3-1-2018

Linking genotype to phenotype in a changing ocean: inferring the genomic architecture of a blue mussel stress response with genome-wide association

S. E. Kingston
Bowdoin College

P. Martino
Bowdoin College

M. Melendy
Bowdoin College

F. A. Reed
University of Hawai'i at Mānoa

D. B. Carlon
Bowdoin College

JOURNAL OF Evolutionary Biology

doi: 10.1111/jeb.13224

Linking genotype to phenotype in a changing ocean: inferring the genomic architecture of a blue mussel stress response with genome-wide association

S. E. KINGSTON*† , P. MARTINO*, M. MELENDY*, F. A. REED‡ & D. B. CARLON*†

*Department of Biology, Bowdoin College, Brunswick, ME, USA
†Schiller Coastal Studies Center, Bowdoin College, Brunswick, ME, USA
‡Biology Department, University of Hawai'i at Mānoa, Honolulu, HI, USA

Keywords:

calcification;
climate change;
genome-wide association;
genomic architecture;
Mytilus;
ocean acidification.

Abstract

A key component to understanding the evolutionary response to a changing climate is linking underlying genetic variation to phenotypic variation in stress response. Here, we use a genome-wide association approach (GWAS) to understand the genetic architecture of calcification rates under simulated climate stress. We take advantage of the genomic gradient across the blue mussel hybrid zone (*Mytilus edulis* and *Mytilus trossulus*) in the Gulf of Maine (GOM) to link genetic variation with variance in calcification rates in response to simulated climate change. Falling calcium carbonate saturation states are predicted to negatively impact many marine organisms that build calcium carbonate shells – like blue mussels. We sampled wild mussels and measured net calcification phenotypes after exposing mussels to a 'climate change' common garden, where we raised temperature by 3°C, decreased pH by 0.2 units and limited food supply by filtering out planktonic particles >5 μm, compared to ambient GOM conditions in the summer. This climate change exposure greatly increased phenotypic variation in net calcification rates compared to ambient conditions. We then used regression models to link the phenotypic variation with over 170 000 single nucleotide polymorphism loci (SNPs) generated by genotype by sequencing to identify genomic locations associated with calcification phenotype, and estimate heritability and architecture of the trait. We identified at least one of potentially 2–10 genomic regions responsible for 30% of the phenotypic variation in calcification rates that are potential targets of natural selection by climate change. Our simulations suggest a power of 13.7% with our study's average effective sample size of 118 individuals and rare alleles, but a power of >90% when effective sample size is 900.

Introduction

Climate change is applying enormous evolutionary pressure on natural populations (Parmesan, 2006). The biological response to climate change is of basic and applied interest. From a basic research point of view, the selective gradients driven by changes in the physical and biological environment provide tests of ecological and evolutionary theory. For example, how well does niche theory predict changes in species range limits (Pearson & Dawson, 2003; Wiens *et al.*, 2009)? Or, what are the relative roles of phenotypic plasticity vs. adaptation by natural selection in persistence (Hoffmann & Sgrò, 2011; Merilä & Hendry, 2014)? From the applied point of view, understanding the limits of plasticity and adaptation to climate change in specific systems has important economic consequences. For example, will crop yield decline in areas with historically high productivity (Lobell *et al.*, 2008)? A parallel question in the sea is whether a local fishing economy is likely to persist in the next 100 years, or migrate

Correspondence: Sarah E. Kingston, Bowdoin College, 6500 College Station, Brunswick, ME 04011, USA.
Tel.: +1 207 798 7128; fax: +1 207 725 3405;
e-mail: skingsto@bowdoin.edu

with target fish species that seek more favourable environmental conditions (Cheung *et al.*, 2010).

Whether a population can adapt to a changing environment ultimately depends on an interaction between individual-based phenotypic plasticity in traits that buffer stress, and the amount of standing genetic variation within populations for stress-reducing traits. In general, phenotypically plastic responses that buffer climate change appear to be commonplace in many systems, and plasticity is often used as the null hypothesis to test the rarer alternative of adaptation by natural selection (Merilä & Hendry, 2014). On the other hand, there are a number of examples where populations have adapted to spatial gradients in climate variables, but only a few examples of populations that have adapted to temporal gradients in climate variables (Gienapp *et al.*, 2008; Hoffmann & Sgrò, 2011). Demonstrating the potential for natural selection is technically and intellectually challenging, requiring three conditions hold in the system of interest: (i) that trait variation has a genetic basis, (ii) that trait variation has fitness consequences along the climatic gradient and (iii) trait evolution is not constrained by genetic architecture, including the effects of pleiotropy, epistasis and genetic correlations on the rate of evolution (Endler, 1986; Hoffmann & Sgrò, 2011). As many stress-related traits that affect organism fitness in changing environments are quantitative, the approach for determining the genetic basis and architecture of such 'climate change' traits is the framework of quantitative genetics. Unfortunately, the traditional tools of quantitative genetics may be difficult to apply to many natural systems of ecological or economic interest because relatedness cannot be easily manipulated or observed. For example, delayed maturity in long-lived species or early life history phases in which environments are very different from adults (i.e. many marine species) make laboratory crosses and the measurement of adult traits technologically difficult or logistically unfeasible. Further, collecting observational data on mating in the wild for pedigree construction is near impossible in a large widely distributed population if pollination/fertilization occurs by transport of microscopic gametes in wind or water. Establishing the potential for adaption to climate change by natural selection in these systems calls for new approaches.

Fortuitously, marker-based methods to estimate quantitative genetic parameters in natural populations have evolved rapidly during the last 10 years in parallel with advances in genomic technology. For example, 'marker-based heritability in the wild' (Lynch & Ritland, 1999) has enjoyed an increase in the precision of estimates of narrow-sense heritability (h^2 – the proportion of phenotypic variation due to additive genetic variation), as estimates of relatedness from genome-wide sampling are approaching the accuracy of estimates derived from observed pedigrees (e.g. Bérénos *et al.*, 2014). New 'omic' technology is also useful in identifying quantitative trait loci (QTL) that are important in climate change. Genome-wide association methods that were first employed to understand the genetic underpinnings of complex human diseases (GWAS) are now technically feasible in any system with the advent of RADseq and genotype by sequencing (GBS). These anonymous, next-generation sequencing-derived markers provide high-density population genomic sampling in any organism (Savolainen *et al.*, 2013). 'GWAS in the wild' has facilitated the establishment of the genetic basis of ecologically and evolutionarily important quantitative traits in systems with few genetic resources, for example cone serotiny as an adaptation to different fire regimes in lodgepole pine (Parchman *et al.*, 2012) and beak morphology as an adaption to food resources in Darwin's finches (Chaves *et al.*, 2016). These examples suggest that GWAS in the wild also has great promise for unravelling the genetic basis of new phenotypes under selection by climate change in many systems.

In this study, we use a GWAS in the wild approach to reveal the genetic basis of a trait predicted to be under strong, multifarious selection in the next 100 years. Our system includes two species of blue mussels, *Mytilus edulis* (Linnaeus, 1758) and *M. trossulus* (Gould, 1850), native to the Gulf of Maine (GOM) ecosystem. Few other marine environments on Earth are experiencing environmental change as rapidly as the GOM, a shallow, cold and productive body of water bordered by the Canadian Maritime Provinces in the north and New England in the south. Historically, the GOM has been the locus of some of the most productive fisheries in the world, once the centre of thriving groundfish fisheries – including cod, haddock, halibut and flounder (Christensen *et al.*, 2003) – is now dominated by the American lobster (Steneck *et al.*, 2011). In the last 20 years, the historically cold GOM has been experiencing one of the fastest temperature changes on Earth (Pershing *et al.*, 2015). This trend is consistent with global climate models that predict that the temperate oceans of the Northern Hemisphere will be the most strongly impacted by change in a suite of physical and biological variables, including temperature, ocean acidity, dissolved oxygen and productivity (Mora *et al.*, 2013). Within the GOM, as in other temperate marine environments, blue mussels are an important ecological system, historically forming dense monocultures in both intertidal and subtidal habitats (Lubchenco & Menge, 1978; Witman, 1987; Dudgeon & Petraitis, 2001) but now on the decline in rocky intertidal habitats during the last 40 years for reasons that are unclear (Sorte *et al.*, 2016). Economically, blue mussels are a substantial focus of aquaculture in the Canadian Maritimes, generating $49.5 million in value in 2013 (Fisheries and Oceans Canada). Despite the ecological and economic importance of blue mussels, there are relatively few existing genomic tools for *Mytilus edulis* or *M. trossulus*: transcriptomic resources have been

sequenced from a variety of *M. edulis* tissues (e.g. Philipp *et al.*, 2012; Freer *et al.*, 2013; Lesser & MacManes, 2016). Across the genus, there are greater resources: a draft genome is available for the Mediterranean species *M. galloprovincialis* (Murgarella *et al.*, 2016), as well as a transcriptome (Moreira *et al.*, 2015).

Our focal phenotypic trait is the net rate of calcification. Calcification is a fundamental physiological process in many marine organisms that provides the raw material for the organismal bauplan via shells and tests. In addition to providing the body plan, calcified structures buffer organismal physiology from change in the physical environment and provide defence from predators. Thus, we expect natural selection to maximize net calcification (calcification minus any $CaCO_3$ lost through dissolution) under increasing environmental stress. In the GOM, as in other temperate marine environments, calcifying marine organisms are facing multiple stressors as the planet warms. Perhaps most fundamentally, ocean acidity (OA) is increasing (J. Salisbury, pers. comm.) in conjunction with rapidly warming temperatures (Pershing *et al.*, 2015) and reduced ocean productivity (Balch *et al.*, 2012). These three factors have been shown to interact in complex ways to negatively affect growth rates (Kroeker *et al.*, 2013, 2016; Thomsen *et al.*, 2013). We therefore focus on the phenotypic variation in net calcification rates that is expressed under stress induced by these three environmental factors.

