

1-16-2008

Collaborative Research: Extreme Discordance between Allozyme and Non-allozyme Introgression in Baltic Mussels. Selection on Allozymes?

Paul D. Rawson

Principal Investigator; University of Maine, Orono, prawson@maine.edu

Follow this and additional works at: https://digitalcommons.library.umaine.edu/orsp_reports



Part of the [Population Biology Commons](#)

Recommended Citation

Rawson, Paul D., "Collaborative Research: Extreme Discordance between Allozyme and Non-allozyme Introgression in Baltic Mussels. Selection on Allozymes?" (2008). *University of Maine Office of Research and Sponsored Programs: Grant Reports*. 159.
https://digitalcommons.library.umaine.edu/orsp_reports/159

This Open-Access Report is brought to you for free and open access by DigitalCommons@UMaine. It has been accepted for inclusion in University of Maine Office of Research and Sponsored Programs: Grant Reports by an authorized administrator of DigitalCommons@UMaine. For more information, please contact um.library.technical.services@maine.edu.

Final Report for Period: 08/2006 - 07/2007**Submitted on:** 01/16/2008**Principal Investigator:** Rawson, Paul D.**Award ID:** 0315891**Organization:** University of Maine**Title:**

Collaborative Research: Extreme Discordance between Allozyme and Non-allozyme Introgression in Baltic Mussels. Selection on Allozymes?

Project Participants**Senior Personnel****Name:** Rawson, Paul**Worked for more than 160 Hours:** Yes**Contribution to Project:**

As a PI on this collaborative, multi-institution project, Paul Rawson was responsible for overseeing the work conducted at the University the Maine. This work included the isolation and sequencing of coding sequence for three allozyme genes (Gpi, Mpi, and Lap), the analysis of the resulting data and the attempts to construct informative families for determining linkage among marker loci.

Post-doc**Name:** Harper, Fiona**Worked for more than 160 Hours:** Yes**Contribution to Project:**

Dr. Harper held a post-doctoral research appointment in the lab of Dr. Paul Rawson and during year 2 of this project she assisted with the isolation of RNA, construction of cDNA and initial sequencing of Glucose phosphate isomerase (Gpi) alleles from mussels sampled from allopatric populations of *M. edulis* and *M. trossulus*. Dr. Harper's work was instrumental in helping streamline the sequencing protocols that were eventually used to obtain sequence information from Baltic mussels. Dr. Harper also helped to supervise Beth Ann Pomerlau's undergraduate research project. Dr. Harper's salary was covered by NSF project #DEB0133349; she did not receive direct funding from DEB0315891.

Graduate Student**Name:** Caponera, Jay**Worked for more than 160 Hours:** Yes**Contribution to Project:**

Jay Caponera was a graduate research assistant in the lab of Dr. Paul Rawson. For his M.S. thesis, Jay was responsible for isolating full-length coding sequence for the mannose phosphate isomerase (MPI) locus in *M. edulis*. As part of his work, he determined that there are two independent Mpi genes in *Mytilus* and his thesis focused on the patterns of variation and the impact of selection at these two genes in populations of *M. edulis*. Jay also contributed substantially to the isolation and analysis of Gpi and Lap.

Undergraduate Student**Name:** Pomerlau, Beth**Worked for more than 160 Hours:** Yes**Contribution to Project:**

Beth Ann was an undergraduate student at the University of Maine who completed her senior thesis research project in the lab of Dr. Paul Rawson. Beth Ann's thesis investigated the degree of gamete compatibility between *M. edulis* and *M. trossulus* and used a proteomics approach to isolate of genes regulating gamete recognition in these two species. Beth Ann also provided general logistical support to this proposal and received support from this award.

Name: McGowen, Afton**Worked for more than 160 Hours:** Yes**Contribution to Project:**

Afton worked as an intern in the lab of Dr. Paul Rawson starting with the summer between her junior and senior year in high

school. During the summer of 2005 Afton assisted in the development and testing of scnDNA markers for differentiating *M. trossulus* and *M. edulis*.

Name: Kranich, Lisa

Worked for more than 160 Hours: No

Contribution to Project:

Lisa Kranich provided technical assistance to this project. She assisted with the isolation of RNA, construction of cDNA and the amplification, cloning and sequencing of Gpi PCR products for targeted individuals.

Name: Gettings, Rachel

Worked for more than 160 Hours: No

Contribution to Project:

Rachel Gettings prepared a senior capstone paper on the unique environmental challenges posed by the chronic low salinities in the Baltic Sea and the physiological responses exhibited by two major inhabitants of the Baltic, *Mytilus* and *Fucus*. Her capstone has provided the foundation for our efforts to develop materials for a webpage on the evolution and ecology of Baltic Sea fauna and flora.

Technician, Programmer

Name: Feindel, Scott

Worked for more than 160 Hours: No

Contribution to Project:

Scott Feindel is the hatchery manager for the oyster broodstock development program at the University of Maine's Darling Marine Center. Scott provided assistance with algal and larval mussel culture as part of our efforts to construct pair matings among individual mussels during year 2 of the project.

Other Participant

Research Experience for Undergraduates

Organizational Partners

DUKE UNIVERSITY

This was a collaborative research project with Drs. Cliff Cunningham and Cynthia Riginos of Duke University. We have worked closely with Drs. Cunningham and Riginos to complete the proposed work.

University of Queensland, Australia

During the period covered by this award, Dr. Cynthia Riginos accepted a faculty position at the University of Queensland. After her move from Duke University we have continued to work closely to complete the proposed research.

Other Collaborators or Contacts

I have established contact with Dr. Phil Yund (University of New England) and Dr. Michael McCartney (University of North Carolina, Wilmington) regarding collaborative research projects investigating the role of gamete compatibility, intrinsic selection and extrinsic selection in structuring *Mytilus* hybrid zones.

Activities and Findings

Research and Education Activities: (See PDF version submitted by PI at the end of the report)

Findings: (See PDF version submitted by PI at the end of the report)

Training and Development:

Two high school interns and three undergraduate students at the University of Maine received training under this award. Brendan Horton interned in the Rawson lab for one summer (2004) during which time he received hands-on training in molecular biological methodology. As a result of his experience Brendan has decided to pursue undergraduate studies in molecular biology and its application at Boston University and has participated in two NSF-sponsored REU programs since entering college. Afton McGowen worked as an intern in the Rawson lab nearly every summer from her senior year in high school till she graduated from the University of New Hampshire. Her involvement in this project included assisting in the development and testing of scnDNA markers for differentiating *M. trossulus* and *M. edulis*. Afton has since entered a Ph.D. program in biomedical sciences at Duke University. Lisa Kranich, an undergraduate at the University of Maine, provided assistance with the isolation of RNA, construction of cDNA and the amplification, cloning and sequencing of Gpi PCR products for targeted individuals. She received training in a variety of molecular biological techniques that she now applies in her job with the University of Maine Wildlife Forensics Laboratory. Rachel Gettings prepared a senior capstone paper on the unique environmental challenges posed by the chronic low salinities in the Baltic Sea and the physiological responses exhibited by *Mytilus* and *Fucus*, two major inhabitants of the Baltic, as part of her undergraduate studies at the University of Maine. Rachel received training in library research methods to help her complete her project. Beth Ann Pomerlau investigated the evolution of gamete recognition proteins in *Mytilus* for her senior thesis research at the University of Maine. Beth Ann received training in molecular biology, population genetics and animal husbandry and is currently applying to graduate programs in the biological sciences.

