barfield 2011). Yet, the details of how non-neutral polymorphism is generated and maintained within populations in nature remain inadequately understood (Turelli and Barton 2004; Mitchell-Olds et al. 2007). This study tests for the presence of an underlying mechanism that potentially acts to facilitate the maintenance of genetic variation within populations, particularly for traits under selection during ecological invasions.

At a locus under selection, observed levels of variation are attributable to two major mechanisms:mutation-selection balance (Lande 1975; Houle et al. 1996) and balancing selection (Gillespie and Turelli 1989; Turelli and Barton 2004). The relative importance of these two mechanisms remains unclear (Houle et al. 1996; Turelli and Barton 2004, Mitchell-Olds et al. 2007). Balancing selection tends to lead to the maintenance of alleles of intermediate frequencies, whereas mutation-selection balance tends to lead to a prevalence of rare alleles (Turelli and Barton 2004). Balancing selection refers to any type of selection that maintains

genetic variation, such as overdominance, frequency dependent selection, and antagonistic selection. In particular, alleles that are subject to opposing selection (e.g., that are beneficial in some environmental contexts and detrimental in others) could be maintained in a population by antagonistic selection between spatially or temporally varying environments (Levene 1953; Haldane and Jayakar 1963; Wallace 1968; Felsenstein 1976; Hedrick 1974, 1976, 1986; Curtsinger et al. 1994; Dean 2005; Epinat and Lenormand 2009; Carter and Nguyen 2011; Connallon and Clark 2012a). Temporally varying selection is of particular interest as a mechanism that maintains variation, because a large number of successful invaders appear to have originated from disturbance-prone temporally varying environments, more than what might be expected due to transport opportunity alone (Lee and Gelembiuk 2008).

In diploid organisms, balancing selection *via* marginal overdominance can maintain polymorphism for alleles that are antagonistically-selected across spatially and temporally varying environments. Marginal overdominance refers to the case when heterozygotes possess higher mean fitness across environments relative to that of any homozygote, even when fitness of the heterozygotes does not exceed that of the best-fit homozygote within any specific environment (Wallace 1968). Marginal overdominance could operate over both temporal and spatial scales. Geometric mean overdominance is sufficient for maintaining protected polymorphisms under a basic model of temporal variation in fitness with random mating (Haldane and Jayakar 1963; Gillespie 1973, 1974, 1998; Hedrick 1976, 1986 and 2006; Hoekstra et al. 1985). On the other hand, harmonic mean overdominance is sufficient for maintaining protected polymorphism under a basic model of spatial variation in fitness with random mating (Levene 1953). As the geometric mean is larger than the harmonic mean, conditions for maintaining protected polymorphisms under temporal variation appear more stringent than under

spatial variation. Additionally, in the presence of recurrent mutation, antagonistic selection can inflate genetic variance to a much greater extent and over a far wider parameter range than could be produced by protected polymorphisms alone (Bürger and Gimelfarb 2002; Connallon and Clark 2012b; Delph and Kelly 2014).

Under antagonistic selection, beneficial reversal of dominance can greatly increase the magnitude of marginal overdominance (i.e., by increasing the mean fitness of heterozygotes) (Kidwell et al. 1977; Curtsinger et al. 1994; Fry 2010; Connallon and Clark 2012b). Beneficial reversal of dominance refers to the case where dominance switches at a locus across distinct traits (e.g. fitness in different environments) such that an allele is always dominant for the trait where it is beneficial and recessive where harmful. For example, in the context of a metabolic pathway, such a switching of dominance across environments might arise due to the more-fit allele in a given environment compensating for the lower function of the allele maladapted to that environment (Wright 1934; Kacser and Burns 1981). Beneficial reversal of dominance greatly increases the parameter range under which polymorphism is favored and the efficiency with which antagonistically selected alleles are maintained (see Discussion) (Curtsinger et al. 1994; Connallon and Clark 2012b). Without beneficial reversal of dominance, the strength of balancing selection for a pair of antagonistically selected alleles is on the order of $N_e s_1 s_2$, whereas with beneficial reversal of dominance, the strength of balancing selection is on the order of $N_e(s_1+s_2)(1-2h)$ (where N_e is effective population size, s_1 and s_2 are selection coefficients, and h is a dominance parameter), which might be orders of magnitude greater (Connallon and Clark 2012b). Masking of conditionally deleterious alleles in heterozygotes can thus dramatically strengthen balancing selection generated by marginal overdominance.

Yet, despite the potential for beneficial reversal of dominance to facilitate adaptation to changing environments, empirical studies demonstrating beneficial reversal of dominance are still rare (Caspari 1950; Kohn et al. 2003; Roux et al. 2004; López et al. 2010; Hund et al 2012) (see Discussion for details). Many studies have focused either on testing for marginal overdominance or on detecting allele frequency fluctuations without testing specifically for beneficial reversal of dominance (Wills 1975; Watt 1983; Hedrick 2002; Mojica et al. 2012; Bergland et al. 2013). Thus, the goal of this study was to test explicitly whether beneficial reversal of dominance for salinity tolerance is operating in the copepod *Eurytemora affinis*. This copepod has invaded fresh water from saline habitats multiple times independently from genetically-distinct sources (Lee 1999). Interestingly, saline (brackish) populations of *E. affinis* that are able to invade freshwater habitats appear to originate from habitats marked by large seasonal fluctuations in salinity (though never completely fresh), whereas saline populations of *E. affinis* that reside in environments with less variation in salinity have not invaded (Lee 1999; Winkler et al. 2008). Thus, an evolutionary history in fluctuating environments potentially corresponds to invasiveness in this species (Lee and Gelembiuk, 2008).

Beneficial reversal of dominance in its most extreme form would result in freshwater tolerance being completely dominant under freshwater conditions and saltwater tolerance being completely dominant under saline conditions. The less fit alleles would always be masked from selection in the heterozygous state, preventing erosion of genetic variation for salinity tolerance in both salt- and freshwater environments (Wallace 1968; Hoekstra et al. 1985; Curtsinger et al. 1994). Such reversal of dominance would explain results from previous studies on *E. affinis*, which found that alleles favoring high-salinity tolerance were apparently maintained in a decades-old freshwater population and alleles favoring freshwater tolerance were apparently

retained in a saline population, despite the presence of tradeoffs between fresh and higher salinity tolerance (Lee et al. 2003, 2007). Moreover, temporally varying antagonistic selection in the presence of reversal of dominance could greatly inflate genetic variation for salinity tolerance in the native, fluctuating estuarine environment. Reversal of dominance should (by rendering any freshwater-beneficial alleles dominant in fresh water) increase initial rates of survival for stocks of *E. affinis* transplanted from the native estuarine environment to a freshwater environment. Most importantly, maintenance of high levels of genetic variation for salinity tolerance in the native estuarine population would increase the ability of invading freshwater-transplanted stocks to rapidly adapt to the novel freshwater environment (Barrett and Schluter 2008; Lee and Gelembiuk 2008).

In order to test for reversal of dominance, and consequently marginal overdominance, I compared survival across salinities of F1 crosses between inbred lines generated from saline and freshwater populations (salt x fresh), relative to control F1 crosses made between saline (salt x salt) or between freshwater lines (fresh x fresh). Support for beneficial reversal of dominance would be evident if survival of F1 crosses between the saltwater and freshwater inbred lines were equivalent to that of saltwater F1 crosses (salt x salt) under saltwater conditions (15 PSU), and to that of freshwater F1 crosses (fresh x fresh) under freshwater conditions (0 PSU). Relatively high survival of saltwater x freshwater F1 crosses across all environments, resulting in the average advantage of the heterozygotes over homozygotes, would provide evidence for marginal overdominance.

Materials and Methods

EXPERIMENTAL INBRED LINES

To test for the presence of beneficial reversal of dominance (and marginal overdominance), four inbred lines derived from saline and freshwater populations of *E. affinis* were crossed, and survival of between-salinity F1 crosses was compared to that of within-salinity F1 crosses (see next paragraph for details on experimental design). The four inbred lines were generated independently, two each from two populations of *E. affinis* through full-sibling mating for 30 generations (2.5 years). Two independent saltwater inbred lines (SW_1 and SW_2) were derived from the ancestral saline population in Baie de L'Isle Verte, St. Lawrence marsh, Quebec, Canada (Fig. 1; 48°00'14" N, 69°25'31" W), whereas two freshwater inbred lines (FW_1 and FW_2) were derived from the freshwater invading population in Lake Michigan at Racine Harbor, Wisconsin, USA (Fig. 1; 42°43'46" N, 87°46'44" W). The two saltwater inbred lines were generated and reared at their native salinity of 15 PSU (PSU ≈ parts per thousand salinity), whereas the two freshwater inbred lines were generated and reared in Lake Michigan water (0 PSU, conductivity ≈ 300 μS/cm).

EXPERIMENTAL DESIGN

In order to test for the presence of beneficial reversal of dominance and marginal overdominance, crosses between and within the saline and freshwater inbred lines were performed and their survival was compared across salinities. Specifically, a full diallel mating scheme (Table 1) was performed to generate three types of F1 offspring: **(a) matings within parental inbred lines** (Fig. 2B-(a); Table 1, diagonal**)** as controls, **(b) within-salinity F1 crosses,** namely matings between inbred lines derived independently from the same population

(i.e., Fig. 2B-(b); Table 1, light grey cells, crosses between FW_1 and FW_2 and between SW_1 and SW_2; denoted hereon as "**SW-F1**" and "**FW-F1**"), to account for effects of heterosis that might arise from mating between genetically-distinct inbred lines, and **(c) between-salinity F1 crosses** (Fig. 2B-(c), Table 1, dark grey cells, denoted hereon as "**SWxFW-F1**"), referring to matings performed between saline (SW_1 or SW_2) and freshwater (FW_1 or FW_2) inbred lines. The F1 offspring were then reared across three salinities (0, 2.5, 15 PSU) to measure survival and infer changes in dominance of salinity tolerance across salinities (Fig. 2C, described below).

The between-salinity F1 crosses (c, above; Fig. 2B-(c)) would be heterozygous at loci that would confer saline or freshwater tolerance. Comparing survival of these between-salinity crosses to survival of the parental inbred lines (a, above; Fig. 2B-(a)) and to that of the within-salinity F1 crosses (b, above; Fig. 2B-(b)) allowed assessing the switching of dominance in salinity tolerance between saline and freshwater habitats. This comparison allowed observing whether the heterozygotes (i.e. between-salinity SWxFW-F1 crosses) exhibited survival that was not significantly different from that of the more-fit homozygote in each habitat. If this were the case, the more-fit allele of the heterozygote would be exhibiting dominance in each habitat (see below).

The within-salinity F1 crosses (mating between inbred lines derived independently from the same population; Fig. 2B-(b), SW-F1 and FW-F1 crosses) were created to account for increased survival that might arise from heterosis (hybrid vigor). The inbred lines that were independently derived from the same populations are likely to have become fixed for different recessive or partially recessive deleterious alleles at some loci across their genomes. F1 crosses between inbred lines fixed for deleterious alleles at different loci would display heterosis due to masking of expression of deleterious alleles in the heterozygous state (Charlesworth and

Charlesworth 1987, 1999). Therefore, within-salinity F1 crosses (Fig. 2B-(b), SW-F1 and FW-F1) could be used to account for heterosis that might arise from crossing genetically-distinct inbred lines. Thus, higher survival of the FW-F1 and SW-F1 crosses relative to the parental inbred lines (Fig. 2B-(a), matings within SW_1, SW_2, FW_1, and FW_2) would be the result of heterosis.

Given that the within-salinity F1 crosses (Fig. 2B-(b), SW-F1 and FW-F1) could account for heterosis, their survival was compared to that of the between-salinity F1 crosses (Fig. 2B-(c), SWxFW-F1) to test for reversal of dominance and marginal overdominance. Support for beneficial reversal of dominance would be evident if the between-salinity F1 crosses (SWxFW-F1) showed no significant difference in survival relative to that of within-saltwater crosses (SW-F1) under saltwater conditions and relative to that of within-freshwater crosses (FW-F1) under freshwater conditions. Such a result would indicate the switching of dominance in salinity tolerance between saline and freshwater conditions. On the other hand, significantly higher survival of the between-salinity F1 crosses (SWxFW-F1) relative to FW-F1 crosses under freshwater conditions and relative to SW-F1 crosses under saltwater conditions would provide support for overdominance at the salinity tolerance loci. To test for the presence of marginal overdominance, marginal survival (arithmetic mean survival across all three salinities, 0, 2.5 and 15 PSU) of the between-salinity F1 crosses (SWxFW-F1) was compared to that of both within-salinity F1 crosses (SW-F1 and FW-F1). Marginal overdominance would be evident if marginal survival of the SWxFW-F1 crosses were significantly higher than that of the SW-F1 and FW-F1 crosses.

PERFORMING THE COMMON-GARDEN EXPERIMENT

In order to assess the presence of reversal of dominance and marginal overdominance, the different types of F1 crosses (Fig. 2B) were reared under three common-garden salinities and their survival was compared at the different salinities (Fig. 2C). The common-garden experiments were conducted under controlled standard conditions, in a 13°C environmental chamber on a 15L:9D photoperiod. To ensure that individuals were virgins prior to crossing, ca. 200 post-metamorphic juveniles from each inbred line were isolated and placed individually into 20 ml scintillation vials filled with 10 ml of 2.5 PSU water (Fig. 2A). To avoid osmotic shock, juveniles from all four inbred lines (SW_1, SW_2, FW_1, and FW_2) were gradually transferred to a common salinity of 2.5 PSU, and then reared at this salinity until they reached sexual maturity (Fig. 2A). Juveniles were reared at this salinity because it was the least stressful common environment for both the saline and freshwater populations. The experiment started with post-metamorphic juveniles to avoid imposing selection (on any residual genetic variance) in response to salinity in the parental inbred lines prior to mating. Prior experiments revealed that most mortality in response to non-native salinities occurs before metamorphosis, and that the copepods are less sensitive to osmotic stress following metamorphosis to the juvenile (copepodid) stage (Lee et al. 2003; Lee et al. 2007). After 8 – 10 days of acclimation at 2.5 PSU (i.e. when juveniles became sexually mature adults), a full diallel mating scheme as described in the previous section was performed (Table 1 and Fig. 2B). Following successful mating, egg sacs (F1 offspring) were removed from females and each egg sac was split across the three treatment salinities (0, 2.5 and 15 PSU) into separate vials (4 – 6 eggs per vial) and reared until adulthood (Fig. 2C). Hatching and survival were recorded every second day.

The common-garden experiment was conducted in two blocks at different time periods, consisting of 3 – 8 replicate clutches per cross in the first block (July – October 2010) and 5 – 11 replicate clutches per cross in the second block (February – May 2011). In total 184 clutches were used, with 8 – 19 replicate clutches per cross of inbred lines, where each clutch was the product of an independent male x female mating (see Table 2, third column). The freshwater alga *Rhodomonas minuta* was fed to all the copepods in the 0 PSU treatment, and saltwater alga *Rhodomonas salina* to all the copepods in the 15 PSU treatment, while a 1:1 mixture of *R. minuta* and *R. salina* was fed to the copeoeds in 2.5 PSU treatment. To avoid bacterial infection all copepods were treated with antibiotic Primaxin® (20 mg/L) every 3 – 4 days.

STATISTICAL ANALYSIS

In order to make comparisons among survival of different types of F1 offspring, survival data was analyzed *via* a mixed effect logistic regression model using the *lme4* package of the statistical software R (Bates 2007). Survival probabilities for the 16 diallel matings (*cross*) under three salinity treatments (0, 2.5 and 15 PSU) were estimated. Fixed effects included *cross* (genotype effect, 16 levels), *salinity* (3 levels), and *cross-by-salinity* interactions (48 levels), while random effects included *block* (time when experiment was conducted) and *clutch* (the effect of belonging to the same egg sac). The response variable (*survival from hatching to adult*) was treated as binary. The experiment was carried out in two blocks, with 62 clutches in block 1 and 122 clutches in block 2 (184 clutches in total).

In the mixed effect logistic regression model, each survival probability takes the form

$$\frac{1}{1+e^{-(h_{ij}+b_k+c_m)}} \qquad (1)$$

where the fixed effect η_{ij} is the expected log odds of survival for the cross *i* at salinity *j*, b_k is the random effect for *block* and c_m is the random effect for *clutch*. Distribution of random effects for *block* (b_k) and *clutch* (c_m) are assumed to be normal with mean 0 and variance σ^2_{block} and σ^2_{clutch}, respectively ($b_k \sim N(0, \sigma^2_{block})$ for $k = 1,2$ and $c_m \sim N(0, \sigma^2_{clutch})$ for $m = 1,\ldots,184$). The estimated mean survival probability for the cross *i* under salinity *j* is

$$\hat{\pi}_{i\ j} = \frac{1}{1+e^{-\hat{\eta}_i}} \qquad (2)$$

where $\hat{\eta}_{ij}$ is estimated from the model and data.

To make statistical inferences about differences in survival between the F1 crosses (Fig. 2B) and to estimate standard errors for survival probabilities of individual crosses, parametric bootstrapping was implemented (Efron and Tibshirani 1993) with 500 bootstrap replicates. The bootstrap approach was used because formulas for standard errors in mixed effects logistic regression models are unavailable. To test for differences in survival between crosses, 95% confidence intervals were implemented. Specifically, 95% confidence intervals for the difference between means that do not contain 0 indicate a significant difference between two crosses at the 0.05 level ($p < 0.05$).

Results

BENEFICIAL REVERSAL OF DOMINANCE BETWEEN SALINITIES

The results strongly supported the presence of beneficial reversal of dominance in salinity tolerance between salt- and freshwater conditions in *E. affinis*. That is, saltwater tolerance was completely dominant under saltwater conditions, whereas freshwater tolerance was completely dominant under freshwater conditions. This reversal of dominance was supported by the lack of

significant differences when comparing survival of the between-salinity F1 crosses (Fig. 2B-(c), SWxFW-F1) with that of the within-freshwater F1 crosses (Fig. 2B-(b), FW-F1) under freshwater conditions (Figs. 3 and 4B, Table 3) and with that of the within-saltwater F1 crosses (Fig. 2B-(b), SW-F1) under saltwater conditions (Figs. 3 and 4A; Table 4). Furthermore, the between-salinity F1 crosses (SWxFW-F1) showed evidence of higher survival (overdominance) under conditions suboptimal, i.e. at 2.5 PSU, to both saltwater (SW-F1) and freshwater (FW-F1) within-salinity F1 crosses (Table 5), indicating that dominance always shifted in a manner that increased the survival of heterozygotes.

Because I observed the same pattern of dominance in salinity tolerance shifting between salt- and freshwater conditions in all eight SWxFW-F1 reciprocal crosses (Tables 3 and 4, Figs. 3 and 4), these results provided strong evidence for the complete beneficial reversal of dominance in salinity tolerance. Under saltwater conditions, survival of all eight reciprocal SWxFW-F1 crosses was not significantly different from survival of the within-saltwater F1 crosses (SW-F1), but was significantly higher ($p < 0.05$) than survival of the within-freshwater F1 crosses (FW-F1) (Fig. 4A, Table 4). Likewise, under freshwater conditions survival of all eight SWxFW-F1 crosses was not significantly different from that of the within-freshwater F1 crosses (FW-F1), but was significantly higher ($p < 0.05$) than survival of the within-saltwater F1 crosses (SW-F1) (Fig. 4B, Table 3). These replicated results provided robust support for the complete dominance of saltwater tolerance under saltwater conditions (Figs. 3 and 4A, Table 4) and the complete dominance of freshwater tolerance under freshwater conditions (Figs. 3 and 4B, Table 3).

The saltwater inbred lines (Fig. 3a, SW_1 and SW_2) and their within-salinity F1 crosses (Fig. 3b, SW-F1) exhibited reaction norms of opposite slope across salinities compared to their

freshwater counterparts (Fig. 3a, FW_1 and FW_2, and Fig. 3b, FW-F1). The opposing slopes were supported by a significant cross-by-salinity interaction ($\chi^2_{14} = 320.3$, $p < 0.001$). Consistent with this significant interaction, the two freshwater parental inbred lines (FW_1 and FW_2) and their reciprocal crosses (FW-F1, i.e., FW_1 x FW_2 and FW_2 x FW_1) displayed significantly higher survival at their native salinity (0 PSU) than at 15 PSU (Fig. 3, blue lines, $p < 0.05$, Table 2), whereas the saltwater inbred lines (SW_1 and SW_2) and their reciprocal crosses (SW-F1, i.e. SW_1 x SW_2 and SW_2 x SW_1) showed the opposite pattern, of significantly lower survival at 0 PSU than at their native 15 PSU (Fig. 3, red lines, $p < 0.05$, Table 2). In sharp contrast, the between-salinity SWxFW-F1 crosses exhibited high survival across all three salinities (Fig. 3c, purple lines, Table 2), showing much flatter reaction norms and no significant cross by salinity interaction ($\chi^2_{14} = 5.399$, $p = 0.979$). These patterns of survival across salinities supported the presence of beneficial reversal of dominance.

I was able to assess the increase in survival due to heterosis by comparing the survival of the parental inbred lines (Fig. 2B-(a), Fig. 3a) with that of the within-salinity crosses (Fig. 2B-(b), Fig. 3b, SW-F1 and FW-F1). I did find evidence of heterosis in some instances, where crosses between independently-derived lines from a population (SW-F1 and FW-F1) showed higher survival than that of some of the parental inbred lines (Fig. 4). Under freshwater conditions the within-freshwater F1 crosses (FW-F1) showed significantly higher survival relative to the FW_1 parental inbred line only ($p < 0.05$, Fig. 4B; Table 3, difference in survival = 0.49 and 0.45), indicating the presence of heterosis in that particular case. Similarly, under saltwater conditions the within-saltwater F1 crosses (SW-F1) showed significantly higher survival relative to the SW_1 parental inbred line ($p < 0.05$, Fig. 4A; Table 4, difference in survival = 0.254), indicative of heterosis.

I found that heterosis could not fully explain the increased survival in the between-salinity crosses (SWxFW-F1) relative to the parental inbred lines (Fig. 4, Tables 3 and 4). As the within-salinity crosses accounted for heterosis resulting from crossing independently-derived inbred lines, the increase in survival relative to the within-salinity crosses was most likely due to the effects of dominance at loci affecting salinity tolerance. Higher survival of the between-salinity F1 crosses (SWxFW-F1) relative to that of within-saltwater F1 crosses (SW-F1) under freshwater conditions (Fig. 4B, Table 3) and relative to that of within-freshwater F1 crosses (FW-F1) under saltwater conditions (Fig. 4A, Table 4) revealed the effects of reversal of dominance. This elevated survival of the between-salinity F1 crosses (SWxFW-F1) beyond that of the within-salinity F1 crosses (SW-F1 or FW-F1) under maladaptive conditions revealed the effects of dominance of the beneficial alleles (Figs. 3, and 4, purple lines or bars).

When I compared survival at an intermediate salinity (2.5 PSU) that was suboptimal for both the saline and freshwater populations, I found evidence of overdominance in the between-salinity F1 crosses (SWxFW-F1) relative to the within-salinity F1 crosses (SW-F1 and FW-F1) (Fig. 3; Table 5). At this intermediate salinity (2.5 PSU), several between-salinity F1 crosses (SW_1xFW_1, FW_1xSW_2, FW_2xSW_2 and SW_2xFW_2) displayed significantly higher survival relative to that of the freshwater FW_2xFW_1-F1 cross and both saline SW-F1 crosses ($p < 0.05$, Table 5), indicating overdominance (heterozygote advantage). The other four SWxFW-F1 crosses exhibited survival that was not significantly different from that of the SW-F1 and FW-F1 crosses ($p > 0.05$). Thus, overall dominance shifted across salinities in a manner that always optimized survival of the between-salinity (SWxFW-F1) crosses, including at the intermediate salinity.

MARGINAL OVERDOMINANCE ACROSS SALINITIES

Although the presence of complete reversal of dominance between the salt- and freshwater environments would consequently result in marginal overdominance, I also formally tested for marginal overdominance (specifically, arithmetic mean overdominance) in salinity tolerance across all three salinities. The presence of marginal overdominance in salinity tolerance was evident from the higher marginal survival (mean survival across all three salinities) of the between-salinity F1 crosses (SWxFW-F1) relative to the within-salinity F1 crosses (SW-F1 and FW-F1) (Table 6, Fig. 5). I found significantly higher marginal survival in seven out of eight between-salinity crosses (SWxFW-F1) relative to the within-salinity crosses (SW-F1 and FW-F1) (Table 6, $p < 0.05$, based on 95% confidence intervals obtained by 500 bootstrap replicates). In the only case where survival of the between-salinity cross was not significantly higher than survival of the within-salinity F1 crosses, the between-salinity cross (FW_2 x SW_1) showed an average survival probability across salinities of 0.70 ± 0.07 (Table 7), which was higher than marginal survivals of the FW_1 x FW_2 (0.55 ± 0.09) and FW_2 x FW_1 (0.50 ± 0.1) crosses (Table 6 and 7). Overall, these results provided strong evidence for marginal overdominance in Wallace's sense (1968, where marginal overdominance = arithmetic mean overdominance) in salinity tolerance in *E. affinis*.

Discussion

EVIDENCE OF REVERSAL OF DOMINANCE AND MARGINAL OVERDOMINANCE IN SALINITY TOLERANCE

For a locus with pleiotropic effects, dominance relationships between two alleles can vary across traits. Beneficial reversal of dominance is a specific case of such variation in dominance, where

an allele that is beneficial for some traits and detrimental for other traits is always dominant in the traits for which it is beneficial and recessive in the traits for which it is detrimental (Kidwell et al. 1977; Gillespie 1978; Hoekstra et al. 1985; Curtsinger et al. 1994; Fry 2010). Relevant traits can include fitness (or components of fitness) in different environments. In order to provide evidence for beneficial reversal of dominance, I demonstrated that survival of the between-salinity F1 crosses (SWxFW-F1, carrying both fresh and salt water tolerance alleles) was not significantly different from that of the freshwater crosses (FW-F1, carrying only freshwater tolerance alleles) under freshwater conditions (Table 3, Figs. 3 and 4B) and not significantly different from that of the saltwater crosses (SW-F1, carrying only saltwater tolerance alleles) under saltwater conditions (Table 4, Figs. 3 and 4A). Additionally, I showed that survival of the between-salinity F1 crosses (SWxFW-F1) was significantly higher ($p < 0.05$) than that of the freshwater crosses (FW-F1) under saltwater conditions (Table 4, Figs. 3 and 4A) and also higher than that of the saltwater crosses (SW-F1) under freshwater conditions, such that the heterozygote exhibited the phenotype of the more fit allele in each environment (Table 3, Figs. 3 and 4B). These results together provided support for the complete dominance of freshwater tolerance under freshwater conditions and the complete dominance of saltwater tolerance under saltwater conditions.

These results indicated that higher survival of the between-salinity (SWxFW-F1) crosses, relative to that of the within-salinity crosses at their less favored salinities (Fig. 4), was not simply a consequence of heterosis. I accounted for the effects of heterosis by comparing survival of the between-salinity crosses to that of the within-salinity crosses (SW-F1 and FW-F1), which were performed between inbred lines that were independently derived from the ancestral (wild saline or wild freshwater) populations. F1 offspring from the within-salinity crosses reveal the

degree of heterosis arising purely from crossing two different inbred lines (each potentially suffering from some degree of inbreeding depression). These within-salinity crosses did in some instances show higher survival, indicative of heterosis, relative to the parental inbred lines (Figs. 4A and 4B, Tables 3 and 4, see Results). As the within-salinity crosses accounted for the effects of heterosis, the higher survival of the between-salinity F1 crosses beyond that of the within-salinity F1 crosses under maladaptive conditions was likely the result of dominance of the conditionally-beneficial allele (i.e., due to beneficial reversal of dominance), and not the result of simple heterosis.

I also formally estimated the effect of reversal in dominance on marginal survival (the mean survival across all three environments i.e., 0, 2.5 and 15 PSU). I found significantly higher marginal survival of the between-salinity SWxFW-F1 crosses relative to both within-salinity F1 crosses (SW-F1 and FW-F1) (Table 6, Fig. 5). This finding directly provides evidence for marginal overdominance in Wallace's (1968) sense of arithmetic mean overdominance (Fig. 5, Table 6) and presents a model for the protection of polymorphism in a population (see Hoekstra et al. 1985).

In addition to complete reversal of dominance between salt- and freshwater conditions, I also observed overdominance at the intermediate salinity (2.5 PSU), which was suboptimal for both fresh- and saltwater parental inbred lines. Overdominance at the intermediate salinity was indicated by significantly higher survival of four of the between-salinity (SWxFW-F1) crosses relative to the within-salinity F1 crosses (SW-F1 and FW-F1) (Table 5, Fig. 3). Thus, overdominance at the intermediate salinity along with the reversal of dominance between saline and freshwater conditions acted to optimize survival of the between-salinity crosses (SWxFW-F1) across salinities.

Beneficial reversal of dominance for survival to adulthood (i.e. the fitness-related trait we measured) need not necessarily imply beneficial reversal of dominance for total fitness. However, survival to adulthood is one of the most central components of fitness. In addition, in prior studies we have not observed any trade-offs in survival among life history stages (Lee et al. 2003, 2007, 2013). Salinity tolerance in *E. affinis* appears to predominantly reflect appropriate osmoregulation — e.g. freshwater adapted lines (with high survival in fresh water) exhibit a less extreme reduction in blood hemolymph concentration at low salinity (i.e. better uptake and conservation of ions) relative to saline adapted lines (Lee et al. 2012). Lines raised at the salinity to which they are adapted also appear to show more typical behavior (e.g. quicker swimming speed and escape responses) (pers. obs.). I expect that, with respect to salinity, fecundity and total fitness would show a positive correlation with survival to adulthood, because all these aforementioned traits would be similarly affected by the ability to osmoregulate appropriately at each salinity. Thus, total fitness is likely to show a similar pattern of beneficial reversal of dominance as survival to adulthood, though this remains to be formally tested.

This study appears to provide *prima facie* evidence for complete beneficial reversal of dominance for a core fitness-related trait. Alternatively, the data could be explained, without invoking reversal of dominance, by complementation of loss-of-function alleles in a multilocus model (Kawecki 1997). Kawecki (1997) has shown that, for a species initially occupying multiple habitats, deleterious mutations at loci that provide crucial habitat-specific functions could drive specialization. Loss of function mutations at loci that are nonessential in a given habitat could result in different loci retaining function in different habitats. For example, in a freshwater population, functional alleles could be lost at loci required for survival in saline habitats and *vice versa*. In such a case, in freshwater x saline F1 offspring, the functional allele at

each locus would compensate for the nonfunctional allele. Thus, freshwater x saline F1 crosses could then display marginal overdominance, with high fitness across all salinities, due to simple dominance of functional over nonfunctional alleles at each locus (i.e., without reversal of dominance).

However, such a mechanism as described above is unlikely to explain the data presented here. Specialization through the accumulation of nonfunctional alleles would be a protracted process (driven by mutation pressure), whereas freshwater adaptation has occurred very rapidly and recently in invasive populations of *E. affinis*. More importantly, the negative genetic correlations observed between fresh- and saltwater tolerance in *E. affinis* (Lee et al. 2003; Lee et al. 2007) indicate the presence of antagonistic pleiotropy, with the same loci affecting survival in both environments. Such antagonistic pleiotropy is also consistent with the observation that in freshwater-adapted populations of *E. affinis*, levels of ion-motive V-type ATPase activity are elevated in freshwater-reared animals and concomitantly reduced in saline-reared animals, when compared to saline-adapted populations (i.e. evidence of a genetic tradeoff for this crucial ion uptake enzyme) (Lee et al. 2011). Thus, reversal of dominance remains the most plausible explanation for my data.

In general, the extent of dominance is not a fixed value, but can vary as a function of genetic background or environmental factors (Bourguet et al. 1996; Billiard and Castric 2011). Rather than conceptualizing our experiment in terms of dominance at separate traits (where survival at each salinity is viewed as its own trait), my experiment could equivalently be conceptualized as measuring a single trait (survival to adulthood) across a set of environments (range of salinities), with reversal of dominance reflecting plasticity of dominance across environments for the alleles governing the trait. One example of dramatic plastic shifts in

dominance across environmental conditions comes from a study of insecticide resistance in the mosquito *Culex pipiens*, conferred by an allele of the *Ace* locus that encodes an acetylcholinesterase that is insensitive to organophosphorus insecticides (Bourguet et al. 1996). Depending on environmental conditions, levels of dominance of insecticide resistance varied from almost complete dominance to almost complete recessivity (Bourguet et al. 1996). However, although this prior study and various others have demonstrated plasticity of dominance across environmental conditions, I am unaware of prior studies with results analogous to mine, where the heterozygote shows fitness (or a major component of fitness, such as survival to adulthood) equal to, alternately, the better of the two homozygotes across the range of an environmental variable.

OTHER EMPIRICAL EXAMPLES OF BENEFICIAL REVERSAL OF DOMINANCE

There are few good, explicit examples of beneficial reversal of dominance in the literature. A study of a locus with antagonistic pleiotropic effects on viability, developmental time, and mating ability in the moth *Ephestia kuhniella* (Caspari 1950) is sometimes cited as an empirical example of reversal of dominance. However, the inferences in this study were somewhat indirect because, for the locus being studied, only one of the two classes of homozygotes could be distinguished from the heterozygote. Thus, as noted by Curtsinger et al. (1994), this study could not distinguish between overdominance and reversal of dominance in all three fitness components (the heterozygote was, for each component, at least equal to and possibly superior to the better homozygote). A cross between inbred lines of *Zea mays* (corn) with different temperature optima for lateral root growth resulted in a hybrid for which lateral root growth at each temperature was comparable to the better parent (Hund et al 2012). However, the data were

somewhat ambiguous (due to a limited number of replicates), a cross in the reciprocal direction (with the opposite inbred line serving as the male *versus* female parent) did not produce the same result, and crosses between other inbred lines did not appear to behave similarly.

A few examples demonstrate apparent beneficial reversal of dominance more clearly. Curtsinger et al. (1994) point out that sickle cell anemia could be conceptualized as a case of beneficial reversal of dominance (and the same interpretation could be applied to other genetic disorders that confer malaria resistance, such as thalassemia) (López et al. 2010). With sickle cell anemia, the wild-type allele shows near-complete dominance with respect to sickle cell disease, such that the heterozygotes generally do not display the disease, whereas the sickle-cell allele displays dominance with respect to malaria resistance. Warfarin resistance in rats furnishes another example, where warfarin resistance due to mutation of the VKORC1 gene is associated with an increased dietary requirement for vitamin K, and an associated substantial fitness cost in resistant homozygotes (Kohn et al. 2003). The wild-type allele is dominant for low vitamin K requirement (i.e. heterozygotes, like wild-type homozygotes, only require a low level of dietary vitamin K) and the warfarin-resistance allele is dominant for warfarin resistance. Yet another example is the *csr1-1* chlorsulfuron herbicide resistance allele in *Arabidopsis thaliana*, which carries a fitness cost in resistant homozygotes. In this case, the *csr1-1* allele is dominant for herbicide resistance and the wild-type allele shows essentially complete dominance with respect to fitness cost (Roux et al. 2004). In these examples, of sickle cell anemia, VKORC1 mutant, and *csr1-1* alleles, homozygotes carrying the mutant allele (i.e. the allele conferring tolerance or resistance) suffer deleterious effects, which are present even in the environment for which the mutant allele is the optimal allele. For example, a sickle cell homozygote in an environment with

a high incidence of malaria will have increased malaria resistance, but will also bear the large fitness cost of sickle cell disease.

In contrast to the cases above, for salinity tolerance in the copepod *E. affinis,* the optimal alleles in the homozygous state allow near full survival and high absolute fitness in their optimal environment. Moreover, reversal of dominance appears complete in the heterozygote, with the optimal alleles in each environment showing full dominance. I will note that the warfarin system could hypothetically provide results similar to mine depending on the design of the experiment, as the fitness cost in the warfarin system is not absolute, but depends on an environmental variable. An experimental design utilizing two environments, one with normal food and a second with food supplemented with high levels of vitamin K and warfarin, should theoretically show high absolute fitness of the optimal homozygote and the heterozygote in both environments. However, as it currently stands, present study is the only one that fully empirically demonstrates the complete beneficial reversal of dominance of fitness across environments.

ORIGINS AND EVOLUTION OF REVERSAL OF DOMINANCE

Reversal of dominance across environments might arise automatically due to the better-fit allele in a given environment compensating for the reduced function of the allele maladapted to that environment (i.e., where dominance is a simple consequence of nonlinearities in metabolic or developmental systems; Wright 1934; Kacser and Burns 1981; Gilchrist and Nijhout 2001). Alternatively, it is theoretically possible that plasticity in dominance could be selected for, to produce higher dominance of conditionally-beneficial alleles for the conditions under which they are beneficial (Bourguet 1999; Otto and Bourguet 1999; Rice 2002). Such selection for reversal of dominance might be particularly intense in fluctuating environments.

In some cases, the segregation load associated with marginal overdominance could ultimately be resolved by gene duplication. A gene duplication combining two antagonistically-selected alleles that exhibit marginal overdominance in a heterozygote could allow fixation of permanent "overdominance" for a gene that was previously polymorphic (Haldane 1932; Spofford 1969; Otto and Yong 2002; Labbé et al. 2014). However, for many genes, such a duplication might prove detrimental due to disruptions in gene dosage.

EVOLUTIONARY IMPLICATIONS OF BENEFICIAL REVERSAL OF DOMINANCE AND MARGINAL OVERDOMINANCE

Role in Maintaining Genetic Variation

Marginal overdominance in fitness across environments has profound implications for the maintenance of polymorphism and adaptation in temporally varying environments. With marginal overdominance, survival of the heterozygotes increases, such that a population has a higher chance of maintaining genetic variation in the face of changing environments. Here I found arithmetic mean overdominance for a core component of fitness (survival to adulthood). Arithmetic mean overdominance is more stringent than both harmonic mean overdominance (where harmonic mean overdominance for fitness is required to indefinitely maintain a protected polymorphism under a basic model of spatial heterogeneity) and geometric mean overdominance (where geometric mean overdominance for fitness is required to indefinitely maintain a protected polymorphism under a basic model of temporal fluctuation).

Beneficial reversal of dominance, as found in this study, can greatly strengthen the magnitude of marginal overdominance. With changes in the environment, the less favored allele would be masked from negative selection in the heterozygous state (i.e. since the heterozygote

would be close in fitness to the fitter homozygote in each environment). Under a variety of models, theoretical analyses have found that for antagonistically selected alleles, protected polymorphism can be maintained when conditionally-beneficial alleles are dominant in the conditions under which they are beneficial (Kidwell et al. 1977; Curtsinger et al. 1994; Epinat and Lenormand 2009; Fry 2010; Connallon and Clark 2012b). In the case of complete dominance (Table 8, $h_1 = h_2 = 0$), reversal of dominance would satisfy the requirement for arithmetic mean overdominance across environments, that is, marginal overdominance as defined by Wallace (1968). If alleles show partial dominance (Table 8, $0 < [h_1, h_2] < 0.5$), could result in geometric mean or harmonic mean overdominance (Levene 1953; Gillespie 1973; Felsenstein 1976; Hoekstra et al. 1985). Conditions for geometric mean or harmonic mean overdominance are less stringent than for arithmetic mean overdominance, because the geometric mean or harmonic mean is always less than or equivalent to the arithmetic mean (Felsenstein 1976). Reversal of dominance and spatiotemporally varying selection could thus act in concert to maintain protected polymorphisms, as the resulting marginal overdominance would assure that the less favored allele is protected against negative selection during environmental change.

Not only would beneficial reversal of dominance increase the strength of balancing selection for protected polymorphisms, but, in the presence of recurrent mutation, it should also greatly increase genetic variance across a far wider parameter range and to a greater extent than could be produced by protected polymorphisms alone (Connallon and Clark 2012b). The importance of the latter phenomenon has often been overlooked. Although there is copious literature on conditions required to indefinitely maintain protected polymorphisms, there has been relatively scant attention paid to the potential net positive effect of antagonistic selection on

genetic variance when alleles are not preserved indefinitely. Recent theoretical developments suggest that transient balanced polymorphism may be very common in diploids, with antagonistic selection exhibiting a potentially large genetic variance inflating effect under conditions that would not sustain protected polymorphism (Kelly 2006; Bürger and Gimelfarb 2002; Sellis et al 2011; Connallon and Clark 2012b; Delph and Kelly 2014).

Reversal of dominance could also have potential implications for evolution of mating systems (Epinat and Lenormand 2009). In a spatially heterogeneous habitat with partial migration between niches, beneficial reversal of dominance may allow maintenance of high levels of locally-maladaptive alleles, increasing inbreeding depression (as a reflection of segregation load) upon assortative mating. This may disfavor the evolution of assortative mating and inhibit speciation.

Given the exposure of many organisms to continuously changing environments, marginal overdominance as a mode of balancing selection might be widespread across taxa. For example, balancing selection *via* marginal overdominance may contribute to the observation that the frequency of *Drosophila melanogaster* polymorphisms at hundreds of loci oscillates reproducibly across seasons (Bergland et al. 2014). Numerous studies suggest theoretical potential for higher genetic variation in organisms that originate from fluctuating environments (Korol et al. 1996; Kondrashov and Yampolsky 1996; Bürger and Gimelfarb 2002; Hedrick et al. 2002; Lee and Gelembiuk 2008).

It is likely that the spatiotemporal variation in salinity that *E. affinis* experiences in its native estuarine environment (Winkler et al. 2008) combined with marginal overdominance in salinity tolerance would promote the maintenance of polymorphism at salinity tolerance loci in the wild. In the estuarine habitat, salinity levels are spatially heterogeneous and large fluctuations

in salinity occur on both seasonal and shorter timescales. In the case of *E. affinis*, it appears that dominance relationships shift across environmental conditions such that both salt- and freshwater alleles are protected from removal by natural selection in the heterozygous state.The presence of beneficial reversal of dominance and marginal overdominance that we found in *E. affinis* in this study is concordant with previously reported high levels of genetic variance in salinity tolerance observed in both salt- and freshwater populations of *E. affinis* (Lee et al. 2003, 2007). For example, recessivity of saltwater tolerance in fresh water could explain the presence of saltwater tolerant alleles in the freshwater habitat despite the constant freshwater conditions that would select against them.

Implications for Adaptation During Invasions

It is increasingly recognized that, in many instances, biological invasions require adaptation to the new range (Carroll et al. 2001; Lee 2002; Dambrowski and Feder 2007; Prentis et al. 2008; Sultan et al. 2012; Colautti and Barrett 2013; Vandepitte et al. 2013). Populations may successfully invade into "black hole sink" environments, with novel abiotic and biotic conditions, in which they could not persist without evolutionary rescue (Holt et al. 2002; Chevin and Lande 2009). Successful invasions by *E. affinis* (*via* ballast water discharges and canal building) into bodies of fresh water (with ionic concentrations orders of magnitude lower than the native estuarine range) during the last several decades constitutes one such example (Lee 1999; Lee et al. 2003). Such invasions constitute evolutionary events.

Marginal overdominance would provide conditions for the maintenance of variation in temporally and spatially varying environments and would lead to elevated levels of standing genetic variation that could facilitate rapid adaptation during invasions into novel habitats.

Invasions into novel environments would be facilitated, as successful invasion often requires very rapid adaptation. Such adaptation is more likely to occur from standing genetic variation rather than from *de novo* beneficial mutations arising during invasion events (Innan and Kim 2004; Colosimo et al. 2005; Lee et al. 2007; Barrett and Schulter 2008; Prentis et al. 2008). Thus, marginal overdominance in salinity tolerance in *E. affinis* would not only result in higher survival in response to temporally varying environments within the native range, but also in its greater potential for colonization of and adaptation to new environments.

