Chapter 4

Population Genetics of the Soft-Shell Clam
Mya arenaria

Joseph I. Hoffman and David L. J. Vendrami

Introduction

The soft-shell clam *Mya arenaria* is a benthic marine bivalve found in coastal habitats across the northern hemisphere. Population genetic studies of this species began in the 1970s and have been motivated over the years by a variety of questions. Due to the ability of the soft-shell clam to persist over a wide range of water temperatures and salinities, this species was soon recognized as a good candidate for investigating how environmental heterogeneity can shape genetic variation. This, together with its abundance for sampling, led to *M. arenaria* being described as "an excellent organism for the study of selection, migration and genetic differentiation" (Morgan et al. 1978). The soft-shell clam is also commercially important in several regions of North America (Kennedy 2023, this volume) where it has been subject to stocking operations aimed at counteracting declines in landings caused by overfishing (Beal 2023, this volume), habitat degradation, and predation by invasive species (Strasser and Barber 2009). This has promoted an interest in characterizing patterns of *population genetic structure* and *local adaptation* (see Glossary in Appendix 1 for specialized terms indicated in italics), which can inform effective management. Finally, *M. arenaria* is believed to have gone extinct in the eastern Atlantic during the Pleistocene (Hessland 1945) and to have subsequently recolonized European shores as a consequence of repeated anthropogenic introductions from North America dating back to the 13th century (Petersen et al. 1992; Beets et al. 2003; Behrends et al. 2005). Eastern Atlantic populations of *M. arenaria* therefore provide a unique system for investigating how anthropogenic translocations followed by natural dispersal can affect patterns of population structure and genetic diversity in marine organisms.

This chapter will provide a concise overview of studies that have used *genetic markers* to investigate the population genetics of *Mya arenaria*. We will not cover the genetic basis of disease (see Seitz et al. 2023, this volume), including the well-documented role of the retro-element, *Steamer*, in transmissible neoplasia (Arriagada et al. 2014). We will describe how methodological developments over the past few decades have advanced our understanding of three major topics: genetic diversity, population structure, and historical demography. We will identify areas of consensus in the literature as well as key outstanding questions. Finally, we will conclude with the outlook for future studies of the population genetics of soft-shell clams.

Background

The soft-shell clam occurs in numerous intertidal infaunal communities across North America and Europe, where it is often one of the most important producers of benthic biomass (Strasser 1998). This species originated in the Pacific during the Miocene, extended its range to the Atlantic coast of North America, and eventually reached Europe in the late Pliocene (Figure 1; MacNeil 1965; Strauch 1972). Nonetheless, *M. arenaria* is now considered only to be native to the Atlantic coast of North America, while it is regarded as invasive in both the Pacific and in Europe, where it is believed to have disappeared during the Pliocene extinction (Hessland 1945). The species was reintroduced to the Pacific coast of North America from the Atlantic coast prior to 1874, probably as a by-product of oyster transplants for mariculture (Strasser 1998). Following natural dispersal, the distribution of *M. arenaria* along the Pacific coast now extends from California to Alaska (Hanna 1966).

The history of *Mya arenaria* in the eastern Atlantic appears to be even more complex. Fossil evidence documenting the presence of soft-shell clams in Jutland and the North Sea in the 13th century (Petersen et al. 1992; Beets et al. 2003; Behrends et al. 2005) supports the hypothesis that this species was transferred from the Atlantic coast of North America to Europe by the Vikings (Petersen et al. 1992; Essink and Oost 2019). As a consequence of both natural dispersal and further anthropogenic introductions that occurred mainly during the past century (Strasser 1998; De Noia et al. 2020), the present distribution of soft-shell clams in the eastern Atlantic stretches from the White Sea to the Bay of Biscay and includes the Baltic Sea and the coasts of the British Isles (Figure 1). The species can also be found in Iceland and the
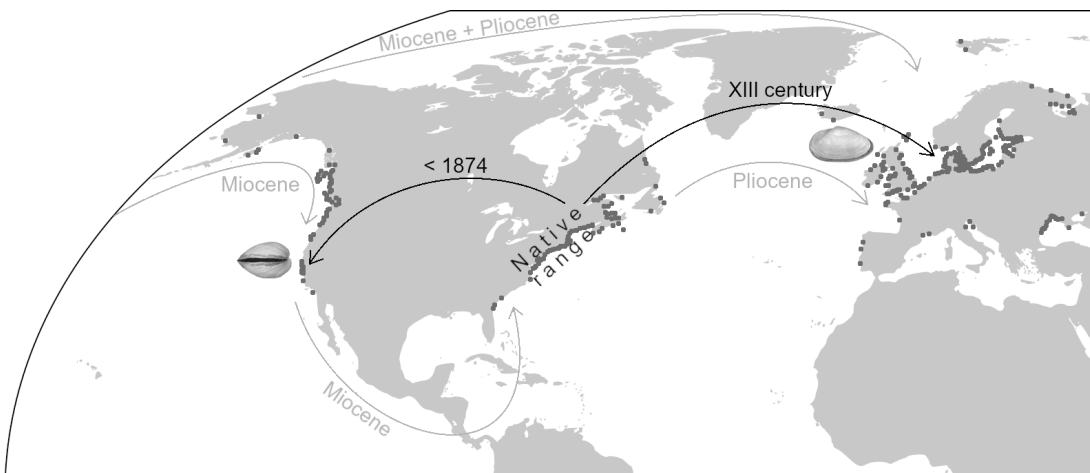

Figure 1. Map showing the contemporary distribution of the soft-shell clam *Mya arenaria* (gray points) together with hypothesized invasion pathways prior to and after the Pleistocene glaciation (gray and black arrows respectively). The species originated in the Pacific during the Miocene and expanded its range eastwards via natural dispersal. It is generally thought to have survived the Pleistocene glaciation only along the Atlantic coast of North America, which is currently regarded as the native part of this species' contemporary geographic range. Soft-shell clams are believed to have been reintroduced into Europe by the Vikings in the 13th century and were later transferred to the Pacific coast of North America prior to 1874 as a by-product of oyster mariculture.

