2017

Some like it cold: the relationship between thermal tolerance and mitochondrial genotype in an invasive population of the European green crab, Carcinus maenas

Aidan Fisher Coyle
Bowdoin College, acoyle907@gmail.com

Some like it cold: the relationship between thermal tolerance and mitochondrial genotype in an invasive population of the European green crab, *Carcinus maenas*

An Honors Paper for the Department of Biology

By Aidan Fisher Coyle

Bowdoin College, 2017

ABSTRACT

Hybrid zones provide natural laboratories to study how specific genes, and interactions among genes, may influence fitness. On the east coast of North America, two separate populations of the European green crab (*Carcinus maenas*) have been introduced in the last two centuries. An early invasion from Southern Europe colonized New England around 1800, and was followed by a second invasion from Northern Europe to Nova Scotia in the early 1980s (Roman 2006). As these populations hybridize, new combinations of genes potentially adapted to different ends of a thermal spectrum are created in a hybrid zone. To test the hypothesis that mitochondrial and nuclear genes have effects on thermal tolerance, I measured response to cold stress in crabs collected from locations between southern Maine and northern Nova Scotia, and then genotyped the mitochondrial *CO1* gene and two nuclear SNPs. Three mitochondrial haplotypes, originally from Northern Europe, had a strong effect on the ability of crabs to right themselves at a temperature of 4.5ºC. Crabs carrying these three haplotypes were 20% more likely to right compared to crabs carrying the haplotype from Southern Europe. The two nuclear SNPs, which were derived from transcriptome sequencing and were strong outliers between Northern and Southern European *C. maenas* populations, had no effect on righting response at low temperature. These results add *C. maenas* to the short list of ectotherms in which mitochondrial variation affects thermal tolerance, and suggests that natural selection is shaping the structure of the hybrid zone between the northern and southern populations This discovery of linkage between mitochondrial genotype and thermal tolerance also provides potential insight into the patterns of expansion for invasive populations of *C. maenas* around the world.

INTRODUCTION

Hybrid zones, regions where distinctive populations meet and mix, have been characterized as “windows into many evolutionary processes” (Harrison 1993). They act as natural laboratories, providing real-world tests of ideas about species boundaries, reproductive isolation, and the impact of recombination on organism fitness. By their nature, observation of these hybrid zones can illuminate the dynamics of “genomic collisions” between distinct species or populations that have adapted to differing environments.

Such a case appears to occur along the eastern coast of North America, from Nova Scotia to Maine. Here, two separate populations of the European green crab, *Carcinus maenas*, meet and mate. *C. maenas* is one of the world’s most prolific and harmful marine invasive species. While its native range spans from Iceland to Morocco, it has colonized every continent but Antarctica, including both the east and west coast of North America (Hidalgo *et al.* 2005). *C. maenas* first displayed its remarkable ability to invade

in 1817, when it was sighted in Massachusetts (Klassen and Locke 2007). Its range slowly expanded north
until 1960, when it reached a stable northern limit near Halifax, Nova Scotia. However, in the early 1980s, *C. maenas* began to be sighted much further north, and in much colder waters (Klassen and Locke 2007). Since 1980, this northern population of *C. maenas* has expanded northward by roughly 450 miles, and now covers the entirety of Nova Scotia and Prince Edward Island, as well as a significant portion of Newfoundland (Klassen and Locke 2007). Analysis of nuclear and mitochondrial DNA (mtDNA) has revealed that this rapid northern expansion was caused by a second invasion. The first invasion appears to have been founded by a very limited number of individuals, and originated from a population in Southern Europe. The second invasion - which took place in the early 1980s - consisted of individuals from the Norwegian population of *C. maenas* (Roman 2006).

The genetic consequences of hybridization between these two very different lineages in this region have been followed for roughly 10 years. It is known that different kinds of genetic markers show different gradients in allele frequencies, or clines, in this region (Darling *et al.* 2014). The prevailing view on these particular genetic markers, including microsatellites and a mitochondrial marker, is that they are largely neutral with respect to the action of natural selection. In this case, theoretical modeling of clines shows that the width of a genetic cline will largely be determined by stochastic demographic processes, such as larval dispersal (Barton and Hewitt 1985). However, an important alternative hypothesis to the neutral view is that nuclear and mitochondrial genomic regions are under natural selection, and are directly related to the ability of *C. maenas* to survive in a remarkable variety of thermal conditions. Thus, it seems likely that for some portions of the genome, strictly neutral thinking will not apply to understanding both the geographic distribution of genetic variation along the hybrid zone and, potentially, how sets of genes may interact to determine organismal fitness.

In *C. maenas*, there is strong indirect evidence that spatially varying natural selection has shaped the nuclear genome. From a physiological point of view, Tepolt and Somero (2014) have found that populations of *C. maenas* from Northern Europe have greater cold tolerance and less heat tolerance than populations from Southern Europe. A similar pattern, albeit weaker, was found in populations sampled from the northern and southern end points of its range on the east coast of North America. An analysis of transcriptomic variation in Europe also supports genomic differentiation in thermal tolerance (Tepolt and Palumbi 2015). In this study, 1563 outlier genes were identified that had greater differentiation than a null expectation based on neutrality. These collective results make the green crab hybrid zone an excellent experimental system to determine the impact of genes and gene interactions on fitness via examinations of physiological performance.

Compared to other well-studied terrestrial models of evolutionary genetics, marine systems can be quite difficult study. Marine species may be difficult to culture, often have complex life-histories, and have fewer genetic tools. Yet, there are several prominent examples of this genotypic-phenotypic linkage within marine systems. Two of the most prominent are the mannose phosphate isomerase (*Mpi*) and glucose phosphate isomerase (*Gpi*) polymorphisms within the northern acorn barnacle (*Semibalanus balanoides*). In the northern extent of its range, the *Mpi* polymorphism plays an essential role in determining survival. In areas with low temperature change and low desiccation (areas low in the tidal zone and those covered by algae), the *S* homozygote has a significantly higher fitness (Schmidt *et al.* 2000). However, in the high intertidal - where desiccation and temperature change are constant threats - the *F* homozygote has a much higher fitness (Schmidt and Rand, 2001). These alleles have been linked to the ability of *S. balanoides* to either process mannose - a carbohydrate found in the cell walls of algae - or resist thermal stress. The *Gpi* polymorphism has a similar impact on thermal stress and carbohydrate processing in the southern extent of the range of *S. balanoides.* This change from *Mpi* to *Gpi* as the polymorphism under selection is due to a shift in diet makeup between the northern and southern extents of the range of *S. balanoides* (Rand *et al.* 2002). Essentially, the intertidal zone within the northern range of *S. balanoides* is marked by a sharp cline in *Mpi* frequency. Further south, this *Mpi* cline disappears due to a lack of selective pressure, and is replaced by an intertidal cline in *Gpi*.

The barnacle examples described above clearly link nuclear variation at a single locus to phenotypic performance along environmental gradients (here, within the intertidal zone), and can explain clinal variation in these particular genes. But both of these genes are nuclear - far less is known about the linkage between mitochondrial genotype and phenotype. Mitochondrial DNA is a relatively small portion of the genome. In mammals, it consists of a mere 15 - 17 kb, and codes for 37 genes (Ballard and Whitlock 2004). These genes deal with the core mechanisms of cellular survival. Of those 37, just under two-thirds are involved in translating the mtDNA, while the remaining one-third code for proteins that play a role in the electron transport chain (Ballard and Whitlock 2004). This means that mtDNA is both small and relatively conserved (Burton and Barreto 2012). Furthermore, while paternal leakage has been documented to occur at low frequencies, mtDNA is typically inherited completely from the maternal line. (Ballard and Whitlock 2004). Recombination of mtDNA is also rare in animals, although it has been documented in a few cases (Ladoukakis and Zouros 2001). Thus, the mitochondrial genome is relatively small compared to the nuclear genome, but clearly plays important functional roles in cellular metabolism, physiology, and organismal-level function. Regardless of these attributes, it has been the target of a body of phylogeographic work that assumes neutrality, because it has been relatively easy to sequence and analyze in a diversity of non-model systems.

However, this assumption of mtDNA neutrality is not always true. Variation in mtDNA has been shown to have effects on fitness, often linked to thermal tolerance, in a variety of different species. These include hares (Slimen *et al.* 2017), tuna (Dalziel *et al.* 2006), and humans (Balloux *et al.* 2009). In the mammals, variance in the *ATP6* mitochondrial gene was connected to shifts in temperature tolerance. Mechanistically, *ATP6* variants reduce the coupling efficiency of oxidative phosphorylation, which increases heat production while simultaneously increasing basal metabolic rate (Ballard and Whitlock 2004). This mechanism results in a tradeoff - a population with an *ATP6* variant sees increased cold tolerance but decreased fitness in more temperate environments. This example illustrates how mtDNA can be under selection, and that thermal tolerance can be the selective factor. Furthermore, the mitochondrial genome has a great deal of interaction with the nuclear genome. Some of this interaction happens during mtDNA replication, transcription, and translation - all of which are reliant on products from the nuclear genome. The mitochondrial proteins, which play an essential role in the oxidative phosphorylation (OXPHOS) system, also interact extensively with nuclear proteins (Burton and Barreto 2012). The mitochondrial genome is also particularly prone to point mutations, and many animal species have a rate of mitochondrial substitution over 20 times the rate of nuclear substitution (Burton and Barreto 2012). This leads to a high rate of deleterious mutations, with the nuclear genome taking over roles opened by mitochondrial knockout mutations. This process of mitonuclear coadaptation may have caused the long-term shrinkage of the mitochondrial genome (Rand *et al.* 2004, Gray 2011) This tight coevolution can cause hybrids to suffer decreased fitness, as their mtDNA and nuclear DNA are no longer performing complementary roles (Burton and Barreto 2012). Thus, such mitonuclear interactions are expected on traits linked to fitness in a broad variety of systems, and should express themselves in hybrid zones, where new combinations of mitochondrial genomes and nuclear genes are possible in F1 hybrids and down-stream crosses.