Reversal of dominance might also increase initial rates of survival and reduce the risk of population extinction upon introduction into a new environment (e.g., when a stock of estuarine *E. affinis* is transplanted to a freshwater environment through ballast water discharge). In a species invasion, the inoculum into the new range typically consists of a relatively limited number of individuals. Reversal of dominance would effectively broaden the tolerance range, providing more individuals with sufficient absolute fitness to survive the new environment. Survival of a lineage during the initial generations of an invasion is a precondition for ultimate evolutionary rescue and persistence in a novel environment (Chevin and Lande 2010; Palmer and Feldman 2012). As the favorable allele should be effectively dominant in the new habitat (and thus visible to selection as heterozygotes), reversal of dominance should also accelerate initial adaptation; though, once the favored allele reaches high frequency, complete adaptation should be retarded, as the unfavorable allele would be masked from selection. It appears that an evolutionary history in variable environments might correspond to invasiveness of species given the large number of invasive populations that originates from temporally varying disturbance-prone environments (Lee and Gelembiuk, 2008). Such varying environments might often be crucial for enabling marginal overdominance to maintain polymorphism and high levels of

standing genetic variance at key traits that might undergo selection during habitat change. Specifically, adaptation might be facilitated along dimensions corresponding to environmental characteristics that are subject to fluctuating selection in the native environment (e.g. salinity in the case of *E. affinis*).

ACKNOWLEDGMENTS

This project was supported by National Science Foundation grants NSF DEB-0448827, DEB-0745828, and DEB-1050565 to Carol E. Lee. G. Gelembiuk, B. Larget, and C. E. Lee provided invaluable contribution with design, analysis, and interpretation. Numerous students assisted with generating the inbred lines over a period of three years, including W. Moss, A. Tritt, B. Metzger, A. Jenstead, K. Lee, A. Bateman, M. Noll, K. Jung, S. Kahn, K. Bridges, T. Pian, and A. Toy. K. Lee, S. Kahn, and J. Hilton provided valuable assistance with conducting common-garden experiment. D. Gulisija provided valuable comments and discussion.

LITERATURE CITED

Barrett, R. D. H., and D. Schluter. 2008. Adaptation from standing genetic variation. Trends Ecol. Evol. 23:38-44.

Bates, D., M. Maechler, B. Bolker and S. Walker. 2013. lme4: Linear mixed effects models using Eigen and S4. R package version 1.0-4. http://CRAN.R-project.org/package=lme4.

Bergland, A. O., E. L. Behrman, E. L., K. R. O'Brien, P. S. Schmidt and D. A. Petrov. 2014. Genomic evidence of rapid and stable adaptive oscillations over seasonal time scales in Drosophila.PLoS Genetics, (in press).

Billiard, S., and V. Castric. 2011. Evidence for Fisher's dominance theory: how many 'special cases'?

Trends in Genetics 27:441-445.

Boulding, E. G., and T. Hay. 2001. Genetic and demographic parameters determining population persistence after a discrete change in the environment. Heredity 86:313–324.

Bourguet, D. 1999. The evolution of dominance. Heredity 83:1–4.

Bourguet, D., M. Prout, and M. Raymond. 1996. Dominance of insecticide resistance presents a plastic response. Genetics 143:407-416.

Bürger, R., and A. Gimelfarb. 2002. Fluctuating environmentsand the role of mutation in maintaining quantitative genetic variation. Genet. Res. 80:31–46

Carroll, S. P., H. Dingle, T. R. Famula, and C. W. Fox. 2001. Genetic architecture of adaptive differentiation in evolving host races of the soapberry bug, *Jadera haematoloma*. Genetica 112: 257–272.

Carter, A.J.R., and A.Q. Nguyen 2011. Antagonistic pleiotropy as a widespread mechanism for the maintenance of polymorphic disease alleles. BMC Med. Genet. 12:160-172.

Caspari, E. 1950. On the Selective Value of the Alleles Rt and rt in *Ephestia kuhniella*. Am. Nat. 84:376-380.

Charlesworth, D.H., and B. Charlesworth.1987. Inbreeding depression and its evolutionary consequences. Annu. Rev. Ecol. Syst. 18:237-268.

Charlesworth, B., and D.H. Charlesworth.1999. The genetic basis of inbreeding depression. Genet. Res., Camb. 74:329-340.

Chevin, L. M., and R. Lande. 2010. When do adaptive plasticity and genetic evolution prevent extinction of a density-regulated population? Evolution 64:1143-1150.

Colautti R. I., and S. C. H. Barrett. 2013. Rapid adaptation to climate facilitates range expansion of an invasive plant. Science 342:364-366.

Colosimo, P. F., K. E. Hosemann, S. Balabhadra, G. Jr. Villarreal, M. Dickson, M. Grimwood, J. Schmutz, R. M. Myers, D. Schluter, and D. M. Kingsley. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. Science 307:1928-1933.

Connallon, T., and A. G. Clark. 2012a. Antagonistic versus nonantagonistic models of balancing selection: characterization of relative timescales and hitchhiking effects of partial selective sweeps. Evolution 67:908-917.

Connallon, T., and A. G. Clark. 2012b. A general population genetic framework for antagonistic selection that accounts for demography and recurrent mutation. Genetics 190:1477-1489.

Crow, J. F., and M. Kimura. 1970. An Introduction to Population Genetics Theory. Harper & Row, Publishers, New York.

Curtsinger, J. W., P. M. Service, and T. Prout. 1994. Antagonistic pleiotropy, reversal of dominance, and genetic-polymorphism. Am. Nat. 144:210–228.

Dambroski, H. R., and J. L. Feder. 2007. Host plant and latitude-related diapause variation in Rhagoletis pomonella: A test for multifaceted life history adaptation on different stages of diapause development. Journal Evol. Biol. 20: 2101-2112.

Dean, A. M. 2005. Protecting Haploid Polymorphisms in Temporally Variable Environments. Genetics 169:1147-1156.

Delph, L. F., and J. K. Kelly. 2014. On the importance of balancing selection in plants. New Phytol. 201:45-56.

Efron, B., and R.J. Tibshirani. 1993. An Introduction to the Bootstrap, Chapter 6, pp. 53-56, Chapman & Hall.

Engel, R. A. 1962. *Eurytemora affinis,* a calanoid copepod new to Lake Erie. Ohio J. Sci. 62:252.

Epinat, G., and T. Lenormand. 2009. The evolution of assortative mating and selfing with in- and outbreeding depression. Evolution 63:2047-2060.

Felsenstein, J. 1976. The theoretical population genetics of variable selection and migration. Annu. Rev. Genet. 10:253–280

Fry, J. D. 2010. The genomic location of sexually anatagonistic variation: some cautionary comments. Evolution 64:1510–1516.

Gilchrist, M. A., and H. F. Nijhout. 2001. Nonlinear developmental processes as sources of dominance. Genetics 159:423–432.

Gelembiuk, G. W., B. D. Eads, M. Posavi, and C. E. Lee. In prep. Evolutionary shifts in gene expression across independent invasions by the copepod *Eurytemora affinis*.

Gillespie, J. H. 1973. Polymorphism in random environments. Theor. Pop. Biol. 4:193-195.

Gillespie, J. H. 1974. Polymorphism in patchy environments. Am. Nat. 108:145-151.

Gillespie, J. H. 1978. A general model to account for enzyme variation in natural populations. V. The SAS-CFF model. Theor. Pop. Biol. 14:1–45.

Gillespie, J. H. 1998. Population Genetics: A Concise Guide. Johns Hopkins University Press, Baltimore, MD.

Gillespie, J. H., and C. H. Langley. 1974. A general model to account for enzyme variation in natural populations. Genetics 76:837-848.

Gillespie, J. H., and M. Turelli. 1989. Genotype-environment interactions and the maintenance of polygenic variation. Genetics 121:129–138.

Gomulkiewicz, R., R. D. Holt, and M. Barfield. 1999. The effects of density dependence and immigration on local adaptation and niche evolution in a black-hole sink environment. Theor. Pop. Biol. 55:283-296.

Haldane, J B S. 1932. The Causes of Evolution. Longmans Green & Co, London.

Haldane, J. B. S., and S. D. Jayakar. 1963. Polymorphism due to selection of varying direction. J. Genet. 58:237-242.

Hartl, D. L., and R. D. Cook. 1973. Balanced polymorphisms of quasi-neutral al1eles. Theor. Pop. Biol 4:163-172.

Hedrick, P. W. 1974. Genetic variation in a heterogeneous environment. I. Temporal heterogeneity and the absolute dominance model. Genetics 78:757–770.

Hedrick, P. W. 1976. Genetic variation in a heterogeneous environment. II. Temporal heterogeneity and directional selection. Genetics 84:145–50.

Hedrick, P.W. 1986. Genetic polymorphism in heterogeneous environments: a decade later. Annu. Rev. Ecol. Syst. 17:535–66.

Hedrick, P. W. 2002. Pathogen resistance and genetic variation at MHC loci. Evolution 56:1902–1908.

Hedrick, P. W. 2005. Genetics of populations. Chapter 4, pp. 184-220, Jones and Bartlett Publishers.

Hedrick, P. W. 2006. Genetic polymorphism in heterogeneous environments: the age of genomics. Annu. Rev. Ecol. Syst. 37:67–93.

Hedrick, P. W., R. N. Lee, and D. Garrigan. 2002. MHC variation in red wolves: evidence for ancestry from coyotes and balancing selection. Mol. Ecol. 11:1905–1913.

Hoekstra, R. F., R. Bijlsma, and A. J. Dolman. 1985. Polymorphism from environmental heterogeneity: models are only robust if the heterozygote is close in fitness to the favored homozygote in each environment. Genet. Res. 45:299-314.

Holt, R. D., R. Gomulkiewicz, and M. Barfield. 2003. The phenomenology of niche evolution via quantitative traits in a 'black-hole' sink. Proc. R. Soc. B 270:215-224.

Holt, R. D., and M. Barfield. 2011. Theoretical perspectives on the statics and dynamics of species' borders in patchy environments. Am. Nat. 178:S6-S25.

Houle, D., B. Morikawa, and M. Lynch. 1996. Comparing mutational variabilities. Genetics 143:1467-1483.

Hund, A., R. Reimer, P. Stamp, and A. Walter. 2012. Can we improve heterosis for root growth of maize by selecting parental inbred lines with different temperature behaviour? Philos. Trans. R. Soc. Lond., B. 367:1580-1588.

Innan, H., and Y. Kim. 2004. Pattern of polymorphism after strong artificial selection in a domestication event. Proceedings of the National Academy of Sciences USA 101:10667-10672.

Kacser, H., and J. A. Burns. 1981 The molecular basis of dominance. Genetics 97:639–666.

Kawecki, T. 1997. Sympatric speciation *via* habitat specialization driven by deleterious mutations. Evolution 51:1751-1763.

Keller, S.R., and D.R. Taylor. 2008. History, chance, and adaptation during biological invasion: separating stochastic phenotypic evolution from response to selection. Ecology Letters 11:852-866.

Kelly, J. K. 2006. Geographical variation in selection, from phenotypes to molecules. Am. Nat. 167:481–495.

Kidwell, J. F., M. T. Clegg, F. M. Stewart, and T. Prout. 1977. Regions of stable equilibria for models of

differential selection in the two sexes under random mating. Genetics 85:171–183.

Kim, Y., and D. Gulisija. 2010. Signatures of recent directional selection under different models of population expansion during colonization of new selective environments. Genetics 184:571-585.

Kohn, M. H., H.-J. Pelz, and R. K. Wayne. 2003. Locus-specific genetic differentiation at Rw among warfarin-resistant rat (*Rattus norvegicus*) populations. Genetics 164:1055–1070.

Kondrashov, A. S. and L.Yu Yampolsky. 1996. High genetic variability under the balance between symmetric mutation and fluctuating stabilizing selection. Genet. Res. 68:157-164.

Korol, A. B., V. M. Kirzhner, Y. I. Ronin, and E. Nevo. 1996. Cyclical environmental changes as a factor maintaining genetic polymorphism. 2. Diploid selection for an additive trait. Evolution 50:1432-1441.

Labbé, P., P. Milesi, A. Yébakima, N. Pasteur, M. Weill, and T. Lenormand. 2014. Gene-dosage effects on fitness in recent adaptive duplications: ace-1 in the mosquito *Culex pipiens*. Evolution 68: 2092–2101.

Lande, R. 1975. The maintenance of genetic variability by mutation in a polygenic character with linked loci. Genet. Res. 26:221–235.

Lee, C. E. 1999. Rapid and repeated invasions of fresh water by the saltwater copepod *Eurytemora affinis*. Evolution 53:1423–1434.

Lee, C. E. 2002. Evolutionary genetics of invasive species. Trends Ecol Evol 17:386-391.

Lee, C. E., J. L. Remfert, and G. W. Gelembiuk. 2003 Evolution of physiological tolerance and performance during freshwater invasions. Integr. Comp. Biol. 43:439-449.

Lee, C. E., J. L. Remfert, and Y. M. Chang. 2007. Response to selection and evolvability of invasive populations. Genetica 129:179–192.

Lee, C. E., and G. W. Gelembiuk. 2008. Evolutionary origins of invasive populations. Evol. Appl. 1:427-448.

Lee, C. E., M. Kiergaard, B. D. Eads, G. W. Gelembiuk, and M. Posavi. 2011. Pumping ions: Rapid parallel evolution of ionic regulation following habitat invasions. Evolution 65:2229-2244.

Lee, C. E., M. Posavi, and G. Charmantier. 2012. Rapid evolution of body fluid regulation following independent invasions into freshwater habitats. J. Evol. Biol. 25:625-633.

Lee, C. E., W. E. Moss, N. Olson, K. F. Chau, Y.-M. Chang, and K. E. Johnson. 2013. Feasting in fresh water: Impacts of food concentration on freshwater tolerance and the evolution of food x salinity response during the expansion from saltwater into freshwater habitats. Evol. Appl. 6:673-689.

Levene, H. 1953. Genetic equilibrium when more than one ecological niche is available. Am. Nat. 87:331-333.

López, C., C. Saravia, A. Gomez, J. Hoebeke, and M. A. Patarroyo. 2010. Mechanisms of genetically-based resistance to malaria. Gene 467:1-12.

Lynch, M., and B. Walsh. 1998. Genetics and Analysis of Quantitative Traits, pp. 610-619. Sinauer Associates, Inc.

Mitchell-Olds, T., J.H. Willis, and D. B. Goldstein. 2007. Which evolutionary processes influence natural genetic variation for phenotypic traits? Nat. Rev. Genet. 8:845–856.

Mojica, J. P., Y. W. Lee, J. H. Willis, and J. K. Kelly. 2012. Spatially and temporally varying selection on intrapopulation quantitative trait loci for a life history trade-off in *Mimulus guttatus*. Mol. Ecol. 21:3718-3728.

Nielsen, J. L., G. T. Ruggerone, and C. E. Zimmerman. 2012. Adaptive strategies and life history characteristics in a warming climate: Salmon in the Arctic? Environ. Biol. Fish. 96:1187-1226.

Otto, S. P., and D. Bourguet. 1999. Balanced polymorphisms and the evolution of dominance. Am. Nat. 153:561–574.

Otto, S.P., and P. Yong. 2002. The evolution of gene duplicates. Advances in Genetics 46:451-483.

Palmer, M. E., and M. W. Feldman. 2012. Survivability is more fundamental than evolvability. PLoS One 7:e38025.

Phillips, B. L., G. P. Brown, J. K. Webb, and R. Shine. 2006. Invasion and the evolution of speed in toads. Nature 439:803.

Prentis, P. J., J. R. U. Wilson, E. E. Dormontt, D. M. Richardson, and A. J. Lowe. 2008. Adaptive evolution in invasive species. Trends Plant Sci. 13:288–294.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Reznick, D. N., and C. K. Ghalambor. 2001. The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. Genetica 112:183–198.

Rice, S. H. 2002. A general population genetic theory for the evolution of developmental interactions. Proc Natl Acad Sci USA 99:15518-15523.

Roux, F., J. Gasquez, and X. Reboud. 2004. The dominance of the herbicide resistance cost in several *Arabidopsis thaliana* mutant lines. Genetics 166:449–460.

Sellis, D., B. J. Callahan, D. A. Petrov, and P. W. Messer. 2011. Heterozygote advantage as a natural

consequence of adaptation in diploids. Proc Natl Acad Sci USA 108:20666-20671.

Simons, A. M. 2007. Selection for increased allocation to offspring number under environmental unpredictability. J. Evol. Biol. 20:813–817.

Spofford, J. B. 1969. Heterosis and the evolution of duplications. Am. Nat. 103:407–432.

Sultan, S. E., T. Horgan-Kobelski, L. M. Nichols, C. E. Riggs, and R. K. Waples. 2013. A resurrection study reveals rapid adaptive evolution within populations of an invasive plant. Evol. Appl. 6:266-278.

Turelli, M., and N. H. Barton. 2004. Polygenic variation maintained by balancing selection: pleiotropy, sex-dependent allelic effects and GxE interactions. Genetics 166:1053–1079.

Vandepitte, K., T. de Meyer, K. Helsen, K. van Acker, I. Roldán-Ruiz, J. Mergeay, and O. Honnay. 2014. Rapid genetic adaptation precedes the spread of an exotic plant species. Mol Ecol. 23:2157-64.

Wallace, B. E. 1968. Selection in favor of heterozygotes. In: Topics in Population Genetics. Norton, New York pp. 203-216.

Watt, W. B. 1983. Adaptation at specific loci. II. Demographic and biochemical elements in the maintenance of the *Colias* PGI polymorphism. Genetics 103:691-724.

Wills, C. 1975. Marginal overdominance in *Drosophila*. Genetics 81:177-189.

Winkler, G., J. J. Dodson, and C. E. Lee. 2008. Heterogeneity within the native range: population genetic analyses of sympatric invasive and noninvasive clades of the freshwater invading copepod *Eurytemora affinis*. Mol. Ecol. 17:415–430.

Wright, S. 1934. Physiological and evolutionary theories of dominance. Am. Nat. 68:25-53.

TABLES AND FIGURES

Table 1. Full diallel mating scheme (Lynch and Walsh, 1998) of four independent inbred lines of *E. affinis* resulting in 16 different F1 offspring used in this study (the first inbred line denotes the male parent). **(a)** Matings within parental inbred lines are on the diagonal. **(b)** Within-salinity F1 crosses (light grey cells, FW-F1 and SW-F1) are crosses between two independent inbred lines derived from the same population. **(c)** Between-salinity crosses F1 (dark grey cells, SWxFW-F1) are crosses between saline and freshwater inbred lines.

INDEPENDENT INBRED LINES (MALES)	INDEPENDENT INBRED LINES (FEMALES)			
	FW_1	**FW_2**	**SW_1**	**SW_2**
FW_1	$FW_1 \times FW_1$	$FW_1 \times FW_2$	$FW_1 \times SW_1$	$FW_1 \times SW_2$
FW_2	$FW_2 \times FW_1$	$FW_2 \times FW_2$	$FW_2 \times SW_1$	$FW_2 \times SW_2$
SW_1	$SW_1 \times FW_1$	$SW_1 \times FW_2$	$SW_1 \times SW_1$	$SW_1 \times SW_2$
SW_2	$SW_2 \times FW_1$	$SW_2 \times FW_2$	$SW_2 \times SW_1$	$SW_2 \times SW_2$

FW_1 and FW_2 – two independent freshwater inbred lines derived from the freshwater Lake Michigan population (Fig. 1, FW)

SW_1 and SW_2 – two independent saltwater inbred lines derived from the saline L'Isle Verte population (Fig. 1, SW)

Table 2. Differences in mean survival (from hatching to adult) at three salinities (0, 2.5, 15 PSU) with standard errors (obtained by 500 parametric bootstrap resamples) for each of sixteen matings. Significant differences ($P < 0.05$) are denoted by *, and are in bold.

Type of F1 offspring	Cross	#Replicates (#Clutches)	Differences in Survival Probablitiy ± SE		
			0 PSU *vs.* 2.5 PSU	**0 PSU *vs.* 15 PSU**	**2.5 PSU *vs.* 15 PSU**
(a) Parental inbred lines	FW_1	14	0.153 ± 0.091	**0.313 ± 0.096***	**0.160 ± 0.071***
	FW_2	16	**0.231 ± 0.100***	**0.606 ± 0.105***	**0.554 ± 0.099***
	SW_1	19	**-0.154 ± 0.066***	**-0.529 ± 0.097***	**-0.375 ± 0.079***
	SW_2	11	**-0.189 ± 0.090***	**-0.550 ± 0.123***	**-0.361 ± 0.110***
(b) Within-salinity F1 crosses	FW_1 x FW_2	10	0.108 ± 0.106	**0.662 ± 0.092***	**0.554 ± 0.111***
	FW_2 x FW_1	8	**0.275 ± 0.128***	**0.607 ± 0.111***	**0.332 ± 0.129***
	SW_1 x SW_2	12	**-0.548 ± 0.103***	**-0.741 ± 0.085***	-0.193 ± 0.106
	SW_2 x SW_1	9	**-0.443 ± 0.114***	**-0.571 ± 0.105***	-0.128 ± 0.123
(c) Between-salinity F1 crosses	FW_1 x SW_1	12	0.039 ± 0.087	-0.003 ± 0.080	-0.042 ± 0.085
	SW_1 x FW_1	11	0.004 ± 0.078	0.071 ± 0.089	0.067 ± 0.089
	FW_1 x SW_2	10	-0.008 ± 0.068	0.032 ± 0.074	0.040 ± 0.074
	SW_2 x FW_1	8	0.080 ± 0.098	0.062 ± 0.093	-0.018 ± 0.103
	FW_2 x SW_1	11	-0.064 ± 0.086	-0.010 ± 0.094	0.054 ± 0.087
	SW_1 x FW_2	12	-0.052 ± 0.095	-0.106 ± 0.091	-0.054 ± 0.090
	FW_2 x SW_2	10	0.034 ± 0.079	0.115 ± 0.084	0.081 ± 0.089
	SW_2 x FW_2	11	-0.016 ± 0.090	-0.022 ± 0.085	-0.006 ± 0.080

Table 3. Differences in maximum-likelihood (ML) estimates of survival (from hatching to adult) in four inbred lines and their reciprocal crosses under freshwater (0 PSU) conditions. Values in cells show pairwise differences in survival between different crosses (row - column values). Bold numbers on the diagonal are ML estimates of survival probabilities for each cross at 0 PSU. Differences in ML estimates of survival probabilities were tested by constructing 95% confidence intervals for mean differences using standard errors (in parentheses) obtained by 500 parametric bootstrap resamples.

Cross	FW_1	FW_2	SW_1	SW_2	FW_1xFW_2	FW_2xFW_1	SW_1xSW_2	SW_2xSW_1	FW_1xSW_1	SW_1xFW_1	FW_1xSW_2	SW_2xFW_1	FW_2xSW_1	SW_1xFW_2	FW_2xSW_2	SW_2xFW_2
FW_1	**0.34**	-0.28 (0.129)	0.34 (0.102)	0.34 (0.102)	-0.49 (0.117)	-0.45 (0.130)	0.29 (0.104)	0.25 (0.108)	-0.45 (0.115)	-0.51 (0.107)	-0.55 (0.109)	-0.52 (0.116)	-0.34 (0.127)	-0.37 (0.121)	-0.53 (0.111)	-0.51 (0.114)
FW_2		**0.62**	0.62 (0.121)	0.62 (0.108)	-0.21 (0.121)	-0.17 (0.137)	0.58 (0.109)	0.54 (0.112)	-0.17 (0.123)	-0.23 (0.112)	-0.26 (0.117)	-0.24 (0.122)	-0.06 (0.127)	-0.09 (0.127)	-0.25 (0.120)	-0.23 (0.123)
SW_1			**0.00**	0.00 (0.000)	-0.83 (0.084)	-0.79 (0.104)	-0.04 (0.031)	-0.08 (0.061)	-0.78 (0.083)	-0.85 (0.076)	-0.88 (0.065)	-0.86 (0.076)	-0.68 (0.104)	-0.71 (0.098)	-0.86 (0.071)	-0.85 (0.082)
SW_2				**0.00**	-0.83 (0.084)	-0.79 (0.104)	-0.04 (0.031)	-0.08 (0.082)	-0.78 (0.083)	-0.85 (0.076)	-0.88 (0.065)	-0.86 (0.076)	-0.68 (0.104)	-0.71 (0.098)	-0.86 (0.071)	-0.85 (0.061)
FW_1xFW_2					**0.83**	0.04 (0.123)	0.78 (0.084)	0.74 (0.093)	0.04 (0.106)	-0.02 (0.100)	-0.06 (0.099)	-0.03 (0.107)	0.15 (0.119)	0.12 (0.120)	-0.04 (0.101)	-0.02 (0.106)
FW_2xFW_1						**0.79**	0.75 (0.105)	0.70 (0.111)	0.00 (0.122)	-0.06 (0.119)	-0.10 (0.112)	-0.07 (0.123)	0.11 (0.136)	0.08 (0.127)	-0.08 (0.121)	-0.06 (0.128)
SW_1xSW_2							**0.04**	-0.04 (0.064)	-0.74 (0.085)	-0.80 (0.080)	-0.84 (0.070)	-0.82 (0.079)	-0.64 (0.102)	-0.67 (0.100)	-0.82 (0.076)	-0.80 (0.086)
SW_2xSW_1								**0.08**	-0.70 (0.094)	-0.76 (0.089)	-0.80 (0.084)	-0.78 (0.088)	-0.59 (0.109)	-0.63 (0.103)	-0.78 (0.086)	-0.76 (0.092)
FW_1xSW_1									**0.79**	-0.06 (0.100)	-0.10 (0.096)	-0.08 (0.103)	0.11 (0.111)	0.08 (0.110)	-0.08 (0.100)	-0.06 (0.102)
SW_1xFW_1										**0.85**	-0.04 (0.089)	-0.01 (0.097)	0.17 (0.113)	0.14 (0.111)	-0.02 (0.093)	0.00 (0.102)
FW_1xSW_2											**0.88**	0.02 (0.094)	0.21 (0.113)	0.17 (0.106)	0.02 (0.094)	0.04 (0.097)
SW_2xFW_1												**0.86**	0.18 (0.114)	0.15 (0.111)	0.00 (0.097)	0.01 (0.109)
FW_2xSW_1													**0.68**	-0.03 (0.122)	-0.19 (0.116)	-0.17 (0.118)
SW_1xFW_2														**0.71**	-0.16 (0.109)	-0.14 (0.112)
FW_2xSW_2															**0.86**	0.02 (0.100)
SW_2xFW_2																**0.85**

Significant ($p < 0.05$) Not significant

Table 4. Differences in maximum-likelihood (ML) estimates of survival (from hatching to adult) in four inbred lines and their reciprocal crosses under saltwater (15 PSU) conditions. Values in cells show pairwise differences in survival between different crosses (row - column values). Bold numbers on the diagonal are ML estimates of survival probabilities for each cross at 15 PSU. Differences in ML estimates of survival probabilities were tested by constructing 95% confidence intervals for mean differences using standard errors (in parentheses) obtained by 500 parametric bootstrap resamples.

Cross	FW_1	FW_2	SW_1	SW_2	FW_1xFW_2	FW_2xFW_1	SW_1xSW_2	SW_2xSW_1	FW_1xSW_1	SW_1xFW_1	FW_1xSW_2	SW_2xFW_1	FW_2xSW_1	SW_1xFW_2	FW_2xSW_2	SW_2xFW_2
FW_1	**0.02**	0.01 (0.029)	-0.51 (0.100)	-0.53 (0.117)	-0.14 (0.085)	-0.16 (0.105)	-0.76 (0.091)	-0.63 (0.117)	-0.77 (0.079)	-0.75 (0.092)	-0.83 (0.072)	-0.78 (0.097)	-0.66 (0.100)	-0.79 (0.084)	-0.73 (0.096)	-0.85 (0.072)
FW_2		**0.01**	-0.52 (0.101)	-0.54 (0.117)	-0.15 (0.084)	-0.17 (0.104)	-0.77 (0.091)	-0.64 (0.114)	-0.78 (0.079)	-0.76 (0.092)	-0.84 (0.072)	-0.79 (0.097)	-0.67 (0.099)	-0.80 (0.083)	-0.74 (0.098)	-0.86 (0.071)
SW_1			**0.53**	-0.02 (0.131)	0.37 (0.111)	0.35 (0.130)	-0.25 (0.121)	-0.13 (0.134)	-0.26 (0.109)	-0.25 (0.116)	-0.32 (0.111)	-0.27 (0.124)	-0.16 (0.121)	-0.29 (0.113)	-0.22 (0.126)	-0.34 (0.109)
SW_2				**0.55**	0.39 (0.123)	0.37 (0.141)	-0.23 (0.133)	-0.11 (0.150)	-0.24 (0.120)	-0.23 (0.133)	-0.30 (0.132)	-0.25 (0.141)	-0.14 (0.136)	-0.27 (0.132)	-0.20 (0.127)	-0.32 (0.123)
FW_1xFW_2					**0.16**	-0.02 (0.121)	-0.62 (0.102)	-0.49 (0.130)	-0.62 (0.101)	-0.61 (0.112)	-0.69 (0.098)	-0.63 (0.123)	-0.52 (0.116)	-0.65 (0.105)	-0.59 (0.114)	-0.71 (0.095)
FW_2xFW_1						**0.18**	-0.60 (0.126)	-0.47 (0.139)	-0.61 (0.119)	-0.59 (0.127)	-0.67 (0.110)	-0.62 (0.135)	-0.51 (0.134)	-0.63 (0.122)	-0.57 (0.134)	-0.69 (0.116)
SW_1xSW_2							**0.78**	0.13 (0.121)	-0.01 (0.108)	0.01 (0.118	-0.07 (0.103)	-0.02 (0.126)	0.10 (0.116)	-0.03 (0.108)	0.03 (0.117)	-0.09 (0.107)
SW_2xSW_1								**0.66**	-0.13 (0.123)	-0.12 (0.126)	-0.20 (0.121)	-0.14 (0.139)	-0.03 (0.141)	-0.16 (0.128)	-0.09 (0.134)	-0.21 (0.121)
FW_1xSW_1									**0.79**	0.01 (0.106)	-0.06 (0.093)	-0.01 (0.116)	0.10 (0.112)	-0.03 (0.099)	0.04 (0.112)	-0.08 (0.090)
SW_1xFW_1										**0.78**	-0.08 (0.103)	-0.02 (0.125)	0.09 (0.119)	-0.04 (0.106)	0.03 (0.117)	-0.09 (0.107)
FW_1xSW_2											**0.85**	0.05 (0.106)	0.16 (0.107)	0.04 (0.093)	0.10 (0105)	-0.02 (0.091)
SW_2xFW_1												**0.80**	0.11 (0.128)	-0.02 (0.118)	0.05 (0.124)	-0.07 (0.111)
FW_2xSW_1													**0.69**	-0.13 (0.114)	-0.06 (0.127)	-0.18 (0.114)
SW_1xFW_2														**0.82**	0.07 (0.111)	-0.05 (0.100)
FW_2xSW_2															**0.75**	-0.12 (0.114)
SW_2xFW_2																**0.87**

Significant ($p < 0.05$) Not significant

Table 5. Differences in maximum-likelihood (ML) estimates of survival (from hatching to adult) in four parental inbred lines and their reciprocal crosses at 2.5 PSU. Differences in ML estimates of survival probability were tested by constructing 95% confidence intervals for mean differences using standard errors (in parentheses) obtained by 500 parametric bootstrap resamples (bold numbers on diagonal are ML estimates of survival for each cross).

Cross	FW_1	FW_2	SW_1	SW_2	FW_1xFW_2	FW_2xFW_1	SW_1xSW_2	SW_2xSW_1	FW_1xSW_1	SW_1xFW_1	FW_1xSW_2	SW_2xFW_1	FW_2xSW_1	SW_1xFW_2	FW_2xSW_2	SW_2xFW_2
FW_1	**0.18**	-0.21 (0.112)	0.03 (0.082)	-0.01 (0.104)	-0.54 (0.117)	-0.33 (0.140)	-0.41 (0.115)	-0.34 (0.132)	-0.56 (0.103)	-0.66 (0.096)	-0.71 (0.091)	-0.60 (0.121)	-0.56 (0.108	-0.58 (0.106	-0.65 (0.102)	-0.68 (0.092)
FW_2		**0.39**	0.23 (0.103)	0.20 (0.116)	-0.33 (0.134)	-0.13 (0.149)	-0.20 (0.130)	-0.14 (0.143)	-0.36 (0.118)	-0.45 (0.110)	-0.50 (0.107)	-0.39 (0.133)	-0.35 (0.120)	-0.37 (0.121)	-0.44 (0.116)	-0.48 (0.110)
SW_1			**0.15**	-0.04 (0.092)	-0.56 (0.115)	-0.36 (0.137)	-0.44 (0.109)	-0.37 (0.134)	-0.59 (0.097)	-0.69 (0.090)	-0.74 (0.078)	-0.63 (0.112)	-0.59 (0.101)	-0.61 (0.099)	-0.68 (0.095)	-0.71 (0.091)
SW_2				**0.19**	-0.53 (0.126)	-0.32 (0.144)	-0.40 (0.122)	-0.34 (0.142)	-0.56 (0.109)	-0.65 (0.104)	-0.70 (0.098)	-0.59 (0.125)	-0.55 (0.117)	-0.57 (0.114)	-0.64 (0.110)	-0.67 (0.098)
FW_1xFW_2					**0.72**	0.21 (0.162)	0.13 (0.136)	0.19 (0.157)	-0.03 (0.129)	-0.12 (0.121)	-0.17 (0.113)	-0.06 (0.136)	-0.02 (0.131)	-0.04 (0.129)	-0.11 (0.120)	-0.15 (0.119)
FW_2xFW_1						**0.51**	-0.08 (0.153)	-0.01 (0.163)	-0.23 (0.145)	-0.33 (0.137)	-0.38 (0.139)	-0.27 (0.164)	-0.23 (0.151)	-0.25 (0.146)	-0.32 (0.146)	-0.35 (0.141)
SW_1xSW_2							**0.59**	0.06 (0.148)	-0.16 (0.124)	-0.25 (0.118)	-0.30 (0.111)	-0.19 (0.133)	-0.15 (0.127)	-0.17 (0.124)	-0.24 (0.124)	-0.27 (0.112)
SW_2xSW_1								**0.53**	-0.22 (0.141)	-0.32 (0.133)	-0.36 (0.135)	-0.25 (0.156)	-0.21 (0.144)	-0.23 (0.144)	-0.30 (0.138)	-0.34 (0.136)
FW_1xSW_1									**0.75**	-0.10 (0.107)	-0.15 (0.095)	-0.03 (0.125)	0.01 (0.113)	-0.02 (0.116)	-0.08 (0.115)	-0.12 (0.105)
SW_1xFW_1										**0.84**	-0.05 (0.085)	0.06 (0.117)	0.10 (0.110)	0.08 -0.11	0.01 (0.102)	-0.02 (0.093)
FW_1xSW_2											**0.89**	0.11 (0.108)	0.15 (0.101)	0.13 (0.092)	0.06 (0.091)	0.03 (0.089)
SW_2xFW_1												**0.78**	0.04 (0.126)	0.02 (0.124)	-0.05 (0.125)	-0.08 (0.118)
FW_2xSW_1													**0.74**	-0.02 (0.116)	-0.09 (0.115)	-0.12 (0.108)
SW_1xFW_2														**0.76**	-0.07 (0.110)	-0.10 (0.107)
FW_2xSW_2															**0.83**	-0.03 (0.107)
SW_2xFW_2																**0.86**

Significant ($p < 0.05$) Not significant

Table 6. Differences in maximum-likelihood (ML) estimates of mean survival (from hatching to adult) across three salinities (marginal survival) in four parental inbred lines and their reciprocal crosses. Values in cells show pairwise differences in mean survival between different crosses (row - column values). Bold numbers on the diagonal are ML estimates of survival probabilities for each cross. Differences in ML estimates of survival probabilities were tested by constructing 95% confidence intervals for mean differences using standard errors (in parentheses) obtained by 500 parametric bootstrap resamples.

Cross	FW_1	FW_2	SW_1	SW_2	FW_1xFW_2	FW_2xFW_1	SW_1xSW_2	SW_2xSW_1	FW_1xSW_1	SW_1xFW_1	FW_1xSW_2	SW_2xFW_1	FW_2xSW_1	SW_1xFW_2	FW_2xSW_2	SW_2xFW_2
FW_1	**0.18**	-0.16 (0.073)	-0.03 (0.059)	-0.04 (0.068)	-0.37 (0.089)	-0.32 (0.098)	-0.25 (0.083)	-0.24 (0.087)	-0.59 (0.067)	-0.63 (0.067)	-0.68 (0.059)	-0.63 (0.069	-0.52 (0.074)	-0.58 (0.069)	-0.63 (0.067)	-0.67 (0.062)
FW_2		**0.34**	0.13 (0.073)	0.11 (0.081)	-0.22 (0.098)	-0.16 (0.107)	-0.09 (0.091)	-0.09 (0.095)	-0.44 (0.079)	-0.48 (0.079)	-0.53 (0.077)	-0.48 (0.082)	-0.36 (0.084)	-0.43 (0.082)	-0.47 (0.079)	-0.52 (0.079)
SW_1			**0.21**	-0.01 (0.065)	-0.34 (0.088)	-0.29 (0.096)	-0.22 (0.078)	-0.21 (0.083)	-0.57 (0.064)	-0.60 (0.062)	-0.66 (0.061)	-0.60 (0.069)	-0.49 (0.072)	-0.55 (0.067)	-0.60 (0.063)	-0.64 (0.060)
SW_2				**0.22**	-0.33 (0.096)	-0.28 (0.104)	-0.21 (0.087)	-0.20 (0.091)	-0.55 (0.075)	-0.59 (0.073)	-0.64 (0.068)	-0.59 (0.075)	-0.48 (0.082)	-0.54 (0.076)	-0.59 (0.074)	-0.63 (0.071)
FW_1xFW_2					**0.55**	0.05 (0.115)	0.13 (0.101)	0.13 (0.109)	-0.22 (0.093)	-0.26 (0.092)	-0.31 (0.090)	-0.26 (0.094)	-0.15 (0.097)	-0.21 (0.097)	-0.26 (0.094)	-0.30 (0.092)
FW_2xFW_1						**0.50**	0.07 (0.110)	0.08 (0.111)	-0.28 (0.102)	-0.32 (0.100)	-0.37 (0.099)	-0.31 (0.104)	-0.20 (0.105)	-0.26 (0.103)	-0.31 (0.098)	-0.35 (0.103)
SW_1xSW_2							**0.43**	0.00 (0.100)	-0.35 (0.088)	-0.39 (0.083)	-0.44 (0.082)	-0.39 (0.089)	-0.27 (0.091)	-0.34 (0.088)	-0.38 (0.082)	-0.43 (0.080)
SW_2xSW_1								**0.42**	-0.35 (0.093)	-0.39 (0.093)	-0.44 (0.088)	-0.39 (0.092)	-0.28 (0.096)	-0.34 (0.090)	-0.39 (0.090)	-0.43 (0.090)
FW_1xSW_1									**0.77**	-0.04 (0.072)	-0.09 (0.068)	-0.04 (0.073)	0.08 (0.078)	0.01 (0.073)	-0.04 (0.070)	-0.08 (0.067)
SW_1xFW_1										**0.81**	-0.01 (0.067)	0.00 (0.074)	0.11 (0.077)	0.05 (0.071)	0.00 (0.067)	-0.04 (0.068)
FW_1xSW_2											**0.87**	0.05 (0.069)	0.17 (0.076)	0.10 (0.068)	0.06 (0.065)	0.01 (0.062)
SW_2xFW_1												**0.81**	0.11 (0.083)	0.05 (0.074)	0.00 (0.073)	-0.04 (0.071)
FW_2xSW_1													**0.70**	-0.06 (0.082)	-0.11 (0.080)	-0.15 (0.079)
SW_1xFW_2														**0.76**	-0.05 (0.072)	-0.09 (0.070)
FW_2xSW_2															**0.81**	-0.04 (0.067)
SW_2xFW_2																**0.85**

Significant ($p < 0.05$) Not significant

Table 7. Maximum likelihood estimates (ML) of probabilities for marginal survival (mean survival across all three salinities, 0, 2.5 and 15 PSU) from hatching to adult. Standard errors (SE) of estimates were obtained by 500 parametric bootstrap resamples.

Type of F1 offspring	**Cross**	**#Replicates (**#clutches)	**ML of Mean Survival ± Bootstrap SE**
(**a**) Parental inbred lines	FW_1	14	0.18 ± 0.051
	FW_2	16	0.34 ± 0.072
	SW_1	19	0.21 ± 0.050
	SW_2	11	0.22 ± 0.062
(**b**) Within- salinity F1 crosses	FW_1xFW_2	10	0.55 ± 0.089
	FW_2xFW_1	8	0.50 ± 0.096
	SW_1xSW_2	12	0.43 ± 0.079
	SW_2xSW_1	9	0.42 ± 0.085
(**c**) Between-salinity F1 crosses	FW_1xSW_1	12	0.77 ± 0.060
	SW_1xFW_1	11	0.81 ± 0.056
	FW_1xSW_2	10	0.87 ± 0.048
	SW_2xFW_1	8	0.81 ± 0.060
	FW_2xSW_1	11	0.70 ± 0.071
	SW_1xFW_2	12	0.76 ± 0.062
	FW_2xSW_2	10	0.81 ± 0.058
	SW_2xFW_2	11	0.85 ± 0.051

Table 8. The simplest model of antagonistic selection involves one locus with two alleles. Beneficial reversal of dominance occurs when $0 \leq [h_1, h_2] < 0.5$ (Curtsinger et al. 1994). If $h_1 = h_2 = 0$ (complete dominance), beneficial reversal of dominance would result in marginal overdominance in Wallace's sense (1968), where the arithmetic mean of heterozygotes is greater than that of homozygotes (arithmetic mean overdominance; Wills 1976; Felsenstein 1976). Marginal overdominance could also occur when the geometric or harmonic mean fitness of the heterozygote across environments is greater than that of both homozygotes (Gillespie 1973; Levene 1953).

	Genotype		
	A_1A_1	A_1A_2	A_2A_2
Environmental (fitness) context 1	1	$1 - h_1s_1$	$1-s_1$
Environmental (fitness) context 2	$1 - s_2$	$1 - h_2s_2$	1
Marginal overdominance if:	$\overline{w}_{A_1A_1}$	$< \quad \overline{w}_{A_1A_2} \quad >$	$\overline{w}_{A_2A_2}$

$\overline{w}_{A_iA_i}$ = arithmetic, geometric, or harmonic mean fitness of genotype A_iA_i; s_i = selection coefficient; h_i = dominance coefficient; i = 1 or 2.

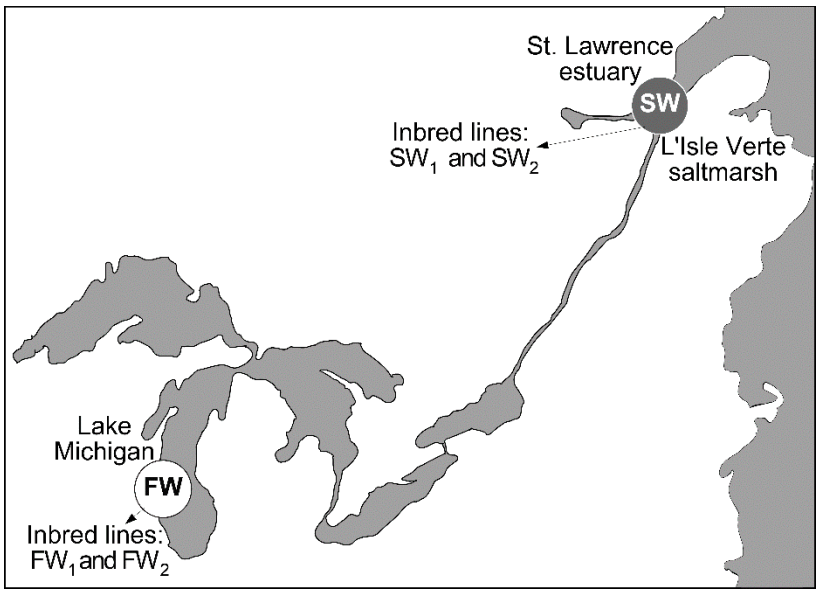

Figure 1. Populations of the copepod *E. affinis* used in this study. The ancestral saltwater population (SW) from Baie de L'Isle Verte salt marsh in the St. Lawrence estuary was used to create two independent saltwater inbred lines (SW_1 and SW_2). The derived freshwater population (FW) from Lake Michigan was established by a recent invasion from the St. Lawrence estuary into the Great Lakes around 1958 (Engel 1962; Lee 1999). Two independent freshwater inbred lines (FW_1 and FW_2) were generated from the Lake Michigan population.