A potential limitation to using GWAS in new systems or traits is statistical power to detect QTL with potentially small effects. A working assumption is that most organisms are well adapted to long-term, stable conditions; however, there may be rare alleles segregating in the population that will be acted upon by selection as conditions change. For sufficient power to detect individual loci with intermediate effect sizes (0.1–0.2) and rare alleles, one needs to use fairly large sample sizes, in the order of hundreds to even thousands of individuals (Spencer *et al.*, 2009; Ball, 2011). We approach this potential limitation in two ways. First, we potentially increase overall power for a given sample size by leveraging a hybrid zone between two morphologically similar species – *M. edulis* and *M. trossulus*. Hybrids and backcrosses can substantially increase power by minimizing the effects of recombination on linkage equilibrium along larger blocks of chromosomes (Hoggart *et al.*, 2004; Patterson *et al.*, 2004). In theory, depending on levels of backcrossing, a larger number of genomic regions with moderate effect sizes can be detected with fewer samples; however, in some cases, population structure can generate spurious associations, so it is important to utilize conservative significance thresholds or cross-validate association inferences across model types (Pritchard *et al.*, 2000). Second, we maximize phenotypic variance and potential effect sizes of individual loci by quantifying calcification under multivariate stress in a climate change common garden, thereby potentially expressing $G \times E$ interactions and increasing the phenotypic variance (Hoffmann & Hercus, 2000). Our approach precedes in four steps: (i) maximizing genomic variation by broadly sampling wild mussels across the hybrid zone in the GOM, (ii) phenotyping net calcification rates in wild caught mussels grown in a climate change common garden, (iii) genotyping common garden mussels with reduced representation sequencing technique and (iv) applying a GWAS framework to the phenotypic and genotypic data This approach allowed us to identify a handful of genomic regions, one with confidence across models, with rare variants responsible for 30% of the phenotypic variation in calcification rates that are potential targets of natural selection by climate change. With simulations, we can show that our design has low but nonzero power to detect these rare variants and that a larger number of variants are likely segregating in wild populations that could be detected with increased sampling.

Materials and methods

Sampling the hybrid zone

In the Northern Hemisphere, *Mytilus trossulus* has a circumpolar distribution and is found on both the Pacific and Atlantic coasts. In contrast, *M. edulis* is restricted to the temperate Atlantic, occurring in subtidal and intertidal habitats on both American and European coasts (Gosling, 1992; Katolikova *et al.*, 2016). In the Gulf of Maine (GOM), the *M. edulis* × *trossulus* hybrid zone has a mosaic structure (reviewed in Riginos & Cunningham, 2005). Whereas populations found south of mid-coast Maine (south of Penobscot Bay) are overwhelmingly *M. edulis*, hybrid and mixed populations exist north of the mid-coast and in Canadian waters (Gosling, 1992; Hayhurst & Rawson, 2009). The mosaic nature of the hybrid interface is heavily related to sea surface temperatures (Yund *et al.*, 2015). Despite substantial genetic differentiation, the two species are difficult to distinguish morphologically (Rawson *et al.*, 1996). Sampling was designed to maximize geographic range within the GOM as well as equitable sampling of both species. In 2014, mussels were collected by hand ($N = 489$) from the intertidal and shallow sublittoral zones across seven sites in the Gulf of Maine (Table 1, Fig. 1) under Maine DMR special licence number ME 2014-41-00. Adult mussels between 25 and 60 mm were targeted to narrow the sampled age class range. Mussels were transported in aerated, chilled seawater from collection locations to the Schiller Coastal Studies Center (SCSC) in Harpswell, ME, within 48 h of collection. Holding tanks at the SCSC were 250-L table aquaria running on raw seawater flow-through. Outflow in holding tanks and from experiments was treated with

Table 1 Sampling locality information and sample sizes.

Sampling location	ID # on map	Latitude	Longitude	*N*	Year	# genotyped
Bailey Island (BI)	1	43.7238	−69.9933	59	2014	30
Phippsburg (PH)	2	43.7984	−69.8520	61	2014	30
New Meadows (NM)	3	43.9072	−69.8719	30	2014	14
Vinalhaven (VH)	4	44.0404	−68.8534	117	2014	82
Mount Desert Rock (MDR)	5	43.9686	−68.1283	74	2014	62
North Lubec, Cobscook Bay (CBL)	6	44.8875	−67.0198	93	2014	62
Eastport, Cobscook Bay (CBE)	7	44.8939	−67.0012	55	2014	42
Total				489		322

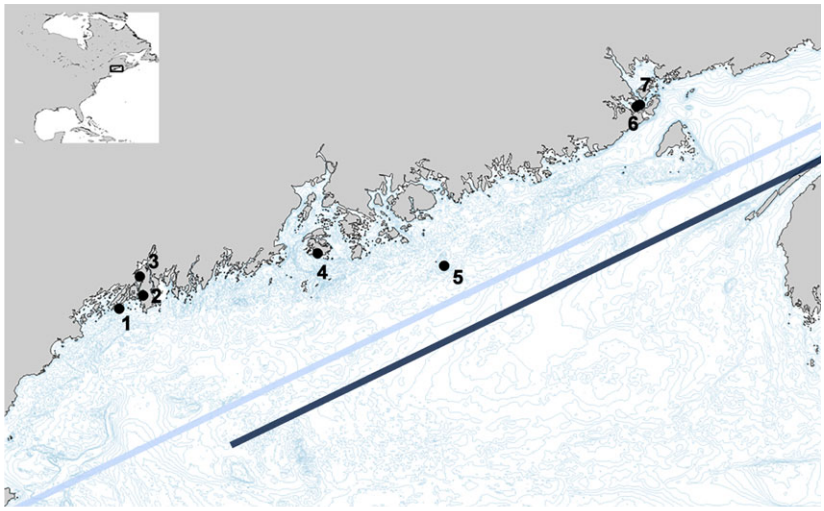

Fig. 1 Map of sampling locations along the coast of the Gulf of Maine (Maine to Canada). Approximate species distribution extents denoted with pale blue (*Mytilus edulis*) and dark blue (*M. trossulus*) lines parallel to shore.

57-W UV irradiation to prevent introduction of foreign biologicals into the local environment. Samples were acclimated for 5–14 days before the start of climate change common garden.

Climate change common garden and phenotyping

We used predictions from the global circulation models of Fogarty *et al.* (2008) and Mora *et al.* (2013) under high-emission scenarios to guide simulated physical and biological conditions likely to occur in the GOM by the year 2100. We used average summertime conditions of the epipelagic GOM as a reference and changed three variables relative to GOM reference points: we increased temperature by 3°C from 17°C, increased acidity by 0.2 units from pH 8.0 and decreased food supply by filtering out microplanktonic cells (>5 μm) from raw seawater (Table 2). To accommodate large sample sizes dictated by statistical power considerations of the GWAS, we conducted two consecutive runs of the common garden during the summer of 2014. Each run consisted of 2–4 individual 250-L sea tables with 4–6 L min^{-1} flow rates. In the climate stress common garden, constant pH was maintained by bubbling CO_2 gas into a 2000 L mixing reservoir controlled by a solenoid valve and electronic pH controller (Azoo; Takikong Corp, Taiwan). The mixing reservoir fed all downstream tables in the common garden. To limit food supply, raw seawater entering sea tables was serially filtered to 5 μm. Temperature was controlled within each 250-L tank with a system that included a 500-W submersible titanium heater (Finnex), submersible pump, temperature probe and electronic controller. To compare the impact of climate change conditions on phenotypic variance versus ambient summer conditions in the GOM, we ran a single 250-L 'ambient' tank during each run of the common garden by controlling temperature to 17°C (1/3-hp Arctica chiller; JBJ Chillers, St. Charles, MO, USA) and using the ambient pH of incoming seawater without modification. In ambient tables, we maintained a constant food supply by adding a diluted supply of Reed Mariculture Shellfish Diet 1800 (Campbell, CA, USA) at a rate of 2.24 g dry weight $tank^{-1}$ day^{-1} delivered with a peristaltic pump

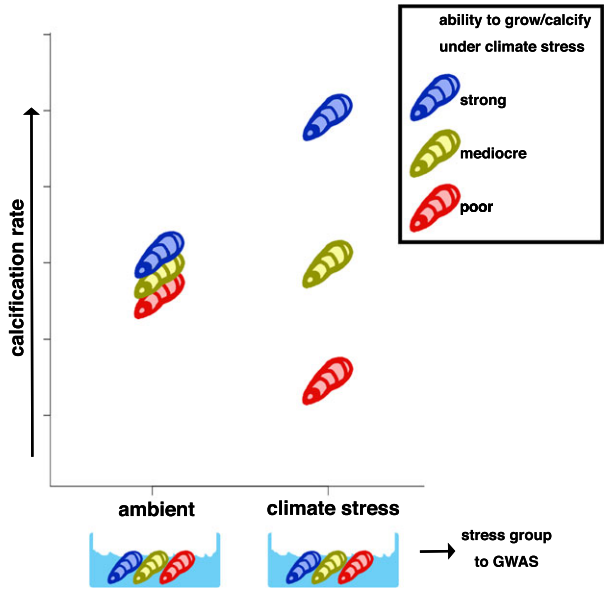

Fig. 2 Conceptual diagram depicting predicted effect of multifarious climate stress on variance in net calcification rate in mussels; greater variance under climate stress should increase power in the subsequent GWAS.

(after serial filtration of incoming seawater to 5 μm). At this concentration, and with a 4–6 L min^{-1} flow rate, the concentration of Chl a in ambient tanks is estimated to be ~250 mg m^{-3} before removal by mussels, which exceeds bloom conditions in Harpswell Sound, where Chl a reaches values of 12–16 mg m^{-3} (S. Drapeau, pers. comm.). In addition, the feeding rate per individual is just above the Reed Mariculture-recommended levels for bivalve aquaculture broodstock. Mussels from different geographic locations were allocated equally among tanks within runs, and the density of mussels within individual 250-L tanks ranged from 0.30 to 0.34 per litre. Combined, the two common garden runs resulted in 331 mussels exposed to climate change conditions and 158 mussels exposed to ambient conditions. Sample sizes for each tank and allocation of mussels among tanks are given in Table S1. We emphasize that only mussels exposed to the multivariate stress of the climate change common garden were used in the downstream GWAS as the evolutionary prediction is that phenotypic variation increases during stress (e.g. Hoffmann & Hercus, 2000) (Fig. 2). The two ambient tanks allowed us to test this prediction.

