Geographic variation in *Megalorchestia californiana* allele frequencies may be caused by winter rather than summer temperatures

John H. McDonald*

Department of Biological Sciences, University of Delaware, Newark, Delaware 19716, USA

ABSTRACT: There are many allozyme polymorphisms with latitudinal clines, for which selection by climate is one obvious explanation. Both warm summer temperatures and cold winter temperatures could plausibly impose selection on enzyme polymorphisms, but because summer and winter temperatures are usually geographically correlated, there has been little attempt to use patterns of geographic variation to discriminate between these possible selective factors. The glucose-6-phosphate isomerase and mannose-6-phosphate isomerase allozyme polymorphisms in the amphipod crustacean *Megalorchestia californiana* have clines that are strongly associated with mean January air temperature but are only weakly associated with mean August air temperature. This suggests that winter temperatures are more likely than summer temperatures to be a selective factor on these polymorphisms.

KEY WORDS: Allozyme · Climate · Cline · Peracarid · Selection

INTRODUCTION

A variety of techniques can be used to test whether a protein polymorphism is affected by natural selection, using evidence such as geographic variation in allele frequency (McDonald 1994, Beaumont 2005, Nosil et al. 2009), selection component analysis (Christiansen & Frydenberg 1973, Siegismund & Christiansen 1985), or DNA polymorphism and divergence data (Nielsen 2005). There are now many examples of protein polymorphisms with evidence of balancing or differentiating selection, but detecting selection is just the first step. A thorough understanding of a protein polymorphism would include a detailed biochemical characterization of the products of different alleles; careful observation of the physiological, developmental, or behavioral phenotypes (Eanes 1999, Zera 2011); and accurate measurement of the relative fitnesses of genotypes in nature, including the effects of different environmental variables (physical, chemical, and biological) on those fitnesses (Joost et al. 2007, Manel et al. 2010, Cox et al. 2011).

Geographic variation in allele frequency is one common source of inferences about environmental factors that may affect fitness; when allele frequencies are correlated with a geographically varying environmental factor, it is evidence that suggests differential selection by that factor. A single cline, such as a latitudinal cline associated with temperature, could also be due to random drift; only a repeated geographic pattern (Oakeshott et al. 1982, McDonald 1987) is strong evidence for selection. However, even a single cline, when it is strongly correlated with an environmental factor, suggests that the factor is worthy of further investigation (Manel et al. 2010). Correlation of allele frequencies with an environmental factor can also provide valuable clues that help focus the search for differences in biochemical and physiological phenotype (Clarke 1975, Coop et al. 2010).

One common type of association of allozyme allele frequency with environment is a latitudinal cline

(Endler 1977, Riddoch 1993, Schmidt et al. 2008). The obvious explanation is selection by temperature, since temperature has pervasive effects on protein stability and function, and higher latitude environments generally have cooler temperatures. Both high temperatures and low temperatures can have profound effects on enzyme function and stability (Somero 1995, Angilletta 2009), so for any particular polymorphism, it is possible that fitness differences among genotypes primarily occur during hot summer temperatures, primarily during cold winter temperatures, or throughout the year. To disentangle summer and winter temperatures as possible selective factors, it would be helpful to sample populations from a set of locations for which summer and winter temperatures are not highly correlated. The Pacific coast of the United States is one area where this is possible. The climate at coastal locations is moderated by the open ocean and cooled in the summer by coastal upwelling, while locations on protected waters such as the Salish Sea and San Francisco Bay experience a broader range of temperatures. As a result, protected beaches are warmer in the summer and cooler in the winter than beaches on the adjacent open coasts (Table S1 in the Supplement at www.int-res.com/articles/suppl/m488p201_supp/), leading to a weak association of summer and winter temperatures (Fig. 1).
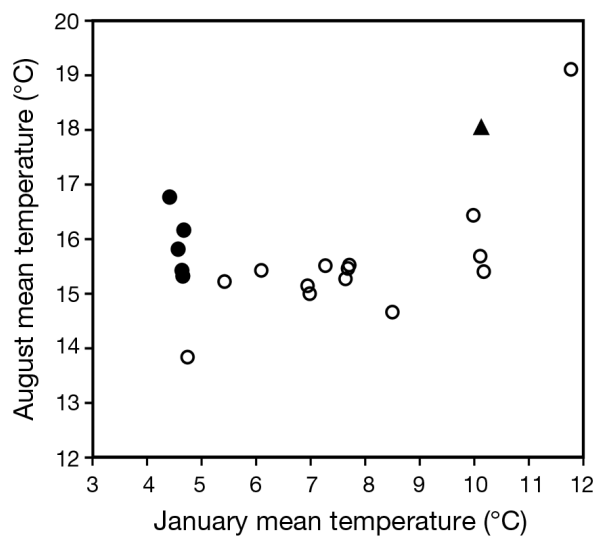