One of these students (Pomerlau) delivered an oral presentation on the work she completed at an international meeting. Two students (Pomerlau and Gettings) are co-authors on manuscripts and webpages stemming from their involvement in this project. We expect that these products will be submitted for peer review by the end of December, 2007.

The work conducted under this award provided extensive training and experience in experimental design, molecular biology, and evolutionary analysis to Jay Caponera, a graduate student in the School of Marine Sciences at the University of Maine. Jay received his M.S. degree in August of 2006 and has since taken a position as a Forensic Scientist with the New York State Police.

Dr. Harper served on a post-doctoral research appointment in the Rawson lab. Dr. Harper assisted with the isolation of RNA, construction of cDNA and initial sequencing of Glucose phosphate isomerase (Gpi) alleles from mussels sampled from allopatric populations of *M. edulis* and *M. trossulus*. Her work was instrumental in helping streamline the sequencing protocols that we used to obtain sequence information from Baltic mussels. She gained valuable research and mentoring experience while at the University of Maine and has since taken a faculty position at Rollins College.

Outreach Activities:

As part of the proposed work we indicated that we would engage an undergraduate student at the University of Maine to begin developing website materials on the evolution and ecology of Baltic Sea flora and fauna. In completing her capstone project at the University of Maine, Rachel Gettings generated much of the basic text and graphical materials to populate this website, which is currently under construction. At the same time, we have been working with middle and high school teachers in Maine to develop hands-on classroom-based lessons that would be coordinated with and link to the website. These efforts are on-going and we hope to have draft lesson plans and at least a working mock-up of the website ready by the end of the current school year (spring 2008).

Journal Publications

Rawson, P.D., "Non-homologous Recombination between the Large Unassigned Region of the Male and Female Mitochondrial Genomes in the Mussel, *Mytilus trossulus*.", *Journal of Molecular Evolution*, p. 717, vol. 61, (2005). Published, 10.1007/s00239-004-0035-6

Caponera, J and P. Rawson, "Highly Divergent Duplicate Mannose-6-Phosphate Isomerase (Mpi) Genes in the Blue Mussel," *Marine Genomics*, p. , vol. , (2007). Submitted,

Rawson, P. and F. Harper, "Historical Patterns of Trans-Arctic Gene Flow in the Blue Mussel, *Mytilus trossulus*.", *Molecular Ecology*, p. , vol. , (2008). reviewed, in revision,

Books or Other One-time Publications

Caponera, J, "Evolutionary Analysis of Duplicate Mannose-6-phosphate Isomerase (MPI) Loci in the Blue Mussel, *Mytilus edulis*", (2006). Thesis, Published

Bibliography: M.S. Thesis, School of Marine Sciences, University of Maine

Web/Internet Site

Other Specific Products

Product Type:

Data or databases

Product Description:

As part of the research under this award we have generated 13 new scnDNA markers for differentiating species of blue mussel. These new markers increase by nearly 3-fold the number of markers tested and available for population genetic studies in *Mytilus*.

Sharing Information:

The protocols associated with these markers will be disseminated in three ways: 1) through publications, 2) through an web-based EST database maintained by C. Riginos, and 3) through a methods page attached to P. Rawson's lab webpage.

Product Type:

Data or databases

Product Description:

We generated 24 new partial or full-length coding sequence expressed sequence tags (ESTs) from *Mytilus* during our work on this project. This includes sequences for the genes encoding leucine aminopeptidase (Lap), glucose phosphate isomerase (Gpi), mannose phosphate isomerase (Mpi), Phosphoglucomutase (Pgm), sodium/potassium ATPase (Na/K ATPase), and arginyl aminopeptidase.

Sharing Information:

We have already submitted and released the sequences for 20 of these ESTs through GenBank; the remaining four will be released upon publication.

Product Type:

Data or databases

Product Description:

We have generated the following gene-specific sequence databases for *Mytilus*:

Glucose phosphate isomerase (Gpi:164 sequences), Mannose phosphate isomerase (Mpi: 64 sequences), Leucine aminopeptidase (Lap:50 sequences), Phosphoglucomutase (Pgm:48 sequences).

Sharing Information:

Most of these sequences will be released upon publication or are being prepared for submission to GenBank.

Contributions

Contributions within Discipline:

Currently, there are three recognized species of blue mussel, *M. edulis*, *M. trossulus*, and *M. galloprovincialis*. Worldwide there are several blue mussel hybrid zones that form wherever the ranges of any two blue mussel species overlap. These hybrid zones provide a natural laboratory for studying how differentiated genomes interact in the face of gene flow. Studies of hybridization in *Mytilus* however, have been hampered by a paucity of genetic markers. Prior to our work, most studies employed 2 to 4 PCR-based scnDNA markers and from 2-8 allozyme markers. Our work has two several important implications for the study of blue mussel hybrid zones. First, we have shown that co-migration of allozyme alleles does not necessarily indicate recent common ancestry of those alleles which can seriously complicate the interpretation of data generated by allozyme electrophoresis. Second we have developed 13 new scnDNA markers, including markers targeting the introns of allozyme genes that can be used in blue mussel hybrid zone studies. Although we have not tested these markers in *M. galloprovincialis*, they are diagnostic for *M. edulis* and *M. trossulus* (preliminary work suggests that most will distinguish between *M. trossulus* and *M. galloprovincialis*). Thus, our work has increased the number of available markers nearly 3-fold which will vastly improve our ability to detect differential introgression and selection across a hybrid zone.

The results of our study will be of interest to evolutionary biologists researchers studying the role of hybridization and introgression in adaptation to extreme environments. Hybridization can be a positive force in that it can generate new genetic combinations and additional phenotypic variation not present in the parental taxa but that is available for selection to act on. On the other hand, continued gene flow in the face of selection can be maladaptive in that it results in repeated re-introduction of alleles that are less fit. We have shown that mussels in the Baltic Sea are of hybrid origin and our on-going analyses will help us elucidate whether gene flow in this system is adaptive or maladaptive. Our data provide definitive evidence that the Baltic mussel system is not a standard hybrid zone involving two parental taxa and limited gene flow between species; rather the Baltic mussel population appears to be a hybrid swarm with extensive gene flow and complete local elimination of one of the parental species (*M. trossulus*).