Svalbard archipelago (Óskarsson 1961; Gulliksen et al. 1985), along the coast of the Iberian peninsula (Conde et al. 2012), and in the Mediterranean, Black, and Azov seas (Gomoiu 1981; Makarevich et al. 2000; Zenetos et al. 2005; Crocetta and Turolla 2011; Figure 1). Nevertheless, *M. arenaria* populations in the Bay of Biscay, the Mediterranean, and along the coasts of Iceland and the Svalbard archipelago are rather sparse and isolated.

As reflected by the broad latitudinal gradient along which this species can be found, the soft-shell clam is a successful colonizer that is able to occupy a broad range of water temperatures and salinities. Yet, population genetic structure is expected to be minimal as *M. arenaria* is a broadcast spawner (Snelgrove et al. 2023, this volume) characterized by high fecundity (Belding 1930; Brousseau 1978) and high dispersal potential conferred by an obligate pelagic larval stage lasting around 2–3 weeks (Strasser 1998). This leads to the expectation that soft-shell clam populations should be connected by extensive *gene flow*, resulting in little or no population structure across genome-wide distributed selectively neutral *loci* (Slatkin 1987).

Nevertheless, the genetic structure of European *M. arenaria* populations is likely to have been profoundly affected by this species' complex demographic history, which is believed to involve multiple extinction, translocation, and recolonization events. Many of these events are likely to have been accompanied by *founder effects*, which can reduce *genetic diversity* (Cross et al. 2016; Lasota et al. 2016; De Noia et al. 2020). However, many aspects of the recent history of this species remain poorly understood. For example, it has not been established with certainty that *M. arenaria* was extirpated from Europe at the beginning of the Pleistocene (Strasser and Barber 2009; Cross et al. 2016).

Genetic Diversity

The first molecular genetic studies of *Mya arenaria* were based on *allozymes*. Levinton (1973) genotyped six marine bivalve species at two allozymes and found that genetic variability at these loci decreased with increasing depth of burial, with *M. arenaria* exhibiting the lowest and second to lowest numbers of alleles, respectively, for the two loci. Levinton (1973) hypothesized that environmental variability could be driving this pattern, with higher levels of polymorphism being favored by *heterozygote advantage* or *diversifying selection* in more heterogeneous environments. According to this argument, the relatively low genetic diversity of *M. arenaria* might be a reflection of the predictability of the deep infaunal habitat where this species is typically found. However, this now seems unlikely given that *M. arenaria* juveniles are found at shallower burial depths (Zwarts and Wanink 1989) where they are exposed to heterogenous and often unfavorable conditions.

Low allozyme diversity in *M. arenaria* as compared to other marine invertebrates was subsequently confirmed by Morgan et al. (1978) and Lasota et al. (2004), who analyzed larger collections of individuals from the Atlantic coast of North America and Europe respectively. These studies reported similarly low levels of *polymorphism* on both sides of the Atlantic as well as the lack of any obvious genotype-environment associations. It was argued that these findings were consistent both with extensive gene flow and with the allozymes being selectively neutral, which is at odds with the hypothesis of Levinton (1973).

Subsequent studies based on *mitochondrial DNA* (*mtDNA*) sequencing of the cytochrome oxidase 1 (COX1) gene confirmed the presence of low levels of genetic variation in *M. arenaria* from the Atlantic and Pacific coasts of North America as well as from Europe (Strasser and Barber 2009; Cross et al. 2016; Lasota et al. 2016). Additionally, phylogenetic analysis of

the *mtDNA haplotypes* revealed a star-shaped pattern indicative of recent population expansion (Slatkin and Hudson 1991). This suggests that comparably low levels of genetic diversity might be a consequence of small effective population sizes during the Pleistocene, with subsequent population expansion having occurred into new habitats that became available as the ice retreated.

A less straightforward pattern emerges from comparisons of mitochondrial diversity between Europe and North America. Specifically, Strasser and Barber (2009) reported higher levels of *haplotype diversity* as well as the presence of *private alleles* in Europe and the Pacific coast of North America, which could be interpreted as evidence for *relict populations* having persisted in these regions during the Pleistocene. Cross et al. (2016) and Lasota et al. (2016) also found that haplotype diversity was higher in Europe than in North America, but the number of haplotypes exhibited the opposite pattern, being substantially lower in Europe. This led the authors to conclude that soft-shell clams show the hallmarks of a pronounced founder effect that may have accompanied their reintroduction into Europe.

Despite these interpretational differences, early studies of *M. arenaria* based on allozymes and mtDNA all uncovered relatively modest levels of genetic polymorphism. However, the subsequent use of *microsatellites* led to the discovery of substantially greater amounts of genetic diversity. For example, St-Onge et al. (2011) developed a panel of 8 microsatellites that each carried up to 20 alleles in a sample of 25 individuals from New Brunswick, while Krapal et al. (2012) developed another panel of 7 loci carrying up to 14 alleles in a sample of 76 individuals from Romania. Microsatellite genotyping also uncovered a latitudinal cline in genetic diversity along the Atlantic coast of North America, consistent with postglacial range expansion from a southerly refugium (St-Onge et al. 2013), as well as substantially lower levels of genetic diversity in Europe, suggestive of a founder effect that may have occurred when European shores were recolonized from North America (Cross et al. 2016). However, these patterns should be treated with caution because microsatellites are prone to *ascertainment bias*, which can inflate genetic diversity estimates and result in artifactual relationships when a panel of microsatellites is inadvertently enriched for polymorphisms that are relatively abundant in the discovery population(s).