Fig. 1. Relationship of January and August mean air temperatures at the sample locations used in this study (see Fig. 2). ○, sites on or near the open coast; ●, sites in the Salish Sea; ▲, site in San Francisco Bay. Data from the PRISM database (PRISM Climate Group 2012) as shown in Table S1 in the Supplement at www.int-res.com/articles/suppl/m488p201_supp/

The amphipod crustacean *Megalorchestia californiana* inhabits sandy beaches on the Pacific coast of North America. Individuals remain burrowed in the sand around the high-tide line during the day, emerging at night to forage for washed-up detritus (Bowers 1964). There is no planktonic larval stage; direct-developing young are released from a brood pouch. Glucose-6-phosphate isomerase (GPI) and mannose-6-phosphate isomerase (MPI) both have allozyme polymorphisms with 2 common alleles, and both polymorphisms have allele frequencies that are correlated with latitude (McDonald 1985). Here data from McDonald (1985) is combined with data from new locations, including several more locations in protected habitats, to test whether *Gpi* and *Mpi* allele frequencies are more strongly associated with summer or winter temperatures.

MATERIALS AND METHODS

Amphipods were collected from the strand line of sandy beaches by 1 of 3 methods. Pitfall traps (plastic drinking cups containing 2 cm of seawater, with the top flush with the sand surface) were used at night; during the day, amphipods were collected by picking them from underneath piles of stranded seaweed or by sieving sand through a 4 mm mesh. With the exception of Santa Cruz (May 31) and some San Francisco individuals (May 2), samples were collected in late March or early April 2009 and 2010. An average of 228 individuals were sampled from each location. Sampling dates and sample sizes are given in Table S2 in the Supplement. Data from 1983 samples were obtained from McDonald (1985) (which erroneously states that samples were collected in 1984).

Amphipods were frozen on dry ice in the field and returned to the laboratory. *Gpi* and *Mpi* genotypes were determined on starch gels using the protocols of McDonald (1987). Individuals with ambiguous results were re-run until a clear genotype was obtained for both enzymes.

Rare alleles (less than 1% in the total of all samples) were pooled with the electrophoretically closer of the 2 common alleles at each locus for all statistical analyses. Allele frequencies from locations sampled in more than 1 yr were compared using G-tests of independence (McDonald 2009, pp. 64–69) on each location separately, and the Cochran-Mantel-Haenszel test (McDonald 2009, pp. 88–94) was used to compare sets of multiple locations sampled in different years.

As a measure of long-term climate, the mean maximum and minimum daily temperatures for January and August for the 1981 to 2010 time period were obtained from the PRISM database (PRISM Climate Group 2012). Mean monthly temperatures were estimated by averaging the mean maximum and minimum daily temperatures. January and August were chosen because at most of the locations, January has the lowest and August has the highest temperatures.

Logistic regression (McDonald 2009, pp. 247–255) of allele frequency on temperature was done separately for each of 6 variables: January and August mean daily maxima, January and August mean daily minima, and January and August means. The strength of each association was measured with McKelvey & Zavoina's (1975) pseudo-R^2; of several possible pseudo-R^2 values, it is most comparable to the familiar R^2 of linear regression, although values of the pseudo-R^2 for strong associations are smaller than R^2 values for strong linear regressions (Veall & Zimmermann 1996). For each locus, the difference between the best-fitting climate variable and each other variable was tested using the method of Vuong (1989) with a Bonferroni correction for the 5 comparisons.

Genotype data have been submitted to Dryad (doi:10.5061/dryad.k172s).