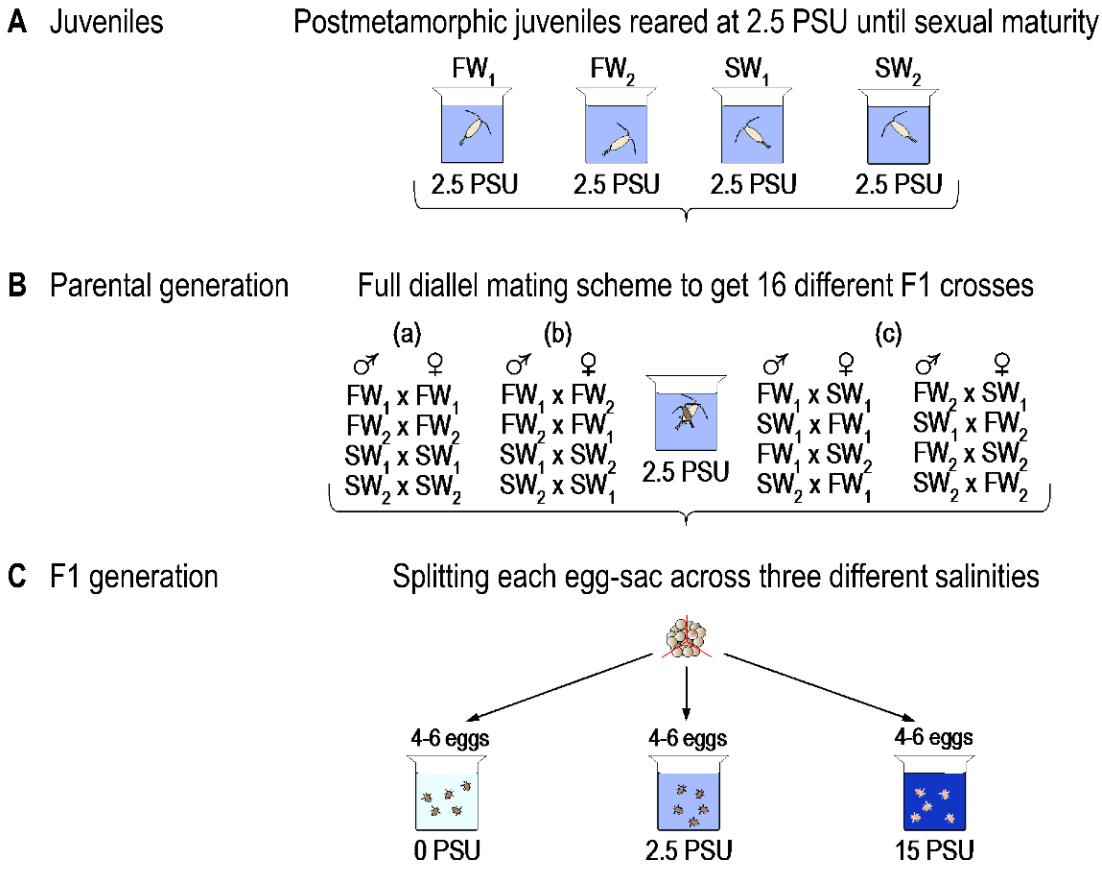

Figure 2. Experimental design to test for reversal of dominance and marginal overdominance in salinity tolerance for the copepod *E. affinis*. **(A)** Juveniles from all four saline and freshwater parental inbred lines (FW_1, FW_2, SW_1 and SW_2) were gradually transferred to a common salinity of 2.5 PSU, and reared at this salinity until they became sexually mature. **(B)** Upon reaching sexual maturity, 16 mating combinations were formed (as shown in Table 1) to obtain three different types of F1 offspring: **(a)** parental inbred lines (SW_1, SW_2, FW_1, and FW_2), **(b)** reciprocal within-salinity F1 crosses (SW-F1 and FW-F1), and **(c)** reciprocal between-salinity F1 crosses (SWxFW-F1). **(C)** After successful mating, egg sacs (F1 offspring) were removed from females and each egg sac was split across three salinities (0, 2.5, and 15 PSU) and reared until adulthood to measure survival and infer changes in dominance of salinity tolerance across salinities. The common-garden experiment was conducted in two blocks at two different time periods.

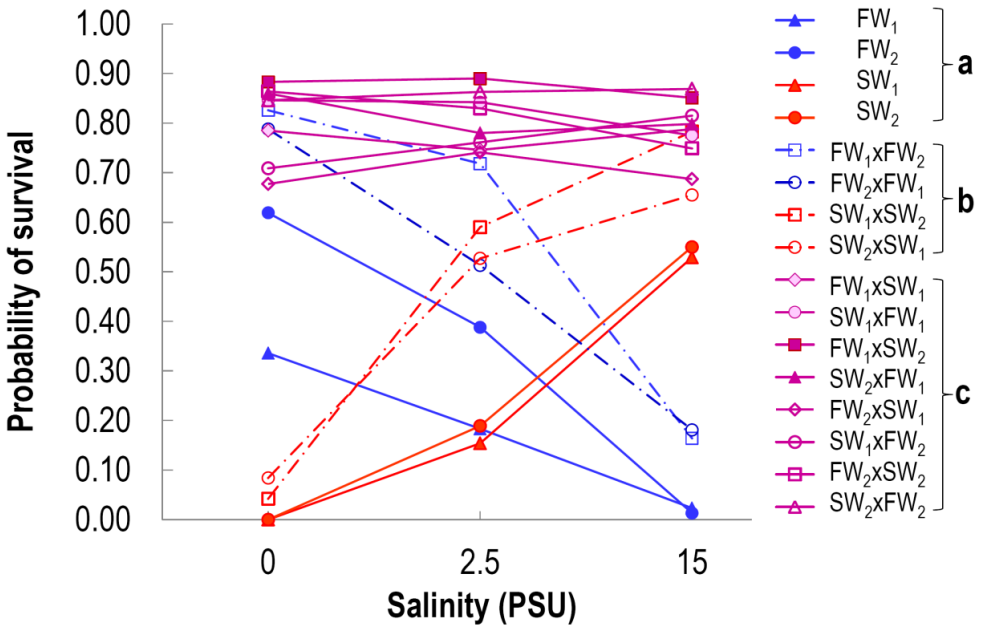

Figure 3. Maximum-likelihood estimates of probabilities of survival from hatching to adult for **(a)** two freshwater and two saltwater parental inbred lines, **(b)** reciprocal within-salinity F1 crosses, and **(c)** reciprocal between-salinity F1 crosses. Survival of the between-salinity F1 crosses (c, purple lines) was not significantly different from survival of freshwater F1 crosses (b, blue dashed lines) under freshwater conditions or survival of saltwater F1 crosses (b, red dashed lines) under saltwater conditions. This pattern of survival strongly supported the presence of beneficial reversal of dominance (also see Tables 3 and 4).

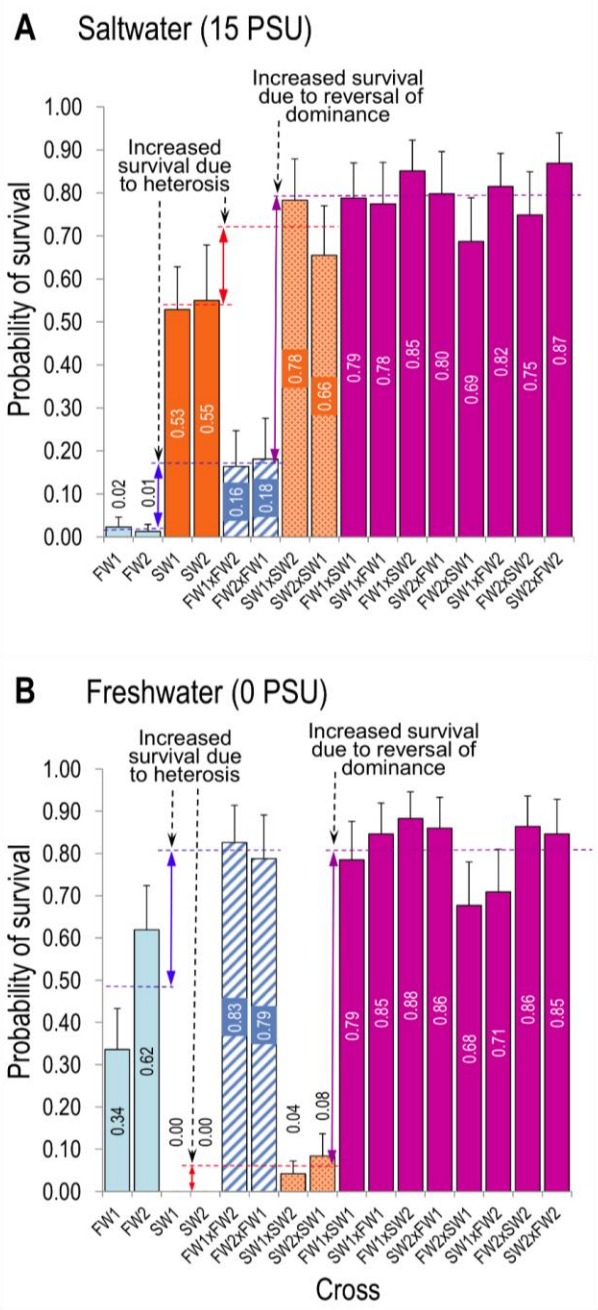

Figure 4. Survival from hatching to adult (maximum-likelihood estimates, numbers in or above the bars) in four inbred lines and their reciprocal crosses under (**A**) saltwater conditions (15 PSU) and (**B**) freshwater conditions (0 PSU). Red arrows (middle) indicate the increase in survival due to heterosis alone in the saltwater within-salinity crosses (SW-F1, light red bars), calculated as the difference between mean survival of saltwater within-salinity crosses (SW-F1) and saltwater parental inbred lines (SW_1 and SW_2, red bars). Blue arrows (left) indicate the increase in survival due to heterosis in the freshwater within-salinity crosses (FW-F1, blue striped bars), calculated as the difference between mean survival of FW-F1 and freshwater parental inbred lines (FW_1 and FW_2, blue bars). Purple arrows (right) indicate the increase in survival in the between-salinity F1 crosses (SWxFW-F1, purple bars) due to reversal of dominance, calculated as the difference between mean survival of the between-salinity F1 crosses (SWxFW- F1) and the within-salinity F1 crosses (SW-F1 and FW-F1).

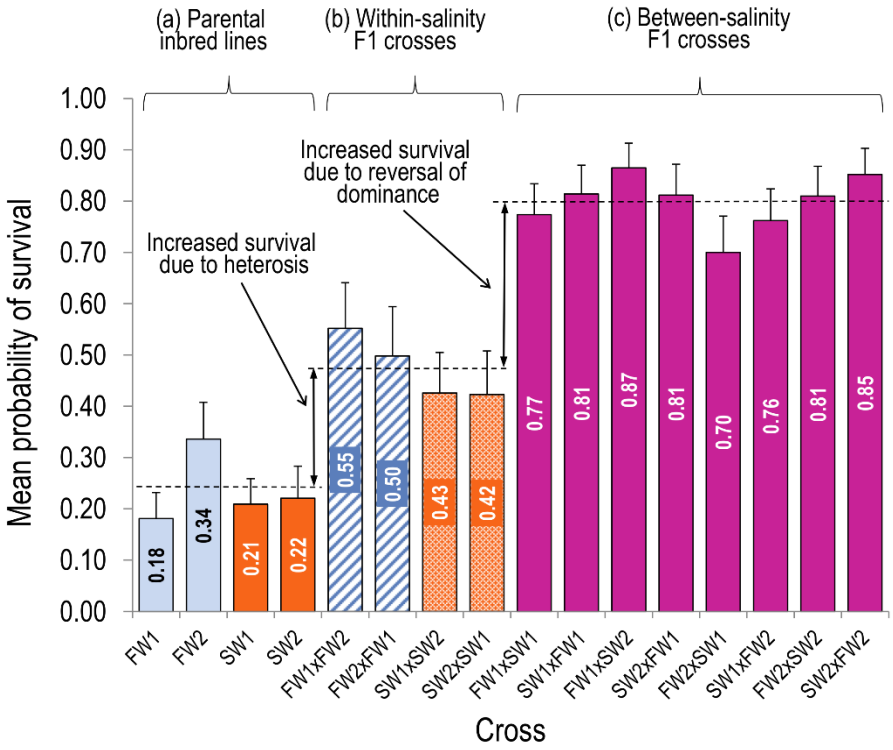

Figure 5. Marginal survival from hatching to adulthood (maximum-likelihood estimates, numbers in the bars) across all three salinities (0, 2.5, and 15 PSU) in parental inbred lines and their reciprocal crosses. Standard error estimates were obtained by 500 parametric bootstrap resamples (see Methods). The crosses shown are **(a)** parental inbred lines, **(b)** within-salinity F1 crosses, between inbred lines independently derived from a population, and **(c)** between-salinity F1 crosses (SWxFW-F1). Differences in survival between the parental inbred lines (a) and the within-salinity F1 crosses (b) were due to heterosis. Differences in survival between the between-salinity F1 crosses (c) and the within-salinity F1 crosses (b) were due to beneficial reversal of dominance in salinity tolerance. Overall higher survival of the between-salinity SWxFW-F1 crosses (c) relative to the within-salinity F1 crosses (b) ($p < 0.05$) provided evidence for marginal overdominance in salinity tolerance.

Chapter 2

Evolutionary shifts in gene expression between ancestral saline and derived freshwater populations of invading copepod *Eurytemora affinis*

ABSTRACT

Evolutionary mechanisms of rapid adaptation during invasions of novel habitats are poorly understood. Comparative gene expression analyses help uncover genes that might be involved during adaptation to novel environments. In this study I characterized evolutionary shifts in gene expression of the dominant zooplankton copepod *Eurytemora affinis* during its invasions from saline into freshwater habitats. I compared patterns of gene expression of inbred lines independently derived from saline and freshwater populations, reared under common garden conditions at both freshwater and saline conditions. For each line and condition, I generated three biological replicates, and used a generalized linear model to analyze RNA-seq data following filtering, normalization, and gene-wise dispersion estimation. I found that relatively small numbers of differentially-expressed genes (FDR < 0.01) between saline and freshwater inbred lines were due to a plastic response to salinity (acclimation), while a majority of differences were due to evolutionary shifts (heritable differences). Carbonic anhydrase, one of the key enzymes involved in sodium uptake, showed an elevated expression in freshwater relative to saline inbred lines. I also found evolutionary shifts in expression of genes involved in energetic metabolism, which might result from high energetic demands required for osmoregulation in freshwater populations of *E. affinis*. Gene sets that were significantly enriched were classified into biological processes such as DNA polymerase activity, DNA integration, enzyme activity, ion transport and homeostasis, membrane and cuticle organization, response to stimulus, amino acid binding and catabolism, and metabolic processes. Plastic genes included those involved in protein metabolism, energy metabolism (including kinase, dehydrogenase and hydrolase activity), structural constituents of cuticle, and transmembrane transport. Most notably, several isoforms of Na^+/H^+ exchangers and Na^+/K^+ -ATPase subunit alpha sowed a strong

acclimatory response. Patterns of gene expression I observed in this study suggest that combined higher activity of carbonic anhydrase, Na^+/H^+ exchangers and Na^+/K^+ -ATPase is necessary for sodium uptake from freshwater environment. It is plausible that regulatory evolution plays important role in rapid adaptation during habitat invasions in *E. affinis*.

Introduction

What allows some populations to colonize novel environments and to persist under drastic changes in environmental conditions? The invasive copepod *Eurytemora affinis*, which is one of the dominant species in many estuaries and saltmarshes worldwide, has independently invaded freshwater habitats multiple times independently over the past 80 years (Lee 1999). During the transition from saline to freshwater conditions, this copepod experiences a drastic change in salinity, crossing the formidable biogeographic boundary that separates saline and freshwater populations of *E. affinis*. This salinity transition poses serious challenges to copepod physiology and has been found to be accompanied by rapid adaptation of underlying ionic regulatory machinery in *E. affinis* (Lee et al. 2011; Lee et al. 2012). However, the mechanisms by which copepods could penetrate the saline to freshwater barrier and adapt to freshwater environments are inadequately understood. Understanding the underlying evolutionary mechanisms of rapid adaptation in *E. affinis* will not only allow us to understand its distribution and invasive success, but also provide a rare opportunity to study rapid adaptation to novel conditions, in general.

Although the physiological mechanisms underlying osmoregulation in aquatic species are not well understood (Towle and Weihrauch 2001, Charmantier et al. 2009, McNamara and Faria 2012), there are a few clear traits that undergo adaptation in response to freshwater conditions in copepods. In particular, selection during freshwater invasions might act on (1) reduction in

integument permeability (Hosfeld 1999), (2) increases in rates of ion uptake (Towle and Weihrauch 2001, Charmantier et al. 2009, Lee et al. 2011, McNamara and Faria 2012), and (3) changes in hemolymph osmolality (Roddie et al. 1984, Thurmann et al. 2010, Lee et al. 2012). These mechanisms are not mutually exclusive and might act in concert, involving multiple ion transporters and enzymes, and specialized structures, such as osmoregulatory organs, to allow rapid adaptation to fresh water by *E. affinis*.

For example, a hypothetical model of ion transport in gills of hyperosmoregulating crabs (Towle and Weihrauch 2001, Rheault et al. 2007, Xiang et al. 2012) proposes a suite of transmembrane transporters and supporting enzymes that might be involved in ionic regulation in crustaceans, namely, Na^+/K^+-ATPase (NKA), V-type H^+ -ATPase (VHA), carbonic anhydrase, $Cl^-/HCO3^-$ exchanger (S4A10), $Na^+/K^+/2Cl^-$ cotransporter (NKCC), and Na^+/H^+ exchanger (NHE) (Fig. 1). Moreover, unlike many larger crustaceans, *E. affinis* does not possess gills. The Lee Lab recently found that ionic and osmotic regulation in *E. affinis* likely occurs primarily in the maxillary glands and newly discovered specialized structures (named "Crusalis organs") located in the swimming legs (Johnson et al. 2014). The Crusalis organs are likely critically important for ionic regulation at low salinities, since they contain ionocytes with high levels of V-type H^+-ATPase expression (Gerber et al., In Prep.). In particular, V-type H^+-ATPase activity and localization in osmoregulatory epithelial cells (apical as opposed to basolateral) have been hypothesized to be critical for freshwater adaptation (Towle and Weihreich 2001, Charmantier et al. 2009, Lee et al. 2011).

Physiological assays revealed evolutionary shifts in activity of ion transporting enzymes (V-type H^+-ATPase and Na^+/K^+-ATPase) between ancestral saline and freshwater derived populations of *E. affinis* (Lee et al. 2011). Furthermore, cDNA microarray analysis suggested

that differential gene expression in Na^+/K^+-ATPase and V-type H^+-ATPase in freshwater relative to saline population might underlie the evolutionary changes of enzymatic activity (Lee et al. 2011). Although, evolutionary shifts in activity of V-type H^+-ATPase, hypothesized to be a crucial enzyme in freshwater adaptation in copepods, did not show a definitive pattern of differential gene expression, V-type H^+-ATPase tended to be upregulated in freshwater populations. The significant difference in V-type-ATPase gene expression between freshwater and saline populations was found for one pairwise comparison only (among 4 pairwise comparisons), and when pooled freshwater populations were contrasted with pooled saline populations (Lee et al. 2011).

While previous research on ion transport evolution in *E. affinis* was informative, it did not offer comprehensive insights into the underlying genetic basis of rapid evolution in invading populations of *E. affinis*. A more comprehensive approach toward elucidating the relative contribution of regulatory adaptation to rapid evolution during invasions is genome wide analysis of expressed genes, a.k.a. RNA-seq or whole transcriptome shotgun sequencing. This method allows comprehensive analysis of the molecular genetic basis underlying phenotypic variation within and between populations (Sîrbu et al. 2012, Zhao et 2014). The powerful feature of RNA-seq method is that it can detect the novel transcripts and isoforms, as well as single nucleotide polymorphisms and indels (Zhao et 2014, Kratz and Carninci 2014).

In this study, I employed an RNA-seq approach to compare transcriptome-wide patterns of gene expression between ancestral saline and derived freshwater populations of *E. affinis* in order to uncover the genetic basis of freshwater adaptation. To this end, I conducted two 2 x 2 factorial common garden experiments — using independently derived saline (SW1 and SW2) and freshwater (FW1 and FW2) inbred lines reared under freshwater (0 PSU) and saline (15

PSU) conditions. The controlled common garden experimental conditions made it possible to distinguish between genetically determined *versus* environmentally induced (acclimation) differences in gene expression between the saline and freshwater populations. Rearing the two populations (in this case, the inbred lines) at common salinities removed the effects of environmental acclimation, and revealed heritable (genetically-based) differences between the populations. On the other hand, detecting differentially expressed genes between conditions (0 *versus* 15 PSU) for a given inbred line revealed genes underlying acclimation to different environments.

Given the evolutionary shifts in ion transport enzyme activity and expression (Lee et al. 2011, Gerber et al., In Prep.) I expect to find evidence of evolutionary shifts between SW and FW inbred lines in expression of V-type H^+-ATPase and Na^+/K^+-ATPase. Specifically I predict upregulation of V-type H^+-ATPase in FW relative to SW inbred lines. In addition, based on proposed models of osmoregulation in freshwater environments (Towle and Weihreich 2001, Charmantier et al. 2009, Ito et al. 2013), I predict a higher expression of carbonic anhydrase, Na^+/H^+ exchanger, and $Na^+/K^+/2Cl^-$ cotransporter in FW relative to SW inbred lines (evolutionary shifts) and under freshwater relative to saline conditions (acclimation).

This study contributes valuable insights into evolutionary shifts in transcriptome-wide gene expression in response to sudden and dramatic environmental change, such as during habitat invasions. This approach could elucidate unknown genes underlying freshwater tolerance and reveal general evolutionary trends that are likely to extend to other many other aquatic invaders. Moreover, understanding rapid adaptation in invasive species might also aid in mitigating their devastating effects, and predicting which populations within an invasive species complex are more likely to invade.

Material and Methods

GENERATING INBRED LINES

To detect evolutionary changes in gene expression associated with colonizations of freshwater habitats inbred lines of *E. affinis* generated from ancestral saline and derived freshwater populations were used. Full-sib mating for 30 generations generated four inbred lines that were used in this study**.** Two saline inbred lines (SW: SW1 and SW2) were independently generated from the saline population in Baie de L'Isle Verte, St. Lawrence marsh, Quebec, Canada (48°00'14" N, 69°25'31" W), whereas two freshwater inbred lines (FW: FW1 and FW2) were derived from the freshwater invading population in Lake Michigan at Racine Harbor, Wisconsin, USA (42°43'46" N, 87°46'44" W). The freshwater inbred lines were generated and reared in filtered lake Michigan water (conductivity ~300 μS/cm, 0 PSU, PSU ≈ parts per thousand salinity), whereas the saline inbred lines were generated and reared in 15 PSU water, made from Instant Ocean and deionized water.

PERFORMING THE COMMON GARDEN EXPERIMENT

In order to disentangle heritable (evolutionary) changes from environmentally induced changes (acclimation) a common garden experiments, comparing four inbred lines across two different, freshwater (0 PSU) and saline (15 PSU), environments (Fig. 2) were performed. Two common-garden experiments (Table 1), each consisting of 2 x 2 factorial design were conducted. In Experiment 1 (September – December 2013) FW2 inbred line was compared to SW1 inbred line, while in Experiment 2 (December 2013 – March 2014) FW1 was compered to SW2 inbred line. Each of four inbred lines was reared under freshwater (0 PSU) and saline (15 PSU) conditions. To infer evolutionary changes between inbred lines, different (saline and freshwater) populations

were compared at the same salinities (0 or 15 PSU) to account for effects of environmental acclimation and observe the evolutionary differences between the populations alone. Comparisons of expression in each of the inbred lines under the contrasting conditions (0 versus 15 PSU) made it possible to infer environmentally induced changes (acclimation). Combined, these comparisons revealed the relative extent to which acclimation and genetically based differences contributed to differences in genomic expression between the saline to freshwater lines.

Prior to the experiment, FW and SW inbred lines were kept at different native salinities (0 and 15 PSU respectively). In the first part of the experiment, copepods were transferred to a common salinity (5 PSU) in order to remove the effects of developmental acclimation in their offspring. The juveniles were reared at 5 PSU because it was the least stressful common environment for both the saline and freshwater inbred lines. Since prior experiments (Lee et al.2003, 2007) revealed that the copepods are less sensitive to osmotic stress following metamorphosis to the juvenile (copepodid) stage the experiment started with postmetamorphic juveniles. About 200 juveniles from each of the two freshwater and two saline inbred lines (FW1, FW2, SW1, and SW2) were gradually transferred to a common salinity of 5 PSU, and reared at this salinity until they became sexually mature, i.e. until they produced offspring (Fig. 2A). Once the offspring (reared at the common salinity of 5 PSU) reached metamorphosis (at age of 14-15 days), cultures were split across two salinities (0 and 15 PSU) and reared under the treatment salinities (Fig. 2B and 2C) until adulthood (age 30-32 days). Thus, samples of copepods from which RNA was extracted (Fig. 2D) were reared at their target salinities (0 or 15 PSU) for ~ 17 days. This amount of time was likely sufficient for acclimation to target salinities since a previous study found that adult *E. affinis* could acclimate to changes in salinity within 12

h (Roddie et al. 1984).

All of the copepods in the 0 PSU treatment were fed the freshwater algae *Rhodomonas minuta*, whereas all of the copepods in the 15 PSU treatment were fed saltwater algae *Rhodomonas salina.* When copepods were reared at 5 PSU they were fed and a 1:1 mixture of *R. minuta* and *R. salina*. All inbred lines were reared under controlled laboratory conditions, at 12°C, and on a 15L:9D photoperiod. To prevent bacterial and fungal infection, copepods were treated with the antibiotic Primaxin® (20 mg/L), D-amino acid cocktail (10 µM of D-methionine, D-leucine and D-tryptophan and 5 µM D-tyrosine), and Voriconazole (0.5 mg/L) every 3 – 4 days. In order to clear the guts of microbes that might contaminate the sequencing data, copepods were starved and treated with 120 µL/L of 6.0 micron copolymer microsphere beads (Thermo Scientific cat# 7505A, Fremont, CA) for the last 24 hours prior to RNA extraction.

RNA EXTRACTION AND SEQUENCING

Once the copepods reached adulthood following the common garden experiment, total RNA was extracted from whole bodies of 50 copepods (25 females and 25 males) per sample (Fig. 1D). Total RNA was extracted with Trizol reagent (Ambion RNA, Carlsbad, CA) and then purified with Qiagen RNeasy Mini Kit (Qiagen cat# 74104, Valencia CA), following the protocol described by Lopez and Bohuski (2007).

Library construction and sequencing were conducted at the Genomics Resource Center (GRC), Institute for Genome Sciences at the University of Maryland School of Medicine. Illumina RNA-seq libraries (strand specific) were prepared with the TruSeq RNA Sample Prep kit (Illumina, San Diego, CA) with modified manufacturer's protocol (adapted from Parkhomchuk et al. 2009). Actinomycin D was added to the first strand synthesis reaction, while

the second strand was synthesized with a dNTP mix containing dUTP. After adapter ligation (6 nucleotide indexes), the second strand cDNA was digested with 2 units of Uracil-N-Glycosylase (Applied Biosystems, Carlsbad, CA). DNA was purified between enzymatic reactions and size selection of the library was performed with AMPure XT beads (Beckman Coulter Genomics, Danvers, MA).

Each library was split across two lanes, and 12 samples were multiplexed per lane on the Illumina HiSeq 2000 (2x101 bp run). Sequencing was conducted in three batches. Replicates 1 and 2 from Experiment 1 (FW2 and SW1 inbred lines) were sequenced in Batch 1, Replicate 1 and 2 of Experiment 2 (FW1 and SW2 inbred lines) were sequenced in Batch 2, and Replicate 3 of both experiments (all four inbred lines) were sequenced in Batch 3 (Table 1). Because experiments were confounded with batch, a separate differential gene expression (DGE) analyses for each experiment were conducted.

DATA PROCESSING

The GRC QC pipeline (see previous section) screened sequence reads against the NCBI nucleotide database in order to evaluate data quality, contamination, and proper laboratory sample tracking. On average 3.5×10^7 fragments, i.e., 7.0×10^7 reads of 101 bp paired-end reads per sample, passed this initial filtering step (Table 1). Further sequence quality filtering and trimming was assessed with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and Trimmomatic (Version 032, Bolger et al. 2014).

In order to quantify transcripts at gene and transcript levels, an expectation maximization approach employing the RSEM (RNA-seq by Expectation Maximization) package (Li and

Dewey 2011) was used. Since the *E. affinis* genome was sequenced recently, only a draft genome with gaps and automated, i.e. incomplete, gene annotation is available. Thus, to improve gene annotation, the Cufflinks Tuxedo protocol (Trapnell et al. 2012) was used. First, the RNA-seq reads of each of 22 samples were mapped to the *E. affinis* genome *via* TopHat (Fig. 3). Then, mapped reads were provided as input into Cufflinks to get a separate transcript assembly for each of 22 samples. Finally, all these assembly files were merged (using Cuffmerge *utility*) with automated gene annotation to create a single gene annotation file. This improved merged gene annotation file was then used as input into RSEM (1) to build reference transcript sequences (using *prepare-reference* script) and (2) to align RNA-seq reads to reference transcripts, and to estimate gene and isoform abundances (using *rsem-calculate expression*). An overview of the protocol used in data analysis is provided in Figure 3.

STATISTICAL ANALYSES

Filtering, normalization, and estimation of dispersion

The gene annotation, created by merging transcriptomes of 22 RNA samples, resulted in 129066 transcripts grouped into 37827 genes from which genes with low counts across all samples were initially filtered out. Since filtering was independent of statistical tests, it acted to increase power of detection of differentially expressed genes (Bourgon et al. 2010). Furthermore, filtering out genes with low counts was a reasonable approach, because genes need to be expressed at some minimum level in order to be translated (Vogel and Marcotte 2012). Thus, genes that had less than one count-per million (CPM) in at least two samples were filtered out. This filtering removed 17271 genes with very low counts while it retained 20556 genes for further analysis.

Next, normalization on the 20566 genes was performed. Initially, two normalization methods, Trimmed Mean of M values (TMM, Robinson and Oshlack 2010) and Upper Quantile (UQ, Bullard et al. 2010) method (Fig. S1) were compared. The two methods found to be equivalent. Thus, the Trimmed Mean of M values (TMM) normalization method, available in Bioconductor's *edgeR* package in the statistical software *R* (McCarthy et al. 2012, Chen et al. 2014, R core team 2013) (Fig. 5C and 5D), was applied.

Finally, a gene-wise dispersion was estimated. An accurate estimation of gene-wise dispersion of the gene counts is crucial for reliable detection of differentially expressed genes (Robinson and Smyth, 2007). Overestimation of dispersion increases the rate of false negative while underestimation increases the rate of false positive tests. Here a common and gene specific dispersion (Fig. S2) was estimated, using a weighted likelihood empirical Bayes method, built into the *estimateDisp* function of the package edgeR (Chen et al. 2014). After obtaining a common dispersion for all genes, a gene-wise empirical Bayes strategy that shrinks each gene's dispersion towards the common dispersion estimate was applied.

To visualize the effect of experimental covariates, the samples were clustered using multidimensional scaling (MDS) and principal component analysis (PCA) (Fig. 6). The MDS function calculates the distance between each pair of samples as the root-mean square distance of the leading fold changes (Chen et al. 2014). Sample clustering was very similar in the both MDS and PCA plots, showing clear separation by genotype (inbred line) in the first dimension and less obvious separation by condition (salinities 0 PSU vs.15 PSU) in second dimension.

Following data filtering, normalization, and estimation of gene-wise dispersion, two statistical analyses, each corresponding to one of the two experiments, were performed *via* a generalized linear model.

Generalized linear model for differential expression analysis

To detect DE genes a generalized linear model (GLM) that accommodated the complex designs of the common garden experiments was used. For this purpose, again Bioconductor's *edgeR* package was employed. The GLM model for the RNA-seq data assumed a negative binomial distribution family, which has the following form (Chen et al. 2014):

$$K_{ij} \sim NB(\mu_{ij}, \phi_i), \quad (1)$$

where K_{ij} represents the counts of reads for gene *i* in sample *j*, μ_{ij} is the gene mean, and ϕ_i is negative binomial dispersion parameter. μ_{ij} is the product of the library size, i.e. sample specific size factor (N_i), and expected proportion reads mapped to gene *i* in sample *j*. A log linear model was used to represent the influence of treatment conditions and library size on the expected counts for each gene,

$$log_2(\mu_{ij}) = x_j^T \beta_i + logN_i, \quad (2)$$

where x_j was the covariate vector specifying treatment for sample *j*, and β_i was the vector of regression coefficients for gene *i*.

To conduct the test for interaction between genotype (inbred lines) and salinity, the read count (K_{ij}) was modeled as the result of the fixed effects of *batch* (2 levels), *genotype* (inbred line effect, 2 levels), *salinity* (2 levels), and *genotype-by-salinity* interactions (4 levels), in each of the experiments. DE genes were detected using FDR (false discovery rate) corrected P values for multiple testing (Benjamini-Hochberg (1995) using a cutoff of 0.01.

Blastx against NCBI's non-redundant and Swiss-Prot (UniProt, Bairoch et al. 2009) protein databases was performed employing the pipeline developed by De Wit et al. (2012). This pipeline combined the blast search results with UniProt flat files in order to complete functional

annotation and extract gene annotation descriptions and gene ontology (GO) categories (Ashburneret al. 2000). For a functional enrichment analysis, Gene Score Resampling analysis (implemented in ErmineJ, Gillis et al. 2010) was employed, using gene annotation containing UniProt and gene ontology (GO) identifiers. This analysis was conducted using log fold-changes (logFC) and TMM normalized (Robinson and Oshlack 2010) gene expression counts from above-mentioned gene expression analysis. Statistically significant GO categories were determined by 10^3 iterations and using an FDR corrected P value < 0.01 (Benjamini and Hochberg 1995) as a cutoff value.

Results

The common-garden experiments revealed both striking evolutionary differences in gene expression between saline and freshwater inbred lines, as well as acclimatory differences between the same inbred lines kept at different salinities. In terms of evolutionary differences between saline and freshwater inbred lines, 17 – 22% genes were differentially expressed (FDR < 0.01) under freshwater (0 PSU) and saline (15 PSU) conditions (Fig. 7). On the other hand, in terms of acclimatory changes between conditions (0 PSU vs. 15 PSU), only 2 – 4% of genes were differentially expressed (FDR < 0.01), indicating a weak plastic response in gene expression in response to salinity change for both saline and freshwater inbred lines (Fig. 10). Combined, these results indicated that the evolutionary changes in gene expression between inbred lines were much greater than the acclimatory changes within inbred lines in response to salinity. The most striking, carbonic anhydrase (involved in proton production) and Na^+/H^+ exchanger 1B (involved in sodium uptake from environment) were both upregulated in freshwater relative to saline inbred lines (Table 4., Fig. 9)

EVOLUTIONARY SHIFTS IN GENE EXPRESSION BETWEEN SALINE- AND FRESHWATER INBRED LINES

Pairwise comparisons between FW and SW inbred lines under freshwater conditions revealed 3916 (FW1 *vs.* SW2, Fig. 7A) and 4312 (FW2 vs. SW1, Fig. 7B) differentially expressed (DE) genes (FDR < 0.01). The comparison between FW and SW inbred lines under saline conditions (15 PSU) revealed 3948 (FW1 *vs.* SW2, Fig. 7C) and 4506 (FW2 *vs.* SW1, Fig. 7D) DE genes (FDR < 0.01). DE genes between saline and freshwater inbred lines under freshwater conditions (0 PSU) are of the special interest because some of them might underlie rapid adaptation to freshwater conditions.

In order to increase the confidence that DE genes point to important adaptive changes, four independent inbred lines were used with primary focus on genes that were differentially expressed in both FW1 vs. SW2 and FW2 vs. SW1 inbred lines pairwise comparisons. The inbreeding process (full-sib mating) imposes strong genetic drift, which could lead to neutral genetic differentiation between the lines (Khaitovich et al. 2005). However, it was assumed that it would be less likely that the same genes were affected in the same manner (either always up or down regulated in all FW compared to SW inbred lines) by genetic drift in both between FW and SW line comparisons. Under freshwater conditions, 1522 genes were DE (FDR < 0.01) across both FW1 *vs.* SW2 and FW2 *vs.* SW1 inbred line comparisons (Fig. 8 A). Under saline conditions, the intersection between both comparisons (FW1 *vs.* SW2 and FW2 *vs.* SW1) revealed 1442 DE genes (Fig. 8 B). Thus, gene expression differences in core sets of 1552 (freshwater conditions) and 1442 (saline conditions) genes identified in both pairwise comparisons represent putative evolutionary shifts due to adaptive evolution in gene expression during freshwater invasion. Note, however, that some adaptive mutations might have been lost in

one of the inbred lines due to genetic drift during the inbreeding process, such that examining only shared DE genes between the comparisons is a conservative estimate of potentially adaptive gene expression changes.

Under freshwater conditions, expression of the cytoplasmic isoform of carbonic anhydrase was 5-fold elevated (log_2 scale) across both FW1 *vs*. SW2 and FW2 *vs*. SW1 inbred lines comparisons (Table 4), while expression of Na^+/H^+ exchanger 1B was 2-fold (log_2 scale) elevated in FW2 relative to SW1 inbred line (Fig. 9). Upregulated carbonic anhydrase, which reversibly catalyzes reaction $H_2O + CO_2 \rightarrow H^+ + HCO_3^-$, indicates higher levels of H^+ in cytoplasm of ionocytes. These higher levels of H^+ can be used for exchange with Na^+ from diluted environment by action of Na^+/H^+ exchanger, which was also upregulated in freshwater relative to saline inbred lines. This evolutionary shift in carbonic anhydrase and Na^+/H^+ exchanger indicates that freshwater inbred lines primarily uptake Na^+ from freshwater environment by exchange for cytoplasmic H^+ produced by carbonic anhydrase catalysis (Fig. 1).

Functional enrichment analysis was conducted using the Gene Score Resampling method (Gillis et al. 2010) and absolute values of log fold-changes (logFC) of gene expression between FW and SW inbred lines. This analysis revealed 40 biologically relevant gene sets that were significantly enriched, classified into biological processes such as DNA polymerase activity, DNA integration, enzyme activity, ion transport and homeostasis, membrane and cuticle organization, response to stimulus, amino acids binding and catabolism, and catabolic and other metabolic processes (Table 3). Functional enrichment analysis revealed 60 genes involved in sodium ion transport, with 8 of them showing differential expression (FDR < 0.01) between both FW1 *vs*. SW2 and FW2 *vs.* SW1 inbred lines under freshwater and saline conditions (Fig. 9). Interesting, Na^+/K^+-ATPase subunit alpha was upregulated (FDR < 0.01) in FW relative to SW

inbred lines under saline conditions (Fig. 9B, Table 6). The subunit alpha is catalytic part of Na^+/K^+-ATPase, the key enzyme involved in transmembrane sodium transport, located in the basolateral membrane of epithelial cells (i.e., ionocytes) where it pumps three Na^+ out of the cell (into blood) in exchange for two K^+ (Towle and Kays 1986) per one ATP consumed. The higher expression of alpha subunit of Na^+/K^+ -ATPase in FW inbred lines relative to SW inbred lines under saline conditions indicates the higher transport of Na^+ from ionocytes into hemolymph. In contrast, two isoforms of the beta subunit of Na^+/K^+ -ATPase were downregulated in FW inbred lines relative to SW inbred lines across both FW1 *vs*. SW2 and FW2 *vs*. SW1 pairwise comparisons, under both freshwater and saline conditions (Fig. 9A and B, Tables 5 and 7). Although beta subunit is not associated with catalytic activity of Na^+/K^+ -ATPase, it is essential for regulation of the transport of sodium to the basolateral plasma membrane (McDonough 1990). Thus, downregulation of beta subunit might suggest lower number of Na^+/K^+ -ATPases in the basolateral (plasma) membrane of ionocytes. This pattern of downregulation in FW relative to SW inbred lines was in concordance with lower enzymatic activity of Na^+/K^+ -ATPase and also lower expression of the alpha subunit of Na^+/K^+ -ATPase, found in a microarray study, in freshwater populations of *E. affinis* relative to saline populations across salinities (Lee et al. 2011). In this study, the alpha catalytic subunit of sodium pump (Na^+/K^+ -ATPase-alpha) was not differentially expressed under freshwater conditions, but was upregulated in FW relative to SW inbred lines under saline conditions (Fig. 9, Table 6).

Another important ion transporter that moves sodium, potassium, and chloride into and out of cells, and plays an essential role in maintenance of cell volume and electrolyte concentration, is the $Na^+/K^+/2Cl^-$ cotransporter also known as *solute carrier family 12 member 2* (Russell 2000). Downregulation of this cotransporter in FW relative to SW inbred lines (Fig. 9,

Table 5) under freshwater conditions suggests that this ion transporter might possibly be basolaterally localized in the osmoregulatory epithelial cell (ionocytes), as this cotransporter has been found to be located on either apical of basolateral membranes of ionocytes (Wilson et al. 2000, Marshall et al. 2002, Luquet et al. 2005). When located in the basolateral membrane, the $Na^+/K^+/2Cl^-$ cotransporter transports sodium and chloride ions from the hemolymph into the ionocytes, from which these ions are then excreted into the surrounding environment (Luquet et al. 2005). Thus, $Na^+/K^+/2Cl^-$ might play an important role in sodium excretion under saline, but not under freshwater conditions. On the other hand, this study did not find any difference in expression of V-type H^+ATPase, which differs from previous results in microarray study reported by Lee et al. (2011) who found the tendency of increased expression of B subunit of V-type H^+ATPase in freshwater relative to saline population.

The other relevant sets of genes for freshwater adaptation are those involved in metabolism, especially energy metabolism, since osmoregulation under freshwater conditions is energetically very costly. For example, in teleost fish osmoregulation consumes about 10 – 15 % of total energy budget (Kirschner 1993, Kidder et al. 2006). Many of the genes from this class, such as N-acetylgalactosamine kinase, tyrosine-protein kinase, mitochondrial aldehyde dehydrogenases, dehydrogenases, and phospholipases were all upregulated in FW relative to SW inbred lines across both pairwise comparisons between saline and freshwater inbred lines (Table 4).