The green crab hybrid zone in the Gulf of Maine presents an excellent opportunity for gaining insight into any potential phenotypic consequences of nuclear variation, mitochondrial variation, or mitonuclear interactions. This thesis is based on the hypothesis that nuclear and/or mitochondrial variation has a strong, measurable impact on physiological performance in different temperatures. Over the course of this project, three genes - one mitochondrial and two nuclear - were examined for linkages to cold tolerance. Along with the reasons already laid out for examining a mitochondrial gene, this particular gene - *Cytochrome oxidase 1*, or *CO1* - was chosen due to the large variety of sequences available in GenBank. The regions of the nuclear genome we examined were identified from previous work by Dr. Carolyn Tepolt and Dr. Stephen Palumbi (Tepolt and Palumbi 2015), who used transcriptome sequencing to examine genetic variance within the native range of *C. maenas*. They identified several sites that showed significant divergence between the northern and southern extents of the native range. These nuclear sites will henceforth be referred to as *SMC* and *283*. Both *SMC*

and *283* have a known SNP that varies between the southern and northern populations. *SMC* also has many characteristics associated with an inversion, or 'supergene'. By examining both physiological performance under cold stress and variance in *SMC*, *283*, and *CO1*, we determined whether variance in genotype could be definitively linked to physiological cold tolerance.

METHODS

There are two main sections of this study. First, individuals from a variety of locations - and presumably with a wide range of genotypes - were collected and tested for physiological cold tolerance. To test physiological cold tolerance, righting response was measured. Measuring righting response - or the time it takes for a crab to right itself after being placed on their carapace - has been shown to be an excellent proxy for physiological stress in a variety of decapods, including *C. maenas* (Ern *et al.* 2015, Cuculescu *et al.* 1998). Next, tissue samples were taken from these individuals for sequencing of two nuclear sites and one mitochondrial gene. These sites were then analyzed to determine whether allelic variation could be linked to differences in cold tolerance.

Testing Cold Tolerance

Crabs were collected from three locations along the east coast of North America - Harpswell, Maine (N 43º47'27", W 69º57'34") Kent Island, New Brunswick, (N 44º33'00", W 66º45'23") and Pomquet, Nova Scotia (N 45º38'27", W 61º48'41") (Fig. 1). The method of crab collection differed between sites, with crabs collected either by hand in the low intertidal or via traps placed in the shallow subtidal. After collection, crabs were transported to the Bowdoin College Coastal Studies Center in coolers filled with either aerated seawater or moist rockweed (*Ascophyllum nodosum*).

To negate for the effects of acclimation, which can have a strong impact on cold tolerance (Tepolt and Somero 2014), crabs were acclimated for a period of 3-5 weeks. Setpoint for acclimation temperature was 15.5ºC. Crab density was maintained at 28 to 60 crabs per tank. To ensure consistency within each site, individuals from each group were exposed to exactly the same seawater. One day before testing, the Kent Island group was unintentionally exposed to water at 10ºC for roughly four hours. Tanks were kept on a 12:12 light:dark cycle. Crabs were fed with herring or shrimp three times a week. Water was constantly aerated throughout the acclimation. Testing began between 7-10 am and ran until 9pm-1am. Light exposure was kept relatively constant. Crabs were noticeably more active later in the day. Crabs were not fed in the days immediately leading up to the experiment, with the last feeding taking place 4-5 days before the start of the testing to ensure that differential food consumption did not impact the experiment.

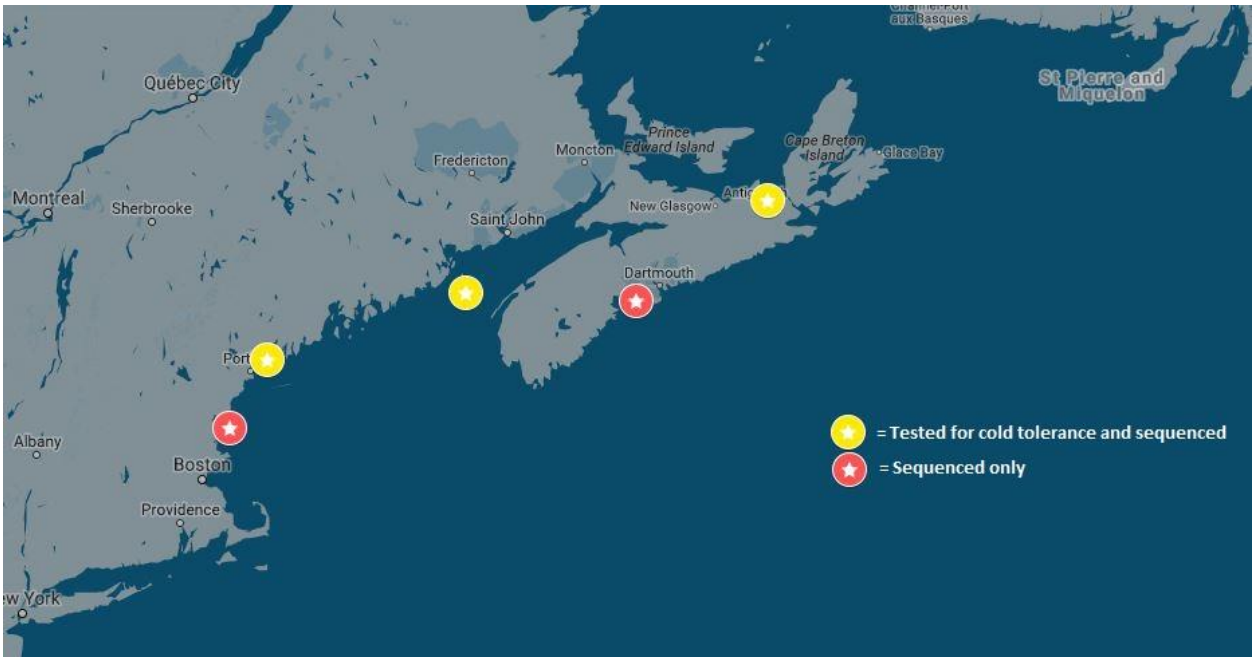


Figure 1. The five sites sampled in this study. South to north, the sites are: Isles of Shoals, ME (n = 12), Harpswell (n = 93), Kent Island (n = 64), Halifax (n = 15), and Pomquet (n = 86). At the yellow sites (Harpswell, Kent Island, and Pomquet), live crabs were brought back and tested for physiological performance under cold stress. Three sites in the genome of each crab were then sequenced. Additional genetic samples were collected from the red sites (Isles of Shoals, Halifax)

To test righting response, crabs were placed in a large plastic tote filled with aerated seawater at 4.5ºC. This tote was contained within an incubator, which allowed the water to stay at a constant temperature throughout the experiment. Crabs were left in the seawater for five minutes to control for size effects. After the five-minute cooling period, a metal ruler was used to flip the crab onto its back, and a stopwatch was used to time how quickly the crabs could right themselves. Righting was defined as when the rearmost dactyls maintained contact with the surface of the tote. Crabs that did not right themselves after two minutes were determined to be unable to right themselves. In these experiments, variance in righting time was related to carapace width (larger crabs took longer) but not with other independent variables. Therefore, whether the crab righted or not was used as the response variable.

After righting response was tested, a variety of morphological parameters - sex, carapace width, number of legs, and morph type (red or green) - were recorded. Morph type has been shown to affect tolerance to a variety of environmental stressors, including salinity shifts, emersion, and pollution (Styrishave *et al.* 2004). A single leg of the crab was then broken off and preserved in 95% ethanol. The crab was then euthanized via freezing.

Genotyping

After a storage period of roughly 1-2 months, muscle tissue was removed from the ethanol-preserved leg and digested. If the leg was too small to adequately remove muscle tissue, the entire leg was cracked in several places with tweezers. In addition to the 251 crabs collected in Part 1, 15 crabs from Halifax, Nova Scotia (N 44º38’28”, W 63º55’45”) were sequenced, along with 22 crabs from Isles of Shoals, Maine (N 42º59’23”, W 70º36’58”) (Fig. 1). These additional samples were used solely to further establish the genetic structure of *C. maenas* along the coast of North America, as all physiological characteristics of crabs from these locations are unknown. Many of the Isles of Shoals samples failed to amplify.

DNA extraction was performed with a Qiagen DNEasy Blood & Tissue Kit (Valencia, CA) following manufacturer's instructions. The purified DNA was diluted to 2 ng/ul and amplified. PCR amplification was performed using Lucigen MasterMix, with the following cocktail: 12.5 ul Lucigen MasterMix 2x, 0.5 ul of forward primer (10 uM), 0.5 ul of reverse primer (10 uM), 1 ul of DNA (2 ng/ul), 10.5 ul water. Primer sequences and PCR profiles for each site are given in Appendix A. Positive amplification and the absence of contamination via non-template DNA was verified with an agarose gel. Amplified products were prepared for sequencing using an exonuclease and shrimp alkaline phosphatase incubation (Exo-SAP). The cocktail consisted of 10 ul amplified product (2 ng/ul) and 2 ul ExoSAP mix. This cocktail was then incubated at 37ºC for 30 minutes and 95ºC for 5 minutes. 5 ul of the forward primer was then added (5 uM). All PCR products were sequenced in the forward direction using the commercial sequencing service GeneWiz. The resulting chromatograms were edited and aligned using the software package Geneious (Biomatters, New Zealand). Each nuclear sequence was characterized by a single nucleotide polymorphism (SNP) which was called by visually inspecting chromatograms for the expected double or single nucleotides. For the mitochondrial *CO1* gene, six haplotypes were identified within the samples collected. Various *CO1* sequences of *C. maenas* within their native range were then downloaded from Genbank (AY616437.1-AY616445.1, JQ305941.1-JQ305943.1, JQ306002.1-JQ306003.1, KF369118.1), and a maximum likelihood tree was constructed using the PHYML module in Genious with 100 bootstraps. A TCS haplotype network was also constructed of *CO1* haplotypes using the software PopART (http://popart.otago.ac.nz/).

The tree and network were then used to determine relation of each *CO1* haplotype to samples from Europe.