RESULTS

Amphipods were collected from 7 locations in spring 2009 and 15 locations in spring 2010, with 4 locations being sampled in both years (Fig. 2, Table S1). There were 2 common alleles at each locus, with 2 rare alleles (less than 1% in the total sample) at *Gpi* and 3 rare alleles at *Mpi* (Table S2). None of the 4 locations that were sampled in both years differed significantly in allele frequency for either locus between years (G-test, p > 0.05), and there was no significant overall trend across the 4 locations at either locus (Cochran-Mantel-Haenszel test, p > 0.05), so the 2009 and 2010 samples were pooled for subsequent analyses. Seven locations were sampled in both 1983 (McDonald 1985) and 2009 to 2010, and they did not differ significantly in allele frequency, either individually or overall, so the 1983 samples were combined with the 2009 to 2010 samples for subsequent analyses. Four samples from McDonald (1985) were not used: Humboldt Bay, Bodega Bay, and Orange County, due to very small sample sizes, and Santa Barbara, which was similar but not identical in location to the larger 2010 Goleta sample.
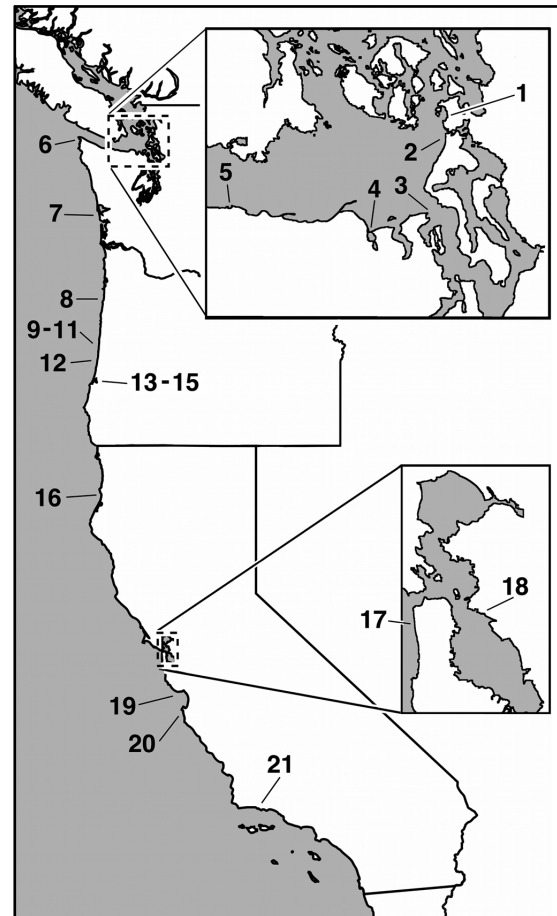

Fig. 2. Sample locations on the Pacific coast of the USA. Exact locations are given in Table S1 in the Supplement available at www.int-res.com/articles/suppl/m488p201_supp/

There was a strong association between *Gpi* allele frequency of the combined samples and January maximum, minimum, or mean temperature, with January maximum temperature having the strongest association as measured by pseudo-R^2 (Table 1). The association of *Gpi* allele frequency and January maximum temperature (Fig. 3) was significant (logistic regression, p < 0.0001). The best association of *Gpi* allele frequency and August temperature was with the August maximum; although the logistic regression was significant (p < 0.0001), it was significantly weaker (Vuong test, p < 0.0001) with a much smaller pseudo-R^2 than any of the January variables. The difference in the associations was largely due to sites in the Salish Sea. Beaches in the Salish Sea are almost 1°C warmer in August but 2 to 3°C cooler in January than most locations on the open coasts of Washington and Oregon (Fig. 1, Table S1), and the *Gpi*[100] allele, which is most common in southern California, is less common inside the Salish Sea than on the open coast.

Table 1. Results of logistic regression of allele frequencies vs. each climate variable. Pseudo-R^2 calculated according to McKelvey & Zavoina (1975)

Locus	Climate variable	Pseudo R^2
GPI	January maximum	0.249
	January minimum	0.210
	January mean	0.243
	August maximum	0.063
	August minimum	0.013
	August mean	0.018
MPI	January maximum	0.255
	January minimum	0.231
	January mean	0.256
	August maximum	0.000
	August minimum	0.100
	August mean	0.022

A similar pattern is seen at San Francisco Bay, where the location inside the bay is about the same January temperature as the coastal location of the same latitude, but the bay site is over 2°C warmer in August; despite the warmer summer temperatures, the Gpi^{100} allele is about the same frequency inside and outside of San Francisco Bay.