During the common garden runs, we monitored conditions within individual sea tables, including temperature, pH and aspects of carbonate chemistry. Temperature was recorded hourly using submersible Hobo Water Temperature Pro Data Loggers. Salinity and pH measurements were taken daily using a YSI 556 Sonde (Yellow Springs, OH, USA). Total alkalinity was measured from daily water samples using a Titrando 905 titration system (Metrohm, Switzerland) calibrated with Dickson Standard Batch 138 (Dickson, 1981; Dickson *et al.*, 2003). Saturation states for both calcite and aragonite were calculated using the software CO_2Sys (Pierrot *et al.*, 2006). A statistical summary of tank parameters is provided in Table S2.

We phenotyped the net calcification rate of individual mussels using buoyant weight, a sensitive technique that can detect small changes in the net calcification over daily intervals. Ries *et al.* (2009) have demonstrated the utility of buoyant weight for detecting physiological changes to climate stress in a diverse set of calcifying organisms, including blue mussels. They also verified a strong isometric relationship (R = 0.99) between dry shell weight and buoyant weight in their experiments with *Mytilus edulis*. Here, to take consecutive buoyant weights on the same individual, mussels were tagged individually with unique combinations of paint colours and numbered bee tags (The Bee Works, Oro-Medonte, ON, Canada) glued to one valve. Each mussel's buoyant weight was measured immediately before the start of the common garden run (day 0) and at the close of the run (day 14). The difference between the day 0 and day 14 measurements was normalized by the starting buoyant weight (to control for size effects) and used as the final scaled phenotype. Thus, net calcification is expressed as the fractional increase of the original starting buoyant weight. Buoyant weights were measured using a hanging weigh basket submerged in seawater suspended from Ohaus Adventurer microbalance model AR3130 (±0.002 g). Repeated weighing of a subset of individual mussels using the hanging basket rig revealed a precision of ±0.004 g or ±0.58% buoyant weight.

DNA extraction and quantification

At the end of each common garden run, we removed adductor muscle tissue from each mussel, froze it immediately at −20°C and then archived all tissues at −80°C. Both shell valves for each mussel were also archived at −20°C after tissue sampling.

Table 2 Targeted set points for physical and biological variables in the climate change common garden vs. ambient sweater conditions. See Table S2 for actual values and alkalinity statistics.

Treatment	Temperature (°C)	pH	Food concentration (mg dry weight mussel^{-1} day^{-1})
Ambient	17	8.0	2.6
Climate change common garden	20	7.7	0*

*Inflow water was filtered for particles >5 μm.

DNA from climate stress common garden tissue samples ($N = 322$) was purified using an E.Z.N.A. mollusc extraction kit (Omega Biotek, Norcross, GA, USA) with slight modifications to the factory protocol. Approximately 10 mg of adductor muscle tissue was subsampled from frozen tissue archives for each individual mussel. The tissue was minced and incubated overnight (shaking) at 60°C in 1× lysis buffer and 1.25 mg/mL final concentration of proteinase K. A 1:1 volume organic extraction was performed prior to kit purification using 24:1 chloroform:isoamyl alcohol; the aqueous layer was carried through to purification. Modifications to the published kit protocol were as follows: using a slightly smaller volume of lysis buffer (300 μL) and removing the second wash step before elution. RNase digestion was used to remove the RNA strands from the purified product. DNA was qualified using gel electrophoresis (Lonza, Rockland, ME, USA) and quantified via spectrophotometry (NanoDrop, Wilmington, DE, USA).

Genotype-by-sequencing (GBS) library preparation

Reduced complexity genomic libraries for each extracted mussel ($N = 322$) were generated using a genotype-by-sequencing (GBS) technique (Elshire *et al.*, 2011; Narum *et al.*, 2013). Our particular technique utilized a dual restriction enzyme digest (one rare, one frequent cutter), synthetic priming sites, inline single-end indices for multiplexing, and size selection to target hundreds of thousands of loci (Gompert *et al.*, 2012; Parchman *et al.*, 2012; Kingston *et al.*, 2017). Each sample was digested with both *Eco*RI (rare cutter) and *Mse*I (frequent cutter) restriction endonucleases; double-stranded adapters consisting of synthetic priming sites and 8- to 10-bp indices (*Eco*RI end), 322 unique index tags to identify each individual, were annealed to the fragmented DNA. Tagged libraries were amplified two separate times using PCR and high-fidelity Taq (iProof; Bio-Rad, Waltham, MA, USA). Size selection targeting 300- to 500-bp fragments from the libraries to reduce number of targeted loci was conducted via electrophoresis on a Pippin Prep platform (Sage Science, Beverly, MA, USA). All samples were pooled for sequencing by synthesis on the Illumina platform; the size distribution of the pooled library was assessed using an Agilent Bioanalyzer (Santa Clara, CA, USA) and quantified via qPCR. The pooled, multiplexed library was run on two lanes of rapid-read 1 × 100 sequencing on an Illumina Hi-Seq 2500 with on-board cluster generation (Harvard/MGH).

Analyses

Phenotypic variation

To test the prediction that phenotypic variance in net calcification increased during climate stress, we used a Bartlett's test for homogeneity of variances. We assessed any impact of sea table assignment (Table S1) and sampling locality using ANOVA. Additionally, a Welch's *t*-test was used to test for differences in mean calcification phenotype between ambient and climate stress common garden while avoiding the assumption of equal variances.

Read processing

Reads were demultiplexed and assigned to a file for each individual mussel based on the unique inline indices using the process_radtags command in the program Stacks (Catchen *et al.*, 2011, 2013). During this process, reads were concurrently filtered for quality (>90%), adapter contamination and appropriate restriction site and truncated to 82 bp.

De novo alignment and SNP calling

There is no publically available genome for *Mytilus edulis* nor *M. trossulus*, so we employed *de novo* assembly methods to create a catalogue of loci using the program Stacks (Catchen *et al.*, 2011, 2013). This catalogue was built from all sequenced individuals using the denovo_map.pl command; the minimum depth of coverage to create a 'stack', or potential locus consisting of the 82-bp sequencing read, for an individual was set at 2 reads. To account for high levels of expected heterozygosity common in molluscan species, the maximum distance allowed between loci was 8 bp. Any highly repetitive loci were removed. The model type used was a single nucleotide polymorphism (SNP) model, and the significance threshold for this model was set at $\alpha = 0.05$. Corrections to genotype calls were made using the rxstacks command and the catalogue of loci rebuilt using the same parameters listed for the denovo_map.pl command. The populations command was executed to further filter the loci used to call the final genotypes: at least two of the populations (designated by geographic sampling location) had to exhibit a locus, at least 20% of individuals had to exhibit calls at a locus, the minor allele frequency cut-off was set at 0.01 and only one SNP per 82-bp locus (stack) was used in the final catalogue (selected randomly if more than one SNP appeared in a stack). These lower filter values for minor allele frequency and per cent missingness were established to retain rare variants in the data set. Robustness of the GWAS using the lower filter values was compared more stringent filtering (to at least 60% of individuals in a population exhibiting calls at a locus).

Population parameters

Genetic differentiation, measured via F_{ST}, among geographic sampling locations was calculated from the final genotype calls (and bootstrapped 100×) in the R package diveRsity (R Core Team, 2013; Keenan, 2015). Inference of ancestral genomic signal with no *a priori*

population designations was performed on the final genotype calls using a variational Bayesian framework implemented in fastSTRUCTURE; 10 repetitions for each $K = n$ from 1 to 5 were run and the chooseK.py option utilized to infer the best K given the data (Raj *et al.*, 2014). Principal component analysis on filtered multilocus SNP genotypes (loci with at least 0.6 of each population exhibiting genotype calls) was performed using the R package SNPRelate (Zheng *et al.*, 2012).

Genome-wide association survey

Genome-wide association survey (GWAS) was performed using both a univariate linear mixed model and Bayesian variable selection regression. We chose the two-pronged analytical approach to leverage the advantage of the Bayesian framework to avoid multiple comparison limitations associated with evaluating one locus at a time while still allowing for comparison of this work to traditional frequentist GWAS analyses (Stephens & Balding, 2009). The univariate linear mixed model was executed in GEMMA (Zhou & Stephens, 2012), whereas the Bayesian variable selection regression was run in piMASS (Guan & Stephens, 2011). The Bayesian model was computed using Markov chain Monte Carlo (MCMC) with 1 000 000 burn-in steps and 100 000 000 sampling steps where every 10th step a sample is recorded. MCMC sampling was executed three times to assure convergence; Gelman and Rubin's convergence diagnostic was calculated using the coda package in R (Plummer *et al.*, 2006). Both models utilized the multilocus SNP genotype in mean genotype file format (dosage of minor allele) and calcification rate phenotype for each individual.

GWAS power analysis and simulations

A power analysis was conducted that focuses on the properties of the identified SNPs with the strongest associations. Very rare alleles of moderate effect are segregating in the population. This raises questions of how many additional variants are expected to exist and what sample sizes are required to discover them.

To be discovered, an allele must be sampled and then found in significant association with the phenotype. The rare nature of these alleles suggests that there is a significant chance of them not being sampled. Also, the allele frequencies are likely overestimated because of this discovery ascertainment bias. There is an average sample size of 117.7 individuals (with both phenotype and genotype) and a minor allele frequency of 0.01403 (Table 3). All of these minor alleles were found as homozygotes, which suggests significant F_{IS}. Therefore, on average the sample contained 1.6513 minor allele homozygotes. Assuming a Poisson distribution, the actual expectation of the number of homozygotes in the sample, h, is expected to be related to the observed number with ascertainment bias, h_a, by

$$h_a = \frac{\sum_{i=1}^{\infty} i \frac{h^i e^{-h}}{i!}}{\sum_{j=1}^{\infty} \frac{h^j e^{-h}}{j!}} = h\left(1 + \frac{1}{e^h - 1}\right).$$

This essentially adjusts the Poisson distribution by keeping track of the nonzero events. This can be solved for h as

$$h = h_a + W\left(-e^{-h_a} h_a\right),$$

where $W()$ is the Lambert-W function. Using this gives an adjusted average of 1.10354 minor allele homozygotes present in the sample and a corrected allele frequency of 0.00937853, suggesting the allele frequency estimate is increased by almost 50% because of ascertainment bias.