Statistical modeling

To analyze the effects of both categorical (sex, mitochondrial and nuclear genotypes, morph color, and number of legs) and continuous (carapace width) independent variables on righting response, we used a multiple logistic regression model implemented by the R-package PerformanceAnalytics. Righting response was categorical, with crabs either righting or failing to right within the two-minute trial. A binary variable for righting response was chosen over a continuous variable such as time to right because the majority of crabs either flipped rapidly within the two-minute trial or remained on lying on their carapace (dorsal side) and failed to right during the two-minute observation period. Two models were run. First, a model was run on all physiological variables (sex, morph type, number of legs, sampling location, and carapace width) to determine whether any of these variables could be associated with righting response. The results of the physiological model were used to select variables for the final model, which included sex and sampling location along with three genetic terms for the mitochondrial and nuclear variants.

RESULTS

Genetic polymorphism in CO1

Six unique mitochondrial haplotypes were sampled from the three Canadian and two US sites (Fig. 2). The two sites in Nova Scotia had a high frequency of two haplotypes that are either closely related or identical to haplotypes commonly found in northern Europe. One of these haplotypes was identical to a Norwegian haplotype (B1, Fig. 2), while another was closely related to haplotypes from both Norway and the Faroe Islands (B2, Fig. 2). These two haplotypes were grouped within B as ‘northern’ haplotypes. The Nova Scotian samples also had a high frequency of a third haplotype that appears to be related to a haplotype previously observed in Portugal (C, Fig. 2). This haplotype is likely a pan-European haplotype found in previous studies to comprise roughly 30-40% of Nova Scotian haplotypes (Darling *et al.* 2008), which agrees with the frequency observed in this study. However, since it was not closely related to the northern haplotype group, it was binned in its own C group. These three haplotypes (B1, B2, and C) combined to dominate the Nova Scotian sites (dark blue and brown shading, Fig. 2). The southern sites (Kent Island, Harpswell, and Isles of Shoals) were dominated by a haplotype identical to a GenBank sample from France (A1, Fig. 2). Two other haplotypes, A2 and A3, were closely related to A1. Haplotype diversity was much lower at the southern sites than at the northern sites (Fig. 3).

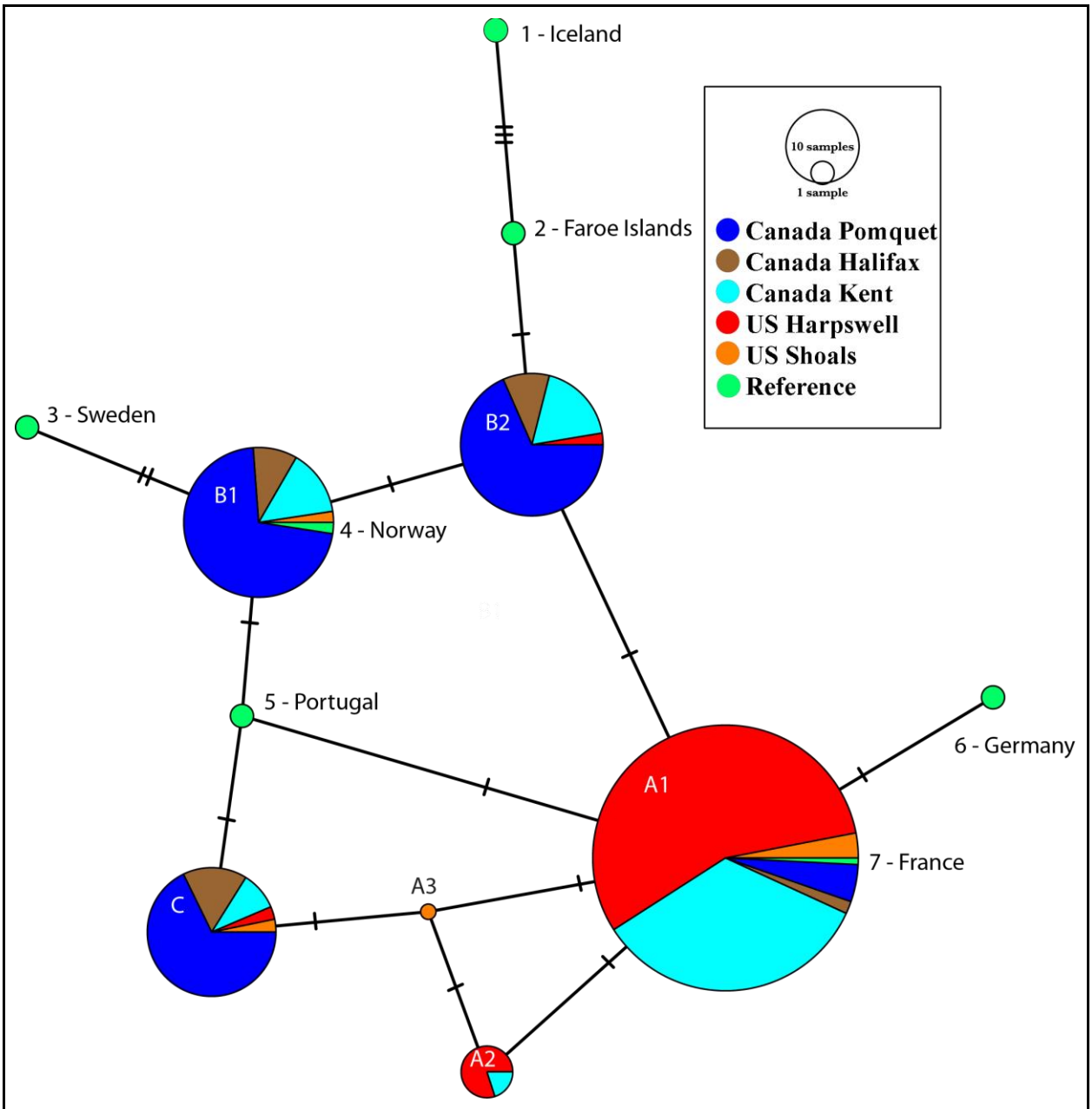


Figure 2. Haplotype network of the *CO1* gene sampled in *C. maenas* from five sites on the east coast of North America along with European reference sequences. Each node is a unique sequence, with the size of the node corresponding to how many samples shared that sequence. Each hash mark corresponds to one mutational step. Color corresponds to the sampling location as described in the legend. Green samples are reference sequences from GenBank. Haplotype labels group three major haplotype groups: A1-3 = Southern Europe, B1+2 = Northern Europe, C = Pan-European.

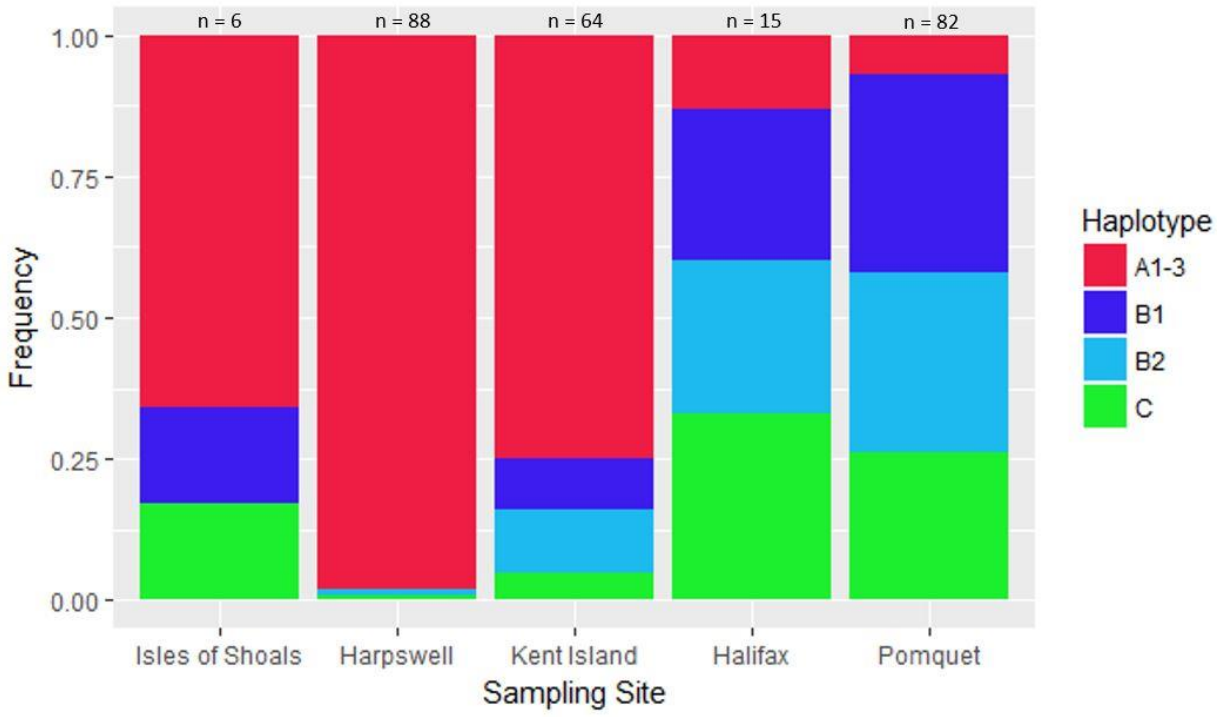


Figure 3. The frequency of the four major mitochondrial haplotype groups at five sampling sites. The sites are ordered by latitude, with the southernmost site on the left side. Number of samples is shown at the top of each bar. Haplotype letters correspond to those in Fig. 2.

Nuclear SNP variation

Hardy-Weinberg expectations for *283* and *SMC* were tested within each site. All sites were within HWE (Appendices C-D). The two loci showed dramatically different cline structure. The *283 A* nucleotide increased in frequency moving from south to north (Fig. 4), while the *SMC T* nucleotide reached the highest frequency on Kent Island, and decreased in the south and north (Fig. 5).