Population Structure

The first study to focus on the population genetic structure of *M. arenaria* found no evidence for genetic differentiation among 12 localities from New England based on sequences of the internal transcribed spacer (ITS-1) ribosomal DNA (rDNA) region (Caporale et al. 1997). Similarly, Lasota et al. (2004) found that seven populations ranging from the Bay of Biscay in northern France to the Gulf of Gdansk in Poland were genetically unstructured at nine allozymes. These authors argued that this pattern could be explained by assuming selective neutrality together with rapid population expansion and extensive gene flow. However, subsequent studies suggest that these early results can be largely explained by the use of minimally informative, slowly evolving nuclear genetic markers.

Strasser and Barber (2009) were the first to deploy more rapidly evolving mtDNA sequences to investigate the population structure of *M. arenaria*. They uncovered weak but statistically significant genetic differences between a European population (Sylt in Germany) and multiple locations from the Pacific and Atlantic coasts of North America. By contrast, minimal population structure was detected within North America despite many of

the samples having been collected along a pronounced environmental cline spanning much of the eastern coast of North America.

These findings were later confirmed by more extensive mitochondrial studies, most notably by Lasota et al. (2016), who extended the coverage of sampling to include three populations from the Pacific coast of North America, multiple European populations, the Mediterranean, the Black and White seas, and Iceland. That study resolved four distinct genetic groups corresponding to the Pacific and Atlantic coasts of North America, Europe, and Iceland. Lasota et al. (2016) also confirmed the known Baltic origin of the Black Sea population (Strasser 1998) and highlighted intriguing similarities between *M. arenaria* sampled from the White Sea, the Mediterranean, and North America, which indicate a probable recent North American origin of these populations.

Subsequent advances were driven by the development of microsatellites (St-Onge et al. 2011; Krapal et al. 2012), which usually offer higher levels of polymorphism than mtDNA markers and provide a biparental nuclear perspective. St-Onge et al. (2013) used seven microsatellites to genotype populations from southern Newfoundland to Virginia together with a population from northern Europe. They not only confirmed the genetic distinctness of European *M. arenaria*, but also resolved five distinct *genetic clusters* within the native range of this species. While these clusters were not congruent with the geographic boundaries of known biogeographic marine ecoregions, evidence was found for *isolation by distance* and the presence of several putative barriers to gene flow. Consequently, extensive population structure within this region appears to result from a combination of northward postglacial range expansion, restricted dispersal and oceanographic barriers that impede the exchange of genetic material.

Arguably, the most definitive perspective on global population structure was provided by Cross et al. (2016), who genotyped populations from northern Europe and both North American coasts at mtDNA and microsatellites. They used a hierarchical clustering approach to uncover patterns of population structure over coarse and fine geographic scales. At the uppermost hierarchical scale, two genetic clusters were identified corresponding to Europe and North America respectively. Moreover, a population sampled from the Netherlands was found to carry genetic contributions from both clusters, which could potentially be indicative of gene flow from North America. At the lowermost hierarchical scale, genetic differences were identified within Europe among populations sampled from Ireland, Wales, and the Netherlands, while within North America, genetic differences were identified among populations from Canada, Maryland, and Oregon. By implication, the species as a whole appears to be extensively structured.