Finally our results will provide additional information regarding the historical and contemporary patterns of colonization by a key member of rocky intertidal communities in the northwest Atlantic. Evolutionary biologists and ecologists have long assumed that near-shore intertidal communities went locally extinct during periods of Pleistocene glaciation. However, new evidence, including evidence from the research conducted under this award, suggests that benthic marine species may have persisted in the north Atlantic, ostensibly in glacial refugia. Thus, our work contributes to the debate regarding the persistence of near-shore communities during large scale climatic changes.

Contributions to Other Disciplines:

The results of our work will likely contribute to the fields of medicine and paleoceanography.

As part of the work conducted under this award we have initiated studies investigating the evolution of the sperm acrosomal protein lysin and vitelline envelope proteins in *Mytilus*.

Fertilization is a complex, multistep process in which haploid sperm and eggs meet and fuse in order to form a diploid zygote. Gamete recognition is the coordinated communication between sperm and eggs and is governed by the complex interaction of complementary molecules expressed on the surface of both types of gametes. Gamete recognition is generally highly species-specific so that fertilization between sperm and eggs from the same species is facilitated at the same time that fertilization between sperm and eggs from different species is limited or blocked. In addition to investigating how the evolution of gamete recognition can lead to speciation, our work on lysin and vitelline envelope proteins enhances our understanding of how gamete recognition changes as variation accumulates in sperm and egg coat proteins and may lead to a greater understanding of the molecular basis of infertility.

Our study of the historical and contemporary patterns of colonization of the North Atlantic by *Mytilus* benefits tremendously from insights on the patterns of Pleistocene glaciation provided by paleoceanographic studies. At the same time, however, the patterns of genetic differentiation that we have documented can help refine paleoceanographic models by providing information on the likelihood and extent of glacial refugia.

Contributions to Human Resource Development:

This project has actively involved and provided training and hands-on research experience to women and minorities. A total of six women (one minority) and three minority men had significant involvement in this project. Nearly all of these participants have continued on with advanced studies or careers in related, scientific fields.

Contributions to Resources for Research and Education:

Contributions Beyond Science and Engineering:

Categories for which nothing is reported:

Any Web/Internet Site

Contributions: To Any Resources for Research and Education

Contributions: To Any Beyond Science and Engineering

Major Findings

The research we conducted under this award tested three central hypotheses regarding the disparate patterns of gene flow for allozyme and non-allozyme markers between North and Baltic Sea populations of blue mussels in the genus *Mytilus*.

Hypothesis 1 – *M. edulis* alleles (both allozyme and non-allozyme) have introgressed into a relict *M. trossulus* population as neutral alleles. Selection is not required to explain the observed allele frequencies for Baltic mussels.

Hypothesis 2 – The apparent absence of *M. trossulus* mtDNA in any Baltic mussel population raises the possibility that Baltic mussels are pure *M. edulis*. This would mean that there has been convergent evolution for electrophoretic mobility of particular allozymes, such that the Baltic alleles comigrate with *M. trossulus* alleles but are descended from *M. edulis* alleles.

Hypothesis 3 – Selection maintains *M. trossulus* allozyme alleles (either through direct selection or selection on linked loci) in Baltic populations despite the asymmetric introgression of *M. edulis* alleles; selection may be acting to retain coadapted *M. trossulus* gene complexes or to retain *M. trossulus* alleles adapted to extreme Baltic environmental conditions.

To date, we have completed work on two approaches seeking to differentiate between these hypotheses. In the first approach, we sampled additional non-allozyme, single copy nuclear DNA (scnDNA) markers to determine whether the disparate patterns of introgression were simply an artefact of sampling to few scnDNA markers. In total, we developed 13 new scnDNA markers that are diagnostic for *M. edulis* and *M. trossulus*. Analyses employing a subset of these scnDNA markers indicate that allopatric populations of *M. trossulus* and *M. edulis* are nearly fixed for alternate alleles at each marker. The patterns of variation for these markers in North and Baltic Sea mussel populations indicate that gene flow does not appear to be impeded for most of these scnDNA markers across the North Sea-Baltic Sea hybrid zone. This work has confirmed the discordant patterns of introgression for allozyme and scnDNA markers. Further, the patterns of variation for each marker type suggest that selection prevents gene flow for at least some allozyme loci, a conclusion that is also supported by Bayesian coalescent estimates of admixture.

The second approach we have employed was to examine sequence variation for allozyme and non-allozyme loci to determine whether Baltic Sea populations are of hybrid origin and whether allozyme alleles with *M. trossulus*-like electrophoretic signatures nest within clades of alleles from allopatric *M. trossulus* populations. We generated a large database of sequences for four allozyme (Gpi, Lap, Pgm, and Mpi) and two non-allozyme genes (twitchin and lysin). With the single exception of twitchin, all loci confirm that Baltic mussels have haplotypes derived from both putative parental types. This observation allows us to confidently reject hypothesis 2 as an explanation for the disparate patterns of introgression between allozyme and non-allozyme markers. Convergence of allozyme electromorphs is insufficient to explain the overall pattern of asymmetric introgression. However, the complex relationships between allozyme and DNA genotypes we have documented imply that convergence to similar electrophoretic mobility can occur between historically differentiated haplotypes. This contrasts with predictions from the infinite alleles model of evolution and has serious implications for any analysis that assumes that allozyme co-migration reflects recent common ancestry of alleles. On-going analyses will help to further elucidate the effects of selection on allozyme genes.

Additional work under this award has investigated the historical patterns of gene flow between *Mytilus* populations in the Northern Hemisphere and the evolution of gamete recognition genes in *Mytilus*. Our work using a number of loci indicates that *M. edulis* has been a long-term resident on both the European and North American sides of the Atlantic Ocean despite extensive glaciation during the Pleistocene. In addition, this work indicates there may have been multiple refugial populations on the North America side of the north Atlantic during periods of Pleistocene glaciation. Separate analyses have investigated the history of *M. trossulus* in the northwest Atlantic. This work has confirmed that *M. trossulus* colonized the northwest Atlantic from source populations and evidence suggests that this colonization predates the last glacial maximum. We have found that that populations of *M. trossulus* on both coasts of North America share unique insertions within the mitochondrial D-loop region that are the result of non-homologous recombination. Recombination is a relatively rare event for the mitochondrial genome in most species but is facilitated by the presence of two highly divergent gender-specific mitochondrial lineages that are unique to some bivalve taxa. Thus, the histories of both species suggest that population survival and persistence are likely in the general vicinity of glaciers.