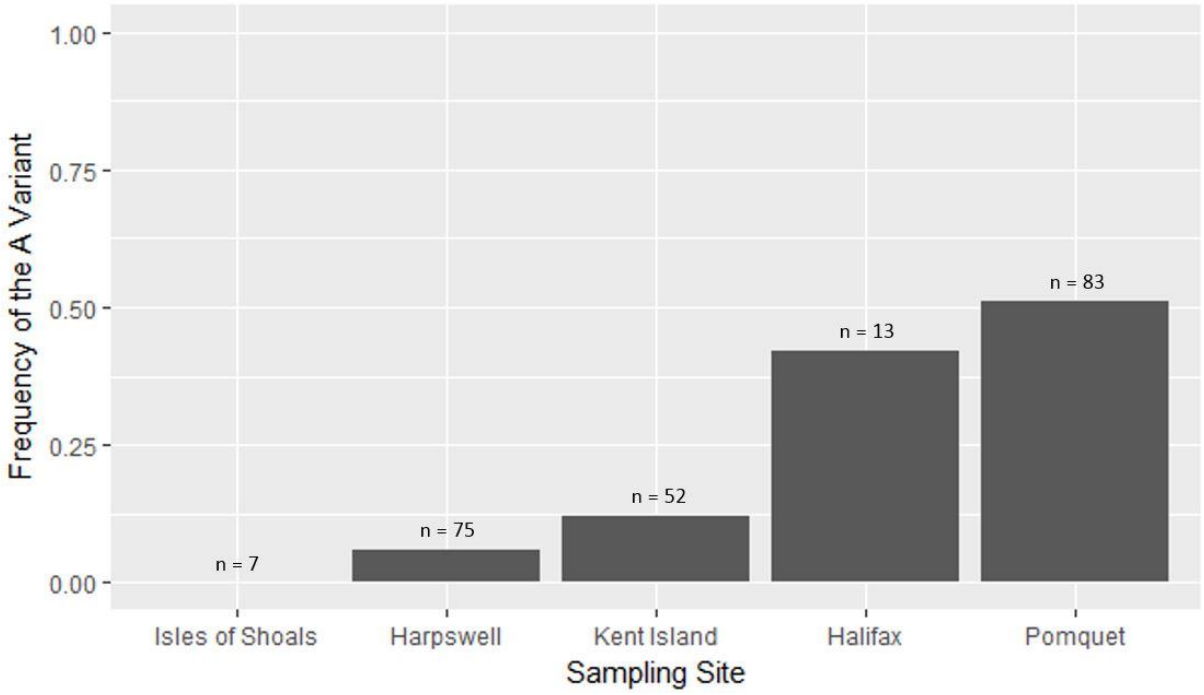


Figure 4. Frequency of the *283 A* nucleotide by sampling location. The sites are ordered by latitude, with the southernmost site on the left. The number of samples per site is shown above each bar.

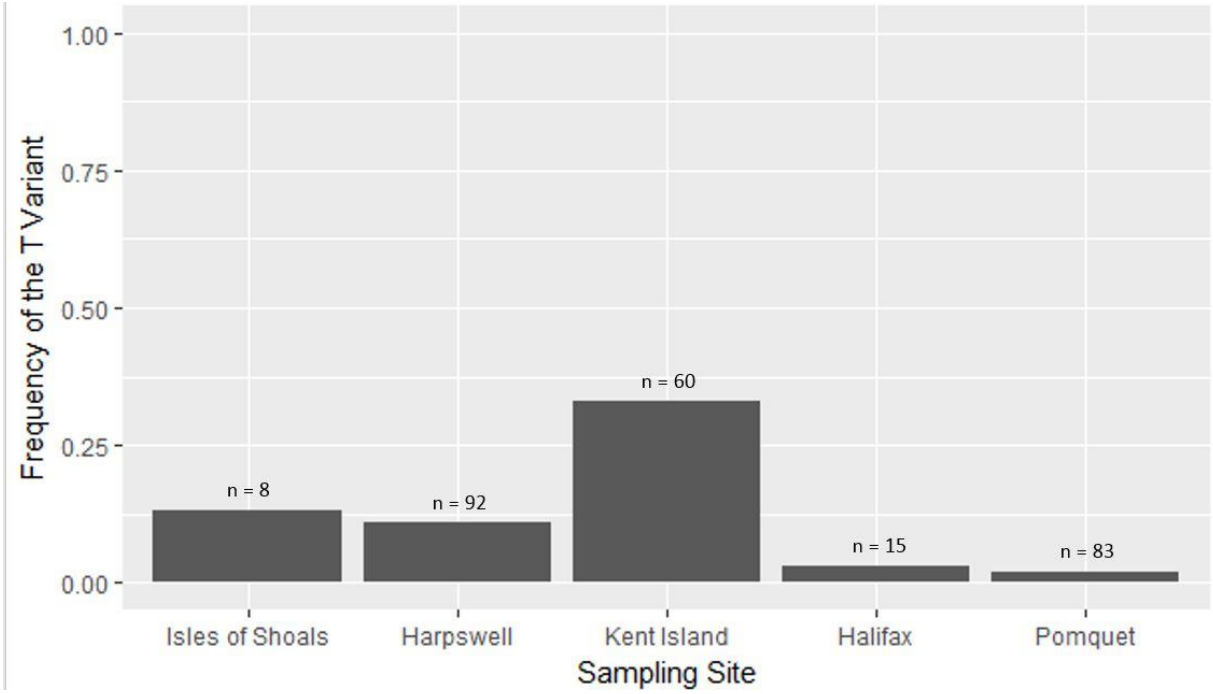


Figure 5. Frequency of the *SMC T nucleotide* by sampling location. The sites are ordered by latitude, with the southernmost site on the left. The number of samples per site is shown above each bar.

Multiple logistic regression models of righting response under cold stress

Only the effects of sex and sampling on righting response were significant in the model considering only physiological variables (Table 1). Therefore, we added these terms to a model with the three genetic terms. In this model, sex was marginally significant, while *CO1* clade explained the most variation in righting response (Table 2). The effect of these two terms is graphically presented in Fig. 6. Northern (B) and pan-European (C) haplotypes increase the probability of righting under cold stress by 20% in males compared to southern haplotypes (A) In fact, every single male with a Northern or pan-European haplotype was able to successfully right itself (Fig. 6). In contrast to the male response, females did not appear to be affected by *CO1* haplotypes, but sample sizes were small (Fig. 6).

Table 1. The results of a multiple regression logistic model of the effects of five physiological variables on righting response. Yellow shading indicates significant variables.

Factor	Deviance	AIC	LRT	Pr (>Chi)
Sex	195.10	201.27	2.169	0.149
Morph Type	197.27	202.96	0.144	0.701
Missing Legs	194.96	202.99	0.111	0.739
Carapace Width	195.09	203.09	0.012	0.913
Sampling site	203.12	207.12	8.015	0.0046

Table 2. The results of a multiple regression logistic model of the effects of site, sex, variance in *283* and *SMC*, and *CO1* clade on righting response. Three mt *CO1* clades were used, corresponding to the A, B, and C groups in Fig. 2.

Factor	**Deviance**	**AIC**	**LRT**	**Pr (>Chi)**
Site	156.73	164.73	0.052	0.819
Sex	160.56	164.56	3.782	0.052
283 (nuclear SNP)	156.46	164.46	0.327	0.573
SMC (nuclear SNP	156.76	164.76	0.0207	0.885
CO1 clade (A/B/C)	165.81	169.81	9.0334	0.00265

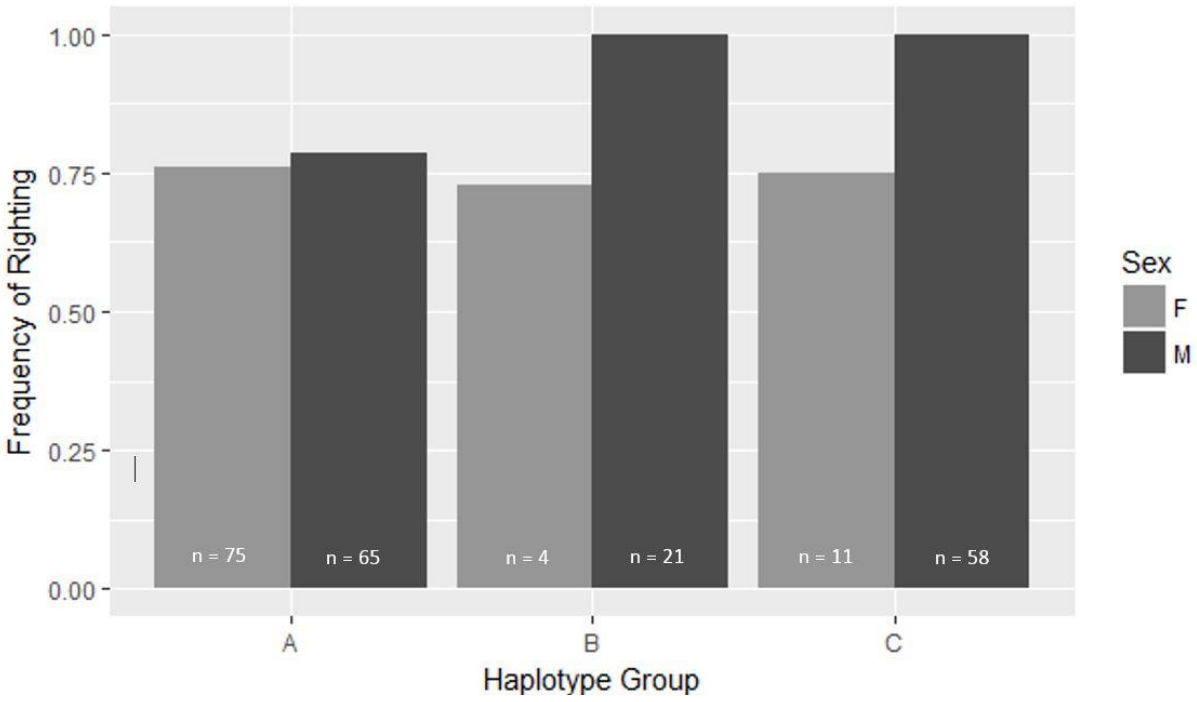


Figure 6. The relationship between mitochondrial haplotype group and righting response for male and female *C. maenas* at 4.5ºC. Haplotype groups are identical to Fig. 2. There is a significant relationship between mitochondrial haplotype and righting response: male crabs carrying C (pan-European) or B (Northern) haplotypes are more likely to right under cold stress than those carrying A (Southern) haplotypes

DISCUSSION

Connecting genotype to phenotype

Our results clearly indicate a strong effect of mitochondrial genotype on organismal fitness. In the combined genetic and physiological logistic regression model, mitochondrial haplotype had the largest impact on righting response (Table 2, LRT). Crabs with pan-European (C) or Northern (B) mitochondrial haplotypes were 20% more likely to right themselves than crabs with Southern haplotypes (A) (Fig. 6). It is important to note that these results do not indicate that the *CO1* gene specifically plays a role in cold tolerance. Instead, since mitochondrial genes are typically inherited entirely from one parent and do not undergo recombination, the entire mitochondrial genome is inherited as a single block. This means that there is effectively complete linkage between the *CO1* gene and all other mitochondrial genes.