Finally, De Noia et al. (2020) sampled more extensively across Europe and combined microsatellite genotyping with the geometric morphometric analysis of shell outlines to investigate the relative contributions of genes and environmental variation to shell shape. They identified three genetic clusters (Figure 2) delimiting populations from the UK and northern France, the Netherlands and Germany, and the Mediterranean respectively. Moreover, their results uncovered a high degree of affinity between *M. arenaria* individuals sampled from Portugal and those found in the North Sea, providing evidence for a possible northern European origin of this population. Shell shape was unrelated to genotype, but the relative thickness of the periostracum, crossed-laminar, and complex crossed-laminar layers of the shell were found to vary with latitude, suggesting that investment in energetically less expensive shell layers may be favored at high latitudes.
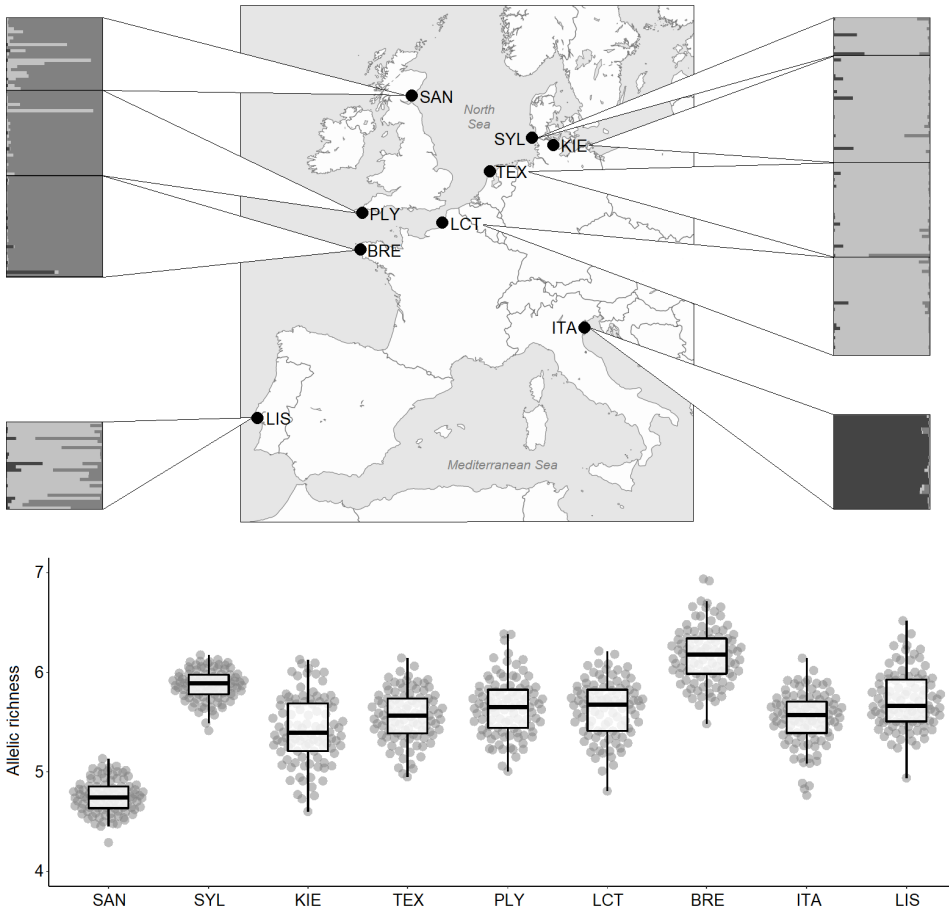
Figure 2. Patterns of population genetic structure and genetic diversity in *Mya arenaria* sampled along a European latitudinal cline inferred from microsatellite data (re-analyzed from De Noia et al. 2020). The top panel shows the results of a genetic cluster analysis, where each horizontal bar represents an individual and the proportions of the different shades of gray indicate proportional assignment to each of three inferred genetic clusters. The data were plotted separately for each sampling location as indicated on the map. Cluster analysis was implemented using the software STRUCTURE (Pritchard et al. 2000). We ran five independent runs for the number of genetic clusters (K) ranging from one to five using 10^6 Markov chain–Monte Carlo iterations after a burn in of 10^5 iterations. The most likely value of K was evaluated using maximal average value of Ln P (D), a model choice criterion that estimates the posterior probability of the data, as well as the ΔK procedure described by Evanno et al. (2005). The bar graphs were produced in R (https://www.R-project.org/). The bottom panel depicts geographical variation in genetic diversity, summarized as allelic richness. To facilitate visual comparisons among populations with different sample sizes, we plotted values derived from 1,000 random subsets of 10 individuals per population as Sinaplots, plotted using the R packages ggplot2 (Wickham 2011) and ggbeeswarm (https://github.com/eclarke/ggbeeswarm). The populations are shown in order of decreasing latitude: St. Andrews, UK (SAN); Sylt, Germany (SYL); Kiel, Germany (KIE); Balgzand, Netherlands (TEX); Plymouth, UK (PLY); Le Crotoy, France (LCT); Brest, France (BRE); Comacchio, Italy (ITA); and Lisbon, Portugal (LIS).

Demographic History

Although recent studies of *Mya arenaria* are mutually corroborative in the sense that they all uncovered pronounced genetic structure over geographic scales of tens to hundreds of kilometers (St-Onge et al. 2013; Cross et al. 2016; De Noia et al. 2020), important questions remain unanswered about the worldwide demographic history of *M. arenaria*. Studies generally agree about the deeper history of the species, with populations across the planet consistently exhibiting star-shaped mitochondrial phylogenies indicative of postglacial expansion (Strasser and Barber 2009; Cross et al. 2016; Lasota et al. 2016). Furthermore, a latitudinal cline in genetic diversity across the native part of the species range in North America is strongly suggestive of northward postglacial expansion from a southern refugium (St-Onge et al. 2013). However, the more recent history of *M. arenaria* is less clear. In particular, although it is widely accepted that soft-shell clams were introduced into northern Jutland by the Vikings, genetic support for *M. arenaria* having been completely extirpated from Europe during the Pleistocene remains equivocal.

In general, anthropogenic introductions are often accompanied by pronounced founder effects, leading to the expectation that introduced populations should possess lower genetic diversity than native ones (Holland 2000; Lallias et al. 2015). Furthermore, where natural dispersal follows anthropogenic translocations, genetic diversity should continue to decline with increasing geographic distance from the original site of translocation due to sequential *genetic bottlenecks* (Hewitt 2000; Holland 2000). Finally, introduced populations would not be expected to carry private alleles unless a sufficient amount of time has passed to allow for the accumulation of new mutations (Slatkin 1985).

Population genetic studies have provided mixed support for these expectations in soft-shell clams from the eastern Atlantic. The number of mitochondrial haplotypes is indeed lower in Europe than in North America (Cross et al. 2016; Lasota et al. 2016), but haplotype diversity tends to be higher in Europe (Strasser and Barber 2009; Cross et al. 2016; Lasota et al. 2016) and two studies detected private alleles in European populations (Strasser and Barber 2009; Cross et al. 2016). Problematically, while these private alleles may represent ancestral polymorphisms that persisted in relict European populations, an *allele discovery curve* for the Atlantic coast of North America suggests that even the largest of these studies (Lasota et al. 2016) may not have sampled sufficient numbers of individuals to capture all of the mitochondrial diversity present in this putative source population. This leaves open the possibility that these haplotypes could have been introduced from North America into Europe, where they subsequently drifted to higher frequencies on a wave of population expansion.

A similarly unclear picture emerges from microsatellite studies. Both St-Onge et al. (2013) and Cross et al. (2016) found that microsatellite diversity quantified as *allelic richness* was lower in European populations than in North America. However, such a pattern could potentially arise from ascertainment bias, as many of the loci screened in these studies were developed from a Canadian discovery population. Furthermore, the geographical distribution of European sampling localities in both studies was too restricted to test for the decline in genetic diversity with distance from Denmark that would be expected under a scenario of extirpation and subsequent reintroduction followed by range expansion. To shed further light on this question, we reanalyzed microsatellite data from nine European populations spanning a latitudinal cline from Scotland to Portugal (De Noia et al. 2020). To facilitate visual comparisons among populations with variable sample sizes, we plotted the allelic richness of

1,000 subsets of 10 randomly sampled individuals per population (Figure 2). We found no evidence for a decrease in genetic diversity either with latitude or with distance from Jutland (Spearman rank correlation tests; latitude: $\rho_s = -0.23$, $p = 0.54$; distance from Jutland: $\rho_s = -0.03$, $p = 0.95$). Instead, the population with the highest allelic richness is Brest in northern France, which is close to the site of a known glacial refugium for many marine species (see Maggs et al. 2008).