Random draws were made to create replicate data sets simulated from the summary information here. The phenotypes of individuals containing major alleles are drawn from a normal distribution matching the average mean and standard deviation of the 'common' phenotype in Table 3. The phenotypes of individuals containing minor alleles, drawn with a probability equal to the allele frequency adjusted to correct for ascertainment bias, are likewise drawn from a normal distribution matched to the mean and standard deviation of the 'rarer' phenotypes. These are simulated

Table 3 A summary of the sampling, genotype and phenotype data of the four significant loci. *N* is the number of samples with both genotypes and phenotypes. MAF is the minor allele frequency. Mean C and SD C are the mean and standard deviation of the phenotype values for the individuals that had the common allele. Mean R and SD R are the corresponding values for the phenotypes of individuals with the rare allele. Individ. are the IDs of the individuals that had the rare allele genotype.

Locus	*N*	MAF	Mean C	SD C	Mean R	SD R	Individ.
72070	119	0.01681	0.04836	0.56481	0.22406	0.76183	GR223, WH264
71586*	24	0.04167	−0.0141	0.22027	1.89985	NA	YE042
104243	151	0.01325	0.0302	0.18058	1.80885	0.12870	OR258, YE042
115310	83	0.01205	0.10457	0.30669	1.56774	NA	WHOR289
Ave*	117.7	0.01403	0.06103	0.35069	1.20022	0.44526	

*The phenotype data of individual YE042, coupled with a minor allele homozygous genotype, are shared between 71586 and 104243. This confounds the inference somewhat, and 71586 was not used in calculating the averages in the table above. However, the remaining three loci share similar properties.

under a range of sample sizes and tested for an association between genotype and phenotype using a standard linear regression and dummy variables for genotype. In a regression framework, this effect size is estimated, from simulation, as $r = 0.29$. A second set of simulations was also run with the effect size halved, $r = 0.15$, and quartered, $r = 0.073$, to assess the power of detecting additional loci of weaker effects.

Results

Climate stress in the common garden

We achieved a moderate level of climate stress in our common garden compared to the ambient conditions: temperature difference was 2.90–3.17 C, pH differential 0.14–0.34 units, difference in Ω_{calc} 0.54–1.17 and Ω_{arag} differential 0.33–0.73 (95% CIs, Table S2). Wilcoxon signed rank tests indicated our climate change common garden maintained significantly higher temperature and lower pH compared to ambient conditions (temperature $V = 2$ $P = 2.2E\text{-}16$, pH $V = 398$ $P = 9.4E\text{-}6$). Saturation states of both calcite and aragonite also differed significantly between ambient and climate stress (Ω_{calc} $t_{22} = 5.609$ $P = 1.5E\text{-}5$, Ω_{arag} $t_{22} = 5.651$ $P = 1.1E\text{-}5$). There was a 4% mortality rate overall.

Phenotypic variation

The climate change common garden resulted in greater phenotypic variation in net calcification rates compared to ambient conditions (Fig. 3; Bartlett's test for homogeneity of variances, $K_1^2 = 4.095$, $P = 0.043$). There was no significant impact of sea table (ambient $F_1 = 2.518$ $P = 0.115$; climate stress $F_3 = 0.217$ $P = 0.884$) or sampling location ($F_6 = 1.264$, $P = 0.277$). The mean net calcification rate within the climate change common garden was slightly greater (0.064, 1Q = −0.037 3Q = 0.115) than mussels in the ambient conditions (0.039, 1Q = −0.052 3Q = 0.104).

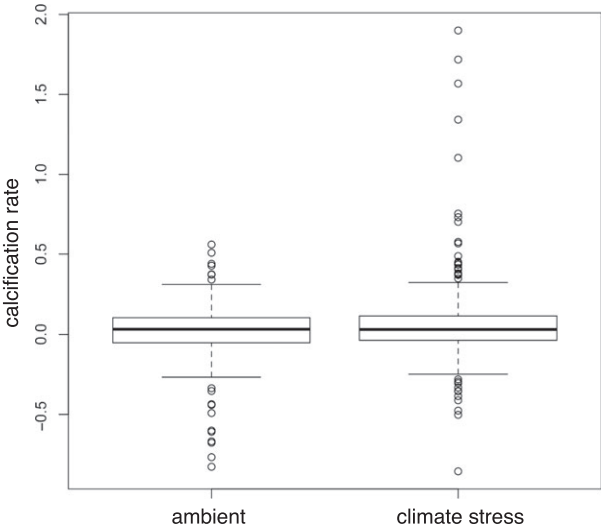

Fig. 3 Boxplot showing phenotypes of mussels (net calcification rate measured as size standardized change in buoyant weight) in the climate stress common garden versus mussel phenotypes in ambient conditions. Boxes represent the first and third quartiles, and whiskers extend to those quartiles ± 1.5 times the interquartile range. Variances are significantly different (Bartlett's test for homogeneity of variances, $K_1^2 = 4.095$, $P = 0.043$), and means are slightly, but significantly, different (Welch's *t*-test, $F_{1;333} = 6.835$, $P = 0.009$).

Sequencing, *de novo* alignment and genotyping

Sequencing rendered 378 249 305 100-base pair (bp) single-end reads. After demultiplexing and quality-filtering, 220 093 969 reads were retained for analysis. *De novo* assembly and downstream locus filtering generated genotype calls at 171 645 SNP loci among the 322 individual mussels. Average depth of coverage was 4.97 ± 1.14 per individual per locus (excluding missing genotype calls, mean ± SD) or 1.18 ± 0.70 (including missing genotype calls, mean ± SD).

Population genomics

Genetic differentiation (F_{ST}) among geographic populations was generally low with the exception of differentiation likely reflecting the species difference (Table 4). Inference of ancestry composition without *a priori* population information revealed two main ancestry clusters likely reflecting the *M. edulis–M. trossulus* split. Whereas there are mixed species populations, there are few intermediate genotype hybrid individuals as would be expected from an F_2 population (Fig. 4). PCA revealed differentiation between two main clusters, again likely reflecting species differences; geographic patterns beyond the species split emerged. Only a handful of individuals appear to be hybrid intermediates on the PCA, all originating from Mount Desert Rock or Cobscook Bay (Fig. 5). After population cluster inference, individuals parsed into majority *M. trossulus* or *M. edulis* ancestry groups (Fig. 4) do not exhibit significant differences in calcification rate by species ancestry (Welch's one-way comparison $F_{1;202} = 0.1543$, $P = 0.6949$).

Genome-wide association survey

The Bayesian variable selection regression (BVSR) model suggests underlying genomic architecture of the calcification rate trait is multilocus, but on the scale of 2–10 influential loci (median = 4, 2–10, 2.5%–97.5% quantiles; Fig. 6). Heritability of the trait was estimated at 0.30 (median, 0.19–0.41, 2.5%–97.5% quantiles; Fig. 6). The four loci with the greatest posterior

Table 4 F_{ST} values among geographic sampling populations of blue mussels in the Gulf of Maine (pairwise estimates bootstrapped, 95% CI).

	BI	PH	NM	VH	MDR	CBL	CBE
Bailey Island (BI)	–	0.0016–0.0025	0.0025–0.0029	0.0019–0.0031	0.0274–0.0280	0.1072–0.1077	0.3608–0.4757
Phippsburg (PH)		–	0.0072–0.0081	0.0051–0.0053	0.0447–0.0448	0.1322–0.1328	0.3916–0.5130
New Meadows (NM)			–	0.0048–0.0053	0.0278–0.0293	0.1071–0.1073	0.367300.4838
Vinalhaven (VH)				–	0.0492–0.0508	0.1440–0.1456	0.4074–0.5324
Mount Desert Rock (MDR)					–	0.0267–0.0278	0.2339–0.3139
North Lubec, Cobscook Bay (CBL)						–	0.1100–0.1527
Eastport, Cobscook Bay (CBE)							–

probability values in the Bayesian variable selection regression model and also significant in the univariate linear mixed model (at a $<4 \times 10^{-7}$) were searched against public databases as well as the *M. galloprovincialis* genome, but could not be annotated (Table 5). One of these four loci has excellent support across both models (104 243, Table 5); the other three do not exhibit as high posterior probabilities in the BVSR or similarly low p-values in the linear model. It is important to note that the two-model approach taken here aims to accomplish two goals: to describe the genomic architecture of the calcification rate trait and to identify genomic regions associated with the trait (and their effect sizes). The BVSR provides estimates of both architecture and association; the univariate linear model does not explicitly address multilocus architecture. Our data describe the genomic architecture – 2–10 influential loci with ~30% heritability of the trait – but likely do not identify all genomic regions associated with the trait. The most conservative estimate of associated regions identified would be the single 104 243 marker; given the small number of regions, however, it may be worth annotating all four loci, despite lower confidence in the other three associations. When the GWAS was repeated (BVSR) with more stringent missing data filters, the genomic architecture posterior probabilities were nearly identical and the top locus (104 243) was again returned with a similar posterior probability and effect size.

Gelman and Rubin's convergence diagnostic indicated good chain convergence among the MCMC runs (heritability 1.01, upper CI 1.02; number of SNPs 1.00, upper CI 1.00). Population structure is often cited as a potential confounding factor for genome-wide association; however, as our *M. edulis*-like genotypes and *M. trossulus*-like genotypes (as determined by fastSTRUCTURE) did not differ significantly in calcification rates, we did not use population designation as a covariate to maximize the power of our analysis given the scale of sampling available (Parchman *et al.*, 2012). Additionally, our two-model approach to screening significance of associated loci and our power analysis are precautions against possible spurious correlations related to species differences. We will need more samples from each taxon, *M. trossulus* in particular, to compare association inferences between pooled and split taxonomic groupings.

GWAS power analysis and simulation

The effective sample size used here, approximately 118 (due to incomplete genotype matrix), has low power, only 13.7% (at a $<4 \times 10^{-7}$ significance level), to detect loci with rare alleles and phenotype simulation parameter values matched to the observed effect size. This suggests that several similar loci, of moderate effect with rare variants, are yet to be discovered. To attain 50% power, approximately 310 individuals (effective size) need to be genotyped and phenotyped; a sample size of approximately 660 yields 90% power (at a $<4 \times 10^{-7}$ significance level). With larger future sample sizes, more stringent significance levels can be used to help filter out false associations; an effective sample size of 900 allows for 90% power with a significance level several orders of magnitude lower (Fig. 7).

Discussion

GWAS in the wild?