A similar pattern was seen for *Mpi* allele frequencies (Fig. 3): there was a strong association with January maximum, minimum, or mean temperatures, with mean temperature showing the strongest association; the association with August minimum temperatures, while significant, was considerably weaker than the associations with January temperatures (Table 1). The Mpi^{90} allele, which was most common in southern California, was less common inside the Salish Sea than on the open coast, and Mpi^{90} was similar in frequency inside and outside of San Francisco Bay.

DISCUSSION

The results show that both *Gpi* and *Mpi* have latitudinal clines that exhibit a stronger association with winter temperatures than with summer temperatures. Of course, a single cline such as seen here could result from random drift in isolated populations, followed by migration. Much of the pattern is due to the 5 sites in the Salish Sea, and it is possible that they are all derived from a single isolated population that differentiated due to drift and happens to
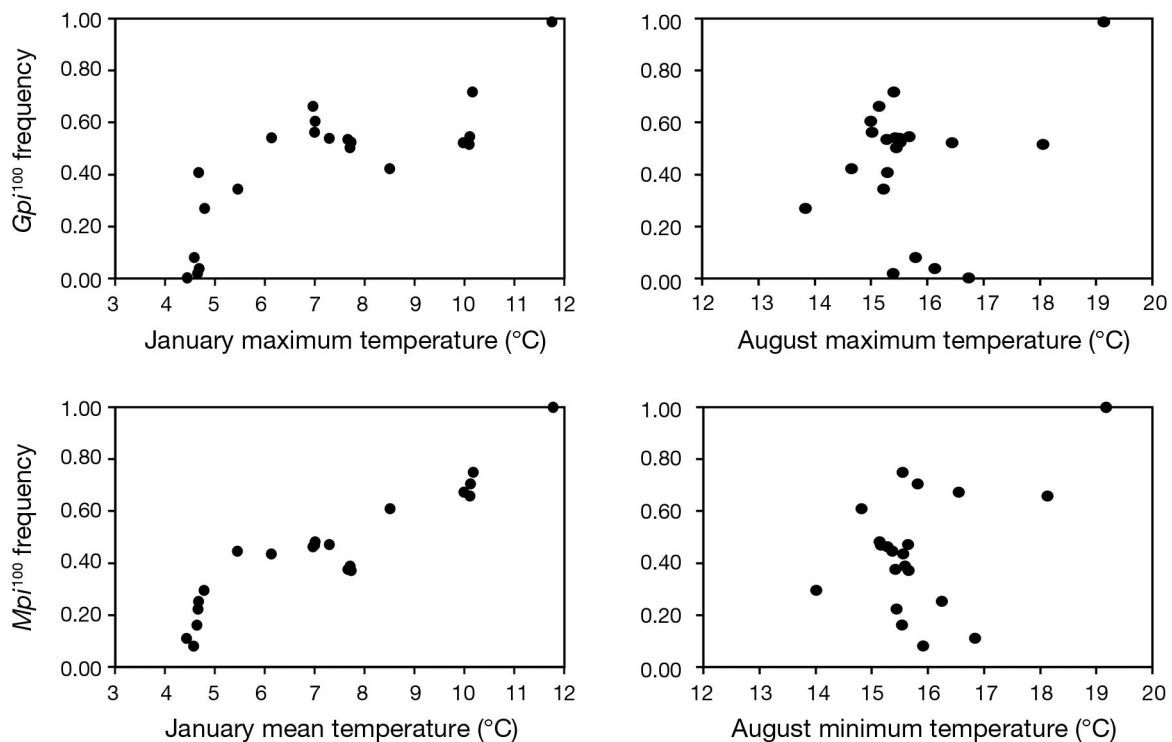

Fig. 3. Association of allele frequency and temperature. The frequency of the allele that becomes more common in the south is shown for *Gpi* and *Mpi*, plotted against January or August temperatures from the PRISM database (PRISM Climate Group 2012). The best-fitting temperature from each month is shown. Complete data are given in Tables S1 & S2 in the Supplement at www.int-res.com/articles/suppl/m488p201_supp/

be found in an area with relatively warm summers and cold winters. However, these results do suggest that if the clines are due to selection, winter temperatures are a worthy candidate for further investigation as a selective factor.