GENE EXPRESSION RESPONSE DUE TO ACCLIMATION TO SALINITY

Comparison of gene expression response of the same genotype (inbred line) under different salinities (0 *vs*. 15 PSU) revealed gene expression changes due to acclimation in response to

salinity. Here, four separate comparisons within inbred lines revealed an order of magnitude lower number of DE genes relative to comparisons between saline and freshwater inbred lines, using an FDR cutoff of 0.01. The weakest response was observed in the FW1 inbred line (Fig. 10A), where only 356 genes were DE between freshwater and saline conditions. Relative to FW1, the two SW inbred lines as well as FW2 showed a much stronger plastic response resulting in 670 – 880 DE genes (Fig. 10. B, C, D). Out of these, a total of 209 genes were differentially expressed between freshwater (0 PSU) and saline (15 PSU) conditions in both of the FW inbred lines (FDR < 0.01) (Fig. 11 A). In the SW inbred lines, 287 genes were differentially expressed between freshwater and saline conditions in both lines (Fig. 11 B and D, Tables 8 and 9).

The most interesting GO categories among DE plastic genes in both FW and SW lines were genes involved in protein metabolism, energy metabolism (including kinase, dehydrogenase and hydrolase activity), structural constituents of cuticle, and transmembrane transport (including antiporter activity). Carboxipeptidase, lipase 3 and dipeptidyl peptidase were upregulated in freshwater (0 PSU) relative to saline (15 PSU) conditions (Table 8) in the FW inbred lines, implying higher protein metabolism under freshwater conditions. Na^+/K^+-ATPase subunit alpha showed higher expression in both SW and FW inbred lines under freshwater relative to saline conditions (Table 8 and 10, Fig. 13). This result is in agreement with higher enzymatic activity of Na^+/K^+-ATPase under 0 PSU relative to 15 PSU, observed in freshwater and saline larvae of *E. affinis* (Lee et al. 2011).

Higher expression of Na^+/H^+ exchanger 9B1 and putative ammonium transporter was observed in both FW and SW inbred lines under freshwater conditions relative to saline conditions (Tables 8 and 10). In addition, carbonic anhydrase was also upregulated in SW inbred lines (Table 10) under freshwater conditions. This increase in expression of carbonic anhydrase

and the Na^+/H^+ exchanger due to acclimation to freshwater conditions is in concordance with findings reported by Towle and Weihrauch (2001 and Ito et al. (2013). Thus, the potential mechanism of sodium uptake from dilute media *via* Na^+/ H^+ exchangers (Fig. 1) is supported by elevated expression of carbonic anhydrase (Table 10).

Discussion

While ancestral saline populations of *E. affinis* inhabit saline environments that fluctuate in salinity both temporally and spatially (Lee 1999, Winkler et al. 2008), the derived freshwater populations live in environments that are always low in salinity, with very low concentrations of essential ions. Thus, invasions of freshwater habitats by this copepod presents a major habitat transition that requires rapid adaptation of ionic regulation in order to maintain steep gradients between body fluids and the dilute freshwater environment (Lee et al. 2011, Lee et al. 2012). Results obtained from this study indicate that gene regulatory evolution occurs during these freshwater invasions, and likely contribute to rapid adaptation to the freshwater environment. Consequently, uncovering the changes in the abundance of gene transcript, i.e. gene expression is crucial for understanding the molecular basis of such rapid adaptation.

EVOLUTIONARY SHIFTS IN GENE EXPRESSION ASSOCIATED WITH FRESHWATER INVASIONS

This study demonstrates that expression of many genes underlies freshwater tolerance. In particular, evolutionary shifts in expression of genes involved in transmembrane ionic regulation and metabolism were detected. Upregulation of many genes involved in metabolism confirm the high energy demands for osmoregulation in freshwater environment. Evolutionary shifts in

expression of carbonic anhydrase and Na^+/ H^+ exchanger 9B1 indicate plausible mechanism of Na^+ uptake form dilute environment previously reported to be involved in freshwater tolerance (Kotlyar et al. 2000, Towle and Weihrauch 2001, Tsai et al. 2007, Lee et all. 2011, Ito et al. 2013).

Differentially expressed (DE) genes involved in transmembrane ion transport are of particular interest because ion transport function has been found to evolve in response to salinity change in prior studies (Lee et al. 2011). Here, most notably, under freshwater conditions carbonic anhydrase was upregulated in FW relative to SW inbred lines (Table 4), while Na^+/H^+ exchanger 9B1 showed higher expression in FW2 relative to SW1 inbred line (FDR < 0.01). Na^+/H^+ exchanger 9B1 also showed evidence of higher expression (FDR < 0.01) in FW relative to SW inbred lines under saline conditions (Table 6). These results are consistent with those of Towle and Weihrauch (2001) and Ito et al. (2013) who reported upregulated carbonic anhydrase and Na^+/H^+ exchanger in euryhaline crab and zebra fish gills under freshwater conditions. Higher expression of carbonic anhydrase suggests higher levels of H^+ in cytoplasm of ionocytes, which could be exchanged for Na^+ from freshwater environment. Carbonic anhydrase reversibly catalyzes the conversion of H_2O and CO_2 into H^+ and HCO_3^- (Henry 1996, Towle and Weihrauch 2001) providing a high level of protons required for transmembrane exchange for sodium ions from dilute environments. Protons accumulated through carbonic anhydrase activity can be transported out of the cell by V-type H^+-ATPase, which would generate an electrochemical gradient on apical side and allow Na^+ uptake from dilute environments. Alternative mechanism of sodium uptake from fresh water is *via* Na^+/ H^+ exchanger (Fig. 1), which also requires elevated H^+ levels within ionocytes (Towle and Weihrauch 2001, Ito et al. 2013). Thus, elevated

expression of carbonic anhydrase is consistent with its important role in sodium uptake from dilute media surrounding freshwater populations of *E. affinis*.

Previous gene expression analysis conducted on larvae of *E. affinis* (Lee et al. 2011) revealed upregulation of the proton pump V-type H^+ ATPase and downregulation of the alpha subunit of Na^+/K^+-ATPase in freshwater relative to saline populations. Recently, *in situ* immunolocalization revealed significantly higher V-type H^+-ATPase expression in the Crusalis organs of the freshwater than saline population of *E. affinis* (Gerber et al., In Prep.). V-type H+ ATPase also showed highly differentiated single nucleotide polymorphisms between fresh- and saltwater rainwater killifish populations (Kozak et al. 2013). Based on these prior results, V-type H^+-ATPase was of special interest in this study. V-type H^+-ATPase pumps protons out of cell creating an electrochemical gradient that could be used to energize sodium uptake against a steep gradient from dilute environments (Fig. 1) (Lee et al. 2011). While this study did not find evolution of regulation for this enzyme, this analysis was limited to differential expression in two freshwater inbred lines that underwent many generations of strong genetic drift due to inbreeding. Thus, this study might have missed some relevant genes that might contribute to adaptive regulatory evolution. Alternatively, higher enzyme activity of V type H^+-ATPase observed in copepods (Lee et al. 2011) might not be due to transcriptional evolution, but due to altered translation, post-translational modification or due to direct regulation of V-type H^+-ATPase. Such an inconsistency between higher enzyme activity and unchanged mRNA abundance for V-type H+ ATPase was found in salmon during freshwater acclimation (Bystriansky and Schulte 2011). Finally, results here were generated from adult copepods, and might not be directly comparable to those obtained from larvae by Lee et al. (2011). It is possible that mechanisms of ionic uptake are altered across different copepod life stages, most likely at

metamorphosis, as observed in some freshwater crustaceans (Charmantier 1998, Bianchini and Wood 2007).

GENE EXPRESSION CHANGES DUE TO ACCLIMATION TO SALINITY

The number of differentially expressed genes between salinities (0 *vs*. 15 PSU) for a given inbred line, indicative of acclimation, was 5 – 7 folds lower than the number of putatively adaptive DE genes (FDR < 0.01). This is interesting because some studies have shown greater plastic transcriptome response to osmotic stress than evolutionary response. For example, Whitehead et al. (2012) found about one third of genes being DE in the killifish gills after salinity challenge, while Lv et al. (2013) found ~1700 DE genes in gills of the Japanese blue crab. However, the stronger response in these studies might be due to more extreme salinities (transfer of animals from 32 p.p.t. to 0.1 p.p.t.) that were applied during the experiment. Besides, both of these studies were conducted on the gill tissues, while in this study, RNA was extracted from the whole bodies of copepods, which might somewhat mask the gene expression response from osmoregulatory organs. Developmental acclimation to 5 PSU, under which postmetamorphic parents to next generation juveniles were kept, could also affect response to salinity.

Comparison of gene expression between freshwater (0 PSU) and saline (15 PSU) conditions revealed a few important salinity acclimation pathways, including protein metabolism, energy metabolism, and transmembrane transport (including antiporter activity). Most notably, both FW and SW inbred lines showed upregulation of the Na^+/H^+ exchangers and putative ammonium transporter under freshwater relative to saline conditions. The ammonium transporter, which is involved in transmembrane movement of ammonia (NH_3), might be upregulated because of higher levels of intracellular ammonia under freshwater conditions. This

higher ammonia level might result from aforementioned higher expression of peptidases, which implies higher protein metabolism and possibly higher levels of ammonia. Alternatively ammonium ions (NH_4^+) can compete with K^+ ions for transport by K^+ transporters, including Na^+/K^+ -ATPase (Towle and Hølleland. 1987).

Carbonic anhydrase, which showed evolutionary shifts in expression, specifically higher expression in FW relative to SW inbred line under freshwater conditions, was also affected by salinity in the SW inbred lines (Table 10). Coupled with higher expression of Na^+/H^+ exchanger, this result suggests effective mechanism of sodium uptake from diluted media *via* Na^+/ H^+ exchangers (Fig. 1). Carbonic anhydrase supplies high levels of protons that are then exchanged for sodium from diluted media Na^+/ H^+ exchangers.

Elevated expression of alpha subunit of $Na^{+/}K^+$-ATPase under freshwater relative to saline conditions in both FW and SW inbred lines (Tables 8 and 10, Fig. 13) is somewhat surprising as many researchers reported that activity and gene expression of $Na^{+/}K^+$-ATPase decreases when salinity decreases abruptly (Scott et al. 2004, Whitehead et al. 2012). However, in many organisms enzymatic activity of $Na^{+/}K^+$-ATPase is usually lowest at the level to which they are adapted and increases as salinity decreases (Lin et al. 2004; Kang et al. 2008). Furthermore, Mo and Greenaway (2001) reported two-fold higher $Na^{+/}K^+$-ATPase concentration in crayfish exposed to distilled water relative to saline conditions. Similar patterns were observed in *E. affinis* larvae (Lee et al. 2011), where enzymatic activity of $Na^{+/}K^+$-ATPase tended to be higher under freshwater relative to saline conditions in both freshwater and saline populations. Patterns of $Na^{+/}K^+$-ATPase expression observed in this study fits into model of sodium uptake from dilute media proposed by Kirchner (2004), who suggested that combined higher activity of

both the Na^+/K^+-ATPase and H^+-ATPase is necessary for sodium uptake from freshwater environment.

Intriguingly, chitinase and chitin binding protein were downregulated under freshwater conditions in both FW and SW inbred lines, which might suggest involvement of chitin metabolic process in salinity acclimation (Lv et al. 2013). Chitinase is an enzyme involved in chitin degradation, which in crustaceans is essential for shedding old and forming new exoskeleton during the molting process. However, there is no evidence that downregulation of these two genes could increase the barrier between copepods internal fluids and environment, and thus reduce ionic loss under freshwater conditions.

The greater evolutionary than acclimatory differences in gene expression is consistent with the selection regime where fluctuating conditions would select for greater genetic variation rather than greater "generalist strategies" such as broad tolerance or plasticity (Lee and Gelembiuk 2008). This is also consistent with greater plasticity in the saline rather than the freshwater populations (Fig. 11, A vs. B), and with the idea that colonization of freshwater environment might be associated with genetic assimilation of the optimal phenotype and decrease in plasticity (Lande 2009, 2015).

Conclusions

In summary, this study revealed evolutionary changes in gene expression that might be associated with rapid adaptation to freshwater. These results are only partially in concordance with previous findings on evolutionary shifts in physiological function and suggest plausible candidate genes associated with freshwater invasions. In particular, gene expression analysis suggests that changes in expression of ion uptake transporters (Na^+/H^+ exchangers) and

supporting enzymes (carbonic anhydrase) might be crucial in colonization of the freshwater environment. These findings, however, need to be validated with functional studies in order to verify the role of candidate genes in rapid adaptation to freshwater conditions. Moreover, differential expression analysis cannot distinguish between gene expression changes that are due to *cis*- (near the gene) and *trans*- (elsewhere in the genome) regulation. To overcome this limitation of gene expression analysis, allele specific expression analysis has also been applied (Chapter 3). Distinguishing *cis*- *vs. trans*-regulatory changes is essential in gene expression analysis and is yet to clarify the role that regulatory adaptation plays in invasion success.

ACKNOWLEDGMENTS

This project was supported by National Science Foundation grants NSF DEB-0448827, DEB-0745828, and DEB-1050565 to Carol E. Lee. Graduate student M. Bontrager provided valuable discussions and bioinformatics support. Numerous students assisted with generating the inbred lines over a period of three years, including W. Moss, A. Tritt, B. Metzger, A. Jenstead, K. Lee, J. Hilton, A. Bateman, M. Noll, K. Jung, S. Kahn, K. Bridges, T. Pian, and A. Toy. J. Remfert and K. Lee provided valuable assistance with conducting the common-garden experiments.

LITERATURE CITED

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

Ashburner, M., C. Ball, J. Blake, D. Botstein, H. Butler, J. Cherry, A. Davis, K. Dolinski, S. Dwight, and J. Eppig. 2000. Gene ontology: tool for the unification of biology. Nat. Genet. 25:25–29.

Bairoch, A., L. Bougueleret, S. Altairac, V. Amendolia, A. Auchincloss, G. Argoud-Puy, K. Axelsen, D. Baratin, M. Blatter, and B. Boeckmann. 2009. The universal protein resource (UniProt) 2009. Nucleic Acids Res. 37:D169–D174.

Barnosky, A. D., N. Matzke, S. Tomiya, et al. 2011. Has the Earth's sixth mass extinction already arrived? Nature 471:51–57.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B 57:289–300.

Bianchini, A., and C. M. Wood. 2008. Sodium uptake in different life stages of crustaceans: the water flea *Daphnia magna Strauss*. J. Exp. Biol. 211:539–547.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics*, btu170.

Bottomly, D., N. A. R. Walter, J. E. Hunter, P. Darakjian, S. Kawane, et al. 2011. Evaluating gene expression in C57BL/6J and DBA/2J mouse striatum using RNA-Seq and microarrays. PLoS ONE 6: e17820.

Bourgon, R., R. Gentleman, and W. Huber. 2010. Independent filtering increases detection e7943. doi:10.1371/journal.pone.0007943

Bron, J. E., D. Frisch, E. Goetze, S. C. Johnson, C. E. Lee and G. A. Wyngaard. 2009. Observing copepods through a genomic lens. Frontiers in Zoology 8:22–36.

Bullard, J. H., E. Purdom E, K. D. Hansen, et al. 2010. Evaluation of statistical methods for normalization and differential expression in mRNA-seq experiments. BMC Bioinformatics 11:94–106.

Bystriansky, J. S., and P. M. Schulte. 2011. Changes in gill H+-ATPase and Na+/K+-ATPase expression and activity during freshwater acclimation of Atlantic salmon (Salmo salar). J. Exp. Biol. 214:2435–2442.

Caroll, S. B. 2000. Endless forms: the evolution of gene regulation and morphological diversity. Cell 101:577–580.

Caroll, S. B. 2005. Evolution at two levels: on genes and forms. PLoS Biology 3:1159-1166.

Charmantier, G. 1998. Ontogeny of osmoregulation in crustaceans: a review. Invertebr. Reprod. Dev. 33:177–190.

Charmantier, G., Charmantier-Daures, M. & Towle, D. 2009. in Osmotic and ionic regulation. Cells and animals (ed D H Evans) 165–230. CRC Press.

Chen, G., C. Wang and T. Shi. 2011. Overview of available methods for diverse RNA-Seq data analyses. Sci China Life Sci 54:1121–1128.

Chen, Y., A. T. L. Lun, and G. K. Smyth. 2014. Differential expression analysis of complex RNA-seq experiments using edgeR. In: Statistical Analysis of Next Generation Sequence Data, Somnath Datta and Daniel S Nettleton (eds), Springer, New York.

Cowles, C.R., J. N. Hirschhorn, D. Altshuler and E. S. Lander. 2002. Detection of regulatory variation in mouse genes. Nat. Genet. 32: 432–437.

De Wit P., M. H. Pespeni, J. T. Ladner, *D. J.* Barshis, F. Seneca, H. Jaris, N. Overgaard Therkildsen, M. Morikawa, and S. R. Palumbi. 2012. The simple fool's guide to population genomics via RNA-Seq: an introduction to high-throughput sequencing data analysis. *Molecular Ecology Resources* 12:1058–1067.

Denver, D. R., K. Morris, J. T. Streelman, S. K; Kim, M. Lynch, and W. K. Thomas. 2005. The transcriptional consequences of mutation and natural selection in Caenorhabditis elegans. Nat. Genet. 37: 544–548.

Dillies, M. A., A. Rau, J. Aubert, et al. 2012. A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. Briefings in Bioinformatics doi:10.1093/bib/bbs046.

Fay, J. C., and C. I. Wu. 2000. Hitchhiking under positive Darwinian selection. Genetics 155: 1405–1413.

Fraser, H. B., A. M. Moses, and E. E. Schadt. 2010. Evidence for widespread adaptive evolution of gene expression in budding yeast. PNAS 107:2977–2982.

Fraser, H. B. 2013. Gene expression drives local adaptation in humans. Genome Res. 23:1089–1096.

Fu, X., N. Fu, S. Guo, Z. Yan, Y. Xu, et al. 2009. Estimating accuracy of RNA-Seq and microarrays with proteomics. BMC genomics 10: 161.

Gerber, L., G. Charmantier, M. Charmantier-Daures, and C. E. Lee. Invasions from saline into freshwater habitats: Localization of osmoregulatory sites and of major ion pumps in the swimming legs of the copepod *Eurytemora affinis*. (In Prep.).

Gillis, J., M. Mistry, and P. Pavlidis. 2010. Gene function analysis in complex data sets using ErmineJ. Nature Protocols 6:1148 –1159.

Gillis, J., and P. Pavlidis. 2011. The Impact of Multifunctional Genes on "Guilt by Association" Analysis. PLoS ONE 6(2): e17258. doi:10.1371/journal.pone.0017258

Hardcastle, T.J., and K.A. Kelly. 2010. baySeq: empirical bayesian methods for identifying differential expression in sequence count data. BMC Bioinformatics 11:422–435.

Havird, J. C., S. R. Santos, and R.P. Henry. 2014. Osmoregulation in the Hawaiian anchialine shrimp Halocaridina rubra (Crustacea: Atyidae): expression of ion transporters, mitochondria-rich cell proliferation and hemolymph osmolality during salinity transfers. J. Exp. Biol. 217:2309–2320.

Hayden, E. C. 2012. RNA studies under fire. Nature 484:428.

Henry, R. P., 1996. Multiple roles of carbonic anhydrase in cellular transport and metabolism. Annu. Rev. of Physiol. 58:523–538.

Henry, R. P., Č. Lucu, H. Onken, and D. Weihrauch. 2012. Multiple functions of the crustacean gill : osmotic/ionic regulation, acid-base balance, ammonia excretion, and bioaccumulation of toxic metals. Front Physiol. 3(431):1–34. doi:10.3389/fphys.2012.00431.

Hermisson, J., and P. S. Pennings. 2005. Soft sweeps molecular population genetics of adaptation from standing genetic variation. Genetics 169:2335–2352.

Hoekstra, H. P., and J. A. Coyne, 2007. The locus of evolution: evo devo and the genetics of adaptation. Evolution 61:995–1016.

Hosfeld, B. 1999. Ultrastructure of ionocytes from osmoregulatory integumental windows of Tachidius discipes and Bryocamptus pygmaeus (*Crustacea, Copepoda, Harpacticoida*) with remarks on the homology of nonsensory dorsal organs of crustaceans. Acta Zool. 80:61–74.

Hudson, R. R., M. Kreitman, and M. Aguadé. 1987. A test of neutral molecular evolution based on nucleotide data. Genetics 116:153–159.

Hwang, P. P., and T. H. Lee. 2007. New insights into fish ion regulation and mitochondrion-rich cells. Comp. Biochem. Physiol. A 148:479–497.

Ito, Y., S. Kobayashi , N. Nakamura, H. Miyagi, M. Esaki, K. Hoshijima, S. Hirose. 2013. Close association of carbonic anhydrase (CA2a and CA15a), Na(+)/H(+) exchanger (Nhe3b), and ammonia transporter Rhcg1 in zebrafish ionocytes responsible for Na(+) uptake. Front Physiol. 2013 4:59. doi: 10.3389/fphys.2013.00059. eCollection 2013.

Jensen, M. K., S. S. Madsen, and K. Kristiansen.1998. Osmoregulation andsalinity effects on the expression and activity of Na+K+-ATPase in the gills of European sea bass, Dicentrarchus labrax (L.). J. Exp. Zool. 282:290–300.

Johnson, K. E., L. Perreau, G. Charmantier, M. Charmantier-Daures and C.E. Lee. 2014. Without Gills: Exploring the localization of osmoregulatory function in the copepod *Eurytemora affinis*. Physiol. Bioch. Zoo. 87:310–324.

Kettner, C., A. Bertl, G. Obermeyer, C. Slayman, and H. Bihler. 2003. Electrophysiological Analysis of the Yeast V-Type Proton Pump: Variable Coupling Ratio and Proton Shunt. Biophys J. 85: 3730–3738.

Khaitovich, P., S. Pääbo, and G. Weiss. 2005. Toward a neutral evolutionary model of gene expression. Genetics 170: 929–939.

Khlebovich, V. V., and E. N. Abramova. 2000. Some problems of crustacean taxonomy related to the phenomenon of Horohalinicum. Hydrobiologia 417:109–113.

Kidder, G. W., C. W. Petersen, and R. L. Preston. 2006. Energetics of osmoregulation: I. Oxygen consumption by Fundulus heteroclitus. J. Exp. Zool. A305 309–317.

Kim, Y., and D. Gulisija D. 2010. Signatures of recent directional selection under different models of population expansion during colonization of new selective environments. Genetics. 184:571–585.

King, M.C., and A. C. Wilson. 1975. Evolution at two levels in humans and chimpanzees.

Science 188:107–116.

Kirschner, L. B. 1993. The energetics of osmoregulation in ureotelic and hypoosmotic fishes. J Exp Zool 267:19–26.

Kirschner, L. B. 2004. The mechanism of sodium chloride uptake in hyperregulating aquatic animals. J. Exp. Biol. 207:1439–1452.

Kotlyar, S., D. Weihrauch, R. S. Paulsen, and D. W.Towle. 2000. Expression of arginine kinase enzymaticactivity and mRNA in gills of the euryhalinecrabs Carcinus maenas and Callinectes sapidus. J. Exp. Biol. 203:2395–2404.

Kozak, G. M., R. S. Reid, E. M. Brennan, E. L. Berdan, R. C. Fuller, and A. Whitehead. 2013. Functional and population genomic divergence within and between two species of killifish adapted to different osmotic niches. Evolution 68:63–80.

Kratz, A. and P. Carninci. 2014. The devil in the details of RNA-seq. Nature biotechnology 32:882–884.

Landau, W. M., and P. Liu. 2013. Dispersion Estimation and Its Effect on Test Performance in RNA-seq Data Analysis: A Simulation-Based Comparison of Methods. PLoS ONE 8(12): e81415. doi:10.1371/journal.pone.0081415.

Lande, R. 2009. Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. J. Evo. Biol. 22:1435–1446.

Lande, R. 2015. Evolution of phenotypic plasticity in colonizing species. Mol. Ecol. 24:2038–2045.

Lee, C. E. 1999. Rapid and repeated invasions of fresh water by the saltwater copepod *Eurytemora affinis*. Evolution 53:1423–1434.

Lee, C. E., J. L. Remfert, and G. W. Gelembiuk. 2003. Evolution of physiological tolerance and performance during freshwater invasions. Integr. Comp. Biol. 43:439–449.

Lee, C. E., J. L. Remfert, and Y. M. Chang. 2007. Response to selection and evolvability of invasive populations. Genetica 129:179–192.

Lee, C. E., M. Kiergaard, B. D. Eads, G. W. Gelembiuk, and M. Posavi. 2011. Pumping ions: Rapid parallel evolution of ionic regulation following habitat invasions. *Evolution* 65:2229–2244.

Lee, C. E., M. Posavi, and G. Charmantier. 2012. Rapid evolution of body fluid regulation following independent invasions into freshwater habitats. J. Evol. Biol. 25:625–633.

Lemos B., C. D. Meiklejohn, M. Caceres, and D. L. Hartl. 2005. Rates of divergence in gene expression profiles of primates, mice, and flies: stabilizing selection and variability among functional categories. Evolution 59:126–137.

Lemos, B., L. O. Araripe, and D. L. Hartl. 2008. Polymorphic Y Chromosomes Harbor Cryptic Variation with Manifold Functional Consequences. Science 319:91–93.

Levine, M. T., M. L. Eckert, and D. J. Begun. 2011. Whole-genome expression plasticity across tropical and temperate *Drosophila melanogaster* populations from Eastern Australia. Mol. Biol. Evol. 28:249–256.

Li, B., and C. N. Dewey 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 12:323–338.

Li, S., P. P. Łabaj, P. Zumbo, P. Sykacek, W. Shi, et al. 2014. Detecting and correcting systematic variation in large-scale RNA sequencing data. Nature Biotechnology 32:888–895.

Lopez, J. A., and E. Bohuski. 2007. Total RNA extraction with Trizol Reagent and purification with Qiagen RNeasy Mini Kit. ©DGC, Indiana University.

Lorin-Nebel, C., V. Boulo, C. Bodinier and G. Charmantier. 2006. The $Na^+/K^+/2Cl^-$ cotransporter in the sea bass Dicentrarchus labrax during ontogeny: involvement in osmoregulation. J. Exp. Biol. 209:4908–4922.

Luquet, C.M., D. Weihrauch, M. Senek, and D.W. Towle. 2005. Induction of branchial ion transporter mRNA expression during acclimation to salinity change in the euryhaline crab Chasmagnathus granulatus. J. Exp. Biol. 208:3627–3636.

Lv, J., P. Liu , Y. Wang , B. Gao , Chen P, et al. 2013. Transcriptome qnalysis of *Portunus trituberculatus* in response to salinity stress provides insights into the molecular basis of osmoregulation. PLoS ONE 8: e82155. doi:10.1371/journal.pone.0082155

Marioni, J. C. Mason, S. Mane, M. Stephens, and Y. Gilad. 2008. RNA-Seq: an assessment of technical reproducibility and comparison with gene expression arrays. Genome Res 18: 1509–1517.

Marshall, W. S., E. M. Lynch, and R. R. F. Cozzi. 2002. Redistribution of immunofluorescence of CFTR anion channel and NKCC cotransporter in chloride cells during adaptation of the killifish Fundulus heteroclitus to sea water. J. Exp. Biol. 205:1265–1273.

McCarthy, D. J., Y. Chen and G. K. Smyth GK. 2012. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Research 40:4288–4297.

McDonough, A. A., K. Geering, and R. A. Farley. 1990. The sodium pump needs its beta subunit. FASEB J. 1:1598–605.

McGettigan, P. A. 2013. Transcriptomics in the RNA-seq era Current Opinion in Chemical Biology 17:4–11.

McNamara, J. C. and Faria, S. C. 2012. Evolution of osmoregulatory patterns and gill ion transport mechanisms in the decapod Crustacea: A review. J. Comp. Physiol. B 182:997–1014.

Mo, J. L. and P. Greenaway, 2001. cAMP and sodium transport in freshwater crayfish, Cherax destructor. Comp. Biochem. Physiol. A 129: 843–849.

Mortazavi, A., B. A.Williams, K. McCue, L. Schaeffer and B. Wold. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5:621–628.

Ogasawara, O., and K. Okubo. 2009. On Theoretical Models of Gene Expression Evolution with Random Genetic Drift and Natural Selection. PLoS ONE 4(11): e7943. doi:10.1371/journal.pone.0007943

Oshlack A. and M. J. Wakefield. 2009. Transcript length bias in RNA-seq data confounds systems biology. Biology Direct 4:14 doi:10.1186/1745-6150-4-14.

Pan, Q., O. Shai, L. J. Lee, B. J. Frey, and B. J.Blencowe. 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nature genetics 40:1413–1415.

Parkhomchuk, D., T. Borodina, V. Amstislavskiy, M. Banaru, L. Hallen, S. Krobitsch, H. Lehrach, and A. Soldatov. 2009. Transcriptome analysis by strand-specific sequencing of complementary DNA. Nucleic Acids Res. 37:e123. doi: 10.1093/nar/gkp596.

Perry, S. F. 1997. The chloride cell: structure and function in the gills of freshwater fishes. Ann. Rev. Physiol. 59:325–347.

Petzel, D. H. 2000. Na(+)/H(+) exchange in mosquito Malpighian tubules. Am. J. Physiol. Regul. Integr. Comp. Physiol. 279:R1996–2003.

Pickrell, J. K., J. C. Marioni, A. A. Pai, J. F. Degner, B. E. Engelhardt, E. Nkadori, J. B. Veyrieras, M. Stephens, Y. Gilad, and J. K. Pritchard. 2010. Understanding mechanisms underlying human gene expression variation with RNA sequencing. Nature. 464:768–772.

Pimentel, D., L. Lach, R. Zuniga and D. Morrison. 2000. Environmental and economic costs of nonindigenous species in the United States. *Bioscience* 50:53–65.

Posavi, M., G. W. Gelembiuk, B. Larget, and C. E. Lee. 2014. Testing for beneficial reversal of dominance during salinity shifts in the invasive copepod *Eurytemora affinis*, and implications for the maintenance of genetic variation. Evolution. 68:3166–3183.

Przeworski, M., G. Coop and J. D. Wall. 2005. The signature of positive selection on standing genetic variation. Evolution 59:2312–2323.

R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL *http://www.R-project.org/*

Rheault, M. R., B. A. Okech, S. B. Keen, M. M. Miller, E. A. Meleshkevitch, P. J. Linser, D. Y. Boudko, AND W.R. 2007. Molec.ular cloning, phylogeny and localization of AgNHA1: the first Na^+/H^+ antiporter (NHA) from a metazoan, *Anopheles gambiae*. *J. Exp. Biol.* 210:3848–3861.

Rienmüller, F., I. Dreyer, G. Schönknecht, A. Schulz, K. Schumacher, R. Nagy, E. Martinoia, I. Marten, and R. Hedrich. 2012. Luminal and Cytosolic pH Feedback on Proton Pump

Activity and ATP Affinity of V-type-ATPase from Arabidopsis. Biol. Chem. 287: 8986–8993.

Robinson, M. D., and A. Oshlack A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biology 11, R25.

Robinson, M. D., and G. K. Smyth. 2007. Moderated statistical tests for assessing differences in tag abundance. Bioinformatics 23:2881–2887.

Robinson, M. D., D. J. McCarthy, and G. K. Smyth. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139:140–154.

Robinson, M.D.M., D. J. D. McCarthy and G. K. G Smyth. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26: 139–140.

Roddie, B. D., R. J. G. Leakey, and Berry, A.J. 1984. Salinity temperature tolerance and osmoregulation in *Eurytemora affinis* (Poppe) (Copepoda: Calanoida) in relation to its distributionin the zooplankton of the upper reaches of the Forth Estuary. J. Exp. Mar. Biol. Ecol. 79:191–211.

Ronald J. and J. M. Akey. 2007. The evolution of gene expression QTL in Saccharomyces cerevisiae. PLoS One 2: e678. doi: 10.1371/journal.pone.0000678.

Ronald, J., R. B. Brem, J. Whittle, J. and L. Kruglyak. 2005. Local regulatory variation in *Saccharomyces cerevisiae*. PLoS Genet. 1(e25):213–222.

Russell, J. M. 2000. Sodium-potassium-chloride cotransport. Physiol. Rev. 80:211–276.

Scott, G. R., J. G. Richards, B. Forbush, P. Isenring, and P. M. Schulte. 2004. Changes in gene expression in gills of the euryhaline killifish Fundulus heteroclitus after abrupt salinity transfer. Am J. Physiol. Cell Physiol. 287:C300–C309.

Shetlar, R. E. and D. W. Towle. 1989. Electrogenic sodium-proton exchange in membrane vesicles from crab *Carcinus maenas* gill. Am. J. Physiol. 257:R924–R931.

Sîrbu, A., G. Kerr, M. Crane, and H. J. Ruskin 2012. RNA-Seq vs. dual- and single channel microarray data: sensitivity analysis for differential expression and clustering. PLoS ONE 7: e50986.

Soneson, C., and M. Delorenzi. 2013. A comparison of methods for differential expression analysis of RNA-seq data. BMC Bioinformatics 14:91–108.

Stefflova, K., D. Thybert, M. D. Wilson, I. Streeter, J. Aleksic, P. Karagianni, A. Brazma, D. J. Adams, I. Talianidis, J. C. Marioni, P. Flicek and D. T. Odom. 2013. Cooperativity and Rapid Evolution of Cobound Transcription Factors in Closely Related Mammals. Cell 154:530–540.

Stern, D. L. 2000. Perspective: evolutionary developmental biology and the problem of variation. Evolution 54:1079–1091.

Stern, D.L., and Orgogozo V. 2008. The loci of evolution: how predictable is genetic evolution? Evolution 62:2155–2177.

Stone, J.R., and G, A. Wray, G.A. 2001. Rapid evolution of cis-regulatory sequences via local point mutations. Mol. Biol. Evol. 18:1764–1770.

Swindell, W., M. Huebner, and A. Weber. 2007. Plastic and adaptive gene expression patterns associated with temperature stress in *Arabidopsis thaliana*. Heredity 99:143–150.

Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585–595.

Tarazona, S., F. Garcia-Alcalde, J. Dopazo, A. Ferrer, and A. Conesa. 2011. Differential expression in RNA-seq: A matter of depth. Genome Res. doi/10.1101/gr.124321.111.

Thurman, C., Hanna, J. and Bennett, C. 2010. Ecophenotypic physiology: osmoregulation by fiddler crabs (Uca spp.) from the northern Caribbean in relation to ecological distribution. Mar. Freshw. Behav. Physiol. 43:339–356.

Towle, D. W., and T. Hølleland. 1987. Ammonium ion substitutes for K^+ in ATP-dependent Na^+ transport by basolateral membrane vesicles. Am. J. Physiol. 52:R479–R489.

Towle, D. W., and D. Weihrauch. 2001. Osmoregulation by gills of eury-haline crabs: molecular analysis of transporters. Am. Zool. 41:770–780.

Trapnell, C., A. Roberts, L. Goff, G. Pertea, D. Kim, D. R. Kelley, H. Pimentel, S. L. Salzberg, J. L. Rinn and L. Pachter. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 7:562-578.

Tsai, J. R., and H. C. Lin. 2007. V-type H+-ATPase and Na+,K+-ATPase in the gills of 13 euryhaline crabs during salinity acclimation. J. Exp. Biol. 210:620–627.

Tse, W. K. F., and C. K. C. Wong. 2011. NBCE1 and H^+-ATPase mRNA expression are stimulated in the mitochondria-rich cells of freshwater acclimating Japanese eels (Anguilla japonica). Can. J. Zool. 89:348–355.

Vogel, C., and E. M. Marcotte. 2012. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nat Rev Genet. 13:227–232.

Wakabayashi, S., M. Shigekawa, and J. Pouysségur.1997. Molecular physiology of vertebrate Na+-H+exchangers. Physiol. Rev. 77:51–74.

Wang Y., N. Ghaffari, C. D. Johnson, U. M. Braga-Neto, H. Wang, R. Chen, and H. Zhou. 2011. Evaluation of the coverage and depth of transcriptome by RNA-Seq in chickens**.** BMC Bioinformatics 12(Suppl 10):S5.

Wang, C., B. Gong, P. R. Bushel et al. 2014. The concordance between RNA-seq and microarray data depends on chemical treatment and transcript abundance. Nature Biotechnology 32: 926–932.

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

Weng, X. H., M. Huss, H. Wieczorek and K. W. Beyenbach. 2003. The V-type H^+ -ATPase in Malpighian tubules of *Aedes aegypti*: localization and activity. J. Exp. Biol. 206:2211–2219.

Whitehead, A., J. L. Roach, S. Zhang, and F. Galvez. 2011. Genomic mechanisms of evolved physiological plasticity in killifish distributed along an environmental salinity gradient. Proc.Natl. Acad. Sci. U.S.A. 108:6193–6198.

Whitehead, A., J. L. Roach, S. Zhang, and F. Galvez. 2012. Salinity- and population-dependent genome regulatory response during osmotic acclimation in the killifish (Fundulus heteroclitus) gill. J. Exp. Biol. 215:1293–1305.

Wilson, J. M., D. J. Randall, M. Donowitz, A. W. Vogl, and A. K. Y. Ip. 2000. Immunolocalization of ion transport proteins to branchial epithelium mitochondria-rich cells in the mudskipper (Periophthalmodonschlosseri). J. Exp. Biol. 203:2297–2310.

Winkler, G., J. J. Dodson, and C. E. Lee. 2008. Heterogeneity within the native range: population genetic analyses of sympatric invasive and noninvasive clades of the freshwater invading copepod *Eurytemora affinis*. Mol. Ecol. 17:415–430.

Wittkopp, P. J., and P. Beldade. 2009. Development and evolution of insect pigmentation: Genetic mechanisms and the potential consequences of pleiotropy. Seminars in Cell and Developmental Biology 20:65–71.

Wittkopp, P. J., B. K. Haerum and A. G. Clark. 2008. Regulatory changes underlying expression differences within and between Drosophila species. Nat. Genet. 40: 346–350.

Wittkopp, P. J., B. K. Haerum and Clark, A. G. 2004. Evolutionary changes in *cis* and *trans* gene regulation. Nature 430:85–88.

Wray, G. A. 2007. The evolutionary significance of cis-regulatory mutations. Nat. Rev. Genet. 8:206–216.

Xiang, M. A., P. J. Linser, D. A. Price, and W. R. Harvey. 2012. Localization of two Na+- or K+-H+ antiporters, AgNHA1 and AgNHA2, in Anopheles gambiae larval Malpighian tubules and the functional expression of AgNHA2 in yeast. *Journal of Insect Physiology* 58:570–579.

Zhao, S., W-P. Fung-Leung, A. Bittner, K. Ngo, and X. Liu. 2014. Comparison of RNA-Seq and Microarray in Transcriptome Profiling of Activated T Cells. PLoSONE 9(1): e78644. doi:10.1371/journal.pone.0078644.

TABLES AND FIGURES

Table 1. Summary of Illumina HISeq 2000 RNA sequencing scheme. In total 22 RNA samples were collected from two 2x2 factorial common garden experiments. Four different inbred lines (two saline: SW1 (VA) and SW1 (VE), and two freshwater: FW1 (RA) and FW2 (RB)) were reared each with replicates in both saline (15 PSU) and freshwater (0 PSU) conditions in order to distinguish between environmentally induced and genetically based differences in gene expression. RNA-seq data were generated by pair end (PE) strand specific sequencing on Illumina HiSeq 2000. Each library was first split in half and then samples were run on two lines, multiplexing 12 samples per line.

Sample ID	Inbred line	Salinity (PSU)	Biological replicate	Experiment	Batch	Insert size (bp)	Total Reads (101 bp)
FW1-0-1	FW1	0	1	2	2	399	64,756,320
FW1-0-2		0	2	2	2	406	65,950,514
FW1-15-1	FW1	15	1	2	2	394	67,511,098
FW1-15-2		15	2	2	2	390	65,665,190
FW2-0-1	FW2	0	1	1	1	289	74,424,014
FW2-0-2		0	2	1	1	373	86,526,182
FW2-0-3		0	3	1	3	392	67,495,870
FW2-15-1	FW2	15	1	1	1	308	77,791,450
FW2-15-2		15	2	1	1	262	67,488,764
FW2-15-3		15	3	1	3	403	66,185,938
SW1-0-1	SW1	0	1	1	1	332	83,510,946
SW1-0-2		0	2	1	1	280	65,894,198
SW1-0-3		0	3	1	3	396	54,635,908
SW1-15-1	SW1	15	1	1	1	383	87,755,840
SW1-15-2		15	2	1	1	315	71,000,680
SW1-15-3		15	3	1	3	394	53,779,200
SW2-0-1	SW2	0	1	2	2	398	87,103,524
SW2-0-2		0	2	2	2	388	82,173,766
SW2-0-3		0	3	2	3	357	67,040,052
SW2-15-1	SW2	15	1	2	2	393	54,128,010
SW2-15-2		15	2	2	2	406	57,246,210
SW2-15-3		15	3	2	3	340	87,084,548

Table 2. RNA-seq library size and relative scaling factors (normalization factors) for 22 samples. Normalization factor smaller than one indicates that a large number of fragments correspond to small number of highly expressed genes in a given library (Chen 2014). All the normalization factors are close to one indicating that all 22 libraries are very similar in composition.

Sample ID	Inbred line	Salinity (PSU)	Library size	Normalization factor
FW1-0-1	FW1	0	20,874,833	0.9896
FW1-0-2		0	20,496,791	1.0167
FW1-15-1		15	21,728,439	0.9414
FW1-15-2		15	20,944,384	0.9989
FW2-0-1	FW2	0	22,194,623	0.9511
FW2-0-2		0	25,674,233	1.0317
FW2-0-3		0	19,722,864	0.9868
FW2-15-1		15	23,000,605	1.0554
FW2-15-2		15	20,135,925	0.9240
FW2-15-3		15	19,340,208	1.0056
SW1-0-1	SW1	0	22,832,531	1.0326
SW1-0-2		0	20,493,609	1.0507
SW1-0-3		0	15,956,566	1.0351
SW1-15-1		15	27,774,732	0.9813
SW1-15-2		15	22,643,922	0.9747
SW1-15-3		15	15,738,738	0.9803
SW2-0-1	SW2	0	27,504,119	1.0253
SW2-0-2		0	26,369,310	1.0074
SW2-0-3		0	22,360,048	1.0541
SW2-15-1		15	17,551,630	0.9936
SW2-15-2		15	18,383,208	0.9975
SW2-15-3		15	29,379,775	0.9794

Table 3. Relevant gene ontology (GO) categories and number of differentially expressed genes (FDR corrected P value < 0.01) in FW *vs.* SW inbred lines comparison identified by ErmineJ (Gillies et al. 2010).

GO ID	GO name	No. of genes in GO category**	Multi-functionality*
GO:0016597	amino acid binding	46	0.842
GO:0006022	aminoglycan metabolic process	49	0.759
GO:0016209	antioxidant activity	46	0.105
GO:0070001	aspartic-type peptidase activity	24	0.145
GO:0006816	calcium ion transport	63	0.889
GO:0046942	carboxylic acid transport	85	0.961
GO:0071554	cell wall organization or biogenesis	24	0.445
GO:0009063	cellular amino acid catabolic process	78	0.889
GO:0006873	cellular ion homeostasis	88	0.988
GO:0009200	deoxyribonucleoside triphosphate metabolic process	10	0.825
GO:0051606	detection of stimulus	64	0.89
GO:0015074	DNA integration	21	0.018
GO:0034061	DNA polymerase activity	47	0.068
GO:0005231	excitatory extracellular ligand-gated ion ch. Activity	47	0.642
GO:0008238	exopeptidase activity	87	0.003
GO:0005230	extracellular ligand-gated ion channel activity	57	0.554
GO:0004930	G-protein coupled receptor activity	97	0.335
GO:0030203	glycosaminoglycan metabolic process	38	0.725
GO:0016798	hydrolase activity, acting on glycosyl bonds	88	0.655
GO:0016298	lipase activity	53	0.337
GO:0055065	metal ion homeostasis	90	0.972
GO:0004497	monooxygenase activity	58	0.0335
GO:0052126	movement in host environment	12	0.657
GO:1903510	mucopolysaccharide metabolic process	28	0.668
GO:0007218	neuropeptide signaling pathway	38	0.004
GO:0009124	nucleoside monophosphate biosynthetic process	46	0.906
GO:0015849	organic acid transport	85	0.961
GO:0005343	organic acid:sodium symporter activity	20	0.654
GO:0030072	peptide hormone secretion	13	0.755
GO:0015833	peptide transport	18	0.798
GO:0030247	polysaccharide binding	15	0.075
GO:0000272	polysaccharide catabolic process	32	0.683
GO:0007204	positive regulation of cytosolic Ca ion concentration	20	0.806
GO:0006278	RNA-dependent DNA replication	10	0.18

Table 3. Continued

GO ID	GO name	No. of genes in GO category**	Multi Functionality*
GO:0003964	RNA-directed DNA polymerase activity	27	0.005
GO:0050954	sensory perception of mechanical stimulus	51	0.893
GO:0015081	sodium ion transmembrane transporter activity	79	0.827
GO:0042302	structural constituent of cuticle	14	0.001
GO:0008484	sulfuric ester hydrolase activity	11	0.006
GO:0015293	symporter activity	88	0.707

* Multifuncionality – indicates the degree to which the group is biased towards multifunctional genes (*i.e.*, towards genes grouped into more than one GO category). A score is ranging from 0 to 1, where 1 is the highest (Gillis and Pavlidis, 2011).