The evolution of gamete recognition systems is likely to play an important role in contributing to or maintaining pre-zygotic isolation between species of free-spawning marine invertebrates, like *Mytilus*. Research examining the evolution of the sperm acrosomal protein, lysin, has shown there is an excess of nonsynonymous (amino acid) substitutions between *Mytilus* species, an observation that is indicative of positive selection. Comparisons of population-specific sequence evolution as well as the distribution of major lysin alleles frequencies among allopatric and sympatric (hybrid) populations indicates that balancing selection maintains divergent lysin alleles within some species. In the first ever attempt to examine the evolution of egg proteins in mussels, we have investigated how evolutionary changes in the molecular structure of the vitelline envelope. The vitelline envelope is a matrix surrounding the egg that plays a substantial role in gamete recognition and affects intra- and interspecific patterns of fertility in mussels. We have used a proteomics approach to compare the vitelline envelope proteins from *M. edulis* females that differ in their receptivity to sperm from *M. trossulus*. We have found substantial variation in the vitelline envelope protein profiles among receptive and non-receptive *M. edulis* females. Detailed analysis of one protein from these profiles, MERP-1, reveals extensive divergence at the amino acid level and suggests that there may also be large differences in the degree of post-translational modification among *M. edulis* females with different mating phenotypes.

Research and Education Activities

One useful way to study selection in natural populations is to document disparate patterns of gene flow between differentiated parental types in a hybrid zone. Genes that fail to introgress are candidates for reducing hybrid fitness and help maintain the identity of the parental species. The collaborative research funded under this award examined an extreme case of differential gene flow between blue mussel populations (*Mytilus spp.*) in the North and Baltic Seas. Previous to our study, analysis of variation at a subset of five allozyme markers, including glucose phosphate isomerase (Gpi), mannose phosphate isomerase (Mpi), leucine aminopeptidase (Lap) and phosphoglucosmutase (Pgm), indicated that mussel populations in the Baltic Sea contain high frequencies of alleles commonly found in allopatric populations of *M. trossulus*. These same populations contain high frequencies of alleles typical of allopatric *M. edulis* populations at other allozyme loci as well as at non-allozyme, single copy nuclear DNA (scnDNA) markers and contain exclusively *M. edulis* mitochondrial DNA haplotypes. Under this award, we tested three competing hypotheses explaining the observed discordance between allozyme and scnDNA marker allele introgression.

Hypothesis 1 – *M. edulis* alleles (both allozyme and non-allozyme) have introgressed into a relict *M. trossulus* population as neutral alleles. Selection is not required to explain the observed allele frequencies for Baltic mussels.

Hypothesis 2 – The apparent absence of *M. trossulus* mtDNA in any Baltic mussel population raises the possibility that Baltic mussels are pure *M. edulis*. This would mean that there has been convergent evolution for electrophoretic mobility of particular allozymes, such that the Baltic alleles comigrate with *M. trossulus* alleles but are descended from *M. edulis* alleles.

Hypothesis 3 – Selection maintains *M. trossulus* allozyme alleles (either through direct selection or selection on linked loci) in Baltic populations despite the asymmetric introgression of *M. edulis* alleles; selection may be acting to retain coadapted *M. trossulus* gene complexes or to retain *M. trossulus* alleles adapted to extreme Baltic environmental conditions.

To differentiate between these hypotheses we proposed to: 1) estimate the patterns of introgression for a larger set of scnDNA markers, 2) determine the historical origin of allozyme alleles in Baltic populations, 3) test for selection on allozyme loci using DNA sequence-based approaches, and 4) determine the linkage groups for allozyme coding genes.

Goal 1: Estimate patterns of introgression for allozymes loci and scnDNA markers.

Hypothesis one predicts that the discordant patterns for allozyme and scnDNA markers will disappear as more scnDNA markers are sampled while hypotheses two and three predict that analysis of additional scnDNA markers will confirm the discordant patterns of variation between allozyme and scnDNA markers. A key component of the work addressing this goal has been the development of new nuclear codominant DNA markers that are diagnostic between the parental species of *M. edulis* and *M. trossulus* and then applying these markers specifically to this hybrid zone to examine patterns of gene flow. This aspect of the project has been particularly challenging for a variety of reasons. First, such an approach relies upon the successful amplification of a discrete locus from both species. Because mussels are extremely variable at the nucleotide level finding reliable priming sites that are preserved across species is often highly problematic. Unfortunately, reducing the stringency of PCR typically results in non-specific amplification for some individuals. Thus, the development of any new marker is time

consuming because it requires multiple rounds of sequencing to determine SNPs and potential priming sites, primer design, and confirmation of priming in both parental taxa.

We have found that the most successful strategy for developing markers has been to take the following approach:

1. Target exons as priming sites rather than anonymous DNA. This strategy has been aided by taking advantage of the thousands of EST sequences available in Genbank (see materials developed) and using these sequences to design primers
2. Test primers on individuals of both species. Approximately 1/3 of primer pairs will reliably amplify in both species.
3. If there appears to be length variation and greater than predicted length, we assume that an intron is being amplified and survey a greater number of individuals per species, resolve fragment lengths on agarose or acrylimide, and score length variants.
4. If there is no length variation and the PCR product size resembles that predicted based on EST sequences, we assume that a large contiguous exon has been amplified. The PCR products are sequenced for SNP discovery. Diagnostic SNPs are sometimes amenable to digestion with restriction enzymes and those that are not have been catalogued for future work using high throughput commercial (fluorescent labeling, multiplexing) systems.

The following table provides a summary of the target genes for which we have been attempting to develop scnDNA markers for *Mytilus*.

Locus	Marker type	Development stage
<i>60S</i>	Exon - Possible SNPs	Sequenced for SNP discovery
<i>40S</i>	Intron – possible SNPs	Sequenced for SNP discovery
<i>Collagen protein</i>	Exon - Possible SNPs	Sequenced for SNP discovery
<i>Creatine kinase</i>	SNP (restriction digest)	Variation being assessed between parental species
<i>Fructose biphosphate</i>	Exon - Possible SNPs	Sequenced for SNP discovery
<i>Gpi</i>	Intron length variation	Complete
<i>Hsc71</i>	Intron length variation	Variation being assessed between parental species
<i>Laminin receptor</i>	Exon - Possible SNPs	Sequenced for SNP discovery
<i>Leucine Aminopeptidase</i>	SNP (restriction digest)	Variation being assessed between parental species
<i>Lysin (Mlx)</i>	SNP (restriction digest)	Complete – Riginos et al. 2006
<i>Mannanase</i>	Intron length variation	Variation being assessed between parental species
<i>Mpi</i>	SNP (restriction digest)	Complete
<i>Mytilus edulis Reproductive Protein-1 (MERP-1)</i>	SNP (restriction digest)	Variation being assessed between parental species
<i>p53</i>	Intron length variation	Variation being assessed between parental species
<i>Pgm</i>	Intron length variation	Complete
<i>Plla</i>	SNP (restriction digest)	Complete
<i>Superoxide dismutase (SOD)</i>	Intron length variation	Complete
<i>Twitchin</i>	SNP (restriction digest)	Variation being assessed between parental species

We originally proposed to develop ten new scnDNA markers. To date, we have completed the analysis of four new scnDNA markers, developed and are in the process of analyzing variation at nine other scnDNA markers, and have finished the SNP discovery phase for four additional loci, thereby exceeding our original goal.