There was no significant change in allele frequency between the 1983 samples and the samples collected a quarter-century later, so unlike some polymorphisms (Umina et al. 2005), there is no evidence that changing climate has had a selective effect. Most of the new samples in this study were collected in late March and early April, and most of the samples in McDonald (1985) were collected in late February through mid-August. One possible follow-up experiment would be to collect large samples from a single location at multiple times throughout the year, in particular before and after the winter, to see if seasonal selection is strong enough to produce measurable changes in genotype frequencies.

Numerous species have *Gpi* polymorphisms in which allozyme allele frequencies are associated with latitude or temperature (Riddoch 1993). For the most part, these geographic patterns of allele frequency change occur in areas such as the eastern coast of the USA, where the high correlation between winter and summer temperatures would make it difficult to use the geographic patterns to separate the roles of summer and winter temperatures as possible selective agents. Here I take advantage of the disconnect between summer and winter temperatures provided by the Salish Sea and San Francisco Bay, whose protected shores have air temperatures that are warmer in summer but colder in winter than the open coast of Washington and Oregon. The much stronger association of *Gpi* allele frequencies with winter temperatures suggests that if temperature-based selection is the cause of the geographic pattern, winter temperatures are more important than summer temperatures.

This analysis used air temperatures from the PRISM database (PRISM Climate Group 2012), which estimates the daily maximum and minimum temperature at locations by interpolating between nearby weather stations and adjusting for the effects of altitude and proximity to bodies of water. During the daytime, the temperature in the sand may be quite different from the outside air temperature (Pollock & Hummon 1971, Milton et al. 1997). Unlike sessile intertidal organisms, whose body temperature can be accurately modeled using air and water temperature data (Wethey et al. 2011), *Megalorchestia californiana* would have some control of their body temperature during the daytime based on the elevation of their burrow on the beach

and the depth in the sand. At night, *M. californiana* could avoid extreme cold air temperatures by remaining in their burrows. Temperature-related selection might result from the occasional extremely hot or cold day; from daytime highs or nighttime lows; or from the daily, monthly, or yearly variability in temperature rather than the monthly mean. Temperatures at other times of the year, such as the main breeding season, might be more important than the coldest and hottest months, and rainfall or cloud cover could conceivably have a selective role. With only 21 geographic locations, a multiple logistic model with all of the many possible climate variables would have risked overfitting, so only the January and August mean, maximum, and minimum temperatures were compared. Considering all the possible selective factors, the close fit of both *Gpi* and *Mpi* allele frequencies to January temperatures is therefore remarkable.

GPI catalyzes the interconversion of glucose-6-phosphate and fructose-6-phosphate that is part of the important glycolysis/gluconeogenesis reversible pathway. GPI competes with several other enzymes for glucose-6-phosphate, so that GPI activity may play a role in allocating resources to lipid storage vs. immediate energy production (Eanes 2011). There have been numerous biochemical studies of *Gpi* polymorphisms with allele frequencies associated with temperature (Watt 1977, Hoffmann 1981, Watt 1983, Hall 1985, Zera 1987, Van Beneden & Powers 1989, Patarnello & Battaglia 1992, Dahlhoff & Rank 2000), but there is no single biochemical property that is consistently associated with alleles that are more common in warmer areas; for example, while the warm-associated allozyme exhibits greater resistance to thermal denaturation in some species (Watt 1977, Van Beneden & Powers 1989), it is less stable in others (Hoffmann 1981, Hall 1985, Zera 1987). Making full use of geographic patterns to infer as much as possible about possible selective factors would be a useful precursor to detailed biochemical and physiological studies of enzyme polymorphisms.

Compared with *Gpi*, relatively few latitudinal clines of *Mpi* allele frequency are known. This may be because fewer people stained for MPI during the bygone era of allozyme surveys, due to the generally weak activity of MPI (Saunders et al. 1969) and higher cost of the MPI stain ingredients. There are, however, several organisms with variation in *Mpi* allele frequency on a smaller geographic scale. The amphipod crustaceans *Platorchestia platensis*, *Orchestia grillus*, and *Hyale plumulosa* exhibit repeated differences in allele frequency between beaches out-

side and inside small bays on Long Island, New York (McDonald 1987, McDonald 1991); temperature, oxygen, and food type are among the plausible selective factors that may differ between the exposed and protected habitats. The barnacle *Semibalanus balanoides* exhibits differences in *Mpi* allele frequency between the upper and lower intertidal in Maine (Schmidt & Rand 1999), but the allele that is favored in the upper intertidal in Maine is favored in the lower intertidal in Rhode Island (Rand et al. 2002).