** Not all the genes in GO category are differentially expressed.

Table 4. Genes showing evolution of increased expression in freshwater (FW) relative to saline (SW) inbred lines under freshwater conditions (0 PSU).

Gene description	FW1 *vs*. SW2		FW2 *vs.* SW1	
	logFC	FDR	logFC	FDR
Zinc metalloproteinase nas-5	10.7	2.97E-65	8.95	1.13E-83
Xanthine dehydrogenase/oxidase	8.84	1.18E-51	4.67	5.92E-45
Thymidine kinase 2, mitochondrial	4.94	1.15E-24	7.74	1.60E-11
Sodium/calcium exchanger 3	9.42	2.13E-36	2.91	9.87E-26
Carbonic anhydrase	5.46	4.13E-14	4.89	2.85E-10
Carboxypeptidase B	3.16	1.88E-10	6.17	2.27E-58
F-box only protein 40	4.22	1.50E-09	4.73	6.81E-14
Phospholipid-metabolizing enzyme A-C1	5.41	1.03E-25	3.48	1.12E-22
Putative ammonium transporter 1	3.63	1.36E-13	2.75	2.41E-08
Tyrosine-protein kinase shark	3.77	7.88E-38	1.21	9.93E-08
Chymotrypsinogen A	1.75	6.71E-06	3.23	6.35E-44
Selenoprotein W	2.45	7.59E-20	1.75	8.41E-09
Aromatic-L-amino-acid decarboxylase	1.66	4.56E-18	1.78	1.46E-28
Arylsulfatase B	1.55	9.11E-04	1.84	1.15E-09
Glucose-dependent insulinotropic receptor	2.37	3.72E-06	1	8.23E-06
LINE-1 reverse transcriptase homolog	2.42	1.73E-33	0.94	3.39E-03
Potassium voltage-gated channel subfamily B member 2	1.09	1.46E-04	2.12	1.88E-15
Transmembrane protease serine 6	1.48	1.74E-09	1.67	6.08E-20
Arylsulfatase B	1.67	4.09E-06	1.36	2.47E-04
Zinc finger protein 518B	1.48	4.80E-03	1.41	4.48E-03
Protein hedgehog	1.3	1.10E-03	1.54	2.86E-04
Medium-chain specific acyl-CoA dehydr. mt	1.32	1.73E-11	1.42	3.33E-06
Putative acyl-CoA dehydrogenase AidB	1.71	1.20E-15	0.91	3.26E-08
4-hydroxyphenylpyruvate dioxygenase	1.73	2.21E-13	0.77	8.42E-05
Sodium-independent sulfate anion transporter	1.25	1.50E-08	1.19	2.17E-08
Neuronal acetylcholine receptor subunit alpha-6	1.66	4.24E-08	0.73	5.25E-04
Cytosolic 10-formyltetrahydrofolate dehydro.	0.96	1.55E-03	1.41	1.17E-09
Solute carrier family 15 member 1	1.03	1.53E-06	1.07	6.75E-03
Solute carrier family 35 member G1	1.11	1.98E-05	0.98	3.50E-07
Zinc finger protein 728	0.69	7.53E-03	1.39	4.92E-10
Homeobox protein HMX1	1.09	3.42E-04	0.98	1.22E-03
Putative phospholipase B-like 1	1.25	5.99E-04	0.82	8.47E-03

logFC = $\log_2$(expression in FW) - $\log_2$(expression in SW)
FDR = false discovery rate corrected P value

Table 4. Continued

FAD-dependent oxidoreductase domain-protein 1	1.32	1.48E-05	0.67	3.26E-03
Intraflagellar transport protein 22 homolog	1.36	2.65E-05	0.62	9.17E-03
Aldehyde dehydrogenase, mitochondrial	1.1	1.04E-04	0.79	3.78E-04
Succinate-semialdehyde dehydrogenase, mitoch.	0.83	1.81E-04	1.02	6.09E-05
Xanthine dehydrogenase/oxidase	0.98	2.02E-04	0.86	4.09E-06
Aldehyde dehydrogenase X, mitochondrial	1.2	1.36E-03	0.57	5.52E-03
Catalase	0.68	4.18E-04	1.02	4.15E-09
Putative phospholipase B-like 2	1.04	3.99E-07	0.64	2.23E-04
Tripeptidyl-peptidase 2	1.05	8.10E-09	0.62	3.92E-03
D-aspartate oxidase	0.69	3.74E-03	0.97	5.15E-07
Adenylosuccinate lyase	0.67	4.84E-03	0.89	4.33E-05
Isobutyryl-CoA dehydrogenase, mitochondrial	0.76	2.89E-05	0.74	4.61E-07
Putative phospholipase B-like 2	0.79	5.37E-04	0.68	1.04E-03
N-acetylgalactosamine kinase	0.7	5.05E-04	0.72	2.96E-04
Heat shock factor-binding protein 1	0.68	4.25E-03	0.68	9.66E-04
UDP-glucose:glycoprotein glucosyltransferase 1	0.74	1.80E-04	0.58	2.41E-04
Excitatory amino acid transporter 3	0.77	1.43E-04	0.52	5.33E-03
Uridine-cytidine kinase 2-B	0.74	5.35E-03	0.53	2.16E-03
Adenylosuccinate lyase	0.58	6.27E-03	0.58	7.15E-04
NAD(P)H-dependent D-xylose reductase	0.58	2.60E-03	0.57	2.21E-03
Presequence protease, mitochondrial	0.6	9.24E-03	0.48	6.75E-03

Table 5. Genes showing evolution of reduced expression in freshwater (FW) relative to saline (SW) inbred lines under freshwater conditions (0 PSU).

Gene description	FW1 *vs*. SW2		FW2 *vs.* SW1	
	logFC	FDR	logFC	FDR
Zinc finger protein 142	-10.5	2.24E-33	-8.33	2.62E-21
Retrovirus-related Pol polypr. transposon TNT 1-94	-9.47	2.28E-34	-6.47	5.46E-21
Retrovirus-related Pol polyprotein type-2 R2DM	-9.81	4.52E-18	-5.36	7.10E-06
Mechanosensory protein 2	-8.92	8.44E-68	-3.51	1.05E-05
LINE-1 retrotransposable element ORF2 protein	-2.97	3.78E-07	-8.43	2.14E-18
Solute carrier family 12 member 2	-9.17	2.01E-26	-1.34	9.00E-03
Retrovirus-related Pol polyprotein – transposon opus	-4.11	4.12E-33	-3.63	1.36E-33
Pro-Pol polyprotein	-3.72	1.37E-46	-3.55	3.61E-47
Solute carrier family 13 member 2	-3.6	2.68E-11	-2.51	5.05E-11
Nuclear hormone receptor HR96	-3.63	4.49E-09	-1.54	4.43E-05
E3 ubiquitin-protein ligase UHRF1	-2.16	1.10E-10	-2.97	2.73E-33
Heat shock protein HSP 90-alpha	-3.3	1.07E-18	-1.1	7.52E-08
Cyclin-dependent kinase 2-associated protein 1	-2.55	3.84E-11	-1.67	1.26E-12
Probable cytochrome P450 6d4	-1.61	7.74E-18	-2.48	7.57E-27
Vitellogenin	-2.4	6.63E-12	-1.64	2.74E-08
$Na^+/K^+/2Cl^-$ cotransporter	-2.39	4.75E-25	-1.56	2.09E-09
Sodium-dependent phosphate transporter 1-A	-1.61	5.68E-07	-2.33	8.06E-15
Fatty acid-binding protein 9	-2.04	7.37E-12	-1.86	8.10E-10
Carboxypeptidase B	-2.16	1.23E-10	-1.72	9.81E-07
Neuronal acetylcholine receptor subunit alpha-10	-2.17	6.77E-21	-1.65	7.08E-20
Cation transport regulator-like protein 2	-1.36	1.56E-03	-2.36	1.44E-10
Apolipoprotein D	-2.14	2.74E-16	-1.53	4.15E-10
Solute carrier family 13 member 5	-1.21	1.86E-06	-2.44	1.04E-20
Alternative oxidase, mitochondrial	-1.41	2.82E-06	-2.15	1.15E-16
Calmodulin	-2.17	1.32E-25	-1.34	8.95E-09
Cytochrome P450 6k1	-2.29	5.76E-23	-1.17	4.51E-09
Electrogenic sodium bicarbonate cotransporter 1	-1.07	1.47E-04	-2.2	5.55E-26
Mitogen-activated protein kinase 1	-2.57	2.19E-21	-0.65	1.35E-03
Sodium/potassium-ATPase subunit beta-2	-1.13	1.89E-05	-1.93	3.48E-21
Sodium-independent sulfate anion transporter	-2.04	5.63E-14	-0.96	9.22E-07
Zinc finger protein 142	-1.57	1.73E-09	-1.29	4.59E-05
Apoptosis inhibitor IAP	-1.28	2.68E-11	-1.58	1.15E-11

logFC = $\log_2$(expression in FW) - $\log_2$(expression in SW)
FDR = false discovery rate corrected P value

Table 5. Continued

Sodium/calcium exchanger 1	-1	1.93E-07	-1.8	4.15E-17
Transmembrane protein 192	-1.28	4.45E-12	-1.5	4.79E-16
Zinc finger protein 142	-1.3	2.15E-04	-1.41	1.73E-06
Serine proteinase stubble	-1.71	5.65E-08	-0.93	8.14E-05
Sodium/potassium-transporting-ATPase subunit beta	-1.4	9.57E-08	-1.2	1.40E-06
Sodium/potassium-ATPase subunit beta	-1.17	4.38E-05	-1.36	1.53E-06
E3 ubiquitin-protein ligase sina	-1.64	3.96E-06	-0.89	5.71E-04
Heat shock 70 kDa protein	-1.57	6.41E-04	-0.89	6.25E-05
Retrovirus-related Pol polyprotein ransposon 412	-0.91	9.12E-04	-1.55	5.57E-13
Polypeptide N-acetylgalactosaminyltransferase 5	-1.3	7.82E-05	-1.13	8.40E-05
Sodium-independent sulfate anion transporter	-1.69	5.49E-15	-0.61	3.55E-03
Inhibitor of apoptosis protein	-0.97	4.03E-05	-1.28	6.66E-13
Acetyl-coenzyme A synthetase 2	-1.11	1.15E-04	-1.12	1.02E-05
Actin	-1.27	8.08E-05	-0.91	9.82E-03
Trypsin-2	-1.11	3.04E-05	-0.77	3.66E-03
Transmembrane channel-like protein 4	-0.97	8.33E-06	-0.87	4.34E-05
Sodium-driven chloride bicarbonate exchanger	-1.09	5.46E-06	-0.68	4.46E-04
Glyoxylate reductase/hydroxypyruvate reductase	-0.88	1.13E-05	-0.72	1.77E-04
Homeobox protein CDX-2	-0.81	4.21E-03	-0.76	2.19E-03
Mitochondrial sodium/hydrogen exchanger 9B2	-0.91	4.70E-03	-0.61	2.19E-03

Table 6. Genes showing evolution of increased expression in freshwater (FW) relative to saline (SW) inbred lines under saline conditions (15 PSU).

Gene description	FW1 *vs*. SW2		FW2 *vs.* SW1	
	logFC	FDR	logFC	FDR
Zinc metalloproteinase nas-5	11.19	2.29E-68	7.45	2.11E-65
Methylcytosine dioxygenase TET2	5.42	1.99E-49	10.27	2.32E-38
ATP-dependent DNA helicase PIF1	9.72	4.43E-15	4.27	1.37E-06
Sodium/calcium exchanger 3	11.01	4.48E-56	2.08	2.19E-17
Xanthine dehydrogenase/oxidase	7.83	4.34E-35	4.63	7.72E-42
Zinc finger protein 846	1.43	3.13E-07	10.15	3.38E-57
Alanine aminotransferase 2	6.57	7.90E-10	4.87	1.74E-08
Chymotrypsinogen A	7.37	8.52E-14	3.67	8.48E-04
Trypsin	2.29	3.92E-06	8.67	5.41E-14
Thymidine kinase 2, mitochondrial	3.66	9.91E-14	7.25	8.62E-10
Chymotrypsin B	5.39	1.66E-47	5.31	7.35E-39
Retrovirus-related Pol polyprotein from type-2 R2DM	5.55	1.90E-24	2.39	1.25E-12
F-box only protein 40	3.03	1.01E-05	4.72	1.32E-13
Heat shock protein 23	3.88	3.40E-04	3.59	5.31E-09
Retrovirus-related polyprotein transposon opus	2.82	5.60E-13	4.64	3.23E-39
Carboxypeptidase B	2.77	2.69E-07	4.61	1.12E-36
Na+/K+ ATPase subunit alpha	2.65	2.92E-10	3.37	7.86E-17
Sodium/hydrogen exchanger 9B1	1.03	3.41E-03	4.99	5.54E-62
Phospholipid-metabolizing enzyme A-C1	4.05	1.42E-14	1.72	7.77E-07
Kelch-like protein 12	3	3.56E-17	2.48	5.04E-06
Arylsulfatase B	2.21	8.98E-06	2.98	1.08E-20
Epithelial chloride channel protein	2.34	4.18E-31	2.69	2.08E-24
Putative ammonium transporter 1	2.73	4.93E-09	2.29	2.76E-07
Tyrosine-protein kinase shark	3.68	6.50E-32	1.15	5.28E-07
Chymotrypsinogen A	1.65	2.47E-05	2.6	7.95E-28
Arylsulfatase B	2.08	6.26E-08	2.08	5.53E-10
F-box/LRR-repeat protein 7	3.11	7.39E-24	0.88	4.89E-03
Serine proteinase stubble	1.84	9.08E-10	1.95	1.75E-06
Retrovirus-related polyprotein transposon 297	1.79	3.25E-08	1.69	1.19E-03
Retrovirus-related polyprotein transposon 412	1.27	2.15E-03	2.16	1.35E-19
Soluble guanylate cyclase 88E	1.53	9.46E-04	1.75	8.53E-06
Tektin-2	1.57	2.03E-06	1.62	1.24E-09
Aromatic-L-amino-acid decarboxylase	1.61	5.16E-17	1.57	6.42E-23
Retrovirus-related polyprotein transposon 412	1.76	1.38E-07	1.34	2.88E-05

logFC = $\log_2$(expression in FW) - $\log_2$(expression in SW)
FDR = false discovery rate corrected P value

Table 6. Contimued

Putative ascorbate peroxidase	2.1	2.33E-07	0.99	5.84E-03
Potassium voltage-gated channel subfamily B2	1.15	1.10E-04	1.89	2.73E-12
Fatty acyl-CoA reductase 2	1.54	2.14E-05	1.47	2.94E-04
Putative fatty acyl-CoA reductase CG5065	1.51	1.77E-05	1.5	2.93E-06
Sodium-independent sulfate anion transporter	1.36	2.96E-09	1.62	3.73E-14
Endoglucanase E-4	1.23	6.29E-07	1.74	7.65E-10
Sorbitol dehydrogenase	1.56	7.65E-11	1.25	8.54E-08
Alpha-amylase	1.25	1.37E-04	1.54	1.37E-08
Kelch-like protein 12	1.64	4.25E-09	1.02	1.22E-05
Sonic hedgehog protein	1.38	8.51E-03	1.26	7.65E-03
Serine protease 52	1.71	1.13E-09	0.84	6.84E-05
FAD-dependent oxidoreductase domain protein 1	1.46	3.12E-06	0.96	9.41E-06
Fatty acid-binding protein, epidermal	1.54	2.07E-12	0.81	8.91E-05
Neuronal acetylcholine receptor subunit alpha-6	1.5	1.33E-06	0.78	1.87E-04
Carboxypeptidase D	1.12	1.51E-09	1.15	8.79E-10
Probable cytochrome P450 12b2, mitochondrial	1.05	3.49E-05	1.16	3.14E-07
Na-dependent neutral amino acid transp. SLC6A17	0.89	2.29E-03	1.24	2.26E-04
RE1-silencing transcription factor	0.99	1.84E-05	1.05	6.40E-08
Homeobox protein HMX1	0.99	6.17E-03	1.01	2.28E-03
Involucrin	0.95	3.21E-03	0.99	1.52E-03
DNA replication complex GINS protein SLD5	1.27	1.20E-04	0.62	3.29E-03
Probable medium-chain acyl-CoA dehydrogenase, mt.	0.69	1.11E-03	1.16	2.00E-04
DNA replication complex GINS protein SLD5	1.14	3.23E-06	0.6	3.11E-03
Solute carrier family 35 member G1	0.95	3.58E-04	0.77	1.11E-04
D-aspartate oxidase	0.64	8.88E-03	1.06	2.67E-08
Chitin synthase 1	0.88	1.82E-05	0.76	2.65E-06
Mitochondrial basic amino acids transporter	0.98	2.59E-05	0.63	4.13E-05
Mannan-binding lectin serine protease 2	0.74	4.62E-03	0.86	8.02E-04
Zinc finger protein 142	0.98	1.15E-06	0.49	9.69E-03
Titin	0.85	4.96E-04	0.57	5.00E-03
Tripeptidyl-peptidase 2	0.97	2.14E-07	0.43	6.31E-03
N-acetylgalactosamine kinase	0.64	1.88E-03	0.63	2.10E-03
MAP kinase-activated protein kinase 2	0.54	9.85E-03	0.71	9.49E-06
Alpha-ketoglutarate-dependent dioxygenase abh1	0.62	2.00E-03	0.57	1.82E-03
Glutamate--cysteine ligase	0.59	5.61E-03	0.57	9.80E-04

Table 7. Genes showing evolution of reduced expression in freshwater (FW) relative to saline (SW) inbred lines under saline conditions (15 PSU).

Gene description	FW1 *vs.* SW2		FW2 *vs.* SW1	
	logFC	FDR	logFC	FDR
Mechanosensory protein 2	-12.07	2.77E-87	-4.37	2.14E-07
Zinc finger protein 142	-11.05	3.17E-39	-8.04	4.32E-18
Retrovirus-related Pol polyprotein type-2 R2DM	-9.7	4.32E-17	-4.73	4.68E-04
LINE-1 retrotransposable element ORF2 protein	-5.02	1.98E-15	-8.69	2.15E-19
Retrovirus-related Pol polypr. ransposon TNT 1-94	-9.47	6.12E-33	-3.5	3.03E-12
Chymotrypsin-like elastase family member 2A	-3.97	1.77E-05	-8.99	1.60E-111
Glutathione S-transferase	-11.18	2.08E-121	-1.52	8.07E-05
Kelch-like protein diablo	-6.51	3.88E-60	-5.84	5.49E-109
RNA-directed DNA polymerase mob. element jockey	-8.26	2.83E-16	-2.06	2.30E-03
Low choriolytic enzyme	-6.71	1.07E-43	-1.74	1.13E-15
Phospholipid scramblase 2	-4.15	9.81E-36	-3.82	6.33E-62
Receptor-type tyrosine-protein phosphatase eta	-4.19	9.33E-11	-3.3	5.32E-09
Solute carrier family 13 member 2	-3.84	1.98E-09	-2.21	2.21E-08
Nuclear hormone receptor HR96	-4.11	1.20E-22	-1.84	3.97E-10
Solute carrier family 22 member 15	-2.95	1.38E-06	-2.71	9.43E-08
2-keto-3-deoxy-L-fuconate dehydrogenase	-2.61	1.92E-08	-2.95	8.37E-17
Putative calmodulin-like protein 6	-1.71	4.62E-05	-3.37	1.96E-08
Carboxypeptidase B	-1.8	2.45E-08	-3.17	4.62E-23
E3 ubiquitin-protein ligase UHRF1	-2.51	1.34E-13	-2.35	6.55E-21
Cytochrome P450 6k1	-2.91	6.68E-33	-1.6	4.07E-16
Solute carrier family 13 member 5	-1.6	2.71E-10	-2.67	2.25E-24
Glutamate-gated chloride channel	-2.07	5.96E-12	-2.19	3.72E-19
E3 ubiquitin-protein ligase UHRF1	-2.09	4.50E-18	-2.09	8.22E-23
Probable pyridoxine biosynthesis SNZERR	-2.02	2.18E-13	-2.08	3.38E-14
15-hydroxyprostaglandin dehydrogenase NAD(+)	-1.75	1.42E-12	-2.33	2.83E-15
N-acetylglucosamine-1-phosphotransferase α/β	-1.97	3.33E-04	-2.06	1.31E-06
Low-density lipoprotein receptor-related protein 4	-1.84	6.59E-05	-2.17	1.55E-07
Retrovirus-related Pol polypr. from transposon 17.6	-1.99	2.64E-07	-1.96	1.65E-04
Lipase 3	-1.29	1.09E-03	-2.57	1.00E-11
Subversion of eukaryotic traffic protein A	-1.79	2.39E-05	-1.94	9.59E-05
Protein Malvolio	-1.91	3.97E-07	-1.8	1.16E-05
Diaminopropionate ammonia-lyase	-2.06	4.41E-06	-1.54	4.99E-03
Alternative oxidase, mitochondrial	-1.49	5.92E-07	-2.1	9.87E-16
Glucoamylase	-1.66	5.95E-08	-1.86	1.86E-08
Acetylcholine receptor subunit alpha-like	-2.04	1.25E-09	-1.45	3.72E-07
Arylsulfatase J	-1.75	1.61E-04	-1.66	2.87E-04

logFC = $\log_2$(expression in FW) - $\log_2$(expression in SW)
FDR = false discovery rate corrected P value

Table 7. Continued

Protein Malvolio	-1.71	4.33E-04	-1.64	2.36E-07
Electrogenic sodium bicarbonate cotransporter 1	-1.1	1.80E-04	-2.19	7.00E-25
Retrovirus-related Pol polyprotein transposon 412	-1.42	1.49E-07	-1.8	2.20E-16
Xanthine dehydrogenase	-2.09	4.07E-24	-1.11	1.18E-04
Probable calcium-binding protein CML29	-1.95	3.36E-08	-1.23	4.09E-07
Calmodulin	-1.61	4.81E-15	-1.51	4.34E-11
Actin-5C	-1.82	3.86E-16	-1.28	6.97E-06
Alpha-galactosidase	-1.63	3.88E-12	-1.47	1.60E-12
Tropomyosin	-1.85	1.68E-07	-1.24	3.09E-06
Lipase member K	-1.53	2.63E-07	-1.56	1.31E-07
Meiotic recombination protein SPO11-1	-1.87	3.88E-05	-1.19	7.70E-03
Insulin-like growth factor-binding protein complex	-1.91	7.25E-20	-1.08	6.93E-07
Selenium-binding protein 1	-1.31	3.34E-05	-1.67	6.78E-13
Zinc finger protein 142	-1.7	1.13E-10	-1.26	9.44E-05
D-amino acid dehydrogenase 2	-1.75	8.31E-05	-1.13	1.58E-03
Purine nucleoside phosphorylase	-1.81	1.84E-10	-1.05	2.70E-06
Rho GTPase-activating protein 15	-1.75	2.00E-08	-1.1	3.85E-04
Lysosomal alpha-mannosidase	-1.64	1.99E-07	-1.2	1.92E-08
Delta-sarcoglycan	-1.81	2.14E-03	-1.02	3.03E-03
Protein sister of odd and bowel	-2.05	2.51E-13	-0.74	1.66E-03
Solute carrier family 35 member F4	-1.47	2.59E-04	-1.32	1.98E-04
Kelch-like protein 31	-1	5.50E-04	-1.76	3.46E-20
Cell division control protein 42 homolog	-1.72	5.94E-09	-0.94	1.05E-04
Sodium/potassium-transporting-ATPase subunit β	-1.08	2.02E-04	-1.58	2.37E-10
Ankyrin repeat domain protein SOWAHC	-1.74	1.13E-11	-0.89	1.81E-05
Solute carrier family 28 member 3	-1.43	1.41E-08	-1.16	4.38E-04
Probable cytochrome P450 6d4	-1.2	2.84E-10	-1.39	1.50E-09
Xanthine dehydrogenase	-1.19	6.77E-05	-1.18	2.46E-03
Alpha-amylase	-1.17	6.22E-06	-1.19	1.93E-07
Protein O-GlcNAcase	-1.8	1.95E-13	-0.51	2.78E-03
Zinc finger protein 407	-1.38	3.11E-05	-0.92	5.83E-05
Kelch-like protein 10	-1.31	5.86E-11	-0.97	9.41E-07
Cytochrome P450 3A16	-1.00	3.97E-04	-1.15	2.04E-09
Probable D-lactate dehydrogenase, mitochondrial	-0.78	3.42E-05	-1.26	1.10E-14
Zinc finger protein 142	-0.96	1.85E-03	-1.05	3.95E-08
Carbonic anhydrase	-1.06	7.13E-07	-0.94	1.65E-05
Apoptosis inhibitor IAP	-0.73	3.76E-04	-1.07	7.18E-06
Lipase 3	-1.05	3.27E-05	-0.7	4.70E-03
Transcriptional regulator Myc	-0.93	7.48E-05	-0.77	3.84E-03

Table 8. Subset of genes showing acclimation *via* increased expression (FDR corrected P value < 0.01) under freshwater (0 PSU) relative to saline (15 PSU) conditions in FW inbred lines.

Gene description	FW1 [0 vs. 15 PSU]		FW2 [0 vs. 15 PSU]	
	logFC	FDR	logFC	FDR
Vitronectin	1.02	8.85E-06	0.71	2.63E-05
Sodium/potassium-ATPase subunit α	0.94	6.09E-03	1.03	1.24E-05
Calpain-1 catalytic subunit	0.85	5.09E-03	1.16	7.29E-09
Retinol dehydrogenase 12	1.11	1.07E-03	0.94	7.36E-03
Probable cysteine proteinase A494	1.03	5.83E-03	1.16	1.70E-08
Na^+/H^+ exchanger 9B1	1.02	2.90E-05	1.18	3.77E-04
Dual specificity phosphatase DUPD1	1.33	1.35E-04	1.04	4.75E-03
Plasma kallikrein	1.12	1.58E-06	1.32	1.33E-17
Dipeptidyl peptidase 1	1.35	1.02E-03	1.1	2.61E-03
Carboxypeptidase B	1.42	1.66E-06	1.12	6.81E-03
Matrix metalloproteinase-14	1.24	1.68E-04	1.47	4.14E-11
Lipase 3	1.4	3.20E-03	1.35	3.25E-03
Protein-methionine sulfoxide oxidase MICAL2	1.35	1.00E-04	1.5	5.84E-10
Tubulin polyglutamylase TTLL4	1.12	8.85E-05	1.78	6.28E-18
Organic cation transporter 1	1.72	2.35E-07	1.46	4.99E-11
Exoglucanase	1.34	1.68E-04	1.84	3.77E-12
Mitochondrial Na^+/H^+ exchanger 9B2	2.06	2.23E-10	1.35	1.93E-04
Serine protease inhibitor dipetalogastin	1.66	3.54E-10	1.78	1.66E-13
SPARC	1.57	1.37E-04	2.3	6.64E-10
Glutathione peroxidase	1.83	4.00E-07	2.26	3.32E-28
Putative ammonium transporter 1	2.5	3.96E-07	2.73	5.43E-16
Oxidative stress-induced growth inhibitor 1	2.75	1.42E-08	2.7	2.11E-09
Na^+/H^+ exchanger beta	1.41	2.22E-03	4.13	8.39E-32
Sodium/potassium-ATPase subunit beta-1	3.86	7.83E-19	4.19	9.12E-30
Formimidoyltransferase-cyclodeaminase	4.43	5.07E-27	5.25	4.84E-128

logFC = $\log_2$(expression in FW at 0 PSU) - $\log_2$(expression in FW at 15 PSU)
FDR = false discovery rate corrected P value

Table 9. Subset of genes showing acclimation *via* reduced expression (FDR corrected P value < 0.01) in freshwater (0 PSU) relative to saline (15 PSU) conditions in FW inbred lines.

Gene description	FW1 [0 *vs.* 15 PSU]		FW2 [0 *vs.* 15 PSU]	
	logFC	FDR	logFC	FDR
Deoxynucleoside kinase	-4.86	2.68E-06	-4.2	8.09E-22
Ovochymase-1	-3.93	8.28E-16	-4.93	3.71E-64
Kelch repeat-containing protein kel-10	-3.6	5.34E-13	-2.73	7.93E-06
Epididymal secretory protein E1	-2.94	2.17E-13	-2.58	7.40E-13
Serine proteinase stubble	-2.47	5.87E-13	-2.24	4.75E-22
Heme-binding protein 2	-1.91	5.39E-03	-2.73	6.76E-06
Heparan sulfate glucosamine 3-O-sulfotrans. 2	-2.42	5.08E-09	-1.88	8.33E-05
Glutamate-gated chloride channel beta	-2.1	2.34E-08	-1.88	4.15E-10
Kunitz-type protease inhibitor 2	-2.24	2.33E-09	-1.63	8.81E-22
Glutamate-gated chloride channel	-1.84	1.34E-06	-1.97	2.13E-13
Protein rhomboid	-2.17	2.76E-10	-1.64	2.08E-04
Heparan sulfate glucosamine 3-O-sulfotrans. 1	-1.88	2.86E-10	-1.89	1.45E-10
Acidic mammalian chitinase	-1.9	2.52E-09	-1.55	6.12E-08
Niemann-Pick C1 protein	-1.88	4.12E-11	-1.55	1.15E-09
Cytochrome P450 3A16	-1.47	2.78E-07	-1.95	7.27E-15
MFS-type transporter SLC18B1	-1.95	1.39E-03	-1.32	3.21E-03
Vitellogenin	-2.21	1.57E-08	-1.02	5.25E-03
Monosaccharide-sensing protein 1	-1.73	6.96E-03	-1.49	1.14E-07
Papilin	-1.94	8.68E-09	-1.21	8.83E-09
Solute carrier family 12 member 2	-1.98	1.73E-15	-1.11	2.12E-04
Transcriptional repressor scratch 1	-1.46	2.86E-10	-1.6	1.11E-18
Sulfhydryl oxidase 1	-1.33	2.30E-04	-1.45	4.47E-06
Na-dependent phosphate transporter 1-A	-1.21	4.22E-03	-1.42	3.33E-05
Heat shock 70 kDa protein	-1.49	5.39E-04	-1.08	4.62E-06
Neural/ectodermal factor IMP-L2	-0.96	2.31E-03	-1.53	7.45E-07
Sodium/calcium exchanger 3	-1.59	5.01E-08	-0.81	6.24E-03
Serine protease easter	-1.02	2.08E-04	-1.31	3.36E-09
Histidine ammonia-lyase	-1.03	3.59E-03	-1.08	9.31E-08
Fatty acid-binding protein homolog 1	-1.24	1.34E-04	-0.82	2.68E-03
Hemocytin	-1.13	3.85E-04	-0.82	8.42E-05
Delta-1-pyrroline-5-carb. dehydrogenase, mt.	-0.97	1.15E-03	-0.97	7.59E-05
Na-driven chloride bicarbonate exchanger	-0.88	8.79E-04	-0.96	2.60E-07
4-aminobutyrate aminotransferase, mt.	-1.03	2.90E-05	-0.74	1.93E-03
Calcium-activated Cl channel regulator 2	-1.02	2.00E-05	-0.71	5.84E-04
Multidrug resistance-associated protein 1	-0.75	8.61E-03	-0.81	1.06E-04
Na^+/H^+ exchanger 2	-0.72	8.89E-03	-0.64	1.23E-03

logFC = $\log_2$(expression in FW at 0 PSU) - $\log_2$(expression in FW at 15 PSU)
FDR = false discovery rate corrected P value

Table 10. Subset of genes showing acclimation *via* increased expression (FDR corrected P value < 0.01) in freshwater (0 PSU) relative to saline (15 PSU) conditions in SW inbred lines.

	FW1 [0 vs. 15 PSU]		FW2 [0 vs. 15 PSU]	
	logFC	FDR	logFC	FDR
Sodium/potassium-ATPase subunit alpha	7.77	2.21E-69	6.68	1.78E-63
Na^+/H^+ exchanger 9B1	5.16	4.65E-66	3.45	8.63E-35
Fibrillin-2	6.25	2.38E-06	3.05	1.19E-05
Chymotrypsinogen A	4.22	3.76E-04	4.74	6.98E-03
Solute carrier family 12 member 2	3.17	1.39E-23	3.4	2.05E-13
Retinol dehydrogenase 12	2.53	4.60E-14	2.53	1.02E-12
Innexin inx2	2.2	2.51E-04	1.71	7.00E-04
Alpha-amylase	1.59	3.28E-08	2.26	2.68E-17
Serine protease inhibitor dipetalogastin	1.63	2.20E-26	2.01	7.69E-32
Dipeptidyl peptidase 1	1.8	3.37E-06	1.68	3.01E-08
Trypsin	1.87	7.23E-13	1.15	3.33E-03
Nuclear hormone receptor nhr-57	1.84	2.77E-27	1.14	3.61E-08
Putative ammonium transporter 1	1.34	6.93E-11	1.62	1.77E-11
Putative ammonium transporter 1	1.24	7.28E-05	1.64	4.34E-05
Mt. Na^+/H^+ exchanger 9B2	1.3	1.44E-08	1.23	1.02E-06
Organic cation transporter protein	1.24	2.38E-07	1.28	3.47E-07
Organic cation transporter protein	1.35	1.36E-08	1.02	4.20E-05
Retrovirus-rel. Pol transposon gypsy	1.2	7.48E-09	1.16	3.49E-03
Chymotrypsin-like elastase 2A	1.01	8.83E-04	1.31	2.93E-06
E3 ubiquitin-protein ligase sina	0.97	8.53E-04	1.24	1.78E-04
Carbonic anhydrase	1.17	1.43E-03	0.98	8.27E-05
Transmembrane protease serine 12	1.13	5.90E-04	0.94	9.59E-03
5-aminolevulinate synthase, nonspecif. mt.	0.87	1.40E-07	1.16	1.42E-11
L-threonine 3-dehydrogenase, mt.	0.89	2.21E-04	1.07	8.16E-06
Oxidative stress-induced growth inhibitor 1	0.77	4.10E-06	1.18	4.83E-12
Cytochrome P450 2B15	1.2	6.19E-11	0.68	3.48E-04
Cysteine proteinase 3	0.86	1.60E-04	0.97	2.57E-04
Monocarboxylate transporter 12	0.66	3.65E-03	1.15	1.53E-04
Membrane metallo-endopeptidase-like 1	0.65	8.49E-03	0.98	1.18E-06
Transmembrane protein 107	0.75	2.09E-04	0.83	2.92E-04
Methylenetetrahydrofolate reductase	0.7	1.20E-04	0.86	1.87E-04
Phosphoglycolate phosphatase	0.57	4.39E-03	0.73	2.46E-03

logFC = $\log_2$(expression in SW at 0 PSU) - $\log_2$(expression in SW at 15 PSU)
FDR = false discovery rate corrected P value

Table 11. Subset of genes showing acclimation *via* reduced expression (FDR corrected P value < 0.01) in freshwater (0 PSU) relative to saline (15 PSU) conditions in SW inbred lines.

Gene description	SW1 [0 vs. 15 PSU]		SW2 [0 vs. 15 PSU]	
	logFC	FDR	logFC	FDR
Kelch repeat-containing protein kel-10	-6.4	5.39E-22	-8.8	9.96E-54
Peptidyl-Asp metalloendopeptidase	-7.18	2.56E-10	-7.89	3.76E-05
High choriolytic enzyme 2	-7.02	1.06E-92	-3.61	3.88E-13
Deoxynucleoside kinase	-4.22	2.29E-21	-6.06	2.04E-14
Heme-binding protein 2	-4.34	1.52E-11	-3.11	6.71E-11
Electroneutral sodium bicarbonate exchanger 1	-1.97	6.44E-04	-3.76	1.40E-27
Gastric triacylglycerol lipase	-3.24	2.96E-18	-2.17	8.23E-19
Chymotrypsin B	-2.78	1.69E-03	-1.95	7.62E-03
Nuclear hormone receptor HR96	-1.86	3.56E-09	-2.24	3.93E-15
Nuclear hormone receptor HR96	-1.69	4.90E-08	-2.28	1.55E-24
Monosaccharide-sensing protein 1	-1.62	7.96E-09	-1.99	5.03E-06
Putative phospholipase B-like 1	-1.93	7.40E-12	-1.62	2.64E-07
Glutamate-gated chloride channel	-1.81	2.57E-11	-1.58	7.17E-08
Glutamate-gated chloride channel subunit beta	-1.87	6.95E-10	-1.42	2.83E-06
Selenium-binding protein 1	-1.13	1.32E-05	-2.11	1.44E-15
Inward rectifier potassium channel 2	-1.77	4.03E-06	-1.44	5.09E-07
Putative phospholipase B-like 2	-1.48	1.07E-05	-1.55	8.59E-12
Inward rectifier potassium channel 4	-1.44	2.49E-05	-1.3	2.29E-07
Putative phosphoenolpyruvate synthase	-1.45	3.41E-08	-1.28	3.16E-06
Histone-lysine N-methyltransferase SETD7	-1.35	9.55E-03	-1.23	2.74E-04
Tubulin polyglutamylase TTLL4	-1.23	1.01E-09	-1.16	8.49E-04
Transcriptional repressor scratch 1	-1.11	3.59E-09	-1.14	5.69E-10
Cytochrome b5-related protein	-0.74	7.16E-04	-1.49	6.81E-13
Transmembrane protease serine 6	-1.21	1.14E-09	-1	3.94E-05
Zinc metalloproteinase nas-14	-1.36	5.24E-05	-0.84	2.43E-05
Cytochrome P450 3A16	-1.37	1.30E-07	-0.82	7.26E-04
G-protein coupled receptor GRL101	-1.4	7.28E-07	-0.65	2.56E-03
Serine protease easter	-1.06	4.11E-06	-0.94	5.58E-06
Putative phosphoenolpyruvate synthase	-1.01	1.56E-03	-0.97	2.92E-04
Histidine ammonia-lyase	-0.89	2.71E-05	-1.09	1.22E-05
Sodium-driven chloride bicarbonate exchanger	-0.78	4.40E-05	-1.1	1.19E-09
Anion exchange protein 2	-0.88	9.30E-06	-0.74	6.99E-05
Probable chitinase 3	-0.75	6.63E-03	-0.79	3.48E-05
Na and Cl-dependent glycine transporter 2	-0.79	2.63E-03	-0.7	4.96E-03
A-kinase anchor protein 9	-0.64	6.42E-03	-0.81	2.58E-03
Carboxypeptidase N catalytic chain	-0.7	1.73E-03	-0.67	5.60E-03

logFC = $\log_2$(expression in SW at 0 PSU) - $\log_2$(expression in SW at 15 PSU)
FDR = false discovery rate corrected P value

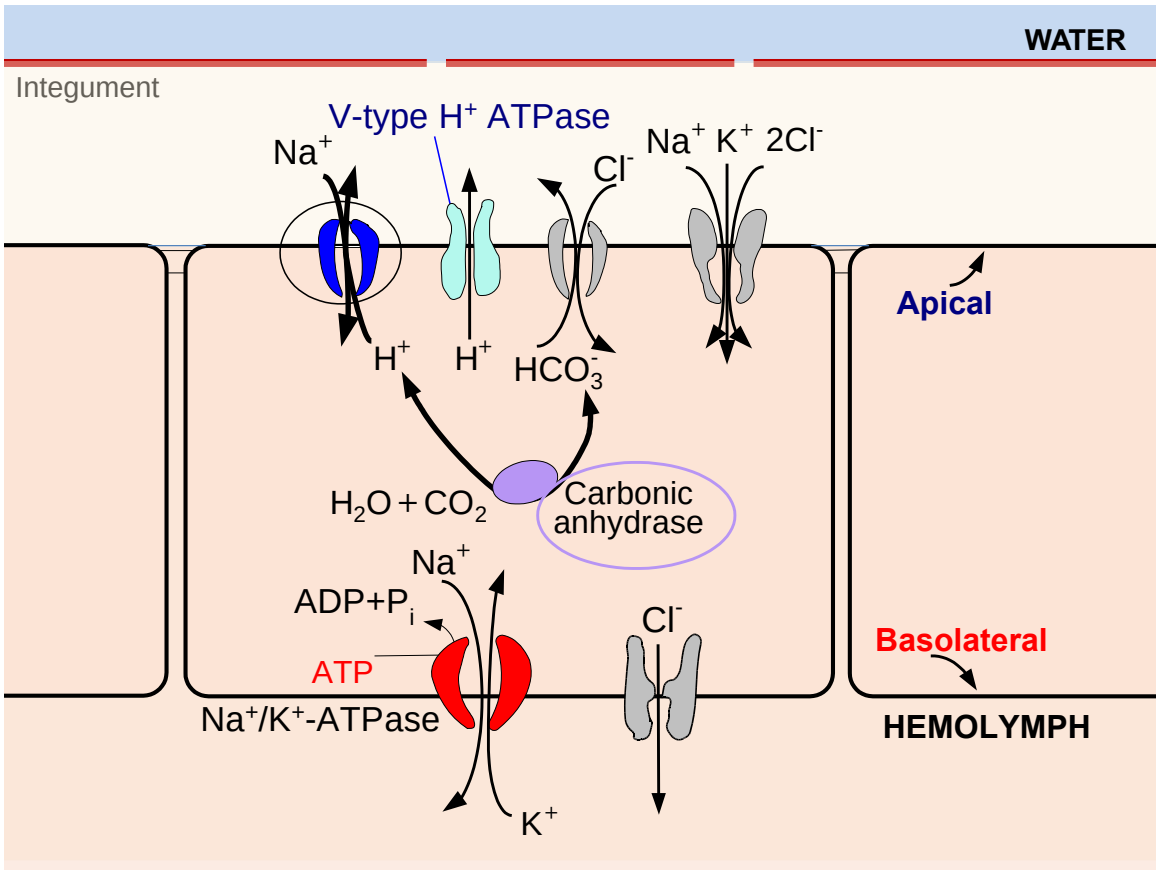

Figure 1. A simplified hypothetical model of ion uptake across epithelial tissue (adapted from Lee et al. 2011 and Towle and Weihrauch, 2001). A suite of transmembrane transporters and supporting enzymes (Na^+/K^+-ATPase, V-type H^+-ATPase, carbonic anhydrase, $Cl^-/HCO3^-$ exchanger, $Na^+/K^+/2Cl^-$ cotransporter, and Na^+/H^+ exchanger) might be involved in ionic regulation in crustaceans. Location of ion transport enzymes (apical *vs.*basolateral) might be crucial for effective osmoregulation. ATP levels needed for higher activity of ion transport activity in freshwater environment might be maintained by arginine kinase. Under saline conditions high ionic concentration of water relative to that of the intracellular fluid allows diffusion of Na^+ into the cell. Na^+/K^+-ATPase (red), located on basolateral membrane than uptakes the Na+ ions. On the contrary, ionic concentration of the freshwater is orders of magnitude lower than that of saline. The crucial transporter and enzyme, involved in osmoregulation, might be Na^+/H^+ exchanger (blue) and V-type H^+-ATPase (teal).