We found that under projected climate stress from multiple variables, blue mussels from the Gulf of Maine (GOM) exhibit extensive variability in calcification rate phenotype, and this variation is linked to a handful of loci of moderate effect. Others have successfully used a GWAS approach to determine the genetic basis of phenotypic traits in wild populations (for plants, see review by Alberto *et al.*, 2013), but to our knowledge, this is a first application of GWAS to hidden genetic variation expressed under simulated climate change. We have therefore established a genetic basis for variation in calcification rate under high temperature, low pH and low food supply. Our estimates of narrow-sense heritability for this key trait are in the order of 30% – indicating that substantial genetic variation for calcification under climate stress exists within these populations. Whether

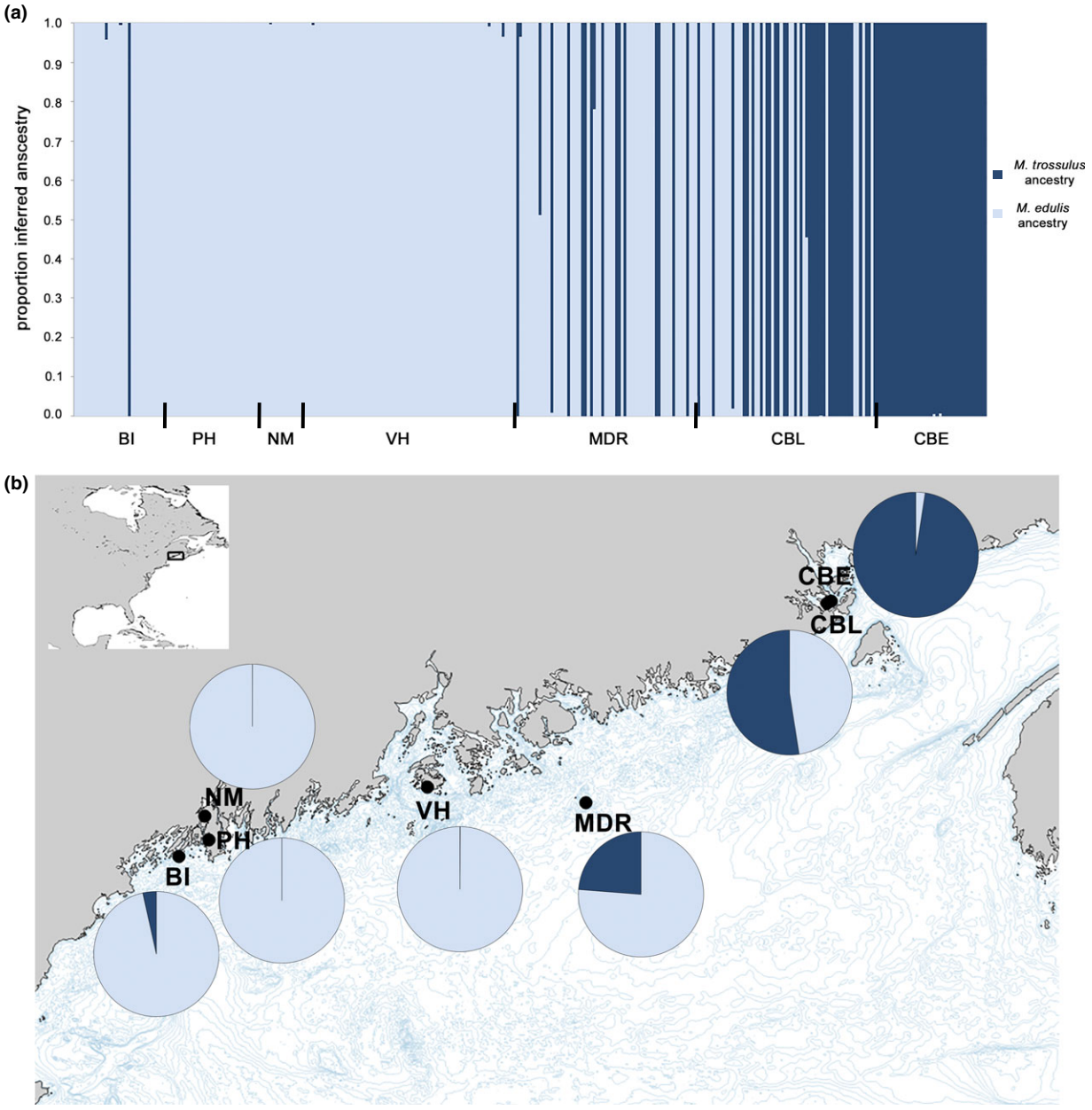

Fig. 4 (a) fastSTRUCTURE plot showing proportion of ancestry inferred from the full multilocus genotype on the *y*-axis where individual mussels are on the *x*-axis, grouped by population and ordered from south (left) to north (right). (b) fastSTRUCTURE-inferred ancestry signals, pooled by population, graphed geographically by collection site in the Gulf of Maine.

the mean calcification rate remains high due to natural selection on this quantitative trait over the next 100 years will depend on epistatic effects and the nature of genetic correlations with other traits that may themselves be under different forms of selection. However, classic breeding programs show that narrow-sense heritability tends to be predictive of response to selection even for complex traits (Hill *et al.*, 2008). We note that our basic approach can be used to identify the set of structural and regulatory genes that control calcification. When combined with gene expression profiling, GWAS can simultaneously discover gene expression QTL, or eQTL (Cookson *et al.*, 2009; Pritchard *et al.*, 2017), thus providing a gene regulatory network for calcification under climate stress. This extension of the basic GWAS concept, when combined with the results

of more traditional gene expression experiments that target subsets of the transcriptome under climate stress (e.g. Hüning *et al.*, 2013; Lesser, 2016), seems promising, because linkages among seemingly disparate metabolic pathways can be revealed.

The power to detect loci of moderate effects with a GWAS will increase when the phenotypic variance is maximal. We have shown that the phenotypic response under multivariate climate stress was significantly more variable than under more ideal control conditions. This pattern of increased variance in response to stress is common and implies environmental stressors can reveal phenotypic differences that would go undetected under ideal conditions (Latta *et al.*, 2015). Related to this increased variance under stress, environmental changes can uncover novel genetically determined phenotypes for selection to act upon (Waddington, 1956). This increased variance has a genetic component and was key to our strategy to detect associated loci; increased effect size leads to increased power. The inferred architecture of this trait – a handful of influential loci with moderate effect sizes – is similar to other genome-wide association inferences on the architecture of adaptive traits. In stickleback fishes, 10–18 loci were identified as significantly influencing the armoured plate

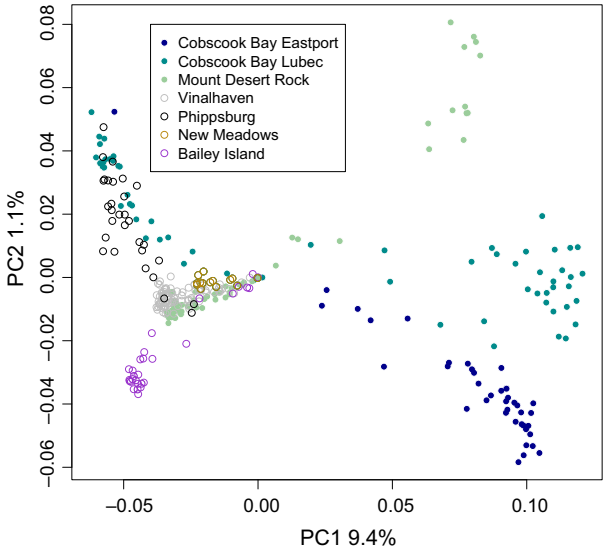

Fig. 5 Principal component analysis (PCA) based on filtered multilocus SNP genotypes (17 313 loci) across 322 mussel individuals, colour- and shape-coded by sampling location.

Table 5 Four loci associated with the stress phenotype calcification rate in blue mussels in the Gulf of Maine. Effect sizes and posterior probabilities (Rao-Blackwellized) are estimated through the Bayesian variable selection regression model, whereas the linear *P*-values are from the univariate linear mixed model (Wald test).

Locus ID	Effect size	Posterior probability	Linear *P*-value
104243	8.08	0.999	1.27E-13
115310	5.18	0.869	3.84E-07
71586	4.10	0.416	2.07E-09
72070	2.82	0.141	4.07E-08

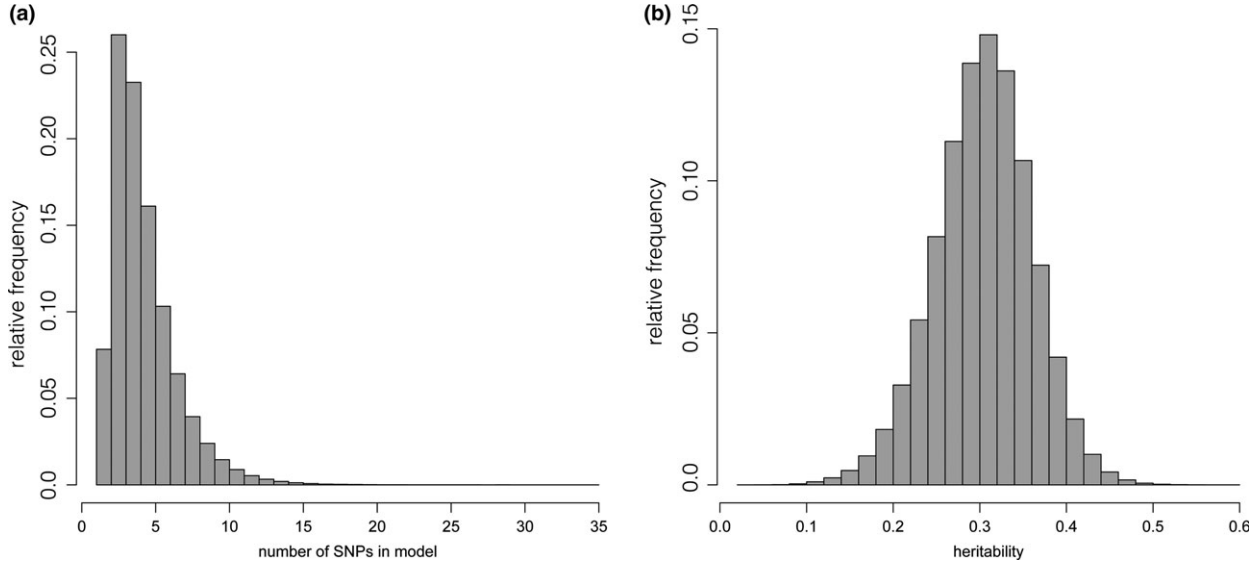

Fig. 6 Posterior distributions from the Bayesian variable selection regression model for the genome-wide association survey (GWAS). (a) Model estimate of the number of genetic loci underlying the calcification rate trait; (b) model estimate of the heritability of the calcification rate trait.