MPI is a reversible enzyme that catalyzes the interconversion of fructose-6-phosphate and mannose-6-phosphate. Many algae contain high levels of mannose polysaccharides, so *Megalorchestia californiana* presumably uses MPI to process mannose into energy. Having sufficient MPI activity is important when food contains high levels of mannose, as phosphorylating mannose to mannose-6-phosphate uses ATP, and if ATP is used faster than it is regenerated, it can be fatal (de la Fuente et al. 1986, Harvey et al. 2012). Under stressful conditions, mannose causes differences in growth and survival between *Mpi* genotypes in *Semibalanus balanoides* (Schmidt 2001). Selection on *Mpi* may be particularly likely in marine organisms that eat algae that are high in mannose.

LITERATURE CITED

Angilletta MJ Jr (2009) Thermal adaptation: a theoretical and empirical synthesis. Oxford University Press, Oxford

Beaumont MA (2005) Adaptation and speciation: What can F_{st} tell us? Trends Ecol Evol 20:435–440

Bowers DE (1964) Natural history of two beach hoppers of genus *Orchestoidea* (Crustacea: Amphipoda) with reference to their complemental distribution. Ecology 45: 677–696

Christiansen FB, Frydenberg O (1973) Selection component analysis of natural polymorphisms using population samples including mother-offspring combinations. Theor Popul Biol 4:425–445

Clarke B (1975) The contribution of ecological genetics to evolutionary theory: detecting direct effects of natural selection on particular polymorphic loci. Genetics 79: 101–113

Coop G, Witonsky D, Di Rienzo A, Pritchard JK (2010) Using environmental correlations to identify loci underlying local adaptation. Genetics 185:1411–1423

Cox K, Vanden Broeck A, Van Calster H, Mergeay J (2011) Temperature-related natural selection in a wind-pollinated tree across regional and continental scales. Mol Ecol 20: 2724–2738

Dahlhoff EP, Rank NE (2000) Functional and physiological consequences of genetic variation at phosphoglucose isomerase: heat shock protein expression is related to enzyme genotype in a montane beetle. Proc Natl Acad Sci USA 97:10056–10061

de la Fuente M, Peñas PF, Sols A (1986) Mechanism of mannose toxicity. Biochem Biophys Res Commun 140: 51–55

Eanes WF (1999) Analysis of selection on enzyme polymorphisms. Annu Rev Ecol Syst 30:301–326

Eanes WF (2011) Molecular population genetics and selection in the glycolytic pathway. J Exp Biol 214:165–171

Endler JA (1977) Geographic variation, speciation, and clines. Princeton University Press, Princeton

Hall JG (1985) Temperature-related kinetic differentiation of glucosephosphate isomerase alleloenzymes isolated from the blue mussel, *Mytilus edulis.* Biochem Genet 23: 705–728

Harvey JA, Cloutier J, Visser B, Ellers J, Wackers FL, Gols R (2012) The effect of different dietary sugars and honey on longevity and fecundity in two hyperparasitoid wasps. J Insect Physiol 58:816–823

Hoffmann RJ (1981) Evolutionary genetics of *Metridium senile.* 1. Kinetic differences in phosphoglucose isomerase allozymes. Biochem Genet 19:129–144

Joost S, Bonin A, Bruford MW, Despres L and others (2007) A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. Mol Ecol 16:3955–3969

Manel S, Joost S, Epperson BK, Holderegger R and others (2010) Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. Mol Ecol 19:3760–3772

McDonald JH (1985) Size-related and geographic variation at two enzyme loci in *Megalorchestia californiana* (Amphipoda: Talitridae). Heredity 54:359–366

McDonald JH (1987) Repeated geographic variation at three enzyme loci in the amphipod *Platorchestia platensis.* Evolution 41:438–441

McDonald JH (1991) Contrasting amounts of geographic variation as evidence for direct selection: the *Mpi* and *Pgm* loci in eight crustacean species. Heredity 67: 215–219

McDonald JH (1994) Detecting natural selection by comparing geographic variation in protein and DNA polymorphisms. In: Golding B (ed) Non-neutral evolution: theories and molecular data. Chapman and Hall, New York, NY

McDonald JH (2009) Handbook of biological statistics, 2nd edn. Sparky House Publishing, Baltimore