In our combined genetic and physiological logistic regression model (Table 2), both sex and mitochondrial haplotype were shown to have an impact on righting response. At first glance, it is not particularly surprising that females were less likely to be able to successfully right themselves, as females deal with additional physiological burdens associated with egg development. However, there is minimal research showing a lower physiological tolerance for stress in female *C. maenas*. In fact, there are studies that directly counter this, finding higher resilience for *C. maenas* females under salinity stress (Pennoyer *et al.* 2016). This could be the result of low sample size, as only 15 females had either a northern or pan-European sequence (Fig. 6). As an alternate possibility, mating-related stress could play a role, as the male to female ratio was disproportionately high in the Pomquet, Nova Scotia tank. Since that tank contributed a large proportion of northern and pan-European haplotypes among the crabs tested for righting response (Fig. 2, Fig. 3), stress associated with mating (and rejection of males) could have affected the physiological response of northern and pan-European females. Alternately, it is possible that *CO1* genotype has differential impacts on male and female crab phenotypes. Furthermore, though sex contributed to the accuracy of the model, the factor-specific p-value was not significant (Table 2). A larger sample size of females with northern and pan-European sequences is needed to determine whether sex truly has an impact on cold tolerance.

While some mitochondrial genes certainly seem to play a role in determining thermal tolerance, this does not mean that nuclear genes do not contribute to this phenotype. Indeed, mitonuclear interactions - or the interaction of mitochondrial and nuclear proteins - may play an essential role in determining thermal tolerance. Mitochondrial genes are all involved in either translating the mtDNA or in producing proteins involved in the electron transport chain (Ballard and Whitlock 2004). However, while there are under 40 mtDNA genes (Ballard and Whitlock 2004), there are roughly 1500 nuclear genes involved in oxidative phosphorylation, or OXPHOS (Gershoni *et al.* 2009). This means that OXPHOS is heavily dependent on mitonuclear interactions (Wolff *et al.* 2014). These protein interactions typically rely on a “lock-and-key” principle, in which the nuclear and mitochondrial proteins must match each other. As a result, it is quite possible that there are nuclear genes that covary with the mitochondrial genes, with both playing an important role in cold tolerance. A nuclear effect on phenotype was not seen for the two nuclear genes chosen in this study, indicating that they are not linked to genes involved in adaptations to cold tolerance, or that our sample sizes were not large enough to detect smaller nuclear effects. Interestingly, both SNPs were outliers in an analysis of transcriptomic variation between Northern and Southern Europe (Tepolt and Palumbi 2015). It is possible that these sites are connected to some adaptive function between regions, but not directly related to righting response. The cline in *283* is quite similar to the cline in *CO1* (Fig. 3, Fig. 4), which hints that temperature may play a role in *283* frequency. Alternately, *283* may be under selection from a factor other than temperature

that has the same pattern of north/south variance, or the cline in *283* could be due to neutral demographic processes.

This connection between mitochondrial genome and thermal tolerance is not unique to *C. maenas.* In fact, there are numerous examples of species where the mitochondrial genome is a site of thermal selection. While the majority of these are endotherms, there are ectotherms where the mitochondrial genome plays a role in thermal tolerance, such as *Drosophila* (Camus *et al*, unpublished). In fact, “among intertidal species, in particular, thermal adaptation is frequently driven by selection on mitochondrial genomes” (Darling 2014). One such example is the abalone (genus *Haliotis*), which shows significant variance between congeners in thermal tolerance. (Somero 2002). As an intertidal and shallow subtidal species, *C. maenas* is forced to deal with extreme temperatures, which certainly provides a selective pressure for increased thermal tolerance.

The main mechanistic hypothesis for how mitochondrial genome impacts cold tolerance in ectotherms is oxygen and capacity-limited thermal tolerance (OCLTT) (Ballard and Melvin 2010, Ern *et al.* 2015). According to this hypothesis, cold-adapted ectotherms have a higher mitochondrial capacity, but also an increased metabolic rate (Ballard and Melvin 2010). This results in a shift in thermal tolerance rather than a broadening. Greater cold tolerance is a trade-off, with lower heat tolerance as its result. We see this pattern in *C. maenas*. Within their native range (and within the east coast of North America), northern populations, that presumably are dominated by B and C haplotypes, are more cold-tolerant than southern populations, presumably dominated by A haplotypes. However, the southern populations are more heat-tolerant than the northern populations (Tepolt and Somero 2014). This see-saw balance of heat and cold tolerance provides evidence that mitochondrial adaptation in *C. maenas* may follow the OCLTT model of thermal tolerance, but to establish this, linkage between specific mitochondrial genes and thermal tolerance is needed. Moreover, we did not examine an effect of mitochondrial genome on heat tolerance in this study.

It remains unknown how mitochondrial haplotypes affect the thermal tolerance of planktonic larvae. In general, *C. maenas* larvae have a much narrower band of thermal tolerance than adults (Hines *et al.* 2004), meaning that the larval thermal tolerance could restrict the range of *C. maenas*. If mitochondrial haplotype has a significant impact on larval thermal tolerance, then the mitochondrial haplotype can be an excellent way to determine the range expansion potential of invasive populations. This would also mean that management policy should be focused on minimizing genetic diversity in invasive populations. To look at the *C. maenas* population on the west coast of the United States as a case study, a larval mitochondrial/thermal tolerance linkage would mean that keeping northern mitochondrial haplotypes out of this population is essential to preventing a dramatic northern expansion.

The linkage between mitochondrial variance and thermal tolerance can be further evaluated in the field. A simple field-based survey could link early-season variability in activity to mitochondrial genotype. Within the Bay of Fundy, *C. maenas* is believed to move to deeper waters in the winter and then migrate to shallower waters as the water warms in the spring (Klassen and Locke 2007). If individuals with a northern or pan-European haplotype (as defined in Figure 2) tend to be active in shallow waters significantly earlier than those with a Southern European haplotype, the linkage between mitochondrial haplotype and cold tolerance will translate to the field.

Phylogeography and neutral clines

Previous studies of the dynamics of the green crab hybrid zone have focused on neutral interpretations of genetic variation (Roman and Palumbi 2004, Pringle *et al.* 2011, Darling *et al.* 2014). Regardless of the mechanisms involved, it is clear that mitochondrial genotype influences cold tolerance in *C. maenas*. This result suggests that differences in cline structure between mitochondrial markers and putatively neutral nuclear markers like microsatellites (Darling *et al.* 2014) must be related to environmental gradients as well as stochastic demographic processes. With regards to gradients in temperature in the Gulf of Maine and the Canadian Maritimes, the influence of the cold Labrador current is particularly strong as it moves southward along the coast of Nova Scotia (Pettigrew *et al.* 2005) and into the upper reaches of the Gulf of Maine. Thus, the thermal transition from colder waters to seasonally much warmer temperatures in the lower gulf could explain the mt *CO1* cline documented here.

While *283* and *CO1* appeared to have a generally similar cline structure, the nuclear SNP *SMC* was dramatically divergent from this north-south cline. Instead, the *T* nucleotide of *SMC* peaked at Kent Island, within the hybrid zone (Fig. 5). This does not seem to be the result of neutral genetic drift, as this would either result in a cline between the two sites of the invasion events (in this case, a general north/south cline) or in no cline. Alternatively, this pattern could be the result of “cryptic” invasions of this genotype to the islands of Grand Manan. It is also possible that *SMC* is under selection. This selection is likely not due to temperature, as Kent Island has lower temperatures than the Maine sites and higher temperatures than the Nova Scotian sites. There are a variety of environmental parameters that may vary between the Kent Island site and other sites, including much more extreme tides, differences in algal canopy makeup and cover, and possible differences in diet or predation. *SMC* could be responsible for adaptation to any of these parameters, could play a role in reproductive processes, or could have some other impact on fitness. However, given that the *SMC* variant clearly appears to be under selection, future research is needed to determine the impact of *SMC* on phenotype.

These findings of non-neutrality for the mitochondrial genome may help to explain questions raised by past research. The population structure of the *Carcinus* genus in Japan may be one such example. This population has the nuclear genes of *C. aestuarii*, a closely-related congener of *C. maenas*, and a mix of *C. maenas* and *C. aestuarii* mitochondrial haplotypes. This likely stems from a single invasion event of *C. maenas/aestuarii* hybrids (Darling 2011). There appears to be some level of spatial ordering in mitochondrial genes, with *C. maenas* haplotypes dominating Southern Japan, while *C. aestuarii* haplotypes dominate more northern areas. Assuming that thermal tolerance is a trade-off, as it is under the OCLTT model, increased thermal tolerance would be unlikely to develop for *C. aestuarii*, as its native range in the Mediterranean is panmitic (Darling 2011). Heat tolerance, however, could certainly develop in the Atlantic North African population of *C. maenas*, as there would likely be relatively low gene flow with northern Europe. An invasive population with haplotypes from both *C. aestuarii* and North African *C. maenas* would, therefore, likely segregate with *C. maenas* haplotypes in the hotter areas while *C. aestuarii* dominates cooler areas. This pattern appears in Japan (Darling 2011), making it an intriguing area for research into mitochondrial impacts on heat tolerance.

Summary

This finding of a clear linkage between mitochondrial haplotype and thermal tolerance helps answer some questions (such as the population structure in the Japanese population of *C. maenas*), but raises far more. Does the linkage to cold tolerance vary according to the OCLTT model? Are there mitonuclear interactions? Is a similar relationship seen in larvae or in the field? The mitochondrial non-neutrality also calls into doubt the current phylogeographical models of *C. maenas* population structure. This means that new models will have to be constructed using truly neutral genes, and that the old models can be analyzed for data on spatial variation in mitochondrial selection. In all, this study opens numerous avenues for future research on mitochondrial genes, population structure, and physiology of *C. maenas*.