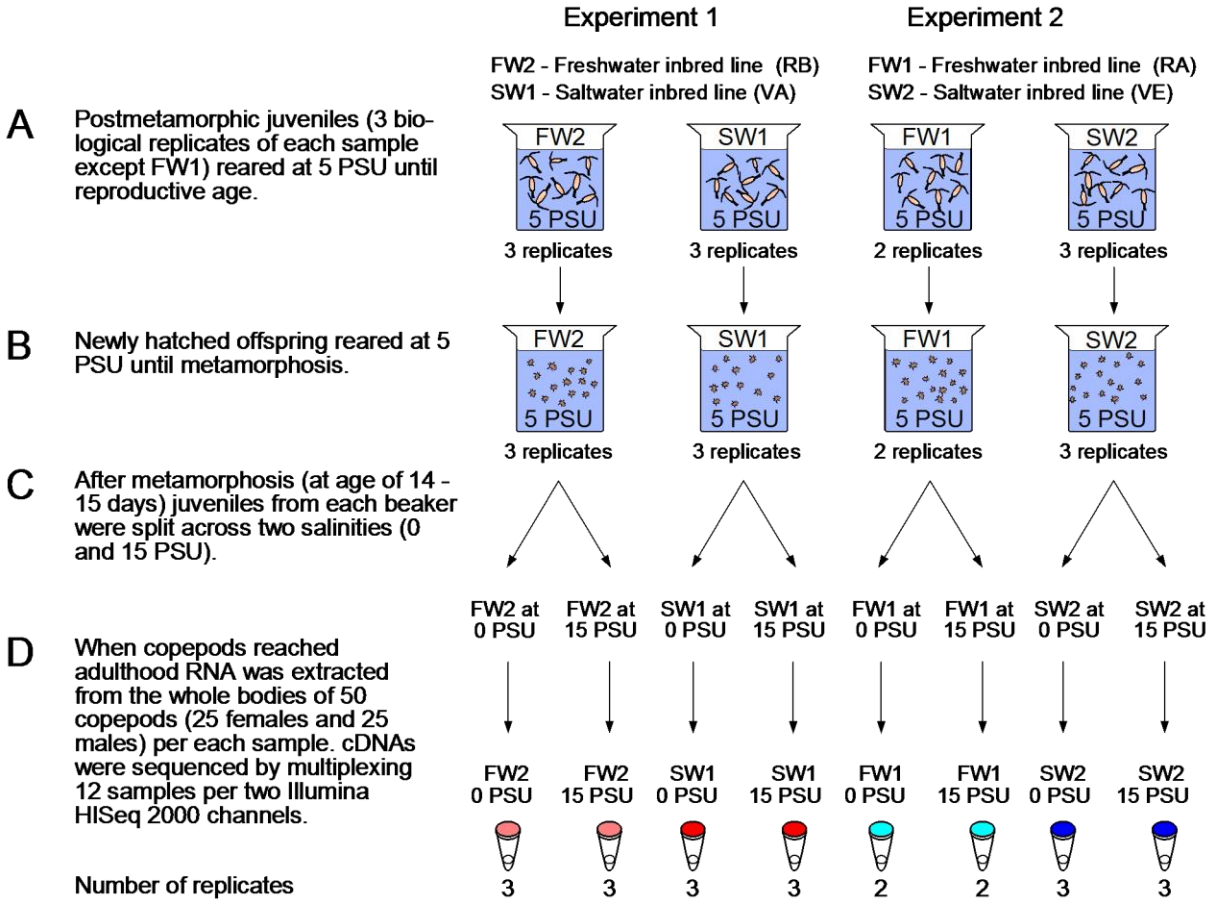

Figure 2. A common garden experiment scheme that allowed distinguishing between environmentally induced and genetically based differences in gene expression between saline and freshwater inbred lines. **(A)** Juveniles from all four saline and freshwater parental inbred lines (FW1, FW2, SW1 and SW2) were gradually transferred to a common salinity of 5 PSU, and reared at this salinity until they became sexually mature. **(B)** The newly produced offspring was separated from parents and reared at 5 PSU until metamorphosis (~15 days of age) **(C)** When offspring reached metamorphosis each sample was split across two salinities (0 and 15 PSU). (**D**) Juveniles were reared at final salinities (either 0 or 15 PSU) for the next 16-18 days i.e., until they became adults. The total of 50 adult copepods (25 females and 25 males) were randomly selected from each sample for total RNA extraction. Samples were sequenced in three batches allowing only FW1 *vs*. SW2 and FW2 *vs*. SW1 pairwise comparisons.

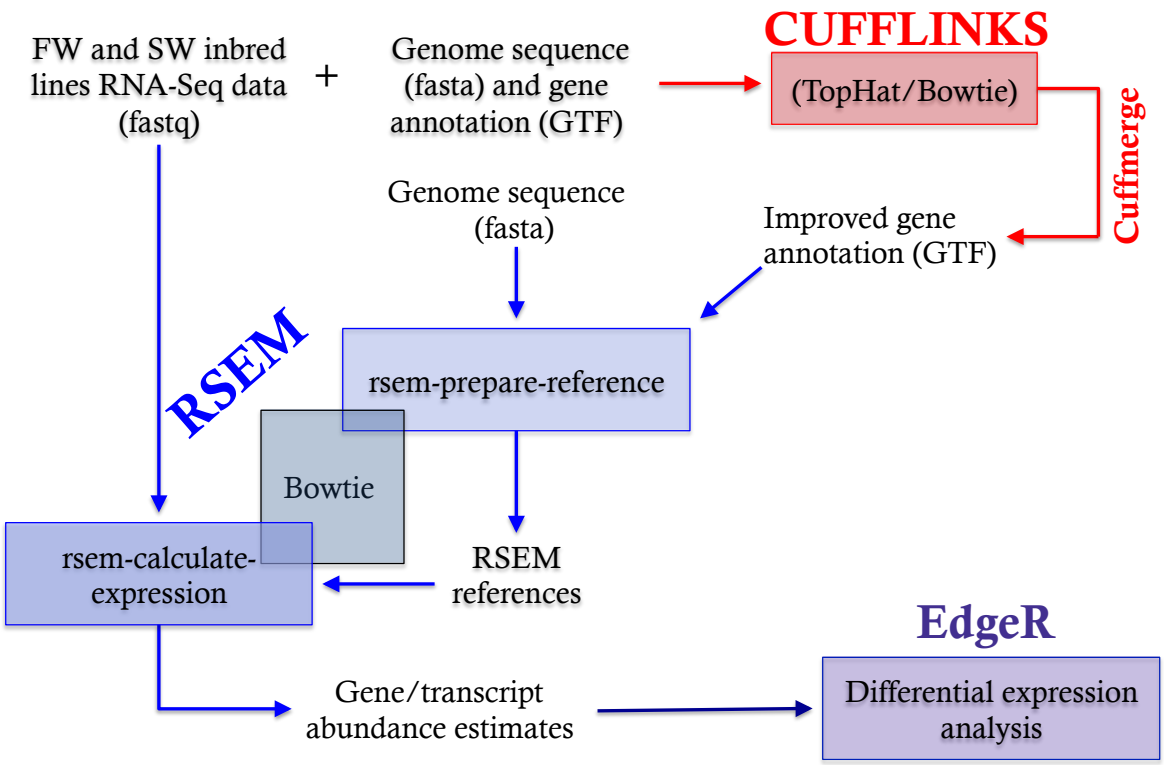

Figure 3. An overview of a protocol used to detect DE genes. RNA-seq data, complete genome sequence (fasta format), and gene annotation (GTF file) were used as input to TopHat which uses Bowtie as alignment engine. Mapped reads were than assembled in Cufflinks (Trapnell et. 2012) producing transcript assemblies for all 22 samples collected from four inbred lines. In the next step, Cuffmerge utility was used to merge all 22 transcript assemblies with original gene annotation to produce single improved GTF file. This improved GTF file was then used as input to RSEM (Li and Dewey 2011) to generate reference transcript sequences and to estimate transcript abundances at gene and isoform-level. RSEM abundance estimates (estimated number of fragments for a given gene or isoform) were used to test for differential expression applying Generalized Linear Models (GLM) framework using Bioconductor Package EdgeR (McCarthy et al. 2012). Employed statistical model accommodated for complex design of common garden experiment.

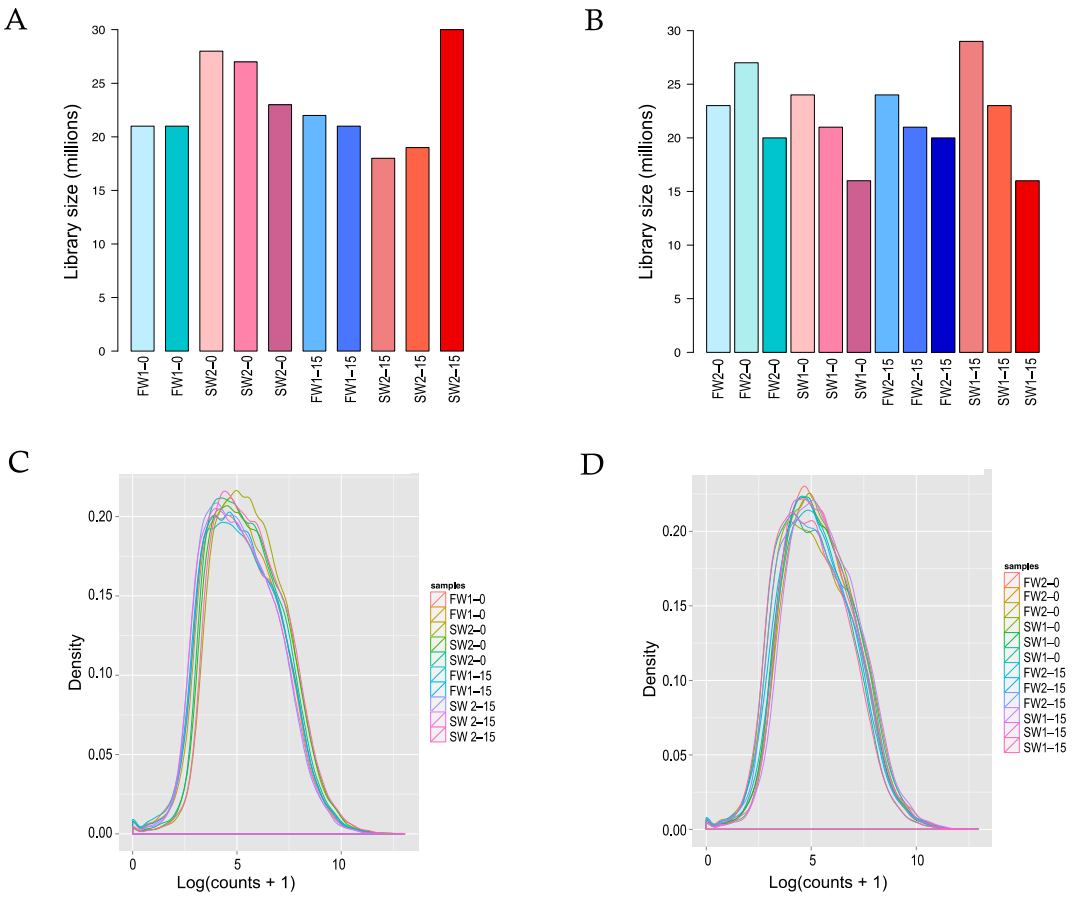

Figure 4. Library size variation and read count distributions (density) across 22 samples of RNA-seq data. The raw read count histograms showing considerable sample-to-sample variation in sequencing depth (library size) in (**A**) FW1 and SW2 and (**B**) FW2 and SW1 inbred lines. Lower panel is showing density function of the log of read counts for (**C**) FW1 and SW2 and (**D**) FW2 and SW1 inbred lines.

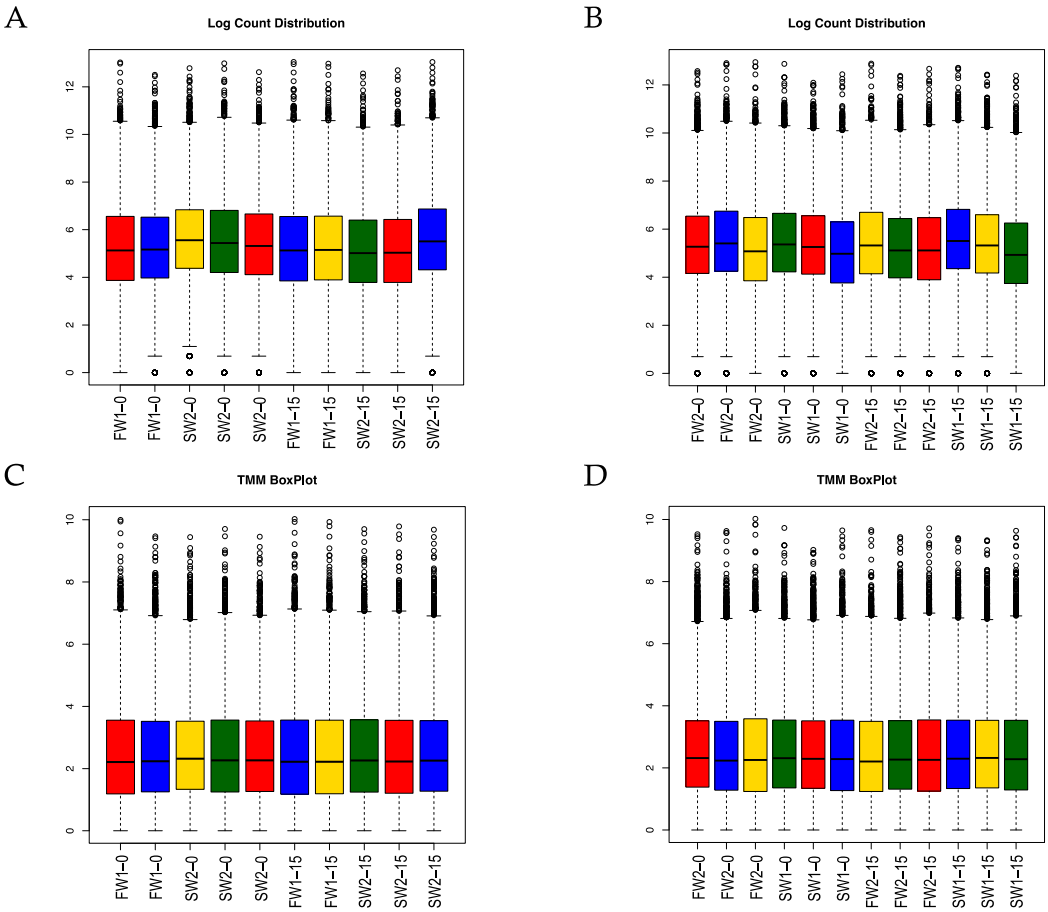

Figure 5. Distribution of raw (**A** and **B**) and TMM (Trimmed Mean of M values) normalized (**C** and **D**) log counts. Box plots of raw read counts show considerable sample-to-sample variation in counts distribution. This bias was corrected in *edgeR* differentially expression analysis by performing TMM normalization. After TMM normalization (**C** and **D**) all the sample count means and all the sample count distributions are almost aligned. This stabilization of read count distributions across samples indicates that TMM provided an effective normalization.

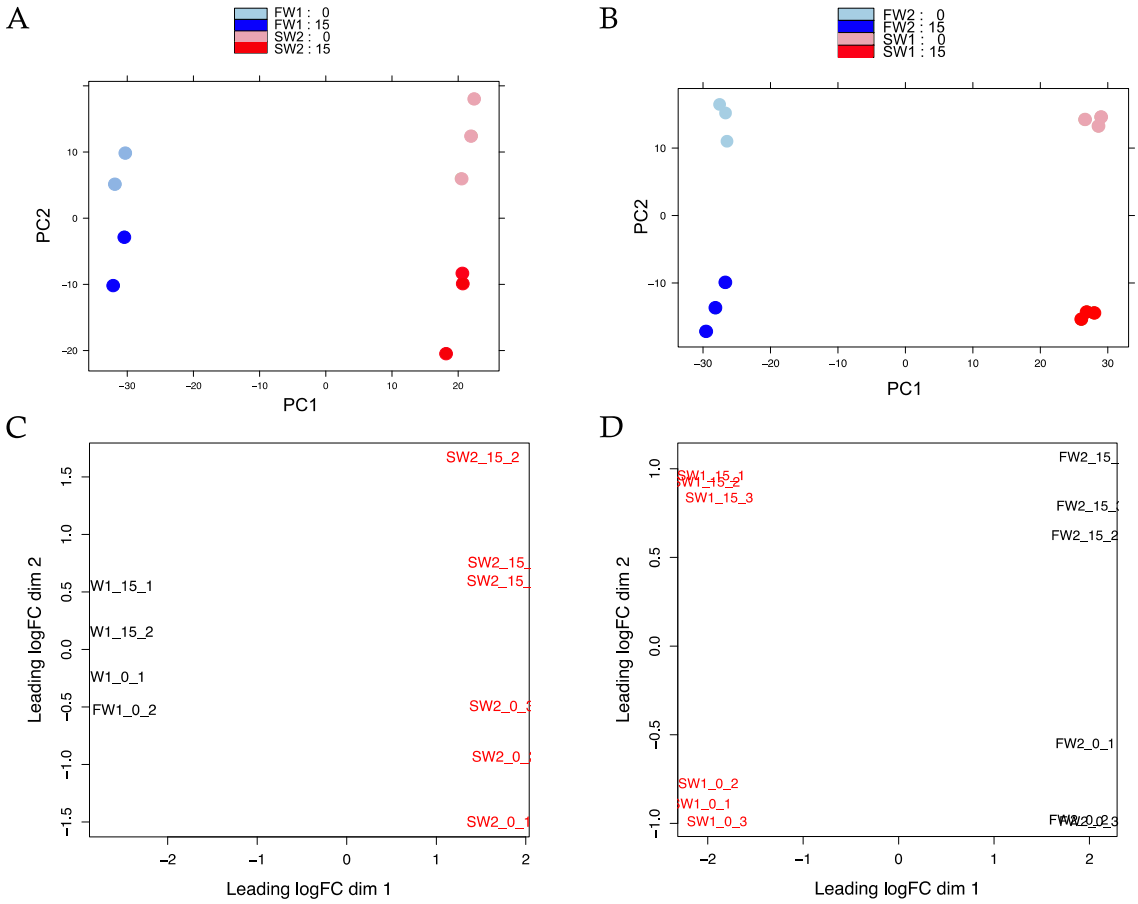

Figure 6. The 22 samples shown in n principle component analysis (PCA) plot (**A** and **B**), and in two-dimensional scaling (MDS) plot (**C** and **D**). Samples are well separated by the genotype (inbred line) in the first dimension (SW *vs.* FW) on both plots. A separation between salinity conditions (0 PSU vs. 15 PSU), in second dimension, is also obvious for all the inbred lines except for FW1.

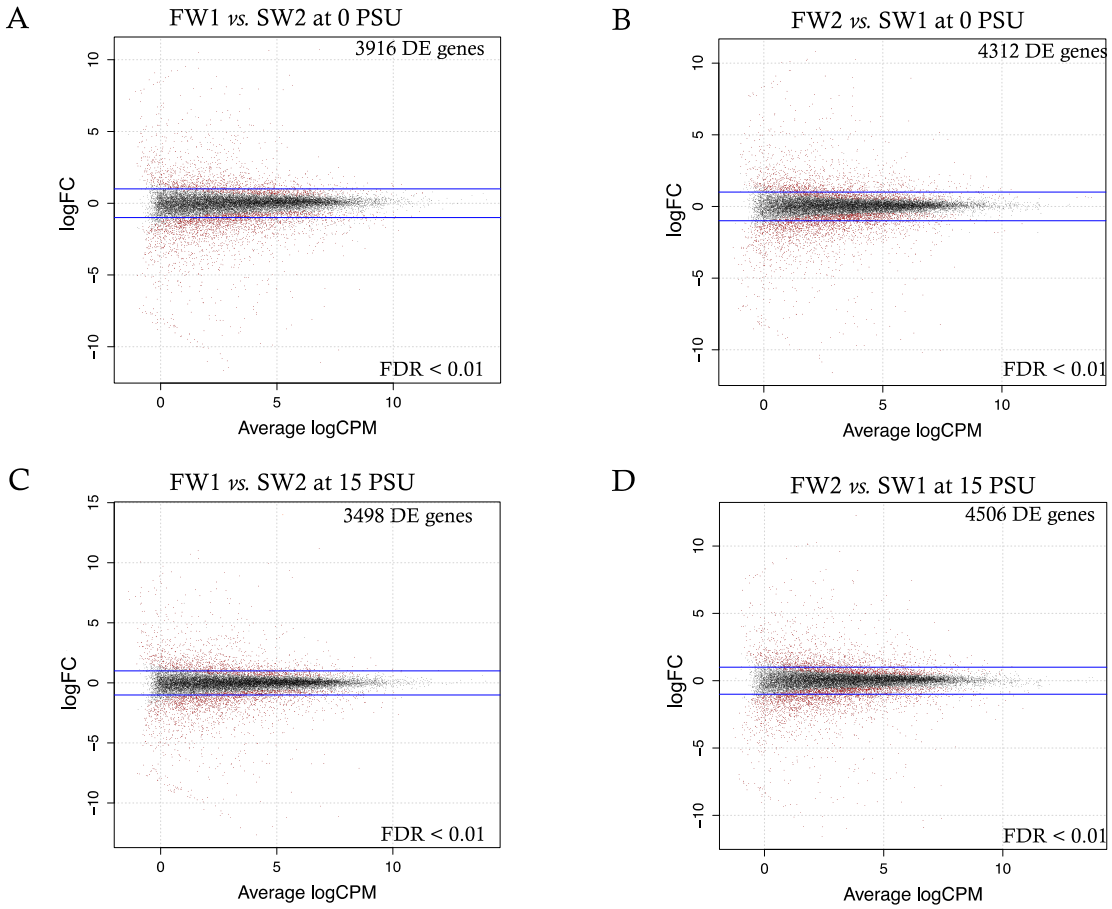

Figure 7. Differentially expressed genes (FDR < 0.01) in pairwise comparisons between fresh (FW) and saline (SW) inbred lines (**A** and **B**) under freshwater (0 PSU) and (**C** and **D**) under saline (15 PSU) conditions. Differentially expressed genes are highlighted, and horizontal blue lines indicate 2-fold changes.

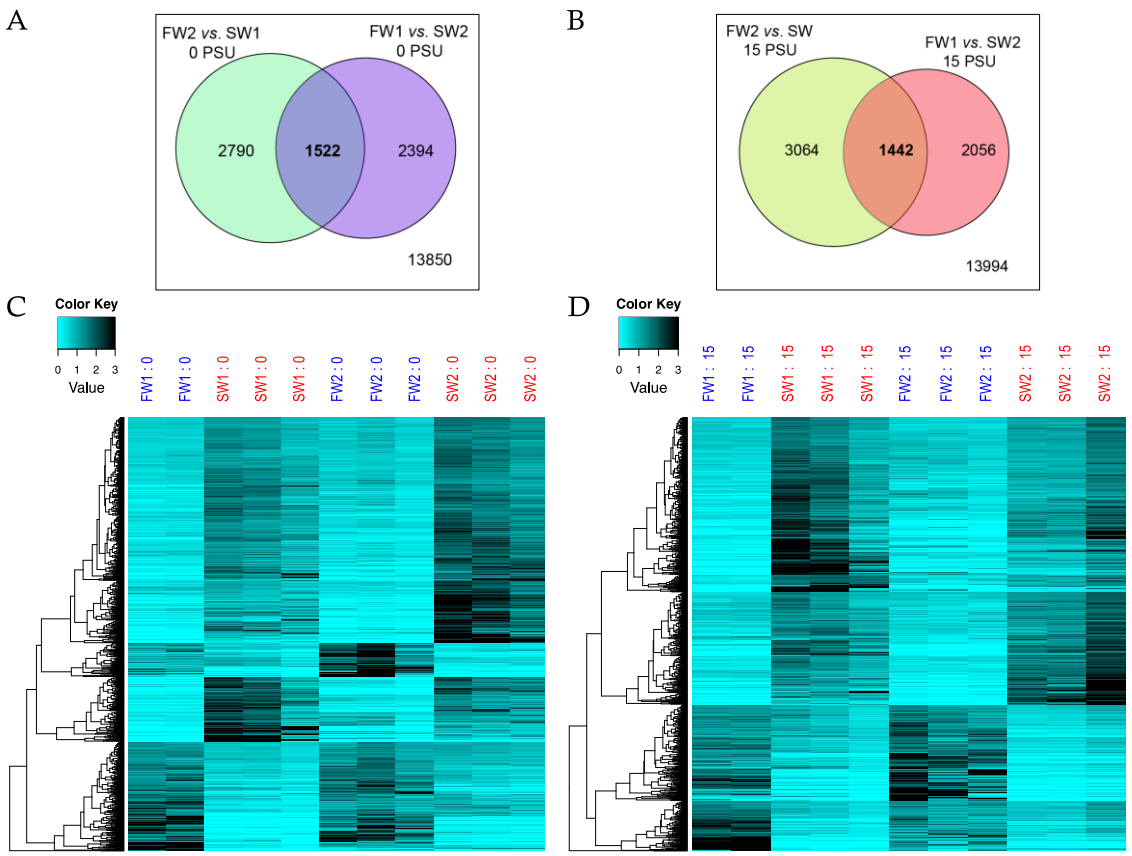

Figure 8. Venn's diagrams showing DE genes (FDR < 0.01) in two separate pairwise comparisons between freshwater (FW1 and FW2) and saline (SW1 and SW2) inbred lines under (**A**) freshwater (0 PSU) and (**B**) saline (15 PSU) conditions. Out of total 20566 genes 1522 were differentially expressed (DE) in both pairwise comparisons under freshwater (0 PSU) conditions (**A**). Similarly, under saline (15 PSU) conditions (**B**) 1422 genes were DE in both pairwise comparisons. Heat map of relative fold differences of genes that were differentially expressed in both pairwise FW versus SW inbred lines comparisons (FDR<0.01) under freshwater (**C**) and saline (**D**) conditions. For heat maps each gene count is normalized to the mean counts for that gene across all the samples.

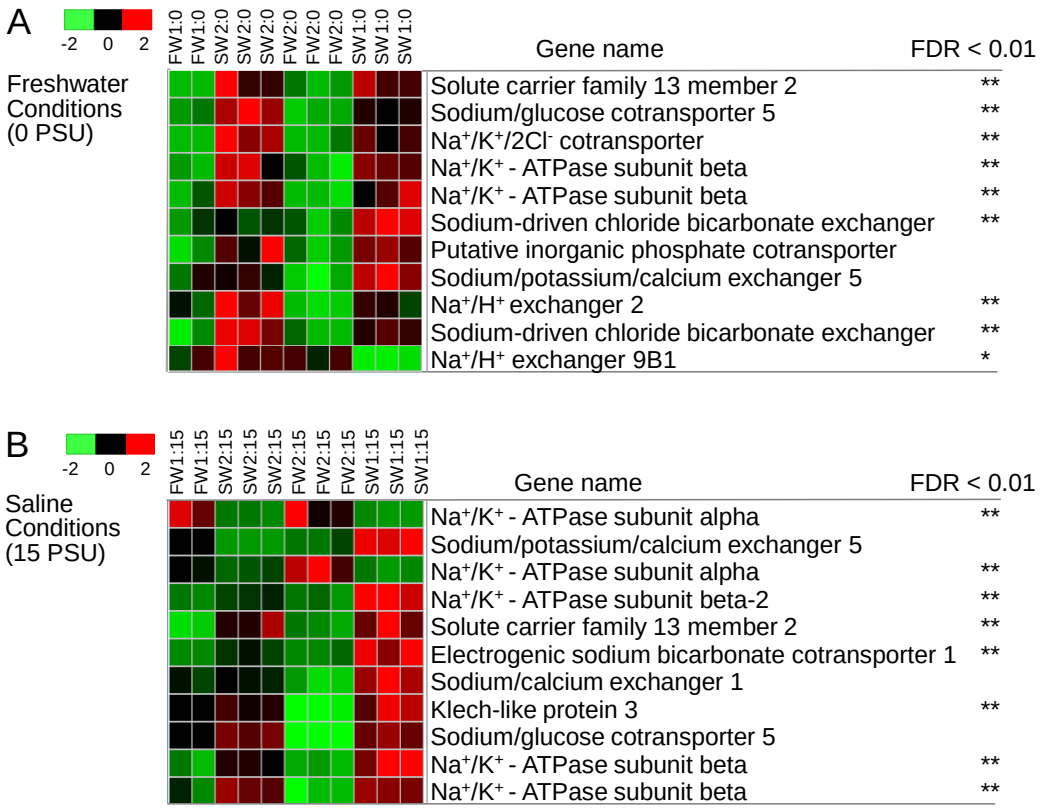

Figure 9. Heat map of TMM normalized gene counts of subset of genes that were grouped into GO category — sodium ion transport (containing 60 genes in total) under (**A**) freshwater and (**B**) saline conditions. Although GO — sodium ion transport was significantly enriched (FDR corrected P value = 0.00036) only 8 genes were differentially expressed between FW and SW inbred lines in freshwater and saline conditions, respectively. Genes that were differentially expressed (based on FDR < 0.01 from negative binomial GLM) across both FW1 *vs*. SW2 and FW2 *vs*. SW1 inbred lines comparisons are denoted by **, while genes that were DE in only one comparison are denoted by *.

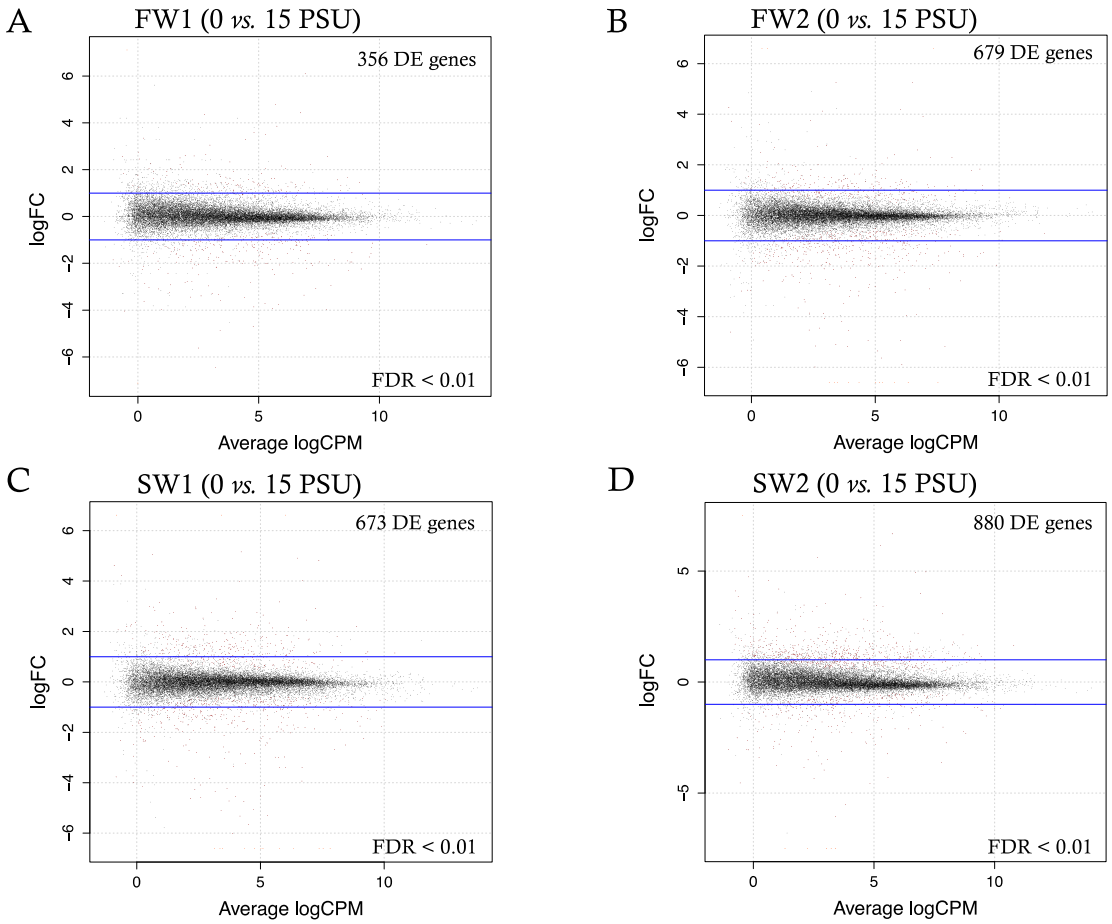

Figure 10. A smear plots showing DE genes (FDR < 0.01) between freshwater (0 PSU) and saline (15 PSU) conditions for two freshwater (A and B) and two saline (C and D) inbred lines. Differentially expressed genes are highlighted and blue lines indicate 2-fold changes in gene expression.

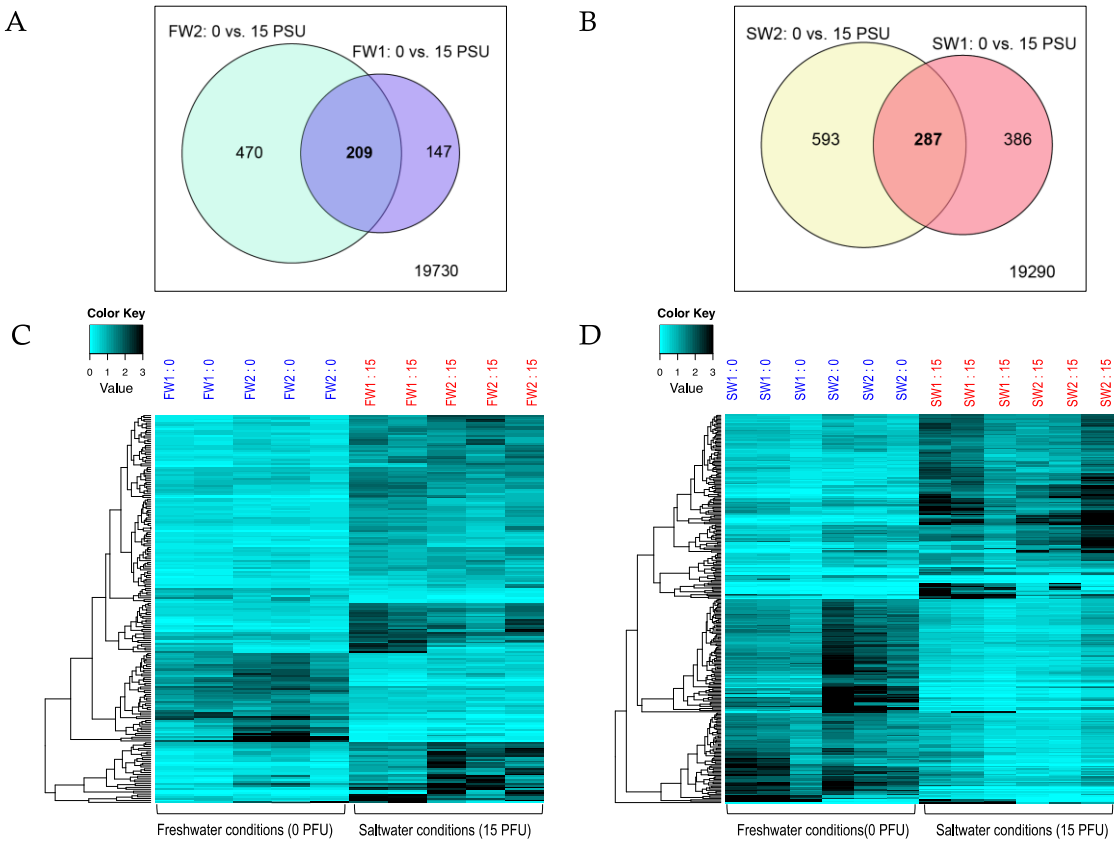

Figure 11. Venn's diagrams showing DE genes due to rearing salinity (aacclimation). (**A**) Comparison between freshwater (0 PSU) and saline conditions (15 PSU) reveled 456 and 697 DE genes in FW1 and FW2 inbred line, respectively, with 209 DE genes in both inbred lines. (**B**) When saline inbred lines were compared under 0 PSU *vs*. 15 PSU then 287 genes were DE in both SW inbred lines. (**C**) Heat map showing relative fold differences of 209 genes that were differentially expressed in freshwater inbred lines in comparison between freshwater (0 PSU) and saline (15 PSU) conditions. (**D**) Heat map of 287 genes that were differentially expressed between freshwater (0 PSU) and saline (15 PSU) conditions in SW inbred lines. Each gene count is normalized to the mean counts for that gene across all the samples.

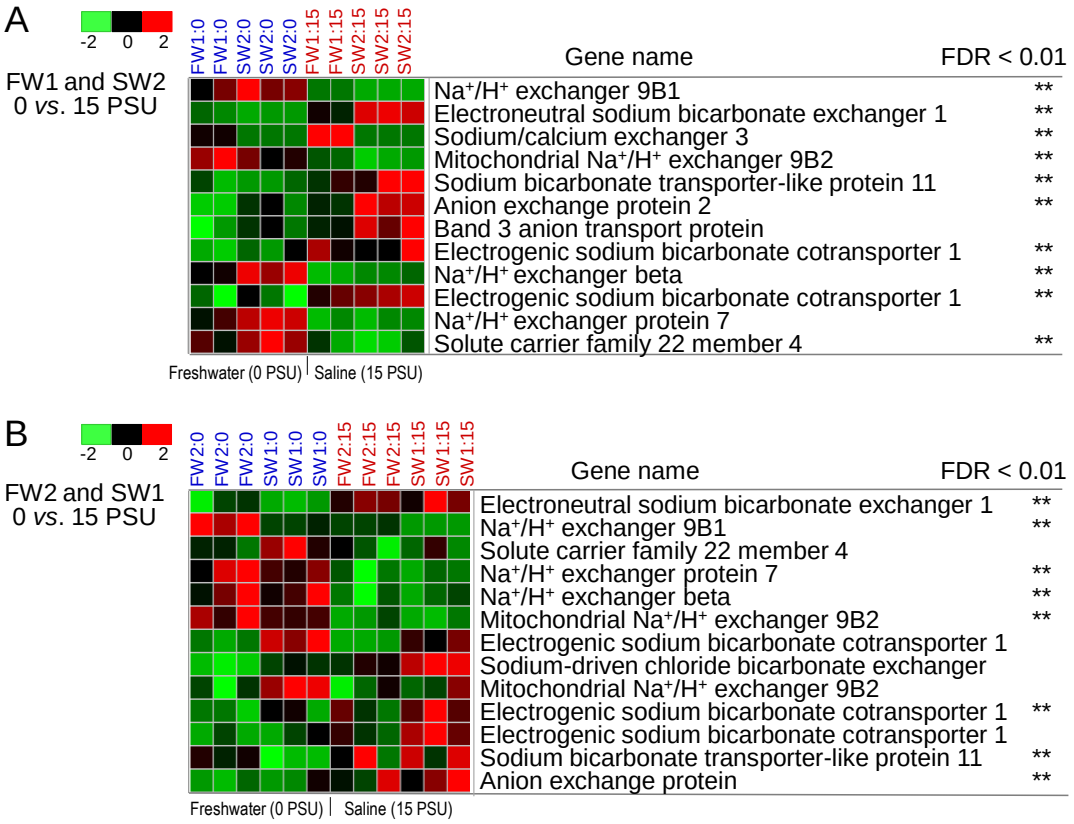

Figure 12. Gene expression response in FW and SW inbred lines of *E. affinis* due to rearing salinity (0 PSU *vs.* 15 PSU). The subset of genes that were grouped into GO category — antiporter activity, containing 50 genes in total) is shown. Genes that were differentially expressed (based on FDR corrected P value < 0.01, from negative binomial GLM) in both FW1 and SW2 (**A**) and FW2 and SW1 (**B**) inbred lines are denoted by **.

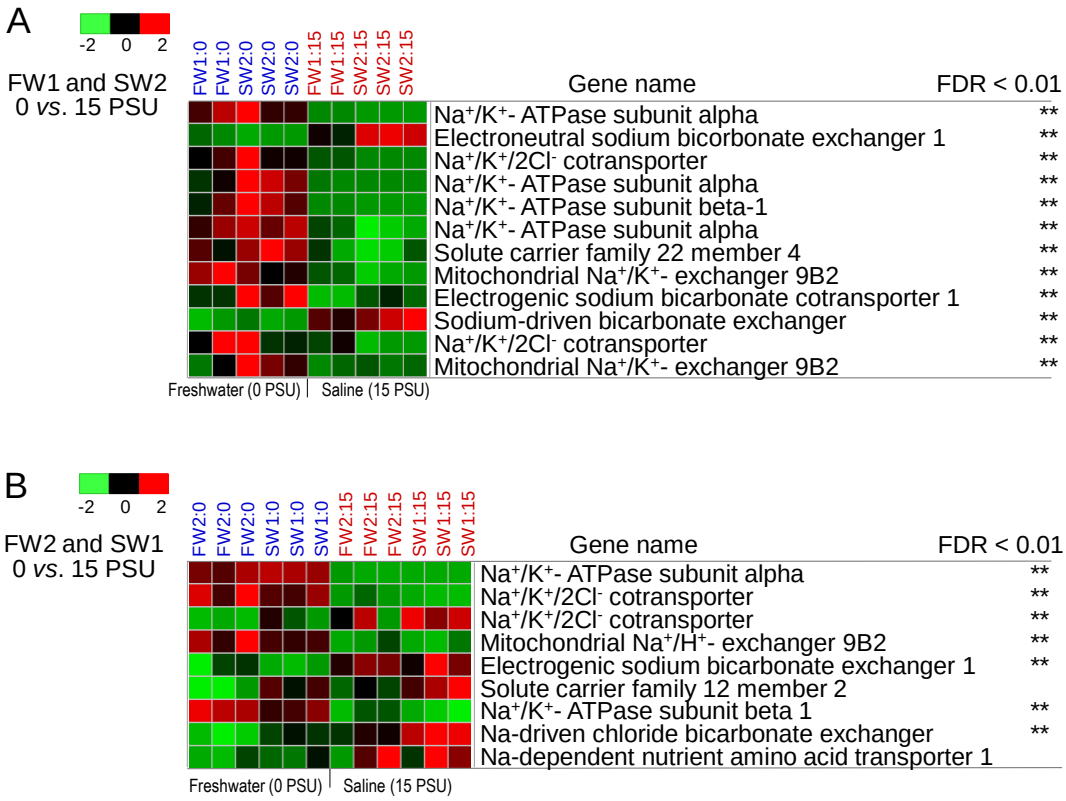

Figure 13. Gene expression response in FW and SW inbred lines of *E. affinis* due to rearing salinity (0 PSU *vs.* 15 PSU). The subset of genes that were grouped into GO category — sodium ion transport, containing 60 genes in total) is shown. Genes that were differentially expressed (based on FDR corrected P value < 0.01, from negative binomial GLM) in both FW1 and SW2 (**A**) and FW2 and SW1 (**B**) inbred lines are denoted by **.