Finally, the overall picture in the eastern Atlantic appears to be complicated by multiple recent translocations. For example, *Mya arenaria* is known to have been introduced into the Black Sea in the 1960s (Leppäkoski 1994; Strasser 1998), to Greece prior to 1984 (Zenetos et al. 2004, 2005), to the west coast of France in the 1990s (Crocetta and Turolla 2011), and to the Iberian Peninsula (Conde et al. 2012) and Italy prior to the 2000s (Crocetta 2011, 2012). While in some cases these introductions may have been unintentional, involving seawater ballast (Leppäkoski 1994) or as a by-product of oyster plantings (Strasser 1998), in other instances soft-shell clams may have been deliberately introduced as illegal plantings (Strasser 1998). This complex history appears to be reflected by the available genetic data, which reveal increased genetic affinity to North American populations in samples from the Mediterranean, the White Sea (Lasota et al. 2016) and the Netherlands (Cross et al. 2016). Consequently, it would seem reasonable to describe the genetic structure of soft-shell clams in the eastern Atlantic as a patchwork of populations of variable genetic ancestry.

Conclusions and Future Perspectives

In common with many species, our understanding of the population genetics of soft-shell clams has gradually unfolded with technological developments that have provided access to increasingly informative genetic markers. Early allozyme studies primarily focused on how natural selection may have shaped patterns of genetic diversity, while subsequent studies based on mtDNA and microsatellites assumed selective neutrality and were more concerned with population structure and historical demography. While it has become clear that *Mya arenaria* populations are extensively structured, conflicting with the notion that larval dispersal should homogenize populations through gene flow, important questions remain unanswered about the role of Pleistocene extinctions, the possible presence of relict populations in the eastern Atlantic, and the effects of anthropogenic translocations on population structure and patterns of genetic diversity. Addressing these questions will require even more extensive sampling in combination with a new generation of approaches capable of genotyping many orders of magnitude more genetic markers.

Fortunately, the past decade has witnessed a revolution in DNA sequencing technologies. Long-read approaches are facilitating the assembly of draft reference genomes (Rhoads and Au 2015) and both mitochondrial (Wilson et al. 2016) and nuclear (Plachetzki et al. 2020) reference genomes are now available for *M. arenaria*. Moreover, increasingly inexpensive short-read sequencing approaches (Tautz et al. 2010; Hohenlohe et al. 2021) are allowing many thousands to millions of *single nucleotide polymorphisms* to be genotyped in multiple individuals. Among these, *reduced representation sequencing* approaches have emerged as a particularly cost-effective means of screening sufficient numbers of genetic markers to uncover subtle, fine-scale population structure (Vendrami et al. 2017; Sherman et al. 2020), to quantify variation in genetic diversity and inbreeding (Hoffman et al. 2014), to investigate patterns of admixture (Hohenlohe et al. 2013; Vendrami et al. 2020), to reconstruct recent de-

mographic histories (Rougemont et al. 2017; Vendrami et al. 2019), and to elucidate patterns of local adaptation (Cayuela et al. 2021) in marine populations. Ultimately, by teasing apart selectively neutral patterns from footprints of natural selection, these and related approaches promise to offer a unified perspective on the mechanisms shaping the genetic diversity of an ecologically and commercially important marine species.

References

Arriagada, G., M. J. Metzger, A. F. Muttray, J. Sherry, C. Reinisch, C. Street, I. W. Lipkin, and S. P. Goff. 2014. Activation of transcription and retrotransposition of a novel retroelement, *Steamer*, in neoplastic hemocytes of the mollusk *Mya arenaria*. Proceedings of the National Academy of Sciences of the United States of America 111:14175–14180.

Beal, B. F. 2023. Soft-shell clam *Mya arenaria* mariculture: history and culture techniques. Pages 495–545 *in* V. S. Kennedy and B. F. Beal, editors. The soft-shell clam *Mya arenaria*: biology, fisheries, and mariculture. American Fisheries Society, Bethesda, Maryland.

Beets, D. J., T. A. M. De Groot, and H. A. Davies. 2003. Holocene tidal back-barrier development at decelerating sea-level rise: a 5 millennia record, exposed in the western Netherlands. Sedimentary Geology 158:117–144.

Behrends, B., G. Hertweck, G. Liebezeit, and G. Goodfriend. 2005. Earliest Holocene occurrence of the soft-shell clam, *Mya arenaria*, in the Greifswalder Bodden, Southern Baltic. Marine Geology 216:79–82.

Belding, D. L. 1930. The soft-shelled clam fishery of Massachusetts. Commonwealth of Massachusetts, Department of Conservation. Division of Fisheries and Game, Marine Fisheries Section, Marine Fisheries Series 1, Boston.

Brousseau, D. J. 1978. Spawning cycle, fecundity, and recruitment in a population of soft-shell clam, *Mya arenaria*, from Cape Ann, Massachusetts. U.S. National Marine Fisheries Service Fishery Bulletin 76:155–166.

Caporale, D. A., B. F. Beal, R. Roxby. and R. J. Van Beneden. 1997. Population structure of *Mya arenaria* along the New England coastline. Molecular Marine Biology and Biotechnology 6:33–39.