Acknowledgements:

I would like to thank my advisor, Dave Carlon, for his expertise and guidance. Without him, this project would not have even been remotely possible. I would also like to thank Dr. Carolyn Tepolt for generously providing transcriptome sequence of outlier SNPs, and for her feedback on my honors research. Dr. Robin Hadlock-Seeley provided green crab

samples from the Isles of Shoals, and Thew Suskiewicz and Ladd Johnson collected green crabs from Halifax, Nova Scotia. Many thanks to Nick Keeney for his work to keep the lab running, and to both Hugh Cipperoni and the members of the 2016 Bowdoin Marine Science Semester for their efforts in crab collection. Thanks also to Sarah Kingston for her guidance in papers, presentations, and border crossings, and to both Sarah and Vlad Douhovnikoff for sharing their lab space. This project was funded by the Quahog Bay Conservancy and the Grua-O'Connell Research Award

References:

Ballard J.O. and Melvin R.G. 2010. "Linking the mitochondrial genotype to the organismal phenotype". *Molecular Ecology* 19(8): 1523-1539

Ballard J.O. and Whitlock M.C. 2004. "The incomplete natural history of mitochondria". *Molecular Ecology* 13: 729-744

Balloux F., Handley L.L., Jombart T., Liu H. and Manica A. 2009. "Climate shaped the worldwide distribution of human mitochondrial DNA sequence variation". *Proceedings of the Royal Society B*. DOI: 10.1098/rspb.2009.0752

Barton, N. H. and G. M. Hewitt. 1985. "Analysis of hybrid zones". *Annual Review of Ecology and Systematics* 16: 113-148.

Burton R.S. and Barreto F.S. 2012. "A disproportionate role for mtDNA in Dobzhansky-Muller incompatibilities?" *Molecular Ecology* 21: 4942-4957

Camus M.F., Wolff J.N., Sgro C.M. and Dowling D.K. Unpublished. "Experimental evidence that thermal selection has shaped the latitudinal distribution of mitochondrial haplotypes in Australian *Drosophila melanogaster*". *bioRxiv*

Cuculescu M., Hyde D. and Bowler K. 1998. "Thermal tolerance of two species of marine crab, *Cancer pagurus* and *Carcinus maenas*". *Journal of Thermal Biology* 23(2): 107-110

Dalziel A.C., Moyes C.D., Fredriksson E. and Lougheed S.C. 2006. "Molecular Evolution of Cytochrome c Oxidase in High-Performance Fish (Teleostei: Scombroidei)". *Journal of Molecular Evolution* 62(3): 319-331

Darling J.A. 2011. "Interspecific Hybridization and Mitochondrial Introgression in Invasive *Carcinus* Shore Crabs". *PLOS ONE* 6(3): e17828. htt
ps://doi.org/10.1371/journal.pone.0017828

Darling J.A., Bagley M.J., Roman J., Tepolt C.K. and Geller J.B. 2008. “Genetic patterns across multiple introductions of the globally invasive crab genus *Carcinus*”. *Molecular Ecology* 17(23): 4992-5007

Darling J.A., Tsai E.Y., Blakesee A.M.H. and Roman J. 2014. “Are genes faster than crabs? Mitochondrial introgression exceeds larval dispersal during population expansion of the invasive crab *Carcinus maenas*”. *Royal Society of Open Science* DOI: 10.1098/rsos.140202

Ern R., Huong D.T.T., Phuong N.T., Madsen P.T., Wang T. and Bayley M. 2015. “Some like it hot: Thermal tolerance and oxygen supply capacity in two eurythermal crustaceans”. *Scientific Reports.* 5, doi:10.1038/srep10743

Geneious [computer software]. 2015. Biomatters Limited, New Zealand.

Gershoni M., Templeton A.R. and Mishmar D. 2009. “Mitochondrial bioenergetics as a major motive force of speciation”. *BioEssays* 31: 642-650

Gray M.W. 2011. “The incredible shrinking organelle”. *EMBO Reports* 12(9): 873

Harrison R.G. 1993. “Hybrid Zones and the Evolutionary Process”. Oxford University Press, New York

Hidalgo F.J., Barón P.J. and Orensanz J.M. 2005. “A prediction come true: the green crab invades the Patagonian coast”. *Biological Invasions* 7: 547-552

Hines A., Ruiz G., Hitchcock N.G. and De Rivera C.E. 2004. “Projecting Range Expansion of Invasive European Green Crab (*Carcinus maenas*) to Alaska: Temperature and Salinity Tolerance of Larvae”. *Smithsonian Environmental Research Center*. Research report for Prince William Sound Regional Citizens’ Advisory Council.

Klassen G. and Locke A. 2007. “A Biological Synopsis of the European Green Crab, Carcinus maenas”. Canadian Manuscript Report of Fisheries and Aquatic Sciences No. 2818

Ladoukakis E.D. and Zouros E. 2001. “Recombination in Animal Mitochondrial DNA: Evidence from Published Sequences”. *Molecular Biology and Evolution* 18(11): 2127-2131

Pennoyer K.E., Himes A.R. and Frederich M. 2016. “Effects of sex and color phase on ion regulation in the invasive European green crab, *Carcinus maenas*”. *Marine Biology* 163(137): doi:10.1007/s00227-016-2910-2

Pettigrew N.R., Churchill J.H., Janzen C.D., Mangum L.J., Signell R.P., Thomas A.C., Townsend D.W., Wallinga J.P. and Xue H. 2005. “The kinematic and hydrographic structure of the Gulf of Maine Coastal current. *Deep Sea Research Part II: Topical Studies in Oceanography*. 52:(19-21): 2369-2391

Pringle J.M., Blakeslee A.M., Byers J.E. and Roman J. 2011. “Asymmetric dispersal allows an upstream region to control population structure throughout a species’ range”. *Proceedings of the National Academy of Sciences* 108: 15288-15293

R Development Core Team. 2008. “R: A language and environment for statistical computing”. R Foundation for Statistical Computing, Vienna, Austria.

Rand D.M., Haney R.A. and Fry A.J. 2004. “Cytonuclear coevolution: the genomics of cooperation”. *Trends in Ecology and Evolution* 19(12): 645-653

Rand D.M., Spaeth P.S., Sackton T.B. and Schmidt P.S. 2002. “Ecological Genetics of Mpi and Gpi Polymorphisms in the Acorn Barnacle and the Spatial Scale of Neutral and Non-Neutral Variation”. *Integrative & Comparative Biology*. 42(4): 825-836

Roman J. 2006. “Diluting the founder effect: cryptic invasions expand a marine invader’s range”. *Proceedings of the Royal Society B*. Doi:10.1098/rspb.2006.3597

Roman J. and Palumbi S.R. 2004. “A global invader at home: population structure of the green crab, *Carcinus maenas*, in Europe” *Molecular Ecology* 13: 2891-2898

Schmidt P.S., Bertness M.D. and Rand D.M. 2000. “Environmental heterogeneity and balancing selection in the acorn barnacle *Semibalanus balanoides*”. *Proceedings of the Royal Society of London* 267: 379-384

Schmidt P.S. and Rand D.M. 2001. “Adaptive Maintenance of Genetic Polymorphism in an Intertidal Barnacle: Habitat- and Life-Stage-Specific Survivorship of Mpi Genotypes”. *Society for the Study of Evolution* 55(7): 1336-1344

Slimen H.B., Schaschl H., Knauer F. and Suchentrunk F. 2017. “Selection on the mitochondrial *ATP synthase 6* and the *NADH dehydrogenase 2* genes in hares (*Lepus capensis* L., 1758) from a steep ecological gradient in North Africa”. *BioMed Central Evolutionary Biology* 17(46): DOI: 10.1186/s12862-017-0896-0

Somero G.N. 2002. “Thermal Physiology and Vertical Zonation of Intertidal Animals: Optima, Limits, and Costs of Living”. *Integrative & Comparative Biology* 42(4): 780-789

Styrishave B., Rewitz K. and Andersen O. 2004. “Frequency of moulting by shore crabs *Carcinus maenas* (L.) changes their colour and their success in mating and physiological performance”. *Journal of Experimental Marine Biology and Ecology* 313: 317-336

Tepolt, C.K. and Palumbi, S.R., 2015, “Transcriptome sequencing reveals both neutral and adaptive genome dynamics in a marine invader”. *Molecular ecology*, 24(16), pp. 4145-58.

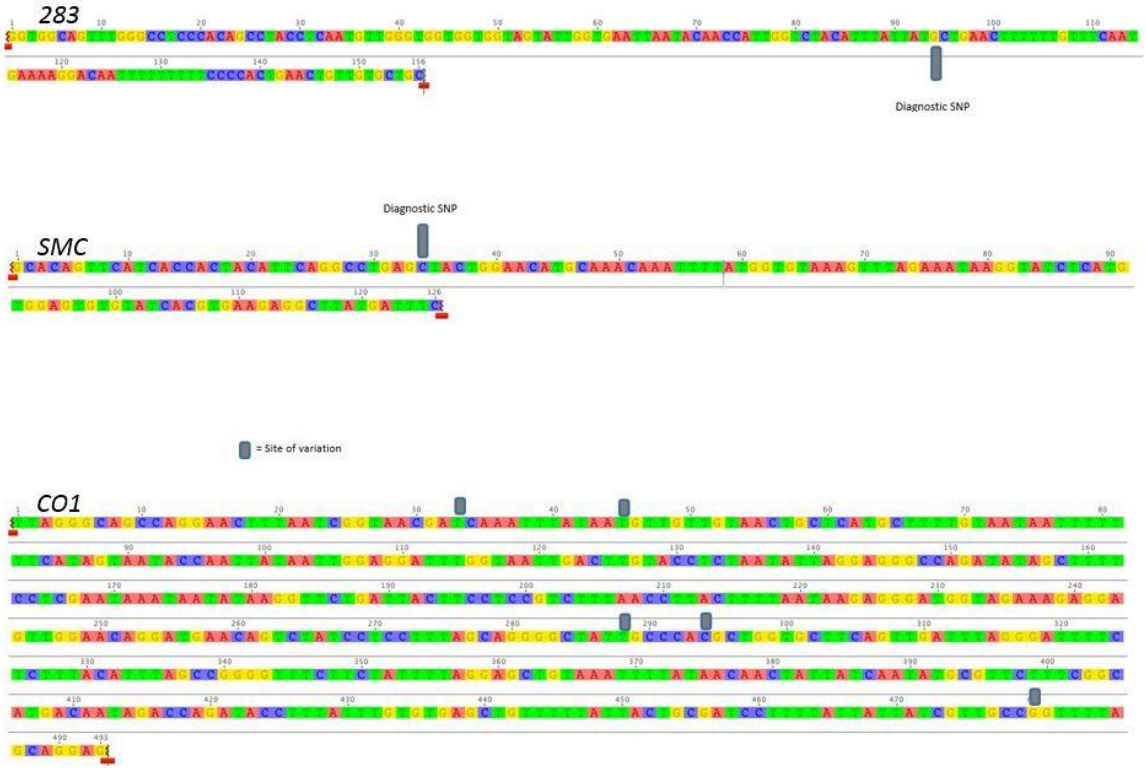
Tepolt C.K. and Somero G.N. 2014. “Master of all trades: thermal acclimation and adaptation of cardiac function in a broadly distributed marine invasive species, the European green crab, *Carcinus maenas*”. *Journal of Experimental Biology* 1129-1138