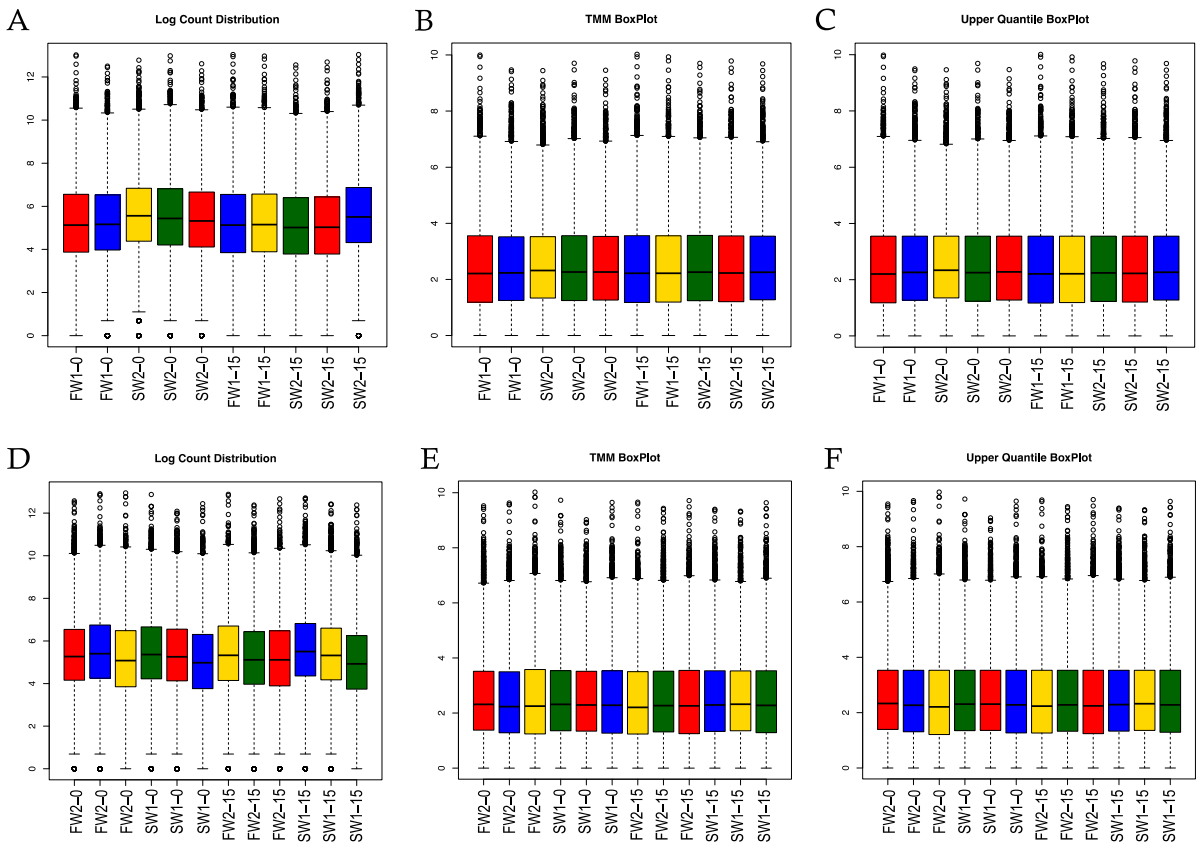

Figure S1. Distribution of raw (**A** and **D**), TMM (Trimmed Mean of M values) normalized (**B** and **E**), and UQ (upper quantiles) normalized (**C** and **F**) log counts. Box plots of raw read counts show considerable sample-to-sample variation in counts distribution. After TMM (**B** and **E**) and UQ normalization (**C** and **D**) all the sample count means and all the sample count distributions are almost aligned. This stabilization of read count distributions across samples indicates that both TMM and UQ provided an effective normalization.

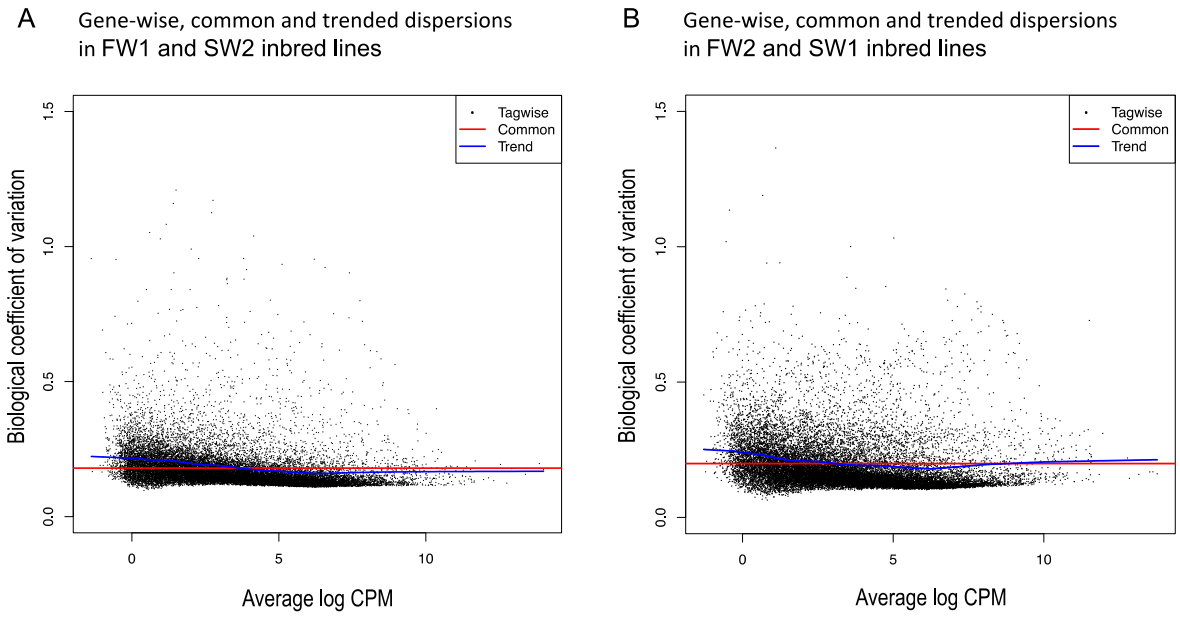

Figure S2. The biological coefficient of variation plot showing gene-wise, common and trended dispersions as a function of average log of counts per million (logCPM) in FW1 *vs*. SW2 (**A**) and FW2 *vs*. SW1 (**B**) comparison. Black dots represent the gene-wise dispersion estimates after empirical Bayes shrinkage.

Chapter 3

Exploring the nature of regulatory evolution in the saline to freshwater invading copepod *Eurytemora affinis*

ABSTRACT

The nature of regulatory adaptation is an unresolved question of fundamental importance in evolutionary biology. Furthermore, distinguishing *cis*- from *trans*-regulatory changes in gene expression could uncover the specific genes underlying freshwater adaptation in the invasive copepod *Eurytemora affinis*, as regulatory changes that are regulated in *cis*- could reveal the loci under selection. Moreover, copepod *E. affinis* has invaded freshwater habitats over the past 60 years, and no study has yet examined regulatory evolution on such a short ecological time scales. Thus, in this study I examined the relative contribution of *cis*- *versus trans*-regulatory evolution during rapid evolutionary shifts underlying saline to freshwater habitat invasions in the copepod *Eurytemora affinis*. Using RNA-seq data from two independent saline (SW) and two independent freshwater (FW) inbred lines and their F_1 hybrids, I found relatively low proportions of genes that were differentially expressed (DE, 9% of expressed genes) between saline and FW inbred lines. Out of the DE genes, I classified 7%, 22%, 6%, 17%, 17% and 30% of genes as *cis* only, *trans* only, *cis + trans* both, *cis x trans* interaction, *compensatory interaction,* and *ambiguous*, respectively. The higher prevalence of *trans*- (22%) relative to *cis*-regulatory changes (7%) might be a consequence of larger mutational target size of *trans* relative that in *cis* regulatory elements. Furthermore, I found several genes regulated in *cis* that might be involved in osmoregulation. Most notable, Na^+/H^+ exchanger and acidic chitinase showed significant *cis*-effect, suggesting that they might be target of selection during freshwater invasions. This study emphasizes the importance of separating *cis*- and *trans*-regulatory

changes, and significantly contributes to our understanding of the nature of regulatory evolution in invasive *E. affinis*.

Introduction

Among the unresolved and central topics in evolutionary biology regards the extent and nature of regulatory evolution. In particular, rapid adaptation to novel environments, such as during habitat invasions, might arise more frequently due to mutations that alter gene regulation rather than to those that change the coding regions of DNA (Stern and Orgogozo 2008, Fraser et al. 2010, Jones et al. 2012, Fraser 2013). In general, mutations that affect gene regulation may occur at the level of *cis*-regulatory elements (CRE), which are located near the regulated allele (e.g. promoter, enhancer), and/or at the level of *trans*-acting factors (TF), which are encoded elsewhere in the genome (e.g. transcription factors, activators, repressors). As it has become increasingly evident that changes in gene regulation play an important role in adaptive evolution (Wray 2007, Stern and Orgogozo 2008), distinguishing between *cis*- and *trans*-regulatory changes become crucially important. This problem is central to understanding of molecular basis of evolution, in part, since changes in gene expression due to variation in CRE imply that selection is directed at the regulated gene, while changes due to TF imply that selection is acting somewhere else in the genome. Thus, partitioning the regulatory variation into "*cis*" and "*trans*" reveals important insights into the specific candidate targets of selection.

Population genetics theory predicts that mutations that have larger pleiotropic effects will also impose a larger fitness cost (Fisher 1930, Orr 2000). As mutations at

trans-acting factors may affect the expression of many genes, they tend to be pleiotropic (Cooper et al. 2007). On the other hand, because of their modular structure, mutations in CRE usually do not show any pleiotropic effects. In addition, *cis*-regulatory mutations show higher additive fitness effects than *trans*-regulatory mutations (Lemos et al. 2008). Consequently such an additive mode of action makes heterozygotes visible to selection at *cis*-regulatory sites, allowing selection to operate more efficiently. Therefore, *trans*-acting factors, such as transcription factors, may be more evolutionary constrained than *cis*-regulatory elements. Thus, adaptive *cis*-regulatory mutations will tend to accumulate over time and contribute to regulatory divergence between populations, to a greater extent than *trans*-acting factors (Stern 2000).

Several interspecific experimental studies demonstrated the prevalence of *cis*-regulatory changes in regulatory evolution (Witkopp et al. 2004, Wittkopp et al. 2008, Graze et al. 2009, Tirosh et al. 2009). However, the contribution of *cis*- and *trans*-regulatory changes to gene expression differences between populations within species is still elusive (Wayne et al. 2004, Osada et al. 2006, Wittkopp et al. 2008, Suvorov et al. 2013). Moreover, it is not clear how rapid adaptation, such as during habitat invasions, might affect nature of regulatory evolution. In the case of rapid evolution, we might predict that a greater proportion of regulatory changes to be due to trans-regulatory evolution, as selection might not have had time to remove novel mutations in *trans* acting factors. Furthermore, copepod *E. affinis* has evolved freshwater tolerance in a very short period of time (50 years). So far, no study has examined regulatory evolution on such short ecological time scale.

Thus, in this study, I examined the relative contribution of *cis*- *vs. trans*-regulatory evolution during rapid evolutionary shifts underlying saline to freshwater habitat invasions in the copepod *Eurytemora affinis*. To infer the relative contribution of *cis*- and *trans*-regulation to adaptation in invading *E. affinis*, I applied next generation technology (RNA-seq data) and allele specific gene expression (ASGE, Fig. 2) analysis. In order to distinguish between *cis*- and *trans*-regulatory changes, I crossed saline and freshwater inbred lines of *E. affinis* and compared gene expression in parental inbred lines to allele specific expression in their F_1 hybrids.

Allele specific expression analysis of genome wide analysis of RNA-seq data reveals the relative contribution of changes in *cis*- and *trans*-regulation by comparing gene expression levels in parental lines to allele specific expression in their F_1 hybrids (Wittkopp et al. 2008, McManus et al. 2010). The fact that a *cis*-regulatory change affects only expression of its target allele, and a change in *trans*-regulatory factors affects both alleles at a locus in diploid organisms allows disentangling *cis*- from *trans*-regulatory change. While in maternal and paternal inbred lines, gene expression is subject to only their line specific *trans*- and *cis*- regulatory elements (assumed to be homozygous in inbred lines), allele specific gene expression in F_1 hybrids is a subject to both maternal and paternal *trans*-acting factors but still only to line specific *cis*-regulatory elements (see Methods, Fig. 2). Consequently, allelic expression in F_1 hybrids that resembles gene expression of their parents indicates *cis* regulation, whereas allelic expression level in hybrids that differs from parental gene expression indicates at least some level of *trans*-regulatory evolution (see Material and Methods for more details).

This study will significantly aid in understanding regulatory evolution on a genomic scale, particularly in response to sudden and dramatic environmental changes, such as habitat invasions. Furthermore, this study contributes information on regulatory evolution in a new model system, adding to the knowledge base regarding general trends in *cis- vs. trans*-regulatory evolution. Detection of *cis*-regulatory changes will provide a pool of putative candidate loci that could be further tested to determine if they serve as the targets of selection during freshwater invasions. As this case study involves invasion occurring on time scales of only a few decades, it could provide novel insights into the evolution of gene expression underlying rapid adaptation to new environments, with relevance for other invasive species.

Material and methods

INBRED LINES AND COMMON GARDEN EXPERIMENT

In order to detect evolutionary changes in gene expression associated with colonization of freshwater habitats, RNA-seq data from *E. affinis* inbred lines derived from ancestral saline and freshwater populations and their F_1 crosses were used. Inbred lines were produced by brother sister mating for 30 generations (~2.5 years) from two wild populations; one sampled from the St. Lawrence salt marsh, at L'Isle Verte (saline, SW) Quebec, Canada, and another from Lake Michigan at Racine, WI, USA (freshwater, FW) (see Lee et al. 2011 for details on exact locations). In order to detect differentially expressed genes and allele specific expression (ASE), two saline (SW1 and SW2) and two freshwater (FW1 and FW2) inbred lines and their F_1 crosses (Fig. 1) were reared

under two common-garden salinities (0 and 15PSU), and their RNA-seq reads abundance was compared.

The common-garden experiments were conducted under controlled standard conditions, at 12°C on a 15L:9D photoperiod. To ensure that individuals were virgins prior to crossing, ca. 200 post-metamorphic juveniles from each inbred line were isolated and placed individually into 20 ml scintillation vials filled with 10 ml of 5 PSU water. After 8-10 days of acclimation at 5 PSU (i.e. when juveniles became sexually mature adults), a mating scheme to produce three different types of offspring (genotypes) (Fig. 1) was performed:

1. Offspring from saline parental inbred lines (SW1 x SW1 and SW2 x SW2),
2. Offspring from freshwater parental inbred lines (FW1 x FW1 and FW2 x FW2), and
3. F_1 crosses (hybrids) ♀SW1 x ♂FW2 and ♀SW2 x ♂FW1.

Following successful mating, offspring were reared under 5 PSU conditions until they reached metamorphosis (age of ~14-15 days). After metamorphosis, at the age of 14-15 days, samples of copepods of each genotype were split across two different salinities (0 and 15 PSU) and reared at these salinities until adulthood (until age of ~30-32 days).

To perform the allele specific expression assay, parental inbred lines SW1 and FW2 were contrasted to their F_1 crosses ♀SW1 x ♂FW2, while parental inbred lines SW2 and FW1 were contrasted to their F_1 crosses ♀SW2 x ♂FW1.

RNA EXTRACTION AND SEQUENCING

When the copepods reached adulthood, total RNA was extracted from whole bodies of 50 copepods (25 females and 25 males) per sample. Two biological replicates per parental line and per each F_1 hybrid were collected (Fig. 1). Total RNA was extracted with Trizol reagent (Ambion RNA, Carlsbad, CA) and then purified with Qiagen RNeasy Mini Kit (Qiagen cat# 74104, Valencia CA), following the protocol of Lopez and Bohuski (2007).

Library construction and sequencing were conducted at the Genomics Resource Center (GRC), Institute for Genome Sciences, at the University of Maryland School of Medicine. A modified strand-specific protocol was adapted from Parkhomchuk et al. (2009). Actinomycin D was added to the first strand synthesis reaction, and the second strand was synthesized with a dNTP mix containing dUTP. The second strand cDNA was digested, after adapter ligation, with 2 units of Uracil-N-Glycosylase (Applied Biosystems, Carlsbad, CA). Purification of DNA was performed with AMPure XT beads (Beckman Coulter Genomics, Danvers, MA) between enzymatic reactions and size selection of the library. Quantity and size of the libraries were assessed on the LabChip GX (Perkin Elmer, Waltham, MA) and with the Library Quantification Kit for Illumina (Kapa Biosciences, Woburn, MA).

RNA-SEQ DATA PROCESSING

RNA-seq data were generated by paired-end (PE) strand specific sequencing on Illumina HiSeq 2000. Initially, to avoid possible batch effects each library was split in half and then samples were run on two lanes, multiplexing 12 samples per lane on the Illumina

HiSeq 2000 (2x101 bp run). On average 7 x 10^7 of paired-end (PE) reads (3.5 x 10^7 fragments) per sample were generated. Sequence quality, trimming and filtering was assessed with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and Trimmomatic (Trimmomatic Version 032, Bolger et al. 2014). In the final step of quality filtering and trimming the following criteria were applied: the leading and trailing bases were removed if their quality was below 5 (Phred-33), the bases in sliding window of 4 were removed if quality score dropped below 15 (Phred-33), and complete reads were removed if average quality of read was below the score of 20 (Phred-33). Based on this filtering criteria about 10 % of reads were discarded and on average ~3.2 x 10^7 of fragments (read pairs) were available for aligning (Table 1).

QUANTIFICATION OF ALLELE SPECIFIC EXPRESSION

In order to partition total gene expression differences into those that were due to *cis*- or *trans*-regulation, the relative allele specific expression of F_1 hybrids (H_{SW} – saline alleles in hybrids and H_{FW} – freshwater alleles in hybrids) were compared to gene expression in parental inbred lines (i.e. to P_{SW} – alleles in saline parent and P_{FW} – alleles in freshwater parent) (Fig. 2, Wittkopp et al. 2004; McManus et al. 2010). If expression differences at a locus between saline and freshwater populations were due to *cis*-regulatory changes, they persisted in the F_1 hybrids. The reason is that the F_1 hybrid retains *cis*-regulatory elements for each of the two alleles, and each allele is regulated by its own *cis* region (Fig. 2). In other words, the F_1 hybrids exhibit unequal levels of transcript in the saline *vs.* fresh allele, reflecting the unequal expression in the parental inbred lines. On the other hand, if differences in expression at the regulated gene in

parental lines are due to *trans*-regulatory differences, the expression level in F_1 hybrids at both chromosomes (alleles) is similar (i.e. no significant difference in expression between alleles is expected) since they will be regulated in an equal manner by both (SW and FW) *trans* elements (see Fig. 2). Degrees of expression between these two extreme cases will imply the relative contribution of the *cis*- and *trans*- regulatory evolution to the change in gene expression at the locus.

Accurate detection of *cis*- *versus trans*-regulatory changes requires unbiased alignment of reads from both alleles at a locus to a reference genome. If only one parental genome is available, reads representing the alleles from that parental line might preferentially map relative to those from other parental line (Degner et al. 2009, Satya et al. 2012, Stevenson et al. 2013). To overcome this bias, several different strategies were proposed, including aligning the reads separately to both parental genomes (McManus et al. 2010), or using additional information from all possible haplotypes during alignment (Skelly et al. 2011).

In order to quantify allele specific expression (ASE) in *E. affinis*, the Allim software package (http://code.google.com/p/allim/, Pandey et. al 2013), which corrects for mapping bias when the reference genome is available for only one parental line (Fig. 3) was employed. Allim first generates genome for both parental lines using RNA-seq data from both parental lines and the reference genome from one of the parental lines. To estimate remaining mapping bias, Allim performs sequence specific simulation and estimates bias factors, which are then used to correct read abundance.

The Allim pipeline encompasses 5 modules (Fig. 3). In Module 1, RNA-seq reads from both SW and FW parental lines were first mapped onto the reference *E.*

affinis genome *via* GSNAP mapper (Wu and Nacu 2010). Next, based on mapped reads and gene annotation, fixed SNPs between the two parental lines were identified. GSNAP then used these fixed SNPs to create genomes for both parental lines, which were further used as a reference in the next round of mapping (Pandey et al. 2013). Parental genomes were created by substitutions of the base in reference genome at detected fixed SNP. Thus, they are identical to the reference genome except at the fixed SNP positions.

The remaining mapping bias was estimated by simulation using the following algorithm (Pandey et al. 2013):

(1) Generating the same number of RNA-seq reads (pair end) for identical genome positions in both parents (recall that parental genomes are different only by fixed SNPs). In the simulation process, the same number of reads per identical genome positions was generated by simulating all reads that span across at least one fixed SNP only once (Pandey et al. 2013).

(2) Mapping simulated reads to both parental genomes simultaneously (Module 3, Fig. 3). Parental expression ratio of mapped simulated reads different from 1 indicated mapping bias.

(3) Calculating the bias correction factor for genes with different number of mapped reads.

The two parental genomes that were created in Module 1 were then used for mapping the parental and F_1 hybrid RNA-seq reads following bias correction. Then, in modules 4 and 5 (Fig. 3), estimation of parental gene expression, allele specific expression in hybrids, and statistical test of allelic imbalance were performed. For RNA-seq read normalization, the trimmed mean of M-values normalization (TMM), implemented in

edgeR (Robinson and Oshlack 2010, Chen et al. 2014) was used. To test for *cis*- and *trans*-regulatory effects, a linear model in ANOVA framework and multiple testing corrections with single-step method, employing *multcomp* package and *glht* function of R were applied (Hothorn et al. 2008, R core team 2013).

CLASSIFICATION OF *CIS*- AND *TRANS*-REGULATORY EVOLUTION

Classification of types of evolutionary changes in gene regulation, between saline and freshwater inbred lines, was performed following the strategy described by Landry et al. (2005) and McManus et al. (2010). Seven different categories of inferred regulation were distinguished applying following criteria:

Cis only: Significant differential allelic expression detected in parental lines ($P_{SW} \neq P_{FW}$) and in hybrid ($H_{SW} \neq H_{FW}$), but no significant difference detected in the ratio of allelic expression in the parental lines relative to that in the hybrid ($P_{SW}/P_{FW} \cong H_{SW}/H_{FW}$).

Trans only: Significant differential allelic expression detected in parental lines ($P_{SW} \neq P_{FW}$) but not in hybrid ($H_{SW} \cong H_{FW}$). However, significant difference detected in the ratio of allelic expression in the parental lines relative to that in the hybrid ($P_{SW}/P_{FW} \neq H_{SW}/H_{FW}$).

Cis + trans: Significant differential allelic expression in parental lines ($P_{SW} \neq P_{FW}$) and in hybrid ($H_{SW} \neq H_{FW}$), and significant difference in the ratio of allelic expression in the parental lines relative to that in the hybrid (($P_{SW}/P_{FW} \neq H_{SW}/H_{FW}$). The log-transformed allelic expression ratios of this class of genes in parental lines and in hybrid have the same signs, indicating that *cis*-

and t*rans*-regulatory changes favored expression of the same allele.

Cis* x *trans: Significant differential allelic expression in parental lines ($P_{SW} \neq P_{FW}$), and in hybrid ($H_{SW} \neq H_{FW}$), and significant difference in the ratio of allelic expression in the parental lines relative to that in the hybrid ($P_{SW}/P_{FW} \neq H_{SW}/H_{FW}$). The log-transformed allelic expression ratios of this class of genes in parental lines and in hybrid have opposite signs. In this scenario *cis*- and t*rans*-regulatory changes favored expression of opposite alleles.

Compensatory: No significant allele specific differential expression in parental lines ($P_{SW} \cong P_{FW}$), but significant allele specific differential expression in hybrid ($H_{SW} \neq H_{FW}$), and significant difference in the ratio of allelic expression in the parental lines relative to that in the hybrid ($P_{SW}/P_{FW} \neq H_{SW}/H_{FW}$). In this type of interaction, *cis*- and t*rans*-regulatory differences compensate each other, resulting in no gene expression difference between the parental lines.

Conserved: Genes in this category do not show evidence of differential allelic expression either in parental lines or hybrids.

Ambiguous: Genes that do not belong to any of aforementioned categories, i.e., without clear biological interpretation.

For functional enrichment analysis ErmineJ version 3.0.2 (Gillis et al. 2010) using gene annotation, gene descriptions, and gene ontology (GO) identifiers (Ashburneret et al. 2000) was employed. Gene Score Resampling analysis (implemented in ErmineJ, Gillis et al. 2010) was conducted using log fold-changes (logFC) and TMM normalized gene counts (Robinson and Oshlack 2010). Statistically significant GO categories were

determined by 10^3 iterations and using FDR corrected P value < 0.01 (Benjamini and Hochberg 1995) as a cutoff value.

Results

To characterize patterns of allele specific gene expression in *E. affinis*, I conducted two separate comparisons: **a**) paternal saline and freshwater inbred lines, SW1 and FW2 *versus* their F_1 hybrid —♀SW1 x ♂FW2, and **b**) two other independent saline and freshwater paternal inbred lines, SW2 and FW1 *versus* their F_1 hybrid — ♀SW2 x ♂ FW1. I reared all parental lines and their hybrids under freshwater conditions (0 PSU). The analysis revealed 99241 fixed SNPs between SW1 and FW2 inbred line and 168907 fixed SNPs between SW2 and FW1 line, with minimum coverage of 10 reads per fixed SNP. On average I detected 4.6 (SW1 *vs.* FW2) and 7.2 (SW2 *vs*. FW1) SNPs per gene. I detected at least one SNP in 57% (SW1 *vs*. FW2) and 62% (SW2 *vs*. FW1) of annotated genes. The average number of SNPs per gene obtained in this study was higher than that reported in similar studies (Suvorov et al. 2013, Quinn et al. 2014) conducted on Drosophila inbred lines.

In total I detected 1986 (9.2% of total number of genes included in analysis) differentially expressed ($P_{adj} < 0.05$) genes in in SW1 *vs*. FW2 and 1960 (8.3% of total number of genes included in analysis) differentially expressed ($P_{adj} < 0.05$) genes in SW2 *vs.* FW1 comparison, respectively (Fig. 4). Among 1986 differentially expressed genes (DE) in SW1 *vs*. FW2 comparison 143 (7.2% of DE genes) genes showed evidence of significant *cis*-regulatory effect alone, whereas 440 (22% of DE genes) showed evidence

of significant *trans*-effect (Fig. 4). Significant evidence of both *cis*- and *trans*-effects (genes classified as *cis + trans, cis x trans* and *compensatory* interactions) was detected in 807 (40% of DE genes) genes (Fig. 4). For 124 DE genes, *cis*- and *trans*-regulatory effects favored expression of the same allele (*cis + trans* effect), whereas for 338 genes, the opposite alleles were favored (*cis x trans* interaction). 345 DE genes showed compensatory interaction, where *cis*-and *trans*-effects compensated each other, resulting in no difference between allelic expression in saline and freshwater parental lines despite significant differential allele specific expression in their F_1 hybrids. A relatively large proportion (30%) of DE genes was classified as ambiguous with no clear biological interpretation for their expression pattern. Similar frequencies of gene regulatory classes were observed in the SW2 vs. FW1 comparison except for *cis + trans* category, were only 11 genes showed evidence of significant effect (Fig. 4). The proportion of differentially expressed genes showing significant evidence of *cis* and *trans*-effects detected in this study was lower than reported by others (McManus et al. 2010, Suvorov et al. 2013).

Gene enrichment analysis performed by ErmineJ (Gillis et al. 2010) revealed no significant GO category for different regulatory types. GO categories — sodium ion transport, glycosaminoglycan catabolic process, and polysaccharide catabolic process were marginally significant.

The Na^+/H^+ exchanger 9B1 was upregulated in *cis* in FW2 relative to the SW1 parental line (Table 2), while the electrogenic sodium bicarbonate cotransporter was upregulated in *trans* in FW1 relative to SW2 parental inbred lines (Table 7). Acidic chitinase involved in catabolism of chitin, which might affect cuticle permeability (Lv et

al. 2013), was downregulated in *cis* in the FW2 relative to SW1 inbred line (Table 4). Other genes that showed evidence for significant *cis* effect were listed in Tables 2 to 5, while genes that showed significant *trans* effect were listed in Tables 6 to 9.

Discussion

Detection of single nucleotide polymorphism (SNP) between fresh and saline populations of *E. affinis* is crucial for unraveling the relationship between genotypes and adaptive phenotypes, and could deepen our understanding of evolutionary mechanisms of invasive success. Employing RNA-seq and Alllim package software (Pandey et al. 2013), I identified 99241 fixed SNPs between SW1 and FW2 and 168907 fixed SNPs between SW2 and FW1 inbred lines. Many of these SNPs might be fixed by genetic drift, because a strong genetic drift might be acting on numerous loci during the process of generating of the inbred lines. Thus, caution is needed when interpreting the significance of the number of reported SNPs.

Genes that showed evidence of significant *cis*-regulatory effects are of special interest because they might have been targets of selection during freshwater invasion of *E. affinis*. However, only a small proportion of DE genes showed evidence of *cis*-regulatory divergence alone between saline and freshwater inbred lines in this study (Fig. 4, Tables 2, 3, 4 and 5), with no genes showing significant *cis*- effect across both SW1 *vs*. FW2 and SW2 *vs*. FW1 comparisons. Again, strong genetic drift during the inbreeding process might have resulted in some relevant adaptive mutations to be lost in one of the lines, so power to detect significant adaptive *cis* regulatory adaptation might be reduced in this study. Despite this limitation, the two genes that showed significant

cis-effect, Na^+/H^+ exchanger and acidic chitinase, might be considered as good candidates that serve as targets of selection during freshwater invasions. The Na^+/ H^+ exchanger 9B1, which mediates uptake of environmental Na^+ across the apical membrane of osmoregulatory epithelia in exchange for intracellular H^+ (Towle and Weihrauch 2001), was upregulated in *cis* in the FW2 relative to SW1 parental line (Table 2). This result suggests that the freshwater population of *E. affinis* might have acquired a mutation(s) in the *cis*-regulatory region of Na^+/H^+ exchanger gene, which results in increased Na^+ uptake activity from dilute media. Acidic chitinase showed a strong *cis* effect of 4.5 fold downregulation (on $\log_2$ bases) in FW2 parental line relative to SW1 parental line (Table 4). This enzyme is involved in chitin degradation, which might affect copepod cuticle permeability to water and ions. Further studies should be conducted to clarify the role of these candidate genes in rapid adaption in *E. affinis.*

Overall, the pattern of *cis*- and *trans*-regulatory effects I observed in the comparison between populations within *E. affinis* species is not in concordance with the pattern observed in comparisons between species of Drosophila (Wittkopp et al. 2004, Wittkopp et al. 2008, Graze et al. 2009) and yeast (Tirosh et al. 2009). These previous studies demonstrated a greater proportion of gene expression differences due to *cis*-relative to *trans*-regulatory changes. However, I observed three-fold greater proportion of genes showing significant *trans*-effects than genes with significant *cis*-effects (Fig. 4). Results of this study are more similar to Suvorov et al. (2013), who conducted ASE assay on intraspecific *Drosophila* inbred lines, and to Tirosh (2009), who studied regulation of gene expression in yeast. In both of these studies a much larger proportion of genes with significant *trans*-effect than with significant *cis*-effect were reported.

The opposite pattern of prevalence of *cis*- and *trans*-regulatory changes between and within species is in agreement with theoretical expectation and experimental findings (Emerson et al. 2010, Schaefke et al. 2013, Suvorov et al. 2013). It was hypothesized (Stern 2000) that *trans*-regulatory factors should show stronger pleiotropic effects than *cis*- regulatory factors, and consequently would be subjected to stronger purifying selection. This would lead to the prevalence of *cis*-regulatory divergence between species. However, because the mutational target size is larger in *trans* relative that in *cis* regulatory elements, the prevalence of *trans*-regulatory changes is expected on shorter time scales of evolutionary divergence between populations within species (Emerson et al. 2010, Schaefke et al. 2013). Thus, prevalence of *trans*- over *cis*-regulatory differences between ancestral saline and derived freshwater populations of *E. affinis* might be consequence of short divergence time (~60 years, Lee 1999).

CAVEATS

In order to infer the relative contribution of *cis*- and *trans*-regulation to freshwater adaptation in *E. affinis* a more sensitive assay might be required. This study relied on draft genome of *E. affinis*, with gaps in sequence and incomplete gene annotation. Using the draft genome in combination with incomplete annotation is challenging and limiting in answering interesting biological questions. Approximately 50% of gene models of draft genome *E. affinis* are missing, or incompletely and incorrectly annotated. Thus it is very likely that this analysis is missing some important genes underlying regulatory evolution of freshwater tolerance. Another significant improvement would be to use specific tissue that is involved in the process studied (e.g. osmoregulatory tissue) instead

of conducting analysis on RNA extracted from whole bodies of copepods. Finally, very strong genetic drift during strict full-sib inbreeding process could have cause many mutations of smaller adaptive effect to be lost in distinct inbred lines. Indeed, if adaptation to freshwater conditions is due primarily to complex, quantitative traits, which involve numerous genes and pathways, many of relevant genes might be lost during the inbreeding process and not be detectable under current study design. Further studies will clarify the nature of regulatory evolution in *E. affinis* as well as which genes are likely targets of selection during rapid adaptions to freshwater conditions.

ACKNOWLEDGMENTS

This project was supported by National Science Foundation grants NSF DEB-0448827, DEB-0745828, and DEB-1050565 to Carol E. Lee. Martin Bontrager provided bioinformatics assistance and valuable help in conducting common garden experiment and RNA extractions. Jane Remfert provided valuable assistance with conducting common garden experiment. Numerous students assisted with generating the inbred lines over a period of three years, including W. Moss, A. Tritt, B. Metzger, A. Jenstead, K. Lee, A. Bateman, M. Noll, K. Jung, S. Kahn, K. Bridges, T. Pian, A. Toy, K. Lee, S. Kahn, and J. Hilton.

LITERATURE CITED

Ashburner, M., C. Ball, J. Blake, D. Botstein, H. Butler, J. Cherry, A. Davis, K. Dolinski, S. Dwight, and J. Eppig. 2000. Gene ontology: tool for the unification of biology. Nat. Genet. 25:25–29.

Barnosky, A. D., N. Matzke, S. Tomiya et al. 2011. Has the Earth's sixth mass extinction already arrived? Nature 471: 51–57.

Benjamini, Y. and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B Stat. Methodol. 57:289–300.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: A flexible trimmer for Illumina Sequence Data. Bioinformatics, btu170.

Bullard, J. H., E. Purdom E, K. D. Hansen et al. 2010. Evaluation of statistical methods for normalization and differential expression in mRNA-seq experiments. BMC Bioinformatics 11:(94).

Caroll, S. B. 2000. Endless forms: the evolution of gene regulation and morphological diversity. Cell 101:577–580.

Caroll, S. B. 2005. Evolution at two levels: on genes and forms. PLoS Biology 3:1159–1166.

Carroll, S. B., J. K. Grenier, and S. D. Weatherbee. 2004. From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design. Blackwell Publishing, Malden.

Chen, G., C. Wang and T. Shi. 2011. Overview of available methods for diverse RNA-Seq data analyses. Sci China Life Sci 54:1121–1128.

Coolon, J. D., C. J. McManus, K. R. Stevenson, B. R. Graveley, and P. J. Wittkopp. 2014. Tempo and mode of regulatory evolution in Drosophila. Genome Res. 24:797–808.

Cooper, T. F., E. A. Ostrowski, and M. Travisano. 2007. A negative relationship between mutation pleiotropy and fitness effect in yeast. Evolution 61:1495–1499.

Cowles, C.R.,J. N. Hirschhorn, D. Altshuler and E. S. Lander. 2002. Detection of regulatory variation in mouse genes. Nat Genet 32: 432–437.

Degner, J. F., J. C. Marioni, A. A. Pai, J. K. Pickrell, E. Nkadori, Y. Gilad, and J. Pritchard. 2009. Effect of read-mapping biases on detecting allele-specific expression from RNA-sequencing data. Bioinformatics 25:3207–3212.

Denver, D. R., K. Morris , J. T. Streelman, S. K; Kim, M. Lynch, and W. K. Thomas. 2005. The transcriptional consequences of mutation and natural selection in *Caenorhabditis elegans*. Nat. Genet. 37: 544–548.

Denver, D. R., K. Morris, J. T. Streelman, S. K. Kim, M. Lynch M, et al. 2005. The transcriptional consequences of mutation and natural selection in Caenorhabditis elegans. Nat Genet 37:544–548.

Dillies, M. A., A. Rau, J. Aubert et al. 2012. A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. Briefings in Bioinformatics doi:10.1093/bib/bbs046.

Emerson, J. J., C. C. Hsieh, H. M. Sung, T. Y. Wang, C. J. Huang, et al. 2010. Natural selection on cis and trans regulation in yeasts. Genome Res. 20: 826–836.

Engström, P., T. Steijger, B. Sipos, G. R. Grant, A. Kahles, G. Rätsch, N. Goldman, T. J. Hubbard, J. Harrow, R. Guigó, P. Bertone; RGASP Consortium. 2013. Systematic evaluation of spliced alignment programs for RNA-seq data. Nature Methods, 10;1185–119.

Fisher, R. A. 1930. The genetical theory of natural selection. Clarendon Press, Oxford. U.K.

Gillis J., and P. Pavlidis. 2011. The Impact of Multifunctional Genes on "Guilt by Association" Analysis. PLoS ONE 6(2): e17258. doi:10.1371/journal.pone.0017258

Gillis J., M. Mistry, and P. Pavlidis. 2010. Gene function analysis in complex data sets using ErmineJ. Nature Protocols 6:1148 –1159.

Graze R. M., L. M. McIntyre, B. J. Main, M. L. Wayne, and S. V. Nuzhdin. 2009. Regulatory divergence in Drosophila melanogaster and D. simulans, a genomewide analysis of allele-specific expression. Genetics 183:547–561.

Hardcastle, T.J. and K.A. Kelly. 2010. baySeq: empirical bayesian methods for identifying differential expression in sequence count data. BMC Bioinformatics, 11:422–435.

Hayden, E.C. 2012. RNA studies under fire. Nature 484:428.

Hoekstra, H. P. and J. A. Coyne, 2007. The locus of evolution: evo devo and the genetics of adaptation. Evolution 61:995¬016.

Hothorn T., F. Bretz, and P. Westfall. 2008. Simultaneous inference in general parametric models. Biometrical Journal 50:346–363.

Jones, F. C., M. G. Grabherr, Y. F. Chan, P. Russell, E. Mauceli, J. Johnson, R. Swofford, M. Pirun, M. C. Zody, S. White, et al. 2012. The genomic basis of adaptive evolution in threespine sticklebacks. Nature 484:55–61.

Landry, C. R., P. J. Wittkopp, C. H. Taubes, J. M. Ranz, A. G. Clark, and D. L. Hartl. 2005. Compensatory *cis-trans* evolution and the dysregulation of gene expression in interspecific hybrids of Drosophila. Genetics 171:1813–1822.

Landry, C. R., B. Lemos, S. A. Rifkin, W. J. Dickinson, and D. L. Hartl. 2007. Genetic properties influencing the evolvability of gene expression. Science 317:118–121.

Lemos B., C. D. Meiklejohn, M. Caceres, and D. L. Hartl. 2005. Rates of divergence in gene expression profiles of primates, mice, and flies: stabilizing selection and variability among functional categories. Evolu tion 59:126–137.

Lemos, B., L. O. Araripe and D. L. Hartl. 2008. Polymorphic Y Chromosomes Harbor Cryptic Variation with Manifold Functional Consequences. Science 319:91–93.

Lv, J., P. Liu , Y. Wang , B. Gao , Chen P, et al. 2013. Transcriptome analysis of *Portunus trituberculatus* in response to salinity stress provides insights into the molecular basis of osmoregulation. PLoS ONE 8: e82155. doi:10.1371/journal.pone.0082155

McGettigan, P. A. 2013. Transcriptomics in the RNA-seq era Current Opinion in Chemical Biology 17:4–11

McManus, C. J., J. D. Coolon, M. O. Duff, et al. 2010. Regulatory divergence in Drosophila revealed by mRNA-seq. Genome Research 20:816¬825.

Mortazavi, A., B. A.Williams, K. McCue, L. Schaeffer and B. Wold. 2008. Mapping and quantifying mammalian transcriptomes by RNA–Seq. Nat Methods 5:621–628.

Orr, H. A. 2000. Adaptation and the cost of complexity. Evolution 54:13¬20.

Osada, N., M. H. Kohn, and C. I. Wu. 2006. Genomic inferences of the cis-regulatory nucleotide polymorphisms underlying gene expression differences between Drosophila melanogaster mating races. Mol Biol Evol 23:1585–1591.

Pan, Q., O. Shai, L. J. Lee, B. J. Frey, and B. J.Blencowe. 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nature genetics 40:1413¬5.

Pandey, R.V., S. U. Franssen, A. Futschik, and C. Schlötterer. 2013. Allelic imbalance metre (Allim), a new tool for measuring allele-specific gene expression with RNA-seq data. Mol Ecol Resour. 13:740¬745.

Parkhomchuk, D., T. Borodina, V. Amstislavskiy, M. Banaru, L. Hallen, S. Krobitsch, H. Lehrach, and A. Soldatov. 2009. Transcriptome analysis by strand-specific sequencing of complementary DNA. Nucleic Acids Res. 37:e123. doi: 10.1093/nar/gkp596

Pimentel, D., L. Lach, R. Zuniga and D. Morrison. 2000. Environmental and economic costs of nonindigenous species in the United States. Bioscience 50, 53–65.

Quinn, A., P. Juneja and F. M. Jiggins. 2014. Estimates of allele-specific expression in Drosophila with a single genome sequence and RNA-seq data. Bioinformatics 30:2603–2610.

R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL *http://www.R-project.org/*

Robinson, M. D., and A. Oshlack. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biology 11, R25.

Robinson, M.D., D. J. McCarthy and G. K. Smyth. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26: 139¬140.

Ronald J. and J. M. Akey. 2007. The evolution of gene expression QTL in *Saccharomyces cerevisiae*. PLoS One 2: e678. doi: 10.1371/journal.pone.0000678.

Satya, R.V., N. Zavaljevski, and J. Reifman. 2012. A new strategy to reduce allelic bias in RNA-Seq read mapping. Nucleic Acids Res. 40:e127.

Schaefke B., J. J, Emerson,T. T. Wang, M. Y. Lu, L.C. Hsieh and W.-H. Li. 2013. Inheritance of gene expression level and selective constraints on trans- and cis-regulatory changes in yeast. Mol Biol Evol 30:2121–2133.

Skelly, D. A., M. Johansson, J. Madeoy, J. Wakefield, and J.M. Akey. 2011. A powerful and flexible statistical framework for testing hypotheses of allele-specific gene expression from RNA-seq data. Genome Res. 21:1728–1737.

Soneson, C., and M. Delorenzi. 2013. A comparison of methods for differential expression analysis of RNA-seq data. BMC Bioinformatics 14:91¬108.

Stefflova, K., D. Thybert, M. D. Wilson, I. Streeter, J. Aleksic, P. Karagianni, A. Brazma, D. J. Adams, I. Talianidis, J. C. Marioni, P. Flicek and D. T. Odom. 2013. Cooperativity and Rapid Evolution of Cobound Transcription Factors in Closely Related Mammals. Cell 154:530–540.

Stern, D. L. 2000. Perspective: evolutionary developmental biology and the problem of variation. Evolution 54:1079–1091.

Stern, D.L. and Orgogozo V. 2008. The loci of evolution: how predictable is genetic evolution? Evolution 62:2155–2177.

Stevenson, K.R., J. D. Coolon, and P. J. Wittkopp. 2013. Sources of bias in measures of allele-specific expression derived from RNA-seq data aligned to a single reference genome. BMC Genomics14:536–548 .