Introgression patterns

Whether or not patterns of gene flow and introgression differ between allozyme and non-allozyme loci is a central focus of this funded work. However, the degree to which loci will be differentiated between allopatric species results from the stochastic process whereby shared ancestral polymorphisms are lost from descendant species. Therefore, it is not sufficient to compare genetic differentiation amongst loci across the hybrid zone; these data need to be considered in light of locus-specific differentiation between species (see Riginos & Cunningham 2005 for an extended discussion on this point).

Between *M. edulis* and *M. trossulus* there are a range of genetic distances amongst allozyme loci, reflected by the range of values along the X axis in Figure 1 (with traditionally “species diagnostic” loci having delta p approaching one on the X axis). The genetic distances between North Sea and Baltic Sea populations roughly match the between species pattern, so that individual points fall approximately on the one-to-one line. In contrast, scnDNA based markers show nearly fixed differences between species (clustering towards delta p values of one, perhaps a consequence of ascertainment bias favoring fixed differences between species). The important difference among the types of loci, however, is based on the observation that these scnDNA markers that are nearly fixed between species show a range of genetic distances across the hybrid zone. Therefore, gene flow does not seem to be impeded for many scnDNA markers. These simple patterns are consistent with selection preventing gene flow for at least some allozyme loci (hypothesis 3) and are also supported by Bayesian coalescent estimates of admixture (see Riginos & Cunningham 2005).

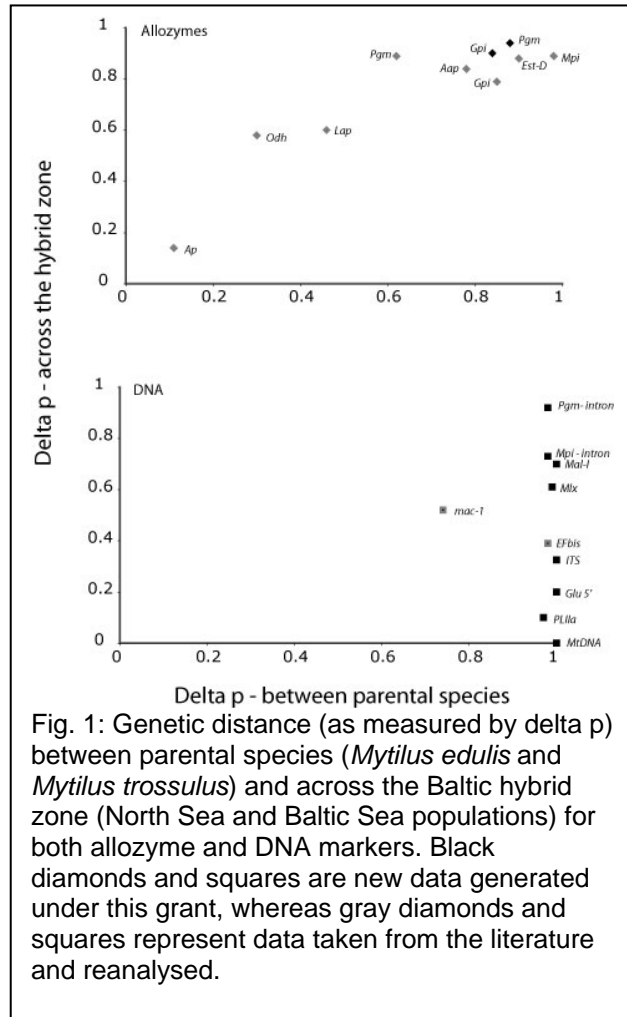


Fig. 1: Genetic distance (as measured by delta p) between parental species (*Mytilus edulis* and *Mytilus trossulus*) and across the Baltic hybrid zone (North Sea and Baltic Sea populations) for both allozyme and DNA markers. Black diamonds and squares are new data generated under this grant, whereas gray diamonds and squares represent data taken from the literature and reanalysed.

Goal 2: Determine the historical origins of Baltic alleles using DNA sequences of allozyme coding loci and scnDNA markers.

Comparison of allozyme and scnDNA sequences can be used to determine whether Baltic Sea populations are of hybrid origin and whether allozyme alleles with *M. trossulus*-like electrophoretic signatures nest within clades of alleles from allopatric *M. trossulus* populations. Hypothesis one predicts that sequences for allozyme and scnDNA genes will nest within both *M. trossulus* and *M. edulis* clades. Hypothesis 2 predicts that sequences of Baltic allozyme and scnDNA alleles should share a common ancestor with *M. edulis* sequences while hypotheses 1 and 3 predict that sequences of scnDNA from Baltic populations will fall into two groups, nesting with *M. edulis* and *M. trossulus* allopatric sequences, respectively, while sequences for Baltic Gpi, Pgm, and Mpi alleles will nest within clades from pure *M. trossulus* populations.

Analysis of sequences for Gpi, Pgm, and Mpi required that we first develop the methods for the amplification of full-length coding sequence for each gene. Using a combination of degenerate primer PCR, 5' and 3' RACE, and GeneWalking methodologies we successfully isolated full-length sequence for all three genes. We also obtained full-length coding sequence for a fourth allozyme locus, leucine aminopeptidase. This locus was not part of our originally proposed research because there are a large number of similar aminopeptidase genes and we felt that the likelihood we could isolate Lap by degenerate primer PCR was low. However, we fortuitously obtained full-length coding sequence for a putative Lap locus while working on Mpi and have included this locus in our analysis. Based on our initial sequences we have developed primer sets for amplifying each gene and have used these primers to assay gene-specific variation among Baltic and neighboring North Sea mussel populations as well as one *M. trossulus* (Washington), and at least one allopatric *M. edulis* (Rhode Island) population.

Below we discuss the results from our analysis of sequence variation at each of these genes.