Wolff J.N., Ladoukakis E.D., Enriquez J.A. and Dowling D.K. 2014. “Mitonuclear interactions: evolutionary consequences over multiple biological scales”. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369(1646): 20130443 doi: 10.1098/rstb.2013.0443

Appendices:

Appendix A. Primers used to sequence the three sites examined in this study, along with amplification protocol. Length is defined as the reliable read obtained in each sample. Due to low quality reads at the ends of some sequences, the ends of the sequences were not included in this length. All loci were sequenced using the forward primer.

Sequence Name	Product Size (bp)	Primers	PCR Profile (annealing temperature highlighted)
283	156	F – GCCATGAGCATTTCTTTGAGTGG R – CGCAGTGGTGGAGTTTGGAA	3 min. 95ºC; 35 cycles (1 min. at 95ºC, 1 min. at 53ºC, 1 min. at 72ºC). 5 min. at 72ºC
SMC	126	F – AGCACAGGAAGGCTGTGG R – ACGAAATCATAAGCCTCTTCACG	3 min. 94ºC; 17 cycles (30 sec. at 94ºC, 30 sec. at 65ºC [-1ºC per cycle], 1.5 min. at 72ºC). 20 cycles (30 sec. at 94ºC, 30 sec. at 48ºC. 1.5 min. at 72ºC). 10 min. at 72ºC.
CO1	493	F – GCTTGAGCTGGCATAGTAGG R – GAATGAGGTGTTTAGATTTCG	3 min. 94ºC; 30 cycles (1 min. at 94ºC, 1 min. at 50ºC, 1 min. at 72ºC). 5 min. at 72ºC.



Appendix B. The sequence of the nuclear and mitochondrial sites examined in this study. Only the reliably high-quality portions of these sites (shown here) were examined for variance. Grey boxes mark sites of variance.

Appendix C. Observed and expected variance in *283* at each sampling site. *283* was in Hardy-Weinberg equilibrium for all sites.

Location	GG (obs.)	GA (obs.)	AA (obs.)	GG (exp.)	GA (exp.)	AA (exp.)	Chi-Square (χ^2)	P Value
Isles of Shoals	7	0	0	7	0	0	0	>0.05
Harpswell	67	7	1	66.27	8.46	0.27	2.2337	>0.05
Kent Island	42	8	2	40.69	10.61	0.69	3.1565	>0.05
Halifax	5	5	3	4.33	6.35	2.33	0.5849	>0.05
Pomquet	24	34	25	20.25	41.49	21.25	2.7073	>0.05

Appendix D. Observed and expected variance in *SMC* at each sampling site. *SMC* was in Hardy-Weinberg equilibrium for all sites.

Location	**CC (obs.)**	**TC (obs.)**	**TT (obs.)**	**CC (exp.)**	**TC (exp.)**	**TT (exp.)**	**Chi-Square (χ^2)**	**P Value**
Isles of Shoals	6	2	0	6.13	1.75	0.13	0.1633	>0.05
Harpswell	73	18	1	73.09	17.83	1.09	0.0088	>0.05
Kent Island	26	28	5	27.12	25.76	6.12	0.4450	>0.05
Halifax	14	1	0	14.02	0.97	0.02	0.0178	>0.05
Pomquet	80	3	0	80.02	2.95	0.03	0.0281	>0.05

Appendix E. Black and white versions of Figure 1 - 3, Tables 1-2, and Appendix A and B.

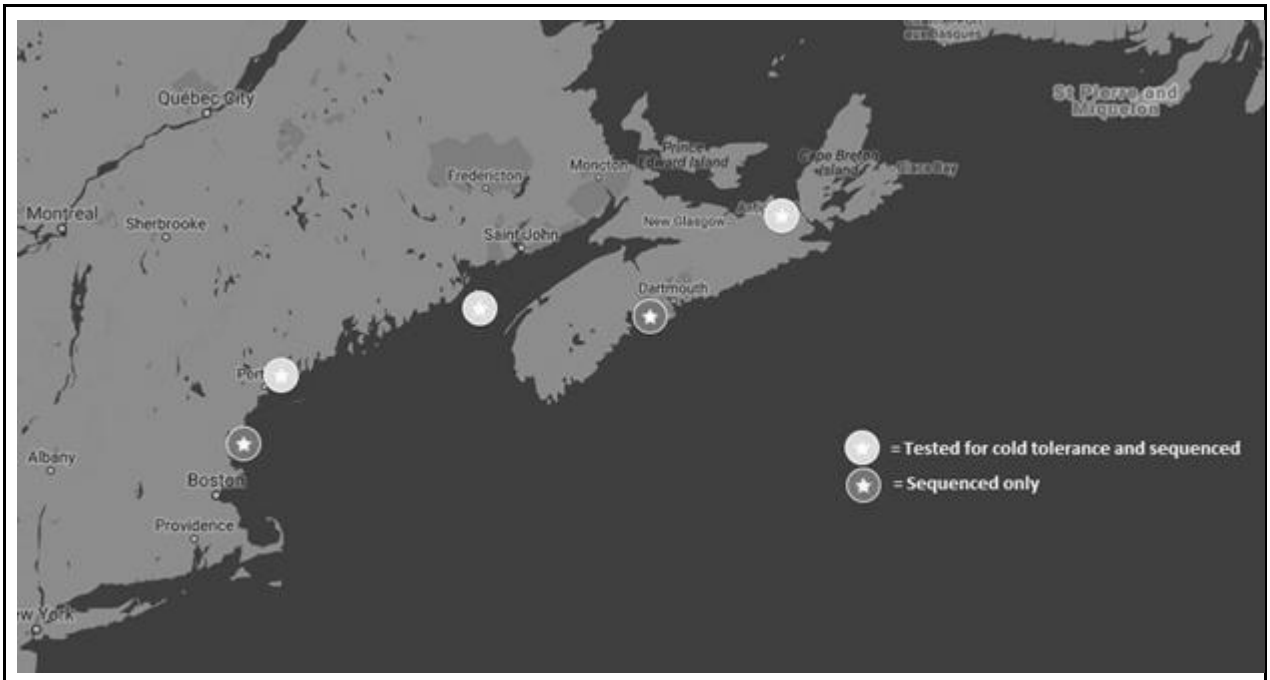


Figure 1. The five sites used in this study. South to north, the sites are: Isles of Shoals, ME (n = 12), Harpswell (n = 93), Kent Island (n = 64), Halifax (n = 15), and Pomquet (n = 86). At the yellow sites (Harpswell, Kent Island, and Pomquet), live crabs were brought back and tested for physiological performance under cold stress. Three sites in the genome of each crab were then sequenced. Additional genetic samples were collected from Halifax and the Isles of Shoals.