Cayuela, H., Y. Dorant, C. Mérot, M. Laporte, E. Normandeau, S. Gagnon-Harvey, M. Clément, P. Sirois, and L. Bernatchez. 2021. Thermal adaptation rather than demographic history drives genetic structure inferred by copy number variants in a marine fish. Molecular Ecology 30:1624–1641.

Conde, A., J. M. Novais, and J. Domínguez. 2012. The presence of *Mya arenaria* in the Ria de Aveiro is the third confirmed record of this invasive clam on the Portuguese coast. Marine Biodiversity Records 5:e88.

Crocetta, F. 2011. Marine alien mollusca in the Gulf of Trieste and neighbouring areas: a critical review and state of knowledge (updated in 2011). Acta Adriatica: International Journal of Marine Sciences 52:247–259.

Crocetta, F. 2012. Marine alien mollusca in Italy: a critical review and state of the knowledge. Journal of the Marine Biological Association of the United Kingdom 92:1357–1365.

Crocetta, F., and E. Turolla. 2011. *Mya arenaria* Linné, 1758 (Mollusca: Bivalvia) in the Mediterranean Sea: its distribution revisited. Journal of Biological Research 16:188–193.

Cross, M. E., C. R. Bradley, T. F. Cross, S. Culloty, S. Lynch, P. McGinnity, R. M. O'Riordan, S. Vartia, and P. A. Prodöhl. 2016. Genetic evidence supports recolonisation by *Mya arenaria* of western Europe from North America. Marine Ecology Progress Series 549:99–112.

De Noia, M., L. Telesca, D. L. J. Vendrami, H. K. Gokalp, G. Charrier, E. M. Harper, and J. I. Hoffman. 2020. Population genetic structure is unrelated to shell shape, thickness and organic content in European populations of the soft-shell clam *Mya arenaria*. Genes 11:article 298.

Essink, K., and A. P. Oost. 2019. How did *Mya arenaria* (Mollusca; Bivalvia) repopulate European waters in mediaeval times? Marine Biodiversity 49:1–10.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14:2611–2620.

Gomoiu, M. T. 1981. Distribution of *Mya arenaria* L. populations in the western part of the Black Sea. Cercetări Marine 14:145–158.

Gulliksen, B., B. Holte, and K. J. Jakola. 1985. The soft bottom fauna in Van Mijenfjord and Raudfjord, Svalbard. Pages 199–215 *in* J. S. Gray and M. E. Christiansen, editors. Marine biology of polar regions and effects of stress on marine organisms: proceedings of the 18th European marine biology symposium, University of Oslo, Norway, 14-20 August 1983. Interscience/Wiley, Chichester, UK.

Hanna, G D. 1966. Introduced mollusks of western North America. Occasional Papers of the California Academy of Science 48.

Hessland, I. 1945. On the quaternary Mya period in Europe. Arkiv För Zoology 37A:1–51.

Hewitt, G. 2000. The genetic legacy of the Quaternary ice ages. Nature (London) 405:907–913.

Hoffman, J. I., F. Simpson, P. David, J. M. Rijks, T. Kuiken, M. A. Thorne, R. C. Lacy, and K. K. Dasmahapatra. 2014. High-throughput sequencing reveals inbreeding depression in a natural population. Proceedings of the National Academy of Sciences of the United States of America 111:3775–3780.

Hohenlohe, P. A., M. D. Day, S. J. Amish, M. R. Miller, N. Kamps-Hughes, M. C. Boyer, C. C. Muhlfeld, F. W. Allendorf, E. A. Johnson, and G. Luikart. 2013. Genomic patterns of introgression in Rainbow and Westslope Cutthroat Trout illuminated by overlapping paired-end RAD sequencing. Molecular Ecology 22:3002–3013.

Hohenlohe, P. A., W. C. Funk, and O. P. Rajora. 2021. Population genomics for wildlife conservation and management. Molecular Ecology 30:62–82.

Holland, B. S. 2000. Genetics of marine bioinvasions. Hydrobiologia 420:63–71.

Kennedy, V. S. 2023. History of fisheries for the soft-shell clam *Mya arenaria*. Pages 423–472 *in* V. S. Kennedy and B. F. Beal, editors. The soft-shell clam *Mya arenaria*: biology, fisheries, and mariculture. American Fisheries Society, Bethesda, Maryland.

Krapal, A. M., O. P. Popa, E. I. Iorgu, M. Costache, and L. O. Popa. 2012. Isolation and characterization of new microsatellite markers for the invasive softshell clam, *Mya arenaria* (L.) (Bivalvia: Myidae). International Journal of Molecular Sciences 13:2515–2520.

Lallias, D., P. Boudry, F. M. Batista, A. Beaumont, J. W. King, J. R. Turner, and S. Lapègue. 2015. Invasion genetics of the Pacific oyster *Crassostrea gigas* in the British Isles inferred from microsatellite and mitochondrial markers. Biological Invasions 17:2581–2595.

Lasota, R., H. Hummel, and M. Wolowicz. 2004. Genetic diversity of European populations of the invasive soft-shell clam *Mya arenaria* (Bivalvia). Journal of the Marine Biological Association of the United Kingdom 84:1051–1056.

Lasota, R., K. Pierscieniak, P. Garcia, B. Simon-Bouhet, and M. Wolowicz. 2016. Large-scale mitochondrial COI gene sequence variability reflects the complex colonization history of the invasive soft-shell clam, *Mya arenaria* (L.) (Bivalvia). Estuarine, Coastal and Shelf Science 181:256–265.

Leppäkoski, E. 1994. The Baltic and the Black Sea–Seriously contaminated by nonindigenous species? Pages 37–44 *in* Proceedings of the Conference and Workshop Nonindigenous Estuarine and Marine Organisms (NEMO). National Oceanic and Atmospheric Administration, Silver Spring Maryland.