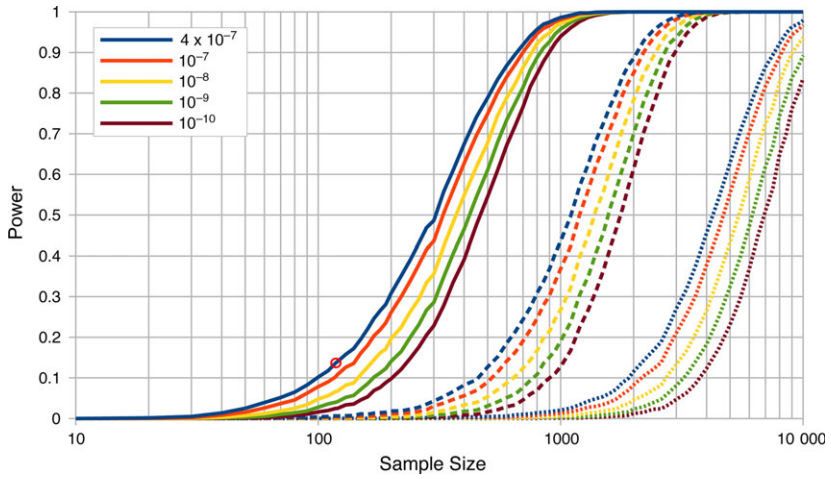

Fig. 7 Power to detect a genotype–phenotype association versus sample size. A total of 10 000 replicate simulations were run at each point by matching the allele frequency (corrected for ascertainment bias), average mean and average variance of the distributions in Table 3. The blue line corresponds to the significance level used here (4×10^{-7}). The red circle over the blue line corresponds to the average sample size (Table 3) and significance level used in this study. A series of increasingly stringent significance levels are indicated by the other colours for illustration. In this study, stress was used to increase the effect size and therefore power. The dashed lines are the power when the effect size is halved; the dotted lines correspond to an effect size that is quartered; this illustrates the types of sample sizes needed to detect loci or without using stress to increase the variance (in the order of 2000–3000 or 8000–10 000, respectively, in these examples). In terms of power to detect associations, effect size and sample size are more important considerations than stringent significance levels.

phenotype (Mazzarella *et al.*, 2016). In the case of the serotiny traits in lodgepole pine, 11 loci were associated with this important reproductive phenotype with effect sizes nearly identical to those estimated here (Parchman *et al.*, 2012). Our inferences, and those of others cited here, contrast with the pattern of one or two loci of large effect often demonstrated for sexually selected phenotypes, like avian plumage characteristics (Shapiro *et al.*, 2013; Bourgeois *et al.*, 2017). Whereas we have demonstrated clear success, it is also important to consider that the genetic architecture of any phenotypic trait lies on a spectrum of a few loci with major effects (as in the avian plumage examples cited above) to many loci of small effects. In the latter extreme, highly polygenic traits that follow Fisher's infinitesimal model (Fisher, 1919) will require other technical approaches to parse their genomic structure (Rockman, 2012).

Our simulations have revealed that the rarity of alleles linked to calcification rate has a major effect on our ability to detect 'climate change' loci. In fact, our simulations suggest that our experimental design had a power of about 10% given an effective sample size of $N = 100$ individuals (from 302 individuals phenotyped and genotyped at >170K SNPs scored across the entire population). Thus, it is not surprising that we found only a few associated SNPs (three with varying cross-model support). On the other hand, this relatively low power also suggests there are loci controlling calcification rate under stress in blue mussel genomes yet to be detected, perhaps in the order of 10s if sample sizes are increased to $N = 900$. Our simulations suggest increasing effective sample size provides the most dramatic boost to power (Fig. 7). Given our emphasis on application of GWAS on wild populations with few genomic resources, it is useful to consider the costs, measured in time and money, of effective sample sizes approaching 1000 individuals. Such an increase in sample size represents a 10× increase in sampling effort, experimental effort, DNA extraction, library preparation and sequencing. If we consider molecular benchwork and sequencing alone, we spent over $10K (USD) to obtain an effective samples size of just over 100 mussels. Even as sequencing costs decline, increasing funding to support this aspect of a project budget alone may not be a trivial task for many molecular ecology laboratories faced with reduced federal funding levels. Further, laboratory and aquarium facilities to conduct climate change simulations also represent significant resource commitments at this sampling scale. On the other hand, if we assume that many traits under selection by climate change are under stabilizing or weak directional selection, our results of rare alleles may be general, and a combination of large sample sizes and field sampling efforts that maximize the probability of including these genotypes in the study population are recommended.

It is also important to consider the downstream steps required to verify that the variation detected in a GWAS is linked to the trait of interest. In our case, we are interested in further development of blue mussels as a model to study links between genotype and quantitative phenotypes under selection by changing

environments. Thus, genome sequencing will be required to map the SNPs discovered here, and annotate linked genes. At present, a draft genome has been assembled for the congener *Mytilus galloprovincialis* (Murgarella *et al.*, 2016), but *Mytilus* genomes are large and repetitive, suggesting assembly problems may arise even among congeners. Linking transcriptome sequencing with GWAS to find expression QTL (eQTL) is another approach to fine scale mapping of the putative loci identified by the GWAS. In this case, targeted sequencing within and near expressed genes that correlate with SNPs can identify the location and possible function of a QTL.

Complex selection seascapes

Unlike the open ocean, coastal and intertidal environments are characterized by greater diel and seasonal physical and biological variability. Climate change is predicted to increase the amplitude of these fluctuations, pushing the limits of physiological acclimatization (Johnson *et al.*, 2013; Helmuth *et al.*, 2016). Consequently, selective pressure on organisms in these coastal ecosystems is layered across complex spatial and temporal scales. For example, some sites may experience different combinations of variables that cause homoeostatic stress. Further, the succession of sets of variables that stress physiology is also likely to be spatially variable. Biogeographic changes in coastal population distributions in response to changing climate are likely to mirror this complexity (Wethey *et al.*, 2011). We have demonstrated that a population of wild blue mussels in the Gulf of Maine exhibits the heritable phenotypic variance required to calcify in projected climate conditions, including high temperature, low pH and reduced productivity. Individual mussels exhibiting the potentially adaptive alleles are not restricted to a single portion of the geographic distribution within the Gulf of Maine; variants are found throughout the geographic sampling span. However, these alleles are associated with majority *M. edulis* ancestry individuals. It is important to note that many of these variants are found in mixed populations and that we ended up sampling fewer majority ancestry *M. trossulus* individuals (only ~88/322; Fig. 4a). We have not fully quantified the physical characteristics of these sites, but mapping the geographic distribution of rare climate change alleles and understanding the local environmental seascape (e.g. Kroeker *et al.*, 2016) is an interesting future direction.

A recent assessment of vulnerabilities of shellfisheries in coastal communities places the Gulf of Maine waters in high-impact categories (Ekstrom *et al.*, 2015). In the state of Maine, lack of diversity in economically important fisheries is cited as a liability in a changing climate (Steneck *et al.*, 2011). Bivalve aquaculture has the potential contribute to sustainable and economically viable food production, but as acidification trends continue, both wild and captive populations of marine calcifiers face settlement and growth challenges (Green *et al.*, 2009). Standing genetic variation, including the loci we have identified here, has the potential to provide the raw material for sustainable aquaculture in the Anthropocene. Yet, we also note that more work will be need to be done that identifies genetic correlations with other desirable traits for aquaculture, for example the ratio of shell to tissue mass, or reproductive allocation. These correlations are also the key to understanding the multivariate response to climate change in quantitative characters (Lande, 1980). With some of the stress-calcification-associated loci identified, we can continue to untangle the molecular architecture of the stress calcification trait in the context of aquaculture application; expression quantitative loci (eQTL) associated with adaptive traits, both climate stress and those relevant to success in culture, can help piece together regulatory networks relevant to increasing success and yield of aquacultured blue mussels in an increasing warmer and more acidic environment (Cookson *et al.*, 2009; Pritchard *et al.*, 2017). In addition to further exploration of the expression network in adults, it will also be important to characterize similarities and differences in juvenile stress response at settlement.

Despite recent development of genomic resources for other bivalve species, little public genome-scale data are available for *Mytilus edulis* and *Mytilus trossulus* (Zhang *et al.*, 2012; Timmins-Schiffman *et al.*, 2013; Lockwood *et al.*, 2015; Moreira *et al.*, 2015; Lesser & MacManes, 2016; Murgarella *et al.*, 2016; Yarra *et al.*, 2016; Sun *et al.*, 2017; Wang *et al.*, 2017). Our GWAS in the wild approach demonstrates viability of this line of investigation for nonmodel species with few genomic resources. We have identified a number of loci that could potentially respond to selective pressures induced by climate change. As none of these identified loci are annotated, it will be important to further investigate functional roles of these genomic regions. Elucidating the genomic architecture underlying calcification under stress is not just an interesting evolutionary exercise, but has much potential for improving aquaculture yields in an increasingly changing world.

Conclusions

Although the power of highly multilocus SNP analysis has greatly expanded the quantitative genetic toolbox (Parchman *et al.*, 2012; Bérénos *et al.*, 2014), our simulation suggests we are working close to the theoretical minimum sampling required to identify any genomic regions associated with our phenotype of interest. As the waters in the Gulf of Maine continue to change rapidly, this success of this genomic tool to efficiently assess adaptive potential among wild populations is notable. Our dual-model assessment approach allows

inference of the genomic architecture of the trait, even if all associated loci are not initially discovered. We provide a power analysis which can be applied to other systems where rare, advantageous alleles are likely to exist in wild populations. We offer an assessment of trade-offs (and pay-offs) involved in balancing resources, time, geographic sampling range and statistical power; our execution of GWAS in the wild exemplifies an analytical framework suitable for both basic research and applied questions.