Stone, J.R., and G, A. Wray, G.A. 2001. Rapid evolution of cis-regulatory sequences via local point mutations. Mol. Biol. Evol. 18:1764–1770.

Suvorov A., V. Nolte, R. V. Pandey, S. U. Franssen, A. Futschik, et al. 2013. Intra-Specific Regulatory Variation in *Drosophila pseudoobscura*. PLoS ONE 8:e83547. doi:10.1371/journal.pone.0083547

Tarazona, S., F. Garcia-Alcalde, J. Dopazo, A. Ferrer, and A. Conesa. 2011. Differential expression in RNA-seq: A matter of depth. Genome Res. doi/10.1101/gr.124321.111.

Tirosh, I., S. Reikhav, A. A. Levy, and N. Barkai. 2009. A yeast hybrid provides insight into the evolution of gene expression regulation. Science 324:659–662.

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

Wayne, M. L., Y. J. Pan, S. V. Nuzhdin, and L. M. McIntyre. 2004. Additivity and transacting effects on gene expression in male Drosophila simulans. Genetics 168:1413–1420.

Wittkopp, P. J., B. K. Haerum and A. G. Clark. 2008. Regulatory changes underlying expression differences within and between Drosophila species. Nat. Genet. 40: 346–350.

Wittkopp, P. J., B. K. Haerum, and A. G. Clark. 2004. Evolutionary changes in *cis* and *trans* gene regulation. Nature 430:85–88.

Wittkopp, P.J., and P. Beldade. 2009. Development and evolution of insect pigmentation: Genetic mechanisms and the potential consequences of pleiotropy. Seminars in Cell and Developmental Biology 20:65–71.

Wray, G.A. 2007. The evolutionary significance of *cis*-regulatory mutations. Nat. Rev. Genet. 8:206–216.

Wu, T. D., and S. Nacu. 2010. Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics, 26:873–881.

TABLES AND FIGURES

Table 1. Summary of Illumina HISeq 2000 RNA sequencing scheme of 12 RNA samples that were collected from four parental inbred lines and their F_1 crosses. Two saline (SW1 and SW2) and two freshwater (FW1 and FW2) highly inbred lines and their F_1 crosses were reared in freshwater conditions (0 PSU) in order to distinguish between *cis*- and *trans*-regulation.

Inbred line/ F1 hybrid	Description	Biol. replicate	Insert size (bp)	Total reads (101 bp)	Total # of pairs (2 x 101 bp)
SW1	SW parental inbred line 1	1	332	83,510,946	37,436,178
SW1	SW parental inbred line 1	2	280	65,894,198	29,675,869
FW2	FW parental inbred line 2	1	289	74,424,014	33,332,865
FW2	FW parental inbred line 2	2	373	86,526,182	38,853,539
SW1 x FW2	F1 Hybrid (♀SW1 x ♂FW2)	1	289	73,042,830	32,804,829
SW1 x FW2	F1 Hybrid (♀SW1 x ♂FW2)	2	276	59,767,008	26,918,432
FW1	FW parental inbred line 1	1	399	64,756,320	29,700,683
FW1	FW parental inbred line 1	2	406	65,950,514	30,182,225
SW2	SW parental inbred line 2	1	398	87,103,524	40,026,799
SW2	SW parental inbred line 2	2	388	82,173,766	37,787,929
SW2 x FW1	F1 Hybrid (♀SW2 x ♂FW1)	1	391	46,491,844	21,331,202
SW2 x FW1	F1 Hybrid (♀SW2 x ♂FW1)	2	384	53,439,568	24,605,631

Table 2. Downregulated gene expression ($P_{adj} < 0.05$) in ***cis*** in parental line SW1 relative to FW2 parental line under freshwater conditions (0 PSU).

Gene description	logFC(P_{SW}/P_{FW})	logFC(H_{SW}/H_{FW})
Cytochrome c	-7.00	-5.70
S-antigen protein	-6.45	-4.44
60S ribosomal protein L31	-5.61	-7.36
Glutamate decarboxylase	-5.55	-4.27
Superoxide dismutase [Cu-Zn]	-4.79	-3.79
Transmembrane protein 192	-4.68	-6.71
Delta-like protein C	-4.24	-2.21
Thiamine transporter 1	-4.03	-5.23
Protein SHQ1 homolog	-3.84	-3.01
Cysteine and glycine-rich protein 2	-3.65	-1.41
Voltage-dependent anion-selective channel protein 2	-3.32	-4.39
Interferon regulatory factor 2-binding protein-like B	-2.97	-1.40
Transcription elongation factor SPT6	-2.72	-2.96
Na+/H+ exchanger 9B1	-2.42	-1.50
Short-chain specific acyl-CoA dehydrogenase, mt.	-2.29	-3.45
Putative L-cysteine desulfhydrase 1	-2.24	-2.34
LON peptidase N-terminal domain, RING protein 2	-1.94	-3.13
Heme-binding protein 1	-1.23	-0.73
Ornithine decarboxylase antizyme 1	-1.12	-1.45
Probable signal peptidase complex subunit 2	-1.11	-0.99
Probable sulfite oxidase, mitochondrial	-0.71	-1.54
Protein white	-0.67	-1.29

P_{adj} = p values adjusted for multiple comparisons by single-step method
P_{SW} and P_{FW} = parental saline and freshwater inbred lines alleles
H_{SW} and H_{FW} = saline and freshwater alleles in F_1 hybrid (♀SW1 x ♂FW2)

Table 3. Downregulated gene expression ($P_{adj} < 0.05$) in ***cis*** in SW2 relative FW1 parental inbred line under freshwater conditions (0 PSU).

Gene description	logFC(P_{SW}/P_{FW})	logFC($H_{SW}/_{FW}$)
FK506-binding protein 15	-6.77	-6.21
N6-adenosine-methyltransferase subunit METTL14	-5.99	-3.69
Xanthine dehydrogenase/oxidase	-4.59	-2.24
GrpE protein homolog, mitochondrial	-4.50	-3.24
Neurogenic locus Notch protein	-3.65	-2.39
Acylamino-acid-releasing enzyme	-2.99	-3.10
Transmembrane protein 41A-A	-2.90	-3.47
DDB1- and CUL4-associated factor 5	-2.45	-2.49
Nuclear cap-binding protein subunit 1	-2.43	-2.21
RE1-silencing transcription factor	-2.20	-1.34
Troponin C, isoform 1	-2.02	-1.71
Lissencephaly-1 homolog	-2.00	-1.08
tRNA methyltransferase 10 homolog A	-1.89	-1.72
Mitochondrial uncoupling protein 4	-1.66	-1.21
Calmodulin-4	-1.54	-1.40
Phosphatase and actin regulator 2	-1.48	-1.78
Conserved oligomeric Golgi complex subunit 7	-1.01	-1.25
ATP-dependent (S)-NAD(P)H-hydrate dehydratase	-0.74	-0.94
Katanin p60-ATPase-containing subunit A1	-0.62	-0.80
Tyrosine-protein phosphatase Lar	-0.61	-0.77
Receptor-type tyrosine-protein phosphatase delta	-0.60	-0.30
Probable aldehyde oxidase 3	-0.51	-0.64

P_{adj} = p values adjusted for multiple comparisons by single-step method
P_{SW} and P_{FW} = parental saline and freshwater inbred lines alleles
H_{SW} and H_{FW} = saline and freshwater alleles in F hybrid (♀SW2 x ♂FW1)

Table 4. Upregulated gene expression ($P_{adj} < 0.05$) in ***cis*** in SW1 parental line relative to FW2 parental line under freshwater conditions (0 PSU).

Gene description	logFC(P_{SW}/P_{FW})	logFC(H_{SW}/H_{FW})
Zinc finger protein 407	11.17	7.06
Probable aconitate hydratase, mitochondrial	10.28	6.40
Sushi, von Willebrand factor type A	8.20	5.15
Atrial natriuretic peptide receptor 2	7.58	6.11
Coiled-coil domain-containing protein 174	6.52	4.84
Paramyosin	5.64	4.39
Galactoside 2-alpha-L-fucosyltransferase 3	5.36	5.34
Glutathione S-transferase 2	5.34	6.29
Vitellogenin	5.07	4.00
Probable cytochrome P450 12b2, mitochondrial	4.94	3.15
Apolipoprotein D	4.84	9.40
Acidic chitinase	4.59	7.96
Leukotriene A-4 hydrolase	4.47	1.68
Serine/threonine-protein kinase 26	4.43	2.77
Carbohydrate sulfotransferase 11	4.40	6.57
Arylsulfatase B	3.93	5.46
Cysteine-rich, acidic integral membrane protein	3.53	2.93
Zinc finger protein 367	3.33	4.49
Cytochrome b5 type B	3.24	5.69
60S ribosomal protein L8	3.09	1.11
Intraflagellar transport protein 46 homolog	2.84	4.58
Xaa-Pro aminopeptidase 1	2.30	2.03
Glutathione reductase	1.94	1.76
Aminopeptidase Q	1.88	3.18
Early endosome antigen 1	1.42	3.66
Cell division cycle protein 20 homolog	1.35	0.82
3-hydroxyisobutyrate dehydrogenase, mitochondrial	1.00	1.38
Apolipoprotein D	0.99	2.20
Chymotrypsinogen A	0.97	0.45

P_{adj} = p values adjusted for multiple comparisons by single-step method
P_{SW} and P_{FW} = parental saline and freshwater inbred lines alleles
H_{SW} and H_{FW} = saline and freshwater alleles in F_1 hybrid (♀SW1 x ♂FW2)

Table 5. Upregulated gene expression ($P_{adj} < 0.05$) in ***cis*** in SW2 relative FW1 parental inbred line under freshwater conditions (0 PSU).

Gene description	logFC(P_{SW}/P_{FW})	logFC(H_{SW}/H_{FW})
Equilibrative nucleoside transporter 3	9.05	8.63
Actin cytoskeleton-regulatory complex protein SLA1	4.64	3.91
Potassium voltage-gated channel subfamily H 7	4.34	4.42
Metal-response element-binding transcription factor 2	3.60	2.99
Heat shock 70 kDa protein cognate 4	3.40	6.95
Coagulation factor V	3.35	5.56
Pro-epidermal growth factor	3.31	7.00
Prolyl 4-hydroxylase subunit alpha-2	3.06	5.69
Aggrecan core protein	2.96	3.79
DCN1-like protein 1	2.79	3.46
tRNA:m(4)X modification enzyme TRM13 homolog	2.63	2.94
Zinc finger protein 878	2.51	2.36
F-box only protein 21	2.25	2.00
Mitogen-activated protein kinase 13-A	2.13	3.28
Receptor-type tyrosine-protein phosphatase gamma	1.64	1.75
Clotting factor B	1.60	1.40
Secreted frizzled-related protein 5	1.59	2.13
Tyrosine-protein phosphatase Lar	1.54	1.15
Katanin p60-ATPase-containing subunit A1	1.54	1.81
Cystinosin homolog	1.52	1.42
Bifunctional glutamate/proline--tRNA ligase	1.51	1.36
Enamelin	1.39	1.30
Motile sperm domain-containing protein 2	1.37	2.37
Guanine nucleotide exchange factor for Rab-3A	1.36	1.32
DnaJ homolog subfamily C member 22	1.34	1.99
Acetylcholinesterase	1.20	2.92
Solute carrier organic anion transporter 4A1	1.17	1.71
Transcriptional coactivator YAP1	1.14	0.99
E3 ubiquitin-protein ligase LRSAM1	1.12	0.87
Receptor-type tyrosine-protein phosphatase zeta	0.86	1.63
N-acetyltransferase ESCO1	0.57	0.56

P_{adj} = p values adjusted for multiple comparisons by single-step method
P_{SW} and P_{FW} = parental saline and freshwater inbred lines alleles
H_{SW} and H_{FW} = saline and freshwater alleles in F_1 hybrid (♀SW2 x ♂FW1)

Table 6. Downregulated gene expression ($P_{adj} < 0.05$) in ***trans*** in SW1 vs. FW2 parental line comparison under freshwater conditions (0 PSU).

Gene description	logFC(P_{SW}/P_{FW})	logFC(H_{SW}/H_{FW})
Vitellogenin	-12.12	-0.06
Polypeptide N-acetylgalactosaminyltrans. 9	-10.35	0.16
Myohemerythrin	-10.06	-0.15
Contactin	-9.40	-0.37
Trans-1,2-dihydrobenzene-1,2-diol dehydrogenase	-7.72	0.08
Zinc finger protein 664	-7.32	0.10
Ubiquitin-conjugating enzyme E2 C	-7.05	-0.53
Probable G-protein coupled receptor Mth-like 10	-7.03	-0.13
E3 ubiquitin-protein ligase RNF185	-6.89	0.03
Nuclear hormone receptor family member nhr-3	-6.87	-0.09
RCC1 and BTB domain-containing protein 1	-6.85	0.25
Proactivator polypeptide	-6.62	0.65
Fatty acid-binding protein, epidermal	-6.42	0.06
Gelsolin	-6.11	0.03
Polysialoglycoprotein	-5.95	0.55
Transmembrane protein 68	-5.74	0.08
Glycine-tRNA ligase	-5.35	-0.42
U5 small nuclear ribonucleoprotein 200 kDa helicase	-5.28	-0.18
E3 ubiquitin-protein ligase NEDD4-like	-5.14	-0.55
Sequestosome-1	-5.13	-0.33
Dynein light chain 1, cytoplasmic	-4.94	-0.27
Hemocytin	-4.88	-0.04
Calcium and integrin-binding family member 3	-4.74	-0.13
Dynein light chain roadblock-type 2	-4.67	-0.32
Kinesin-like protein unc-104	-4.40	0.49
DNA replication licensing factor MCM6	-4.39	0.01
Glucose dehydrogenase [FAD, quinone]	-4.29	-0.26
Dipeptidyl peptidase 1	-4.27	-0.14
Inward rectifier potassium channel 4	-4.09	-0.48
Histone-lysine N-methyltransferase EHMT2	-4.03	0.22
Apoptosis inhibitor IAP	-4.00	-0.06
Unc-112-related protein	-3.98	-0.10
Golgi apparatus protein 1	-3.75	0.34
Neurofilament heavy polypeptide	-3.60	-0.51
Chymotrypsin	-3.59	0.50
Aryl hydrocarbon receptor nuclear translocator	-3.59	0.53
Protein phosphatase 1A	-3.49	0.50
Vitronectin	-3.40	-0.02
Nucleoside diphosphate kinase homolog 5	-3.32	-0.40

P_{adj} = p values adjusted for multiple comparisons by single-step method
P_{SW} and P_{FW} = parental saline and freshwater inbred lines alleles
H_{SW} and H_{FW} = saline and freshwater alleles in F_1 hybrid (♀SW1 x ♂FW2)

Table 6. Continued

Deoxynucleoside triphosphate triphosphohydrolase SAMHD1	-3.15	-0.09
Peroxisome assembly protein 12	-3.08	-0.10
BAI1-associated protein 3	-3.07	0.13
Histone RNA hairpin-binding protein	-3.00	0.45
NADP-dependent malic enzyme	-2.91	-0.02
Protein-methionine sulfoxide oxidase mical3b	-2.76	-0.24
Zinc finger protein 64 homolog, isoforms 1 and 2	-2.59	0.41
Protein SERAC1	-2.51	0.44
Innexin inx2	-2.48	0.12
NADPH-cytochrome P450 reductase	-2.43	-0.38
Alpha-parvin	-2.41	-0.41
G2/M phase-specific E3 ubiquitin-protein ligase	-2.38	-0.61
Heat shock 70 kDa protein cognate 3	-2.37	-0.09
Beta-1,3-glucan-binding protein	-2.34	-0.62
60S ribosomal protein L27a	-2.32	-0.52
Cyclin-L1	-2.27	-0.26
Fibroblast growth factor receptor 4	-2.24	0.04
SWI/SNF-related matrix-associated actin-dependent reg.	-2.24	0.07
TBC domain-containing protein kinase-like protein	-2.20	0.01
Serpin B4	-2.18	-0.33
Mevalonate kinase	-2.16	-0.03
Mini-chromosome maintenance complex-binding protein	3.75	-0.28
Endoglucanase E-4	-2.10	0.05
Transmembrane protein 258	-1.72	-0.02
Protein notum homolog	-1.52	-0.09
Transmembrane protein 198	-1.41	0.37
Glucose-6-phosphate 1-dehydrogenase	-1.32	0.13
Vitellogenic carboxypeptidase	-1.30	-0.60
Protein C19orf12 homolog	-1.25	0.43
Glutamate receptor ionotropic, kainate 5	-1.20	0.31
Bifunctional lysine-specific demethylase and hydroxylase NO66	-1.17	0.40
Endophilin-A1	-1.08	-0.43
Replicase polyprotein 1a	-1.04	-0.45
Corticotropin-releasing factor-binding protein	-0.92	-0.44
Apoptosis inhibitor 5 homolog	-0.82	0.29
Trypsin-1	-0.77	0.25
Methionine aminopeptidase 1D, mitochondrial	-0.75	-0.22
Glutamate carboxypeptidase 2	-0.72	0.12
Importin-7	-0.71	0.44
15-hydroxyprostaglandin dehydrogenase [NAD(+)]	-0.54	-0.18

Table 7. Downregulated gene expression ($P_{adj} < 0.05$) in ***trans*** in SW2 relative to FW1 parental line comparison under freshwater conditions (0 PSU).

Gene description	logFC(P_{SW}/P_{FW})	logFC(H_{SW}/H_{FW}
Protein flightless-1	-7.16	0.19
KIF1-binding protein	-6.35	-0.37
Peroxidasin homolog	-5.99	-0.99
Demethylmenaquinone methyltransferase	-5.19	0.44
Pyridoxal kinase	-4.80	-0.85
Arginine demethylase and lysyl-hydroxylase	-4.71	-0.57
Multiple inositol polyphosphate phosphatase 1	-4.42	0.37
Electrogenic sodium bicarbonate cotransporter 1	-4.40	-0.69
Calmodulin	-4.29	-0.31
TRP cation channel subfamily A member 1 homolog	-4.07	-0.34
Vitellogenin-2	-3.92	-0.19
Zinc finger protein 85	-3.46	0.20
Monoacylglycerol lipase ABHD12	-3.38	0.25
Zinc finger and BTB domain-containing protein 7C	-3.14	-0.29
DNA polymerase beta	-3.13	0.05
UDP-glucose 4-epimerase	-2.82	-0.33
2-amino-3-carboxymuconate-6-semia. decarboxylase	-2.79	-0.05
Transcriptional repressor CTCF	-2.61	-0.26
Serine/threonine-protein phosphatase 6 subunit	-2.48	0.22
Calcium-activated chloride channel regulator 2	-2.39	0.13
Cytochrome c oxidase subunit 7C, mitochondrial	-2.08	0.06
Proteasome subunit beta type-6	-1.89	0.19
ABC transporter G family member 23	-1.88	-0.05
Zinc finger protein 778	-1.88	0.14
Zinc finger protein 395	-1.44	0.05
ATP-binding cassette sub-family F member 2	-1.32	-0.72
RB1-inducible coiled-coil protein 1	-1.24	-0.20
Carbonic anhydrase 13	-1.03	-0.23
NFX1-type zinc finger-containing protein 1	-1.00	-0.01
RNA polyme I-spec. transcription init. factor RRN3	-0.97	-0.16
RE1-silencing transcription factor A	-0.92	-0.43
Rap guanine nucleotide exchange factor 2	-0.85	-0.25
Phosphatidylinositol 3,4,5-trisphosphate	-0.84	-0.04
Carboxypeptidase D	-0.77	-0.37
Filamin-C	-0.73	-0.25
Mannose-binding protein C	-0.67	-0.11

P_{adj} = p values adjusted for multiple comparisons by single-step method
P_{SW} and P_{FW} = parental saline and freshwater inbred lines alleles
H_{SW} and H_{FW} = saline and freshwater alleles in F_1 hybrid (♀SW2 x ♂FW1)

Table 8. Upregulated gene expression ($P_{adj} < 0.05$) in ***trans*** in SW1 relative to FW2 parental line comparison under freshwater conditions (0 PSU).

Gene description	logFC(P_{SW}/P_{FW})	logFC(H_{SW}/H_{FW})
Histone H1-delta	12.79	0.50
Innexin inx2	10.10	0.49
E3 ubiquitin-protein ligase UHRF1	7.98	0.20
Perlucin-like protein	7.84	-0.28
Leukotriene A-4 hydrolase	7.66	-0.02
ADP-ribosylation factor-like protein 5B	7.02	0.15
Probable nuclear transport factor 2	6.94	0.30
Sodium/potassium/calcium exchanger 6, mt.	6.64	-0.16
La-related protein 4	6.35	-0.59
Neuronal calcium sensor 2	6.28	0.11
E3 ubiquitin-protein ligase listerin	6.05	-0.20
Actin-related protein 10	5.99	-0.33
60S ribosomal protein L39	5.98	-0.24
TRPl cation channel subfamily A member 1 homolog	5.52	-0.07
Betaine--homocysteine S-methyltransferase 1	5.51	-0.38
Ethanolaminephosphotransferase 1	5.17	-0.27
Neural cell adhesion molecule 1	5.12	-0.13
Probable protein phosphatase methylesterase 1	5.04	-0.11
L-galactose dehydrogenase	4.94	-0.15
Probable nucleolar GTP-binding protein 1	4.86	-0.54
Dedicator of cytokinesis protein 7	4.77	0.47
Leptin receptor gene-related protein	4.69	-0.08
Repetitive proline-rich cell wall protein	4.57	0.11
Double-stranded RNA-specific editase Adar	4.54	-0.38
Dolichyl-diphosphooligosacch-protein glycosyltrans.	4.30	-0.04
Selenium-binding protein 1	4.29	0.45
Myosin light chain kinase, smooth muscle	4.28	-0.17
Methyl-CpG-binding domain protein 6	4.27	0.16
Eukaryotic translation initiation factor 3 subunit A	4.27	-0.33
Aladin	4.23	0.08
F-box/LRR-repeat protein 20	4.22	0.12
Max-binding protein MNT	4.22	0.01
Neuroligin-2	4.05	0.35
Myosin heavy chain, muscle	4.05	0.29
Pyrroline-5-carboxylate reductase 3	3.99	0.05
Zinc finger protein 142	3.95	-0.37
ATP-dependent RNA helicase DHX8	3.92	-0.25
Murinoglobulin-2	3.88	0.02

P_{adj} = p values adjusted for multiple comparisons by single-step method
P_{SW} and P_{FW} = parental saline and freshwater inbred lines alleles
H_{SW} and H_{FW} = saline and freshwater alleles in F_1 hybrid (♀SW1 x ♂FW2)

Table 8. Continued

Mini-chromosome maintenance complex-binding protein	3.75	-0.28
Protein aubergine	3.67	-0.60
Inositol 1,4,5-trisphosphate receptor	3.52	-0.25
Solute carrier family 22 member 6-A	3.50	0.50
Ubiquitin carboxyl-terminal hydrolase 7	3.44	-0.13
Ankyrin repeat domain-containing protein 29	3.42	0.10
HMG box-containing protein 1	3.28	0.19
Tolloid-like protein 1	3.17	0.27
TRP cation channel subfamily A member 1 homolog	3.08	-0.14
Glutamate-gated chloride channel	2.91	-0.18
Prostaglandin reductase 1	2.90	0.44
DNA topoisomerase 2-binding protein 1-A	2.89	0.04
Eukaryotic initiation factor 4A-III	2.87	0.43
Oxalate:formate antiporter	2.79	-0.54
3-hydroxyacyl-CoA dehydrogenase type-2	2.62	-0.22
Kyphoscoliosis peptidase	2.57	-0.23
Myosin heavy chain, muscle	2.46	0.00
Vacuolar protein sorting-associated protein 54	2.43	-0.05
Enamelin	2.33	0.19
Actin-5C	2.33	-0.21
Disks large 1 tumor suppressor protein	2.33	0.30
Activating molecule in BECN1-regulated autophagy protein 1	2.30	0.10
Tubby-related protein 3	2.25	-0.39
Regenerating islet-derived protein 3-alpha	2.01	0.41
STE20-related kinase adapter protein alpha	1.99	0.22
Short transient receptor potential channel 7	1.98	0.08
Plexin-A2	1.82	-0.19
Steroid receptor seven-up, isoforms B/C	1.78	0.01
Protein transport protein SEC61 subunit alpha	1.76	0.34
Oligosaccharyltransferase complex subunit ostc-B	1.71	-0.30
Chloride intracellular channel protein 6	1.69	0.44
Serine/threonine-protein phosphatase 2A subunit B alpha	1.58	0.14
Multidrug resistance-associated protein 1	1.52	-0.08
Serine/threonine-protein phosphatase PGAM5, mt.	1.50	0.22
Zinc metalloproteinase nas-39	1.45	-0.15
Sodium-dependent phosphate transporter 1-A	1.31	-0.13
Putative acyl-CoA dehydrogenase AidB	1.27	-0.03
Evolutionarily cons. signaling intermediate in Toll pathway, mt.	1.25	-0.53
Synaptobrevin homolog YKT6	1.19	0.15
Tricorn protease-interacting factor F3	1.04	-0.26

Table 9. Upregulated gene expression ($P_{adj} < 0.05$) in ***trans*** in SW2 vs. FW1 comparison under freshwater conditions (0 PSU).

Gene description	logFC(P_{SW}/P_{FW})	logFC(H_{SW}/H_{FW})
Circadian locomoter output cycles protein kaput	3.92	-0.02
Chaoptin	3.71	-0.01
Eukaryotic translation initiation factor 2D	3.62	-0.14
Nuclear pore complex protein Nup88	3.42	0.28
Protein bowel	3.40	0.43
KN motif and ankyrin repeat domain-contain. protein 1	3.10	0.17
Kin of IRRE-like protein 2	3.09	-0.42
Xanthine dehydrogenase	3.05	-0.14
Exocyst complex component 6	2.97	0.18
Lymphoid-specific helicase	2.92	-0.44
Nuclear hormone receptor family member nhr-213	2.73	0.00
RNA pseudouridylate synthase domain-cont. protein 2	2.71	-0.03
Protein FAM186A	2.60	-0.40
Nidogen-1	2.34	-0.14
F-box/LRR-repeat protein 20	2.29	0.06
Transcription factor SPT20 homolog	2.17	-0.34
Solute carrier family 22 member 21	2.15	-0.10
Centrosomal protein of 290 kDa	2.15	-0.05
Lysosomal Pro-X carboxypeptidase	2.14	-0.17
Cell division cycle protein 20 homolog	2.08	-0.02
Serine protease 30	2.06	-0.15
B-cell lymphoma 6 protein	2.02	-0.01
Zinc finger protein ubi-d4	1.98	-0.15
Facilitated trehalose transporter Tret1	1.98	-0.09
Protein dopey-1 homolog	1.87	-0.35
Nephrin	1.87	-0.12
DNA polymerase alpha subunit B	1.83	-0.14
Metalloreductase STEAP2	1.78	-0.39
Pre-mRNA-splicing factor SYF1	1.78	-0.08
Tubulin polyglutamylase complex subunit 2	1.77	-0.11
Nose resistant to fluoxetine protein 6	1.70	-0.40
Facilitated trehalose transporter Tret1	1.68	-0.17
Heat shock factor-binding protein 1	1.62	-0.26
Putative E3 ubiquitin-protein ligase SINAT1	1.57	-0.12
Alpha-tocopherol transfer protein	1.57	-0.03
Suppressor of cytokine signaling 4	1.56	-0.07

P_{adj} = p values adjusted for multiple comparisons by single-step method
P_{SW} and P_{FW} = parental saline and freshwater inbred lines alleles
H_{SW} and H_{FW} = saline and freshwater alleles in F_1 hybrid (♀SW2 x ♂FW1)

Table 9. Continued

CTD small phosphatase-like protein 2	1.54	-0.09
Tyrosine kinase receptor Cad96Ca	1.49	-0.08
60S ribosomal protein L39	1.46	-0.03
Transcriptional regulator ATRX homolog	1.46	-0.22
Serine/threonine-protein kinase PAK 2	1.45	-0.11
Transcriptional activator protein Pur-alpha	1.44	-0.28
Long-chain-fatty-acid--CoA ligase 4	1.40	-0.10
Cytochrome P450 2K4	1.38	-0.06
Delta-like protein A	1.36	-0.10
Mitotic checkpoint serine/threonine-protein kinase BUB1 β	1.33	-0.07
tRNA wybutosine-synthesizing protein 2 homolog	1.33	0.12
Nucleosome assembly protein 1-like 4	1.32	-0.17
28S ribosomal protein S14, mitochondrial	1.25	0.51
Copine-8	1.22	-0.06
Neuropeptide F receptor	1.20	-0.11
Superkiller viralicidic activity 2-like 2	1.20	-0.23
Heat shock protein 23	1.13	-0.10
APOBEC1 complementation factor	1.12	0.04
Aldehyde dehydrogenase family 9 member A1-A	1.12	0.07
Potassium voltage-gated channel protein Shab	1.11	-0.18
N-acetylglucosamine-6-phosphate deacetylase	1.11	-0.13
Peptide methionine sulfoxide reductase A5	1.08	-0.15
L-galactose dehydrogenase	1.05	-0.36
Deoxyribose-phosphate aldolase	1.03	-0.10
Putative glutathione S-transferase DHAR4	1.00	-0.14
LIM domain-binding protein 2	0.99	0.15
Zinc finger C2HC domain-containing protein 1C	0.87	-0.06
Aldehyde dehydrogenase X, mitochondrial	0.86	-0.07
MOB kinase activator 3A	0.85	-0.05
Forkhead box protein J1.2	0.78	-0.01
Longitudinals lacking protein-like	0.78	0.15
Galactoside 2-alpha-L-fucosyltransferase 1	0.78	-0.08
Protein TIPIN homolog	0.77	-0.18
Eukaryotic translation initiation factor 2D	0.76	-0.11
MFS-type transporter SLC18B1	0.74	-0.05
ATP-dependent RNA helicase DDX55	0.73	-0.04
F-box/WD repeat-containing protein lin-23	0.73	-0.15
Solute carrier family 12 member 5	0.71	-0.24
ABC transporter G family member 20	0.70	0.04
Eukaryotic translation initiation factor 4E	0.61	0.14
Inositol polyphosphate multikinase	0.61	0.01

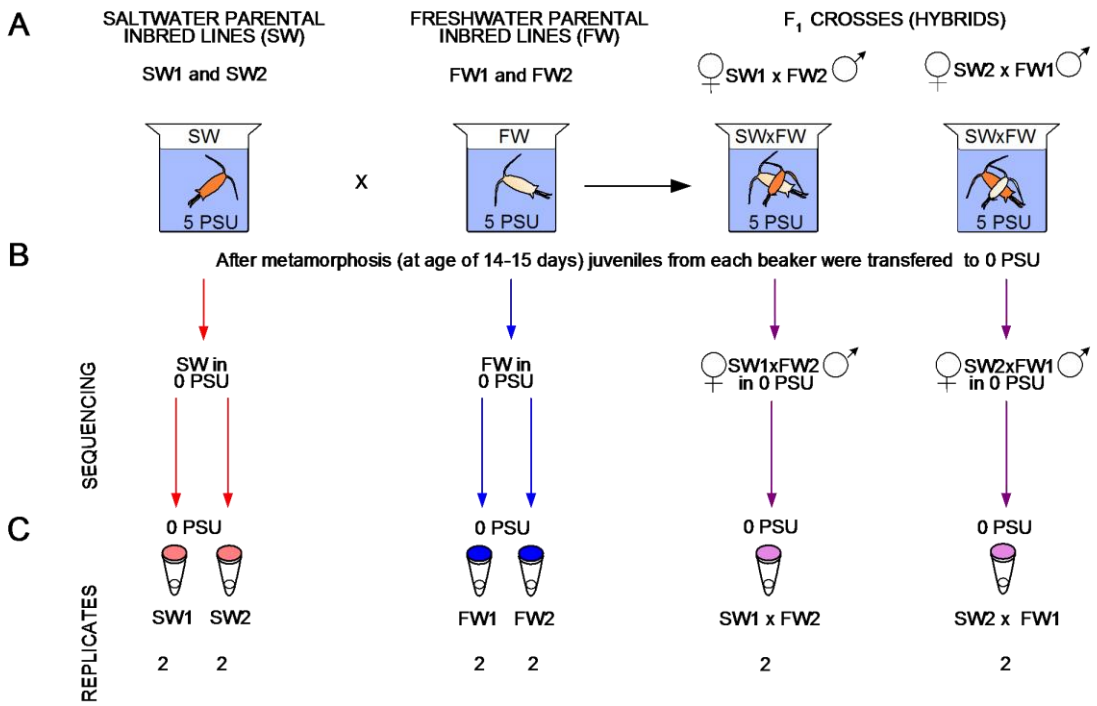

Figure 1. Common garden experiment for allele specific gene expression. **(A)** Juveniles from both saline and freshwater parental inbred lines (SW1, SW2, FW1 andFW2) were individually kept in 20 ml vials in 5 PSU water, and reared at this salinity until they became sexually mature. When their sex was determined females and males were paired to produce three different type of offspring: **SW x SW** (saline parental line), **FW x FW** (freshwater parental line) and ♀**SW** x ♂**FW** (F_1 crosses i.e., hybrids). The newly hatched offspring was separated from parents and reared at 5 PSU until metamorphosis (~15 days of age). **(B)** When offspring reached metamorphosis each sample was transferred to freshwater (0 PSU). (**C**) Juveniles were kept under freshwater conditions for the next 16-18 days i.e., until they became adults. The total of 50 adult copepods (25 females and 25 males) were randomly selected from each sample for total RNA extraction.

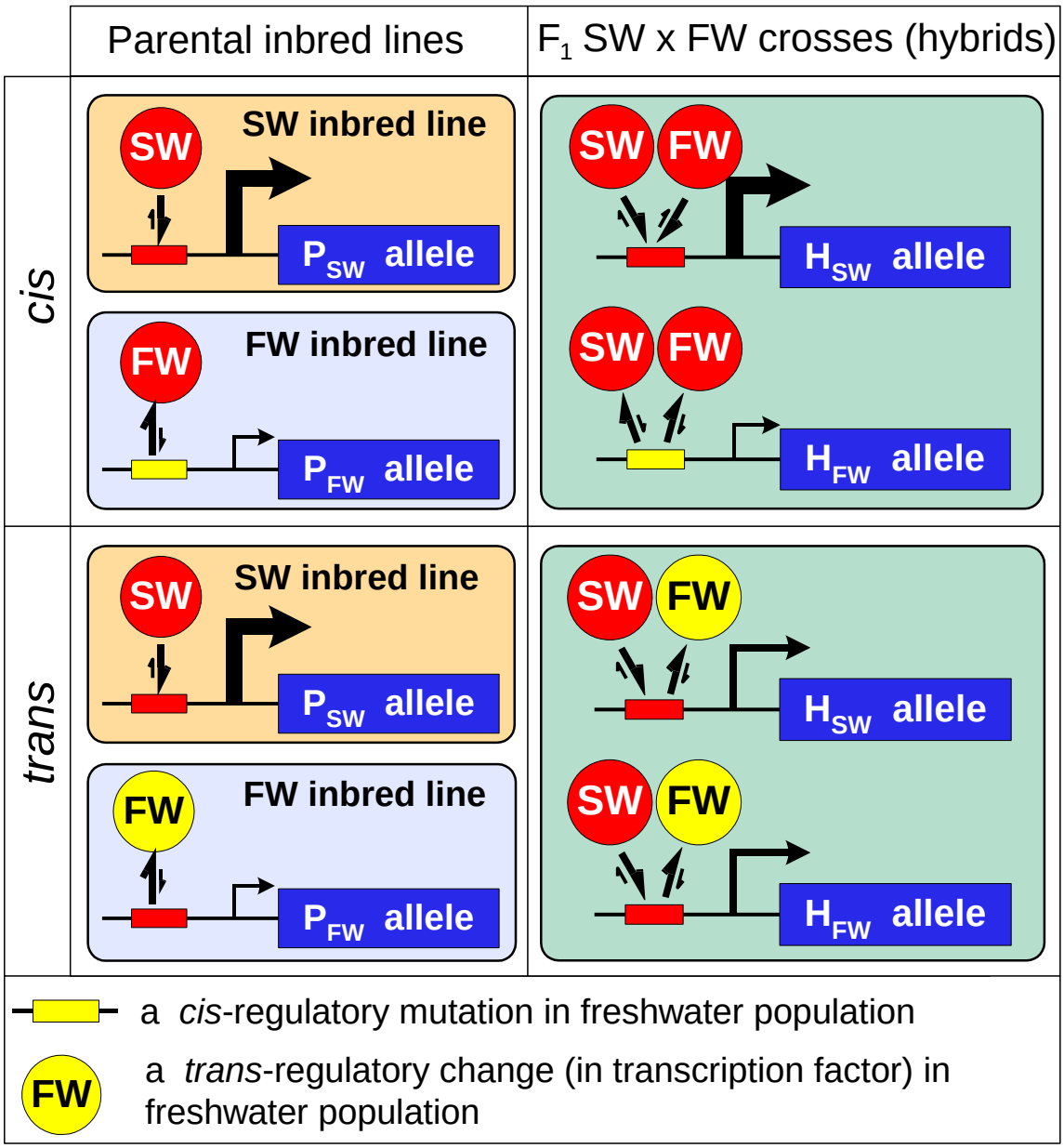

Figure 2. Hypothetical *cis*- and t*rans*-regulatory scenarios in the copepod *E. affinis* (adopted from McManus et al. 2010). Top panel: in hybrids, a *cis*-regulatory mutation in freshwater (FW) population (yellow box) decreases affinity for both the conserved saline (SW) and freshwater (FW) transcription factors (red circles). Bottom panel: a *trans*-regulatory mutation in a freshwater (FW) population transcription factor (yellow circle) reduces its binding affinity for the conserved freshwater (FW) and saline (SW) *cis*-regulatory regions (red boxes).
P_{SW} and P_{FW} – saline (P_{SW}) and freshwater (P_{FW}) allele in parental inbred lines.
H_{SW} and H_{FW} – saline (H_{SW}) and freshwater (H_{FW}) allele in F1 hybrids.

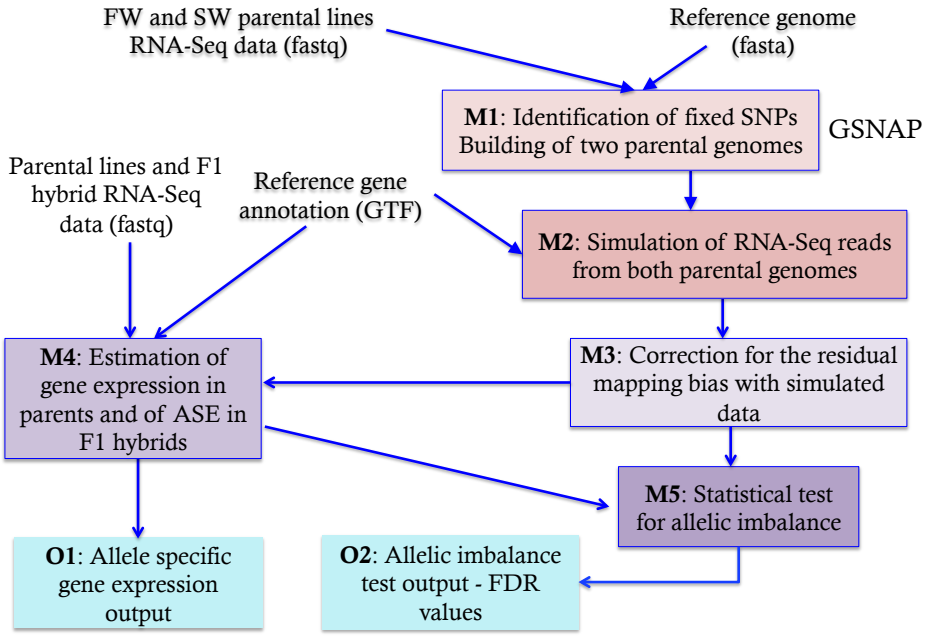

Figure 3. Five modules of Allelic imbalance metre (Allim) pipeline (adopted from Pandey et al. 2013). (1) Identification of fixed SNPs by mapping parental RNA-seq reads to reference genome and building two parental genomes by substitution of fixed SNPs positions in reference genome. The two prenatal genomes are identical to reference genome except at SNPs positions. (2) Simulation of RNA-seq reads from both parental genomes resulting in same number of reads at particular gene in both parents. (3) Estimation of the remaining mapping bias with simulated data (4) Estimation of gene expression in parental lines and allele-specific expression in F_1 hybrids. (5) Statistical test of significant allelic imbalance.

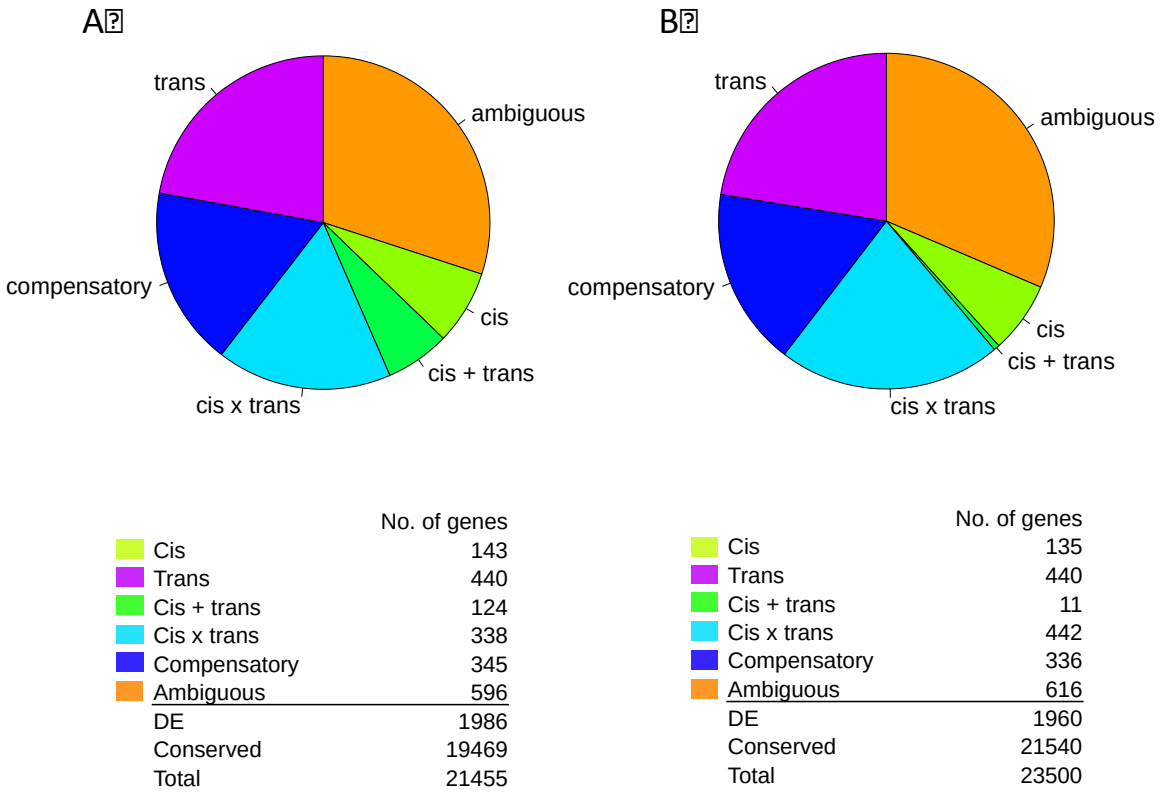

Figure 4. Distribution of gene regulatory categories in (**A**) SW1 *vs*. FW2 and (**B**) SW2 *vs*. FW1 comparison. Genes were classified into different regulatory categories according to Landry et al. (2005) and McManus et al. (2010) criteria (see Methods).