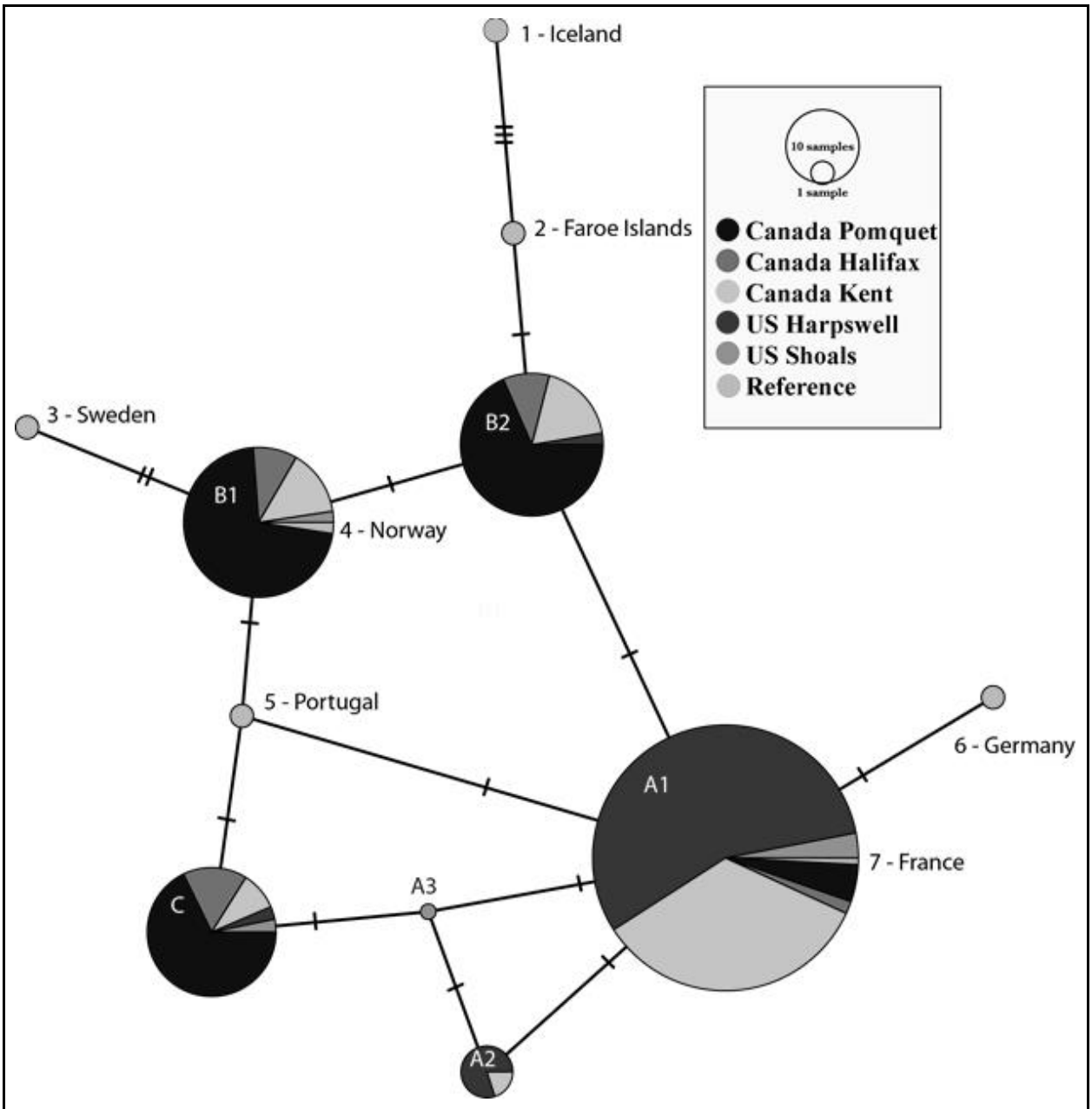


Figure 2. Haplotype network of the *CO1* gene sampled in *C. maenas* from five sites on the east coast of North America along with European reference sequences. Each node is a unique sequence, with the size of the node corresponding to how many samples shared that sequence. Each hash mark corresponds to one mutational step. Color corresponds to the sampling location as described in the legend. Green samples are reference sequences from GenBank. Haplotype labels group three major haplotype groups: A1-3 = Southern Europe, B1+2 = Northern Europe, C = Pan-European.

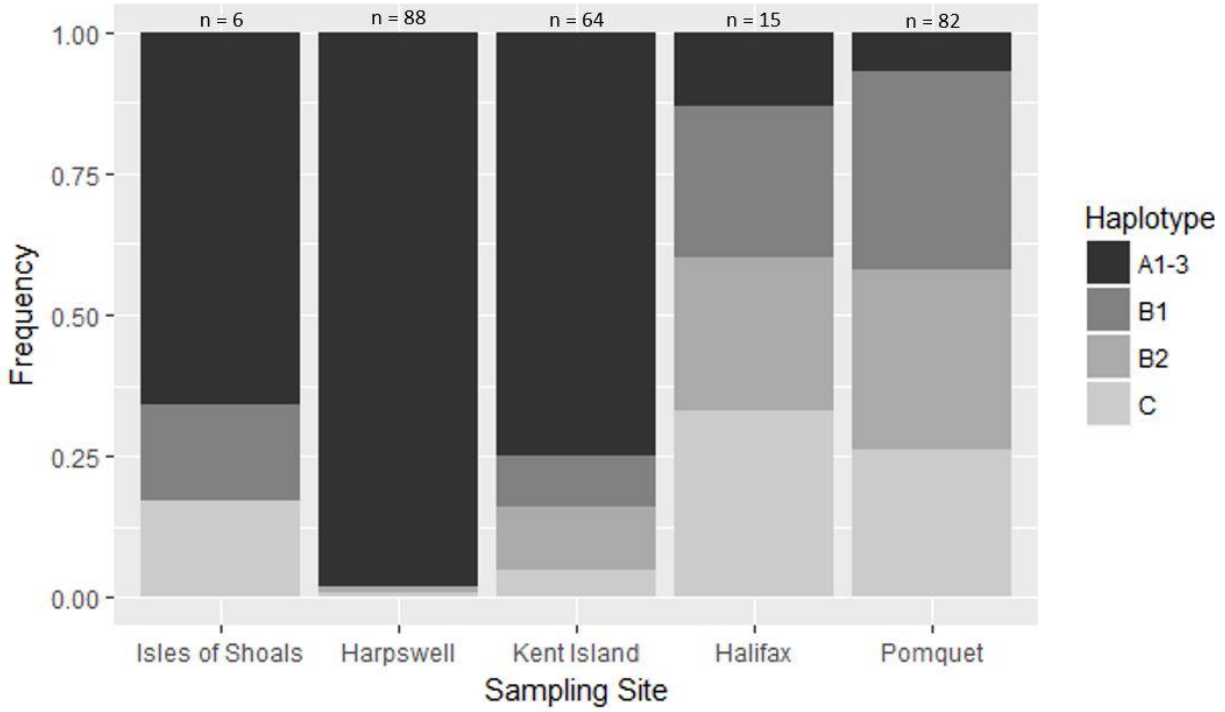


Figure 3. The frequency of the four major mitochondrial haplotype groups at five sampling sites. The sites are ordered by latitude, with the southernmost site on the left side. Number of samples is shown at the top of each bar. Haplotype letters correspond to Fig. 2.

Table 1. The results of a multiple regression logistic model of the effects of five physiological variables on righting response. Yellow shading indicates significant variables.

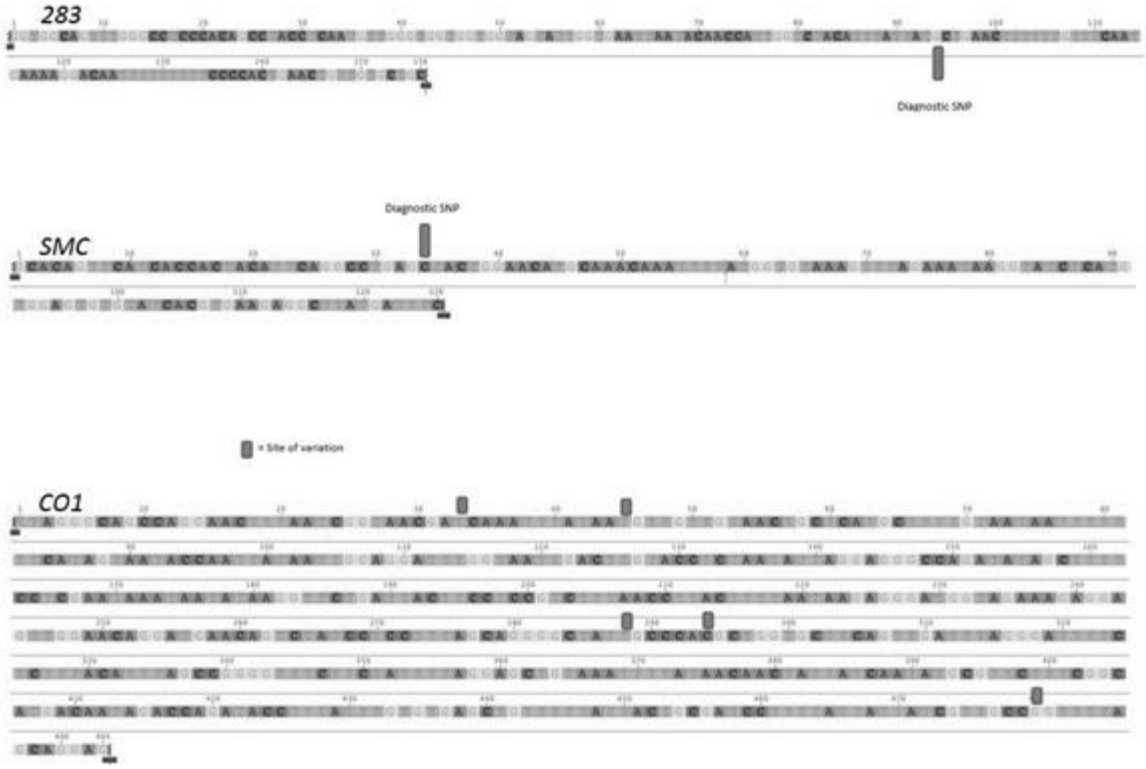
Factor	Deviance	AIC	LRT	Pr (>Chi)
Sex	195.10	201.27	2.169	0.149
Morph Type	197.27	202.96	0.144	0.701
Missing Legs	194.96	202.99	0.111	0.739
Carapace Width	195.09	203.09	0.012	0.913
Sampling site	203.12	207.12	8.015	0.0046

Table 2. The results of a multiple regression logistic model of the effects of site, sex, variance in *283* and *SMC*, and *CO1* clade on righting response. Three mt *CO1* clades were used, corresponding to the A, B, and C groups in Fig. 2.

Factor	**Deviance**	**AIC**	**LRT**	**Pr (>Chi)**
Site	156.73	164.73	0.052	0.819
Sex	160.56	164.56	3.782	0.052
283 (nuclear SNP)	156.46	164.46	0.327	0.573
SMC (nuclear SNP	156.76	164.76	0.0207	0.885
CO1 clade (A/B/C)	165.81	169.81	9.0334	0.00265

Appendix A. Primers used to sequence the three sites examined in this study, along with amplification protocol. Length is defined as the reliable read obtained in each sample. Due to low quality reads at the ends of some sequences, the ends of the sequences were not included in this length. All loci were sequenced using the forward primer.

Sequence Name	Product Size (bp)	Primers	PCR Profile (annealing temperature highlighted)
283	156	F – GCCATGAGCATTTCTTTGAGTGG R – CGCAGTGGTGGAGTTTGGAA	3 min. 95ºC; 35 cycles (1 min. at 95ºC, 1 min. at 53ºC, 1 min. at 72ºC). 5 min. at 72ºC
SMC	126	F – AGCACAGGAAGGCTGTGG R – ACGAAATCATAAGCCTCTTCACG	3 min. 94ºC; 17 cycles (30 sec. at 94ºC, 30 sec. at 65ºC [-1ºC per cycle], 1.5 min. at 72ºC). 20 cycles (30 sec. at 94ºC, 30 sec. at 48ºC. 1.5 min. at 72ºC). 10 min. at 72ºC.
CO1	493	F – GCTTGAGCTGGCATAGTAGG R – GAATGAGGTGTTTAGATTTCG	3 min. 94ºC; 30 cycles (1 min. at 94ºC, 1 min. at 50ºC, 1 min. at 72ºC). 5 min. at 72ºC.



Appendix B. The sequence of the nuclear and mitochondrial sites examined in this study. Only the reliably high-quality portions of these sites (shown here) were examined for variance. Grey boxes mark sites of variance.