Acknowledgments

We express great thanks to those who offered access to sampling locations and corresponding hospitality: Alice Anderson, Caitlin Cleaver, Lisa Dellwo, Dan DenDanto, Damon Gannon, Barney Hallowell, Sam Hallowell, Mark Murray, Bill Schlesinger, Laura Sewall, Toby Stephenson, Sean Todd, and Christopher Tremblay. We are also grateful to those who helped gather mussels in the field: Ben Eisenberg, Damon Gannon, Janet Gannon, Laura Sewall, Aidan Short, and Jenna Watling. We also thank those who contributed to monitoring the tanks and sampled mussel tissues: Steven Allen, Ben Eisenberg, Aidan Short, and Jenna Watling. We thank Michèle LaVigne for guidance on and lab space for the alkalinity titrations. Funding from Bowdoin College made this project possible; funding from Maine INBRE supported equipment and supplies. For the duration of this effort, S. Kingston was funded by the Doherty Marine Biology Postdoctoral Scholar position at Bowdoin College. P. Martino was supported by a Bowdoin Biology Departmental Fellowship. This study is a Bowdoin Schiller Coastal Studies Center Publication (#4).

Data sharing and archiving statement

The data generated by this study are archived in DRYAD Digital Repository (doi:10.5061/dryad.2d8b5): demultiplexed short sequence reads for each sample (.fastq files), the post-assembly genotype calls (.vcf format) and phenotypes with associated metadata for each sample.

References

Alberto, F.J., Aitken, S.N., Alia, R., Gonzalez-Martinez, S.C., Hänninenk, H., Kremer, A. *et al.* 2013. Potential for evolutionary responses to climate change – evidence from tree populations. *Glob. Change Biol.* **19**: 1645–1661.

Balch, W.M., Drapeau, D.T., Bowler, B.C. & Huntington, T.G. 2012. Step-changes in the physical, chemical and biological characteristics of the Gulf of Maine, as documented by the GNATS time series. *Mar. Ecol. Prog. Ser.* **450**: 11–35.

Ball, R.D. 2011. Experimental designs for robust detection of effects in genome-wide case–control studies. *Genetics* **189**: 1497–1514.

Bérénos, C., Ellis, P.A., Pilkington, J.G. & Pemberton, J.M. 2014. Estimating quantitative genetic parameters in wild populations: a comparison of pedigree and genomic approaches. *Mol. Ecol.* **23**: 3434–3451.

Bourgeois, Y.X.C., Delahaie, B., Gautier, M., Lhuillier, E., Malé, P.-J.G., Bertrand, J.A.M. *et al.* 2017. A novel locus on chromosome 1 underlies the evolution of a melanic plumage polymorphism in a wild songbird. *R. Soc. Open Sci.* **4**: 160805.

Catchen, J., Amores, A., Hohenlohe, P., Cresko, W. & Postlethwait, J. 2011. Stacks: building and genotyping loci *de novo* from short-read sequences. *G3 (Bethesda)* **1**: 171–182.

Catchen, J., Hohenlohe, P., Bassham, S., Amores, A. & Cresko, W. 2013. Stacks: an analysis tool set for population genomics. *Mol. Ecol.* **22**: 3124–3140.

Chaves, J.A., Cooper, E.A., Hendry, A.P., Podos, J., De León, L.F., Raeymaekers, J.A. *et al.* 2016. Genomic variation at the tips of the adaptive radiation of Darwin's finches. *Mol. Ecol.* **25**: 5282–5295.

Cheung, W.W., Lam, V.W., Sarmiento, J.L., Kearney, K., Watson, R.E.G., Zeller, D. *et al.* 2010. Large-scale redistribution of maximum fisheries catch potential in the global ocean under climate change. *Glob. Change Biol.* **16**: 24–35.

Christensen, V., Guenette, S., Heymans, J.J., Walters, C., Watson, R., Zeller, D. *et al.* 2003. Hundred-year decline of North Atlantic predatory fishes. *Fish Fish.* **4**: 1–24.

Cookson, W., Liang, L., Abecasis, G., Moffatt, M. & Lathrop, M. 2009. Mapping complex disease traits with global gene expression. *Nat. Rev. Genet.* **10**: 184–194.

Dickson, A.G. 1981. An exact definition of total alkalinity and a procedure for the estimation of alkalinity and total inorganic carbon from titration data. *Deep Sea Res. A Oceanogr. Res. Pap.* **28**: 609–623.

Dickson, A.G., Afghan, J.D. & Anderson, G.C. 2003. Reference materials for oceanic CO_2 analysis: a method for the certification of total alkalinity. *Mar. Chem.* **80**: 185–197.

Dudgeon, S. & Petraitis, P.S. 2001. Scale-dependent recruitment and divergence of intertidal communities. *Ecology* **82**: 991–1006.

Ekstrom, J.A., Suatoni, L., Cooley, S.R., Pendleton, L.H., Waldbusser, G.G., Cinner, J.E. *et al.* 2015. Vulnerability and adaptation of US shellfisheries to ocean acidification. *Nat. Clim. Chang.* **5**: 207–214.

Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S. *et al.* 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* **6**: PMC3087801.

Endler, J.A. 1986. *Natural Selection in the Wild*, illustrated edn. Princeton University Press, Princeton, NJ.

Fisher, R.A. 1919. The correlation between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edinb.* **52**: 399–433.

Fogarty, M., Incze, L., Hayhoe, K., Mountain, D. & Manning, J. 2008. Potential climate change impacts on Atlantic cod (*Gadus morhua*) off the northeastern USA. *Mitig. Adapt. Strat. Glob. Change* **13**: 453–466.

Freer, A., Bridgett, S., Jiang, J. & Cusack, M. 2013. Biomineral proteins from *Mytilus edulis* mantle tissue transcriptome. *Mar. Biotechnol.* **16**: 34–45.

Gienapp, P., Teplitsky, C., Alho, J.S., Mills, J.A. & Merilä, J. 2008. Climate change and evolution: disentangling environmental and genetic responses. *Mol. Ecol.* **17**: 167–178.

Gompert, Z., Lucas, L.K., Nice, C.C., Fordyce, J.A., Forister, M.L. & Buerkle, C.A. 2012. Genomic regions with a history of divergent selection affect fitness of hybrids between two butterfly species. *Evolution* **66**: 2167–2181.

Gosling, E.M. 1992. Systematics and geographic distribution of *Mytilus*. In: *The Mussel Mytilus: Ecology, Physiology, Genetics and Culture*, Developments in Aquaculture and Fisheries Science (E.M. Gosling, ed.), pp. 1–20. Elsevier Science Publishing, Amsterdam.

Green, M.A., Waldbusser, G.G., Reilly, S.L. & Emerson, K. 2009. Death by dissolution: sediment saturation state as a mortality factor for juvenile bivalves. *Limnol. Oceanogr.* **54**: 1037–1047.

Guan, Y. & Stephens, M. 2011. Bayesian variable selection regression for genome-wide association studies and other large-scale problems. *Ann. Appl. Stat.* **5**: 1780–1815.

Hayhurst, S. & Rawson, P.D. 2009. Species-specific variation in larval survival and patterns of distribution for the blue mussels *Mytilus edulis* and *Mytilus trossulus* in the Gulf of Maine. *J. Molluscan Stud.* **75**: 215–222.

Helmuth, B., Choi, F., Matzelle, A., Torossian, J.L., Morello, S.L., Mislan, K. *et al.* 2016. Long-term, high frequency *in situ* measurements of intertidal mussel bed temperatures using biomimetic sensors. *Sci. Data* **3**: 160087.

Hill, W.G., Goddard, M.E. & Visscher, P.M. 2008. Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet.* **4**: e1000008.

Hoffmann, A.A. & Hercus, M.J. 2000. Environmental stress as an evolutionary force. *Bioscience* **50**: 217–226.

Hoffmann, A.A. & Sgrò, C.M. 2011. Climate change and evolutionary adaptation. *Nature* **470**: 479–485.

Hoggart, C.J., Shriver, M.D., Kittles, R.A., Clayton, D.G. & McKeigue, P.M. 2004. Design and analysis of admixture mapping studies. *Am. J. Hum. Genet.* **74**: 965–978.

Hüning, A.K., Melzner, F., Thomsen, J., Gutowska, M.A., Krämer, L., Frickenhaus, S. *et al.* 2013. Impacts of seawater acidification on mantle gene expression patterns of the Baltic Sea blue mussel: implications for shell formation and energy metabolism. *Mar. Biol.* **160**: 1845–1861.

Johnson, Z.I., Wheeler, B.J., Blinebry, S.K., Carlson, C.M., Ward, C.S. & Hunt, D.E. 2013. Dramatic variability of the carbonate system at a temperate coastal ocean site (Beaufort, North Carolina, USA) is regulated by physical and biogeochemical processes on multiple timescales. *PLoS One* **8**: e85117.

Katolikova, M., Khaitov, V., Väinölä, R., Gantsevich, M. & Strelkov, P. 2016. Genetic, ecological and morphological distinctness of the blue mussels *Mytilus trossulus* Gould and *M. edulis* L. in the White Sea. *PLoS One* **11**: e0152963.

Keenan, K., McGinnity, P., Cross, T.F., Crozier, W.W. & Prodöhl, P.A. 2013. An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods Ecol. Evol.* **4**: 782–788. https://doi.org/10.1111/2041-210X.12067

Kingston, S.E., Parchman, T.L., Gompert, Z., Buerkle, C.A. & Braun, M.J. 2017. Heterogeneity and concordance in locus-specific differentiation and introgression between species of towhees. *J. Evol. Biol.* **30**: 474–485.

Kroeker, K.J., Kordas, R.L., Crim, R., Hendriks, I.E., Ramajo, L., Singh, G.S. *et al.* 2013. Impacts of ocean acidification on marine organisms: quantifying sensitivities and interaction with warming. *Glob. Change Biol.* **19**: 1884.

Kroeker, K.J., Sanford, E., Rose, J.M., Blanchette, C.A., Chan, F., Chavez, F.P. *et al.* 2016. Interacting environmental mosaics drive geographic variation in mussel performance and predation vulnerability. *Ecol. Lett.* **19**: 771–779.