Gpi: Allozyme studies have shown that *Mytilus* populations typically harbor substantial variation at the glucose phosphate isomerase (Gpi) locus. These studies have detected upwards of nine allozyme alleles per population. This variation is reflected in the DNA sequences we obtained for Gpi. We have completed full-length sequencing of over 164 Gpi haplotypes from five mussel populations. Among these sequences we have found evidence for several (n=8) large in-frame deletions as well as numerous (n=33) single base deletions. Among the full-length coding sequences, nucleotide diversity within populations ranged from 1.3 to 2.0% with the lowest diversity observed among Baltic Sea sequences and the highest diversity among Pacific *M. trossulus* sequences. The sequences for *M. edulis* and *M. trossulus* were, on average, 3.1% divergent. Phylogenetic analysis suggests there are two major clades among the Gpi sequences (Fig. 2). One clade



Fig. 2. Neighbor-joining tree of Gpi coding alleles.

contains predominantly *M. trossulus* and Baltic Sea sequences while the second clade contains mostly *M. edulis* and North Sea sequences. However, there is low statistical support for these clades and reciprocal monophyly between *M. edulis* and *M. trossulus* is incomplete. Even so, the Baltic Sea sequences have predominantly *M. trossulus* historical affinities. In some cases, however, the Baltic Sea sequences group with North Sea sequences. Although these latter groupings could indicate an *M. edulis* origin for some Baltic Gpi sequences, they are also consistent with gene flow out of the Baltic and into neighboring North Sea populations. We are continuing to proof read additional Gpi sequences. However, our initial analysis indicates that nucleotide variation at Gpi supports hypothesis three.

Pgm: Although Pgm has fewer allozymes alleles than Gpi (six versus nine), nucleotide diversity within Pgm is also very high (1.2-1.3% within species). Whereas our reference populations of *M. edulis* (RI) and *M. trossulus* (WA) share alleles (100, 106, and 111), at the nucleotide level we find many fixed differences between species (average pairwise divergence = 4.6%), accompanied by reciprocal monophyly (Fig. 3). Moreover, we have direct evidence for convergence of electrophoretic mobility as all four haplotypes from two *M. edulis* individuals with Pgm^{111/100} genotypes (where Pgm¹¹¹ is typical of *M. trossulus*) cluster with other *M. edulis* haplotypes and do not share any amino acid changing substitutions with *M. trossulus*-derived Pgm¹¹¹ alleles. In the hybrid zone, DNA sequencing has confirmed that North Sea populations consist of *M. edulis* derived alleles (Fig. 3). Baltic individuals, however, can have alleles with either *M. edulis* or *M. trossulus* historical affinities (Fig. 3). Thus, we can confirm that Baltic individuals indeed have hybrid or mixed ancestry as implied by frequency-based markers.

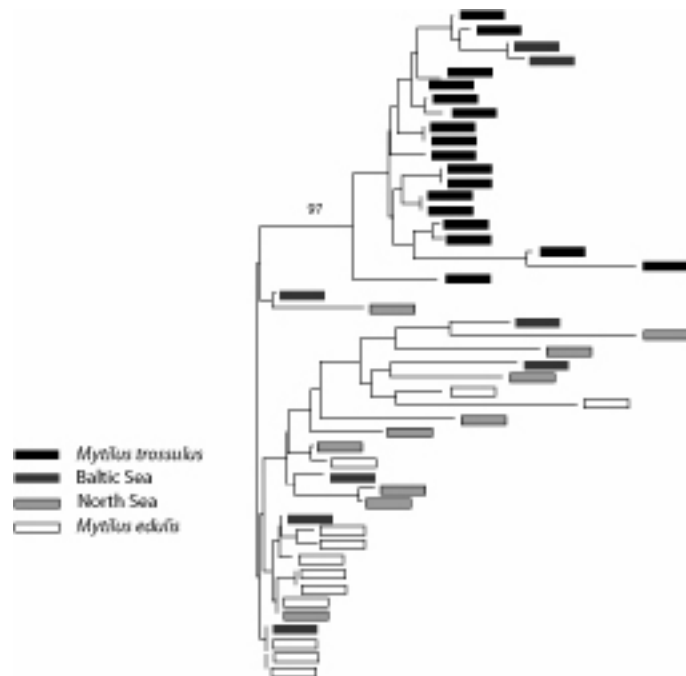


Fig. 3. Neighbor-joining tree of Pgm haplotypes

Mpi – To investigate the relative importance of selection in structuring inter- and intraspecific nucleotide variation for Mpi in marine mussels belonging to the genus *Mytilus*, we attempted to isolate and characterize full-length coding sequence for MPI from *M. edulis* and *M. trossulus*. As a result of this work, we detected duplicate Mpi genes, Mpi-A and Mpi-B, in *Mytilus*. The amino acid composition of Mpi-A and Mpi-B consisting of 438 and 427 residues, respectively, was found to be 52% identical and 71% similar and the estimated divergence time between the genes (170 MYA) substantially predates the evolution of modern *Mytilus* spp. A real-time quantitative RT-PCR (rtqRT-PCR) assay revealed significant changes in Mpi-A and Mpi-B transcript levels between tissue types, with Mpi-A showing strongest expression in mantle tissue while Mpi-B was preferentially expressed in gill and hepatopancreas tissues. Our observations are consistent with the hypothesis that different functional roles have evolved for these two Mpi genes subsequent to gene duplication. Previous studies examining allozyme level variation at Mpi in mussels have generally reported variation at only a single Mpi locus. However, many of

these studies also indicate there are *M. trossulus* and *M. edulis* species-specific alleles at Mpi. We have no direct evidence for which of the two loci we isolated is the one commonly scored by allozyme electrophoresis. However, of the two Mpi loci we have isolated, initial analysis of sequences for Mpi-B suggested that sequences from allopatric populations of *M. edulis* and *M. trossulus* were reciprocally monophyletic. Thus, we felt that Mpi-B was most likely the locus assayed by allozyme studies. Based on this assumption, we collected 65 additional Mpi-B sequences from four populations of mussels including Baltic and North Sea populations. Nucleotide diversity was an order of magnitude lower for Mpi than for either Gpi or Pgm. Within population level of diversity ranged from 0.2% to 0.9%. In this larger dataset we failed to observe any significant population-level differences in Mpi sequences (divergence ranged from 0.3 to 0.8%). Thus, these sequences provide no resolution with respect to our three main hypotheses (data not shown).

Lap: We have collected approximately 50 full-length sequences for an *Lap* gene from 24 individuals belonging to four different mussel populations. Estimates of nucleotide diversity were similar to those obtained from Gpi and Pgm. The lowest sequence diversity within a population was observed for the Baltic sequences (1.5%) while the highest was observed among the Pacific *M. trossulus* sequences (3.2%). The sequences from allopatric populations of *M. edulis* and *M. trossulus* were nearly 4.0% divergent but are sequences from the two species are not reciprocally monophyletic. Similar to what we observed for Gpi, North Sea *Lap* sequences generally have closest affinity with *M. edulis* sequences while Baltic Sea *Lap* sequences are most closely related to *M. trossulus* sequences. One notable exception is a North Sea sequence that is part of a statistically supported clade with several Baltic Sea sequences, a relationship best explained by Baltic to North Sea gene flow (Fig. 4). Preliminary analysis of our *Lap* sequences suggests that variation at this locus supports hypothesis 3.