McKelvey R, Zavoina W (1975) A statistical model for the analysis of ordinal level dependent variables. J Math Sociol 4:103–120

Milton SL, Schulman AA, Lutz PL (1997) The effect of beach nourishment with aragonite versus silicate sand on beach temperature and loggerhead sea turtle nesting success. J Coast Res 13:904–915

Nielsen R (2005) Molecular signatures of natural selection. Annu Rev Genet 39:197–218

Nosil P, Funk DJ, Ortiz-Barrientos D (2009) Divergent selection and heterogeneous genomic divergence. Mol Ecol 18:375–402

Oakeshott JG, Gibson JB, Anderson PR, Knibb WR, Anderson DJ, Chambers GK (1982) Alcohol dehydrogenase and glycerol-3-phosphate dehydrogenase clines in *Drosophila melanogaster* on different continents. Evolution 36:86–96

Patarnello T, Battaglia B (1992) Glucosephosphate isomerase and fitness: effects of temperature on genotype dependent mortality and enzyme activity in two species of the genus *Gammarus* (Crustacea: Amphipoda). Evolution 46:1568–1573

Pollock LW, Hummon WD (1971) Cyclic changes in interstitial water content, atmospheric exposure, and temperature in a marine beach. Limnol Oceanogr 16:522–535

PRISM Climate Group (2012) PRISM Data Explorer 1971–2000 normals. www.prism.oregonstate.edu (accessed 5 July 2012)

Rand DM, Spaeth PS, Sackton TB, Schmidt PS (2002) Ecological genetics of *Mpi* and *Gpi* polymorphisms in the acorn barnacle and the spatial scale of neutral and non-neutral variation. Integr Comp Biol 42:825–836

Riddoch BJ (1993) The adaptive significance of electrophoretic mobility in phosphoglucose isomerase (PGI). Biol J Linn Soc 50:1–17

Saunders SA, Gracy RW, Schnackerz KD, Noltmann EA (1969) Are honeybees deficient in phosphomannose isomerase? Science 164:858–859

Schmidt PS (2001) The effects of diet and physiological stress on the evolutionary dynamics of an enzyme polymorphism. Proc Biol Sci 268:9–14

Schmidt PS, Rand DM (1999) Intertidal microhabitat and selection at *Mpi*: interlocus contrasts in the northern acorn barnacle, *Semibalanus balanoides.* Evolution 53:135–146

Schmidt PS, Serrão EA, Pearson GA, Riginos C and others (2008) Ecological genetics in the North Atlantic: environmental gradients and adaptation at specific loci. Ecology 89:S91–S107

Siegismund HR, Christiansen FB (1985) Selection component analysis of natural polymorphisms using population samples including mother-offspring combinations III. Theor Popul Biol 27:268–297

Somero GN (1995) Proteins and temperature. Annu Rev Physiol 57:43–68

Umina PA, Weeks AR, Kearney MR, McKechnie SW, Hoffmann AA (2005) A rapid shift in a classic clinal pattern in *Drosophila* reflecting climate change. Science 308:691–693

Van Beneden RJ, Powers DA (1989) Structural and functional differentiation of two clinally distributed glucosephosphate isomerase allelic isozymes from the teleost *Fundulus heteroclitus.* Mol Biol Evol 6:155–170

Veall MR, Zimmermann KF (1996) Pseudo-R^2 measures for some common limited dependent variable models. J Econ Surv 10:241–259

Vuong QH (1989) Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica 57:307–333

Watt WB (1977) Adaptation at specific loci. 1. Natural selection on phosphoglucose isomerase of *Colias* butterflies: biochemical and population aspects. Genetics 87:177–194

Watt WB (1983) Adaptation at specific loci. 2. Demographic and biochemical elements in the maintenance of the *Colias* PGI polymorphism. Genetics 103:691–724

Wethey DS, Brin LD, Helmuth B, Mislan KAS (2011) Predicting intertidal organism temperatures with modified land surface models. Ecol Modell 222:3568–3576

Zera AJ (1987) Temperature-dependent kinetic variation among phosphoglucose isomerase allozymes from the wing-polymorphic water strider, *Limnoporus canaliculatus.* Mol Biol Evol 4:266–285

Zera AJ (2011) Microevolution of intermediary metabolism: evolutionary genetics meets metabolic biochemistry. J Exp Biol 214:179–190