Lande, R. 1980. Sexual dimorphism, sexual selection, and adaptation in polygenic characters. *Evolution* **34**: 292–305.

Latta, L.C., Peacock, M., Civitello, D.J., Dudycha, J.L., Meik, J.M. & Schaack, S. 2015. The phenotypic effects of spontaneous mutations in different environments. *Am. Nat.* **185**: 243–252.

Lesser, M.P. 2016. Climate change stressors cause metabolic depression in the blue mussel, *Mytilus edulis*, from the Gulf of Maine. *Limnol. Oceanogr.* **61**: 1705–1717.

Lesser, M.P. & MacManes, M. 2016. Transcriptomic resources for the rocky intertidal blue mussel *Mytilus edulis* from the Gulf of Maine. *J. Shellfish Res.* **35**: 435–645.

Lobell, D.B., Burke, M.B., Tebaldi, C., Mastrandrea, M.D., Falcon, W.P. & Naylor, R.L. 2008. Prioritizing climate change adaptation needs for food security in 2030. *Science* **319**: 607–610.

Lockwood, B.L., Connor, K.M. & Gracey, A.Y. 2015. The environmentally tuned transcriptomes of *Mytilus* mussels. *J. Exp. Biol.* **218**: 1822–1833.

Lubchenco, J. & Menge, B. 1978. Community development and persistence in a low rocky intertidal zone. *Ecol. Monogr.* **59**: 67–94.

Lynch, M. & Ritland, K. 1999. Estimation of pairwise relatedness with molecular markers. *Genetics* **152**: 1753–1766.

Mazzarella, A.B., Boessenkool, S., Østbye, K., Vøllestad, L.A. & Trucchi, E. 2016. Genomic signatures of the plateless phenotype in the threespine stickleback. *Ecol. Evol.* **6**: 3161–3173.

Merilä, J. & Hendry, A.P. 2014. Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evol. Appl.* **7**: 1–14.

Mora, C., Wei, C.-L., Rollo, A., Amaro, T., Baco, A.R., Billett, D. *et al.* 2013. Biotic and human vulnerability to projected changes in ocean biogeochemistry over the 21st century. *PLoS Biol.* **11**: e1001682.

Moreira, R., Pereiro, P., Canchaya, C., Posada, D., Figueras, A. & Novoa, B. 2015. RNA-Seq in *Mytilus galloprovincialis:* comparative transcriptomics and expression profiles among different tissues. *BMC Genom.* **16**: 728.

Murgarella, M., Puiu, D., Novoa, B., Figueras, A., Posada, D. & Canchaya, C. 2016. A First insight into the genome of the filter-feeder mussel *Mytilus galloprovincialis*. *PLoS One* **11**: e0151561.

Narum, S.R., Buerkle, C.A., Davey, J.W., Miller, M.R. & Hohenlohe, P.A. 2013. Genotyping-by-sequencing in ecological and conservation genomics. *Mol. Ecol.* **22**: 2841–2847.

Parchman, T.L., Gompert, Z., Mudge, J., Schilkey, F.D., Benkman, C.W. & Buerkle, C.A. 2012. Genome-wide association mapping of an adaptive trait in lodgepole pine. *Mol. Ecol.* **21**: 2991–3005.

Parmesan, C. 2006. Ecological and evolutionary responses to recent climate change. *Annu. Rev. Ecol. Evol. Syst.* **37**: 637–669.

Patterson, N., Hattangadi, N., Lane, B., Lohmueller, K.E., Hafler, D.A., Oksenberg, J.R. *et al.* 2004. Methods for high-density admixture mapping of disease genes. *Am. J. Hum. Genet.* **74**: 979–1000.

Pearson, R.G. & Dawson, T.P. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Glob. Ecol. Biogeogr.* **12**: 361–371.

Pershing, A.J., Alexander, M.A., Hernandez, C.M., Kerr, L.A., Bris, A.L., Mills, K.E. *et al.* 2015. Slow adaptation in the face of rapid warming leads to collapse of the Gulf of Maine cod fishery. *Science* **350**: 809–812.

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

Pierrot, D.W., Lewis, E. & Wallace, D. 2006. *MS Excel Program Developed for CO_2 System Calculations: ORNL/CDIAC-105a. C. D. I. A*. Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tennessee.

Plummer, M., Best, N., Cowles, K. & Vines, K. 2006. CODA: convergence diagnosis and output analysis for MCMC. *R. News* **6**: 7–11.

Pritchard, J.K., Stephens, M., Rosenberg, N.A. & Donnelly, P. 2000. Association mapping in structured populations. *Am. J. Hum. Genet.* **67**: 170–181.

Pritchard, V.L., Viitaniemi, H.M., McCairns, R.J.S., Merilä, J., Nikinmaa, M., Primmer, C.R. *et al.* 2017. Regulatory architecture of gene expression variation in the Threespine Stickleback *Gasterosteus aculeatus*. *G3 (Bethesda)* **7**: 165–178.

R Core Team 2013. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Raj, A., Stephens, M. & Pritchard, J.K. 2014. fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics* **197**: 573–589.

Rawson, P.D., Secor, C.L. & Hilbish, T.J. 1996. The effects of natural hybridization on the regulation of doubly uniparental mtDNA inheritance in blue mussels (*Mytilus* spp.). *Genetics* **144**: 241–248.

Ries, J.B., Cohen, A.L. & McCorkle, D.C. 2009. Marine calcifiers exhibit mixed responses to CO_2-induced ocean acidification. *Geology* **37**: 1131–1134.

Riginos, C. & Cunningham, C.W. 2005. Local adaptation and species segregation in two mussel (*Mytilus edulis* × *Mytilus trossulus*) hybrid zones. *Mol. Ecol.* **14**: 381–400.

Rockman, M.V. 2012. The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution* **66**: 1–17.

Savolainen, O., Lascoux, M. & Merilä, J. 2013. Ecological genomics of local adaptation. *Nat. Rev. Genet.* **14**: 807–820.

Shapiro, M.D., Kronenberg, Z., Li, C., Domyan, E.T., Pan, H., Campbell, M. *et al.* 2013. Genomic diversity and evolution of the head crest in the rock pigeon. *Science* **339**: 1063–1067.

Sorte, C.J.B., Davidson, V.E., Franklin, M.C., Benes, K.M., Doellman, M.M., Etter, R.J. *et al.* 2016. Long-term declines in an intertidal foundation species parallel shifts in community composition. *Glob. Change Biol.* **2016**: 341–352.

Spencer, C.C., Su, Z., Donnelly, P. & Marchini, J. 2009. Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip. *PLoS Genet.* **5**: e1000477.

Steneck, R.S., Hughes, T.P., Cinner, J.E., Adger, W.N., Arnold, S.N., Berkes, F. *et al.* 2011. Creation of a gilded trap by the high economic value of the Maine lobster fishery. *Conserv. Biol.* **25**: 904–912.

Stephens, M. & Balding, D.J. 2009. Bayesian statistical methods for genetic association studies. *Nat. Rev.* **10**: 681–690.

Sun, J., Zhang, Y., Xu, T., Zhang, Y., Mu, H., Zhang, Y. *et al.* 2017. Adaptation to deep-sea chemosynthetic environments as revealed by mussel genomes. *Nat. Ecol. Evol.* **1**: 0121.

Thomsen, J., Casties, I., Pansch, C., Körtzinger, A. & Melzner, F. 2013. Food availability outweighs ocean acidification effects in juvenile *Mytilus edulis*: laboratory and field experiments. *Glob. Change Biol.* **19**: 1017–1027.

Timmins-Schiffman, E.B., Friedman, C.S., Metzger, D.C., White, S.J. & Roberts, S.B. 2013. Genomic resource development for shellfish of conservation concern. *Mol. Ecol. Resour.* **13**: 295–305.

Waddington, C.H. 1956. Genetic assimilation of the bithorax phenotype. *Evolution* **10**: 1–13.

Wang, S., Zhang, J., Jiao, W., Li, J., Xun, X., Sun, Y. *et al.* 2017. Scallop genome provides insights into evolution of bilaterian karyotype and development. *Nat. Ecol. Evol.* **1**: 0120.

Wethey, D.S., Woodin, S.A., Hilbish, T.J., Jones, S.J., Lima, F.P. & Brannock, P.M. 2011. Response of intertidal populations to climate: effects of extreme events versus long term change. *J. Exp. Mar. Biol. Ecol.* **400**: 132–144.

Wiens, J.A., Stralberg, D., Jongsomjit, D., Howell, C.A. & Snyder, M.A. 2009. Niches, models, and climate change: assessing the assumptions and uncertainties. *Proc. Natl Acad. Sci.* **106**: 19729–19736.

Witman, J. 1987. Subtidal coexistence: storms, grazing, mutualism and the zonation of kelps and mussels. *Ecol. Monogr.* **57**: 167–187.

Yarra, T., Gharbi, K., Blaxter, M., Peck, L.S. & Clark, M.S. 2016. Characterization of the mantle transcriptome in bivalves: *Pecten maximus*, *Mytilus edulis* and *Crassostrea gigas*. *Mar. Genomics* **27**: 9–15.

Yund, P.O., Tilburg, C.E. & McCartney, M.A. 2015. Across-shelf distribution of blue mussel larvae in the northern Gulf of Maine: consequences for population connectivity and a species range boundary. *R. Soc. Open Sci.* **2**: 150513.

Zhang, G., Fang, X., Guo, X., Li, L., Luo, R., Xu, F. *et al.* 2012. The oyster genome reveals stress adaptation and complexity of shell formation. *Nature* **490**: 49–54.

Zheng, X., Levine, D., Shen, J., Gogarten, S., Laurie, C. & Weir, B. 2012. A High-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28**: 3326–3328.

Zhou, X. & Stephens, M. 2012. Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **44**: 821–824.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Table S1 Sea table assignments by sampling location and run date.

Table S2 Summary of temperature, pH, salinity, and saturation states in ambient and climate stress common garden sea tables.

Table S3 Univariate linear model parameter summary of effect sizes and *P*-values (testing the null hypothesis that effect size = 0; Wald test, likelihood ratio test, and Score test).

Data deposited at Dryad: doi: https://doi.org/10.5061/dryad.2d8b5

Received 27 September 2017; revised 11 December 2017; accepted 13 December 2017