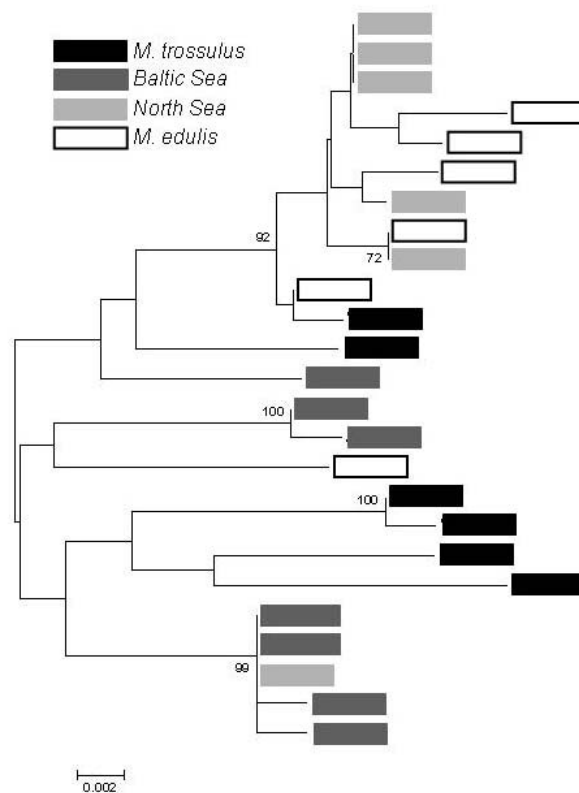


Fig. 4. Neighbor-joining tree for 25 *Lap* haplotypes

Efforts are still on-going to understand the relationships between allozyme alleles and nucleotide evolution for Gpi, Pgm and *Lap*.

Non-allozyme loci

We have sequenced individuals from all reference populations for two non-allozyme loci, twitchin and lysin. For lysin, there are many fixed differences between *M. edulis* and *M. trossulus*. Baltic lysin alleles cluster with both groups, consistent with a history of admixture at this locus. For twitchin also there are many fixed differences between the parental species. In

sharp contrast to lysin, however, the Baltic haplotypes (4 sequenced to date) are all clearly derived from *M. trossulus*. This is the only scnDNA locus so far to indicate fixation of *M. trossulus* alleles in the Baltic; increasing the sample sizes for estimating frequencies of parental alleles via SNP genotyping will increase (or decrease) our confidence in this unusual result.

Overall conclusions from sequence analyses

With the single exception of twitchin, all loci confirm that Baltic mussels have haplotypes derived from both putative parental types. This observation allows us to confidently reject hypothesis 2 as an explanation for observed patterns. Whereas convergence of allozyme electromorphs is insufficient to explain the overall pattern of asymmetric introgression, complex relationships between allozyme and DNA genotypes imply that convergence to similar electrophoretic mobility can occur between historically differentiated haplotypes (with Pgm being the clearest example of this). This apparent violation of the infinite alleles model could have serious consequences for any analysis that assumes that allozyme co-migration reflects recent common ancestry of alleles. Because of small sample sizes, DNA sequencing is not adequate to distinguish between hypotheses 1 and 3.

Goal 3: Test for selection at the DNA level.

The three hypotheses we set out to test under this award make different predictions regarding the nature of selection on allozyme alleles in mussel populations within the Baltic. Hypothesis one assumes that introgression is neutral and posits no role for selection in generating the patterns of variation observed among North and Baltic Sea populations. On the other hand, hypothesis two predicts there should be evidence of convergent selection toward common electrophoretic mobility on allozyme alleles in the Baltic while hypothesis three predicts that positive selection at the allozyme loci could be evident if the alleles are involved in local adaptation.

Because we have only recently finished the actual sequencing for many of the loci in our study the selection analyses are ongoing.

Goal 4: Determine linkage groups for allozyme-coding genes.

Our fourth goal was to determine whether the loci with different patterns of introgression are independently assorting and thus whether linkage (e.g. genetic hitchhiking) explains the shared patterns of introgression among markers. To this end, we aimed to cross individuals with various hybrid genotypes and score marker segregation and linkage among their offspring. In year 1 of this project we sampled mussels from several populations in the Gulf of Maine where we had previously observed appreciable frequencies of hybrid mussels. However, we few individuals with hybrid genotypes and were not able to construct the crosses we had intended. To circumvent this problem, we began constructing F1 crosses between *M. edulis* and *M. trossulus* during year 2 of this project with the intention of mating F1 mussels and examining marker segregation and linkage among their F2 progeny during year 3 of the proposed work. Unfortunately, mortality was extremely high for the F1 generation making this approach untenable. Finally, we resampled natural populations of mussels in eastern Maine toward the end of year 2 and attempted to construct crosses between apparent F1 hybrids and individuals with more introgressed genotypes (as determined using the three nuclear markers we had available at the time). Offspring for several families did not develop normally and mortality was quite high before we could sample. For the families that survived (n=3), the parents were not

informative for more than one or two markers and thus were generally not useful for linkage analysis.

Other Activities and Findings

Pleistocene history and gene flow in Mytilus

Our understanding of contemporary events (particularly gene flow and locus-specific selection) is greatly augmented by determining the historical context. This is especially true in a hybrid zone situation, where the inherent properties of the parental species can influence hybrid dynamics. The current range distribution of *M. edulis* includes populations both in the northwest Atlantic (including our reference population in RI) and populations in Europe. There is considerable debate in the literature regarding the history of these two major regions; there are two prominent competing scenarios. The first suggests a glacial refugia in Europe for *Mytilus* followed by recent range expansion into North America while the second proposes long-term (greater than 100,000 years) residency on both coastlines. Our work using a number of loci has supported the long-term residency scenario (albeit with some cross-ocean gene flow: Riginos et al. 2004), and also raises the possibility of multiple refugial populations within North America (Riginos & Henzler, in review).