Levinton, J. 1973. Genetic variation in a gradient of environmental variability: marine bivalvia (Mollusca). Science 180:75–76.

MacNeil, F. S. 1965. Evolution and distribution of the genus *Mya*, and tertiary migrations of Mollusca. Geological Survey Professional Paper 483-G.

Maggs, C. A., R. Castilho, D. Foltz, C. Henzler, M. T. Jolly, J. Kelly, J. Olsen, K. E. Perez, W. Stam, R. Väinölä, F. Viard, and J. Wares. 2008. Evaluating signatures of glacial refugia for North Atlantic benthic marine taxa. Ecology 89:S108–S122.

Makarevich, P. R., P. A. Lyubin, and V. V. Larionov. 2000. Basic tendencies in the structural dynamics of phytoplanktonic and benthic communities in the Sea of Azov. Russian Journal of Ecology 31:411–414.

Morgan, R. P., S. B. Block, N. I. Ulanowicz, and C. Buys. 1978. Genetic variation in the soft-shelled clam, *Mya arenaria*. Estuaries 1:255–258.

Óskarsson, I. 1961. Note on some rare and new species of mollusca off the coast of Iceland. Náttúrufræðingurinn 30:176–187. (In Icelandic with English summary.)

Petersen, K., L. Rasmussen, J. Heinemeier, and N. Rud. 1992. Clams before Columbus? Nature 359:679–679.

Plachetzki, D. C., M. S. Pankey, M. D. MacManes, M. P. Lesser, and C. W. Walker. 2020. The genome of the softshell clam *Mya arenaria* and the evolution of apoptosis. Genome Biology and Evolution 12:1681–1693.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

Rhoads, A., and K. F. Au. 2015. PacBio sequencing and its applications. Genomics, Proteomics and Bioinformatics 13:278–289.

Rougemont, Q., P. A. Gagnaire, C. Perrier, C. Genthon, A. L. Besnard, S. Launey, and G. Evanno. 2017. Inferring the demographic history underlying parallel genomic divergence among pairs of parasitic and nonparasitic lamprey ecotypes. Molecular Ecology 26:142–162.

Seitz, N. B., C. F. Dungan, C. W. Walker, and S. A. Boettger. 2023. Diseases, pathogens, and parasites of the soft-shell clam *Mya arenaria*. Pages 271–291 *in* V. S. Kennedy and B. F. Beal, editors. The soft-shell clam *Mya arenaria*: biology, fisheries, and mariculture. American Fisheries Society, Bethesda, Maryland.

Sherman, K. D., J. R. Paris, R. A. King, K. A. Moore, C. P. Dahlgren, L. C. Knowles, K. Stump, C. R. Tyler, and J. R. Stevens. 2020. RAD-Seq analysis and in situ monitoring of Nassau Grouper reveal fine-scale population structure and origins of aggregating fish. Frontiers in Marine Science 7:157.

Slatkin, M. 1985. Rare alleles as indicators of gene flow. Evolution: International Journal of Organic Evolution 39:53–65.

Slatkin, M. 1987. Gene flow and the geographic structure of natural populations. Science 236:787–792.

Slatkin, M., and R. R. Hudson. 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics 129:555–562.

Snelgrove, P. V. R., B. F. Beal, and V. S. Kennedy. 2023. Reproduction and larval biology of the soft-shell clam *Mya arenaria*. Pages 51–87 *in* V. S. Kennedy and B. F. Beal, editors. The soft-shell clam *Mya arenaria*: biology, fisheries, and mariculture. American Fisheries Society, Bethesda, Maryland.

St-Onge, P., E. Parent, J. M. Sevigny, R. Tremblay, and R. Rioux-Paré. 2011. Isolation and characterization of eight novel microsatellite markers for the softshell clam (*Mya arenaria*). Molecular Ecology Resources Primer Database Mer-11-0122.

St-Onge, P., J. M. Sévigny, C. Strasser, and R. Tremblay. 2013. Strong population differentiation of softshell clams (*Mya arenaria*) sampled across seven biogeographic marine ecoregions: possible selection and isolation by distance. Marine Biology 160:1065–1081.

Strasser, C. A., and P. H. Barber. 2009. Limited genetic variation and structure in softshell clams (*Mya arenaria*) across their native and introduced range. Conservation Genetics 10:803–814.

Strasser, M. 1998. *Mya arenaria*—an ancient invader of the North Sea coast. Helgoländer Meeresuntersuchungen 52:309–324.

Strauch, F. 1972. Phylogenese, adaption und migration einiger nordischer mariner Mollusken-gen-era (*Neptunea*, *Panomya*, *Cyrtodaria* und *Mya*). Abhandlungen Der Senckenbergischen Natur-forschenden Gesellschaft 531:1–211.

Tautz, D., H. Ellegren, and D. Weigel. 2010. Editorial: next generation molecular ecology. Molecular Ecology 19:1–3.

Vendrami, D. L. J., M. De Noia, L. Telesca, E. M. Brodte, and J. I. Hoffman. 2020. Genome-wide insights into introgression and its consequences for genome-wide heterozygosity in the *Mytilus* species complex across Europe. Evolutionary Applications 13:2130–2142.

Vendrami, D. L. J., M. De Noia, L. Telesca, W. Handal, G. Charrier, P. Boudry, L. Eberhart-Phillips, and J. I. Hoffman. 2019. RAD sequencing sheds new light on the genetic structure and local adaptation of European scallops and resolves their demographic histories. Scientific Reports 9:7455.

Vendrami, D. L. J., L. Telesca, H. Weigand, M. Weiss, K. Fawcett, K. Lehman, M. S. Clark, F. Leese, C. McMinn, H. Moore, and J. I. Hoffman. 2017. RAD sequencing resolves fine-scale population structure in a benthic invertebrate: implications for understanding phenotypic plasticity. Royal Society Open Science 4:160548.