We have also examined the history of *M. trossulus* on the coast of North America. We obtained a large set of mtDNA D-loop sequences from *M. trossulus* mussels sampled from the Pacific and Atlantic coasts. This analysis (Rawson et al. in prep) confirmed that *M. trossulus* populations in the Pacific are the source of *M. trossulus* in the northwest Atlantic and indicates that colonization of the northwest Atlantic by *M. trossulus* may predate the last glacial maximum. We have also found that the D-loop of North American *M. trossulus* contains two large sequence insertions not present in the D-loop of other blue mussel species (Rawson 2005). Populations of all three species of blue mussel typically contain two highly divergent and gender-specific mitochondrial DNA lineages. Comparison of the D-loop region from the male and female lineages indicated that one of the insertions is the result of non-homologous recombination between the gender-specific lineages.

Evolution of gamete recognition genes

The evolution of gamete recognition systems are likely to play an important role in contributing to or maintaining pre-zygotic isolation between species of free-spawning marine invertebrates. Thus, the evolution of gamete recognition genes is relevant to questions related to speciation and hybridization. Moreover, the repeated observation of positive selection acting on such genes (with examples from abalone, snails, and sea urchins) underscores their potentially important evolutionary role.

We have sought to examine the evolution of such genes and gamete interactions in *Mytilus* through several different approaches. Sequence variation for the gene coding for the acrosomal sperm protein M7 lysin has been examined in particular detail. Earlier work had established an excess of nonsynonymous substitutions between *Mytilus* species indicative of positive selection (Riginos & McDonald 2003). This work has been followed up by comparing sequence evolution and major allele frequencies among allopatric and sympatric (hybrid) populations (Riginos et al. 2006). We find no differences in M7 lysin evolution either at the nucleotide level or based on allele frequencies between population types. However, we have uncovered evidence of balancing selection maintaining divergent types of alleles within some species.

Rawson et al. (2003) documented a pattern of asymmetric gamete compatibility between *M. edulis* and *M. trossulus* wherein *M. trossulus* eggs are always unreceptive to sperm from *M. edulis*, and though the eggs from many *M. edulis* females are similarly unreceptive and cannot be fertilized by sperm from *M. trossulus*, the eggs from nearly 40% of *M. edulis* females are as receptive to *M. trossulus* sperm as they are to the sperm from conspecific males. Further, we demonstrated that the receptive phenotype in *M. edulis* is clearly a property of the eggs from specific *M. edulis* females and is not dependent on which *M. trossulus* males contribute sperm to compatible heterospecific crosses. We have expanded upon this earlier study by increasing the number of intra and interspecific crosses between *M. trossulus* and *M. edulis* in order to better estimate the frequency of the *M. trossulus*-receptive mating phenotype in *M. edulis*. This additional sampling confirmed that the receptive mating phenotype occurs at a frequency of nearly 40-50%. Gamete recognition is often determined by a series of complex interactions between proteins expressed on the surface of both gamete types. Mature metazoan eggs are surrounded by a compact, organized extracellular matrix (ECM) or vitelline envelope that plays a major role in gamete recognition. Evolutionary changes in the molecular structure of this matrix can impact the ability of sperm to bind to and penetrate the matrix and reach the egg plasma membrane and ultimately can result in barriers to fertilization and speciation. We have initiated a research program exploring how intraspecific variation in egg-born proteins among marine mussels in the genus *Mytilus* impacts sperm-egg interactions and intra- and interspecific patterns of fertility. To date, our work has focused on using a proteomics approach to compare the vitelline envelope proteins from *M. edulis* females that are receptive to *M. trossulus* sperm to those that are not. We have found substantial variation in the vitelline envelope protein profiles among receptive and non-receptive *M. edulis* females. Detailed analysis of one protein, MERP-1, reveals extensive divergence at the amino acid level and suggests that there may also be large differences in the degree of post-translational modification among *M. edulis* females with different mating phenotypes.

Major presentations pertinent to this research (Rawson)

* student author

Rawson, P. D. Hybridization and Incipient Speciation in Blue Mussels. Invited seminar, Department of Biological Sciences, University of South Carolina. April 10, 2006.

Pomerlau, E. A.*, F. M. Harper, and P. D. Rawson. Isolation of Species-Specific Egg-Borne Proteins from the Blue Mussels, *Mytilus edulis* and *M. trossulus*. Oral presentation, Benthic Ecology Meeting. Quebec City, Quebec, CA. March 2006.

Harper, F. M., E. A. Pomerlau and P. D. Rawson. Species-specific Egg-borne Proteins in the blue mussels, *Mytilus edulis* and *M. trossulus*. Poster presentation, Annual Meeting of the Society for Integrative and Comparative Biology. Orlando, FL. January 2006.

Rawson, P.D. Hybridization among Blue Mussels in the genus *Mytilus*. Invited seminar, School of Marine Sciences, University of Maine. December 16, 2005.

Rawson, P. D. and F. M. Harper. Trans-Arctic Gene Flow in the Blue Mussel *Mytilus trossulus*. Poster presentation, Coordinating Research in the North Atlantic (NSF: CORONA). Roscoff, France. July 2005.

Rawson, P. D. Male Lineage D-Loop Sequence Insertion in the Female Lineage Mitochondrial DNA Control Region of the Blue Mussel *Mytilus trossulus*. Poster presentation, Annual Meeting of the Society for the Study of Evolution. Ft. Collins, CO. June 2004.

Major presentations pertinent to this research (Riginos and Cunningham)

Poster presentations (*student)

2005 *Wang, D and C. Riginos. Gene duplication and natural selection on a gamete recognition gene (Society for the Study of Evolution)

2005 Riginos, C. Gene duplication and natural selection on a gamete recognition gene (Coordinating Research Network on the North Atlantic)

Conference talks (*student)

2005 Riginos, C. Pleistocene refugia and routes of post-glacial colonization in *Mytilus edulis* (Benthic Ecology Meeting, Williamsburg, Virginia)

2004 *Shaffner, R., Rawson, P. R., Riginos, C. Linking Pgm allozyme and nucleotide variation in blue mussels (Society for the Study of Evolution)

2004 Sex-biased gene flow in blue mussels and its effect on estimating trans-Atlantic divergence time (Society for the Study of Evolution)

Invited talks (Riginos)

2006 Natural selection, hybridization, and gamete recognition in mussels (*Mytilus* spp.)
University of Maryland, Baltimore County

2006 Natural selection, hybridization, and gamete recognition in mussels (*Mytilus* spp.)
University of North Carolina, Wilmington

2006 Natural selection, hybridization, and gamete recognition in mussels (*Mytilus* spp.)
Northeastern University

2005 Natural selection, hybridization, and gamete recognition in mussels (*Mytilus* spp.)
University of Queensland

2005 Natural selection, hybridization, and gamete recognition in mussels (*Mytilus* spp.)
Syracuse University

2005 Gene flow and natural selection in blue mussels
University of Georgia

2005 Gene flow, local adaptation, and responses to climate change in marine animals
University of Queensland

2005 Causes of differential gene flow across the mussel genome
Central Florida University