Wickham, H. 2011. ggplot2. Wiley Interdisciplinary Reviews: Computational Statistics 3:180–185.

Wilson, J. J., M. Hefner, C. W. Walker, and S. T. Page. 2016. Complete mitochondrial genome of the soft-shell clam *Mya arenaria*. Mitochondrial DNA Part A 27:3553–3554.

Zenetos, A., M. E. Cinar, M. A. Pancucci-Papadopoulou, J. G. Harmelin, G. Furnari, F. Andaloro, N. Bellou, N. Streftaris, and H. Zibrowius. 2005. Annotated list of marine alien species in the Medi-terranean with records of the worst invasive species. Mediterranean Marine Science 6:63–118.

Zenetos, A., D. Koutsoubas, and E. Vardala-Theodorou. 2004. Origin and vectors of introduction of exotic molluscs in Greek waters. Belgian Journal of Zoology 134:161–168.

Zwarts, L., and J. Wanink. 1989. Siphon size and burying depth in deposit- and suspension-feeding benthic bivalves. Marine Biology 100:227–240.

Appendix 1: Glossary

Allele discovery curve: a plot of the cumulative number of different alleles discovered for an increasing sample size of individuals. The curve reaches an asymptote when the genetic diversity of a population has been completely sampled.

Allelic richness: a measure of genetic diversity calculated as the average number of alleles per locus in a given population.

Allozymes: variant forms of an enzyme coded by different alleles at the same locus. Differenc-es in protein configuration and/or charge allow these variants to be resolved by gel or capillary electrophoresis. Allozymes were among the first forms of genetic variation to be analyzed in wild populations, but their resolution tends be low because only regions of protein-coding DNA can be evaluated and because many DNA variants in these regions do not cause detect-able changes in protein motility.

Ascertainment bias: to achieve maximal levels of polymorphism, microsatellite discovery is often biased towards a subset of the most polymorphic loci. This can give rise to ascertain-ment bias, whereby a locus chosen to be maximally polymorphic in one population (or spe-cies) is then deemed to be less polymorphic in another population (or species). Ascertainment bias affects most types of genetic marker and can produce a biased picture of genetic diversity.

Diversifying selection: diversifying (or disruptive) selection occurs when extreme values of a trait are favored over intermediate values.

Founder effect: the reduction of genetic variation that occurs when a new population is established by a small number of founding individuals. This reduction occurs because of the loss of alleles through genetic drift in very small populations.

Gene flow: the transfer of genetic material from one population to another, either via the dispersal of gametes or through the movement of individuals.

Genetic bottleneck: the stochastic reduction of genetic diversity that accompanies a severe reduction in population size. Bottlenecks can be caused by environmental change, hunting, or habitat destruction. They reduce the gene pool of a given population or species because of the loss of alleles through genetic drift.

Genetic cluster: in population genetics, a genetic cluster refers to a group of genetically related individuals within a population or species. A variety of statistical approaches have been developed to identify the number of genetic clusters from a set of individual genotypes and to assign individuals to those clusters.

Genetic diversity: genetic variation among individuals of a population or species. Genetic diversity reflects the balance between the accumulation of new alleles through mutation and the loss of alleles through natural selection and drift. Genetic diversity is a fundamental aspect of biodiversity that impacts individual and population fitness as well as evolutionary potential.

Genetic drift: random changes in allele frequency from generation to generation within a population. The effects of genetic drift tend to be stronger in small populations.

Genetic marker: an informative DNA sequence that can be used to identify individuals, populations, or species.

Haplotype: a group of alleles in an organism that are inherited together from a single parent.

Haplotype diversity: a measure of genetic diversity representing the probability that two haplotypes sampled at random within a population are different. It is also known as gene diversity.

Heterozygote advantage: a phenomenon whereby heterozygotes at a given locus have superior fitness to both homozygotes.

Isolation by distance: a positive correlation between genetic and geographic distance that arises due to restricted dispersal.

Local adaptation: the process whereby a population evolves to become better suited to its local environment.

Locus: a specific location in the mitochondrial or nuclear genome where a particular gene or genetic marker is located.

Microsatellite: an abundant form of genetic variation in eukaryotic genomes comprising tandem repeats of short DNA motifs. Microsatellites have a sufficiently high mutation rate to generate and maintain extensive polymorphism, making them ideal for characterizing population structure. Microsatellite length polymorphisms are typically resolved by separating polymerase chain reaction (PCR) products according to their size using capillary electrophoresis.

Mitochondrial DNA (*mtDNA*): DNA located in the mitochondria of eukaryotic cells. mtDNA is a very popular genetic marker owing to its ease of PCR amplification and sequencing. It is also predominantly maternally inherited and nonrecombining, simplifying the analysis and visualization of patterns of intraspecific variation. Finally, its high evolutionary rate in comparison to allozymes and its small effective population size relative to nuclear markers allow mtDNA to capture signals of recent historical events without the need for extensive sequencing effort.

mtDNA haplotype: a group of alleles within the mitochondrial genome that are inherited together as a single unit.

Population genetic structure: systematic geographical variation in allele frequencies resulting from the processes of mutation, genetic drift, gene flow, and natural selection.

Polymorphism: the presence of two or more genetic variants at a given locus.

Private allele: an allele that is only found in one of multiple genotyped populations.

Reduced representation sequencing: a family of approaches that reduce genome complexity to allow large numbers of genetic markers to be sequenced in multiple individuals.

Relict population: a population that persists as a remnant of a historically more widespread and diverse population.

Single nucleotide polymorphism (*SNP*): a type of DNA polymorphism involving variation among individuals at a single base position. Single nucleotide polymorphisms are the most common form of variation in vertebrate genomes.